SUPPLEMENTARY INFORMATION for:

The cytochrome P450 2AA gene cluster in zebrafish (*Danio rerio*); expression of *CYP2AA1* and *CYP2AA2* and response to phenobarbital-type inducers

Akira Kubota^{a,1}, Afonso C.D. Bainy^{a,b,1}, Bruce R. Woodin^a, Jared V. Goldstone^a and John J. Stegeman^{a,*}

^aBiology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, 02543, USA.
^bDepartamento de Bioquímica, CCB, Universidade Federal de Santa Catarina, Florianopolis, SC, 88040-900, Brazil.

¹These authors had equal contribution. *Corresponding author: John Stegeman (jstegeman@whoi.edu) Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA Phone: 508-289-2320 Fax: 508-457-2169

Number of pages: 5 pages

Gene	Primer sense	Primer antisense									
CYP2AA1	5'-TTCCATTTTCACTGGGACCG-3'	5'-CGAACAAGACCCATGATGCC-3'									
CYP2AA2	5'-GCCTTTTGTGGGAAACTTAC-3'	5'-AGCCAGTTGGATTGTATTGATGC-3'									
ARNT2	5'-CACCTTTGGATCACATCTCATTG-3'	5'-TCACCCTCCTTAGACGGACC-3'									
EF1a	5'-CAACCCCAAGGCTCTCAAATC-3'	5'-AGCGACCAAGAGGAGGGTAGGT-3'									

