Transgenerational endocrine disruptor effects of cadmium in zebrafish and contribution of standing epigenetic variation to adaptation

Fabien Pierron, Guillemine Daffe, Flore Daramy, Débora Heroin, Aurélien Barré, Olivier Bouchez, Christelle Clérendeau, Alicia Romero-Ramirez, Macha Nikolski

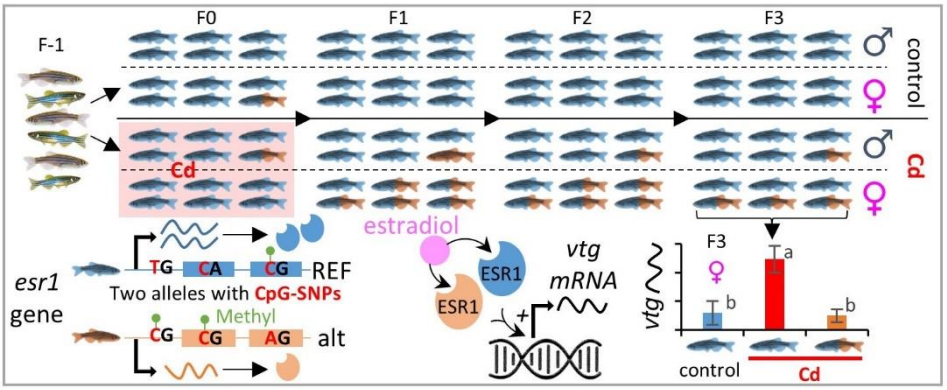
Title: Transgenerational endocrine disruptor effects of cadmium in zebrafish and contribution of standing epigenetic variation to adaptation

List of authors:
Fabien Pierron[a*], Guillemine Daffe[b], Flore Daramy[a], Débora Heroin[a], Aurélien Barré[c], Olivier Bouchez[d], Christelle Clérendeau[a], Alicia Romero-Ramirez[a] and Macha Nikolski[c,e]

Authors' affiliation adresses:
[a]Univ. Bordeaux, CNRS, Bordeaux INP, EPOC, UMR 5805, F-33600 Pessac, France
[b]Univ. Bordeaux, CNRS, INRAE, La Rochelle Univ., UMS 2567 POREA, F-33615 Pessac, France
[c]Univ. Bordeaux, Bordeaux Bioinformatics Center, Bordeaux, France
[d]INRAE, US 1426, GeT-PlaGe, Genotoul, Castanet-Tolosan, 31326, France
[e]Univ. Bordeaux, CNRS, IBGC, UMR 5095, Bordeaux, 33077, France

***Corresponding author:**
Fabien Pierron
UMR 5805 EPOC - OASU
Station Marine d'Arcachon
Université de Bordeaux
Place du Docteur Bertrand Peyneau
33120 ARCACHON
France
Phone : +335 56 22 39 33
fabien.pierron@u-bordeaux.fr
Orcid
Fabien Pierron 0000-0001-6176-3050

Graphical abstract



Abstract

Evidence has emerged that environmentally-induced epigenetic changes can have long-lasting effects on gene transcription across generations. These recent findings highlight the need to investigate the transgenerational impacts of pollutants to assess their long term effects on populations. In this study, we investigated the transgenerational effect of cadmium on zebrafish across 4 generations. A first whole methylome approach carried out on fish of the first two generations led us to focus our investigations on the estradiol receptor alpha gene (*esr1*). We observed a sex-dependent transgenerational inheritance of Cd-induced DNA methylation changes up to the last generation. These changes were associated with single nucleotide polymorphisms (SNPs) that were themselves at the origin of the creation or deletion of methylation sites. Thus, Cd-induced genetic selection gave rise to DNA methylation changes. We also analyzed the transcription level of various sections of *esr1* as well as estrogen responsive genes. While Cd triggered transgenerational disorders, Cd-induced epigenetic changes in *esr1* contributed to the rapid transgenerational adaptation of fish to Cd. Our results provide insight into the processes underpinning rapid adaptation and highlight the need to maintain genetic diversity within natural populations to bolster the resilience of species faced with the global environmental changes.

Key words: Transgenerational inheritance, DNA methylation, endocrine disruptor, rapid adaptation, genetic selection

Environmental implication

Emerging evidence supports the fact that pollutants can have long-lasting effects on the phenotype of organisms and their offspring by inducing changes in epigenetic marks. Thus, short term exposures limited to specific life stages appear insufficient to assess the long term effects of pollutants on populations. Epigenetics may help to develop new biomarkers that can link molecular events to population-level consequences. In our study, a first generation of zebrafish was life-long exposed to an environmentally relevant concentration of cadmium, a widespread non-essential metal with endocrine disruptive effects. We analyzed DNA methylation changes associated with zebrafish phenotypes across 4 generations.

1. Introduction

Earth is experiencing rapid environmental changes due to human activities. It has become important to determine how organisms and their offspring will respond to such changes. Increasing evidences support the fact that environmentally induced epigenetic changes can be inherited across generations, leading to heritable phenotypes and transgenerational diseases (Xin et al., 2015; Nilsson et al., 2018; Cavalli and Heard, 2019; Fitz-James and Cavalli, 2022). In an ecotoxicological context, these recent findings undoubtedly highlight the need to assess the multi- and transgenerational impacts of pollutants to get a better understanding of their long term effects on organisms and populations (Vandegehuchte and Janssen, 2014).

Epigenetics can be defined as "the study of molecules and mechanisms that can perpetuate alternative genetic activity states in the context of the same DNA sequence" (Cavalli and Heard, 2019). Among the various epigenetic mechanisms, DNA methylation refers to the addition of methyl groups to cytosines and is commonly associated with gene repression (Moore et al., 2013). Despite the fact that much more emphasis was put on disease etiology (particularly in vertebrates (Skinner et al., 2010; Nilsson et al., 2018)), epigenetics, by being at the interface between the environment (that is changing) and the genome (that must be more stable), can also allow rapid adaptation to environmental changes by acting on the phenotype of organisms (as documented in plants (Thiebaut et al., 2019)). Thus, epigenetics could also play a non-negligible role in the persistence of populations in changing environments by providing a mechanism of phenotypic plasticity, i.e. the ability of a genotype to generate different phenotypes (Duncan et al., 2014; Kilvitis et al., 2017). Despite being well recognized in plants, transgenerational epigenetic inheritance is still a matter of debate in vertebrates, especially in mammals (Heard and Martienssen, 2014; Horsthemke, 2018). First, whereas there is abundant evidence for intergenerational epigenetic inheritance (from F0 to F1), examples of true epigenetic inheritance remain scarce in vertebrates (Fitz-James and Cavalli, 2022). Indeed,

as individuals of the F1 generation are exposed to the stressor as gametes in their parents (F0), true epigenetic inheritance only arises when the signal is maintained at least in the F2/F3 generations (Nilsson et al. 2018; Fitz-James and Cavalli, 2021). This appears especially important in an ecotoxicological context as contaminants can be themselves transferred between the F0 and F1 generation via the contaminants accumulated in the parental gametes (Pierron et al., 2008). Moreover, it still remains hard to prove whether certain phenotypic variations have a purely epigenetic origin and are not the consequence of genetic variations (Heard and Martienssen, 2014; Fitz-James and Cavalli, 2022), highlighting the need to investigate both genetic variations and epigenetic changes during transgenerational studies. Further complicating the matter, the epigenome is not constant throughout an animal's life but undergoes changes at defined stages of development which raises the question of how stress-induced changes in epigenetic marks can be transgenerationally inherited (Heard and Martienssen, 2014; Zeng and Chen, 2019).

In this context, the aims of the present study were to test in a vertebrate model whether (i) pollutant-induced DNA methylation changes can be transmitted to the unexposed offspring at least until the fourth generation and (iii) such epigenetic modifications can be linked to adverse and/or adaptive phenotypic changes at different levels of biological organization, from gene expression to population. To test these hypotheses, we performed a transgenerational experiment comprising 4 generations of zebrafish. Only fish from the first generation (F0) were exposed throughout their life to an environmentally realistic concentration of cadmium (Cd), a widespread nonessential metal with endocrine disrupting effects (Takiguchi and Yoshihara, 2006). Cd is an element naturally found in the environment. Despite environmental concentrations of Cd vary widely, concentrations in freshwaters are typically below 0.1 $\mu g.L^{-1}$ but can exceed 1 $\mu g.L^{-1}$ in industrialized and mining-impacted environments (Giguère et al., 2004; Alves Peixoto and Jadán-Piedra, 2022). Previous works carried out in diverse species,

from plants to humans, suggested that Cd is able to alter DNA methylation marks (Wang et al., 2012; Pierron et al., 2014; Niekerk et al., 2021; Saintilnord et al., 2021) with sex-specific effects reported in humans (Kippler et al., 2013). Some classical endpoints were measured at each generation such as fish length, weight, mortality, sex ratio and fecundity. Cd exposure triggered transgenerational growth disorders (see Pierron et al., 2022) and led to a progressive transgenerationnal feminization of the population (see Pierron et al., 2021). To investigate the potential Cd-induced DNA methylation changes, we used two complementary approaches. First, in order to identify without *a priori* the regions of DNA that were differentially methylated (called Differentially Methylated Regions, DMRs) in response to Cd exposure, we used a whole genome analysis on fish of the first two generations (F0 and F1). Among the Cd-induced DMRs, three DMRs located along the *esr1* gene were identified by MeDip-Seq in both larvae and females. The *esr1* gene encodes for the estradiol receptor alpha, a receptor activated by the primary female sex hormones, estrogens. ESR1 is essential for sexual development and reproductive function (Chen et al., 2018; Wu et al., 2020). As (i) Cd exposure led to a transgenerational feminization of the population (Pierron et al., 2021) and (ii) Cd is already known to have a potent estrogen-like activity (Johnson et al., 2003), *esr1* appeared as a good candidate gene to link (i) epigenetic changes to inherited phenotypes and (ii) molecular events to population consequences. Thus, we then focused our investigations on this gene by analyzing more precisely its methylation pattern throughout the 4 generations by using a targeted and base-resolution method. Finally, we assessed the transcription levels of various sections of the *esr1* gene as well as targets of ESR1.

