OXFORD
SOT | Society of Toxicology
academic.oup.com/toxsci

Toxicological Sciences, 2026, **209(1)**, kfaf146
https://doi.org/10.1093/toxsci/kfaf146
Advance Access Publication Date: October 23, 2025
Research article

Functional genomic analysis of non-canonical DNA regulatory elements of the aryl hydrocarbon receptor

Shayan Shahriar[1,†], Tajhal D. Patel[2,†], Manjula Nakka[1], Sandra L. Grimm[1,2,3], Cristian Coarfa[1,2,3,*], Daniel A. Gorelick[1,3,*]

[1]Center for Precision Environmental Health, Baylor College of Medicine, Houston, TX 77030, United States
[2]Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, TX 77030, United States
[3]Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, United States

[†]These authors contributed equally to this manuscript.

*Corresponding authors: Cristian Coarfa, 1 Baylor Plaza, Houston, TX 77030, United States. Email: coarfa@bcm.edu; Daniel A. Gorelick, 1 Baylor Plaza, Houston, TX 77030, United States. Email: gorelick@bcm.edu

Abstract

The aryl hydrocarbon receptor (AHR) is a ligand-dependent transcription factor activated by environmental toxicants like halogenated and polycyclic aromatic hydrocarbons, which then binds to DNA and regulates gene expression. AHR is implicated in numerous physiological processes, including liver and immune function, cell cycle control, oncogenesis, and metabolism. Traditionally, AHR binds a consensus DNA sequence (GCGTG), the xenobiotic response element (XRE), recruits coregulators, and modulates gene expression. Yet, recent evidence suggests AHR can also regulate gene expression via a non-consensus sequence (GGGA), termed the non-consensus XRE (NC-XRE). The prevalence and functional significance of NC-XRE motifs in the genome have remained unclear. Although chromatin immunoprecipitation (ChIP) and reporter studies hinted at AHR-NC-XRE interactions, direct evidence for transcriptional regulation in a native context was lacking. In this study, we analyzed AHR binding to NC-XRE sequences genome-wide in the mouse liver, integrating ChIP-seq and RNA-seq data to identify candidate AHR target genes containing NC-XRE motifs in their regulatory regions. We found NC-XRE motifs in 82% of AHR-bound DNA, significantly enriched compared with random regions, and present in promoters and enhancers of AHR targets. Functional genomics on the Serpine1 gene revealed that deleting NC-XRE motifs reduced TCDD-induced Serpine1 upregulation, demonstrating direct regulation. These findings provide the first direct evidence for AHR-mediated regulation via NC-XRE in a natural genomic context, advancing our understanding of AHR-bound DNA and its impact on gene expression and physiological relevance.

Keywords: toxicogenomics; aryl hydrocarbon receptors; polycyclic aromatic hydrocarbons; ChIP-seq; RNA-seq

The toxic effects of secondhand cigarette smoke, fossil fuels, and byproducts of industrial combustion are caused by halogenated and polycyclic aromatic hydrocarbons such as TCDD. Understanding the signaling pathways by which such compounds cause toxicity is crucial for predicting tolerable exposure levels and for reversing the effects of adverse exposure. Aromatic hydrocarbons activate aryl hydrocarbon receptors (AHRs), a ligand-dependent transcription factor. Upon activation, AHR translocates to the nucleus, where it recruits protein coregulators and directly regulates gene expression (Denison et al. 2011). AHR forms a dimer with the aryl hydrocarbon receptor nuclear translocator protein (ARNT, also known as HIF1β) to regulate transcription (Reyes et al. 1992; Hankinson 2005). The AHR–ARNT complex binds to a consensus DNA sequence (GCGTG) termed the xenobiotic response element (XRE) (Denison et al. 1988; Yao and Denison 1992).

Recent evidence complicates these findings. AHR may regulate gene expression by binding to non-consensus XRE (NC-XRE) sequences. AHR was shown to physically interact with RelB, a member of the NF-κB family, and bind to a distinct RelB/AHR-responsive element (RelBAhRE) within the IL-8 promoter. Furthermore, TCDD induced a time-dependent recruitment of AHR to the RelBAhRE site (nucleotides GGGTGCAT) (Vogel et al. 2007). A chromatin immunoprecipitation (ChIP) study found that approximately half of TCDD-AHR complexes in mouse liver were bound to DNA lacking a consensus XRE (Dere et al. 2011). Exploring the genome more closely, Huang and Elferink (2012) identified a non-consensus DNA sequence to which AHR binds. ChIP, gel shift (electrophoretic mobility shift assay), and reporter assays support the hypothesis that AHR can bind to a non-consensus XRE DNA sequence characterized by the nucleotides GGGA (referred to as the NC-XRE motif) (Huang and Elferink 2012; Wilson et al. 2013; Jackson et al. 2014; Joshi et al. 2015), though this has not been functionally tested in a natural genomic context. Here, we sought to analyze AHR binding to non-canonical DNA elements on a genome-wide scale and, in the case of a single target gene, directly test whether NC-XRE motifs are required for AHR-dependent target gene expression.

Results

Frequency and distribution of non-canonical AHR-binding motifs

To determine the genome-wide distribution of AHR binding to DNA following TCDD exposure, we analyzed published AHR

ChIP-seq data from livers of C57BL/6 male and female mice (28 days old) following 2-h exposure to 30 μg/kg TCDD (GEO accession numbers GSE97634, GSE97636) (Fader et al. 2017, 2019). These data were previously assayed for AHR binding to consensus XRE DNA (GCGTG) but was not interrogated for AHR binding to non-consensus DNA such as NC-XRE (GGGA) or RelBAhRE (GGGTGCAT). Given that the XRE, NC-XRE, and RelBAhRE motifs differ in length, which influences their frequency of occurrence across the genome (shorter motifs being more common), we assessed the likelihood of finding XRE, NC-XRE, or RelBAhRE motifs in AHR ChIP peaks versus in random genomic intervals matched in size and chromosome distribution to observed peaks. Both XRE and NC-XRE motifs were significantly enriched in AHR-bound regions, whereas RelBAhRE was not (Fig. 1A). These random regions were scanned with the same motif definitions as the observed peaks, ensuring that motif length-dependent differences in background occurrence were directly incorporated into the enrichment test.

We also found consistent patterns of AHR-bound motif enrichment in the livers of male and female mice following TCDD exposure, based on AHR ChIP-seq data. In males, 29.74% of AHR peaks contained only NC-XRE motifs, compared with 25.79% in females. Peaks containing only XRE motifs were slightly more frequent in females (11.75%) than in males (8.99%). The majority of peaks—51.72% in males and 53.96% in females—contained both XRE and NC-XRE motifs, indicating frequent co-occurrence. Despite these minor differences, the overall distribution of motif combinations was remarkably similar between sexes. RelBAhRE motifs were rare in both groups, appearing in fewer than 1% of peaks (Fig. 1B and Fig. S1A). These findings indicate that AHR binding to canonical and non-canonical DNA motifs occurs at comparable frequencies in male and female mouse livers, supporting a robust and sex-independent pattern of AHR–DNA interactions in response to TCDD. They also highlight the substantial role of NC-XRE motifs in mediating AHR binding. Because of the low number of RelBAhRE in AHR peaks, we only focused on the XRE and NC-XRE motifs in the subsequent studies.

We next examined how AHR-binding sites containing XRE and NC-XRE motifs are distributed across the genome. To do this, we analyzed AHR ChIP-seq peaks in adult mouse liver (GEO accession numbers GSE97634, GSE97636) and overlaid them with genomic regions annotated using a 10-state ChromHMM model (van der Velde et al. 2021). We generated this model using 8 histone modifications profiled via publicly available ChIP-seq datasets (GEO accession numbers: GSM1000140, GSM918718, GSM1000150, GSM1000151, GSM1000152, GSM1000153, GSM769014, GSM769015, GSM751035) (Mouse ENCODE Consortium et al. 2012; Tennant et al. 2013). These included histone marks associated with active promoters (H3K4me3, H3K9ac), enhancers (H3K4me1, H3K27ac), gene bodies and active transcription (H3K36me3), polycomb-mediated repression (H3K27me3), and heterochromatin (H3K9me3) (Fig. S2E).

We found that AHR binding was enriched in active promoters (e.g. active TSS) and enhancers (Fig. 1C and D). Specifically, AHR peaks containing both XRE and NC-XRE motifs were predominantly located in active transcription start sites (TSSs) and active enhancers. On the other hand, most NC-XRE-only peaks were found in the enhancer regions, including both active and poised enhancers. Active enhancers have high chromatin accessibility and confer strong enhancer activity, whereas poised enhancers are marked by intermediate levels of enhancer-associated histone modifications and are thought to be primed for activation. We also noticed that AHR binding was rare in repressive chromatin states, such as those enriched for H3K27me3 (polycomb repression) or H3K9me3 (heterochromatin), as well as in the quiescent regions lacking significant histone modification signals. These patterns were consistent across male and female samples, supporting a sex-independent mechanism of AHR–DNA interaction (Fig. 1C and D). Notably, promoter- and enhancer-associated regions comprise only ∼5.3% of the mouse genome, whereas repressive and quiescent regions account for approximately 84% (van der Velde et al. 2021). This highlights a strong preference of AHR for binding within the active regulatory elements, consistent with its function as a transcription factor. Specifically, NC-XRE motifs may play a particularly important role in enhancer-associated AHR binding.

Repeated NC-XRE motifs associated with AHR ChIP peaks

We searched for NC-XRE motifs present near AHR-binding sites (ChIP peak) within 10 kb of a gene body. The 10 kb parameter allows us to include promoter and gene proximal regulatory DNA peaks, important because our analysis of ChIP-seq data suggested that AHR binds NC-XRE at enhancers (Fig. 1C and D). Among the differentially expressed genes in TCDD vs vehicle, 520 had AHR-binding peaks at XRE, NC-XRE, or XRE + NC-XRE motifs within 10 kb of a gene body. Twenty-eight of these genes had AHR-binding peaks at NC-XRE but not at XRE or XRE + NC-XRE motifs (Table S1).

The preceding analysis is for any single instance of XRE or NC-XRE DNA sequence. However, for some transcription factors, repeated DNA-binding sites produce a more robust response (Klein-Hitpass et al. 1988). To explore the functional importance of repeated runs of XRE or NC-XRE motifs, we sought to examine the association between repeated runs of XRE or NC-XRE motifs and AHR target gene expression.

We defined tandem XRE and NC-XRE motifs based on 2 parameters: The number of repeats and the maximum allowed spacer length (in base pairs) between adjacent motifs, as illustrated in Fig. 2A. The number of repeats refers to the minimum number of XRE or NC-XRE motifs within a given region, whereas the spacer defines the maximum distance permitted between any 2 motifs in a repeat run. Using this framework, we analyzed AHR target genes for varying numbers of XRE or NC-XRE repeats and spacer lengths. Notably, the previously characterized AHR target gene *Serpine1* (encodes the protein PAI1) contains 5 NC-XRE motifs within its putative promoter region. This observation aligns with previous findings that AHR upregulates *Serpine1* via NC-XRE-mediated binding (Huang and Elferink 2012), prompting us to further investigate the regulatory significance of five-repeat NC-XRE clusters.

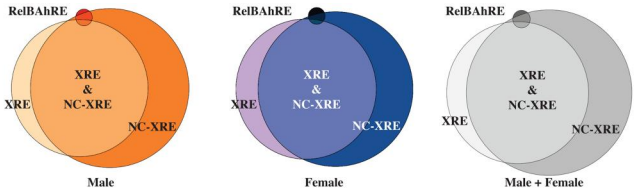
Figure 2B shows the number of AHR target genes with AHR peaks containing at least 5 NC-XRE motifs separated by up to 100 basepairs (bp), compared with target genes found near AHR peaks with at least 5 XRE motifs. Notably, we observed a 9-fold increase in the number of AHR target genes linked to peaks with 5 NC-XRE motifs separated by 50 bp or less, compared with those with 5 XRE motifs under the same spacing constraint. This subset includes *Serpine1*, a well-established AHR target regulated via NC-XRE motifs, as previously discussed (Fig. 2B).

