PNAS
www.pnas.org

Supplementary Information for

Unraveling the genotype by environment interaction in a thermosensitive fish with a polygenic sex determination system

Benjamin Geffroy[a,1], Mathieu Besson[b,c], Núria Sánchez-Baizán[h], Frederic Clota[a,c], Alexander Goikoetxea[a], Bastien Sadoul[a,d], François Ruelle[e], Marie-Odile Blanc[e], Hugues Parrinello[f], Sophie Hermet[g], Eva Blondeau-Bidet[g], Marine Pratlong[f], Francesc Piferrer[h], Marc Vandeputte[a,c] & François Allal[a]

[1]Benjamin Geffroy
Email: bgeffroy@ifremer.fr

This PDF file includes:

Supplementary text
Figures S1 to S11
Tables S1 to S4
Legends for Datasets S1 to S4
SI References

Other supplementary materials for this manuscript include the following:

Datasets S1 to S4

Supplementary Information Text

Supplementary Materials and Methods

Genotyping

SNP calling was done using the ThermoFisher AxiomAnalysisSuite software (v5.1.1.1, Affymetrix, Santa Cara, CA) and quality controls with PLINK 1.9 (Purcell and Chang, 2015).
A first quality control of the genotypes excluded 79 individuals having a genotyping rate lower than 90%: 24 larvae sampled at 77 $DD_{10\,°C}$, 53 larvae sampled at 242 $DD_{10\,°C,}$ and 2 fish sexed at one year. Using the remaining genomic data set, we applied classical quality control by removing SNPs with a minor allele frequency (MAF) inferior to 5%. The resulting data set included 49,543 SNPs on the parents, grandparents, and the remaining 1951 offspring. Using a subset of 1,000 highly polymorphic SNPs, we reconstructed the pedigree of the offspring using the R package APIS (1). The number of fish assigned to a single parental pair was 1929 out of the 1951 genotyped (98.9%).

eGST, QTL presence and heritability assessment

For this estimation, we only kept the offspring with a sex phenotype, a genotype and a pedigree, for a total of 1107 individuals, from now on referred to as the "training" dataset. The following model was used to fit the sex phenotype:
$y = \boldsymbol{X}b + \boldsymbol{Z}u + e$ (Model 1)
where *y* is the vector of binary sex phenotypes (1 = male and 2 = female), *b* is a vector of the fixed effects (intercept and temperature). **X** is the incidence matrix relating phenotypes with the fixed effects. **Z is** the incidence matrix relating phenotypes with the random animal genetic effects. *u* is the vector of random animal genetic effects with the following distribution *N* (0, $\mathbf{G}\sigma^2$g), where **G** is the genomic relationship matrix, σ^2g the additive genetic variance, and **e** a vector of residuals following a *N*(0,$\mathbf{I}\sigma^2$e) distribution, where σ^2e is the residual variance and **I** the identity matrix. Model 1 was applied for sex as a single trait (considering intercept and temperature as fixed effects). A multi-trait animal model (Model 2) with sex considered a different trait at LT and at HT (sex_LT and sex_HT), with intercept as the only fixed effect, was used to estimate the genetic correlation between sex_LT and sex_HT.
The variance components (σ^2g and σ^2e) were estimated using a Gibbs sampler with 500,000 iterations, 100,000 burn-in, and one sample was kept every 20 iterations for posterior analysis. The residual variance σ^2e was set to a value of 1. The posterior distributions were analyzed with the R package boa (2).

The heritability of sex on the PST scale was estimated as:

$$h_u^2 = \frac{\sigma^2{}_g}{\sigma^2{}_g + \sigma^2{}_e}, \quad (2)$$

Using the sex single-trait threshold model (Model 1) built from the training dataset, and the Genomic-based Estimated Breeding Values (GEBV) of PST, we could obtain an estimated Genetic Sex Tendency (eGST) for the rest of the offspring dataset (offspring with a genotype and a pedigree, but no sex phenotype N= 844) by using the POSTGIBBSF90 software. The accuracy of the model was assessed by applying a leave-one-out cross-validation method on the training population. For this cross-validation, the phenotype of individual ***i*** was masked, and the resulting threshold model was used to predict its $eGST_i$, which was then compared to its actual sex phenotype y_i, using the Receiver Operating Characteristic curves (ROC curves) methodology to assess the efficacy of classifying the animals as male or female. The Area Under Curve (AUC) metric (3) was used to interpret the performance of the genomic prediction models, with values of 1 representing the perfect classifier.

Following the same model as described above (Model 1), we used the posterior SNP solutions to apply a weighted genome wide association scan (wGWAS)(4). At the first iteration, **G** was calculated as $\mathbf{G} = \mathbf{ZDZ'}/q$, with **D** equal to the identity matrix (5). Then, direct genomic values (DGVs) were obtained from GEBVs as $DGV_i = -(\Sigma_{j,\, j \neq i}\, g^{ij} GEBV^j / g^{ji})$ with g^{ij} elements of the $\mathbf{G}^{-1}$ matrix (6), and converted to SNP effects as $a_i = \mathbf{DZ'G^{-1}}DGV_i$. A new **D** matrix was computed from the genetic variance explained by each SNP. After two rounds, no further modification of the variance explained by the SNPs was seen. The LOD scores of SNPs were computed as -log(*pvalue*) and a regional variance explained was computed by summing variance explained in 50 SNPs sliding windows. We considered as QTLs the genomic segments that explained a proportion of genetic variance higher than 2%, as done in (7).

RNA extraction, RNA-seq and RNA-seq data analysis

RNA extraction

The gonads of juveniles were particularly small, and the first trial of mRNA extraction with a dedicated kit failed (n = 27). We thus performed mRNA isolation using QIAzol lysis reagent (Quiagen) on remaining gonad samples (n = 43, 22 for LT and 21 for HT) and on individuals at the “flexion” and “all fins” stages. All samples were mechanically disrupted using ball mills using the appropriate QIAzol quantity, following manufacturer instructions. The RNA concentration was assessed by measuring the A260/A280 ratio using the NanoDrop® ND-1000 V3300 Spectrophotometer (Nanodrop Technology Inc., Wilmington, DE, USA) and RNA quality was checked using an Agilent bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) via the

electrophoretic trace method. The RNA integrity number was between 8.5 and 10 for all samples (Table S4). One extraction failed at the “all fins” stages, and one sample presented the lowest RNA quantity, and was removed from the analysis (n total = 68).

RNA-Seq

All RNAs available at the “all fins” stage (n = 68) and at the juvenile stage (n = 43) were used for RNA-Seq. In addition, because of spare space on the sequencer, RNA-Seq was also performed on 10 individuals at the “flexion” stage. We selected five individuals from the LT group and five from the HT group based on their eGST, choosing those which were genetically “weak” females (i.e. with a positive, but not extreme eGST: $0 < eGST < 0.5$).
Because of the difference in RNA quantities extracted from whole bodies or from gonads, cDNA library construction was performed using two separate methods. For whole bodies, libraries construction was described in (8) at the “all fins” stages and was performed similarly at the “flexion” stage.

cDNA library construction of gonads at the juvenile stage

The 43 libraries from juvenile gonad samples were prepared using a NuGEN Universal Plus mRNA-Seq kit (Tecan, Redwood city, CA, USA) according to the manufacturer’s instructions. Briefly, polyadenylated RNAs were selected on oligo(dT) magnetic beads. Chemical fragmentation of selected RNAs was done. The first-strand cDNA synthesis started (random primers) in the presence of Actinomycin D, in order to prevent spurious DNA-dependent synthesis, improving strand specificity. Following this, the second strand cDNA synthesis and repairing the ends of double-stranded cDNA fragments followed. Then, adapters were ligated at both ends of fragmented nucleic acid, preparing them for hybridization onto a flow cell. Next, a selection of the strand to be sequenced and purification were done. Ligated cDNAs were amplified following a 15-cycle PCR. This step also leads to a selection, since only fragments at both ends will be enriched. PCR products were purified using AMPure XP Beads (Beckman Coulter Genomics, Brea, CA, USA).
Libraries were validated using a Fragment Analyzer (Agilent Technologies, Santa Clara, CA, USA) and quantified using the KAPA Library quantification kit (Roche, Bâle, CHE).

Libraries sequencing:

The 10 “flexion”, 68 “all fins” and 43 gonad samples libraries were pooled in equimolar amounts, denatured with NaOH, and diluted to 20-22 pM before the clustering. This was performed on a cBot system (Illumina, San Diego, CA, USA) and sequencing was performed on an HiSeq 2500

(Illumina, San Diego, CA, USA) using the single-end 1*50 nt protocol on 9 lanes of a flow cell V4 or V2. Sequencing produced between 5 and 34 million passed filter clusters per library (9).

