

MOLECULAR AND FUNCTIONAL PROPERTIES OF THE ATLANTIC COD (*GADUS MORHUA*) ARYL HYDROCARBON RECEPTORS AHR1A AND AHR2A

Libe Aranguren-Abadía¹, Roger Lille-Langøy¹, Alexander K. Madsen, Sibel I. Karchner², Diana G. Franks², Fekadu Yadetie¹, Mark E. Hahn², Anders Goksøyr¹, and Odd André Karlsen^{1*}

¹*Department of Biological Sciences, University of Bergen, Bergen, Norway;* ²*Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, United States.* * *Corresponding author*

Corresponding Author

Odd André Karlsen

Department of Biological Sciences

University of Bergen

P.O. box 7803

N-5020, Bergen

Norway

*E-mail: Odd.Karlsen@uib.no

S2-5: Supplementary methods

S6-22: Supplementary tables and figures

S6-9: Table S1-S4

S10-22: Figure S1-13

S23: References

Supplementary methods

RNA isolation and cloning of *gmahr1a*, *gmahr2a*, *gmarnt1*.

gmahr1a, *gmahr2a* and *gmarnt1* were amplified in duplicate or triplicate reactions using gene specific primers (Table S1) with PrimeStar GXL (Takara BIO Europe). Full coding sequences for *gmahr1a*, *gmahr2a* and *gmarnt1* were amplified and inserted into pcDNA3.1/Zeo(+) (Thermo Fisher). The use of forward primers with the Kozak consensus sequence introduced these upstream of the transcription start sites of both *gmahr1a* and *gmahr2a* (Table S1).

Synteny mapping, sequence alignments, and phylogenetic analyses.

GenBank accession numbers for AHR proteins used for phylogenetic analyses: American alligator (*Alligator mississippiensis*) Ahr (XM_006261671.3), *Caenorhabditis elegans* AHR (NM_001025865.3), zebrafish (*Danio rerio*) Ahr1a (AF258854.1), Ahr1b (BC163508.1), Ahr2 (NM_131264.1), mummichog (*Fundulus heteroclitus*) Ahr1a (AY454298.1), Ahr1b (KX372715.1), Ahr2a (AAC59696.3), Ahr2b (KX372716.1), chicken (*Gallus gallus*) Ahr1 (NM_204118.2), Ahr1b (NM_001318993.1), Ahr2 (NM_001319008.1), human (*Homo sapiens*) AHR (NM_001621.4), Atlantic tomcod (*Microgadus tomcod*) Ahr2 (FJ215753.1), mouse (*Mus musculus*) AHR (NM_013464.4), Japanese medaka (*Oryzias latipes*) Ahr1a (XM_011489981.3), Ahr1b (NM_001104678.1), Ahr2a (XM_023950680.1), Ahr2b (XM_011481447.3), red seabream (*Pagrus major*) Ahr1 (AB197787.1), Ahr2 (AB197788.1), Japanese puffer (*Takifugu rubripes*) Ahr1a (NM_001037962.1), Ahr1b (NM_001037959.1), Ahr2a (NM_001037960.1), Ahr2b (NM_001037963.1), frog (*Xenopus laevis*) AHR1A (NM_001171912.2), AHR1B (AY635783.1). Lake sturgeon (*Acipenser fulvescens*) Ahr1 (KM236089), Ahr2 (AIW39681). White sturgeon (*Acipenser transmontanus*) Ahr1 (KJ420394) Ahr2 (KJ420395). Atlantic sturgeon (*Acipenser oxyrinchus*) Ahr1 (MH925108) Ahr2 (MH223597). Shortnose sturgeon (*Acipenser brevirostrum*) Ahr1 (MH925109) Ahr2 (MH223598). The sequences were aligned with MUSCLE in Geneious. Alignment positions with gaps were not included in the subsequent analyses. Bayesian

inference analysis was conducted in MrBayes v3.2.7a using a BLOSUM substitution model¹. *Caenorhabditis elegans Ahr* was used as out-group. For bootstrapping, Markov chain Monte Carlo (MCMC) analysis was run for 300,000 generations for each 1000 samples with a 25% burn-in. Four chains were used with a heating parameter of 0.1.

***In vitro* protein expression and velocity sedimentation assays.**

TNT reactions (100 µl in total) of each gmAhr protein were diluted 1:1 with MEEDMGA buffer and incubated overnight at 4°C with either [³H]TCDD (2 nM) or [³H]BNF (10 nM). Samples were fractionated in 10-30% sucrose gradients by centrifugation, and radioactivity was measured using a scintillation counter (Beckman LS6500).

Transfection, exposure and luciferase reporter gene assay.

COS-7 simian kidney cells were maintained in Dulbecco's modified Eagle medium (DMEM) with phenol red, supplemented with 10% fetal bovine serum (FBS), 4 mM L-glutamate, 1 mM sodium pyruvate and 100U/mL penicillin-streptomycin at 37 °C with 5% CO₂ for 24 hours. The cells were transfected either using Roche XtremeGene 9/HP transfection reagent or Mirus TransIT LT-1 transfection reagent, according to the recommendations of the suppliers. Following transfection, cells were exposed to FICZ (0.001-10 nM), TCDD (0.03-1000 nM), benzo(a)pyrene (B[a]P) (0.3-10000 nM), PCB126 (0.05-10000 nM), BNF (0.3-10000 nM) or to solvent control (0.5-1% final DMSO concentration) for 12 hours. FICZ, TCDD, B[a]P and PCB126 assays were repeated three times and there were four technical replicates. BNF assay was repeated twice and there were three technical replicates. Non-linear regression analyses of dose-response data were performed in Prism v7. Effective concentration 50 (EC₅₀) values and statistical differences between EC₅₀ values, as well as statistical differences between the maximum effects (E_{max}) were obtained using the dose-response analyses drc package in RStudio v1.2.1335 software². Cytotoxicity of the different tested compounds was evaluated with two fluorescent dyes, resazurin and 5-carboxyfluorescein diacetate

acetoxymethyl ester (5-CFDA-AM), which monitor cell membrane integrity and metabolic activity, respectively, essentially as described in Pérez-Albaladejo et al. ³. In short, COS7 cells were seeded, cultivated and exposed to the highest concentrations of each compound. Cells were incubated in resazurin and 5-CFDA-AM solutions at 37 °C for one hour. Fluorescence was measured at 530/590 nm for resazurin and 485/530 nm for CFDA-AM with an EnSpire plate reader (Perkin Elmer, Massachusetts, United States) (Fig. S1, Fig. S2). The assay was repeated twice. Statistical significance was assessed with One-way ANOVA.

Tissue-specific expression of *ahr1a*, *ahr2a*, *arnt1* and *arnt2*.

RNA extractions were conducted using the TRI Reagent® protocol. RNA integrity was assessed by agarose gel electrophoresis and RNA concentrations were measured with a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Massachusetts, USA). Samples possessing $A_{260}/_{280}$ ratios lower than 1.7 were further purified by ethanol precipitation. 500 ng of RNA was reverse transcribed to cDNA using the iScript™ cDNA Synthesis Kit (Bio-Rad, California, USA) following the provider's protocol. Quantitative real-time polymerase chain reaction analyses were performed using 5 µl of cDNA diluted (1:20) in a 20 µl PCR reaction mix containing 0.5 µM of each of the forward and reverse primers and SYBR Green Master I (Roche Applied Sciences, Basel, Switzerland). The reactions were run in a CFX96 Touch Real-Time PCR Detection System (Bio-Rad) following these conditions: 3 min 95°C, 40 cycles (15 sec 95°C, 30 sec 55°C, 30 sec 72°C), 30 sec 72°C, and melt curve. A no reverse transcriptase control (NRT) and a no template control (NTC) for each primer combination were included. Primers for *gmahr1a*, *gmahr2a*, *gmarnt1* and *gmarnt2* are shown in Table S3.

***Ex vivo* exposure assays with precision-cut liver slices (PCLS) and analyses of *cyp1a* expression.**

Fish were anaesthetized with MS-222 (100 mg/L) (Sigma-Aldrich, Missouri, United States) dissolved in saltwater and killed with a blow to the head prior to dissection. The hepatic portal vein

was perfused to remove blood using cold buffer (NaCl (122 mM), KCl (4.8 mM), MgSO₄ (1.2 mM), Na₂HPO₄ (11 mM) and NaHCO₃ (3.7 mM), pH 8.4) and a 10 ml syringe with a 23G needle as described in Ellesat et al.,⁴. Liver cores were cut with a 8 mm cylinder-shaped coring tool and embedded in 3 % ultra-low melting temperature agarose (Sigma-Aldrich) as described previously in Yadetie et al.⁵. Slices of 250 µm in thickness were exposed to the different ligands after a 2-hour acclimatization period, and cultivated in 24-well plates at 10°C for 48 hours.

cDNA synthesis and qPCR analyses were performed as described above. Expression data were analyzed using GLMM models with gamma log₁₀ distribution and fish as a random effect in RStudio v1.2.1335⁶. Level of significance is expressed with * (p< 0.05) or *** (p<0.001).

The viability of the liver slices was assessed with the 3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide (MTT) assay that measures metabolic activity. Slices (n=3) were exposed to selected concentrations of each compound in duplicates following the protocol as described in Yadetie et al.⁵. Absorbance was measured at 590 nm with an EnSpire plate reader. The wet-weight of each liver slice (mg) was used for normalization, and cytotoxicity was expressed relative to the metabolic activity of solvent exposed slices (Fig. S3). Additionally, the viability of the slices was monitored with the Cytotoxicity Detection Kit (LDH) (Sigma-Aldrich, Missouri, United States) according to the manufacturer's protocol and described in Bizarro et al.⁷, which assesses the membrane integrity as a result of the release of lactate dehydrogenase (LDH) into the culture media. Liver slices (n=3) were exposed to the different compounds and culture medium was collected after 48 hours. The assay was repeated two times. Absorbance was measured at 490 nm and 650 nm with an EnSpire plate reader (Fig. S4). Two-way ANOVA was used to test for statistical significance.