2. Methods

2.1 General rearing and breeding program

At the beginning of the experiment, we used 40 randomly mating pairs of wild-type (domesticated) zebrafish. Reproduction took place in small breeding tanks (BT, 2L) and 20 BTs were filled with clean water (Control, C) while 20 other BTs were filled with Cd-contaminated water at a nominal concentration of 1 µg.L^{-1}. Their offspring (F0 generation) were respectively reared in clean or Cd-contaminated water throughout their life (0-169 dpf). Cd concentration in water was determined by electrothermic atomic absorption spectrophotometry using a graphite tube atomizer (GTA120 Agilent). The average concentrations were 0.049 ± 0.004 µg.L^{-1} and 0.987 ± 0.025 µg.L^{-1} (mean ± SE, n = 260) in control and contaminated tanks, respectively. Control and Cd fish were sampled at 5 days post-fertilization (dpf, larvae) and at 124 dpf (adult mature fish). Whole larvae and gonads, liver and brain from adults were fixed in RNAlater solution before being stored at -20°C until analyses. At 170 dpf, we used 20 pairs per condition to produce the next generation (F1). Mating was carried out in Cd-free water for both experimental groups (C or Cd). Pairs were not randomly selected to limit family/genetic biases. For each condition, fish were reared in 4 different large tanks (75L). As all individuals from one pair were raised in the same large tank, one male from one tank was mated with a female from another tank to avoid inbreeding. The same protocol was used for mating and sample collection for the F1, F2 and F3 generations. For both groups of animals (C and Cd), spawning and rearing were carried out in Cd-free water. The experimental set-up is presented in Fig.1 and a detailed description is available in Pierron et al. (2021).

At each generation, all larvae were measured (5 dpf) and all adults sampled at 124 dpf as well as all remaining adults after mating (170 dpf) were measured, weighted and sexed. Moreover, some individuals were randomly sampled at different time points to investigate for potential stage-dependent Cd effects on fish growth. Mortality was recorded daily throughout the experiment. Spawning success and fecundity was measured at each generation. No

significant differences were observed regarding the survival, spawning or fecundity rates between the control and Cd group (whatever the generation considered).

All procedures used in this experiment were approved by the Aquitaine fish-birds ethic committee (APAFIS#7535-2016111009351504).

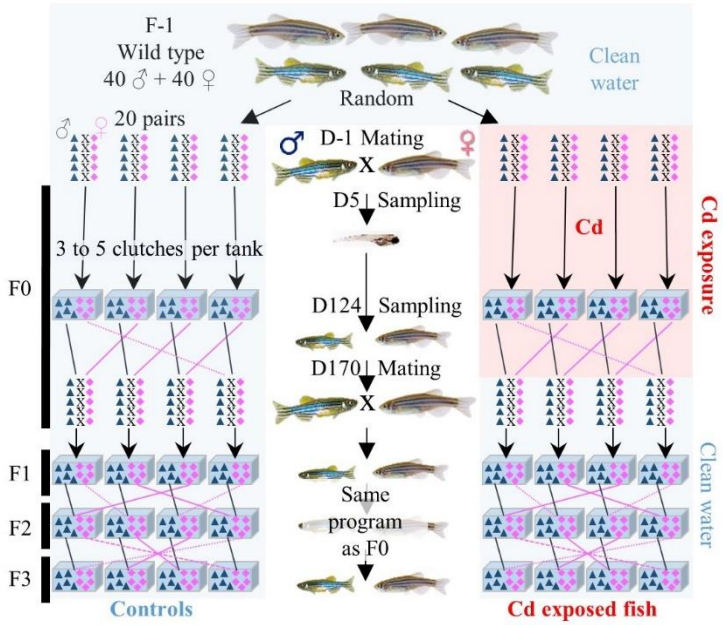


Figure 1: Synoptic of the experimental design
Only fish of the F0 generation were exposed throughout their life to cadmium (Cd). At each generation and for each condition (control or Cd), 20 mating pairs (pairwise mating) and 4 large tanks were used. Whereas random mating was used to produce the F0 generation, for the generations F1 to F3 males from one tank were mated with females from another tank to limit inbreeding and potential genetic/family biases.

2.2 RNA/DNA extraction

For each generation and condition, we used 6 biological replicates. In order to limit potential genetic and sampling biases, we used 6 larvae (5 dpf) from 6 different mating pairs and 6 males (124 dpf) and 6 females (124 dpf) from 3 different tanks (i.e. individuals from different parents). For larval samples used for the MeDip-Seq analysis, we used 6 pools of 2 larvae in order to obtain a sufficient amount of DNA (see Pierron et al. 2022). Total RNA and DNA were purified using the AllPrep DNA/RNA kit (Qiagen) as described in Pierron *et al.* (2019).

2.3 Identification of intergenerational Cd-induced DMRs by MeDip-Seq

DNA samples from larvae and adult zebrafish of the F0 and F1 generation were analyzed by MeDip-Seq, for the total of 168 samples. Library preparation, quality checks, sequencing and bioinformatic analyses were already described in Pierron *et al.* (2022). Briefly, libraries were prepared using the Biooscientific NEXTflex™ Methyl-seq Library Prep Kit before to be immunoprecipitated using the MagMeDIP Kit (Diagenode) and sequenced on two lanes of an Illumina NovaSeq 6000. Raw sequencing data were deposited in the European Nucleotide Archive with accession number PRJEB52137. Sequences were mapped to the *Danio rerio* reference genome GRCz11 and DMRs were identified using the MEDIPS package.

2.4 Analysis of methylation levels and single nucleotide polymorphisms

Genomic sequences were obtained from the Genome Reference Consortium Zebrafish 11 (GRCz11). Specific primers used for PCR amplification and bisulfite sequencing were designed in the promoter (pro) region, exon 3 (ex3) and 5 (ex5) of the *esr1* gene by means of the Pyromark assay design software (Table S1).

For each sample, quantification of single cytosine percent methylation at specific sites was performed as previously described in Pierron *et al.* (2019) by bisulfite-pyrosequencing using the PyroMark Q48 Autoprep instrument (Qiagen). SNP detection was carried out during the run of bisulfite pyrosequencing by modifying the analyzed sequence (AQ mode, Fig. S1). In order to assess the conversion efficiency of the bisulfite treatment, the average methylation level of one CpT site was measured in ex3. Its methylation level was < 4 %, indicating very good conversion efficiency. The methylation level of at least 3 different CpG sites was quantified in the different sections under analysis. For each section, the mean methylation level of all CpG sites was used for statistical analyzes. Examples of pyrograms are provided in Fig. S1.

2.5 Cloning and sequencing of genetic variants

Since (i) the bisulfite treatment converts unmethylated cytosines into thymines after PCR and (ii) the SNP identified in the *esr1* promoter corresponded to the transition TG→CG, we also amplified and sequenced the promoter region of *esr1* from some individuals using native (untreated DNA) genomic DNA as template in order to dissociate genetic *versus* epigenetic changes. Specific primers were designed using the primer3plus software (Untergasser et al., 2007) or the Pyromark assay design software (Table S1). Amplified products (Pyromark PCR kit Qiagen) were sequenced by pyrosequencing (using the PyroMark Q48 Autoprep instrument, Fig. S2) or cloned into pGEM-T Easy vector according to the manufacturer's instructions (Promega) before to be sequenced by Sanger (GATC Biotech).

2.6 Quantitative analysis of transcription levels

RT-qPCR analyses were carried out as previously described in Pierron *et al.* (2022). First-strand cDNA was synthesized from 500 ng of total RNA from female liver or gonads and cDNA were diluted 40-fold. Relative quantification of gene transcription was achieved by concurrent amplification of the *eef1a1a* (see Pierron et al. 2022 for details). Amplification efficiencies for all primer sets were calculated and used to determine transcription levels of targets. Primers used are available in Table S1.

2.7 Statistical analyses

Comparisons among fish groups were performed by two-way analysis of variance (ANOVA), after checking the assumptions of normality (Kolmogorov-Smirnov) and homoscedasticity of the error terms (Levene). If significant effects were detected, the LSD test was used to determine whether means between pairs of samples were significantly different. When the assumptions were not met as deduced from the ad-hoc tests, we used box-cox data

transformations or the nonparametric Kruskal−Wallis test. Using Kruskal−Wallis, the post-hoc Conover-Iman test was used. Comparisons of proportions were performed using the two proportions Z-test. Comparisons of allelic frequencies between groups were carried out by means of the homogeneity chi-square test. Comparisons of slopes were realized using a Student's t-test. Computations were performed using STATISTICA version 6.1 software (StatSoft) and XLSTAT (Addinsoft version 2020.1.1). Numerical results are provided as mean ± SE (standard error).

3. Results and Discussion

3.1 Genome-wide identification of Cd-induced differentially methylated regions

In order to identify the DNA regions that could be differentially methylated in response to Cd exposure, we first used the MeDip-Seq approach on both larvae and adult zebrafish of the F0 and F1 generation. We used a MeDip-Seq (Methylated DNA immunoprecipitation coupled with high-throughput Sequencing) approach rather than WGBS (Whole-Genome Bisulfite Sequencing) to obtain a better coverage of the genome and maximize the number of analyzed samples (Beck et al., 2022), allowing to investigate two stages of development, both sexes, and different organs. Only few Cd-induced DMRs were common to the two generations (i.e. carrying the intergenerational inheritance). Moreover, the number of Cd-induced DMRs was greater in F1 fish (i.e. not directly exposed to Cd) in comparison to F0 fish (i.e. directly exposed to dissolved Cd). For example, while 3908 Cd-induced genome-wide DMRs were identified in larvae of the F1 generation, only 800 were detected in larvae of the F0 generation (Fig. 2). The main result was that more than 12 % of the Cd-induced DMRs identified in larvae of the F1 generation were found in the three organs (liver, brain and gonads) of F1 adult females

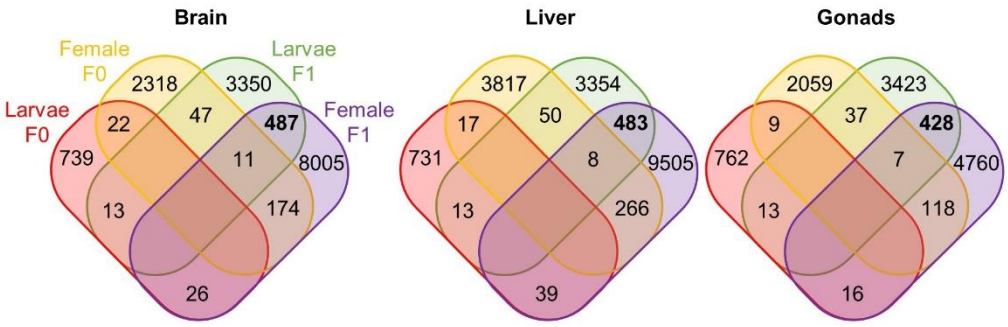


Figure 2: Cd-induced DMRs identified by genome-wide DNA methylation analysis in females and larvae

Venn diagram presenting the number of Cd-induced differentially methylated regions (DMRs) in larvae (whole individual) and female (brain, liver and gonads) zebrafish of the F0 and F1 generation. For each generation, organ or stage of development, Cd-induced DMRs were identified by comparing Cd-exposed fish to their respective controls (MEDIPS, $P < 0.01$). Numbers in bold correspond to the DMRs that were common to larvae and females of the F1 generation (ranging from 428 to 487), 300 of them were found to be common to the 3 organs.