Sequencing quality control

Image analyses and base calling were performed using the Illumina HiSeq Control Software and the Real-Time Analysis component. Demultiplexing was performed using Illumina's conversion software (bcl2fastq 2.20). The quality of the raw data was assessed using FastQC from the Babraham Institute and the Illumina software SAV (Sequencing Analysis Viewer). Potential contaminants were monitored with the FastQ Screen software from the Babraham Institute.

RNA-Seq data analysis

A splice junction mapper, TopHat 2.1.1 (10) using Bowtie 2.3.4.3 (11), was utilized to align the RNA-Seq reads to the *Dicentrarchus labrax* genome (NCBI, reference GCA_000689215.1) with a set of gene model annotations (12). Final read alignments having more than three mismatches were discarded. Samtools (v.1.9)(13) was used to sort the alignment files. Then, the counting was performed with Featurecounts (v.1.6.4 (14)). For both the "flexion" and the "all fins" samples (libraries with high starting material), reads were mapped to the opposite strand of the gene (-s2 option), while for gonad samples (libraries with low starting material), the reads were mapped to the strand of the gene (-s1 option). Before statistical analyses, genes with less than 30 reads (mean of all the analyzed samples) were filtered out.

RRBS sequencing and preliminary analysis

RRBS pools were sequenced on a HiSeq2500 (Illumina, San Diego, CA) using 50 base pair (bp) single-read sequencing to obtain at least 22.2 million reads/sample (~15x coverage). The adapter-trimmed sequence reads were aligned to the reference genome (seabass_V1.0.104, GCA_000689215.1) (12), and methylation values were extracted using Bismark v0.20.0 (15) with default parameters. Following, MethylKit v1.18.0 (16) was used to remove data with low coverage (<10x) and data with extremely high coverage (>99.9th percentile of coverage in each sample) to discard reads with possible PCR bias. Furthermore, data were normalized for read coverage distribution between samples.

Data Analysis

Gene Ontology selection

We selected several GO biological processes based on our hypotheses: Sex determination (GO:0007530), Sex differentiation (GO:0007548), Steroids metabolic process (GO:0008202); Regulation of growth rate (GO:0040009), Lipid biosynthetic process (GO:0008610), Execution phase of apoptosis (GO:0097194); Glucocorticoid metabolic process (GO:0008211); Cellular response to heat (GO:0034605); Regulation of gene expression, Epigenetic (GO:0040029); Histone modification (GO:0016570); DNA methylation or demethylation (GO:0044728); Gene silencing by miRNA (GO:0035195); Gene silencing by RNA (GO:0031047); Regulatory RNA binding (GO:0061980) and Germ cell development (GO:0007281). The presence of the genes in these GOs was then tested in regards to the list of genes previously identified as significantly expressed in the models described in the M&M of the main file.

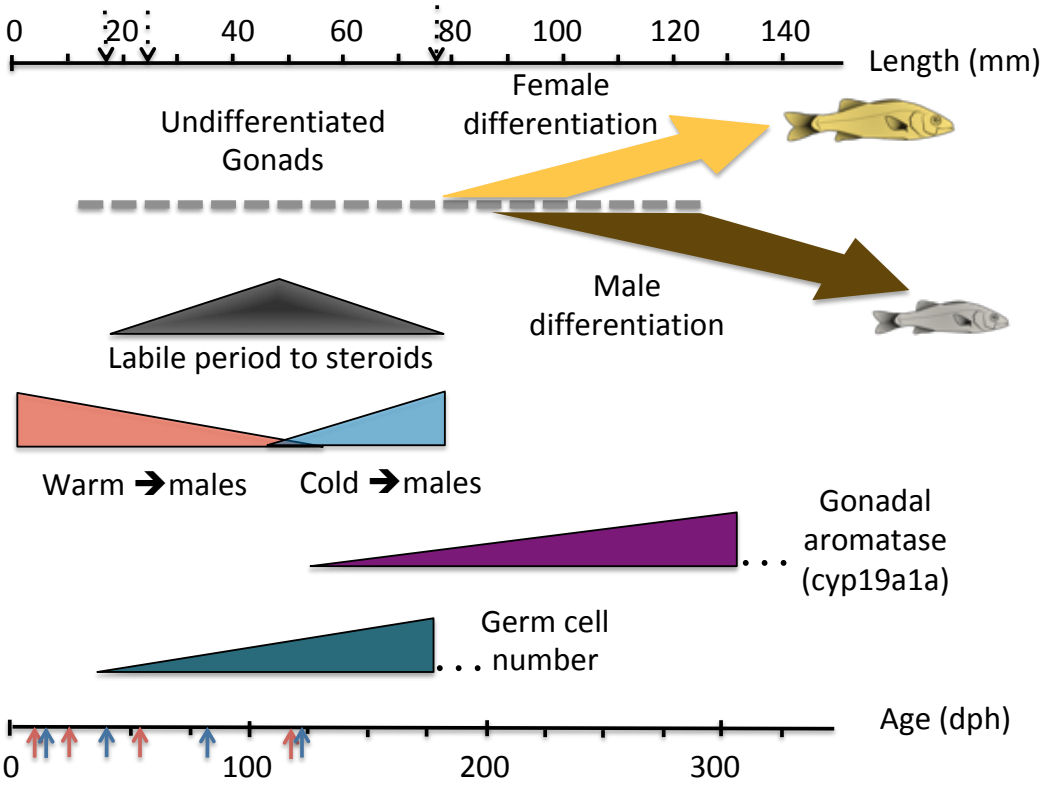

Fig. S1. Scheme of the gonadal development of the European Sea bass, modified from (17). The dotted black arrows indicate the size at which larvae were sampled for molecular analyses. The first sampling is not represented since no size (in length) is available and that these larvae only served for genotyping. The blue and red arrows indicate at which age LT and HT larvae were collected, respectively. Red triangle means masculinization when individuals are exposed too early to high temperature based on age at the start of experiment (17), and blue triangle means masculinization when animals are exposed to low temperature for a relatively long period (based on age)(18).

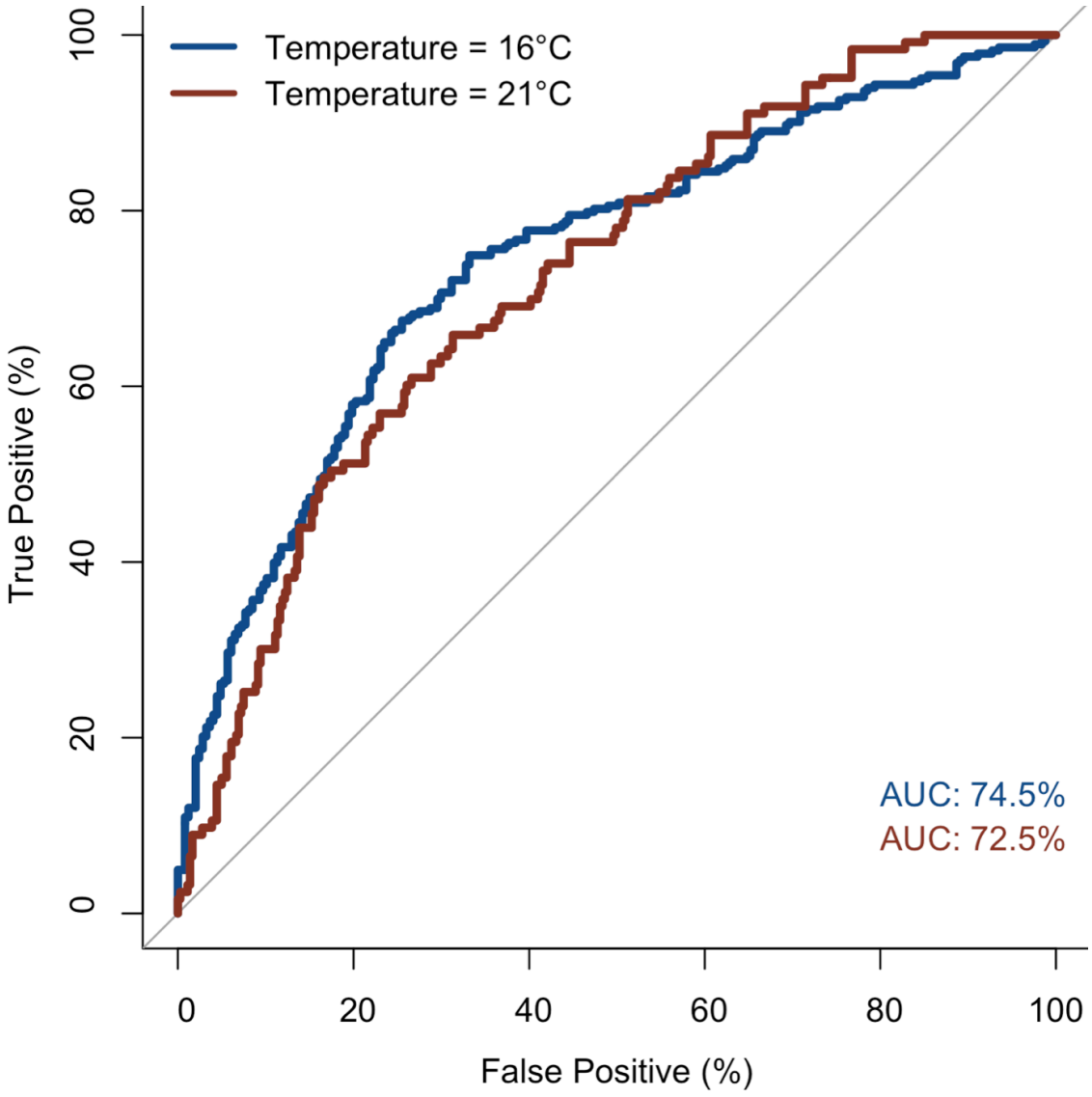

Fig. S2. Receiver operator characteristic (ROC) curve and corresponding area under the curve (AUC) metric for genetic sex tendency (eGST) of European sea bass.