Supplementary tables and figures

Table S1. Overview of primers used for amplification of cod *gmarnt1* and *gmahrs* from Atlantic cod cDNA

Primer name	Primer sequence (5'>3')	Target
Ahr1a cod fwd	ATGTATGCCGGGCGAAAAAGGAGAAAAC	Cod <i>ahr1a</i>
Ahr1a cod rev	TCAAAGGTAGAAGTCTGTGACCTGTGGG	
AHR2a COD Fwd EcoRI	ATCGGAATTCATGTTGGGTAACGCTGGG	Cod <i>ahr2a</i> , part 1, 1-2612
CodAhr2a part5 rev	CATTGTGAGGGAGGTTGGACG	
CodAhr2a part5 fwd	GGAGAGCCTGCTCAGTAATGAC	Cod <i>ahr2a</i> , part 2, 1659-3381, 5'BamHI site
AHR2a COD Rev BamHI	ATCGGGATCCTCAGAAGTTGCAACAGTTGAG	
Cod arnt1 F	ATGGACATTCCTTCTCCAAACC	Cod <i>arnt1</i>
Cod_arnt1_R	TCACTCGTIGAAAGAAGGGTAGATG	
Cod_ahr1a_NotI_fwd	CCGGCGGCCGCATGTATGCCGGGCGAAAAAGG	<i>Ahr1a</i> with flanking NotI and XbaI sites and kozak sequence.
Cod_ahr1a_NotI_kozak_fwd	CCGGCGGCCGCACCATGGGGTATGCCGGGCGAAAAAGG	
Cod_ahr1a_XbaI_rev	GGCTCTAGATCACTTGTTCATCGTCGTCCTTGTAGTCAAGGTAGAAGTCTGTGACC	
Cod_ahr2a_NotI_fwd	CCGGCGGCCGCATGTTGGGTAACGCTGGGAC	<i>Ahr2a</i> with flanking NotI and XbaI sites and kozak sequence.
Cod_ahr2a_NotI_kozak_fwd	CCGGGTACCACCATGGGGGACATTCCTTCTCCAAACC	
Cod_ahr2a_XbaI_rev	GGCTCATCTAGACTTGTTCATCGTCGTCCTTGTAGTCGAAGTTGCAACAGTTGAGTTG	
Cod_arnt1_KpnI_fwd	CCGGGTACCATGGACATTCCTTCTCCAAACC	Amplification of <i>arnt1</i> for insertion to pcDNA3.1/Zeo(+) using KpnI and XhoI sites
Cod_arnt1_XhoI_rev	GGCCTCGAGTCACTCGTTGAAAGAAGGGTAG	

Table S2: Overview of transfection of COS-7 cells using Roche X-tremeGene 9/HP or Mirus TransIT LT-1 transfection reagents

	Roche X-tremeGene HP/9 Reagent	Mirus TransIT LT-1 Transfection reagent
Number of cells/well (*10 ³)	54	10
pcDNA3.1_codAhr (ng/well)	3	3
pcDNA3.1_codArnt (ng/well)	8	6
pGudLuc6.1 (ng/well)	7	30
pRT-TK (ng/well)	3	-
pCMV- βGAL (ng/well)	-	20
pcDNA3.1/Zeo(+)	80	41
Transfection (duration, hours)	5	24
Exposure (duration, hours)	24	24
Final [DMSO] (%)	1	0.5
Measurements (Perkin Elmer EnSpire plate reader)	Promega Luciferase System Kit	Dual-Glo Assay
Exposures	TCDD, FICZ, B[a]P and PCB-126	BNF
		Performed essentially as described in ⁸

Table S3. Overview of the different primers used in tissue-specific quantitative real time PCR analysis in Atlantic cod. Gene short names, GenBank accession or Ensembl numbers, primer sequences and amplicon sizes are shown.

Gene	Accession number	Forward primer (5'-3')	Reverse primer (5'-3')	Product size (bp)
<i>ahr1a</i>	MN329012	CAAGGGCGTCTCAAGTTCCTACAT	CAGCACTATCTCCCCTTTGCATCAC	207
<i>ahr2a</i>	MN329013	ACAAACTGTCCGTGCTCCGACTTA	TCCATTTCGGGCCATTGTGCTTCT	92
<i>arnt1</i>	MN329014	TAGCTACGAGGCTGGACAAATCAC	AGAGTCGGGTAAAGCTCTGCTTTC	111
<i>arnt2</i>	ENSGMOG00000014557	AAGACCAGAGTCACCTGAGGGAAA	GATCAACATCCACTCCCAGTTCTT	110

Table S4: Summary of effective concentration 50 (EC₅₀) and maximum effect (E_{max}) values of the different compounds and receptors tested in the *in vitro* luciferase reporter gene assay. Values were obtained from RStudio v1.2.1335. Significant statistical differences between gmAhr1a and gmAhr2a are indicated with *(p< 0.05) or *** (p<0.001).

Compound	Receptor	EC₅₀ (nM)	Maximal activation (E_{max}) (RLU)
TCDD	Ahr1a	1.10	456.0
	Ahr2a	11.27***	323.4
FICZ	Ahr1a	0.04	238.5
	Ahr2a	0.07	243.5***
B[a]P	Ahr1a	3.87	291.1
	Ahr2a	4.60	337.6***
PCB126	Ahr1a	25.23	136.8
	Ahr2a	78.69	85.0*
BNF	Ahr1a	51.39	214.9
	Ahr2a	11.45	43.3***

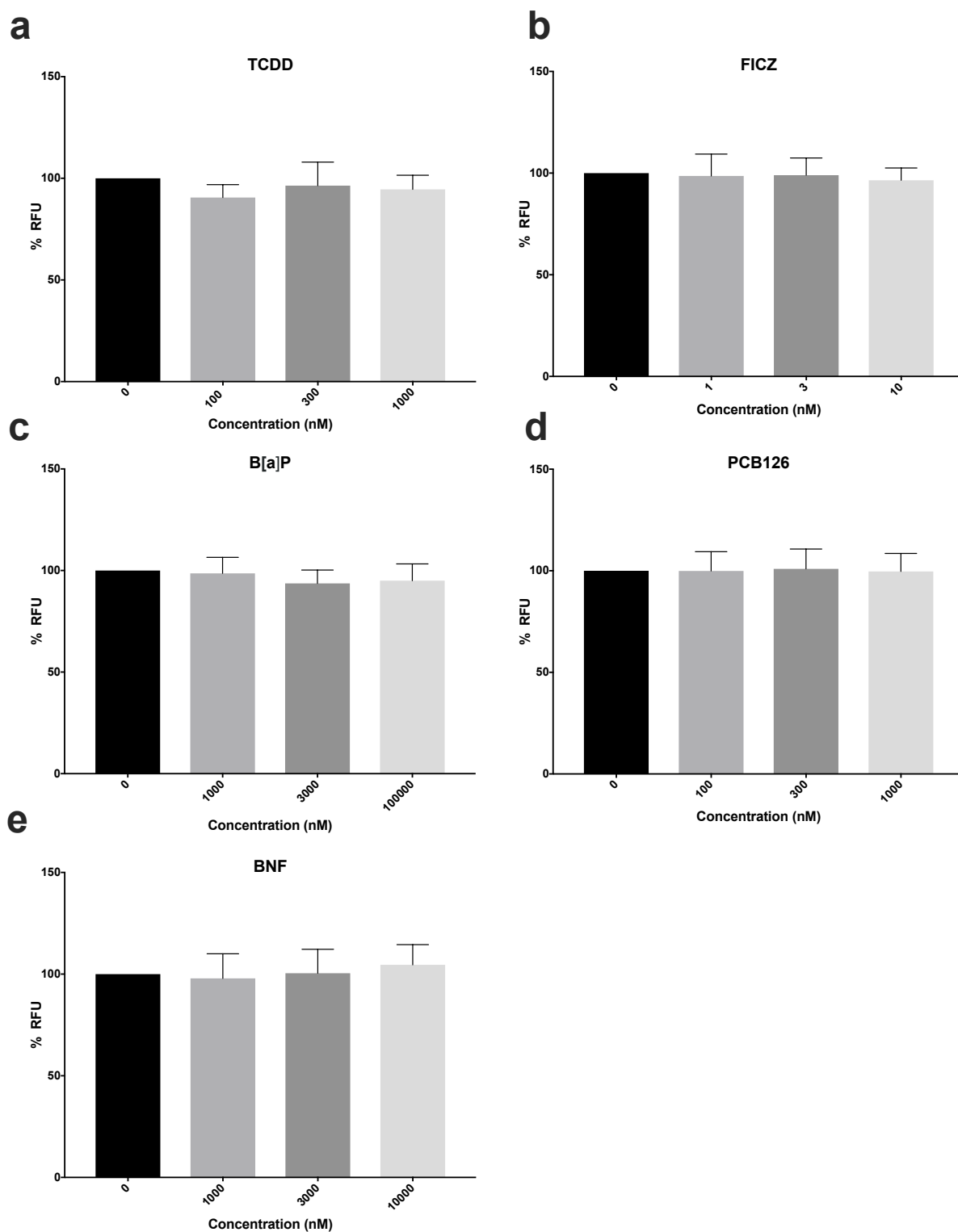


Figure S1. Resazurin activity in COS7 cells exposed to FICZ, TCDD, B[a]P, PCB126 and BNF. COS7 cells were seeded, cultivated for 24 hours, and exposed to the three highest concentrations of TCDD (a), FICZ (b), B[a]P (c) PCB126 (d) and (e) BNF used in the gmAhr reporter gene assay for 12 hours. Fluorescence was measured at 530/590 nm. The cytotoxicity assay was repeated twice and there were four technical replicates in each assay. The data are presented as relative fluorescent units (RFUS) % normalized against the solvent control. One-way ANOVA was used to test for statistical significance.