(reaching only 2 % in the case of males). Moreover, among the DMRs that were common to both larvae and females of the F1 generation, more than 60 % were common to the three organs under study, suggesting that Cd-induced DMRs at the larval stage can be "propagated" and conserved in different fully developed organs of females (Fig. 2). Among them, three different DMRs were identified along the *esr1* gene, encoding for the estrogen receptor alpha. These DMRs overlapped the exon 3-intron 3 (position 26377141-26377825 in NC_007131.7, GRCz11), ex5-in5 (26368432-26369194) and ex6-in6 (26365426-26366936) regions. We then decided to assess the methylation level of these exons in zebrafish from all 4 generations by means of a targeted and base-resolution method, i.e. targeted bisulfite pyrosequencing (BS-Seq). However, we failed to develop sequencing primers for the ex6. As (i) promoter DNA methylation is an important regulatory component of vertebrate gene expression (Lou et al., 2014) and (ii) previous works reported that changes in the promoter methylation level of *esr1* can be linked to endocrine diseases (Tsuboi et al., 2017; Mastoraki et al., 2018), we also determined the methylation level of the *esr1* promoter.

3.2 Changes in methylation levels were associated with synonymous SNPs

We initially encountered some difficulties with the targeted BS-Seq method as poor quality results were obtained for some individuals. After investigation, this was explained by the presence of SNPs. We then modified the sequencing parameters (Fig. S1) to analyze both DNA methylation levels of individual CpG sites as well as the SNPs (i.e. genotyping) during the same BS-Sequencing run.

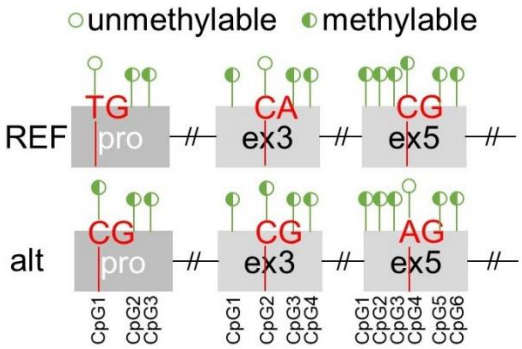


Figure 3: The two alleles of the *esr1* gene
Schematic representation of the CpG methylation sites (green) and single nucleotide polymorphisms (SNPs, red) analyzed by targeted bisulfite sequencing in the promoter region (pro), exon 3 (ex3) and exon 5 (ex5) of *esr1* for the reference (REF) and alternative (alt) allele.

We identified a SNP in both exons 3 and 5 (ex3, position 26377759; ex5 pos. 26368809 in NC_007131.7, GRCz11, Fig. 3). In both cases, while these two SNPs do not trigger change in the protein sequence (i.e. synonymous SNPs), they both affected the mean methylation level of the corresponding exon. Indeed, in the ex3, the A base, that is referenced in the RefSeq sequence (NC_007131.7, GRCz11), was replaced by a G base. The presence of the alternative G base is at the origin of an additional site of methylation as this SNP is preceded by a C (CA→CG), methylated cytosines being primarily found at the location of CpG dinucleotides (Feng et al., 2010; Yong et al., 2016). Inversely, the SNP observed in the ex5 triggered the loss of a CpG site (CG→AG, Fig. 3). These genetic changes, called CpG-SNPs, thus affected the mean methylation level of the exons and significant differences in the methylation level of both ex3 and 5 were observed between control and Cd females from the F1 to the F3 generation (Fig. 4A,B). While methylation levels were significantly higher in Cd females in comparison to

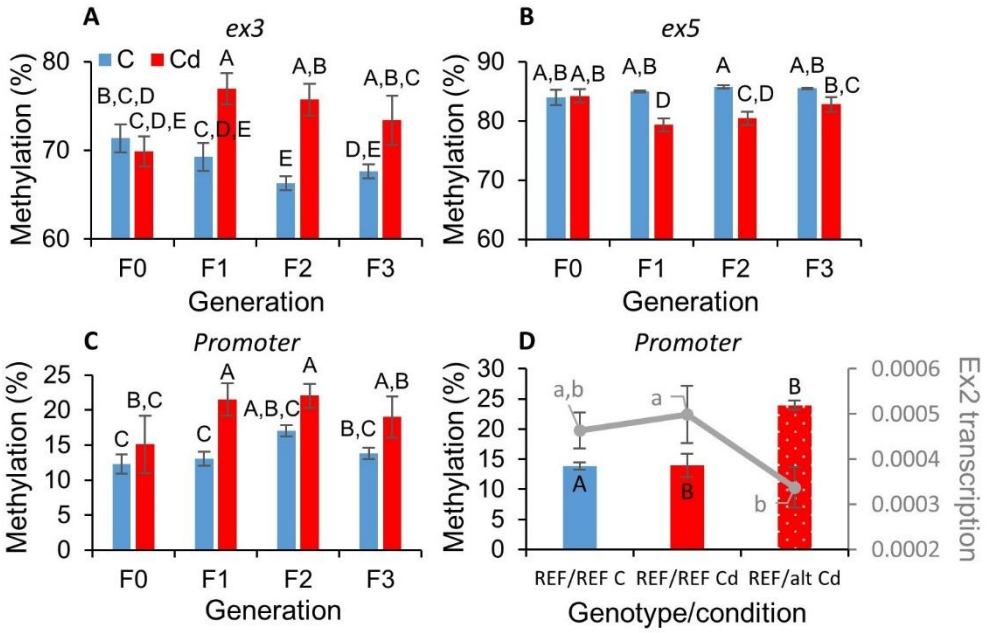


Figure 4: Change in the methylation level of *esr1* was associated with single nucleotide polymorphisms.
(**A**,**B**,**C**) Change in the methylation level (mean ± SE, n = 5-6) of the (**A**) exon 3 (ex3), (**B**) exon 5 (ex5) or (**C**) promoter region in gonads of female zebrafish according to the condition (Control, blue; Cadmium, Cd, red) and generation (F0-F3). For each variable, means designated by different letters are significantly different (LSD, P <0.05). (**D**) Change (mean ± SE, n = 10-23) in the promoter methylation level (bars) and in the transcription level (curves, a.u.) of the exon 2 (ex2) in gonads of female zebrafish according to the genotype (REF/REF or REF/alt) and the condition (C, control; Cd, cadmium). Data from the four generations were combined. No homozygous alt/alt females were observed and only control female was heterozygous. For each variable, means designated by different letters are significantly different (LSD, P <0.05).

controls for the ex3, an inverse trend was observed for the ex5. Interestingly, we also observed an SNP in the promoter region (pos. 26383521 in NC_007131.7, GRCz11). Again, this SNP was related to a CpG site (TG→CG, Fig. 3). However, in this latter case, a significant effect of Cd on the promoter methylation level was observed only at the F1 generation (Fig. 4C). Interestingly, when the genotype (T/T or T/G referred as REF/REF and REF/alt, respectively) of individuals is considered, significant differences were observed in the methylation level of the promoter region (Fig. 4D). This was observed either from the methylation level determined from the 3 CpG sites or from the CpG#2-3 (i.e. by excluding the CpG-SNP, Fig. S3), suggesting that an additional CpG site can affect the methylation level of the neighboring CpG sites. As

all the identified SNPs trigger DNA methylation changes, these changes will be thereafter referred as "epi-genetic" variations to underline the fact that changes in DNA methylation levels are supported by genetic variations.

We must add that when individuals (either females, larvae or males) were heterozygous for the ex3, they were also, and systematically, heterozygous for the ex5 and the promoter region. Moreover, all these SNPs were found to be present in the alternative sequence NW_018395074.1 in GRCz11 (in the alternate locus group ALT_DRER_TU_1 ALT_CTG20_1_1). Thus, the observed SNPs were not just point mutations, but rather corresponded to the presence of two alleles. In the present article they are referred as "REF" for the allele corresponding to the RefSeq (NC_007131.7) and as "alt" for the alternative allele (Fig. 3). We must add that the sequence of the alt allele in our data showed some differences, at least in the promoter region, with respect to the alternative sequence NW_018395074.1 in GRCz11 (Fig. S4). Indeed, using native (non-bisulfite-treated DNA) genomic DNA as template, we found that the alt sequence identified in our work presented one SNP and one deletion (3 bases) compared to the alt GRCz11 sequence and presented higher homologies with the sequences from the wild-derived strains Nadia (NA) and especially Cooch Behar (CB); reaching 100 % homology with CB (Fig. S4). Moreover, in the case of CB and NA, *esr1* sequences of both ex3 and ex5 contain the alternative SNPs (i.e. C and A for ex3 and 5, respectively, Fig. 3 and S4).

3.3 Cd-induced epi-genetic changes were sex-dependent

In order to get more insight into the effect of Cd on fish genotypes, we then analyzed the percentage of REF/REF animals according to the condition (C or Cd), sex/stage of development and generation (Fig. 5A). Despite the low number of biological replicates, some

significant differences were observed. The percentage of homozygous REF/REF females was significantly lower in Cd in comparison to the control condition from the F1 until the F3 generation. In contrast, no difference was observed between control and Cd males throughout the 4 generations. More surprisingly, a significant effect of Cd was observed on larvae from the first (F0) generation. As for females, the percentage of REF/REF larvae was significantly lower in Cd in comparison to the control condition for generations F0 through F2. In addition, considering animals from all 4 generations, we found a significant difference in genotype frequencies between Cd females and larvae on the one hand and Cd males and control individuals (females, males and larvae) on the other hand (Fig. 5B). Cd females and Cd larvae were preferentially heterozygous (% of REF/alt reaching 56.2 % (n = 47) and even 66.7 % when considering individuals from the F1 to the F3 generation (n = 36)), whereas almost all control individuals were homozygous REF/REF (% of REF/REF reaching 97.2 %, n = 72). Cd males were mostly homozygous REF/REF (% of REF/REF reaching 83.3 %, n = 24). Moreover, only one Cd male and one Cd larva were homozygous alt/alt. Thus, the Cd effect was sex-dependent and was associated with an increased heterozygosity in females and larvae, but not in males. We have to note that a particular effort was carried out to limit as much as possible potential genetic/family biases. At each generation, we used 20 different mating pairs for each condition. From the F1 generation, spawning pairs were not random to avoid inbreeding (Fig. 1) and individuals from different parents were sampled and analyzed at each generation, thus strengthening the link between Cd and genetic changes. The most surprising result is that Cd exposure affected the genotype of larvae from the first F0 generation, whereas such an effect was observable only starting from the F1 generation in the case of females. We therefore analyzed more precisely the effect of Cd on mortality and fitness- related traits in the F0 generation.