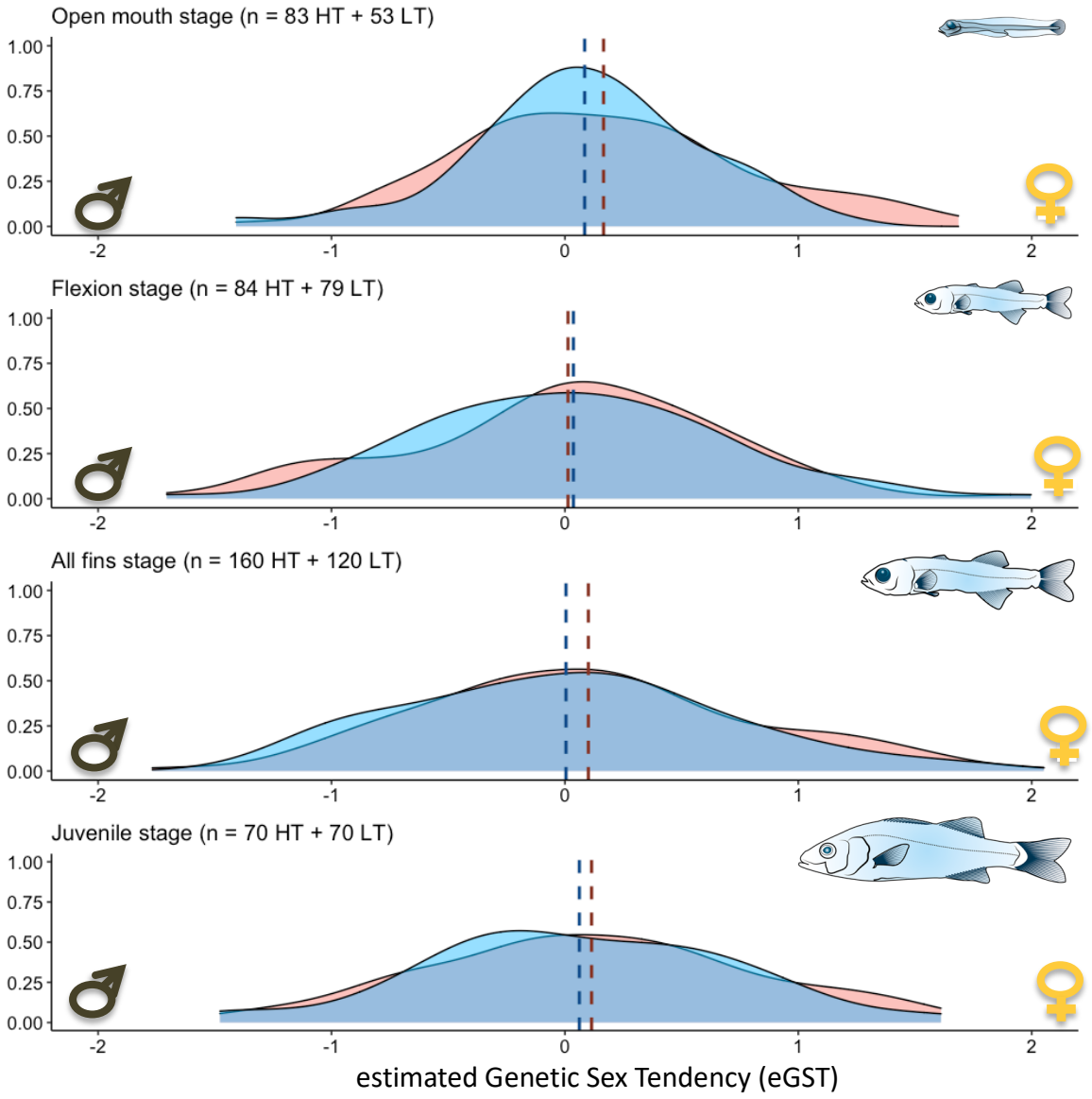

Fig. S3. Density plot of the genetic sex tendency (eGST) of European sea bass at four distinct developmental stages and two temperatures: HT= high temperature 21 °C; LT= low temperature 16 °C. The number (n) of genotyped individuals is indicated. The blue (LT) and red (HT) dotted lines indicate the mean genetic sex tendency at each temperature.

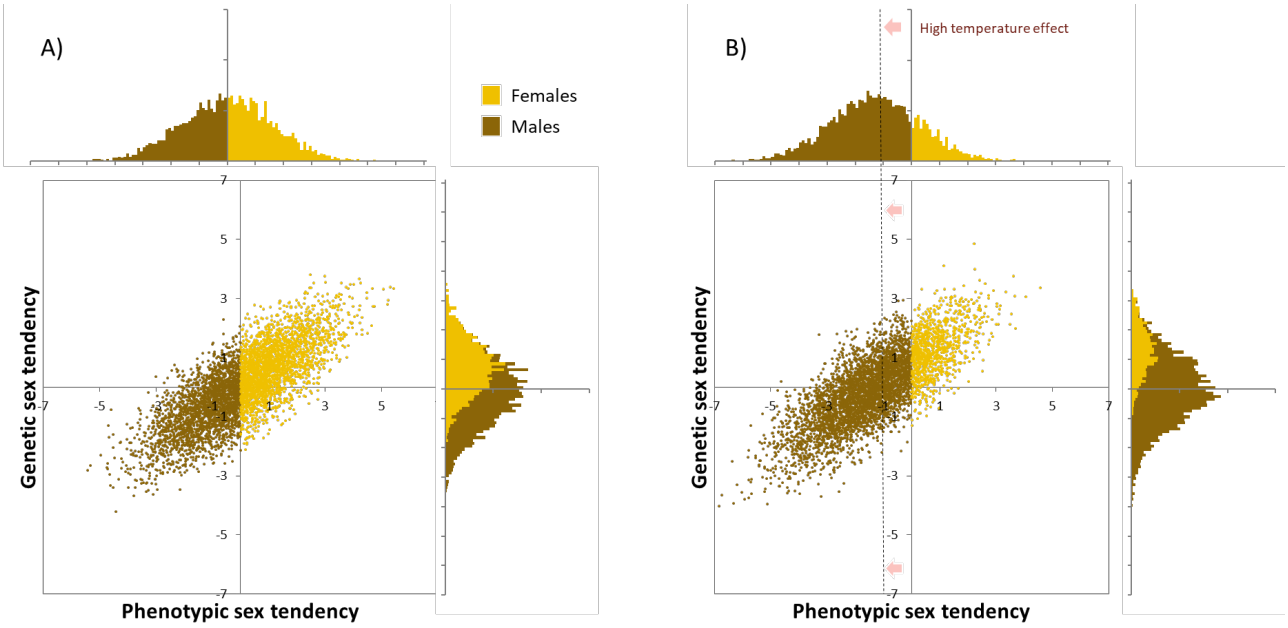

Fig. S4. Simulated marginal and joint distributions of phenotypic and genetic sex tendency under the threshold model of polygenic sex determination in A) a sex-neutral environment and B) a masculinizing, high temperature environment. Heritability (h^2) of phenotypic sex tendency was set to 0.60. When phenotypic sex tendency is positive, all animals differentiate as females, while they differentiate as males when phenotypic sex tendency is negative. In a sex-neutral environment, the mean phenotypic sex tendency of the population is zero, and thus the sex-ratio is 50:50. High temperature displaces the mean phenotypic sex tendency of the population to the left, resulting in lower proportion of females (here 30%). Note that genetic sex tendency has a lower variance than phenotypic sex tendency, as $h^2<1$. Also note that $h^2<1$ also implies that only animals with high genotypic sex tendency always differentiate as females, this constraint being even higher in the high temperature environment (B). A total of 5000 individuals were simulated, with environmental variance of sex tendency $V_E=1.0$, additive genetic variance $V_A=1.5$ and phenotypic variance $V_P=V_A+V_E=2.5$, which corresponds to $h^2=V_A/V_P=0.60$.