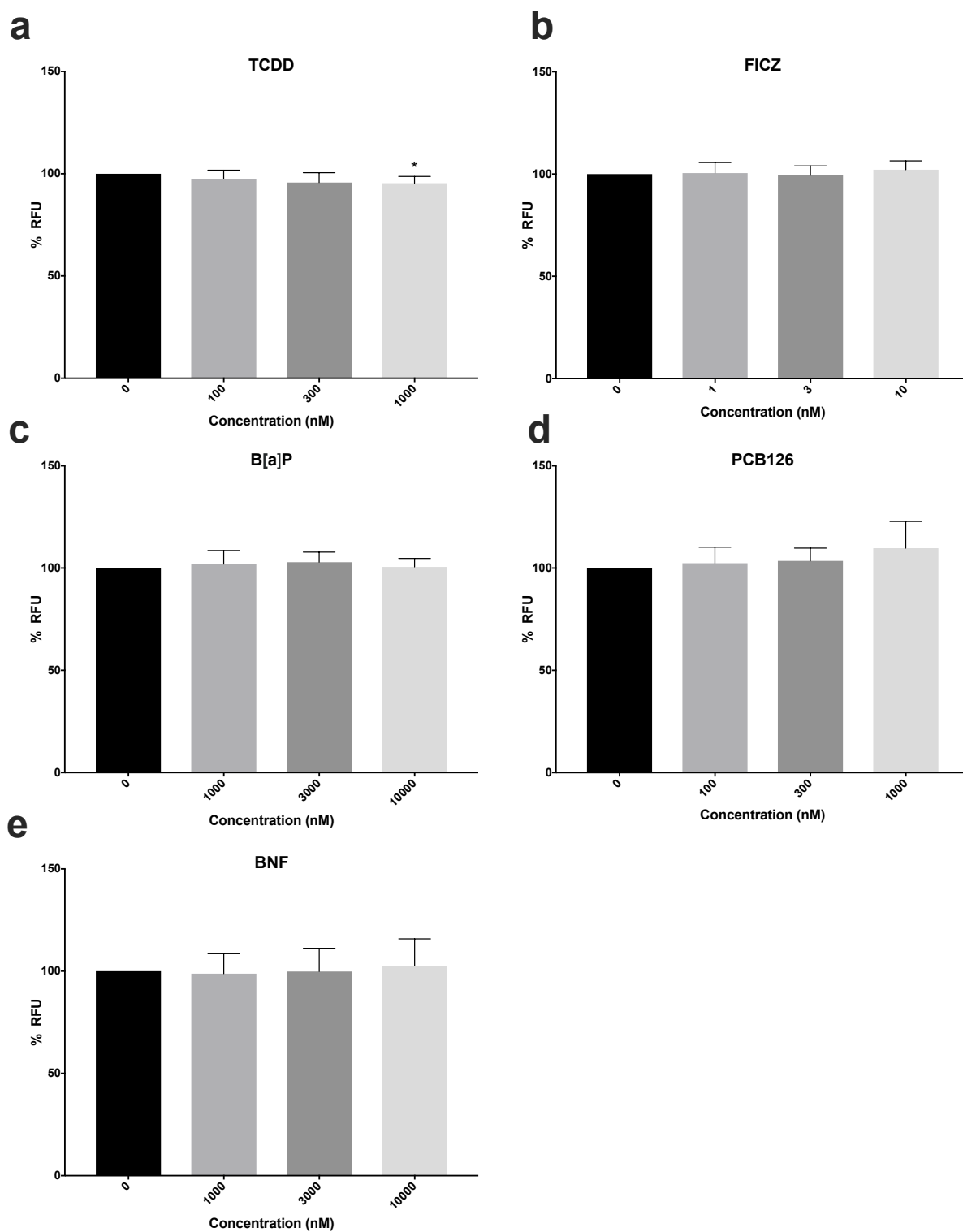


Figure S2. 5-carboxyfluorescein diacetate acetoxyethyl ester (5-CFDA-AM) activity in COS7 cells exposed to FICZ, TCDD, B[a]P, PCB126 and BNF. COS7 cells were seeded, cultivated for 24 hours and exposed to the three highest concentrations of TCDD (a), FICZ (b), B[a]P (c) PCB126 (d) and (e) BNF used in the gmAhr reporter gene assay for 12 hours. Fluorescence was measured at 485/530 nm. The assay was repeated twice and there were four technical replicates in each assay. The data are presented as relative fluorescent units (RFUS) % normalized against the solvent control. One-way ANOVA was used to test for statistical significance. Level of significance is expressed with * ($p < 0.05$).

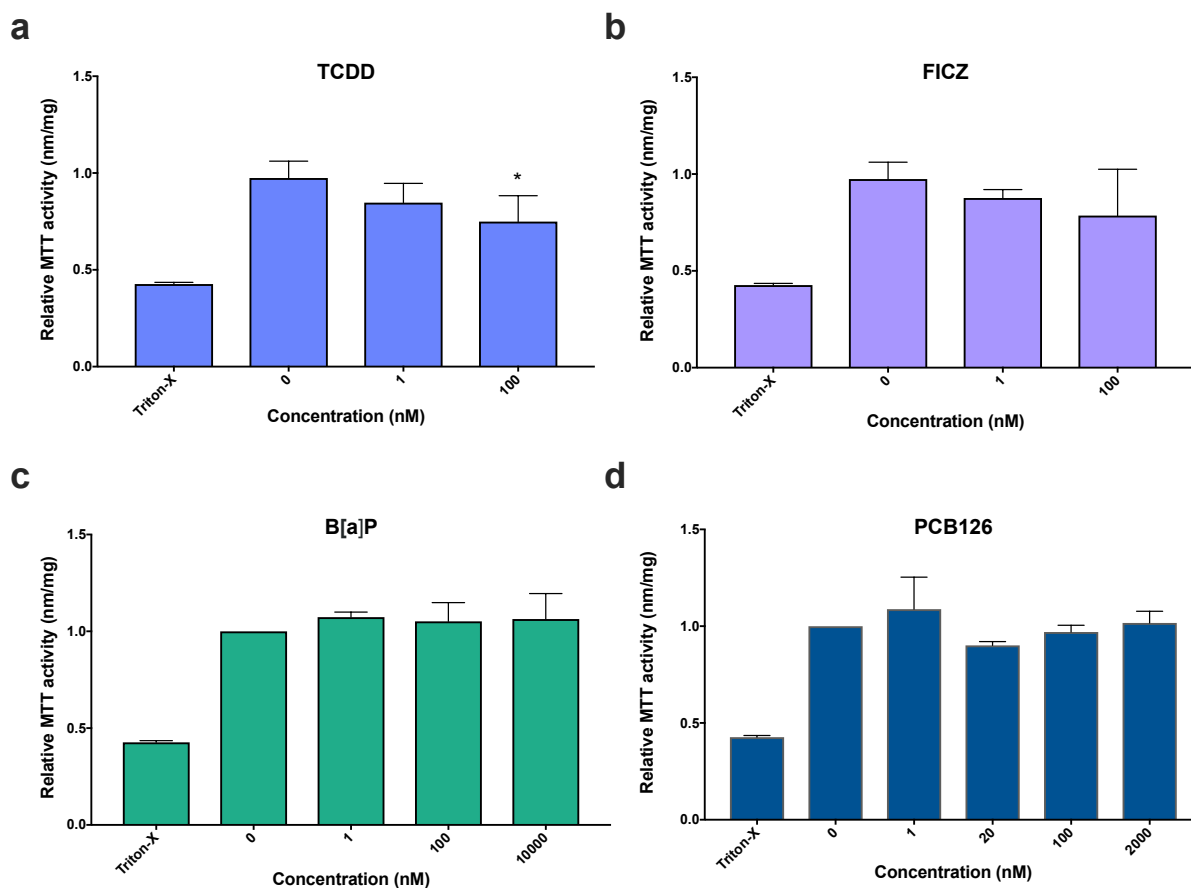


Figure S3. Metabolic activity in precision-cut liver slices (PCLS) exposed to FICZ, TCDD, B[a]P and PCB126. Liver slices (n=3) were prepared and exposed to TCDD (a), FICZ (b), B[a]P (c) and PCB126 (d) at 10°C for 48 hours. Viability was tested in a mix of slices exposed to selected concentrations of each compound in duplicates as described in Yadetie et al. ⁵. Absorbance was measured in 100 μ L aliquots at 590 nm. Relative MTT activity was calculated by normalization of absorbance values (nm) against the weight of each slice (mg) and the solvent control (nm). Triton X-100 was used as a positive control for cell death. The data are presented as mean \pm SEM. Statistical significance was analyzed with Two-way ANOVA. Level of significance is expressed with * ($p < 0.05$).