Further analysis revealed that although 82% of all AHR peaks contain at least 1 NC-XRE motif, only 5.2% harbor a cluster of at least 5 NC-XRE motifs spaced by 50 bp or less (Fig. 2C, Fig. S2A). We next characterized the genomic distribution of these 5× NC-XRE clusters. Of the 426 identified clusters, over 71% were located at or near TSSs, indicating a strong promoter association.

A

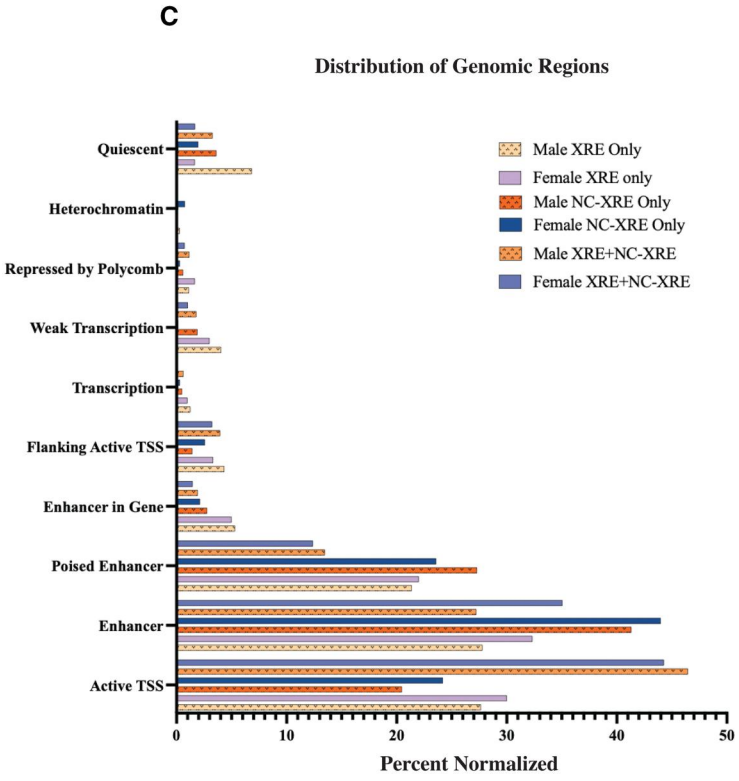
Feature	Sequence	Number in AHR Peaks	Number of Random Peaks	Odds Ratio	P value
XRE	GCGTG	4909	666	16.988	$<10^{-40}$
NC-XRE	GGGA	6664	6021	1.583073	9.53×10^{-34}
RelBAhRE	GGGTGCAT	39	32	1.219784	0.4757227

B



Feature	Number of AHR Peaks					
	Male	Percent Male	Female	Percent Female	Male + Female	Percent Male + Female
XRE Only	714	8.99	297	11.75	719	8.81
NC-XRE Only	2362	29.74	652	25.79	2462	30.18
RelBAHRE Only	14	0.18	9	0.36	16	0.20
XRE + NC-XRE	4108	51.72	1364	53.96	4188	51.34
XRE + RelBAHRE	25	0.31	8	0.32	25	0.31
NC-XRE + RelBAHRE	35	0.44	10	0.40	37	0.45
XRE + NC-XRE + RelBAHRE	23	0.29	5	0.20	23	0.28
Peaks without XRE, NC-XRE or RelBAHRE	746	9.39	207	8.19	773	9.47
Total AHR Peaks	7943	100	2528	100	8158	100

C



D

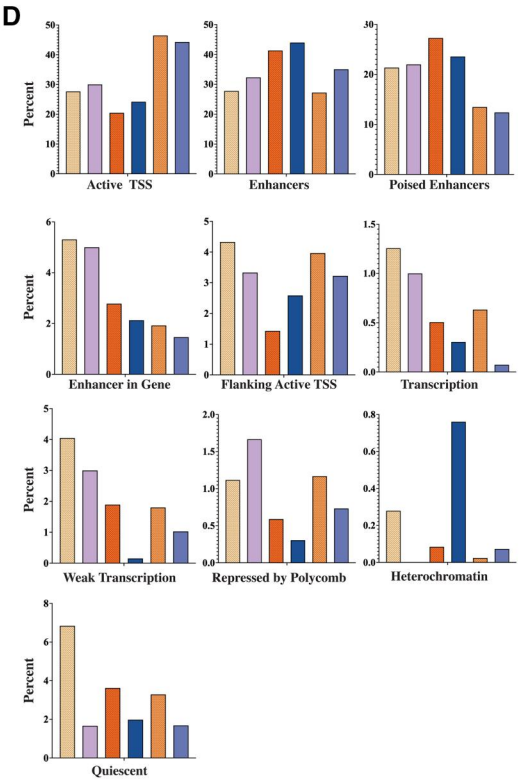


Fig. 1. Enrichment and distribution of NC-XRE, XRE, and RelBAhRE motifs in AHR ChIP-seq Peaks. Following TCDD exposure, we analyzed mouse liver AHR ChIP-seq peaks from 5 male and 5 female mice for the presence of known AHR motifs: Canonical xenobiotic response element (XRE), non-canonical xenobiotic response element (NC-XRE), and RelB AHR response element (RelBAHRE). A) Likelihood of finding XRE, NC-XRE, or RelBAHRE motifs in AHR ChIP peaks versus in random genomic regions. The table shows the number of AHR ChIP peaks containing at least 1 XRE, NC-XRE, or RelBAHRE motif. Odds ratios and *P*-values derived from Fisher's exact test indicate significant enrichment of XRE and NC-XRE motifs in AHR peaks. B) Proportion of AHR ChIP peaks containing XRE DNA sequence, NC-XRE, XRE + NC-XRE, or none in male and female mice. NC-XRE motifs are more prevalent than XRE motifs in both sexes. C) Genomic distribution of AHR ChIP-seq peaks mapped to 10 chromatin states defined by ChromHMM using 8 histone modification profiles. The highest percentages of AHR peaks containing XRE and NC-XRE motifs are found in promoter and enhancer regions. D) Enlarged view of panel C for detailed comparison of chromatin state distributions. This panel highlights the differences in the distribution of AHR peaks in various chromatin states. Overlapping XRE and NC-XRE motifs are slightly more prominent in promoters; NC-XREs without XREs are slightly more enriched in enhancers and poised enhancers; XREs without NC-XREs are more enriched at enhancers in genes and at gene bodies with active or weak transcription.

An additional 27.8% were found in enhancer regions, with 21% in active enhancers and 6.8% in poised enhancers. These clusters were rarely found in repressed chromatin states (e.g. heterochromatin, polycomb-repressed, or quiescent regions) or within gene bodies associated with transcriptional activity (transcription or weak transcription) (Fig. 2D, Fig. S2C and D). Similarly, clusters of 5 XRE motifs were predominantly located in promoters and enhancers, with a small fraction (6.5%) found within gene bodies. We compared the binding strengths of different histone marks at AHR peaks containing 5× XRE or NC-XRE (50 bp spacer). NC-XREs

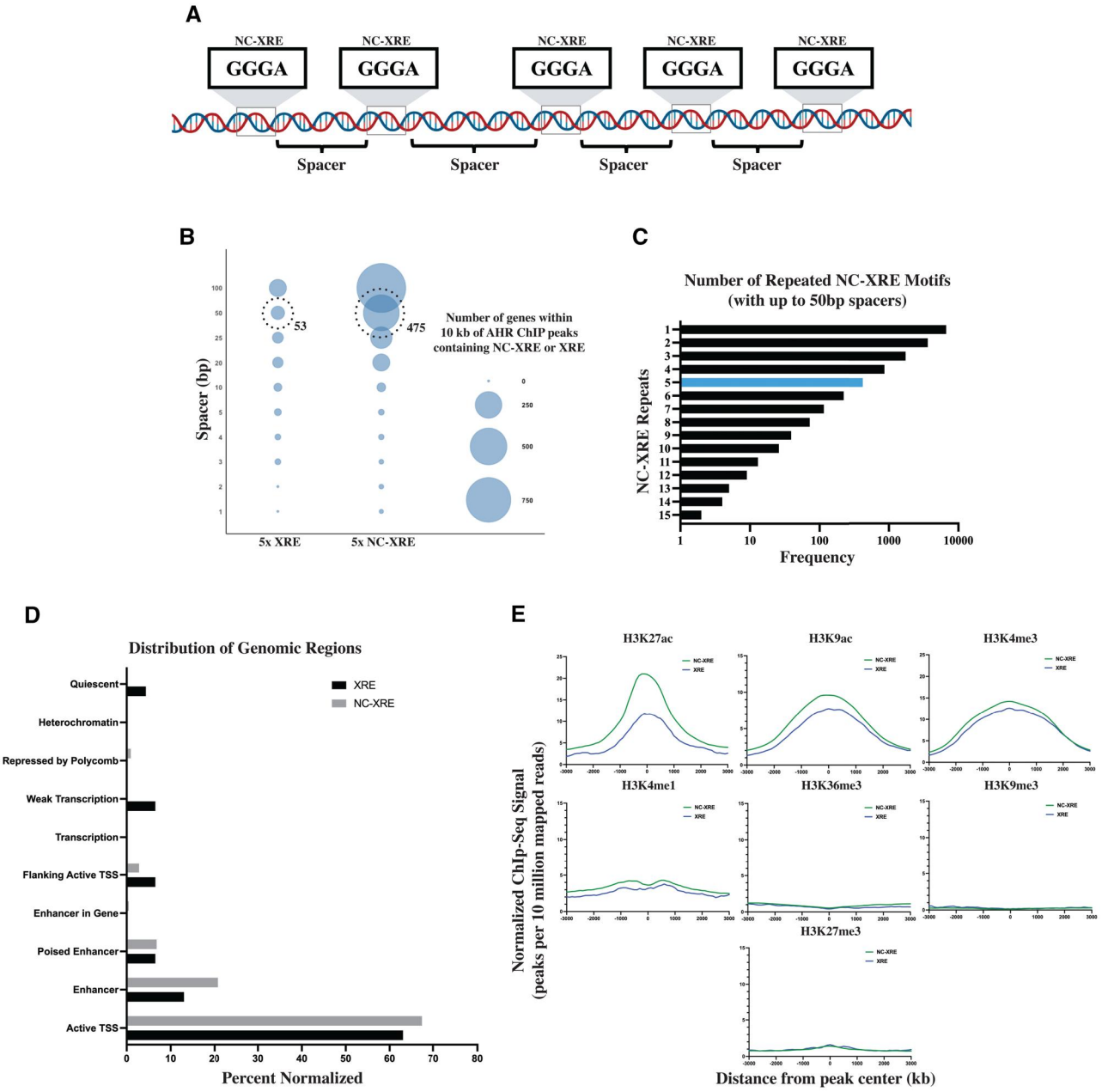


Fig. 2. Analysis of NC-XRE repeats and their genomic context. A) We categorized NC-XRE DNA motifs based on the number of repeated NC-XRE motifs and the distance separating those motifs (spacers) from each other. For example, 5 motifs separated by 50 basepairs (bp) would indicate 5 NC-XRE motifs where each spacer is up to 50 bp long. B) Number of genes within 10 kb of AHR ChIP-seq peaks containing either 5 repeats of NC-XRE or XRE motifs separated by spacers ranging from 1 to 100 bp. The size of the circles represents the number of AHR target genes. Highlighted circles denote that there are 9-fold more AHR target genes associated with AHR ChIP peaks containing 5 NC-XRE motifs versus 5 XRE motifs separated by 50 bp or less. C) Frequency of repeated runs of NC-XRE motifs in AHR ChIP peaks. It shows the total number of AHR ChIP peaks containing different numbers of NC-XRE repeats (1 to 15) with a maximum of 50 bp spacer between any 2 motifs. D) Genomic distribution of AHR ChIP-seq peaks containing 5 XRE or NC-XRE motifs separated by 50 bp or less. These are mapped to 10 chromatin states defined by ChromHMM using 8 histone modification states. The highest percentages of AHR peaks containing XRE and NC-XRE motifs are found in promoter and enhancer regions. NC-XRE motifs are prevalent in enhancers and promoters compared with XRE motifs but are absent in gene bodies (transcription, weak transcription, enhancer in gene). E) Binding strength of indicated histone modifiers at AHR ChIP peaks containing 5 or more NC-XRE or XRE motifs separated by 50 bp or less. AHR-NC-XRE peaks are more associated with activating histone marks (H3K27ac, H3K9ac, H3K4me3, H3K4me1).

are enriched at AHR-binding sites proximal to TSSs (H3K4me3, H3K27ac) and enhancers (H3K27ac) (Fig. 2E, Fig. S2E). Consistent with the broader genomic distribution of NC-XREs, 5× NC-XRE clusters were enriched in enhancer regions compared with 5× XRE clusters. This also suggests that NC-XREs may be more potent when they occur in tandem repeats.