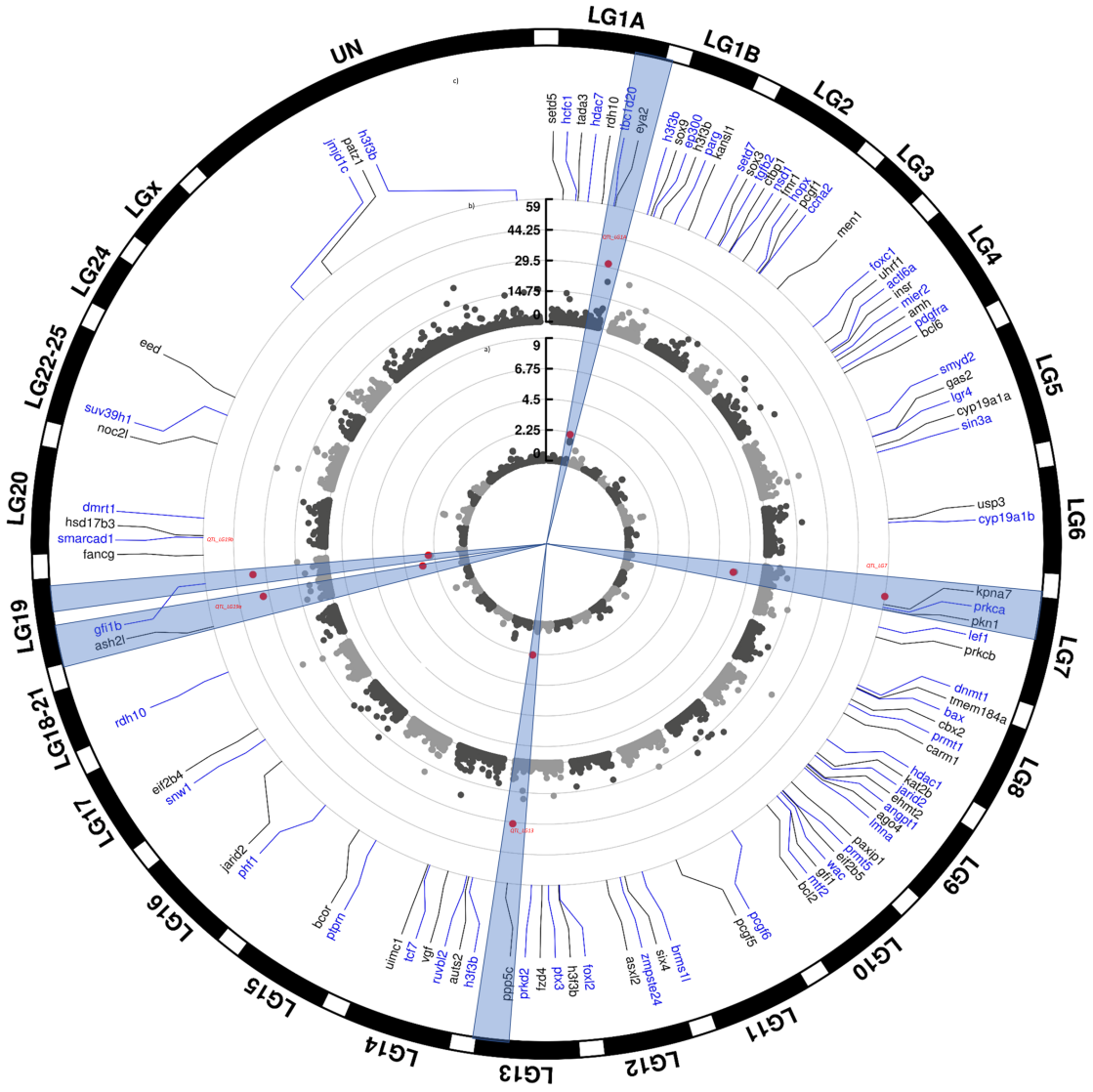

Fig. S5. Circular Manhattan plots of genome wide association scan (GWAS) for sex (male/female) with, from the center to the border: the percentage of variance explained in a sliding window of 50 adjacent SNPs, the SNP effects in *-log(pvalue)* computed after 2 iteration of weighted GWAS, and the location of the genes previously reported as having a significant differential expression level in link with “sex differentiation”, “sex determination” and “histone modification” in the present study. Red dots represent SNP markers explaining over 2% of the variance, and light blue shaded areas represent the regions including the putative quantative trait loci.

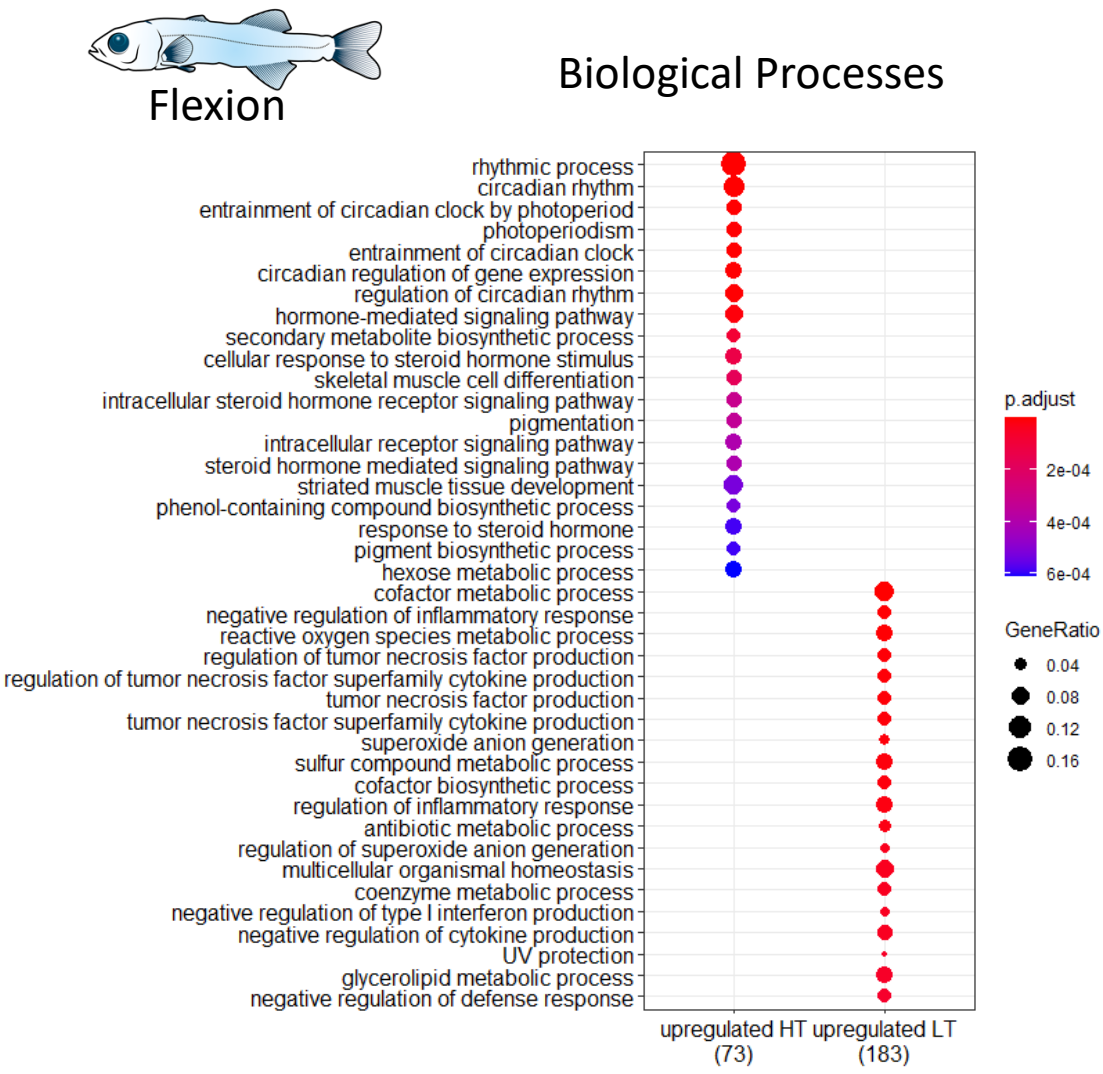

Fig. S6. Dotplot showing enrichment of Gene Ontology biological processes for up and downregulated DEGs between low (LT) and high temperature (HT) of whole fish at the "flexion" stage. Data are from the RNA-Seq. The size of the circle represents the number of genes enriched in the entry, and the color indicates the significance of the p-value.