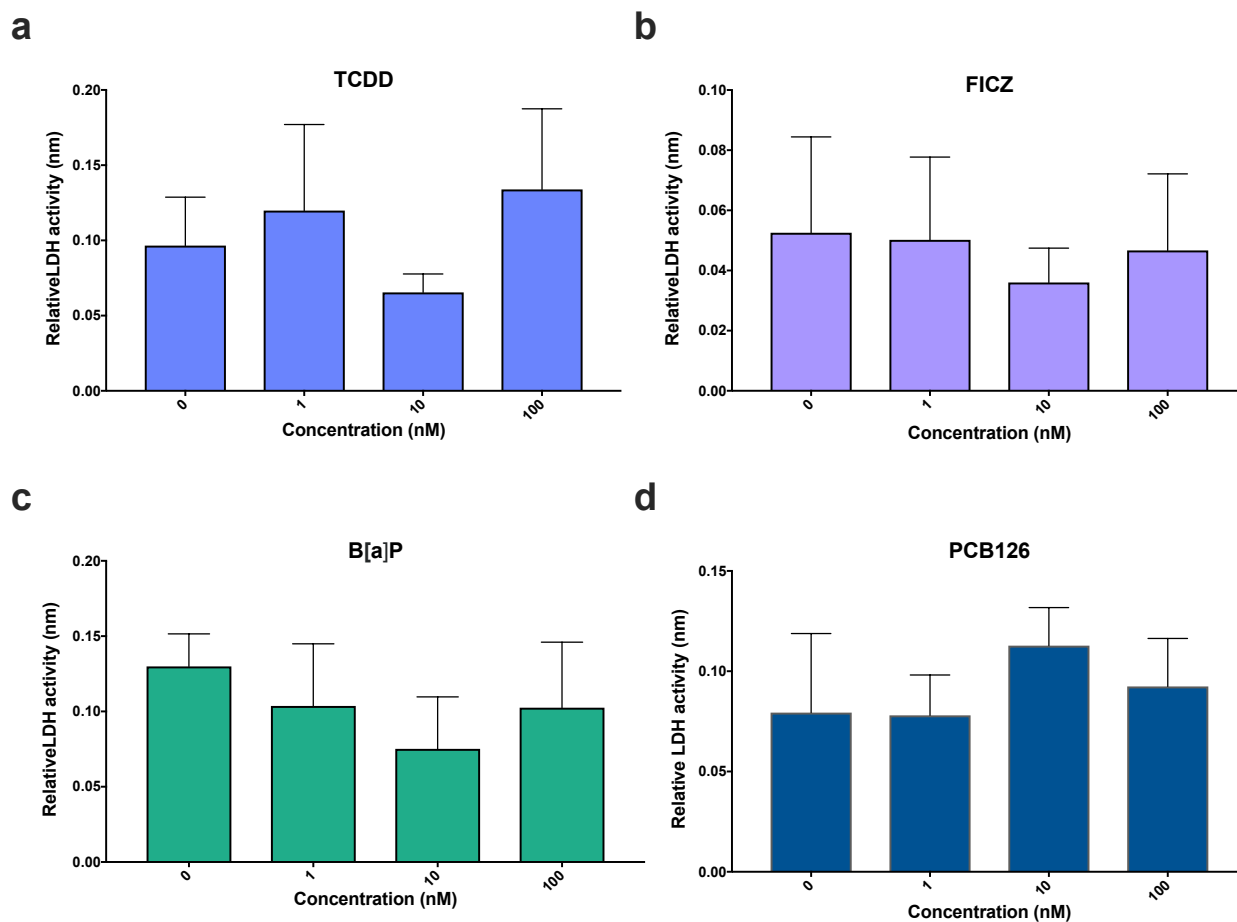


Figure S4. LDH activity in culture medium from precision-cut liver slices (PCLS) exposed to FICZ, TCDD, B[a]P and PCB126. Liver slices (n=3) were prepared and exposed to TCDD (a), FICZ (b), B[a]P (c) and PCB126 (d) at 10°C for 48 hours. Viability was tested in slices culture medium as described in Bizarro et al. ⁷. Absorbance was measured in 50 μ L aliquots 490 nm and 650 nm. The assay was repeated two times. Relative LDH activity was calculated as absorbance (nm) in each exposed group normalized to the solvent control (nm). The data are presented as mean \pm SEM. Statistical significance was analyzed using Two-way ANOVA.

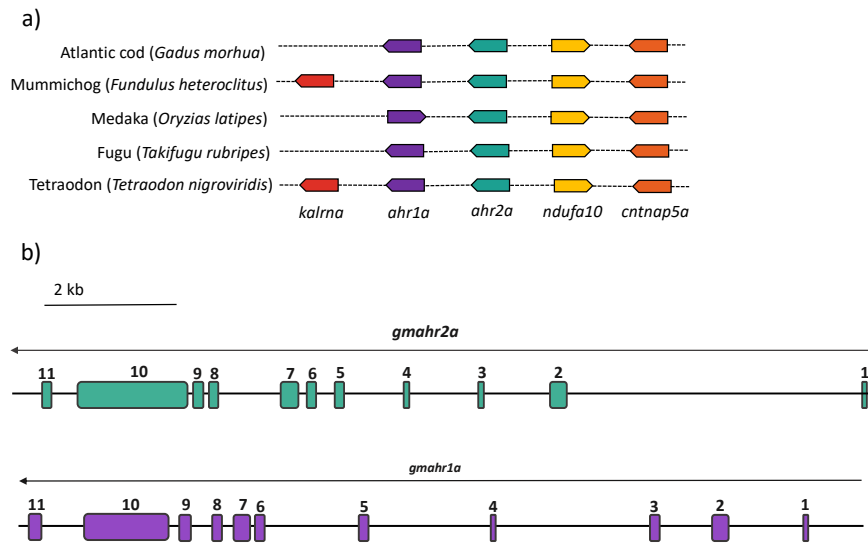


Figure S5. Synteny analyses of the Atlantic cod *gmahr* genes. (a) Synteny analyses of *ahr* genes in selected fish species. *kalrna* (Kalirin RhoGEF kinase a), and *ndufa10* (NADH dehydrogenase ubiquinone 1 alpha subcomplex subunit 10) and *cntnap5a* (contactin-associated protein like 5). Data was obtained from Ensembl. **(b)** Exon/intron distribution of *gmahr1a* and *gmahr2a*. Exons are drawn as colored boxes and their number is indicated above. Arrows indicate the position of the *ahr* genes in the reverse strand (5'-3' direction).

```

gmAhr1a      1  - - - - - M Y A G R K R R K P V Q R A V K Q V S N E G - S K S N P S K R H R D R L N G E L E R L A S L L P F P E E V T T S L D K L S I L R L S V S Y L R      70
gmAhr2a      1  M L G N A G T Y A M K K R K K P V Q K P K K L P G V D G V I K S N P S K R H R D R L N G E L D R L T D L L P F S E D I R T R L D K L S V L R L S V G Y L R      77

gmAhr1a     71  A K N F F S V A L K T S K C N S V V P S G P S S D S N V A V G L S D A W L P E G E L L L Q A L N G F V L V L S A D G T I F Y S S H T I Q D Y L G F H Q T D      147
gmAhr2a     78  V K G F F K A T M K K H N - - - - - G P N G - - Q G R N G V D V A A L S E G D L L L Q A L N G F V I V V T A E G L V F Y S S T I Q D Y L G F H Q S D      145

gmAhr1a    148  V M H Q S V F E L V H T E D Q L E L R K N L H W A L N P P A A T V V A M N Q A S P T E M E Q E S G T A V V T Y N P E Q L P P E N S S F L E R N F V C R F R      224
gmAhr2a    146  V V H Q S V Y E L I H T D D R G M F R E Q L H F A L N P K L Y A T E Q G G D V - - - S A L Q C S S D Q M K Y D P E R L P P E N S S F L E R S F V C R F R      218

gmAhr1a    225  C L L D N S S G F L A L T M Q G R L K F L H G Q H Q R Q D N G K A P P Q L A L F A I A T P L Q P P S I L E I R T K N M I F R T K H K L D F T P M A C D A      301
gmAhr2a    219  C L L D N S S G F L A L K F Q G R L K Y L H G Q S M M G E D G S R V Q S Q L A L F S I A V R V Q T P S I L E I R T K T L I F Q T K H Q L D F T P I G I D N      295

gmAhr1a    302  K G K I V L G Y T E A E L R V R G S G Y Q F I H A A D M L Y C A E N H V R M I K T G E S G L T V F R L L T K D N R W K W Q A N A R L V Y K N G K P D Y I      378
gmAhr2a    296  R G K V V L G Y S E L E L C M R G S G Y Q F I H A A D M M Y C A D N H I R M I K T G E S G L S V F R L L S K S S G W W W Q A N A K L V Y K G R P D F I      372

gmAhr1a    379  I A T Q R P L V E E E G G E H L R K R S M H L P F T Y A T G E A L Y Q S N H P I P G F S D G I H E K N G S K S K K C R S E R A A E G L D P S S L L G A      455
gmAhr2a    373  I A R Q R A L V N A E G E E H L R Q R R L Q L P F S F T T G E A M L Y E V G P S L D V T Q I E - - - - - T S Q - - - - - S F T S G Q Q E E V G G L L G C      438

gmAhr1a    456  L M S Q D E S V Y V Q C P A L E P K L A F H S S F L G D Q L G F S E S D S H P G V D N G W D A - - - - - P G N G I N G A G L P P P A N S F D P L L S T      525
gmAhr2a    439  F L N Q D K N V Y V Q D S E A Q - - L P V D Q V F M E S R A L V - - - - - N V P S D P W Q A L R L Q G D D G G N M I K E E G V T S V S A - - - M M N -      502

gmAhr1a    526  L D S L S L G D Q N S A T A E E E G C S N G E L F R A L E S L G L S A E D L E L L L D E R M I R V E M D - - - - - P D R V - P S L D D I L T N D E I L S      596
gmAhr2a    503  - - - - - A L E D F V E N G E L V S A L E G L D V D A G E L M E - - W E N T L K K L S Q E E N G E N A D Q T K Y E L E S - L L S N D I F A      563

gmAhr1a    597  Y I H D S L E A K S D E A E E - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -      636
gmAhr2a    564  Y V D N V L F K E I A E A N L N T S Q S S C F S P V N N N Q S D L F G Q R A H Y S G S G D T C D M M M F Q S P A V G A N A V S H A K G L S P A V A R Q P M      640

gmAhr1a    637  A V S P A Y V R - T Q L P H Q K A P L - - - - - V - - - - - G Q A P I T L L S Q Q M Q Q H - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -      676
gmAhr2a    641  H N R P A S V A A K G L P P Q T P A L F N S T Q K L S H Y G P A I P E A V P R L S A T P Q - L L T D F F N P S V N L P G L N L P K L P L A S N D L R T F D      716

gmAhr1a    677  K V G Q D W N H N H S C L N G D Y P A Q - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -      731
gmAhr2a    717  P C G Q A S I - - S H Y Q G L A G N A M S N Q M L S N Q T P S K Q T L S N Q T L - A N Q M L S N Q T L S N Q T L S N Q T L S N Q T L A N Q M L S N Q T L S      790

gmAhr1a    732  - Q L D G E Q Q R Q H Q A H P L Y L Q P Q P R P T D T Q G P L S L Q C H A K Q N G E N L N G P S N G T G T D H P L G K S P W Q D F G F G H P P R E N L C      807
gmAhr2a    791  N Q M L S N Q M L S N Q M L S N H T L P N Q T L S T T T L S P Q S L Q P C P L - - - - - T G G P A - - - - - A P M G A N G H F L Q G S I Q Q P A V H M A      856

gmAhr1a    808  G L S N G Q A P V T A P S L D A C I D F A M P E L Q S M D Y - - - - - P V N G G G L Y Q G V E A A V S S Y Q R M S H K R Q Q E Q H F Q P P L A H T A L      877
gmAhr2a    857  P N - - - - - V V A P A P S N L P H N D F S M P A N P S E N S A L F T G N C M V Q G G A P F Q T H S N - - - - - H R A P Q W Q P D L Q R Q H Q P L A H A S V      924

gmAhr1a    878  E - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -      918
gmAhr2a    925  A Q N S H T L P A G S H S Q A F E S Q R L A G L W A Q N Y N G M N Q P P P H R G L A G P - - - - - L R - - - - - T N P S S C M L D K P L H P P P A T H T L      991

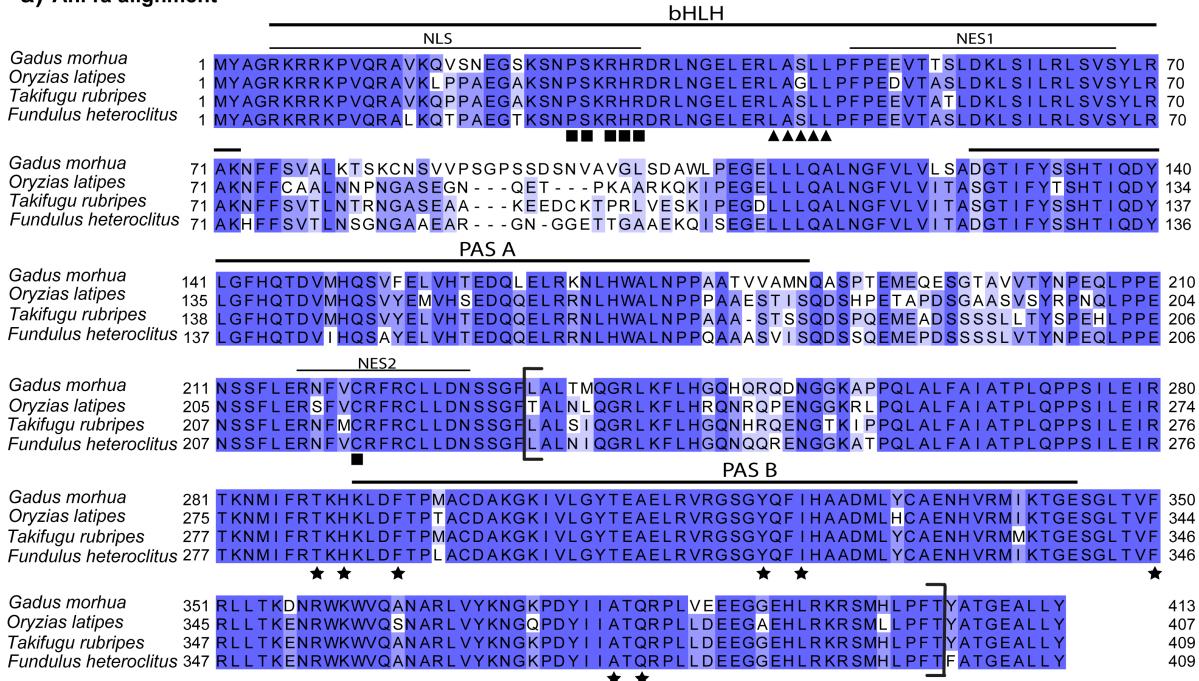
gmAhr1a    919  - - - - - Y Q G N C A L P N G N G A A P P N G Q M P G P D S L P S L A D P Q - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -      951
gmAhr2a    992  G H P H P H T N G N L A S A N G T G L V P A M Q L C Q R G N E A P A L H Q S P P K G Y V V Q W G Q G M P P M G T A A T G Q E N A A F G A T P R Q L L P A      1068

gmAhr1a    952  - - - - - V T D F Y L * - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -      958
gmAhr2a    1069 N I S S G A P D D M A A M P H Y L D G N K H T Q M L S L P T E D N D L L A I P P L V D G N I Y F S D Q S Q L N C C N F *      1128