Putative AHR-NC-XRE target genes and pathways

The preceding analysis examined the presence of NC-XRE motifs in all genomic regions where AHR was bound to DNA. However, transcription factors like AHR may bind DNA without necessarily affecting gene expression. Therefore, to characterize functional AHR-binding events, we sought to identify AHR ChIP-seq peaks

Table 1. ChIP-Seq datasets analyzed in this study.

GEO series	PMID	GEO samples	Target	Sex	Strain	Tissue	Age at assay	Conditions
GSE97634	28213091 31015483	GSM2573805 GSM2573806 GSM2573807 GSM2573808 GSM2573809	AHR	Male	C57BL/6	Liver	Adult, 28 days	Exposed to 30 μg/kg TCDD via oral gavage for 2 h, $n = 5$ mice
GSE97636	26582802	GSM2574037 GSM2574038 GSM2574039 GSM2574040 GSM2574041	AHR	Female	C57BL/6	Liver	Adult, 28 days	Exposed to 30 μg/kg TCDD via oral gavage for 2 h, $n = 5$ mice
GSE31039	22889292	GSM1000140	H3K27ac	Male	C57BL/6	Liver	Adult, 8 weeks	No treatment
		GSM1000150	H3K27me3	Male	C57BL/6	Liver	Adult, 8 weeks	
		GSM1000151	H3K36me3	Male	C57BL/6	Liver	Adult, 8 weeks	
		GSM1000152	H3K79me2	Male	C57BL/6	Liver	Adult, 8 weeks	
		GSM1000153	H3K9ac	Male	C57BL/6	Liver	Adult, 8 weeks	
		GSM769014	H3K4me3	Male	C57BL/6	Liver	Adult, 8 weeks	
		GSM769015	H3K4me1	Male	C57BL/6	Liver	Adult, 8 weeks	
GSE36027	22889292	GSM918718	Control	Male	C57BL/6	Liver	Adult, 8 weeks	No treatment
GSE30298	23238790	GSM751035	H3K9me3	N/A	C57BL/6	Liver	Adult, 8 to 10 weeks	No treatment

Table 2. RNA-Seq datasets analyzed in this study.

GEO series	PMID	GEO samples	Treatment duration (h)	Sex	Strain	Tissue	Age at assay	Conditions
GSE109863	29752288	GSM2971788	2	Male	C57BL/6	Liver	Adult 28 days	On postnatal day 28, mice were orally gavaged with sesame oil vehicle or 30 μg/kg TCDD
		GSM2971789						
		GSM2971790						
		GSM2971791						
		GSM2971792						
		GSM2971793						
		GSM2971794	4					
		GSM2971795						
		GSM2971796						
		GSM2971797						
		GSM2971798						
		GSM2971799						
		GSM2971800	8					
		GSM2971801						
		GSM2971802						
		GSM2971803						
		GSM2971804						
		GSM2971805						
		GSM2971806	12					
		GSM2971807						
		GSM2971808						
		GSM2971809						
		GSM2971810						
		GSM2971811						
		GSM2971812	24					
		GSM2971813						
		GSM2971814						
		GSM2971815						
		GSM2971816						
		GSM2971817						

associated with AHR target genes. For this, we integrated the AHR ChIP-seq data with published RNA-seq data from mouse liver (Nault et al. 2015, 2018; Fader et al. 2017). To maximize identification of AHR target genes, we included RNA-seq data from mouse livers exposed to 30 μg/kg of TCDD, a potent AHR ligand, across multiple durations (2, 4, 8, 12, 24 hr) of exposure (Tables 1 and 2). We defined AHR target genes by differential expression in TCDD-treated mouse liver at 2 h vs DMSO, fold change ≥1.5, FDR < 5%. We focused our analysis on the 2-h timepoint because this is when direct AHR target gene regulation is most prominent before the onset of the secondary effects. As illustrated in Fig. S5C, gene set enrichment analysis (GSEA; discussed in detail later) of AHR-NC-XRE-regulated pathways shows a clear enrichment at 2 h of TCDD exposure, followed by a decline at later time points. This pattern is consistent with the biology of AHR signaling: TCDD rapidly induces AHR activity and upregulates its direct targets within 2 h, but by 4 h and beyond, negative feedback suppresses direct AHR activity. This leads to reduced enrichment of

direct AHR target pathways and an increasing contribution of secondary or indirect effects, including pathways regulated by downstream AHR targets.

We examined whether AHR-dependent differential gene expression correlated with the number of NC-XRE motifs by plotting the fold changes of all upregulated AHR target genes after 2 and 4 h of TCDD exposure against the number of NC-XREs within 10 kb of each gene body, followed by linear regression analysis (Fig. 5A and B). No significant correlation was observed. Given that the known AHR-NC-XRE target *Serpine1* harbors 5 repeated NC-XRE motifs in its putative promoter, we next focused on AHR ChIP-seq peaks containing at least 5 NC-XRE motifs spaced ≤50 base pairs apart to identify additional putative AHR targets. Using this approach, we identified such NC-XRE clusters near 28 AHR target genes, among which 22 genes were upregulated in response to TCDD exposure. These differentially upregulated genes with 5 or more NC-XRE motifs are shown in Fig. 3A. As validation, *Serpine1* was among these genes, consistent with prior findings that AHR upregulates *Serpine1* via NC-XRE DNA (Huang and Elferink 2012). AHR binding was also detected at NC-XRE motifs near genes like *Tiparp*, another known AHR target gene induced by TCDD (Ma et al. 2001) and reported to suppress hepatic gluconeogenesis (Diani-Moore et al. 2010). These findings suggest that NC-XRE motifs may contribute to the regulation of *Tiparp* expression, pointing to a potential novel mechanism underlying TCDD-induced hepatotoxicity.

To further explore the biological relevance of the 22 upregulated AHR target genes containing 5× NC-XRE motifs (50 bp spacers), we performed an over-representation analysis (ORA) to identify significantly related Gene Ontology (GO) biological processes (FDR < 5%). This analysis revealed 14 pathways (Fig. 3B). However, many of these pathways are broad and involve large gene sets—e.g. the GO term "embryo development" includes over 1,000 genes. To contextualize these findings, we conducted a GSEA using a global ranked list of genes based on differential expression in mouse liver following 2-h TCDD treatment compared with DMSO. This approach allowed us to assess whether the identified pathways were enriched across the entire transcriptome. Thirteen out of the 14 pathways identified in the ORA were also significantly enriched in the GSEA (FDR < 25%) (Fig. 3C).

The enriched pathways span a range of biological functions, including development, morphogenesis, oxygen response, and enzymatic activity—processes that align with known physiological roles of AHR. These findings reveal a compelling link between AHR-NC-XRE target gene expression and key biological pathways, suggesting that AHR may regulate these biological processes, at least in part, through NC-XRE-mediated transcriptional control.

Putative transcription factors regulated via NC-XRE motifs

Transcription factors bind regulatory DNA in a complex composed of multiple different transcription factors. For example, AHR is associated with the transcription factor estrogen receptor alpha at regulatory DNA (Klinge et al. 2000; Beischlag and Perdew 2005; Madak-Erdogan and Katzenellenbogen 2012). Another transcription factor, ARNT, is a co-regulator of AHR (Reyes et al. 1992; Lo and Matthews 2012).

Previous studies have focused on transcription factor binding at XRE motifs. To identify transcription factors that might form a complex with AHR at NC-XRE motifs, we identified known transcription factor DNA-binding sites flanking NC-XRE motifs. We focused on AHR ChIP peaks containing at least 2 NC-XRE motifs within 25 bp of each other. This threshold was chosen to maximize the detection of potential transcription factors adjacent to NC-XRE-bound AHR peaks. It includes all high-density 5× NC-XRE (50 bp spacer) clusters and captures additional AHR-NC-XRE-adjacent regions, making it more permissive for potential transcription factor screening. The top 12 most frequently occurring known transcription factor DNA-binding sites are shown in Fig. 4A.

As a complementary approach to identify transcription factors that might form a complex with AHR at NC-XRE motifs, we took the 22 upregulated genes mentioned in Fig. 3A and asked whether any combination of these genes is known to be targets of additional transcription factors, besides AHR. We queried the CollecTRI database—a curated resource of transcriptional regulators and their experimentally validated targets, commonly used to infer regulon activity (Müller-Dott et al. 2023). A regulon is defined as a group of genes regulated by a transcription factor through direct binding to specific regulatory DNA motifs. This analysis revealed 20 significantly overrepresented regulons (FDR < 5%), indicating that multiple transcription factors may form a transcription complex with AHR at NC-XRE DNA (Fig. 4B).

To further investigate potential transcription factors that might form a complex with AHR at NC-XRE motifs, we searched for transcription factors that physically bind to the same genomic regions as AHR. We focused on AHR ChIP-seq peaks containing 5 NC-XRE motifs spaced by 50 base pairs or less. Using the CistromeDB Toolkit, we calculated GIGGLE scores to identify ChIP-seq datasets from the Cistrome Data Browser that significantly overlap with AHR-NC-XRE peak regions (Fig. 4C and D, Fig. S4C) (Mei et al. 2017; Zheng et al. 2019). The GIGGLE score integrates statistical significance (via Fisher's exact test) and enrichment (odds ratio) to determine how likely the observed overlap is beyond chance (Layer et al. 2018). A higher GIGGLE score between 2 different ChIP-seq datasets (targeting different proteins for the pull-down) indicates a strong and statistically meaningful overlap in their peak regions, suggesting that the 2 different proteins of interest tend to bind in similar genomic locations.

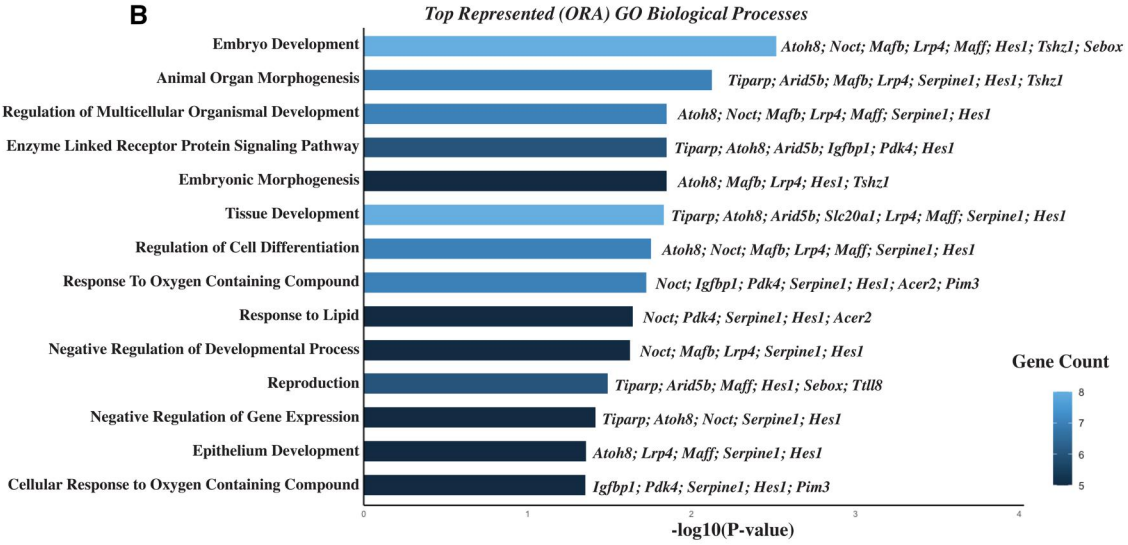
As expected, GIGGLE scores confirmed strong overlap between our AHR-NC-XRE ChIP peak set and existing AHR ChIP-seq datasets, including those originally used to define the 5× NC-XRE peaks. This served as an internal control, validating the sensitivity of our overlap analysis (Fig. 4C; Fig. S4B).

