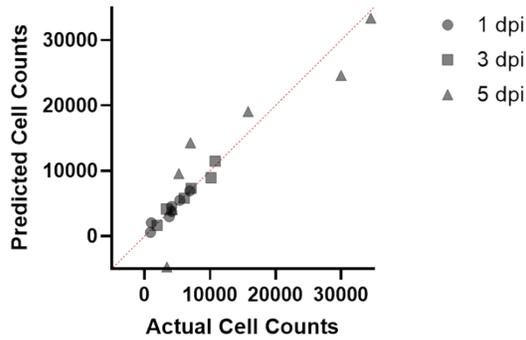


Supplementary Data

Table S1. Primers for qPCR

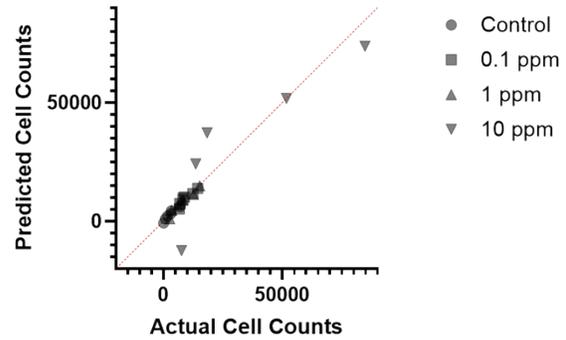
Target gene	Purpose	Direction	Sequence
<i>hprt1</i>	Housekeeping for the standard curve	Forward	5'-ATC AGA CTG AAG AGC TAT TGT AAT GAC CA-3'
		Reverse	5'-TGG CTT ATA TCC AAC ACT TCG TG-3'
<i>tbp</i>	Housekeeping for the standard curve	Forward	5'-GTG GTG TTG TGA GAA GAT G-3'
		Reverse	5'-GGC AGC GCT GCC CAG ATA GCA G-3'
<i>cdh1</i>	Cell metastatic marker	Forward	5'-GGA CAG GGA GGA TTT TGG C-3'
		Reverse	5'-GTG AAG GGA GAT GTA TTG GG-3'
<i>mmp9</i>	Cell metastatic marker	Forward	5'-AGA CGG GTA TCC CTT CGA CG-3'
		Reverse	5'-AAA CCG AGT TGG AAC CAC GAC-3'

a HepG2 cells *in vivo* w/o treatment



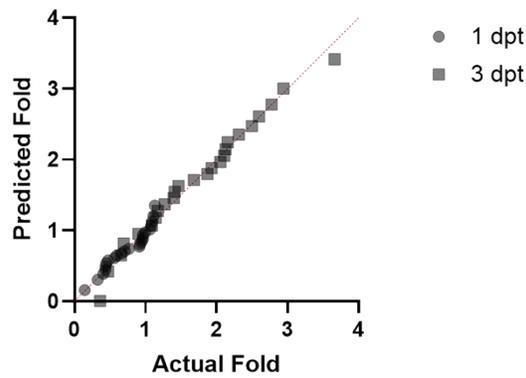
Shapiro-Wilk test (w/o PFOA)	1 dpi	3 dpi	5 dpi
W	0.9194	0.9403	0.8127
P value	0.4649	0.6619	0.0546
Passed normality test (alpha=0.05)?	Yes	Yes	Yes
P value summary	ns	ns	ns

b HepG2 cells *in vivo* w/ PFOA treatment



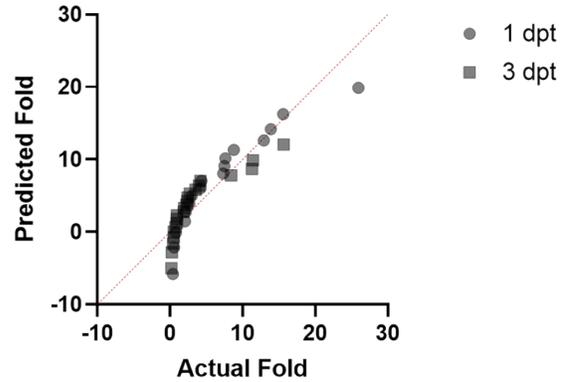
Shapiro-Wilk test (w/ PFOA)	Control	0.1 ppm	1 ppm	10 ppm
W	0.9097	0.8747	0.9260	0.7985
P value	0.4347	0.2457	0.5493	0.0569
Passed normality test (alpha=0.05)?	Yes	Yes	Yes	Yes
P value summary	ns	ns	ns	ns