```

Figure S6. Multiple sequence alignment of full length Atlantic cod Ahr1a and Ahr2a amino acid sequences. gmAhr1a and gmAhr2a were aligned using Clustal Omega. The alignment was edited in Jalview and % identity was used for coloring.

a) Ahr1a alignment



b) Ahr2a alignment

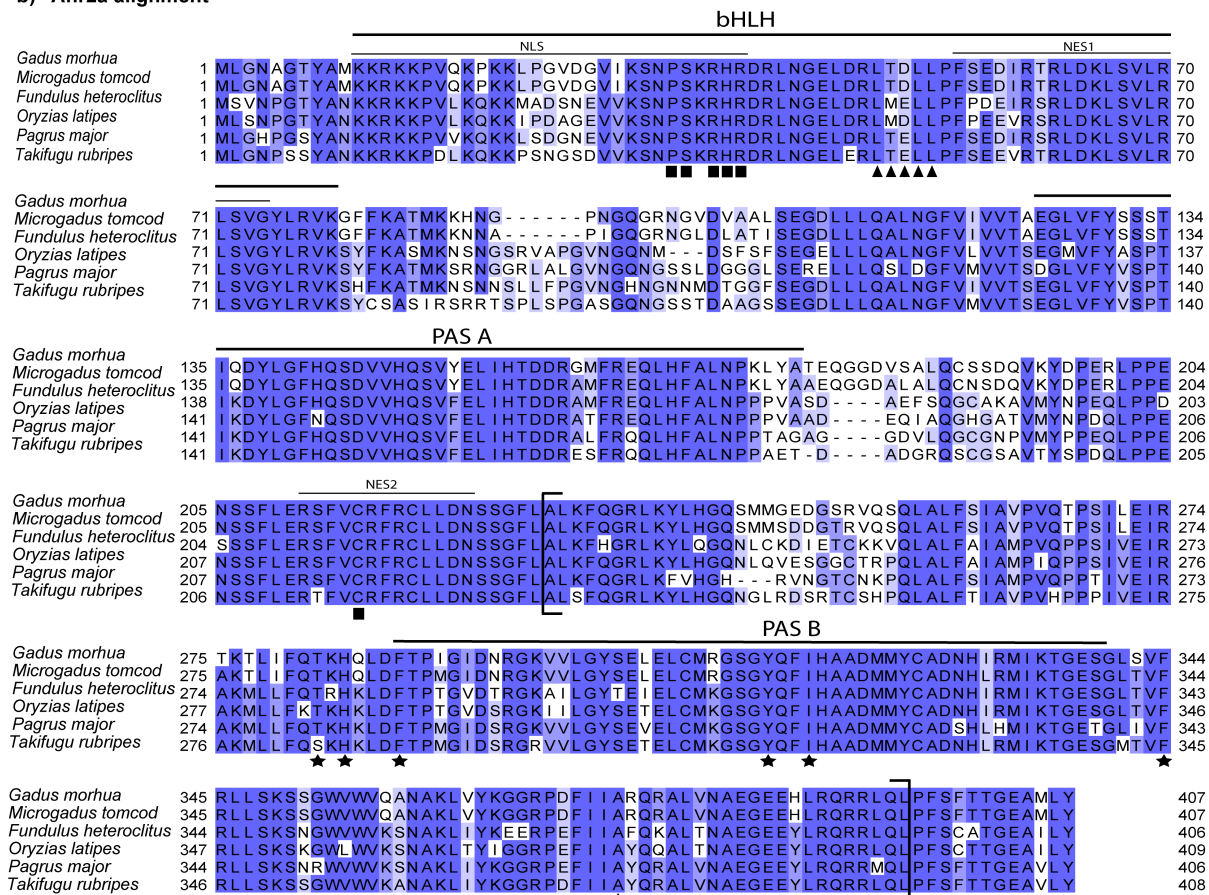


Figure S7. Multiple sequence alignments of N-terminal regions of Ahr1a and Ahr2a from selected fishes. N-terminal amino acid sequences of (a) Ahr1s and (b) Ahr2s were aligned using Clustal Omega and colored by % identity in Jalview. The basic helix loop helix (bHLH) region, nuclear export signals (NES) 1 and 2, nuclear localization signal (NLS), PAS A and B domains are indicated above the alignments, and the ligand binding domain is defined by

brackets. Amino acids present in the LxxLL motif are indicated with triangles. Amino acids involved in binding to xenobiotic response elements (XRE) are indicated in squares and important amino acids for TCDD binding are labelled with stars (see papertext for citations).

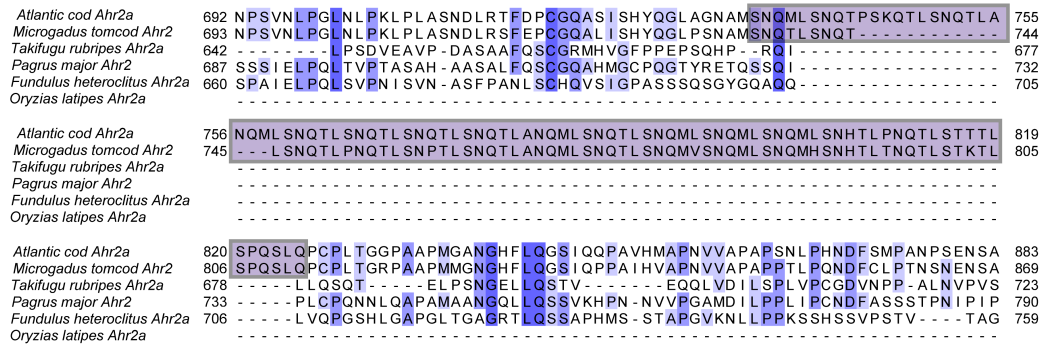


Figure S8. C-terminal repetitive sequence region in Atlantic cod Ahr2a and tomcod Ahr2a. The repetitive sequence region (SNQTL-SNQML) amino acid present (735-804) in gmAhr2a and *Microgadus tomcod* Ahr2 is indicated in a box. Amino acid sequences were aligned using Clustal Omega and colored by % identity using Jalview.