Table S1. Gene specific primers for quantitative real time PCR

	10 20 30 40 (A) 50 80 70 80															90																
Hs CYP2B Rn CYP2B Mm Cyp2b Fh CYP2N	86 M 81 M 910 M 910 M - W	LYNH	LLV	 	L S V P S I P S V D L K A		LAL LAL LAL FIF	LTG LVG LVG	FL FL LL	L L L L L L L L L I A D	V Q R V R G A R G F L R	H P H P H P N R	N T H K S R K S R K P A	DRL GNF GNF N-F	PP PP PP PP	GPR GPR GPR GPK	PL PL PL	PLL PLL PLL PFV	G N I G N I G N I G N I	LQ LQ LQ 4LN	MDR LDR MDR LDS	RG GG GG QH	L L K L L N L L K P H I	SFL SFM SFI FFS	R F F Q L F Q L F K L A	E K E K E K	YGD YGD YGD YGN	VF VF VF VF	ГVН ГVН ГVН SFR	LGP LGP LGP LGK	R P V R P V R P V E S M	VML VML VML
Fh CYP2P Dr CYP2A Dr CYP2A	и и и и и и и и и и и и и и и и и и и		LGL VKL LKL	GW I I D L A D L A	SVGL SVGL	TLF TLF	LGL	IF- IF-		L A D L L E L F E	V L K I F R I F R	I N I Y	S Y K S Y K	FRF GRF	P P P P P P	G P U G P I G P I	PLI	P E V P F V P F V	G N I G N I G N I	P H P H P H	IEA FLK LLK	SK S- N-] Н L - Р М - Р М	Q F K E F I G F K	S S N R S I	A GK A P Q S E	YGN YGE YGG	MT LA	SLR FIF FVF	F G R I G R	GRI KPV KPA	AII IML ISI
Hs CYP2B Rn CYP2B Mm Cyp2b	86 CGV 81 CGT 910 CGT	EAIF DTIF DTIF	100 REAL SEAL REAL	VDK/ VGQ/ VGQ/	E A F ED F E A F	SGR SGR SGR	G K I G T I G T V	A M V A V I A V V	DP EP EP	120 FFR IFK TFK	G E E	Y G Y Y G Y Y G Y	VIF VIF VIF	ANC ANC ANC	N R E R E R	WKV WKA WKT		o RFS RFS RFS	V T 1 L A 1 L A 1	Г M R Г M R Г M R	150 D F G D F G D F G	M G M G M G	KRS KRS KRS	V E E V E E V E E	160 RIC RIC RIC	Q E E Q E E O E E	AQC AQC AQC		170 E E L E E L E E L	RKS RKS RKS	K G A Q G A Q G A	180 LMD PLD PLD
Fh CYP2N Fh CYP2P Dr CYP2A Dr CYP2A	NI SGH 21 NSY NAISTI NA2 NTI	K <u>LV</u> KFMT QLAB QLAB	EAI EAL EAY EAL	V TQ V Q R V Q - V Q - V Q -	GENF GEDF DAF DVF	V D R T D R S G R S G R	P P N P S I P A I P A L	A I A P L F P L F P I F	E R E D DW DW	FYT VFG ITN ISH	E P S N G - L G - L	GG RG GI GI	LFF LVG VMV IMV	NNC SSC TFN TFN	E I Y P N S H S	WKR WKC WRC WRC	Q R I Q R I Q R I Q R I	R F A R F A R F A R F A		F L R F L R F L R F L R	T F G N F G N F G N F G	LG LG LG	KNT KKT KKT KKT	LEL LER VED VED	SIC SIC RVI RVI	E E E E	IRH CQY SRY SQY		E E I E A F A E M A E M	E N E A D Q L K E L K D	K G K Q G Q E G K E G K	PFS PFN SMN SMN
Hs CYP2B Rn CYP2B	6 PTF	LFQS	190 ITA ITA	NIIO	C S I V	200 FGK	RFH	YQD		210 FLK	MLN	LF	Y Q T Y R T	F S I	IS	S V F	230 G Q I	0 LFE VFE	LF	GF	240 LKY LKY	FP	GAH	RQV	250 Y K N	V L Q	EIN	AY	260 IGH	SVE	KHR	270 ETL ATL
Mm Cyp2b Fh CYP2N Fh CYP2P Dr CYP2A Dr CYP2A	10 PTF 1 PAG 1 AQK 11 PQH			NIIO NIIO NIIO DIIO	SIV QLV CLV SIV	F G E MGR F G N F G D	RFE	YTD YHD YSD YDN	R Q Q S K Q K R	F L R F Q T F Q T F E Y		L F YM L L T L	Y Q T S E A N E T N E N	F S L L W L L Y L I M L	I S E G A G	S F S S I W T V W S A A		MFE LYQ MYN IFN		GF QV WL - F	L K Y M K Y M R W I K H	F P I P I P F P	GAH GPH GPH GPH	RQI NKL QRI QKI	SK FS FS KQ	V L Q V F T I T N V A D		DY ELI SF GF	I G H L Q E V K V I R D	SVE EIE RIN EAK	K H R K H K E H R E H K	ATL KDL ENL QTL
DI CIFZA	w2 [] Q II	n L Q .				101	/KIL	10.		1- 1	LLIN		., E.,		(D)	511				- L	1 1 1		5111	QAI	IN QL					- <u>-</u>		
Hs CYP2B Rn CYP2B	36 D P S 31 D P S	A P K I A P R I	280 DLID DFID	T Y L I T Y L I	. H M E . R M E	290 KEK	S N A	H S E H T V	FS FH	300 HQN HEN	L N L L M I	N T S L	LSL LSL	F F A	(D) GT GT	ETT ETS	320 ST ST	o T L R T L R	YGI		330 MLK MLK	Y P Y P	H V A H V A	E R V E K V	340 Y R H Q K H	EIE EID	Q V I Q V I	G P I G S I	350 HRP HRL	P E L P T L	HD R	360 AKM SKM