Beyond AHR itself, we identified 23 transcription factors with ChIP-seq peaks from mouse liver or from cultured mouse hepatocytes that significantly overlapped with the 5× NC-XRE (50 bp spacer) AHR peaks (Fig. 4D). This suggests that these transcription factors may co-occupy NC-XRE-rich DNA regions alongside AHR. These included factors such as BMAL1 and CLOCK, which are central regulators of circadian rhythm, as well as SMAD3 and TGIF1, which mediate TGF-β signaling. This co-localization suggests that AHR may coordinate with multiple transcription factors at NC-XRE-rich regions, potentially forming regulatory complexes that extend beyond canonical AHR–ARNT interactions. Interestingly, 5× NC-XRE motifs exhibited stronger giggle overlap with AHR ChIP-seq datasets than 5× XRE motifs. We interpret this not as evidence that AHR globally prefers NC-XREs, but rather as a consequence of the relatively high frequency of clustered NC-XREs at known AHR-binding sites in mouse liver. These findings reinforce the idea that NC-XRE clusters are enriched in regulatory DNA where AHR and its potential partners co-occupy, highlighting their functional importance.

A

Gene Name	Fold Change	NC-XRE Location	XRE within 300 bp on an NC-XRE
Serpine1	92.62	promoter	No
Tiparp	22.70	promoter and after exon 1	Yes
Maff	4.92	promoter and between alternatively spliced exon 1	Yes
Ttll8	2.53	near exon 9	No
Sebox	2.39	promoter	No
Igfbp1	2.31	promoter	Yes
Pim3	2.17	promoter, exon1, intron between exon 4 & 5	Yes
Mafb	2.05	promoter	Yes
Lrp4	1.90	between exon 1 & 2	Yes
Gcc1	1.89	between exon 1 & 2	Yes
Hes1	1.87	promoter	Yes

Gene Name	Fold Change	NC-XRE location	XRE within 300 bp on an NC-XRE
Slc20a1	1.81	promoter	Yes
Fam222a	1.79	promoter, after exon1, after exon 2	Yes (promoter and after exon 1) No (after exon 2)
Acer2	1.73	between exons 3 & 4 and exons 4 & 5	Yes
Pdk4	1.71	promoter	Yes
Iffo2	1.69	after exon 3	Yes
Arid5b	1.66	between exons 3 & 4 and after exon 7	Yes
Atoh8	1.60	between exons 1 & 2 and 2 & 3	No (exons 1 &2), Yes (exon 2 & 3)
Noct	1.60	promoter and between exons 1 & 2	Yes
Rfx7	1.60	promoter	Yes
Hjurp	1.58	promoter and near exon 7	Yes (promoter), No (near exon 7)
Tshz1	1.52	promoter	Yes



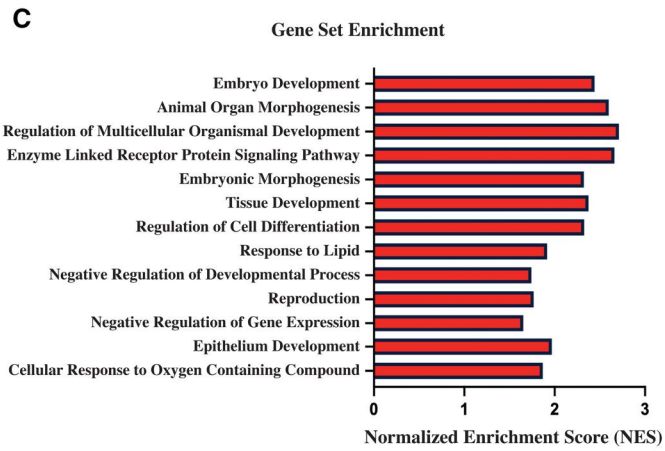


Fig. 3. Analysis of AHR target genes. A) Upregulated AHR target genes (fold change TCDD vs vehicle, 2-h exposure, mouse liver) containing AHR ChIP peaks with 5 or more NC-XRE motifs separated by 50 base pairs or less. AHR target genes are defined as differentially expressed genes (RNA-seq, fold difference in TCDD vs vehicle > 1.5×, FDR < 5%) containing AHR ChIP peaks within 10 kb of a gene body. B) Over-representation analysis (ORA) of Gene Ontology Biological Processes (GOBP) using the AHR target genes from panel A. Enriched pathways were ranked by their statistical significance (*P*-value) along with the number of hit genes and their names. For example, the embryo development GOBP pathway contains 8 upregulated genes (Atoh8, Notch1, Mafb, Lrp4, Maff, Hes1, Tsgz1, Sebox) that have 5 repeats of NC-XRE motifs with a maximum of 50 bp spacers between them. C) Gene Set Enrichment Analysis (GSEA) displaying normalized enrichment scores (NES) for mouse liver global RNA-seq (TCDD vs vehicle) ranked genes. Only the pathways identified in the ORA (panel B) are shown here. All except the "response to oxygen-containing compound" pathway are enriched in the TCDD-exposed mouse liver.

Additionally, we observed significant overlap with several transcription factors identified through ORA and motif enrichment analyses in ChIP-seq datasets derived from non-liver mouse tissues (Fig. S4C), pointing to broader regulatory interactions that may extend beyond liver-specific contexts. These 3 complementary approaches—motif enrichment (Fig. 4A), regulon analysis (Fig. 4B), and ChIP-seq co-localization via giggle (Fig. 4C and D)—provide independent lines of evidence that

A

Motif name	Motif sequence	Percent of AHR peaks with flanking motif & NC-XRE	Percent of AHR peaks with flanking motif & XRE	NC-XRE -log10(p value)	XRE -log10(p value)
Nr1a2	TRAGGTCA	50.62	46.52	215	71
AR halfsite	CCAGGAACAG	44.61	0	27	not significant
RARa	TTGAMCTTTG	40.3	36.23	418	158
Arnt:Ahr	TBGCACGCAA	36.39	67.08	3704	5093
Tgif2	TGTCANYT	33.88	0	36	not significant
ERRa	CAAAGGTCAG	33.42	30.63	307	112
Smad3	TWGTCTGV	31.28	0	17	not significant
Hif1b	RTACGTGC	30.78	50.69	1036	1566
Tgif1	YTGWCADY	30.33	0	15	not significant
Klf14	RGKGGGCGK-GGC	29.25	32.36	25	39
Hic1	TGCCAGCB	27.12	0	26	not significant
COUP-TFII	AGRGGTCA	26.32	23.42	365	119

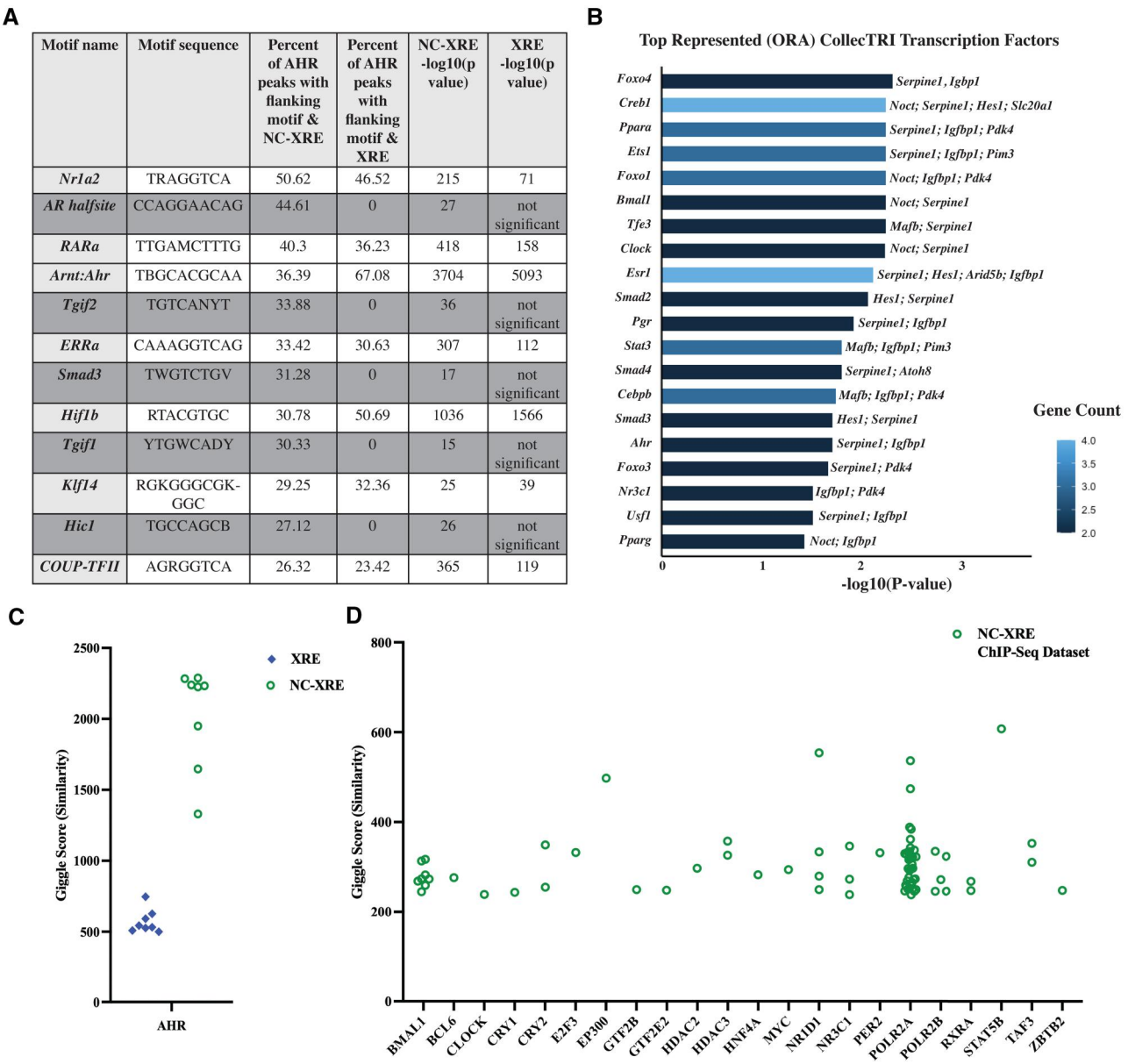


Fig. 4. Putative transcription factors acting with AHR at NC-XRE DNA. A) Top 12 most frequently occurring known transcription factor motifs in the 50 basepairs flanking repeated NC-XRE and XRE motifs (2 or more motifs within 25 basepairs of each other) in genomic regions where AHR binds. Highlighted motifs are enriched flanking NC-XRE but not XRE. Significance defined as $P < 10^{-11}$. B) Over-representation analysis (ORA) using the CollecTRI comprehensive gene regulatory network to identify transcription factors known to upregulate AHR targets (described in Fig. 3A), where the AHR target genes are associated with 5 NC-XRE repeats separated by 50 bp or less. Top represented transcription factors ranked by their statistical significance (−log10 [P-value]). For example, FOXO4 has 2 hits (*Serpine1*, *Igfbp1*), indicating that the *Serpine1* and *Igfbp1* genes are (i) targets of the FOXO4 transcription factor protein, (ii) are targets of AHR (RNAseq from TCDD in mouse liver), and (iii) contain 5 NC-XRE repeats with 50 bp spacers in AHR ChIP peaks within 10 kb of the gene body (ChIP-seq from TCDD in mouse liver). C, D) Giggle scores of AHR peaks with 5 or more XRE or NC-XRE repeats with up to 50 bp spacers compared with experimentally determined transcription factor peaks from mouse liver or liver cell ChIP-seq. Each point represents how strongly XRE or NC-XRE containing AHR-binding sites overlap with the known transcription factor binding sites. A higher giggle score indicates a higher likelihood of these transcription factors co-occupying the AHR peaks. For example, DNA sequence to which STAT5B binds is enriched in AHR ChIP peaks containing NC-XRE (5 or more repeats with 50 bp spacers). Positive control (C) shows high enrichment of AHR DNA binding motifs within AHR ChIP peaks.

converge on overlapping sets of transcription factors. This convergence increases confidence that the identified transcription factors (TFs) are potential partners of AHR at NC-XRE loci, rather than artifacts of any single analysis.