<i>gmAhr2a</i>	1	--MLGNAGTYAMKRRKPKVQKPKKLP	PGVDGVI	KSNPSKRHRDRLNGELDR	LDLPPFSEDIRTR	LDKLSVLR	LSVGLVLR	KGFF	KATMKKH	90
<i>gmAhr1a</i>	1	-----MYAGRRKRPVQRAVKQV	SNELG	SKSNPSKRHRDRLNGELER	LSLPPPEEVT	SLDKLS	ILRLSV	SYLRAK	NF	91
<i>mmAHR</i>	1	MSSGANIT	YASRKRKPKVQKTVP	KIPAE	G	KSNPSKRHRDRLNT	ELDR	LSLPPFQD	VINKL	92
<i>hsAHR</i>	1	MSSSANIT	YASRKRKPKVQKTVP	KIPAE	G	KSNPSKRHRDRLNT	ELDR	LSLPPFQD	VINKL	93
<i>gmAhr2a</i>	91	----GPNQGR	-----NGVDVAA	SEGDL	LLQALNGF	IVVTA	EGLV	FYSST	I	94
<i>gmAhr1a</i>	84	CNSVVP	SGPSSD	SNVAVGL	SDAW	PEGEL	LLQALNGF	VLVLSAD	GT	95
<i>mmAHR</i>	91	ADNRGG	QDCRAQ	-I	---	RDWQD	LOEGEF	LLQALNGF	VLVV	96
<i>hsAHR</i>	92	TERNGG	QDCRAANF	---	---	REGLN	LOEGEF	LLQALNGF	VLVV	97
<i>gmAhr2a</i>	174	KLYATE	QGQGDV	---	---	SALQ	SSDQ	KYD	PER	175
<i>gmAhr1a</i>	176	PAATV	VAMNQAS	PTE	MEQES	GTAV	VY	Y	NP	176
<i>mmAHR</i>	179	D	---	SAQGV	---	DEAH	GPQAA	V	Y	177
<i>hsAHR</i>	181	SQCTE	SGQG	---	---	EEAT	GLPQT	V	C	178
<i>gmAhr2a</i>	262	AVFVQT	PSILE	IRTK	TL	FQTK	HL	D	F	263
<i>gmAhr1a</i>	268	ATPLQPP	SILE	IRTK	KNFI	FRTK	HK	L	D	264
<i>mmAHR</i>	253	ATPLQPP	SILE	IRTK	KNFI	FRTK	HK	L	D	265
<i>hsAHR</i>	269	ATPLQPP	SILE	IRTK	KNFI	FRTK	HK	L	D	266
<i>gmAhr2a</i>	354	VWQAN	AKL	VYK	GGR	PDFI	I	A	R	355
<i>gmAhr1a</i>	360	KWQAN	ARL	VYK	NGK	PDYI	I	A	T	356
<i>mmAHR</i>	355	RWQSN	ARL	I	YR	NGR	PDYI	I	A	357
<i>hsAHR</i>	361	TWQSN	ARL	I	YR	NGR	PDYI	I	A	358
<i>gmAhr2a</i>	432	VGG	LL	GC	F	L	N	D	K	433
<i>gmAhr1a</i>	449	PSS	LL	G	A	L	M	S	O	440
<i>mmAHR</i>	447	PSS	LL	M	S	A	L	I	Q	441
<i>hsAHR</i>	453	PSS	LL	A	M	M	Q	D	E	442
<i>gmAhr2a</i>	504	LED	V	N	C	E	L	V	S	505
<i>gmAhr1a</i>	540	EEG	C	S	N	G	E	L	F	506
<i>mmAHR</i>	522	L	F	P	D	N	K	N	D	507
<i>hsAHR</i>	530	L	F	Q	D	S	K	N	D	508
<i>gmAhr2a</i>	594	S	D	L	F	G	R	A	H	595
<i>gmAhr1a</i>	622	A	---	V	I	N	T	A	A	623
<i>mmAHR</i>	603	S	---	C	---	M	L	Q	E	604
<i>hsAHR</i>	610	S	---	C	---	M	V	Q	E	605
<i>gmAhr2a</i>	683	T	P	L	L	T	D	F	F	684
<i>gmAhr1a</i>	688	H	S	G	L	N	G	D	Y	689
<i>mmAHR</i>	646	H	T	I	N	G	T	F	A	647
<i>hsAHR</i>	644	H	M	Q	V	N	G	M	F	648
<i>gmAhr2a</i>	766	N	Q	T	L	S	N	Q	T	767
<i>gmAhr1a</i>	762	L	S	---	---	L	Q	H	A	768
<i>mmAHR</i>	698	---	---	Y	T	Q	N	F	---	699
<i>hsAHR</i>	696	---	---	Y	T	Q	N	F	---	700
<i>gmAhr2a</i>	843	F	L	Q	S	I	Q	P	A	844
<i>gmAhr1a</i>	816	---	---	V	T	A	P	S	L	817
<i>mmAHR</i>	738	---	---	D	F	V	S	C	L	739
<i>hsAHR</i>	736	---	---	E	D	F	V	T	C	740
<i>gmAhr2a</i>	922	---	---	A	S	V	A	N	S	923
<i>gmAhr1a</i>	885	Q	P	C	R	L	G	P	Y	886
<i>mmAHR</i>	788	---	---	Q	S	S	E	I	P	789
<i>hsAHR</i>	787	---	---	Q	N	V	P	L	R	790
<i>gmAhr2a</i>	1011	V	P	A	M	Q	L	C	R	1012
<i>gmAhr1a</i>	932	A	P	P	N	G	M	P	G	933
<i>mmAHR</i>	824	T	T	H	L	Q	L	H	H	825
<i>hsAHR</i>	824	T	T	H	L	Q	L	H	H	826
<i>gmAhr2a</i>	1103	L	L	A	I	P	P	L	V	1104
<i>gmAhr1a</i>										
<i>mmAHR</i>										
<i>hsAHR</i>										