3.4 Cd exerted an epi-genetic selection

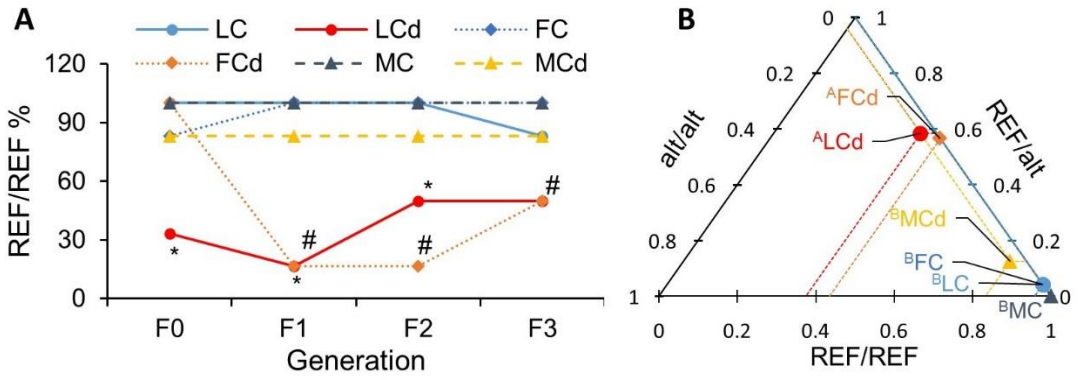


Figure 5: Cd exposure induced sex-dependent transgenerational changes in allele frequencies. (**A**) Change in the percentage of homozygous REF/REF fish (control, C or cadmium, Cd) at each life stage (larvae (L), adult males (M) and females (F)) across generations (F0-F3). For each generation, symbols denote a significant effect of Cd exposure on larvae (*) or females (#), Z-test (n = 6, $P < 0.05$). (**B**) Change in genotype frequencies according to the life stage, sex and condition. Data from the 4 generations were combined to enlarge sample size. Different letters denote significant differences, homogeneity chi-square test ($P < 0.05$).

The frequency (f) of the alt allele in the F0 generation decreased from the larval (f_{alt} = 0.42) to the adult (f_{alt} = 0.04) stage (Fig. 5A), suggesting an effect of selection on the alt allele during zebrafish development. Thus, while the alt allele was initially favored by Cd at early stages of development, it became disadvantageous thereafter. In other words, epi-genotype-dependent mortalities were temporally dynamic (Durland et al., 2021). In support of this assumption, while no difference was observed between the mortality rates of control and Cd fish at the F0 generation (and at each generation thereafter), temporal variations were observed. In general, the highest mortalities were observed during early development (0-14 dpf, Fig. S5). Mortality rates did not exceed 1% at the juvenile-adult stage, but cannibalism was observed until 60 dpf (Fig. S5). Whereas Cd exposure was initially associated with lower mortalities in comparison to controls at the embryo stage (1-2 dpf), Cd triggered a significant increase in mortality from 8 to 11 dpf (i.e. larval stage, Fig. 6A). These mortalities coincided with the stage of gonad differentiation (Tzung et al., 2015; Ortega-Recalde et al., 2019) and more precisely

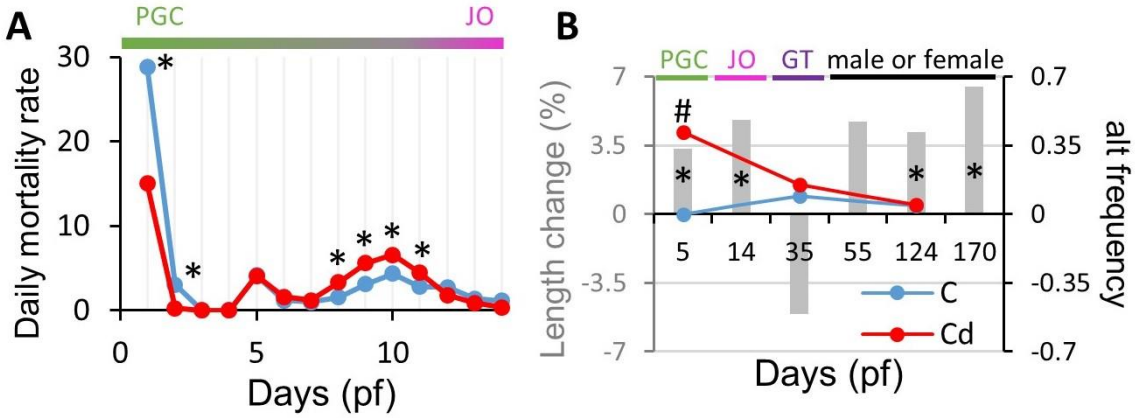


Figure 6: Cd-induced effects on fish mortality and growth varied according to the developmental stage.
(**A**) Daily mortality rate (%) of control (C, blue) and cadmium (Cd, red) fish of the F0 generation from 0 to 14 dpf. For each day, * denotes significant difference between C and Cd (Z test, $P < 0.01$). (**B**) Cd-induced changes in total fish length of the F0 generation (bars, expressed as percentage as compared to controls, n = 48-1014) and change in the frequency of the alternative allele (curves, n = 6-27) according to fish age. * denotes a significant effect of Cd on length (LSD, $P < 0.05$) and # on the frequency of the alt allele (Z-test, $P < 0.05$). (**A,B**) Timeline of gonad differentiation is reported : primordial germinal cells determination and growth (PGC, 0-11 dpf), "juvenile ovary" (JO, 11-21 dpf) and gonad transformation (GT, 21-45 dpf).

with the differentiation of primordial germ cells (PGCs) into oocytes. In wild-type zebrafish, the ovarian development is considered to be the default pathway and differentiation of PGCs into oocytes occurs from 7 to 11 dpf (Uchida et al., 2002; Orban et al., 2009; Tzung et al., 2015). Thereafter, juveniles (from 11 to 21 dpf) develop an immature ovarian tissue (called "juvenile ovary"), regardless of their genetic background (Liew and Orban, 2014; Santos et al., 2017). After 21 dpf, the ovarian tissue is either maintained or eliminated by apoptosis, allowing testis development. In general, gonad transformation (GT; juvenile ovary → testis) begins from around 21 dpf to be completed at 45 dpf.

To get more insight, we then investigated the length of fish as a fitness-related trait (Bitter et al., 2019). Again, temporal changes were observed. While Cd significantly boosted the growth of F0 fish at both the embryo-larval and adult stages (see Pierron et al., 2022), at 35 dpf, when GT occurs, Cd negatively affected the growth of juvenile fish (or at least Cd did not

significantly boost the growth of fish as this negative effect was not statistically significant; Fig. 6B). Accordingly, Brion et al. (2004) have reported that whereas exposure to estradiol at either the embryo-larval (0-21 dpf) or adult stage (85-160 dpf) does not affect the length of fish, exposure to estradiol at the juvenile stage (21-42 dpf) on the contrary triggers a decrease in fish length. Moreover, using an agonist of estrogen receptors, Maack and Segner (2004) showed that the GT is the most susceptible stage to estrogen exposure. At this stage (35 dpf), the frequency of the alt allele in Cd fish had already decreased to a level that was almost comparable to that of controls (Fig. 6B).

Thus, change in the frequency of the alt allele in Cd-exposed fish of the F0 generation was associated with significant temporal changes in embryo-larval mortality and juvenile growth rates; these changes coincided with key events of sexual determination. The pattern was however different in the following generations. Indeed, from the subsequent generation F1, an equal frequency of the alt allele was observed in Cd larvae (f_{alt} = 0.417) and Cd females (f_{alt} = 0.417; Fig. 5A), suggesting that the alt allele remains more advantageous throughout the development of zebrafish, especially in the case of females. A main difference was that zebrafish from the F1 generation were not exposed to dissolved Cd. In view of these considerations, several studies reported that the potent estrogen-like activity of Cd comes from, at least in part, its capacity to bind with high affinity to different domains of ESR1, affecting directly the function of ESR1 (Stoica et al., 2000; Chouchene et al., 2016).

3.5 Cd-induced epi-genetic changes in females were adaptive

In order to get information on the potential impacts associated with synonymous CpG-SNPs on the phenotype of individuals, we then investigated the effect of epi-genetic changes on gene transcription, an intermediate between epigenetic variation and complex diseases or adaptive processes. In accordance with the consensus view that promoter methylation leads to

gene repression in vertebrates, the hypermethylation of the *esr1* promoter observed in heterozygous Cd females was associated with a significant decrease in the transcription level of the ex2 of *esr1* (Fig. 4D). The transcription level of the ex2 in female gonads was significantly lower in REF/alt Cd individuals (promoter methylation = 24.0 ± 0.7 %, mean ± SE, n = 13) in comparison to REF/REF Cd individuals (promoter methylation = 14.0 ± 2.0 %, mean ± SE, n = 10). In order to reinforce this first result, we then investigated the transcription levels of various sections of *esr1* as methylation of exons can induce alternative splicing events (Shayevitch et al., 2018). We also investigated the main target of ESR1, namely genes encoding for vitellogenins (*vtg*$_s$, the *vtg* repertoire of zebrafish including 8 different genes (Chen et al., 2019)). VTGs are specialized transfer proteins that are produced by the liver and transported via the bloodstream to the ovary where they are catabolized to form the egg yolk. ESR1 is considered to be the major mediator of vitellogenesis in zebrafish (Chen et al., 2019). Analyses were thus carried out on female liver. In the case of controls, for which 96 % of females were homozygous REF/REF, highly significant correlations were observed among the different sections of *esr1* as well as with the different genes encoding for VTGs (Fig. 7A). While similar results were observed for heterozygous REF/alt Cd fish, only few significant correlations were observed for homozygous REF/REF Cd fish. In addition, even when significant correlations were observed, the magnitude of the relationship was different between REF/REF Cd and fish from the other conditions (Fig. 7B). For example, despite the fact that significant correlations were observed between the transcription level of the ex 2 and 5 of *esr1* for all genotypes and conditions (control and Cd), the slope coefficient of the linear relationship was significantly different between REF/REF Cd fish and REF/alt Cd or control fish. No difference was observed between REF/alt Cd fish and controls.