c *cdh1*_PFOA



Shapiro-Wilk test (<i>cdh1</i> PFOA)	1 dpt	3 dpt
W	0.9170	0.9786
P value	0.0503	0.8555
Passed normality test (alpha=0.05)?	Yes	Yes
P value summary	ns	ns

d *mmp9*_PFOA



Shapiro-Wilk test (<i>mmp9</i> PFOA)	1 dpt	3 dpt
W	0.8506	0.7387
P value	0.0109	<0.0001
Passed normality test (alpha=0.05)?	No	No
P value summary	*	****

Figure S1. Shapiro-Wilk test for the data generated from the PFOA treated samples. The *in vivo* cell numbers of HepG2-LuciaTM AhR cells without treatment of PFOA (a) and the ones with the treatment of PFOA (b) followed the normal distribution ($p > 0.05$ in all columns). Similarly, the *cdh1* expression levels in engrafted HepG2-LuciaTM AhR cells were also normally distributed ($p > 0.05$ at both 1 and 3 dpt). In contrast, the *mmp9* expression data from qPCR analysis failed to fit the normal distribution ($p = 0.0109$ at 1 dpt and $p < 0.0001$ at 3 dpt).

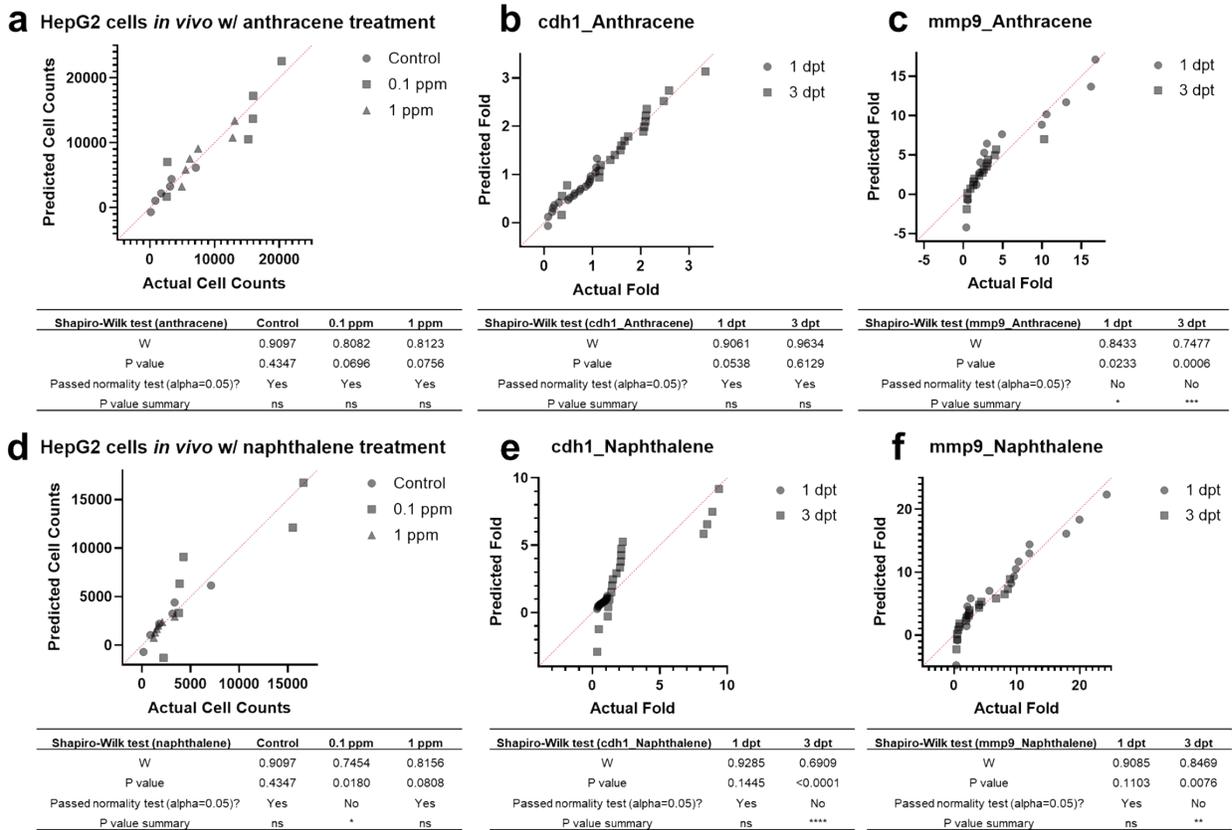
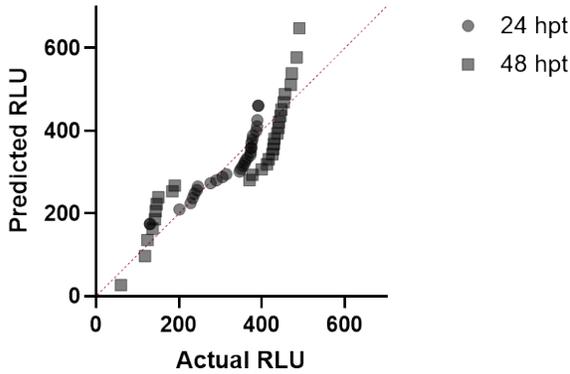
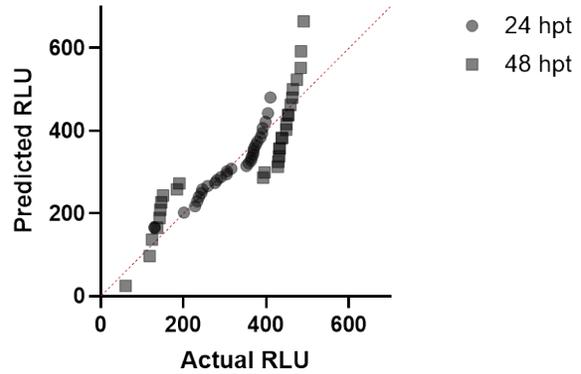