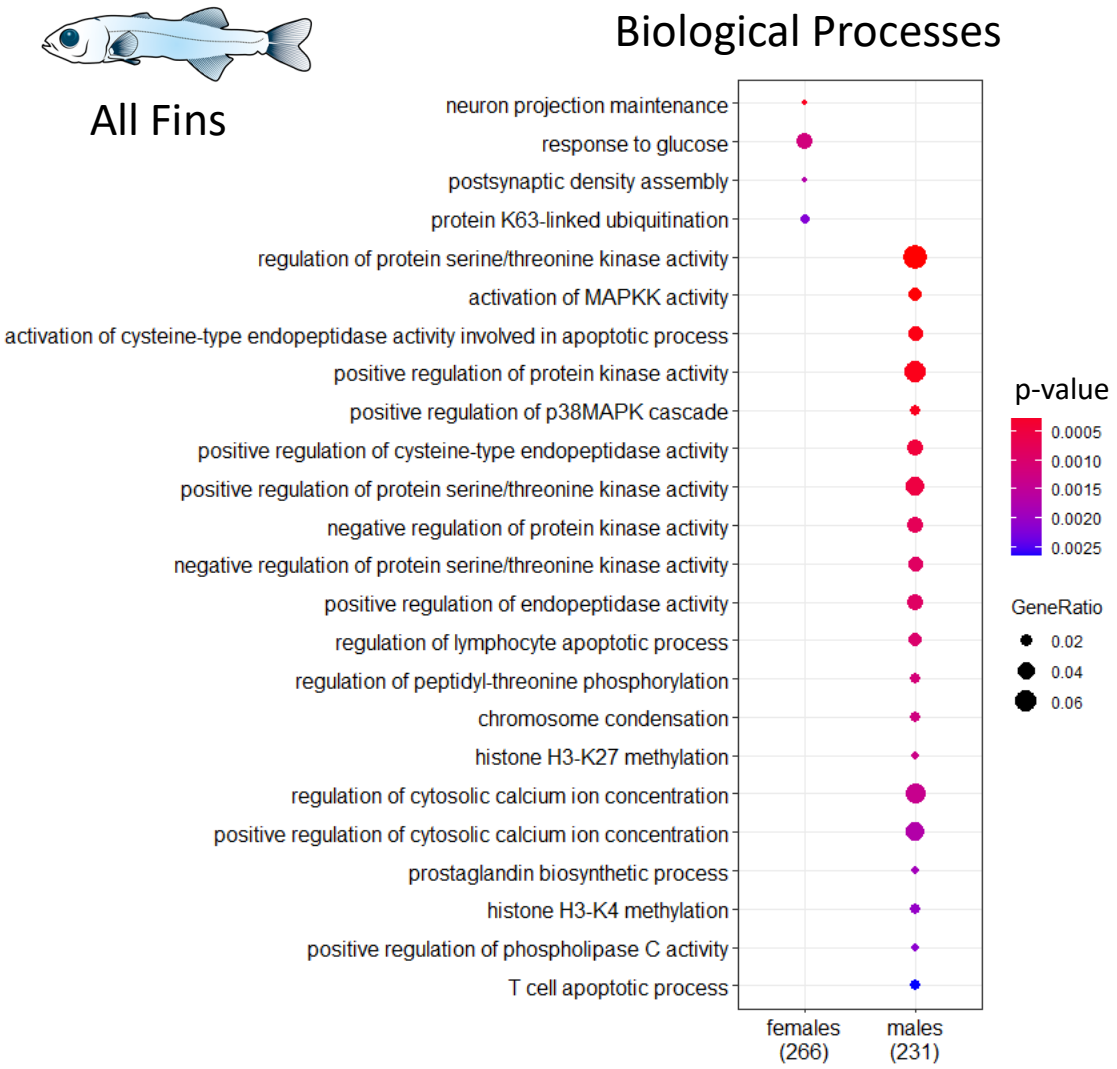

Fig. S7. Dotplot showing enrichment of Gene Ontology biological processes for up and downregulated DEGs between females (positive correlation with eGST) and males (negative correlation with eGST) in whole fish at the "all fins" stage (53 and 78 dph, all temperature treatments cofounded). Data are from the RNA-Seq. The size of the circle represents the number of genes enriched in the entry, and the color indicates the significance of the p-value.

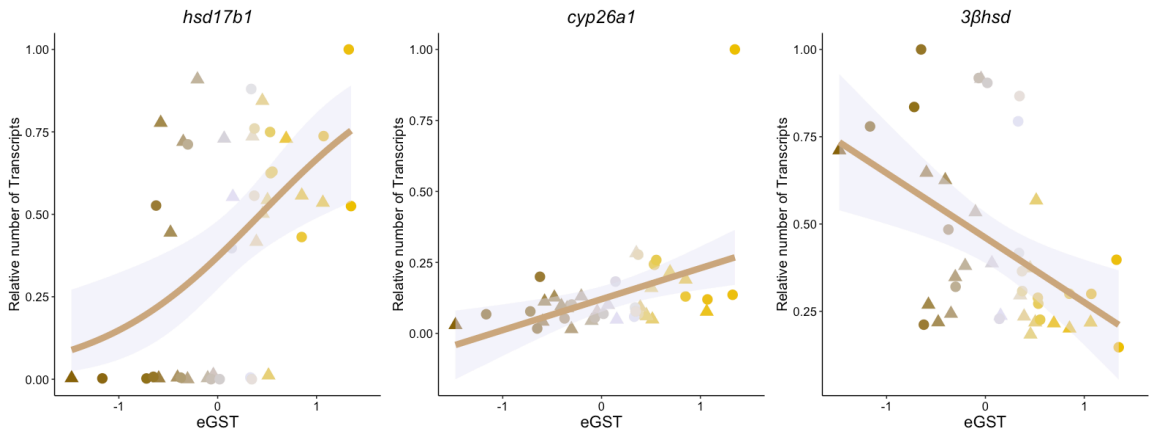

Fig. S8. Significant ($p < 0.01$) linear correlation between the estimated genetic sex tendency (eGST) and 3 genes (number of transcripts on the y axis) involved in the GO steroids metabolic process, namely *hsd17b1, cyp26a,* and *3β-hsd,* from the RNA-Seq analysis of the gonads of fish at the juvenile stage. Circles represented fish kept at high temperature (HT) = 21 °C; and triangles those kept at low temperature (LT) = 16 °C. Individuals are represented with a color gradient, from maroon to yellow, representing their eGST.

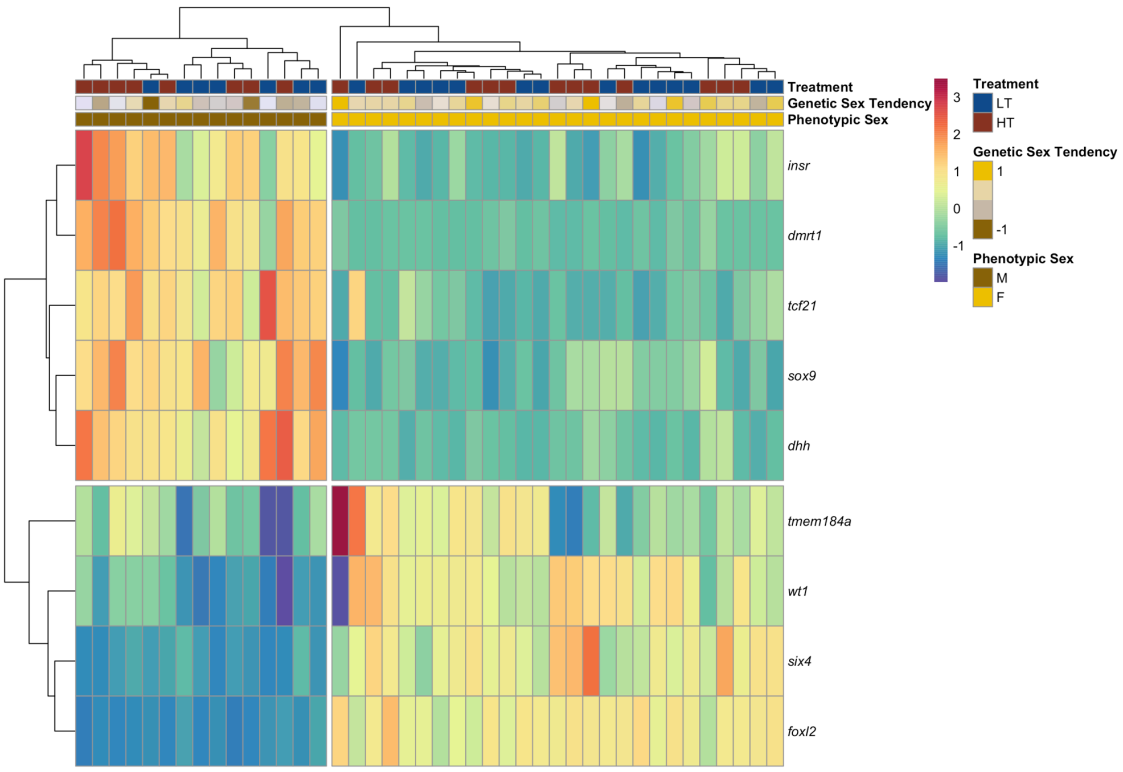

Fig. S9. Heatmap of all genes differentially expressed in gonads of fish at the juvenile stage between genetic males (eGST < 0) and genetic females (eGST > 0) in the gene ontology "sex determination". Individuals are represented with a color gradient, from maroon to yellow, representing their eGST. The phenotypic sex is assessed based on the automatic threshold proposed by the pheatmap package, with indivividuals represented in maroon considered as phenotypic males and those in yellow considered as phenotypic females. The fish from the high temperature (HT) tretament = 21 °C are represented in red; and those from the low temperature (LT) treatment = 16 °C are represented in blue.