In order to investigate more precisely the transgenerational effect of Cd-induced epigenetic changes on gene transcription, we take advantage to the fact that Cd female zebrafish

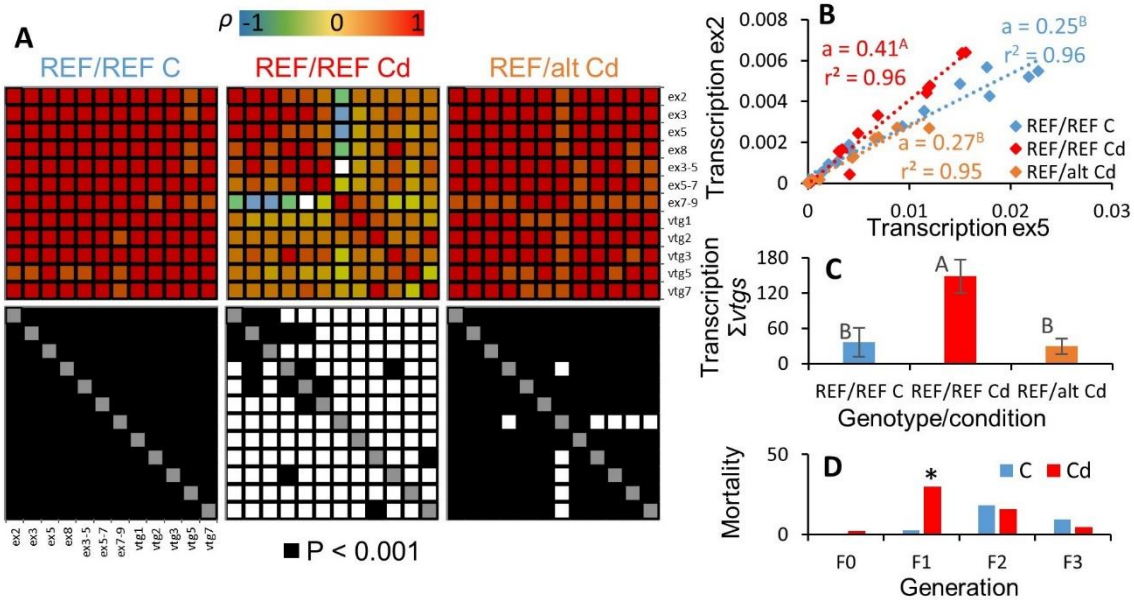


Figure 7: Heterozygosity enabled to restore transcriptional liver homeostasis in response to Cd. (**A**) Correlation matrices obtained from REF/REF control (C, n = 23), REF/REF cadmium (Cd, n = 10) or REF/alt cadmium (Cd, n = 13) fish. Transcription level of different sections of the *esr1* gene (ex2,3,5,8,3-5,5-7 and 7-9) and different genes encoding for vitellogenins (*vtg*1,2,3,5 and 7) were analyzed in female liver. Data for all the 4 generations were combined to enlarge the sample size. The upper panels refer to the Spearman coefficients (ρ) and lower panels to the associated P values. (**B**) Relationship between the transcription level of the ex2 and ex5 in female liver for each genotype (REF/REF or REF/alt) and condition (C or Cd). Data for all the 4 generations were combined. The slope coefficients (a) and Pearson correlation coefficients (r^2, P < 0.0001) are reported. The slope coefficients indicated by different letters are significantly different (Student's t-test, P < 0.01). (**C**) Change in the sum of the transcription levels (mean ± SE, n = 3-6) of the different genes (*vtg*1,2,3,5 and 7) encoding for vitellogenins in female liver of the F3 generation according to the genotype (REF/REF or REF/alt) and the condition (C or Cd). Means designated by different letters are significantly different (LSD, P <0.05). (**D**) Mortality (%) induced by handling stress at the larval stage (5 dpf) according to the generation and condition. * denotes a significant effect of Cd (Z-test, P < 0.05).

of the F3 generation were equally composed of REF/REF and REF/alt individuals. As presented in figure 7C, after 4 generations, REF/REF Cd fish presented a transcription level of vtg_s genes that was significantly 4.6-fold higher in comparison to REF/alt Cd or control fish. Again, no difference was observed between REF/alt Cd and REF/REF controls. Thus, in Cd-exposed fish, heterozygosity allowed to restore transcription levels close to that of controls. Conversely, this result highlights a transgenerational effect of Cd on vtg_s transcription levels in REF/REF Cd females. Thus, notwithstanding Cd induced transgenerational disorders until the F3 generation,

heterozygosity allowed fish to adapt and to maintain "normal" transcription levels of both *esr1* and vtg_s genes as well as their homeostasis.

Previous works reported that change in maternal VTGs can affect the quality and survival of zebrafish eggs (Yilmaz et al., 2021) and larvae (Yilmaz et al., 2019). Despite no observed difference in embryo-larval mortality rates between controls and Cd at each generation, we observed that the offspring (F1 larvae) of the F0 Cd females (i.e. when Cd females where all homozygous REF/REF; Fig. 5A) were significantly more sensitive to handling/"predator" stress than controls. At 5 dpf, some larvae were sampled to carry out behavioral tests (data not shown). During the sampling, we noticed a significantly higher mortality of F1 Cd larvae in comparison to controls (Fig. 7D). Such an effect of Cd was no longer observed at the F2 and F3 generation, i.e. when the mothers were mostly heterozygous (% of REF/alt Cd females reaching 83.3% at both the F1 and F2 generation), reinforcing the fact that Cd-induced epi-genetic changes were adaptive.

3.6 CpG-SNPs supported a rapid adaptation to Cd

We observed drastic and sex-dependent genetic changes in response to Cd exposure. Cd induced a significant increase in the frequency of the alt allele in larvae and females. Considering all genotyped individuals (larvae and adults of all the 4 generations, n = 72), the frequency of the alt allele reached 0.014 in the control condition. Thus, the alt allele pre-existed and was maintained at a low frequency throughout the experiment; heterozygous control individuals were observed at both F0 and F3 generation (Fig. 5A). As firstly evoked, the alt allele shows high homology with the sequences from the wild-derived zebrafish strains NA and CB. NA is considered as a newest laboratory strain (8^{th} generation in captivity) and CB as a wild-derived (2^{nd} generation in captivity). Zebrafish domestication (> 30 generations in captivity) led to a severe reduction in heterozygosity and a loss of (i) genetic polymorphism,

(ii) GC-biased gene conversion and (iii) genetic sex determinant (Wilson et al., 2014; Suurväli et al., 2020). Our results underscore that Cd-induced epi-genetic changes relied on an increased heterozygosity by increasing the frequency of an ancestral and "wild" version of the allele. While the alt wild allele may be slightly deleterious or not conferring any fitness advantage under control laboratory conditions, the alt allele became advantageous in embryos and female zebrafish exposed either directly or transgenerationally to Cd. In other words, the pre-existing epi-genetic diversity allowed fish, at least the females, to rapidly adapt to Cd exposure. Recent works highlighted the importance of standing genetic variation (SGV) in the rapid adaptation of organisms to environmental changes (Bitter et al., 2019; Lai et al., 2019). While novel mutations are commonly considered to be the main genetic source of adaptation, SGV can allow a faster response than waiting for the potential emergence and fixation of adaptive mutations (Lai et al., 2019). In our case, we found that genetic changes were in fact not simple synonymous SNPs as they also triggered the creation or deletion of potential sites of methylation. This can explain how synonymous SNPs can be advantageous or deleterious, i.e. can be under selection. Since selection acts on the phenotype, CpG-SNPs can be a mechanism linking synonymous SNPs and adaptive molecular phenotypes. In our case, differential methylation of the promoter and exons through genetic changes led to different transcriptional signatures. Cotter al. (2013) have already shown how the splicing of *esr1* is complex and diverse, mRNA variants being differentially expressed by tissue-type, sex, stage of development and estrogen exposure. At least in female liver, these changes were clearly adaptive, allowing a homeostatic transcription of *esr1* exons and vtg_s genes, as observed in controls. As the offspring of oviparous animals are entirely dependent on stored egg yolk during early development, vitellogenesis plays a key role in offspring survival and thus, in population viability (Reading et al., 2017). Finally, the equal magnitude of mortality observed in the two

conditions (Control and Cd) throughout the 4 generations shows that standing epi-genetic variation allowed the population to avoid any significant decline in response to metal exposure.

4. Conclusion

While investigating for DNA methylation changes in zebrafish transgenerationally exposed to Cd, we found that Cd exerted a selection on the *esr1* gene. This selection was sex dependent and mortalities coincided with key events of sexual determination, reinforcing the role of Cd as an endocrine disruptor. Since the two identified alleles differ by CpG-SNPs, this genetic selection was associated with DNA methylation changes. In female liver, these Cd-induced epi-genetic changes were furthermore associated with adaptive changes in gene expression. As the CpG-SNPs identified in Cd-exposed fish were found in wild-derived strains of zebrafish, our results support the fact that adaptation proceeded by the exploitation of epi-genetic variants that existed prior to domestication and that were maintained at a low frequency in the laboratory population. Using zebrafish as a vertebrate model, our results thus highlight the importance of standing epi-genetic diversity for the adaptive capacity of populations faced with anthropogenic environmental changes. Concomitantly, our results underline the importance (i) of epigenetics in the ability of animals to rapidly adapt as well as (ii) in maintaining genetic diversity within natural populations to bolster the resilience of species faced with the global environmental changes. Further studies are required to explore the implications of such processes in other species as well as in wild populations exposed to different types of pollution and more generally to human-driven environmental changes. It appears important since our findings support the fact that populations with low genetic variation can be more sensitive to anthropogenic environmental changes, which could have important implications in restocking or environmental remediation programs. More generally, our results

support conservation efforts that aim at maintaining genome-wide genetic variation within wild populations to prevent loss of adaptive potential.

CRediT authorship contribution statement

Fabien Pierron: Conceptualization, Methodoly, Validation, Formal analysis, Investigation, Writing - Original Draft, Supervision, Project administration, Funding acquisition. **Guillemine Daffe**, **Flore Daramy**, **Débora Heroin**, **Olivier Bouchez** and **Christelle Clérendeau**: Investigation. **Alicia Romero-Ramirez** : Formal analysis. **Aurélien Barré** : Data curation, Formal analysis. **Macha Nikolski**: Formal analysis, Writing - Review & Editing.

Data Availability

Raw sequencing data were deposited in the European Nucleotide Archive with accession number PRJEB52137.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Christian Portier, Kévin Lebaillif and Jacques Martin are thanked for their technical help with the experimental setup. This work was supported by the Agence Nationale de la Recherche of France (ANR JCJC TRACE ANR-16-CE34-0008) and by the France Génomique National infrastructure, funded as part of “Investissement d’avenir” program managed by the Agence Nationale pour la Recherche (contrat ANR-10-INBS-09).

References

Alves Peixoto, R.A., Jadán-Piedra, C. Cadmium pollution of water, soil, and food: a review of the current conditions and future research considerations in Latin America. *Environ. Rev.* 30:110-127 (2022).

Beck, D., Ben Maamar., M, Skinner, M.K. Genome-wide CpG density and DNA methylation analysis method (MeDIP, RRBS, and WGBS) comparisons. *Epigenetics*. 17(5):518-530 (2022).

Bitter, M.C., Kapsenberg, L., Gattuso, JP. *et al.* Standing genetic variation fuels rapid adaptation to ocean acidification. *Nat Commun.* 10, 5821 (2019).

Brion, F., Tyler, C.R., Palazzi, X., Laillet, B., Porcher, J.M., Garric, J., Flammarion, P. Impacts of 17beta-estradiol, including environmentally relevant concentrations, on reproduction after exposure during embryo-larval-, juvenile- and adult-life stages in zebrafish (*Danio rerio*). *Aquat Toxicol.* 68(3):193-217 (2004).

Cavalli, G., Heard, E. Advances in epigenetics link genetics to the environment and disease. *Nature.* 571, 489–499 (2019).

Chen, Y., Tang, H., Wang, L., He, J., Guo, Y, Liu, Y., Liu, X., Lin, H. Fertility Enhancement but Premature Ovarian Failure in *esr1*-Deficient Female Zebrafish. *Front Endocrinol (Lausanne).* 9:567 (2018).