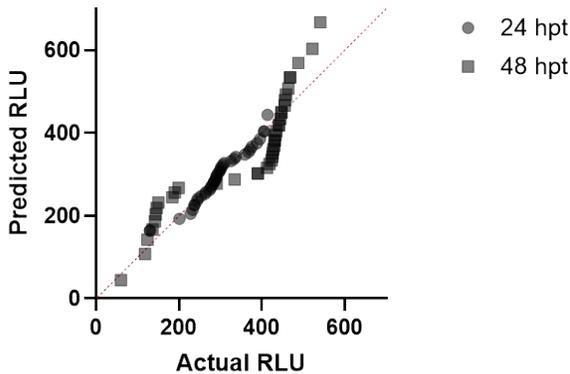
Figure S2. Shapiro-Wilk test for the data generated from the anthracene and naphthalene treated samples. The numbers of HepG2-LuciaTM AhR cells inside the zebrafish larvae with treatment anthracene (a) as well as the *cdh1* expression levels in HepG2-LuciaTM AhR cells *in vivo* (b) were found to be normally distributed ($p > 0.05$ in all groups). The *mmp9* expression values from qPCR analysis (c) followed a non-normal distribution pattern ($p = 0.0233$ at 1 dpt and $p = 0.0006$ at 3 dpt). Accordingly, the HepG2-LuciaTM AhR cell numbers in naphthalene treated zebrafish larvae (d) followed the normal distribution only in 1 ppm treatment group ($p = 0.0808$), but not 0.1 ppm treatment group ($p = 0.0180$). The expression values of *cdh1* (e) and *mmp9* (f) were both normally distributed at 1 dpt ($p > 0.05$) but non-normally distributed at 3 dpt ($p < 0.0001$ for *cdh1* and $p = 0.0076$ for *mmp9*).