We found several motifs enriched in peaks containing both XRE and NC-XRE motifs. As expected, the ARNT motif was enriched in both, supporting the previously known association with XRE motifs and suggesting that the AHR–ARNT complex could be associated with NC-XRE motifs (Fig. 4A). However, ARNT was not enriched in our regulon ORA involving AHR peaks containing 5× NC-XRE motifs. This observation aligns with the findings of Huang and Elferink (2012), who originally observed AHR bound to the NC-XRE motif in the promoter region of *Serpine1* but did not detect ARNT. This might indicate that AHR can function independently of ARNT at NC-XRE-binding sites.

AHR is known to function through activation complexes involving various coregulators, and our findings suggest that several transcription factors may co-occupy AHR-bound genomic

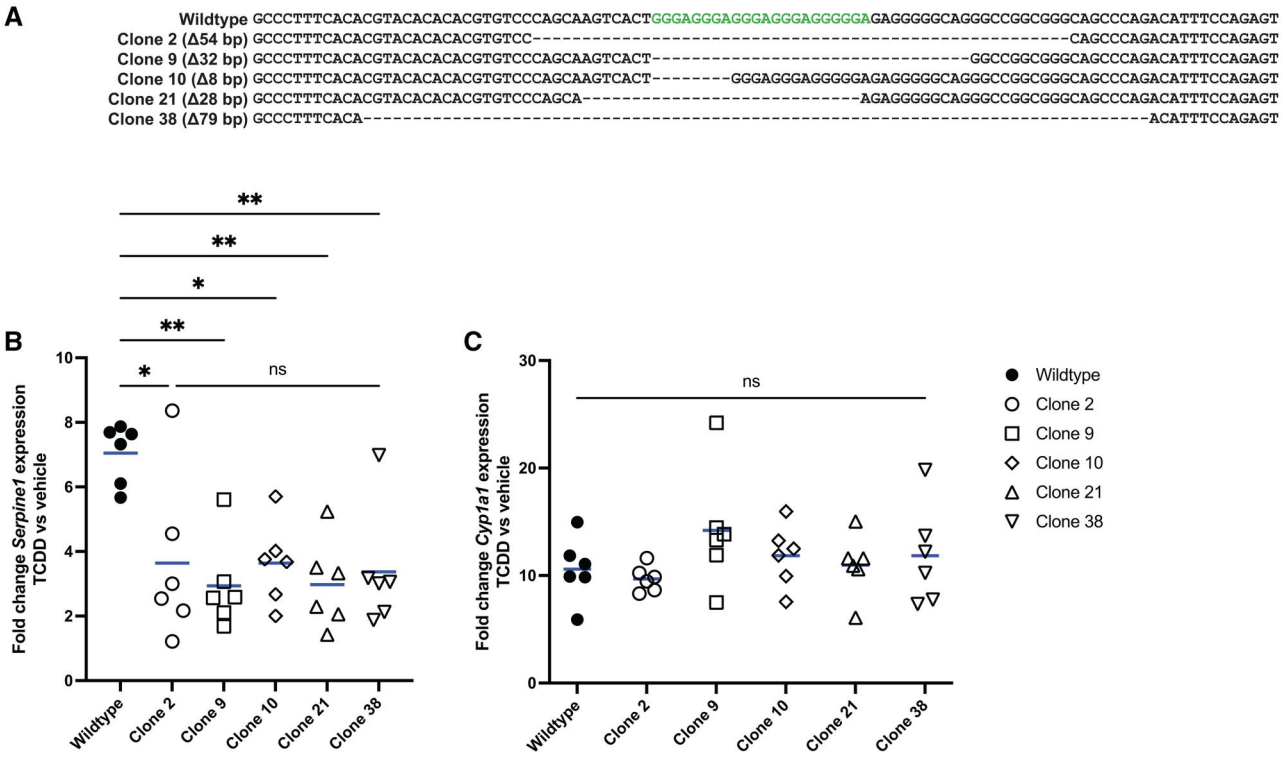


Fig. 5. Mutations in NC-XRE DNA in *Serpine1* promoter reduced TCDD-dependent expression of *Serpine1*. A) Schematic of mouse *Serpine1* gene showing AHR ChIP peak ~150 bp upstream of transcription start site, which contains 5× repeated NC-XRE (nucleotides in red) and no XRE sequences. Hepa 1 to 6 cell clones containing deletions in the NC-XRE region (dashed lines) are shown below the wild-type sequence. B, C) Hepa 1 to 6 wild-type or mutant cells were exposed to 10 nM TCDD or vehicle for 2 h, followed by RNA extraction and qPCR for *Serpine1* (A) or *Cyp1a1* (B). Mutant cell lines with full or partial deletion of NC-XRE exhibit reduced increase in *Serpine1* following TCDD exposure compared with wild-type cells. Mutant cells showed similar increase in *Cyp1a1* following TCDD exposure as wild-type cells. Gene expression normalized to vehicle, 18S rRNA used as reference gene. One-way ANOVA with Tukey's multiple comparisons test, $^{*}P < 0.05$, $^{**}P < 0.01$, ns, not significant ($P > 0.05$).

regions. Among these, BMAL1 (also known as ARNTL), a basic helix-loop-helix (bHLH) transcription factor like AHR, emerged as a particularly intriguing example. BMAL1 is well known for forming heterodimers with CLOCK to regulate circadian rhythm genes such as *Period* (*Per1*, *Per2*, *Per3*) and *Cryptochrome* (*Cry1*, *Cry2*) (Hogenesch et al. 1998; Takumi et al. 1998a, 1998b; Kume et al. 1999). We observed co-occupancy of AHR ChIP-seq peaks containing 5× NC-XRE motifs with both BMAL1- and CLOCK-binding sites (Fig. 4D). We also saw that 2 BMAL1 target genes, *Noct* and *Serpine1*, are significantly overrepresented in the CollecTRI ORA analysis. These 2 target genes are also upregulated in the mouse liver following TCDD exposure (Figs 3A and 4B). Moreover, these AHR-5× NC-XRE peak regions also overlapped with ChIP-seq peaks for *Per2*, *Cry1*, and *Cry2*, suggesting that AHR may be involved in regulating components of the circadian clock. Our findings are consistent with a prior study demonstrating that activated AHR can interact with BMAL1 at E-box sequences and disrupt CLOCK-BMAL1-mediated transcription in cultured human hepatoma cells (Xu et al. 2010). Our data support and extend this work by showing genome-wide co-localization of AHR- and BMAL1-binding sites in the mouse liver following TCDD exposure and suggest that AHR–BMAL1 interactions and subsequent signaling may be mediated through binding at NC-XRE motifs.

Several transcription factor motifs were specifically enriched in the regions flanking NC-XREs, but not in regions flanking canonical XREs. Among these, binding sites for TGIF1 and SMAD3 were particularly prominent (Fig. 4A). Consistent with this, multiple targets of SMAD family transcription factors were upregulated following TCDD exposure in mouse liver (Figs 3A and 4A). Furthermore, both SMADs and TGIF1 non-liver ChIP-seq datasets showed significant overlap with AHR-NC-XRE peaks (Fig. S4C). These findings point to a potential crosstalk between AHR and other known signaling pathways, in this particular case, the TGF-β/SMAD signaling pathway at NC-XRE-enriched regulatory regions. Notably, such crosstalk has not been previously reported, highlighting its potential as a novel and meaningful observation. It suggests that AHR may integrate environmental signals with developmental and homeostatic pathways through non-canonical DNA-binding mechanisms, expanding the scope of AHR's regulatory network.

Other transcription factors such as KLF6 have been reported to associate with AHR at NC-XREs to regulate *Serpine1* expression (Wilson et al. 2013). Our findings are consistent with the idea that NC-XREs may provide a platform for diverse AHR–transcription factor interactions. For example, EP300, which is already known to interact with the AHR:ARNT complex (Kobayashi et al. 1997), showed strong peak overlap with AHR at 5× NC-XRE sites. We also saw overlap with other known co-regulators like E2F and HNF4A (Marlowe et al. 2004; Cholico et al. 2022). Although these proteins were already linked to AHR, their potential binding to the NC-XRE-rich regions has not been studied. Even more interesting, we found transcription factors, which have not been previously linked to AHR as a potential coregulator. These include factors involved in oncogenic signaling (e.g. MYC) and core transcriptional machinery (e.g. POLR2A, POLR2B). Their presence at

the AHR-NC-XRE sites suggests that AHR may regulate cell cycle and RNA production using a previously uncharacterized non-canonical pathway.

Testing NC-XRE function in a natural genomic context

Our bioinformatics analysis supports the hypothesis that AHR regulates target gene expression via NC-XRE DNA. To functionally test this hypothesis at a specific AHR target gene, we focused on *Serpine1*. ChIP studies in mouse liver identified an AHR-binding site ~150 bp upstream of the TSS, which contains a run of 5 NC-XRE motifs (Fig. 5A) (Huang and Elferink 2012). Indirect evidence suggests that AHR upregulates *Serpine1* expression via NC-XRE motifs, but this has never been tested directly.

We cultured mouse Hepa 1 to 6 liver cells, used CRISPR-Cas9 to delete the 5× NC-XRE motif in the *Serpine1* promoter, generated clonal mutant cell lines, and directly tested whether the 5× NC-XRE motif contributes to TCDD-dependent *Serpine1* expression. To control for heterogeneity in gene expression between wild-type cells (Giuliano et al. 2019; Westermann et al. 2022), we first derived a population of wild-type Hepa 1 to 6 cells from a single cell (see Materials and methods section). These clonal wild-type cells were used for all subsequent experiments. We transfected Hepa 1 to 6 cells with Cas9 protein together with a guide RNA targeting the NC-XRE sequence, picked single cells, and established multiple cell lines containing different deletion mutations in the target NC-XRE (Fig. 5A).

We exposed mutant and wild-type cells to TCDD or vehicle for 2 h and assayed *Serpine1* expression by RT-qPCR. In all clones, deletion of the NC-XRE reduced upregulation of *Serpine1* in response to TCDD (Fig. 5B). To test whether AHR function was normal, we also examined expression of *Cyp1a1*, a prototypical AHR target gene thought to be regulated by XRE motifs. All cell lines exhibited increased *Cyp1a1* expression following TCDD exposure, and there was no difference in mean *Cyp1a1* expression levels between wild-type and any mutant cell line (Fig. 5C). Our results suggest that AHR transcription factor function is normal in NC-XRE mutant cell lines. We conclude that AHR upregulates *Serpine1* via this NC-XRE sequence 150 bp upstream of the TSS.

Discussion

We performed a genome-wide assessment of NC-XRE motifs and found they are prevalent in regulatory regions associated with AHR target genes. We also provide functional evidence that NC-XRE motifs in the *Serpine1* promoter are necessary for full upregulation of *Serpine1* in response to TCDD. Although deletion of all 5 NC-XRE motifs in the *Serpine1* promoter reduced TCDD-induced *Serpine1* expression by ~50%, residual induction remained. This suggests that additional AHR-responsive elements, such as the 2 canonical XREs located ~1,000 bp upstream of the Serpine1 promoter, may also contribute to its regulation. Thus, NC-XREs and XREs likely act together to shape the overall transcriptional response of *Serpine1* to TCDD. However, we have only a basic understanding of the physiologic relevance of AHR signaling at NC-XRE motifs. Only 2 target genes (*Serpine1* and *Cdkn1a*) are known to be directly regulated by AHR at NC-XRE DNA, and these interactions were explored in only 1 cell type, hepatocytes, in adulthood (Huang and Elferink 2012; Jackson et al. 2014). The degree to which other AHR target genes are regulated by non-consensus DNA sequence, in additional tissues and cell types and at different stages of organismal development, is not known.