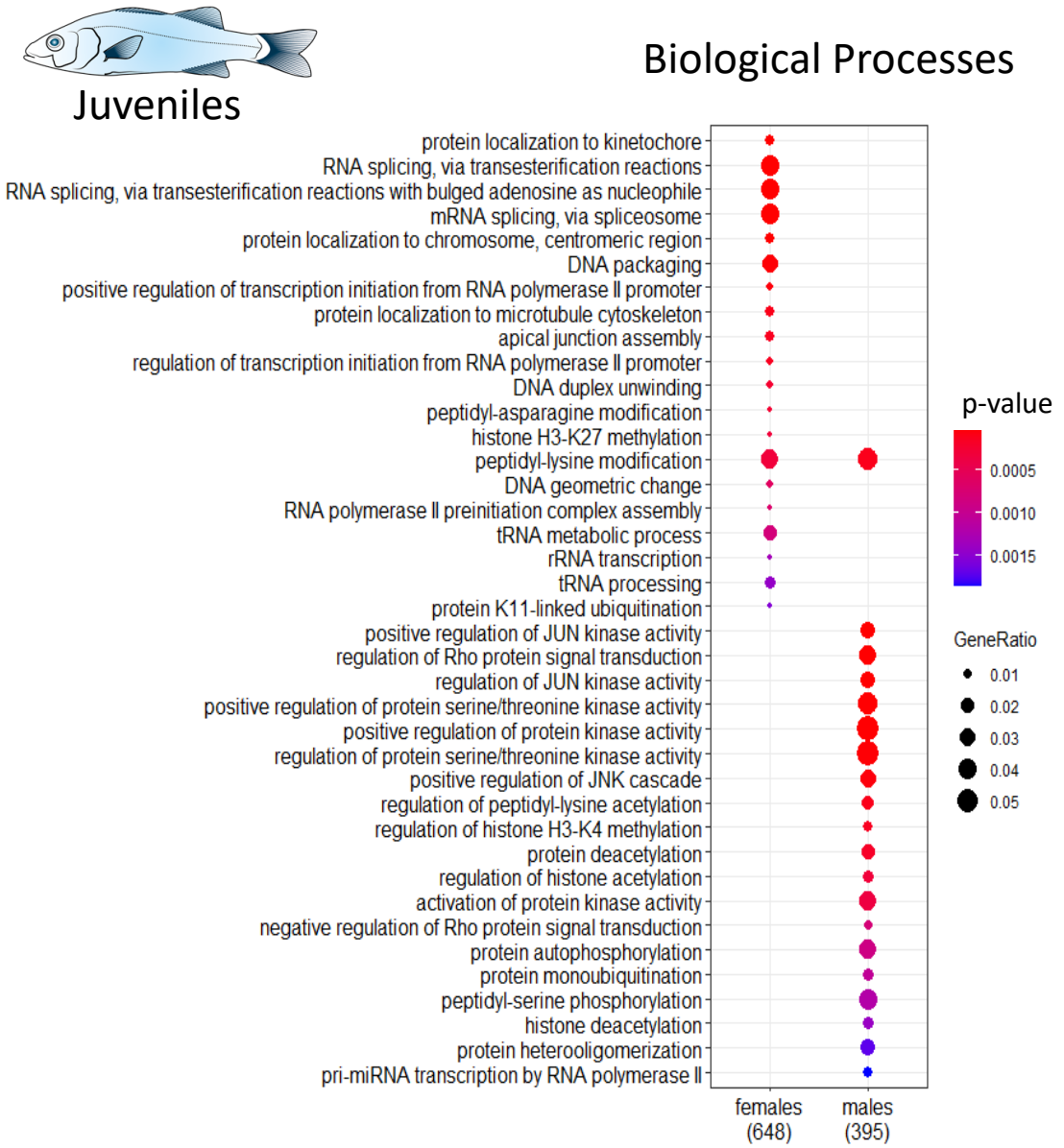

Fig. S10. Dotplot showing enrichment of Gene Ontology biological processes for up and downregulated DEGs between females (positive correlation with eGST) and males (negative correlation with eGST) in the gonads of fish at the juvenile stage (117 and 124 dph, all temperature treatments cofounded). Data are from the RNA-Seq. The size of the circle represents the number of genes enriched in the entry, and the color indicates the significance of the p-value.

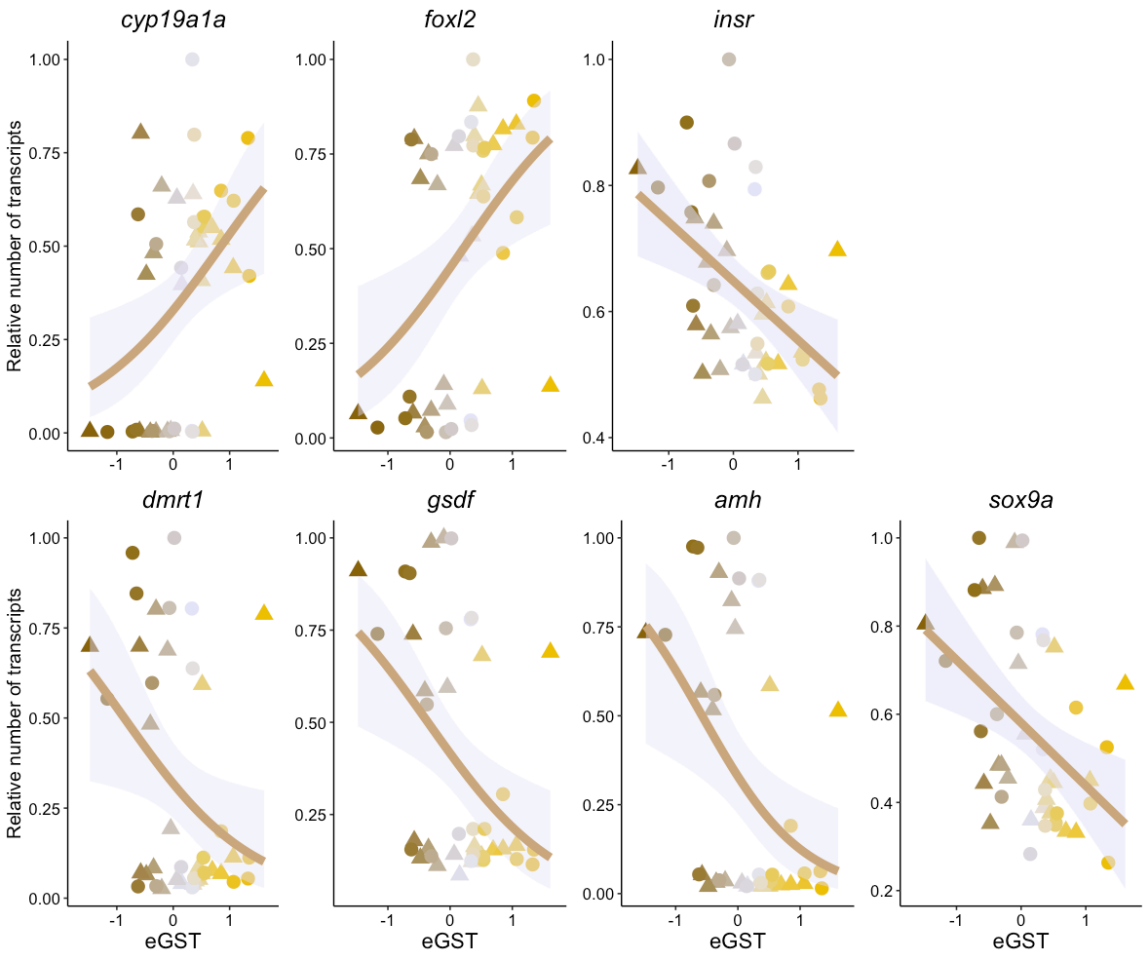

Fig. S11. Significant ($p < 0.01$) linear correlation between the estimated genetic sex tendency (eGST) and both insr and sox9a (relative number of transcripts on the y axis). For five genes, *cyp19a1a, foxl2, dmrt1, amh, gsdf*, and the PC1 axis, a dichotomic distribution was observed and modelled with a "quasibinomial" function. Note that the correlation is still detectable with the outlier (extreme positive value at LT). Circles represented fish kept at high temperature (HT = 21 °C; n = 21); and triangles those kept at low temperature (LT = 16 °C; n = 22). Individuals are represented with a color gradient, from maroon to yellow, representing their lower or higher eGST.