Figure S9. Multiple sequence alignment of N-terminal regions of Atlantic cod, human and mouse Ahrs. N-terminal amino acid sequences of, gmAhr1a, gmAhr2a, mmAHR and hsAHR were aligned using Clustal Omega and colored by % identity using Jalview. The amino acids constituting the “TCDD-binding-fingerprint” in mammalian AHRs, are indicated with stars.


```

gmahr1a 1 -----MYAGRRKRKPVQRVAVKQVSNESG-KSNPSKRHRDRNLNSELERLASLLPFPEEVTSLDKLSILRLSVSYLRAKNFFSVALKTSKCNVSVVPSG 91
drahr1b 1 -----MYAGRRKRKPVQRTVKQPPAEGV-KSNPSKRHRDRNLNSELDRLSLLPFPEEVTSLDKLSILRLSVSYLRAKNFFSVTLKNHNCSNGLSSNN 91
atahr1 1 -----MYASRRKRKPVQKSAKPSL-PEA-KSNPSKRHRDRNLNTELERLASLLPFPPQDVI SKLDKLSVLRSLVSYLRAKSFFTASINSKSCRKRPAGNG 90
drahr1a 1 -MSSSN IYASRRKRKPVQKSVKQPL-TCV-KSNPSKRHRDRMNEVLASLAREI PPFQEI VSTLDKLT IRLSVSYLRAKSHFK I TLNSKSSSQANNT 95
gmahr2a 1 MLGNAGTYAMKKRKKPVQKPKLPGVDGVI-KSNPSKRHRDRNLNSELDRLTDLLPFSEDIRTRLDKLSVLRSLVSYLRAKSFFTATMKKHNG----- 91
atahr2 1 MLATGGIYAVKKRKKR I QK I P K P P P T D G V - K S N P S K R H R D R N L N S E L D K L T S L L P F S E D V C A R L D K L S V L R L S V G Y L K V K S F N A T M K K - N N V G W L R E K 96
drahr2 1 MSAGIGTYAVKKRKKR I QK I P K P P P D G V - K S N P S K R H R D R N L N S E L D K L T N L L P F S E D V R A R L D K L S V L R L S V G Y L K V K S F N A T I K K T G G N G W L N D R 97

gmahr1a 92 PS---SDSNVAVGLSDAWLPEGELLQALNGFVLLVLSADGTIFYSSTIQDYLGFGHQTDMVHQSVFELVHTEDQLELRKNLHWALNPPAATVVMNQ 186
drahr1b 92 SN---HDNSKATGLVDGCLREGELLQALNGFVLLVVTAEGLIFYSSTIQDYLGFGHQTDMVHQSVFELVHTEDQAFRRNLHWALNPPASTQTES- 185
atahr1 91 LD-----VNKLPCEDELLEGELLQALNGFVLLVVSADGSVFYVSPPTLQDYLGFGHNSDVIHQSVYELIHTEDRAEFQRQLHWALNPPGPPDSGQIVP 180
drahr1a 96 K-----QAQELPEGEFLLOAINGFLLVVTSSGTVFYVSSNI EDYLGFGHNSDVIHQSVYELIHTEDRHEFQRQLHWALYFGFTPDSRQLVQ 180
gmahr2a 92 ----PNGQGRNGVDVAALSEGDL LQALNGFVLLVVTAEGLVFYSSSTIQDYLGFGHNSDVIHQSVYELIHTEDRGMFREQHLFALNPKLYATEEQGGDV 184
atahr2 97 PGTFGGNGQVSPQTNGTSFSEGDL LQALNGFVLLVVTAEGLVFYSSSTIQDYLGFGHNSDVIHQSVPEL IHTDDRAMFRRQLHFALNPNPFEEQGTTEL 194
drahr2 98 SGTFGGNGQTASSLDGVNFSSEGDL LQALNGFVLLVVTAEGLVFYSSSTIQDYLGFGHNSDVIHQSVFEL IHTDDRAMFRRQLHFALNPNSSDS-ADGSEA 194

gmahr1a 187 SPTMEQESGTA VV T N P E Q L P P E N S S F L E R N F V C R F R C L L D N S S G F L A L T M Q G R L K F L H G Q H Q R Q D N G K A P P Q L A L F A I A T P L Q P P S I L E I R T K N M 284
drahr1b 186 -SEGDGPANMNSLVL CNP D Q L P P E N S S F L E R N F V C R F R C L L D N S S G F L A L N F Q G R L K F L H G Q N R R L D G G Q M P P Q L A L F A I A T P L Q P P S I M E I R T K N M 282
atahr1 181 AA---DNGLSLPVTYNPEKLPENS AFLERNFVCRFRCLLDNSSGFLALNFQGR LKFLHGQNTKSKDGSTIPPOLALFVVATP LQPPSILEIRTNF 275
drahr1a 181 AS---PD---ASRTCYSPEQLSLENSTCLERNFICRLRCLLDSTSGFLAVNFQGR LKFLYQNESTADGKRIPPOLALFALACLP LQPPSILEIRTKNL 272
gmahr2a 185 SAL---QCSDDQVKYDPERLPPENSSFLERSFVCRFRCLLDNSSGFLALNFQGR LKFLHGQSMMGEDGSRVQSALALFSI AVPVQTPSILEIRTKTL 278
atahr2 195 MQSGSNSTGLSNVVDYDQLVPPENSSFLERSFCCRFRCLLDNSSGFLALNFQGR LKFLHGQNKVSEDTGLVPSQALFAIATP LQPPSILEIRTKTL 292
drahr2 195 MQSSSD--ITRDMVNNPQH I P P E N S S F L E R S F C C R F R C L L D N S S G F L A L N F Q G R L K Y L H G Q N K A E D G T L A H P Q L A L F I I A T P L Q P P S I L E I R S K T L 290

gmahr1a 285 IFR TKHKLDFTPMACDAKGI VLGYTEAE LRVRGSGYQF IHAADMLYCAENHVRMIKTGESLTVFRLLTKDNRWVWQANARLVYKNGKPDY I I A T Q 382
drahr1b 283 IFR TKHKLDFTPMACDAKGI VLGYTEAE LRVRGSGYQF IHAADMLYCAENHVRMIKTGESLTVFRLLTKDNRWVWQANARLVYKNGKPDY I I A T Q 380
atahr1 276 IFR TKHKLDFTP TACDAKGI VLGYTEAE LCVRGTYQF IHAADMLYCAENHVRMIKTGESGMTVFRLLTKQNRWVWQANARLVYKNGRPEY I I A T Q 373
drahr1a 273 MFKITKYKLDFTPIACD TNWNF VLGYTEAE LCNVSGSGYQF IHAADMMYCAEGHMRMRTGETGLTVFRLLTKQNRWVWQANARLVYKNGQPDY I I T S H 370
gmahr2a 279 I F Q T K H Q L D F T P I G I D N R G K V V L G Y S E L C M R G S G Y Q F I H A A D M M Y C A D N H I R M I K T G E S G L S V F R L L S K S S G W W W Q A N A K L V Y K G G R P D F I I A R Q 376
atahr2 293 I F Q T K H K L D F T P M G C D T R G K V V L G Y T E T E L C M R G T G Y Q F I H A A D M M Y C A D N H V R M I K T G E S G L T V F R L L T K N G S W W W Q A N A R L I Y K G G R P D F I I A R Q 390
drahr2 291 L F Q T K H K L D F T P M G I D T R G K V V L G Y T E I E L C M R G S G Y Q F I H A A D M M Y C A D N H I R M I K T G E S G L T V F R L L S K G G T W I W W Q A N A R L V Y K A G R P D F I I A R Q 388

gmahr1a 383 RPLVEEEGGEHLRKRSMHLPFTYATGEALYQSNHPIPGFSDG IHEKNG-SKSKKCRSERAAREGLDFSSLLGALMSDDES VYVCQPA ★★ 469
drahr1b 381 RPLVEEEGGEHLRKRSMHLPFTFATGEALYQINYPMLGFPTLQDKGKNNKTKSKVNKSKDDLDPSLLGAMMRQDES VYVCQPA 468
atahr1 374 RALSDNEGL ENLRKRNLKLPFN FATGEAVLYETTFPLAMTQPM-HAKAGKTSATGASSKARDQESLDPN SLLGAILKQDES IYVCAPA 460
drahr1a 371 RVI TAEEGEENLRNRRAMMLPFSFTTGDVLCAMNCP TSSDPA-----PSDNGTQPKTTRTVNPD SLLVSLKQPKS IYVSPGD 448
gmahr2a 377 RALVNAEEGEEHLRQRRLQLPFSFTTGEAMLVEVGPSLDVTQ IETSQS-----F---TSGQQEEVGGLLGCF LNODKNVYVQD-- 450
atahr2 391 RAL TNEEGEHLRQRRLQLP FN FATGEAVLYETGSTGDDFSNIFPAT-----NKKGAEQKS I D P N S L L G A R L K O D Q S I Y V S H A A 469
drahr2 389 RAL TNEEGEHLRQRRLQLP FN CATGEGVLYEVGPTLDVAEIQNQS K-----GQKMLNPPSLD PD SLLG SMLKQDHS LY SQNND 467

```

Figure S10. Multiple sequence alignment of N-terminal regions of Atlantic cod, white sturgeon and zebrafish Ahrs. N-terminal amino acid sequences of Atlantic cod Ahrs (gmAhr1a and gmAhr2a), zebrafish Ahrs (drAhr1b, drAhr1a, drAhr2)), and white sturgeon Ahrs (atAhr1 and arAhr2) were aligned using Clustal Omega. Percent identity coloring was done in Jalview. Amino acid residues shown to be important in TCDD binding are indicated with stars.

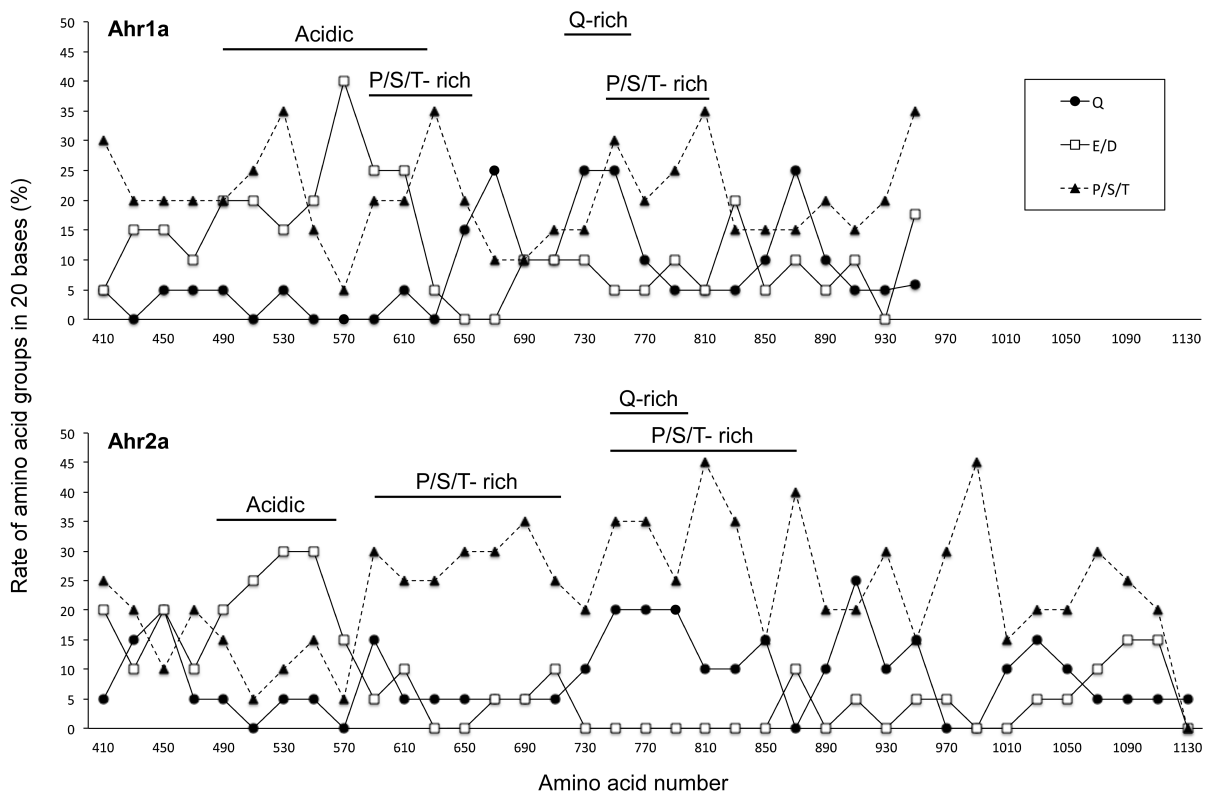


Figure S11 Transactivational sub-domains in Atlantic cod Ahrs C-terminal region. The percentage of the three amino acid groups: Q-rich (glutamine), acidic (aspartic acid and glutamic acid) and P/S/T-rich (proline, serine and threonine) was plotted every 20 bases in the C-terminal region of the sequences. The amino acid groups with the highest percentage were anticipated as transactivational sub-domains. The bars at the top of the figures indicate the predicted sub-domains of the gmAhrs.