Mm Cyp2b Fh CYP2N Fh CYP2P	010 D P S 11 DH S 01 D P S	V P R I N P R I S P R I	OFID OYID OYID	I Y L I T F L S F L	R M E K M E E M C	KEK NQC EKE	(SNQ) - EA - DK	HTE ELG DSG	F H FT FD	H Q N E R N L D N	L M M L A F L C F	CS CV	LSL LDL LDL	FFA FLA FVA	G T G T G T	ETT ETT ETT	S T A T T T	T L R T L L T L H	YGI WAI WGI	LLF	MLK LIK MIC	Y P Y P N P	H V A E V Q Q I Q	EKV EKV ERV	QKH HAH QAH		QVI RVI AVI	G S I G Q T G P 1	HRL FRL SRP	PTL PSM PSM	DDR ADR SDR	TKM PNL DNM
Dr CYPZA Dr CYPZA	A1 D P D A2 D P D	SPRI	OFID OFID	AYLI AYLI	EIE	KQK	FN K	DST	FH	EEN		S A S T		FLA. FLA	GT	DTI DTI	ET	F I R T I R	WGI	IN		N P N P	D V Q D V Q	ERR ERC	HEE		Q V L	G Y I G Y I	DRL DRL	PSM PSM		DKL
			370			380				390			4	100	_		410	0			420				430		_		440		(C)	450
Hs CYP2B Rn CYP2B Mm Cyp2b	B6 P Y T 81 P Y T 910 P Y T	EAVI DAVI DAVI	HE I HE I HE I	QRF QRF QRF		PMC PIC PIC	V P H V P H V P H	IVT RVT RVT	KD KD	FSF FMF FMF	RGY RGY RGY		PKD PKN PKN	TEV TEV TEV	F L Y P Y P			LHD LHD LHD	PH PQ PQ	YFE YFD YFE	K P D H P D Q P D	A F S F S F	N P D N P E N P D	H F L H F L H F L	DAN DAN DAN	NGA NGA NGA	L K K L K K L K K	TE/ SE/ SE/	AFI AFM AFL	PFS PFS PFS	LGK TGK TGK	RIC RIC RIC
Fh CYP2N Fh CYP2P Dr CYP2A Dr CYP2A	1 PYT P1 PYT W1 PYT W2 PYT	DAVI DAVI LATV LATV	HEI HEI YEI HEI	QRM QRM QRC QRY	SNIV GNII ANIA GNIA	PLN PLN P-N P-K	GLR VAR VMH LLH	VAS MAN QTI ETI	KD KD LP RR	TTL TTV TRL TKL	GGY DQY HGY HGY	F I T I D I D I	PKG PKG PQG PQG	TAV TMN TII TTI	M P L A L T I A	MLT TLE NLA NFT	SVI SVI AII AMI	LFD LHD FSN FSD	K T I E S M K D I K E I	W E W E W K W K	T P D T P N H P D H P D	T F A F A F	N PG N P E N P E N P E	H F L H F L N F L N F L	DAN EKI DEN DEN	VGK DGT VGH VGQ	FVK FRK FSK FSK	KEA REA PE PE	AFL AFL SFI YFF	PFS PFS PFS PFS	AGK AGK LGP LGP	RVC RVC RVC RAC
Hs CYP2B	6 LGE	GIAF	460	FLFI	ידד	470 L O N	FSM		VA	480 PED	IDL	ТР	0 E C	190 GVC	KI	PPT	500 Y O	o TIR F	L PI	<u> </u>	510											
Rn CYP2B Mm Cyp2b Fh CYP2N	81 LGE 10 LGE 11 LGE	GIAF SIAF GLAF	N E L N E L M E L	F L F I F L F I F L F I	TTTI TTSI		F S V F S V F S F	SSH ASH SAP	LA	PKD PKD VE-	IDL IDL LST	T P T P E G	KES KES	GIC GIC GIT	K I K I L V	PP1 PP1 PHF	Y Q Y Q Y K	ICF ICF VSA	SAI LAI KAI	2												
Fh CYP2P Dr CYP2A Dr CYP2A	P1 LGE VA1 LGE VA2 LGE	Q L A F T L A F T L A F	MEL TEL TEL	FLF FLF FLF	TSL TAL TSL	LQR LQR LQR	FKF IRF IRF	S P P SW P SW P	PGI PD PN	EQ- AKP AKP	PSL IDM IDM	E Y D G D G	- K L - I M - I V	G V 1 G L V G I V	H C R S R S	PKF PQT PEF	YR I FN FN	L C A V V C I I C	V S H H S H H S H	2 2 D N 2 D T	V E K D											

Fig. S1. Alignment of the zebrafish CYP2AA1, CYP2AA2, and other representative CYP2 deduced amino acid sequences using ClustalW. The (A) proline-rich region, (B) center of α-helix I which contains threonine residue, and (C) heme-binding motif are boxed. Accession numbers for sequences used are as follows: human CYP2B6 (ENSP00000324648), mouse Cyp2b10 (ENSMUSP00000072264), rat CYP2B1 (ENSRNOP00000045196), killifish CYP2P1 (AF117341), and killifish CYP2N1 (AF090434). Prefixes used: Dr; *Danio rerio*, Hs; *Homo sapiens*, Mm; *Mus musculus*, Rn; *Rattus norvegicus*, and Fh; *Fundulus heteroclitus*.



Fig. S2. The four most parsimonious reconstructions of the tandem duplication pattern for the CYP2AA cluster. Duplication patterns were computed using DILTAG (see Methods).



Fig. S3. Spearman's rank correlations of gene expression between *PXR* and *CYP2AA1* (A, B) and between *PXR* and *CYP2AA2* (C-E) in liver of zebrafish treated with DMSO, TCPOBOP or PB.