Chen, Y., Tang, H., He, J., Wu, X., Wang, L., Liu, X., Lin, H. Interaction of nuclear ERs and GPER in vitellogenesis in zebrafish. *J Steroid Biochem Mol Biol.* 189:10-18 (2019).

Chouchene, L., Pellegrini, E., Gueguen, M.M., Hinfray, N., Brion, F., Piccini, B., Kah, O., Saïd, K., Messaoudi, I., Pakdel, F. Inhibitory effect of cadmium on estrogen signaling in zebrafish brain and protection by zinc. *J Appl Toxicol.* 36(6):863-71 (2016).

Cotter, K.A,. Yershov, A., Novillo, A., Callard, G.V. Multiple structurally distinct ERα mRNA variants in zebrafish are differentially expressed by tissue type, stage of development and estrogen exposure. *Gen Comp Endocrinol.* 194:217-29 (2013).

Duncan, E.J., Gluckman, P.D., Dearden, P.K. Epigenetics, plasticity, and evolution: How do we link epigenetic change to phenotype? *J Exp Zool B Mol Dev Evol.* 322(4):208-20 (2014).

Durland, E., De Wit, P., Langdon, C. Temporally balanced selection during development of larval Pacific oysters (*Crassostrea gigas*) inherently preserves genetic diversity within offspring. *Proc Biol Sci.* 288(1958):20203223 (2021).

Feng, S., Cokus, S.J., Zhang, X., Chen, P.Y., Bostick, M., Goll, M.G., Hetzel, J., Jain, J., Strauss, S.H., Halpern, M.E. Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl. Acad. Sci. USA.* 107, 8689–8694 (2010).

Fitz-James, M.H., Cavalli, G. Molecular mechanisms of transgenerational epigenetic inheritance. *Nat Rev Genet.* (2022).

Giguère, A., Campbell, P.G.C., Landis Hare, L., D Gordon McDonald, D.G., and Joseph B Rasmussen, J.B. 2011. Influence of lake chemistry and fish age on cadmium, copper, and zinc concentrations in various organs of indigenous yellow perch (*Perca flavescens*). *Can. J. Fish. Aquat. Sci.* 61(9):1702-1716 (2004).

Heard, E., Martienssen, R.A. Transgenerational epigenetic inheritance: myths and mechanisms. *Cell.* 157(1):95-109 (2014).

Hoffmann, A.A., Sgrò, C.M. Climate change and evolutionary adaptation. *Nature.* 470(7335):479-85 (2011).

Horsthemke, B. A critical view on transgenerational epigenetic inheritance in humans. *Nat Commun.* 9, 2973 (2018).

Johnson, M.D., Kenney, N., Stoica, A., Hilakivi-Clarke, L., Singh, B., Chepko, G., Clarke, R., Sholler, P.F., Lirio, A.A., Foss, C., Reiter, R., Trock, B., Paik, S., Martin, M.B. Cadmium mimics the in vivo effects of estrogen in the uterus and mammary gland. *Nat Med.*, 9(8):1081-4 (2003).

Kilvitis, H.J., Hanson, H., Schrey, A.W., Martin, L.B. Epigenetic Potential as a Mechanism of Phenotypic Plasticity in Vertebrate Range Expansions. *Integr Comp Biol.* 57(2):385-395 (2017).

Kippler, M., Engström, K., Mlakar, S.J., Bottai, M., Ahmed, S., Hossain, M.B., Raqib, R., Vahter, M., Broberg, K. Sex-specific effects of early life cadmium exposure on DNA methylation and implications for birth weight. *Epigenetics*. 8(5):494-503 (2013).

Lai, Y.T., Yeung, C.K.L., Omland, K.E., Pang, E.L., Hao, Y., Liao, B.Y., Cao, H.F., Zhang, B.W., Yeh, C.F., Hung, C.M., Hung, H.Y., Yang, M.Y., Liang, W., Hsu, Y.C,. Yao, C.T., Dong, L., Lin, K., Li, S.H. Standing genetic variation as the predominant source for adaptation of a songbird. *Proc Natl Acad Sci USA*. 116(6):2152-2157 (2019).

Liew, W.C., Orbán, L. Zebrafish sex: a complicated affair. *Brief Funct. Genomics*. 13(2):172-187 (2014).

Lou, S., Lee, H.M., Qin, H., Li, J.W., Gao, Z., Liu, X., Chan, L.L., Lam V.K.L., So, W.Y., Wang, Y., Lok, S., Wang, J., Ma, R.C., Tsui, S.K., Chan, J.C., Chan, T.F., Yip, K.Y. Whole-genome bisulfite sequencing of multiple individuals reveals complementary roles of promoter and gene body methylation in transcriptional regulation. *Genome Biol.* 15(7):408 (2014).

Maack, G., Segner, H. Life-stage-dependent sensitivity of zebrafish (*Danio rerio*) to estrogen exposure. *Comp Biochem Physiol C Toxicol Pharmacol.* 139(1-3):47-55 (2004).

Mastoraki, S., Strati, A., Tzanikou, E., Chimonidou, M., Politaki, E., Voutsina, A., Psyrri, A., Georgoulias, V., Lianidou, E. *ESR1* Methylation: A Liquid Biopsy-Based Epigenetic Assay for the Follow-up of Patients with Metastatic Breast Cancer Receiving Endocrine Treatment. *Clin Cancer Res.* 15;24(6):1500-1510 (2018).

McCurley, A.T., Callard, G.V. Characterization of housekeeping genes in zebrafish: male-female differences and effects of tissue type, developmental stage and chemical treatment. *BMC Mol. Biol*. 9:102 (2008).

Moore, L.D., Le, T., Fan, G. DNA methylation and its basic function. *Neuropsychopharmacology*. 38(1):23-38 (2013).

Niekerk, L.A., Carelse, M.F., Bakare, O.O., Mavumengwana, V., Keyster, M., Gokul, A. The Relationship between Cadmium Toxicity and the Modulation of Epigenetic Traits in Plants. *Int. J. Mol. Sci.* 30;22(13):7046 (2021).

Nilsson, E.E., Sadler-Riggleman, I., Skinner, M.K. Environmentally induced epigenetic transgenerational inheritance of disease. *Environ Epigenet.* 4(2):dvy016 (2018).

Orban, L., Sreenivasan, R., Olsson, P.E. Long and winding roads: testis differentiation in zebrafish. *Mol Cell Endocrinol.* 312(1-2):35-41 (2009).

Ortega-Recalde, O., Day, R.C., Gemmell, N.J., Hore, T.A. Zebrafish preserve global germline DNA methylation while sex-linked rDNA is amplified and demethylated during feminisation. *Nat Commun.* 10:3053 (2019).

Pierron, F., Baudrimont, M., Dufour, S., Elie, P., Bossy, A., Baloche, S., Mesmer-Dudons, N., Gonzalez, P., Bourdineaud, J.-P., Massabuau, J.-C. How cadmium could compromise the completion of the European eel's reproductive migration. *Environ. Sci.Technol.* 42 :4607-4612 (2008).

Pierron, F., Baillon, L., Sow, M., Gotreau, S., Gonzalez, P. Effect of low-dose cadmium exposure on DNA methylation in the endangered European eel. *Environ. Sci. Technol*. 48(1):797-803 (2014).

Pierron, F., Daffe, G., Lambert, P., Couture, P., Baudrimont, M. Retrotransposon methylation and activity in wild fish (*A. anguilla*): a matter of size. *Environ Pollut.* 245:494-503 (2019).

Pierron, F., Lorioux, S., Heroin, D., Daffe, G., Etcheverria, B., Cachot, J., Morin, B., Dufour, S., Gonzalez, P. Transgenerational epigenetic sex determination: environment experienced by female fish affects offspring sex ratio. *Environ Pollut.* 277:116864 (2021).

Pierron, F., Heroin, D., Daffe, G., Daramy, F., Barré A., Bouchez, O., Romero-Ramirez, A., Gonzalez, P., Nikolski, M. Genetic and epigenetic interplay allows rapid transgenerational adaption to metal pollution in zebrafish. *Environ Epigenet.,* 8(1):dvac022 (2022).

Reading, B.J., Sullivan, C.V., Schilling, J. Vitellogenesis in Fishes. In *Reference Module in Life Sciences*, Elsevier, 2017.

Saintilnord, W.N., Tenlep, S.Y.N., Preston, J.D., Duregon, E., DeRouchey, J.E., Unrine, J.M,, de Cabo, R., Pearson, K.J., Fondufe-Mittendorf, Y.N. Chronic Exposure to Cadmium Induces Differential Methylation in Mice Spermatozoa. *Toxicol. Sci.* 180(2):262-276 (2021).

Santos, D., Luzio, A., Coimbra, A.M. Zebrafish sex differentiation and gonad development: A review on the impact of environmental factors. *Aquat Toxicol*. 191:141-163 (2017).

Shayevitch, R., Askayo, D., Keydar, I., Ast, G. The importance of DNA methylation of exons on alternative splicing. *RNA*. 24(10):1351-1362 (2018).

Skinner, M.K., Manikkam, M., Guerrero-Bosagna, C. Epigenetic transgenerational actions of environmental factors in disease etiology. *Trends Endocrinol Metab.* 21(4):214-22 (2010).

Stoica, A., Katzenellenbogen, B.S., Martin, M.B. Activation of estrogen receptor-alpha by the heavy metal cadmium. *Mol Endocrinol.* 14(4):545-53 (2000).

Suurväli, J., Whiteley, A.R., Zheng, Y., Gharbi, K., Leptin, M., Wiehe, T. The Laboratory Domestication of Zebrafish: From Diverse Populations to Inbred Substrains. *Mol Biol Evol*. 37(4):1056–1069 (2020).

Takiguchi, M., Yoshihara, S. New aspects of cadmium as endocrine disruptor. *Environ Sci.* 13(2):107-16 (2006).

Thiebaut, F., Hemerly, A.S., Ferreira, P.C.G. A Role for Epigenetic Regulation in the Adaptation and Stress Responses of Non-model Plants. *Front Plant Sci.* 10:246 (2019).

Tsuboi, K., Nagatomo, T., Gohno, T., Higuchi, T., Sasaki, S., Fujiki, N., Kurosumi, M., Takei, H., Yamaguchi, Y., Niwa, T., Hayashi, S.I. Single CpG site methylation controls estrogen receptor gene transcription and correlates with hormone therapy resistance. *J Steroid Biochem Mol Biol.* 171:209-217 (2017).

Tzung, K.W., Goto, R., Saju, J.M., Sreenivasan, R., Saito, T., Arai, K., Yamaha, E., Hossain, M.S., Calvert, M.E.K., Orbán, L. Early depletion of primordial germ cells in zebrafish promotes testis formation. *Stem Cell Reports.* 4(1):61-73 (2015).

Uchida, D., Yamashita, M., Kitano, T., Iguchi, T. Oocyte apoptosis during the transition from ovary-like tissue to testes during sex differentiation of juvenile zebrafish. *J Exp Biol.* 205(Pt 6):711-8 (2002).

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., Leunissen, J.A.M. Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res.* 35:W71-W74 (2007).