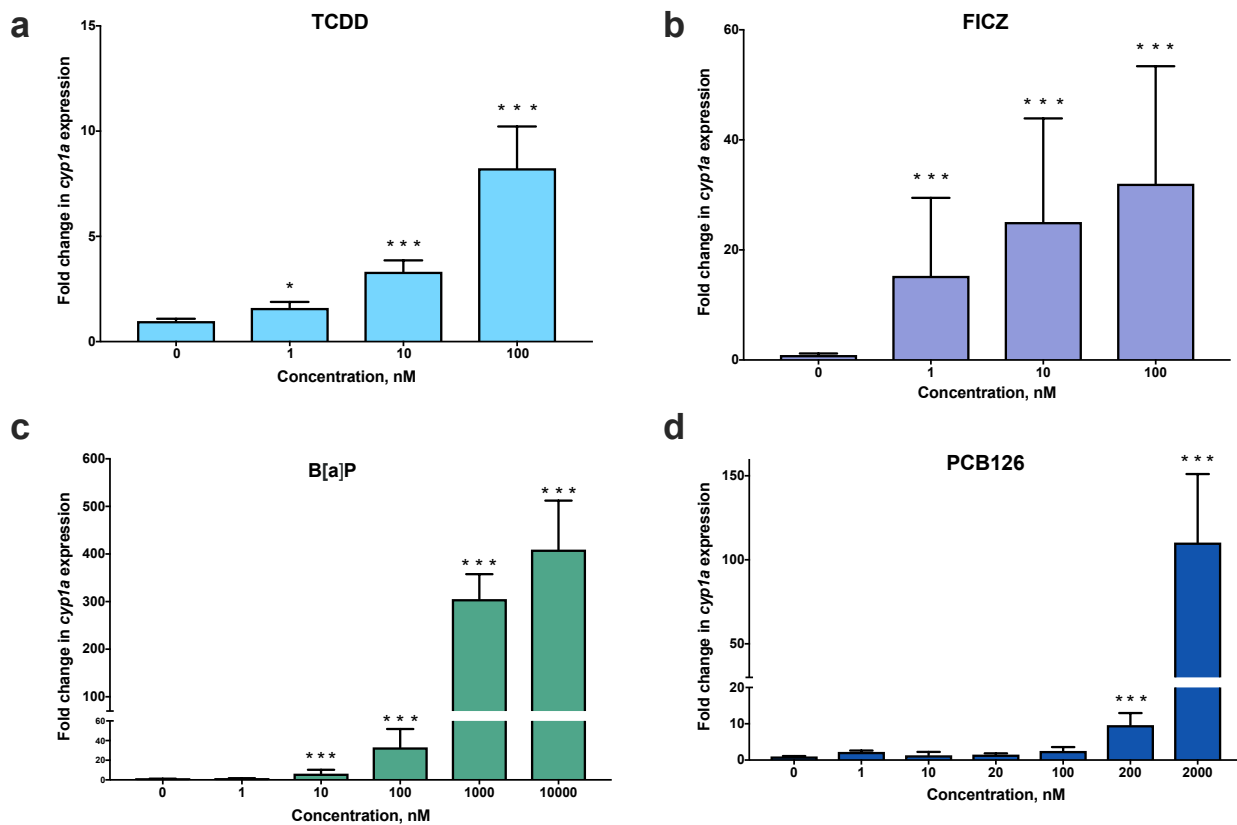


Figure S12. Induction of *cyp1a* from precision-cut liver slices (PCLS) exposed to FICZ, TCDD, B[a]P and PCB126. Liver slices were prepared and exposed to TCDD (1,10,100 nM) (n=5) (a), FICZ (1,10,100 nM) (n=6) (b), B[a]P (1,10,100,1000,10000 nM) (n=6) (c) and PCB126 (1,10,20,100,200,2000 nM) (n=7) (d) at 10°C for 48 hours. Expression of *cyp1a* was measured by qPCR analyses and normalized against the reference gene *arp2*. Fold induction was calculated in comparison with the solvent control. The data are presented as mean \pm SEM. Expression data were analyzed using GLMM models with gamma log10 distribution and fish as a random effect in RStudio v1.2.1335. Level of significance is expressed with * ($p < 0.05$) or *** ($p < 0.001$).

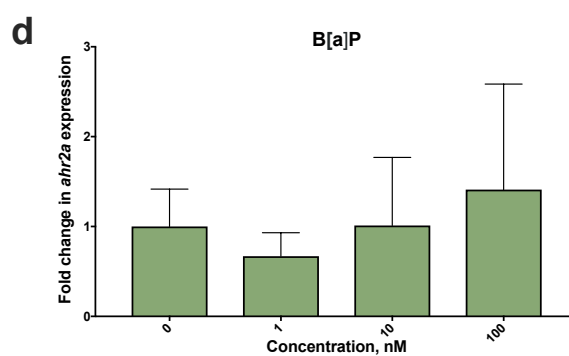
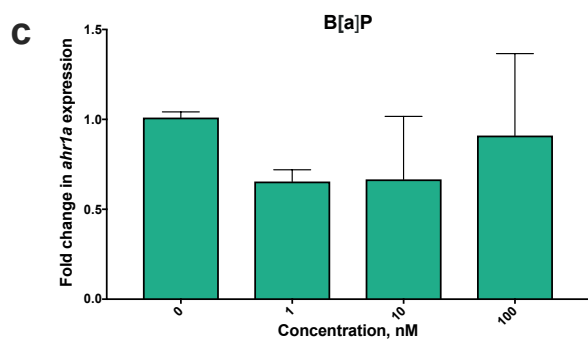
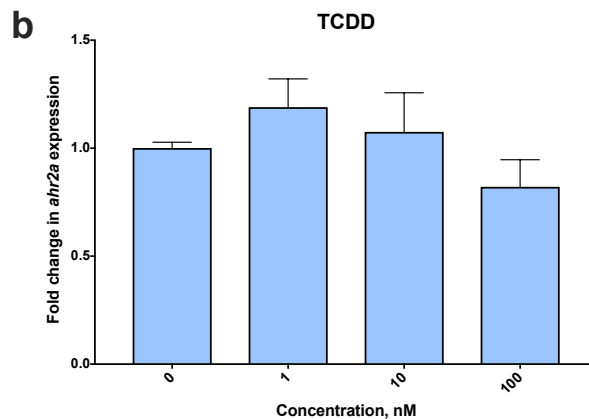
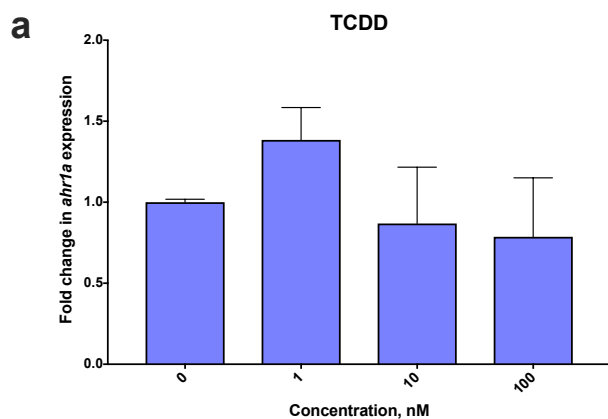


Figure S13. Expression of Atlantic cod *ahr1a* and *ahr2a* in PCLS exposed to TCDD or B[a]P. Liver slices (n=3) were prepared and exposed to TCDD (**a,b**) or B[a]P (**c,d**) at 10°C for 48 hours. Expression of *ahr1a* (**a,c**) and *ahr2a* (**b,d**) was measured by qPCR analyses and normalized against the reference gene *arp2*. Fold induction was calculated in comparison with the solvent control. The data are presented as mean ± SEM.

References

- (1) Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D. L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M. A.; Huelsenbeck, J. P. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Syst. Biol.* **2012**, *61* (3), 539–542. <https://doi.org/10.1093/sysbio/sys029>.
- (2) Ritz, C.; Baty, F.; Streibig, J. C.; Gerhard, D. Dose-Response Analysis Using R. *PLoS One* **2015**, *10* (12), 1–13. <https://doi.org/10.1371/journal.pone.0146021>.
- (3) Pérez-Albaladejo, E.; Rizzi, J.; Fernandes, D.; Lille-Langøy, R.; Karlsen, O. A.; Goksøyr, A.; Oros, A.; Spagnoli, F.; Porte, C. Assessment of the Environmental Quality of Coastal Sediments by Using a Combination of in Vitro Bioassays. *Mar. Pollut. Bull.* **2016**, *108*, 54–61. <https://doi.org/10.1016/j.marpolbul.2016.04.063>.
- (4) Ellesat, K. S.; Yazdani, M.; Holth, T. F.; Hylland, K. Species-Dependent Sensitivity to Contaminants: An Approach Using Primary Hepatocyte Cultures with Three Marine Fish Species. *Mar. Environ. Res.* **2011**, *72* (4), 216–224. <https://doi.org/10.1016/J.MARENVRES.2011.09.003>.
- (5) Yadetie, F.; Zhang, X.; Hanna, E. M.; Aranguren-Abadía, L.; Eide, M.; Blaser, N.; Brun, M.; Jonassen, I.; Goksøyr, A.; Karlsen, O. A. RNA-Seq Analysis of Transcriptome Responses in Atlantic Cod (*Gadus Morhua*) Precision-Cut Liver Slices Exposed to Benzo[a]Pyrene and 17 α -Ethinylestradiol. *Aquat. Toxicol.* **2018**, *201*, 174–186. <https://doi.org/10.1016/j.aquatox.2018.06.003>.
- (6) Crawley, M. J. *The R Book*, Second.; Wiley: New York, 2012.
- (7) Bizarro, C.; Eide, M.; Hitchcock, D. J.; Goksøyr, A.; Ortiz-Zarragoitia, M. Single and Mixture Effects of Aquatic Micropollutants Studied in Precision-Cut Liver Slices of Atlantic Cod (*Gadus Morhua*). *Aquat. Toxicol.* **2016**, *177*, 395–404. <https://doi.org/10.1016/J.AQUATOX.2016.06.013>.
- (8) Grün, F.; Venkatesan, R. N.; Tabb, M. M.; Zhou, C.; Cao, J.; Hemmati, D.; Blumberg, B. Benzoate X Receptors a and b Are Pharmacologically Distinct and Do Not Function as Xenobiotic Receptors *. *J. Biol. Chem.* **2002**, *277* (46), 43691–43697. <https://doi.org/10.1074/jbc.M206553200>.