To address this, we performed integrative pathway and transcription factor analyses of AHR target genes associated with 5× NC-XRE clusters. We identified 22 genes upregulated by TCDD that are proximal to these clusters, and pathway enrichment analyses revealed significant associations with biological processes such as development, morphogenesis, oxygen response, and enzymatic activity. These findings suggest that AHR may regulate a broader array of physiological processes through NC-XRE-mediated transcriptional control than previously appreciated.

ARNT is considered an obligate binding partner of AHR (Reyes et al. 1992), but these studies focused on AHR–ARNT interactions at XRE DNA. Whether ARNT binds AHR at NC-XRE DNA and whether NC-XRE-mediated transcriptional responses require ARNT has not been determined. To our knowledge, no studies have directly tested NC-XRE-dependent gene regulation in ARNT-deficient cell lines, and this remains an important question for future research. Studies suggest that under certain conditions, AHR can act independently of ARNT to upregulate gene expression. In mouse liver, AHR was bound to NC-XRE in the promoter of the *Serpine1* gene following TCDD exposure, but no ARNT was detected (Huang and Elferink 2012). In a rat hepatoma cell line containing a Serpine1:luciferase reporter construct, TCDD increased luciferase levels, but knockdown of ARNT had no effect (Huang and Elferink 2012). Our analysis supports this observation. Although ARNT motifs were enriched in AHR peaks containing both XRE and NC-XRE motifs, ARNT did not show co-occupancy at AHR peaks containing 5× NC-XRE motifs. This raises the possibility that AHR may function independently of ARNT at NC-XRE sites, potentially through interactions with alternative co-regulators.

A limitation of the previous studies is that none considered ARNT2. ARNT2 shares significant sequence homology with ARNT and was shown to interact with AHR and XRE DNA in vitro (Hirose et al. 1996; Tanguay et al. 2000; Rowatt et al. 2003), but to our knowledge, there is an absence of evidence of whether ARNT2 interacts with AHR in vivo. DNA binding, coregulator recruitment, and target gene expression could be different between AHR-ARNT and AHR-ARNT2 transcriptional complexes. Given our findings that ARNT1 does not appear to co-occupy AHR-bound NC-XRE regions, it is plausible that ARNT2 may serve as an alternative co-regulator in these contexts. Notably, we were unable to assess ARNT2 co-localization at AHR-NC-XRE peaks due to the absence of publicly available ARNT2 ChIP-seq datasets. As such, the potential involvement of ARNT2 in AHR-NC-XRE-mediated transcription remains an open question that warrants further study. This raises the intriguing possibility that AHR may form distinct transcriptional complexes depending on the DNA motif it binds—partnering with ARNT1 at XREs and potentially with ARNT2 at NC-XREs. Such differential dimerization could influence target gene specificity, chromatin recruitment, and transcriptional outcomes, warranting deeper investigation into the role of ARNT2 in non-canonical AHR signaling. However, expression of Arnt2 in mice is largely restricted to neuronal tissues and the developing kidney, with little to no detectable expression in mouse liver (Keith et al. 2001; Aitola and Pelto-Huikko 2003). Consistent with this expression pattern, overexpression of Arnt2 in hepatic cells produces minimal TCDD responsiveness compared with Arnt (Dougherty and Pollenz 2008). Together, this suggests that although Arnt2 could influence AHR-NC-XRE-mediated transcription, it is unlikely to function as an AHR-binding partner in the liver.

Our integrative analysis uncovered a broader network of transcriptional regulators co-localizing at AHR-NC-XRE sites, including factors involved in circadian rhythm (e.g. Arntl/Bmal1, Clock), TGF-β/SMAD signaling (e.g. SMAD3, TGIF1), and oncogenic or epigenetic regulation (e.g. MYC, EP300, BCL6). These findings suggest that AHR may orchestrate complex transcriptional programs through non-canonical DNA binding, extending its influence into diverse biological pathways.

Limitations

The evidence for transcription factor co-occupancy at AHR-bound NC-XRE clusters is indirect. We compared AHR peaks from TCDD-treated liver with transcription factor peaks obtained from untreated liver or hepatocyte datasets, which do not provide temporal or spatial resolution of binding events. It, therefore, remains possible that nearby transcription factors occupy adjacent loci without functionally interacting with AHR. For example, a canonical E-box motif (CACGTG, recognized by CLOCK-BMAL1 and other bHLH factors) lies immediately upstream of the NC-XRE cluster in the *Serpine1* promoter. Yet deletion of this sequence in clone 38 had no additional effect on TCDD-induced *Serpine1* expression beyond NC-XRE deletion alone. This indicates that, at least for *Serpine1*, the E-box is dispensable for AHR-dependent regulation and that co-occupancy with CLOCK-BMAL1 is unlikely to explain the observed transcriptional response. These findings underscore that functional validation at individual loci is essential. Direct demonstration of simultaneous AHR and co-factor binding will require future experiments such as sequential ChIP, proximity ligation assays, or single-molecule imaging.

The computational analyses performed here suggest many hypotheses, such as that AHR target genes may be regulated by NC-XRE DNA. However, the algorithms employed here make predictions but lack experimental validation. We demonstrated that the repeated NC-XRE motifs in the promoter region of *Serpine1* regulate AHR-dependent *Serpine1* expression in a mouse liver cell line. A recent study in human lung cancer cells showed upregulation of SERPINE1 by AHR agonists, suggesting SERPINE1 is also an AHR target in humans (Gorelik et al. 2025). Nevertheless, it remains unclear whether AHR regulates gene expression in humans via NC-XRE DNA motifs, as the human *SERPINE1* promoter contains only 3 GGGA repeats separated by 50 bp spacers. However, we appreciate that DNA regulatory elements may be incompletely conserved between species. Showing that AHR binds NC-XRE DNA to regulate gene expression in any species is important foundational knowledge. Although the specific number or configuration of NC-XRE sequence at a given promoter or enhancer may differ between strains of animals or between species, it is extremely unlikely that AHR regulates gene expression by binding to NC-XRE in mice but not in humans. Considering that AHR binds identical XRE DNA sequence in zebrafish, mice, and humans (Denison et al. 1988; Yao and Denison 1992; Tanguay et al. 1999; Andreasen et al. 2002), we argue that the same is likely true for NC-XRE.

Materials and methods

Re-analysis of published murine liver ChIP-Seq datasets

ChIP-Seq datasets of AHR in murine livers after 2 h of TCDD treatment in male (GSE97634) and female (GSE97636) mice were downloaded from NCBI Gene Expression Omnibus (GEO) (Table 1) (Nault et al. 2016; Fader et al. 2017, 2019). Data were trimmed using trim_galore (Krueger 2015), mapped to the mouse genome build UCSC mm10 using bowtie2 (Langmead and Salzberg 2012), and peaks were called using MACS2 using the provided controls (Zhang et al. 2008; Langmead and Salzberg 2012). Peaks were merged using bedtools into male, female, and male + female merged datasets (Quinlan and Hall 2010). Each peak was searched for AHR motifs using an in-house script. To test the likelihood of finding AHR motifs anywhere in the genome, random DNA peaks were generated matching the chromosome and size distribution as previously described (Coarfa et al. 2020), then searched for AHR motifs using a custom Ruby code. We analyzed AHR ChIP-seq peaks by extracting the underlying genomic sequences and scanning them for both XRE and NC-XRE motifs. Motifs provided in IUPAC notation were converted into regular expressions, allowing systematic identification and annotation of motif occurrences within peak regions (doi.org/10.5281/zenodo.17211132). For each motif of interest (XRE, NC-XRE, RelBAhRE), we compared the total number of motif-containing peaks in the observed dataset against the matched random set using a Fisher's exact test. Because only 3 tests were performed, we did not perform multiple testing correction and reported raw Fisher's exact test *P*-values with significance defined at $P < 0.05$. Venn diagrams of AHR motifs in the merged AHR peaks were derived using *eulerr* package (Larsson 2024) in R Studio (RStudio Team 2025), and pie charts were made using GraphPad Prism version 10.2.2, MacOS, GraphPad Software, Boston, MA, United States, www.graphpad.com.

To investigate the genomic distribution of AHR-binding sites, we used ChromHMM, a chromatin state annotation tool (Ernst and Kellis 2012). We utilized ChIP-Seq datasets for 8 histone modifications—H3K27ac, H3K27me3, H3K36me3, H3K79me2, H3K9ac, H3K4me3, H3K4me1, and H3K9me3—along with an input control, all derived from 8-week-old mouse liver tissue. These datasets were downloaded from the NCBI GEO (accession numbers: GSM1000140, GSM918718, GSM1000150, GSM1000151, GSM1000152, GSM1000153, GSM769014, GSM769015, GSM751035) (Table 1) (Nault et al. 2016; Fader et al. 2017, 2019). Raw reads were trimmed using *trim_galore* (Krueger 2015) and aligned to the UCSC mm10 mouse genome build using *bowtie2* (Langmead and Salzberg 2012). The aligned reads were converted into binned signal representations alongside genomic windows of signal abundance, which were then used to train a hidden Markov model to segment the genome into 10 chromatin states (Ernst and Kellis 2017) using the ChromHMM tool (Ernst and Kellis 2012). We overlaid these chromatin states onto AHR peaks containing XRE and NC-XRE motifs, as shown in Figs 1C and D and 2D, using a custom in-house script. Additionally, we used the same histone ChIP-Seq datasets to compute signal distributions across genomic regions (Fig. 2E) using the HOMER software (Heinz et al. 2010). All figures were generated using GraphPad Prism version 10.2.2.

Runs of AHR motifs, e.g. sequences of successive motifs within a maximum distance from each other, were determined using an in-house Python script and bedtools. Flanking sequences were determined using the bedtools software (Quinlan and Hall 2010). Enriched motifs within the flanking sequences were determined using the HOMER software (Heinz et al. 2010).

To identify transcription factors with potential co-binding at AHR-bound NC-XRE or XRE clusters, we performed giggle score analysis using the CistromeDB Toolkit (Mei et al. 2017; Zheng et al. 2019). As input, we used mouse liver AHR ChIP-seq peaks (GSE97634, GSE97636) containing either 5 canonical XRE motifs or 5 NC-XRE motifs with a maximum spacer of 50 base pairs.

Giggle calculates an integrated score that combines statistical significance (Fisher's exact test *P*-value) and enrichment (odds ratio) to quantify whether 2 ChIP-seq peak sets overlap more than expected by chance (Layer et al. 2018).

The giggle score thus reflects both the strength and significance of overlap: Higher scores indicate stronger co-localization between 2 datasets, consistent with the possibility that the corresponding proteins bind the same genomic loci and may participate in transcriptional complexes.

Although giggle automatically compares query peaks against all publicly available datasets in the Cistrome Data Browser, for interpretation, we filtered the results to include only ChIP-seq datasets derived from mouse liver tissue or hepatocyte-derived cell lines. This restriction ensured that the comparisons were performed in a biologically relevant context.

The giggle scores were visualized in Fig. 4C and D to illustrate overlap with AHR datasets (serving as an internal control) and with additional transcription factors that may act as potential AHR co-regulators.

Re-analysis of published murine liver AHR RNA-Seq datasets and integration with ChIP-seq datasets

Several RNA-seq datasets of mouse liver after treatment with TCDD at different doses and for multiple timepoints were downloaded from NCBI GEO: GSE109863, GSE62902, and GSE87519 (Nault et al. 2015, 2018; Fader et al. 2017) (Table 1). Data were trimmed using trim_galore (Krueger 2015), then mapped onto the mouse genome build UCSC mm10 using STAR (Dobin et al. 2013), and gene expression was quantified using feature_counts (Liao et al. 2014). Differentially expressed genes between TCDD and vehicle treatment were inferred using the EdgeR and RUVr R packages (Robinson et al. 2010; Risso et al. 2014), with significance achieved at a fold change exceeding 1.5× and FDR-adjusted *P*-value < 0.05. We used the ranked gene list to perform GSEA (Subramanian et al. 2005) using datasets from the mouse molecular signature database version 7.5.1 (Castanza et al. 2023). Genomic locations (represented as bed files) and gene signatures were integrated using bedtools (Quinlan and Hall 2010). To capture both promoter and proximal enhancer effects, peaks were considered as associated with genes if the gene body was within 10,000 bp of a peak. ORA was performed using a custom Python script, which applies the hypergeometric test used by MSigDB to assess enrichment of gene sets from the gene ontology biological process gene set version 7.5.1 (Castanza et al. 2023) and CollecTRI version 2 regulons (Müller-Dott et al. 2023).