Table S1. Genes from different GOs that are significantly correlated to the eGST in fish at the "all fin" stage. (the direction of the effect provide the direction of the correlation (negative = male and positive = females)

Linear Model								
Sex Determination			Sex differentiation			Histone modifications		
gene	direction of the effect	p-value	gene	direction of the effect	p-value	gene	direction of the effect	p-value
sox9a	-8,06	0,05	eif2b5	17,90	0,042	pcgf5	7,11	0,039
sox3	-7,64	0,05	fzd4	-26,86	0,014	prkd2	-11,64	0,046
			h3f3b	-90,28	0,032	uimc1	14,56	0,012
			vgf	9,29	0,038	phf1	10,35	0,028
			sox9a	-8,06	0,045	jarid2	4,24	0,025
			fancg	-8,50	0,016	snw1	18,47	0,029
			hsd17b3	-10,61	0,009	gfi1b	-3,29	0,047
			ptprn	12,83	0,031	eya2	-6,52	0,033
			foxc1	-4,43	0,043	setd7	28,07	0,031
			pdgfra	-30,60	0,019	prkca	4,29	0,045
			bax	-16,71	0,044	hcfc1	6,05	0,042
			patz1	-11,47	0,008			
lipid biosynthetic process						Growth Regulation		
gene	direction of the effect	p-value	gene	direction of the effect	p-value			
cds2	13,53	0,034	ppp1r9b	10,21	0,049			
ptges	-13,84	0,017	ptprs	24,41	0,037			
st6galnac4	-11,85	0,043	egfr	-26,74	0,001			
ptgis	-9,04	0,006	lamtor2	-19,86	0,046			
hsd17b3	-10,61	0,009	jarid2	4,24	0,025			
ddx20	8,25	0,036	gas1	-8,16	0,050			
acot7	-12,01	0,028	myl2	-42,51	0,030			
cyp27a1	14,26	0,033	apbb2	15,69	0,009			
prox1	7,40	0,016	nf2	-8,84	0,050			
prkag3	-7,22	0,015	prox1	7,40	0,016			
pip5k1c	8,55	0,022	mmp14	-55,87	0,047			
lcat	50,85	0,011	foxc1	-4,43	0,043			
sphk1	-4,37	0,032	rgs2	-56,13	0,008			
dgkd	-3,94	0,039	unc13a	15,28	0,041			
ceacam1	-11,12	0,009	bbs2	6,09	0,031			
abcd2	4,53	0,040	sphk1	-4,37	0,032			
			cspg5	18,36	0,047			
			sema7a	-5,82	0,018			
			gh	17,56	0,009			

			Quadratic Model	
gene	direction of the effect	p-value ST	direction of the quadratic effect	p-value quadratic T°C
sgsm2	79,48	0,009	109,57	0,014
map3k3	-182,14	0,001	-162,59	0,039
entpd2	39,85	0,005	43,65	0,035
hspg2	-698,83	0,000	-631,23	0,026
bmp7	-50,72	0,004	-55,32	0,031
rasgrf1	69,51	0,010	77,17	0,050
eno2	186,27	0,006	228,42	0,023

Table S2. Genes from different GOs that are significantly correlated to the eGST in the gonad of juveniles fish. (the direction of the effect provide the direction of the correlation (negative = male and positive = females)

Sex differentiation			Sex Determination			Histone modifications		
gene	**direction of the effect**	**p-value**	**gene**	**direction of the effect**	**p-value**	**gene**	**direction of the effect**	**p-value7**
bcl2	-631,78	0,0084	insr	-263,0	0,0003	paxip1	44,34	0,003
insr	-263,00	0,0003	six4	35,9	0,0018	prmt5	223,46	0,006
six4	35,88	0,0018	foxl2	246,0	0,0066	wac	-128,71	0,008
foxl2	245,97	0,0066	sox9	-350,0	0,0067	mtf2	165,80	0,007
ptx3	-223,55	0,0037	dmrt1	-292,7	0,0065	pcgf6	21,08	0,008
tcf7	367,51	0,0066	tmem184a	32,5	0,0032	parg	145,11	0,006
tgfb2	-779,96	0,0092				pcgf5	-38,01	0,003
eif2b4	55,26	0,0038				hdac1	208,51	0,002
rdh10	256,98	0,0023				zmpste24	173,91	0,000
sox9	-349,98	0,0067				asxl2	-115,91	0,006
dmrt1	-292,73	0,0065				prkd2	-85,95	0,002
amh	-13130,47	0,0047				ppp5c	88,36	0,008
lgr4	-81,47	0,0096				ruvbl2	74,33	0,003
gas2	-330,17	0,0035				nsd1	-74,40	0,007
cbx2	105,19	0,0047				lmna	230,97	0,001
tbc1d20	61,47	0,0037				brms1l	29,99	0,003
angpt1	-438,16	0,0057				ash2l	86,93	0,001
ago4	-95,22	0,0055				setd5	-268,17	0,005
jmjd1c	-221,58	0,0015				tada3	45,48	0,010
aromatase	1991,61	0,0045				hdac7	-166,63	0,002
						auts2	-260,43	0,002
						kansl1	-224,33	0,010
						fmr1	82,03	0,006
						hopx	-191,18	0,009
						pcgf1	42,06	0,010
						ccna2	616,58	0,004
						ctbp1	-223,82	0,001
						smarcad1	54,69	0,002
						noc2l	92,77	0,009
						suv39h1	25,43	0,002
						eed	63,74	0,000
						men1	57,21	0,008
						uhrf1	339,56	0,002
						actl6a	140,25	0,001
						mier2	34,26	0,009
						bcl6	-144,20	0,004
						gfi1	-20,71	0,008

	smyd2	131,17	0,002
	usp3	41,95	0,007
	sin3a	-194,65	0,003
	kpna7	660,98	0,007
	pkn1	188,11	0,001
	lef1	211,69	0,002
	prkcb	-99,81	0,008
	ep300	-287,62	0,003
	dnmt1	163,76	0,002
	prmt1	311,66	0,006
	carm1	124,62	0,000
	kat2b	-91,65	0,001
	ehmt2	70,12	0,004
	jmjd1c	-221,58	0,002
	bcor	-25,94	0,003

Table S3. Descriptive information regarding sampling

	Stages :		***Open mouth***	***Flexion***	***All Fins***	***Juveniles (early)***	***Juveniles (late)***	***Values for all late juveniles with SD***
	DD10 °C		77	242	550	1254 at HT; 1054 at LT		
	weight (g)		n/a	0.004 (0.002)	0.077 (0.024)	4.88 (1.17)	102.3	97.0 (34.4)
	length (cm)		n/a	1.3 (0.1)	2.2 (0.17)	7.33 (0.57)	19.96	19.47 (2.09)
HT	age (dph)		10	25	53	117	392	
	Number sampled per family	Fam1	13	12	28	15		
		Fam2	10	9	17	6		
		Fam3	15	8	16	7		
		Fam4	4	10	18	6		
		Fam5	18	16	30	20		
		Fam6	2	3	7	3		
		Fam7	13	13	17	5		
		Fam8	5	13	27	8		
	weight (g)		n/a	0.008 (0.004)	0.134 (0.064)	4.67 (1.21)	81.8	79.5 (25.9)
	length (cm)		n/a	1.4 (0.13)	2.4 (0.35)	7.31 (5.9)	18.4	18.24 (1.74)
	age (dph)		13	40	78	124	392	
LT	Number sampled per family	Fam1	9	13	25	10		
		Fam2	5	5	20	10		
		Fam3	5	10	12	10		
		Fam4	4	9	6	0		
		Fam5	9	20	29	10		
		Fam6	3	5	7	6		
		Fam7	3	10	2	16		
		Fam8	9	6	17	8		
	Genotyping		Yes	Yes	Yes	Yes	Yes	
	RNA-seq		No	Whole Body	Whole Body	Gonads	No	
	Energy		No	No	Yes	Yes	No	
	DNA methylation		No	No	No	No	Yes	
	Sexing		No	No	No	Gonad histology	Visual inspection	

Table S4. RNA integrity, A260/A280 and A260/A230 for all samples used in the RNA-seq