Vandegehuchte, M., Janssen, C. Epigenetics in an ecotoxicological context. *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* 764-765:36-45 (2014).

Wang, B., Li, Y., Shao, C., Tan, Y., Cai, L. Cadmium and its epigenetic effects. *Curr Med Chem.* 19(16):2611-20 (2012).

Wilson, C.A., High, S.K., McCluskey, B.M., Amores, A., Yan, Y.L., Titus, T.A., Anderson, J.L., Batzel, P., Carvan, M.J., Schartl, M., Postlethwait, J.H. Wild sex in zebrafish: loss of the natural sex determinant in domesticated strains. *Genetics.* 198(3):1291-308 (2014).

Wu, K., Song, W., Zhang, Z., Ge, W. Disruption of *dmrt1* rescues the all-male phenotype of the *cyp19a1a* mutant in zebrafish - a novel insight into the roles of aromatase/estrogens in gonadal differentiation and early folliculogenesis. *Development.* 147(4):dev182758 (2020).

Xin, F., Susiarjo, M., Bartolomei, M.S. Multigenerational and transgenerational effects of endocrine disrupting chemicals: A role for altered epigenetic regulation? *Semin. Cell Dev. Biol.* 43:66-75 (2015).

Yilmaz, O., Patinote, A., Com, E., Pineau, C., Bobe, J. Knock out of specific maternal vitellogenins in zebrafish (*Danio rerio*) evokes vital changes in egg proteomic profiles that resemble the phenotype of poor quality eggs. *BMC Genomics.* 28;22(1):308 (2021).

Yilmaz, O., Patinote, A., Nguyen, T., Com, E., Pineau, C., Bobe, J. Genome editing reveals reproductive and developmental dependencies on specific types of vitellogenin in zebrafish (*Danio rerio*). *Mol Reprod Dev.* 86(9):1168-1188 (2019).

Yong, W.S., Hsu, F.M., Chen, P.Y. Profiling genome-wide DNA methylation. *Epigenetics & Chromatin.* 9:26 (2016).

Zeng, Y., Chen, T. DNA Methylation Reprogramming during Mammalian Development. *Genes (Basel).* 10(4):257 (2019).

Supplementary Material for:

Title: Transgenerational endocrine disruptor effects of cadmium in zebrafish and contribution of standing epigenetic variation to adaptation

List of authors:
Fabien Pierron, Guillemine Daffe, Flore Daramy, Débora Heroin, Aurélien Barré, Olivier Bouchez, Christelle Clérendeau, Alicia Romero-Ramirez and Macha Nikolski

***Corresponding author:**
Fabien Pierron
UMR 5805 EPOC - OASU
Station Marine d'Arcachon
Université de Bordeaux
Place du Docteur Bertrand Peyneau
33120 ARCACHON
France
Phone : +335 56 22 39 33
fabien.pierron@u-bordeaux.fr

Table S1: Primers used for bisulfite sequencing (BS-Pyro-Seq), sequencing (Seq), and qPCR analyses (rt-qPCR). F forward; R reverse; S sequencing, sequences in bold are biotinylated in 5'. All sequences used to design primer in *esr1* were obtained from the Genome Reference Consortium Zebrafish (GRCz11) on CH20 (NC_007131.7). For *vtg* genes, primers were obtained from Yilmaz et al. (2019).

Region amplified	Analyze	Sequence position on NC_007131.7	Primers
esr1 Exon3	BS-Pyro-Seq	26377595 - 26377832	F: GTGAGGAGTTGTGTAGTGTTAT **R: CACATCCCTACTAATATTCAAAACAACAA** S: GTAGGAGTTGTATATTGGA
esr1 Exon5	BS-Pyro-Seq	26368680 - 26368851	**F: AGTGTAGTTGATTTGTGTTGAAGGTATT** R: ACCACTACCCTTCTTCCTCTTATCCT S: CATTACTACTTCTCCTTC
esr1 Promoter	BS-Pyro-Seq	26383526-26383454	F: AGAGGAGGTAAATAAATTAAAGATAGTTAG **R: TACTCCTTTAACATATAATTTCCCATAACA** S: GGTAAATAAATTAAAGATAGTTAGG
esr1 Promoter	Seq (Sanger)	26383322- 26383602	F: *TGAGTCCCATTACAGGCTGG* R: *ATTTCCCATGGCAGCAGCA*
esr1 Promoter	Seq (Pyro)	26383438-26383593	F: ATTAAAGACAGCCAGGTCGATACT R: **TGGCAGCAGCATGTAAAGTG** S: CAGGTAAACAAGAGATAGGA
esr1 Exon2	rt-qPCR	26381906-26381988	F: TGGACGAGGAATCCAGGTAG R: ACAGCTGTCCCCGTATCTGA
esr1 Exon3	rt-qPCR	26377668-26377733	F: AGAAAGCTTTGCATCCCTCA R: TGTGCAGTGACTACGCCTCT
esr1 Exon5	rt-qPCR	26368649-26368710	F: CAGGCGACATGCATAAAGTG R: CCCACAGGACAAGAGGAAGA
esr1 Exon8	rt-qPCR	26362327 - 26362381	F: TGTGTGTGTGACTCTCAATGTACC R: CTGCTGCTGCTCTCACACAT
esr1 Exon3-Exon5	rt-qPCR	26368687 -26377705	F: TTACGGAGTCTGGTCGTGTG R: CGCTTCTTCCTCTTGTCCTG
esr1 Exon5-Exon7	rt-qPCR	26364793 - 26368548	F: CACCAACATGGCTGACAAAG R: ATGCCCTCAACACATTCTCC
esr1 Exon7-Exon9	rt-qPCR	26361496 - 26364738	F: GCTCAAGCTGGAGGAATTTG R: ATACAGCGGCACTCGATTCT
vtg1	rt-qPCR	/	F: GATTAAGCGTACACTGAGACCA R: AGCCACTTCTTGTCCAAATACT
vtg2	rt-qPCR	/	F: TGCCGCATGAAACTTGAATCT R: GTTCTTACTGGTGCACAGCC
vtg3	rt-qPCR	/	F: GGGAAAGGATTCAAGATGTTCAGA R: ATTTGCTGATTTCAACTGGGAGAC
vtg5	rt-qPCR	/	F: ATTGCCAAGAAAGAGCCCAA R: TTCAGCCTCAAACAGCACAA
vtg7	rt-qPCR	/	F: TTGGTGTGAGAACTGGAGGA R: TTGCAAGTGCCTTCAGTGTA

Figure S1: Pyrograms obtained for the exon3 of the *esr1* gene using bisulfite-treated DNA.

The sequence analyzed after bisulfite treatment was as follow:

TTT/AAGAGYGG[C]TTGGAGGTTTTGATTT/CA/GGGGAAAGAGATTYGTTTTTGTG
YGGTGTG

The sequences analyzed after bisulfite treatment for the exon 5 and promoter were as follow:

>exon 5
TCTCACRCCTAACAAAACRCCCG/A[T]CCRCRATCTTACRAATACCTTC
>promoter
TYGATATTGAYGGTTATTTTTTAGAG[C]TAGGTTATGGTAATTAGGTAAATAAGA
GATAGGAGGTTATT/CA/GTTTGA

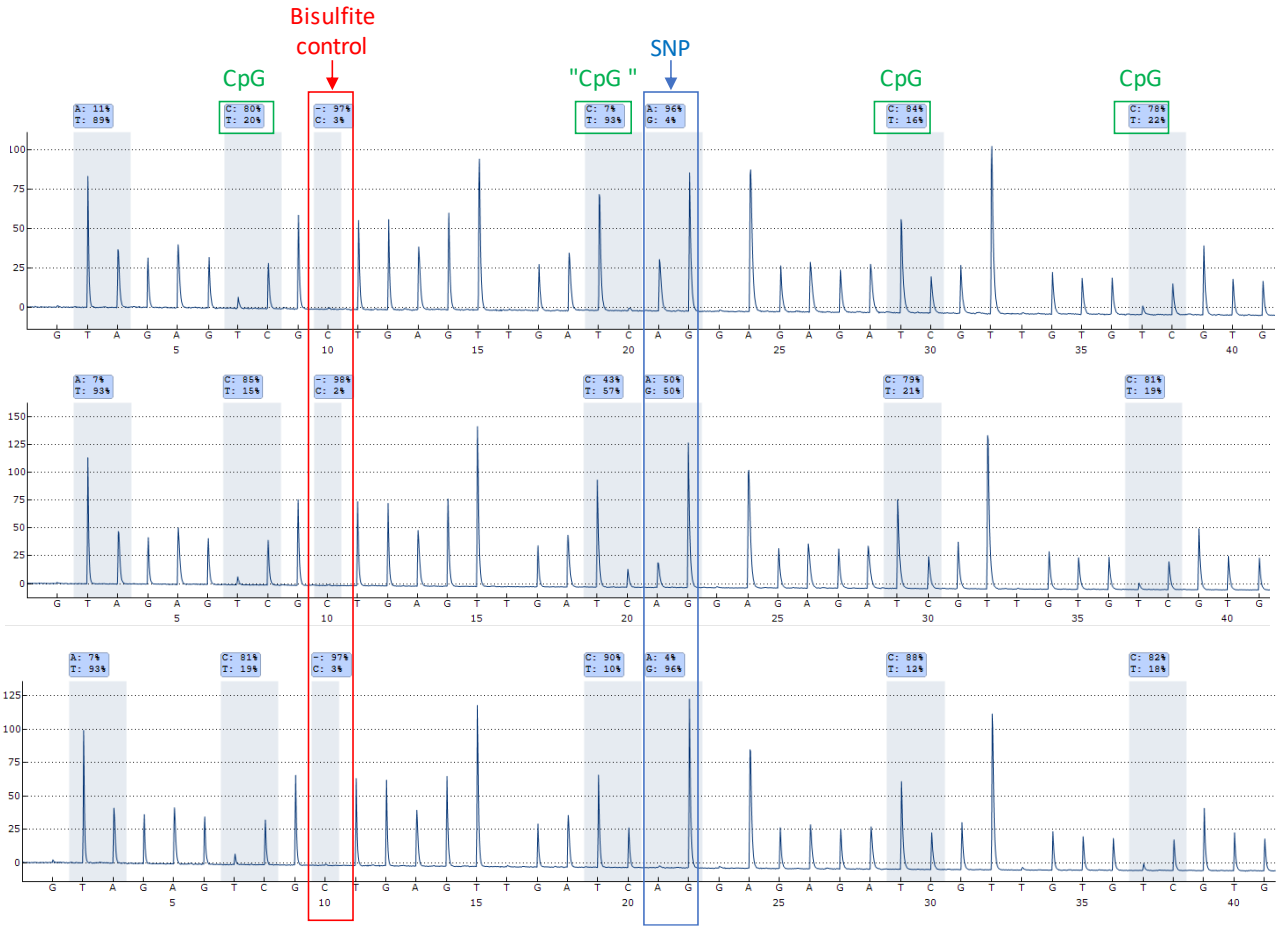


Figure S2: Pyrograms obtained for the promoter of the *esr1* gene using genomic DNA (non-bisulfite-treated DNA).