Cell culture

The mouse hepatoma cell line Hepa 1 to 6 was obtained from ATCC (Manassas, VA) (Darlington et al. 1980). Hepa 1 to 6 cells were cultured in 10 cm plates in Dulbecco's Modified Eagle's medium supplemented with 1% (v/v) Penicillin/Streptomycin and 10% (v/v) fetal bovine serum (FBS) in a humidified incubator (5% CO_2, 37 °C). To derive wild-type clones for CRISPR mutagenesis, Hepa 1 to 6 cells were separated using BD FACSAria II cell sorter (Cytometry and Cell Sorting Core, Baylor College of Medicine) and 1 cell per well was plated into a 96-well plate. Cells grew for 2 to 3 weeks until colonies were visible and selected for further expansion. Wild-type clone 12 was used for subsequent studies.

CRISPR mutagenesis of Hepa 1 to 6 cells

We used CHOPCHOP (Labun et al. 2019) to design guide RNAs targeting the 5× NC-XRE motif region 150 bp upstream of transcription start in the mouse Serpine1 gene (target sequence CAGCAAGTCACTGGGAGGGAGGG, PAM motif underlined). Synthetic single-guide RNA—modified with 2′-O-methyl at the first 3 and last bases and 3′-phosphorothioate bonds between the first 3 and last 2 bases—and purified SpCas9-2NLS protein was purchased from Synthego (Redwood City, CA). We mixed guide RNA and Cas9 (1.3:1 ratio) to form ribonucleoproteins according to Synthego's recommended protocol. Guide RNA and Cas9 ribonucleoprotein complexes in serum-free medium were transfected into Hepa 1 to 6 WT clone 12 cells (see above) using Lipofectamine CRISPRMAX Cas9 Transfection reagent (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's protocol; 100,000 Hepa-1 to 6 WT-12 cells in standard cell culture media (above, containing FBS and antibiotics) were mixed with the ribonucleoprotein-transfection solution and split into 2 wells for genomic analysis and clonal expansion. Cells were incubated in a humidified incubator (5% CO_2, 37 °C) for 24 h. We then changed the media and allowed the cells to incubate for another 3 days before genotyping or deriving single-cell colonies for clonal expansion.

To derive mutant colonies from single cells, we separated and plated transfected WT-12 cells using BD FACSAria II cell sorter as above. Forty-eight single-cell clones were expanded and genotyped, focusing on mutant cell lines from clones 2, 9, 10, 21, and 38.

Genotyping cell lines

Genomic DNA was extracted from cells using phenol–chloroform extraction. Cells were lysed with DNA lysis buffer (10 mM Tris, 200 mM NaCl, 5 mM EDTA, 1% SDS, 0.4 mg/ml proteinase K). Equal volume of phenol:choloroform:isoamyl alcohol was added for DNA extraction into the upper phase, followed by ethanol–sodium acetate precipitation of DNA. We used PCR to amplify the region flanking the NC-XRE motifs using primers 5′-AAGCCAGG CCAACTTTTCCT and 5′-CGGTCCTCCTTCACAAAGCT. PCR products were Sanger sequenced to identify mutations. To identify mutations from mosaic cell populations, we used the ICE CRISPR Analysis Tool (Conant et al. 2022) to deconvolute Sanger sequencing results.

Mutant clones (all derived from WT clone 12), selected for further experiments, were clone 2 (54 bp deletion in NC-XRE region of Serpine1 promoter), clone 9 (32 bp deletion), clone 10 (8 bp deletion), clone 21 (28 bp deletion), and clone 38 (74 bp deletion). The deletions in each clone were further validated by next-generation sequencing. We used PCR to amplify a 450 bp region of DNA encompassing the NC-XRE. After PCR amplification, DNA was quantified by Qubit dsDNA HS fluorometric quantitation (ThermoFisher Q32851). Cytation 5 imaging reader (BioTek; Santa Clara, CA) was used to read the samples; 500 ng of each amplicon was sent for Amplicon E-Z Next Generation Sequencing (GENEWIZ from Azenta Life Sciences, South Plainfield, NJ), paired-end Illumina MiSeq 2X250 bp, 50,000 reads per sample. Each clone contained >99% of the expected mutation.

Quantitative reverse-transcription PCR

Hepa 1-6 cells (parental WT12 cells and WT12 mutant clonal cell lines 2, 9, 10, 21, and 38) at 70% to 80% confluency were treated with 10 nM TCDD ($n = 6$ per treatment group) or 0.1% DMSO (vehicle control, $n = 6$ per treatment group) for 2 h each. We selected 10 nM TCDD because this concentration has been shown to robustly activate AHR target genes while avoiding overt cytotoxicity (Humphrey-Johnson et al. 2015). Cells were then lysed in

TRIzol followed by RNA extraction according to the manufacturer's protocol using Direct-zol RNA Miniprep Kit, including the optional on-column DNase digestion (cat no. 11-331; Zymo Research, Irvine, CA). RNA concentration was measured using NanoDrop 2000 Spectrophotometer (Thermo Scientific, Waltham, MA). The purity of the RNA (A260/A280) was ≥2 with the yield of 60 ng/μl or higher; 1,000 ng of the RNA were converted into cDNA using iScript Reverse Transcription Supermix for RT-qPCR (Biorad, Hercules, CA; cat no 1708841) in a total 20 μl volume and incubated for 5 min at 25 °C followed by incubation at 46 °C for 20 min, at 95 °C for 1 min in C1000 Touch Thermal Cycler (Biorad, Hercules, CA).

To validate differential gene expression, we used real-time quantitative PCR (qPCR) by the PrimePCR Assay (Biorad, Hercules, CA) based on PCR amplification using the CFX Opus 96 Real-Time PCR System (Biorad, Hercules, CA) via the $2^{-\Delta\Delta Ct}$ method under default reaction conditions (95 °C for 2 min; 95 °C for 5 s; 60 °C for 30 s for 40 cycles). An aliquot of 4 μl of 1:10 dilution of cDNA was mixed with 10 μL BIORAD SsoAdvanced Universal Probes Supermix (cat no. 1725280) and target gene-specific PrimePCR Assays (1 μl) and nuclease-free water up to a total volume of 20 μl (3 technical replicates were run for each sample). cDNA samples were also used for measurement of Gapdh levels for normalization. Ct values were calculated by the Bio-Rad instrument using the CFX Maestro Software version 2.3.

The following PrimePCR Assays were used. Note that BioRad does not provide specific primer/probe sequences. BioRad provides the amplicon sequence with additional base pairs added to the beginning and/or end of the sequence (amplicon context sequence). This is in accordance with the minimum information for the publication of real-time qPCR experiments (MIQE) guidelines (Bustin et al. 2011).

Serpine1, Mouse (qMmuCIP0037358)

Amplicon context sequence: CTGTCGGGTTGTGCCGAACCACAAAGAGAAAGGATCGGTCTATAACCATCTCCG

TGGGGGCCATGCGGGCTGAGATGACAAAGGCTGTGGAGGAAGACGCCACTGTG

CCGCTCTCGTTTACCTCGATCCTGACCTTTTGCAGTGCCTGTGCT

***Gapdh*, Mouse (qMmuCEP0039581):**

TGGGAGTTGCTGTTGAAGTCGCAGGAGACAACCTGGTCCTCAGTGTAGCCCAAGATGCCCTTCAGTGGGCCCTCAGATGCCTGCTTCACCACCTTCTTGATGTCA

***Cyp1a1*, Mouse (qMmuCEP0041980):**

CAGCCACCTAGATCATGCCTTCCATGTATGGACTTCCAGCCTTCGTGTCAGCCA

CAGAGCTGCTCCTGGCTGTCACCGTATTCTGCCTTGGATTCTGGGTGGTCAGAG

CCACAAGAACCTGGGTTCCCAAAGGCCTGAAGACTCC

The results shown were the average of 3 or more biologically independent experiments. The cumulative results were statistically analyzed by using Prism version 10 (GraphPad Software, Boston, MA).

Acknowledgments

We dedicate this manuscript to our dear friend and colleague Professor Cornelis Elferink, who provided invaluable advice and encouragement as we pursued these studies. Professor Elferink, who discovered the association between AHR and NC-XRE DNA, died unexpectedly during the preparation of this manuscript.

Supplementary material

Supplementary material is available at *Toxicological Sciences* online.

Funding

T.D.P., S.L.G., and C.C. were partially supported by The Cancer Prevention Institute of Texas (CPRIT) [RP210227, RP200504], National Institutes of Health (NIH)/National Cancer Institute P30 shared resource grant [CA125123], NIH/National Institute of Environmental Health Sciences (NIEHS) center grants [P30 ES030285] and [P42 ES027725], and NIH/National Institute on Minority Health and Health Disparities [P50MD015496]. D.A.G. was supported by NIEHS R01 ES026337.

Conflicts of interest. None declared.

References

Aitola MH, Pelto-Huikko MT. 2003. Expression of Arnt and Arnt2 mRNA in developing murine tissues. J Histochem Cytochem. 51:41–54.

Andreasen EA, Hahn ME, Heideman W, Peterson RE, Tanguay RL. 2002. The zebrafish (*Danio rerio*) aryl hydrocarbon receptor type 1 is a novel vertebrate receptor. Mol Pharmacol. 62:234–249.

Beischlag TV, Perdew GH. 2005. ER alpha-AHR-ARNT protein-protein interactions mediate estradiol-dependent transrepression of dioxin-inducible gene transcription. J Biol Chem. 280:21607–21611.

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, *et al.* 2011. Primer sequence disclosure: a clarification of the MIQE guidelines. Clin Chem. 57:919–921.

Castanza AS, Recla JM, Eby D, Thorvaldsdóttir H, Bult CJ, Mesirov JP. 2023. Extending support for mouse data in the Molecular Signatures Database (MSigDB). Nat Methods. 20:1619–1620.

Cholico GN, Nault R, Zacharewski TR. 2022. Genome-wide ChIPseq analysis of AhR, COUP-TF, and HNF4 enrichment in TCDD-treated mouse liver. Int J Mol Sci. 23:1558.

Coarfa C, Grimm SL, Katz T, Zhang Y, Jangid RK, Walker CL, Moorthy B, Lingappan K. 2020. Epigenetic response to hyperoxia in the neonatal lung is sexually dimorphic. Redox Biol. 37:101718.

Conant D, Hsiau T, Rossi N, Oki J, Maures T, Waite K, Yang J, Joshi S, Kelso R, Holden K, *et al.* 2022. Inference of CRISPR edits from sanger trace data. CRISPR J. 5:123–130.

Darlington GJ, Bernhard HP, Miller RA, Ruddle FH. 1980. Expression of liver phenotypes in cultured mouse hepatoma cells. J Natl Cancer Inst. 64:809–819.

Denison MS, Fisher JM, Whitlock JP. 1988. Inducible, receptor-dependent protein-DNA interactions at a dioxin-responsive transcriptional enhancer. Proc Natl Acad Sci U S A. 85:2528–2532.

Denison MS, Soshilov AA, He G, DeGroot DE, Zhao B. 2011. Exactly the same but different: promiscuity and diversity in the molecular mechanisms of action of the aryl hydrocarbon (dioxin) receptor. Toxicol Sci. 124:1–22.

Dere E, Lo R, Celius T, Matthews J, Zacharewski TR. 2011. Integration of genome-wide computation DRE search, AhR ChIP-chip and gene expression analyses of TCDD-elicited responses in the mouse liver. BMC Genomics. 12:365.