Sample name		RIN	A260/A280	A260/A230
flexion stage - whole body				
P2	16 04	8,7	2,043	2,131
P2	16 06	8.3	na	na
P2	16 13	8,9	2,032	2,214
P2	16 28	8,1	na	na
P2	16 29	8.6	na	na
P2	21 11	8.4	na	na
P2	21 14	8.2	na	na
P2	21 27	8,8	1,966	2,163
P2	21 28	8,3	2,007	2,042
P2	21 31	9.4	na	na
all fins stage - whole body				
P3	16 01	9,6	2,076	2,010
P3	16 03	9,1	2,044	2,045
P3	16 04	9,6	2,045	1,881
P3	16 05	9,1	2,032	2,033
P3	16 06	9,7	2,056	1,835
P3	16 07	9,7	2,059	1,856
P3	16 08	9,1	2,055	2,045
P3	16 09	9,0	2,033	2,030
P3	16 10	9,6	2,092	1,962
P3	16 11	10,0	2,059	2,008
P3	16 12	9,8	2,096	1,950
P3	16 13	9,6	2,053	2,040
P3	16 14	9,7	2,088	1,967
P3	16 15	9,1	2,030	2,034
P3	16 16	9,0	2,040	2,220
P3	16 17	9,7	2,051	1,989
P3	16 18	8,9	2,041	2,151
P3	16 19	9,4	2,037	2,141
P3	16 20	9,4	2,078	1,976
P3	16 21	9,0	2,040	2,274
P3	16 22	9,6	2,075	1,818
P3	16 23	10,0	2,071	1,823
P3	16 24	9,0	2,069	1,846
P3	16 25	9,2	2,039	2,202
P3	16 26	9,5	2,072	1,747
P3	16 27	9,0	2,099	2,117

P3	16 28	9,3	2,055	2,237
P3	16 29	9,2	2,059	2,022
P3	16 30	9,5	2,060	2,073
P3	21 01	9,2	2,038	2,248
P3	21 02	9,1	2,065	2,270
P3	21 03	9,5	2,069	2,153
P3	21 04	9,5	2,056	2,381
P3	21 05	9,2	2,055	2,234
P3	21 06	8,9	2,052	2,298
P3	21 07	9,4	2,012	2,158
P3	21 08	9,2	2,026	2,214
P3	21 09	9,3	2,045	1,957
P3	21 10	9,2	2,055	2,228
P3	21 11	9,1	2,050	2,195
P3	21 12	9,4	2,044	2,011
P3	21 13	8,0	2,029	2,212
P3	21 14	9,4	2,038	2,115
P3	21 16	9,1	1,994	2,023
P3	21 17	9,2	2,060	2,174
P3	21 18	9,1	2,043	2,291
P3	21 19	9,4	2,051	2,245
P3	21 20	9,0	2,066	2,254
P3	21 21	9,0	2,062	2,224
P3	21 22	9,1	2,045	2,185
P3	21 23	9,3	2,014	2,223
P3	21 24	9,1	2,040	2,162
P3	21 25	9,1	2,017	2,343
P3	21 26	8,9	2,019	2,360
P3	21 27	9,2	2,030	2,261
P3	21 28	9,1	2,041	2,296
P3	21 29	9,0	2,066	2,029
P3	21 30	9,7	2,010	2,186
P3	21 31	8,9	2,044	2,295
P3	21 32	9,2	2,041	2,060
P3	21 33	9,5	2,047	2,125
P3	21 34	9,2	2,043	2,273
P3	21 35	9,4	2,054	2,230
P3	21 36	8,9	2,053	2,294
P3	21 37	9,3	2,049	2,198
P3	21 38	9,2	2,048	2,303
P3	21 39	9,2	2,058	2,112
P3	21 40	9,3	2,060	1,927

juveniles stage - gonades

P4	16 CHNS 13	8,7	1,778	1,585
P4	16 CHNS 10	9,2	1,865	1,588
P4	16 CHNS 11	9,1	1,619	1,132
P4	16 CHNS 12	9,2	1,744	1,615
P4	16 CHNS 14	8,7	1,852	2,008
P4	16 CHNS 15	8,5	1,874	1,320
P4	16 CHNS 4	8,4	1,810	1,747
P4	16 CHNS 5	8,7	1,814	2,051
P4	16 CHNS 6	9,1	1,659	1,054
P4	16 CHNS 7	9,2	1,838	1,398
P4	16 CHNS 9	9,4	1,810	1,685
P4	16 CORT 10	8,8	1,812	1,900
P4	16 CORT 11	8,8	1,782	1,815
P4	16 CORT 12	9,0	1,925	1,524
P4	16 CORT 13	8,8	1,734	0,848
P4	16 CORT 14	9,8	1,746	0,648
P4	16 CORT 15	8,6	1,847	1,201
P4	16 CORT 5	9,9	1,731	1,114
P4	16 CORT 6	8,7	1,869	1,428
P4	16 CORT 7	8,3	1,632	1,097
P4	16 CORT 8	8,9	1,841	1,907
P4	16 CORT 9	9,2	1,831	1,092
P4	21 CHNS 10	9,2	1,944	1,266
P4	21 CHNS 11	9,1	1,808	1,926
P4	21 CHNS 12	8,6	1,858	1,874
P4	21 CHNS 13	9,4	1,825	1,884
P4	21 CHNS 14	9,2	1,857	1,729
P4	21 CHNS 15	8,7	1,814	1,591
P4	21 CHNS 5	8,8	1,763	1,559
P4	21 CHNS 6	8,4	1,767	1,829
P4	21 CHNS 8	8,5	1,792	2,072
P4	21 CHNS 9	9,0	1,826	2,134
P4	21 CORT 10	8,8	1,849	1,921
P4	21 CORT 11	8,5	1,892	0,716
P4	21 CORT 12	8,4	1,794	2,059
P4	21 CORT 13	8,9	1,773	1,758
P4	21 CORT 14	9,0	1,944	0,850
P4	21 CORT 15	9,0	1,812	1,813
P4	21 CORT 5	8,5	1,798	1,673
P4	21 CORT 6	9,3	1,852	1,304
P4	21 CORT 7	8,8	1,802	1,212
P4	21 CORT 8	8,8	1,782	1,964
P4	21 CORT 9	8,4	1,738	1,216

Legends for Datasets S1 to S4

Dataset S1: DESeq2 Normalized counts of fish at the flexion stage

Dataset S2: DESeq2 comparison between 16 and 21 genes differentially expressed for fish at the flexion stage

Dataset S3: DESeq2 Normalized counts of fish at the all fin stage

Dataset S4: DESeq2 Normalized counts of fish at the early juvenile stage

SI References

Sample References:

1. R. Griot, *et al.*, APIS: An auto-adaptive parentage inference software that tolerates missing parents. *Molecular Ecology Resources* **20**, 579–590 (2020).
2. B. J. Smith, boa: An R Package for MCMC Output Convergence Assessment and Posterior Inference. *Journal of Statistical Software* **21**, 1–37 (2007).
3. N. R. Wray, J. Yang, M. E. Goddard, P. M. Visscher, The Genetic Interpretation of Area under the ROC Curve in Genomic Profiling. *PLOS Genetics* **6**, e1000864 (2010).
4. X. Zhang, D. Lourenco, I. Aguilar, A. Legarra, I. Misztal, Weighting Strategies for Single-Step Genomic BLUP: An Iterative Approach for Accurate Calculation of GEBV and GWAS. *Frontiers in Genetics* **7**, 151 (2016).
5. P. M. VanRaden, Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science* **91**, 4414–4423 (2008).
6. D. A. L. Lourenco, *et al.*, Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus1. *Journal of Animal Science* **93**, 2653–2662 (2015).
7. S. Faggion, M. Vandeputte, B. Chatain, P.-A. Gagnaire, F. Allal, Population-specific variations of the genetic architecture of sex determination in wild European sea bass Dicentrarchus labrax L. *Heredity* **122**, 612–621 (2019).
8. A. Goikoetxea, *et al.*, Genetic pathways underpinning hormonal stress responses in fish exposed to short- and long-term warm ocean temperatures. *Ecological Indicators* **120**, 106937 (2021).
9. B. Geffroy, RNA-Seq juvéniles de bar: Projet 3S (Seabass, Sex and Stress). *Sextant* (2018) (November 26, 2021).
10. D. Kim, *et al.*, TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* **14**, R36 (2013).
11. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
12. M. Tine, *et al.*, European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nature Communications* **5**, 5770 (2014).
13. H. Li, *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
14. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
15. F. Krueger, S. R. Andrews, Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).
16. A. Akalin, *et al.*, methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology* **13**, R87 (2012).
17. F. Piferrer, M. Blázquez, L. Navarro, A. González, Genetic, endocrine, and environmental components of sex determination and differentiation in the European sea bass (Dicentrarchus labrax L.). *General and Comparative Endocrinology* **142**, 102–110 (2005).
18. M. Vandeputte, *et al.*, Low temperature has opposite effects on sex determination in a marine fish at the larval/postlarval and juvenile stages. *Ecology and Evolution* **10**, 13825–13835 (2020).