a Normal QQ plot

Shapiro-Wilk test (<i>in vivo</i> Anthracene)	24 hpt	48 hpt
W	0.8298	0.7839
P value	0.0002	<0.0001
Passed normality test (alpha=0.05)?	No	No
P value summary	***	****

b Normal QQ plot

Shapiro-Wilk test (<i>in vivo</i> Naphthalene)	24 hpt	48 hpt
W	0.9047	0.7582
P value	0.0110	<0.0001
Passed normality test (alpha=0.05)?	No	No
P value summary	*	****

c Normal QQ plot

Shapiro-Wilk test (<i>in vivo</i> PFOA)	24 hpt	48 hpt
W	0.9547	0.8092
P value	0.1097	<0.0001
Passed normality test (alpha=0.05)?	Yes	No
P value summary	ns	****

Figure S3. Shapiro-Wilk test for the data generated from the *in vivo* luciferase assays. In general, all data were found to be non-normally distributed in the samples treated by anthracene (a), naphthalene (b), and PFOA (c), except one PFOA treated group at 24 hpt ($p = 0.1097$).

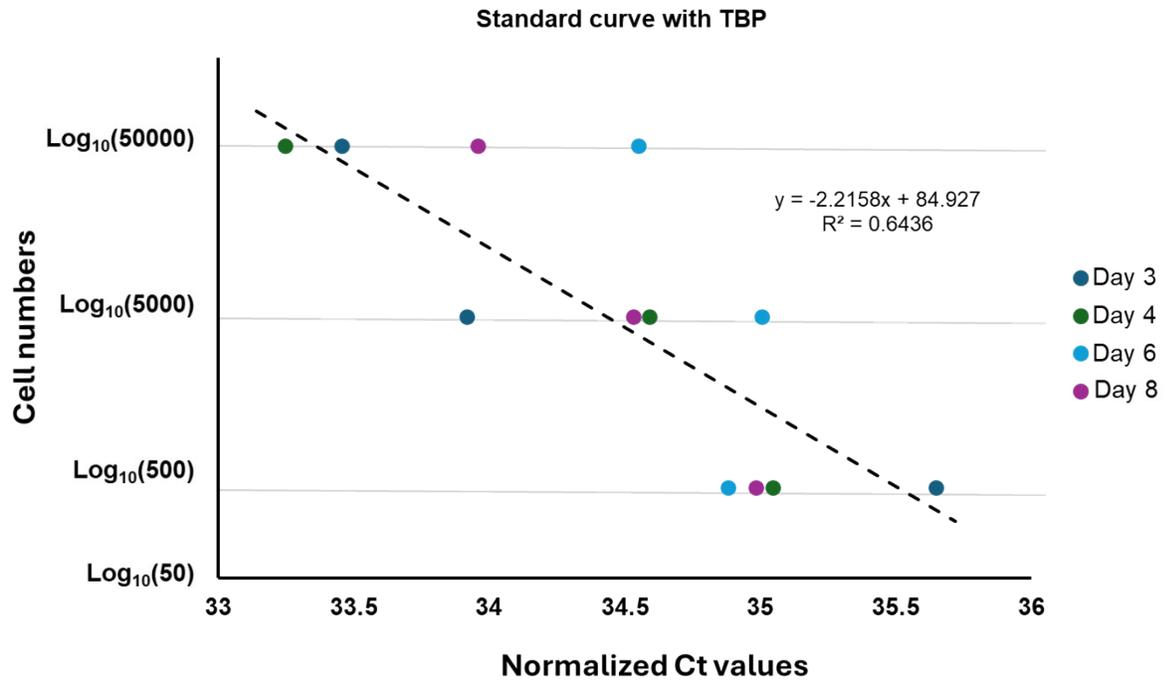


Figure S4. The standard curve for in vivo HepG2 cell quantification using the expression values of *tbp* gene. The Ct values normalized by the actual concentrations of total RNA used for cDNA synthesis of each sample were used for X-axis. The log10 values of different cell numbers used to mix with the embryos/larvae were used for Y-axis.