Diani-Moore S, Ram P, Li X, Mondal P, Youn DY, Sauve AA, Rifkind AB. 2010. Identification of the aryl hydrocarbon receptor target gene TiPARP as a mediator of suppression of hepatic gluconeogenesis by 2,3,7,8-tetrachlorodibenzo-p-dioxin and of

nicotinamide as a corrective agent for this effect. J Biol Chem. 285:38801–38810.
Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 29:15–21.
Dougherty EJ, Pollenz RS. 2008. Analysis of Ah receptor-ARNT and Ah receptor-ARNT2 complexes in vitro and in cell culture. Toxicol Sci. 103:191–206.
Ernst J, Kellis M. 2012. ChromHMM: automating chromatin-state discovery and characterization. Nat Methods. 9:215–216.
Ernst J, Kellis M. 2017. Chromatin-state discovery and genome annotation with ChromHMM. Nat Protoc. 12:2478–2492.
Fader KA, Nault R, Doskey CM, Fling RR, Zacharewski TR. 2019. 2,3,7,8-Tetrachlorodibenzo-p-dioxin abolishes circadian regulation of hepatic metabolic activity in mice. Sci Rep. 9:6514.
Fader KA, Nault R, Kirby MP, Markous G, Matthews J, Zacharewski TR. 2017. Convergence of hepcidin deficiency, systemic iron overloading, heme accumulation, and REV-ERBα/β activation in aryl hydrocarbon receptor-elicited hepatotoxicity. Toxicol Appl Pharmacol. 321:1–17.
Giuliano CJ, Lin A, Girish V, Sheltzer JM. 2019. Generating single cell-derived knockout clones in mammalian cells with CRISPR/Cas9. Curr Protoc Mol Biol. 128:e100.
Gorelik A, Paulo JA, Schroeter CB, Lad M, Shurr A, Mastrokalou C, Siddiqi S, Suyari O, Brognard J, Walter D, *et al.* 2025. CRISPR screens and quantitative proteomics reveal remodeling of the aryl hydrocarbon receptor-driven proteome through PARP7 activity. Proc Natl Acad Sci U S A. 122:e2424985122.
Hankinson O. 2005. Role of coactivators in transcriptional activation by the aryl hydrocarbon receptor. Arch Biochem Biophys. 433:379–386.
Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 38:576–589.
Hirose K, Morita M, Ema M, Mimura J, Hamada H, Fujii H, Saijo Y, Gotoh O, Sogawa K, Fujii-Kuriyama Y. 1996. cDNA cloning and tissue-specific expression of a novel basic helix-loop-helix/PAS factor (Arnt2) with close sequence similarity to the aryl hydrocarbon receptor nuclear translocator (Arnt). Mol Cell Biol. 16:1706–1713.
Hogenesch JB, Gu YZ, Jain S, Bradfield CA. 1998. The basic-helix-loop-helix-PAS orphan MOP3 forms transcriptionally active complexes with circadian and hypoxia factors. Proc Natl Acad Sci U S A. 95:5474–5479.
Huang G, Elferink CJ. 2012. A novel nonconsensus xenobiotic response element capable of mediating aryl hydrocarbon receptor-dependent gene expression. Mol Pharmacol. 81:338–347.
Humphrey-Johnson A, Abukalam R, Eltom SE. 2015. Stability of the aryl hydrocarbon receptor and its regulated genes in the low activity variant of hepa-1 cell line. Toxicol Lett. 233:59–67.
Jackson DP, Li H, Mitchell KA, Joshi AD, Elferink CJ. 2014. Ah receptor-mediated suppression of liver regeneration through NC-XRE-driven p21Cip1 expression. Mol Pharmacol. 85:533–541.
Joshi AD, Mustafa MG, Lichti CF, Elferink CJ. 2015. Homocitrullination is a novel histone H1 epigenetic mark dependent on aryl hydrocarbon receptor recruitment of carbamoyl phosphate synthase 1. J Biol Chem. 290:27767–27778.
Larsson J. 2024. eulerr: Area-Proportional Euler and Venn Diagrams with Ellipses [Computer software]. R Package. https://cran.r-project.org/web/packages/eulerr/index.html
Keith B, Adelman DM, Simon MC. 2001. Targeted mutation of the murine arylhydrocarbon receptor nuclear translocator 2 (Arnt2) gene reveals partial redundancy with Arnt. Proc Natl Acad Sci U S A. 98:6692–6697.
Klein-Hitpass L, Kaling M, Ryffel GU. 1988. Synergism of closely adjacent estrogen-responsive elements increases their regulatory potential. J Mol Biol. 201:537–544.
Klinge CM, Kaur K, Swanson HI. 2000. The aryl hydrocarbon receptor interacts with estrogen receptor alpha and orphan receptors COUP-TFI and ERRalpha1. Arch Biochem Biophys. 373:163–174.
Kobayashi A, Numayama-Tsuruta K, Sogawa K, Fujii-Kuriyama Y. 1997. CBP/p300 functions as a possible transcriptional coactivator of Ah receptor nuclear translocator (Arnt). J Biochem. 122:703–710.
Krueger F. 2015. Trim Galore!: a wrapper around Cutadapt and FastQC to consistently apply adapter and quality trimming to FastQ files, with extra functionality for RRBS data. Babraham Institute. https://github.com/FelixKrueger/TrimGalore
Kume K, Zylka MJ, Sriram S, Shearman LP, Weaver DR, Jin X, Maywood ES, Hastings MH, Reppert SM. 1999. mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. Cell. 98:193–205.
Labun K, Montague TG, Krause M, Torres Cleuren YN, Tjeldnes H, Valen E. 2019. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. Nucleic Acids Res. 47:W171–W174.
Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods. 9:357–359.
Layer RM, Pedersen BS, DiSera T, Marth GT, Gertz J, Quinlan AR. 2018. GIGGLE: a search engine for large-scale integrated genome analysis. Nat Methods. 15:123–126.
Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 30:923–930.
Lo R, Matthews J. 2012. High-resolution genome-wide mapping of AHR and ARNT binding sites by ChIP-Seq. Toxicol Sci. 130:349–361.
Ma Q, Baldwin KT, Renzelli AJ, McDaniel A, Dong L. 2001. TCDD-inducible poly(ADP-ribose) polymerase: a novel response to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Biochem Biophys Res Commun. 289:499–506.
Madak-Erdogan Z, Katzenellenbogen BS. 2012. Aryl hydrocarbon receptor modulation of estrogen receptor α-mediated gene regulation by a multimeric chromatin complex involving the two receptors and the coregulator RIP140. Toxicol Sci. 125:401–411.
Marlowe JL, Knudsen ES, Schwemberger S, Puga A. 2004. The aryl hydrocarbon receptor displaces p300 from E2F-dependent promoters and represses S phase-specific gene expression. J Biol Chem. 279:29013–29022.
Mei S, Qin Q, Wu Q, Sun H, Zheng R, Zang C, Zhu M, Wu J, Shi X, Taing L, *et al.* 2017. Cistrome data browser: a data portal for ChIP-Seq and chromatin accessibility data in human and mouse. Nucleic Acids Res. 45:D658–D662.
Mouse ENCODE Consortium; Stamatoyannopoulos JA, Snyder M, Hardison R, Ren B, Gingeras T, Gilbert DM, Groudine M, Bender M, Kaul R, Canfield T, *et al.* 2012. An encyclopedia of mouse DNA elements (Mouse ENCODE). Genome Biol. 13:418.
Müller-Dott S, Tsirvouli E, Vazquez M, Ramirez Flores RO, Badia-I-Mompel P, Fallegger R, Türei D, Lægreid A, Saez-Rodriguez J. 2023. Expanding the coverage of regulons from high-confidence prior knowledge for accurate estimation of transcription factor activities. Nucleic Acids Res. 51:10934–10949.
Nault R, Doskey CM, Fader KA, Rockwell CE, Zacharewski T. 2018. Comparison of hepatic NRF2 and aryl hydrocarbon receptor binding in 2,3,7,8-tetrachlorodibenzo-p-dioxin-treated mice

demonstrates NRF2-independent PKM2 induction. Mol Pharmacol. 94:876–884.
Nault R, Fader KA, Kirby MP, Ahmed S, Matthews J, Jones AD, Lunt SY, Zacharewski TR. 2016. Pyruvate kinase isoform switching and hepatic metabolic reprogramming by the environmental contaminant 2,3,7,8-tetrachlorodibenzo-p-dioxin. Toxicol Sci. 149:358–371.
Nault R, Fader KA, Zacharewski T. 2015. RNA-Seq versus oligonucleotide array assessment of dose-dependent TCDD-elicited hepatic gene expression in mice. BMC Genomics. 16:373.
Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 26:841–842.
Reyes H, Reisz-Porszasz S, Hankinson O. 1992. Identification of the Ah receptor nuclear translocator protein (Arnt) as a component of the DNA binding form of the Ah receptor. Science. 256:1193–1195.
Risso D, Ngai J, Speed TP, Dudoit S. 2014. Normalization of RNA-seq data using factor analysis of control genes or samples. Nat Biotechnol. 32:896–902.
Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 26:139–140.
Rowatt AJ, DePowell JJ, Powell WH. 2003. ARNT gene multiplicity in amphibians: characterization of ARNT2 from the frog *Xenopus laevis*. J Exp Zool B Mol Dev Evol. 300:48–57.
RStudio Team. 2025. RStudio: integrated development for R. Boston (MA): RStudio, Inc.
Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, *et al.* 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 102:15545–15550.
Takumi T, Matsubara C, Shigeyoshi Y, Taguchi K, Yagita K, Maebayashi Y, Sakakida Y, Okumura K, Takashima N, Okamura H. 1998a. A new mammalian period gene predominantly expressed in the suprachiasmatic nucleus. Genes Cells. 3:167–176.
Takumi T, Taguchi K, Miyake S, Sakakida Y, Takashima N, Matsubara C, Maebayashi Y, Okumura K, Takekida S, Yamamoto S, *et al.* 1998b. A light-independent oscillatory gene mPer3 in mouse SCN and OVLT. EMBO J. 17:4753–4759.
Tanguay RL, Abnet CC, Heideman W, Peterson RE. 1999. Cloning and characterization of the zebrafish (*Danio rerio*) aryl hydrocarbon receptor. Biochim Biophys Acta. 1444:35–48.
Tanguay RL, Andreasen E, Heideman W, Peterson RE. 2000. Identification and expression of alternatively spliced aryl hydrocarbon nuclear translocator 2 (ARNT2) cDNAs from zebrafish with distinct functions. Biochim Biophys Acta. 1494:117–128.
Tennant BR, Robertson AG, Kramer M, Li L, Zhang X, Beach M, Thiessen N, Chiu R, Mungall K, Whiting CJ, *et al.* 2013. Identification and analysis of murine pancreatic islet enhancers. Diabetologia. 56:542–552.
van der Velde A, van der Fan K, Tsuji J, Moore JE, Purcaro MJ, Pratt HE, Weng Z. 2021. Annotation of chromatin states in 66 complete mouse epigenomes during development. Commun Biol. 4:239.
Vogel CFA, Sciullo E, Li W, Wong P, Lazennec G, Matsumura F. 2007. RelB, a new partner of aryl hydrocarbon receptor-mediated transcription. Mol Endocrinol. 21:2941–2955.
Westermann L, Li Y, Göcmen B, Niedermoser M, Rhein K, Jahn J, Cascante I, Schöler F, Moser N, Neubauer B, *et al.* 2022. Wildtype heterogeneity contributes to clonal variability in genome edited cells. Sci Rep. 12:18211.
Wilson SR, Joshi AD, Elferink CJ. 2013. The tumor suppressor Kruppel-like factor 6 is a novel aryl hydrocarbon receptor DNA binding partner. J Pharmacol Exp Ther. 345:419–429.
Xu C-X, Krager SL, Liao D-F, Tischkau SA. 2010. Disruption of CLOCK-BMAL1 transcriptional activity is responsible for aryl hydrocarbon receptor-mediated regulation of Period1 gene. Toxicol Sci. 115:98–108.
Yao EF, Denison MS. 1992. DNA sequence determinants for binding of transformed Ah receptor to a dioxin-responsive enhancer. Biochemistry. 31:5060–5067.
Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, *et al.* 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9:R137.
Zheng R, Wan C, Mei S, Qin Q, Wu Q, Sun H, Chen C-H, Brown M, Zhang X, Meyer CA, *et al.* 2019. Cistrome data browser: expanded datasets and new tools for gene regulatory analysis. Nucleic Acids Res. 47:D729–D735.

Toxicological Sciences, 2025, 209, 1–15
https://doi.org/10.1093/toxsci/kfaf146
Research article