The sequence analyzed was as follow:

GGTCACCA/GCTTGATTCACTTGCCAAAA

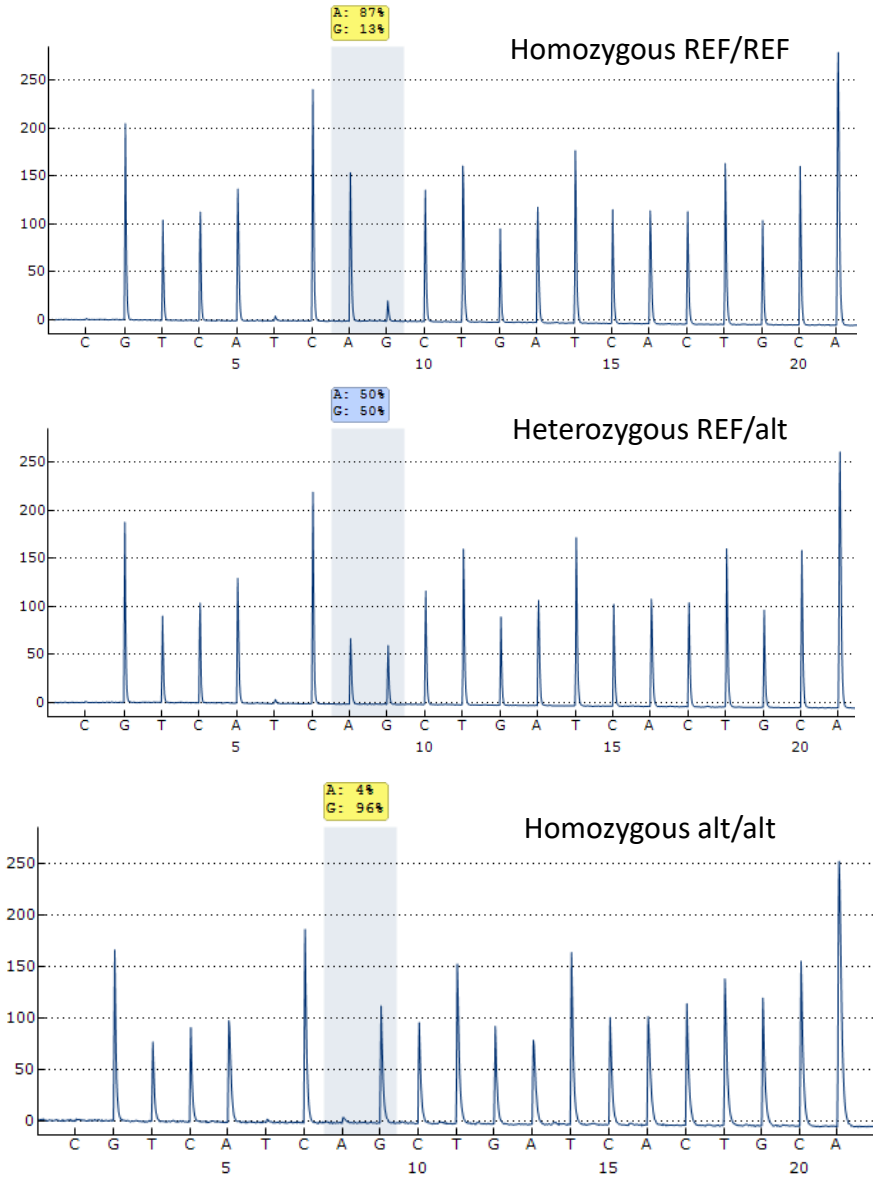


Figure S3: Change in the methylation level (mean ± SE, n = 10-23) of the *esr1* promoter in gonads of female zebrafish (124 dpf) determined from the CpG1-3 (solid bars) or CpG2-3 (dashed bars) according to the genotype (REF/REF or REF/alt) and the condition (C, control; Cd, cadmium). Data from the four generations were combined to enlarge sample size. No homozygous alt/alt females were observed and only one control female was heterozygous. For each variable, means designated by different letters are significantly different (LSD, P <0.05).

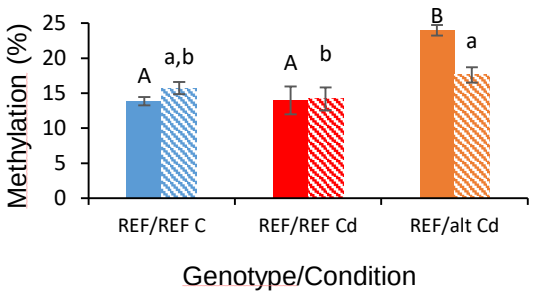


Figure S4: The alternative allele showed high homologies with the sequences from the wild-derived strains Nadia and Cooch Behar. CpG sites analyzed by bisulfite pyrosequencing are framed in red.

(**A**) Multiple sequence alignment of the *esr1* promoter (region located on the chromosome 20, (NC_007131.7) position 26383322-26383602 in the reference genome GRCz11). (**B**) Phylogram of the promoter sequences analyzed. (**C**) Multiple sequence alignment of the exon 3 of *esr1* (region located on the chromosome 20, (NC_007131.7) position 26377732- 26377785 in the reference genome GRCz11) (**D**) Multiple sequence alignment of the exon 5 of *esr1* (region located on the chromosome 20, (NC_007131.7) position 26368787-26368823 in the reference genome GRCz11)

GRCz11_REF: reference sequence in the reference genome; GRCz11_alt: alternative sequence (NW_018395074.1) in the reference genome; REF/REF: data obtained from REF/REF homozygous individuals used in our study; alt/alt: data obtained from homozygous alt/alt individuals used in our study; NA: Nadia (LR812588.1), the eighth generation of animals taken from nature in Nadia, India, in 2000; CB: Cooch Behar (LR812563.1), a new strain derived from wild fish (2nd generation in captivity) collected in 2012 from Cooch Behar, India.

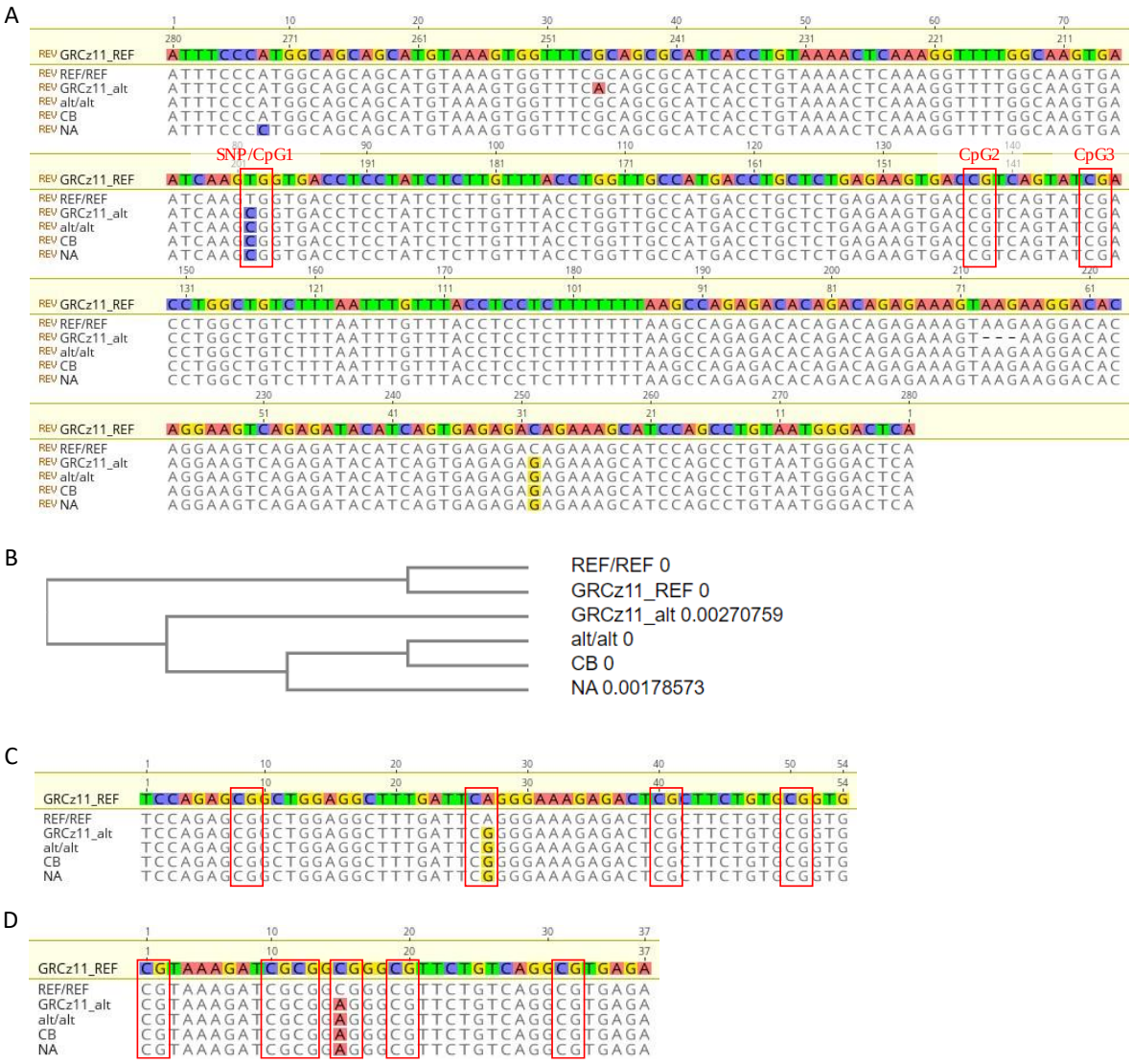


Figure S5: Mortality rates across the 4 generations.

(**A**) Change in the embryo-larval (0-14 dpf) mortality rates across the 4 generations (F0 to F3) for control or Cd fish. Means designated by different letters are significantly different (LSD, P <0.05). At least 14 spawns (n = 14-20) were followed per condition at each generation. (**B**) Mortality was recorded throughout the experiment at least once a day. At each generation, all individuals were counted at different stages of development: eggs (0 dpf), larvae (14 dpf) and adults (170 dpf). Cannibalism (mean ± SE, n = 4 aquaria) was thus estimated by inventorying the number of unexplained disappearances (i.e. by subtracting the number of animals that were found dead in the aquaria and sampled for analysis). Means designated by different letters are significantly different (LSD, P <0.05). We much add that we observed zebrafish with swollen belly, mainly at the juvenile stage, when interindividual variability in size was marked (see figure C below). By photographing the aquaria at different times and counting animals, we estimate that cannibalism did not occur beyond 60 dpf. (**C**) Juvenile (35 dpf) control (blue) or Cd (red) zebrafish of the F0 generation.