

High-Resolution Genome-wide Mapping of AHR and ARNT Binding Sites by ChIP-Seq

Raymond Lo and Jason Matthews¹

Department of Pharmacology and Toxicology, University of Toronto, Toronto, Ontario M5S 1A8, Canada

¹To whom correspondence should be addressed at Department of Pharmacology & Toxicology, University of Toronto, Medical Sciences Building, Room 4336, 1 King's College Circle, Toronto, ON M5S 1A8, Canada. Fax: (416) 978-2723. E-mail: jason.matthews@utoronto.ca.

Received June 26, 2012; accepted August 7, 2012

The aryl hydrocarbon receptor (AHR) and AHR nuclear translocator (ARNT) activated complex regulates genes in response to the environmental contaminant 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). AHR has also emerged as a potential therapeutic target for the treatment of human diseases and different cancers, including breast cancer. To better understand AHR and ARNT signaling in breast cancer cells, we used chromatin immunoprecipitation linked to high-throughput sequencing to identify AHR- and ARNT-binding sites across the genome in TCDD-treated MCF-7 cells. We identified 2594 AHR-bound, 1352 ARNT-bound, and 882 AHR/ARNT cobound regions. No significant differences in the genomic distribution of AHR and ARNT were observed. Approximately 60% of the cobound regions contained at least one core an aryl hydrocarbon response element (AHRE), 5'-GCGTG-3'. AHR/ARNT peak density was the highest within 1 kb of transcription start sites (TSS); however, a number of AHR/ARNT cobound regions were located as far as 100 kb from TSS. *De novo* motif discovery identified a symmetrical variation of the AHRE (5'-GTGCGTG-3'), as well as FOXA1 and SP1 binding motifs. Microarray analysis identified 104 TCDD-responsive genes where 98 genes were upregulated by TCDD. Of the 104 regulated genes, 69 (66.3%) were associated with an AHR- or ARNT-bound region within 100 kb of their TSS. Overall our study identified AHR/ARNT cobound regions across the genome, revealed the importance but not absolute requirement for an AHRE in AHR/ARNT interactions with DNA, and identified a modified AHRE motif, thereby increasing our understanding of AHR/ARNT signaling pathway.

Key Words: Aryl hydrocarbon receptor (AHR); AHR nuclear translocator (ARNT); Chromatin immunoprecipitation (ChIP); ChIP with next-generation sequencing (ChIP-Seq).

The aryl hydrocarbon receptor (AHR) belongs to the basic helix-loop-helix (bHLH) PER-ARNT-SIM (PAS) family of transcription factors (McIntosh *et al.*, 2010). Other members of this family include hypoxia-inducible factor 1 alpha (HIF-1 α), AHR nuclear translocator (ARNT; HIF-1 β), single-minded (SIM), and the AHR repressor (AHRR). The AHR mediates the toxic effects of environmental contaminants, such as

2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). AHR binds and is activated by numerous structurally diverse natural and synthetic chemicals (Denison and Nagy, 2003). The AHR regulates an array of physiological responses including xenobiotic metabolism, vasculature development, immunosuppression, T-cell differentiation, reproduction, and cell cycle progression (Bock and Köhle, 2006). AHR has also recently emerged as a potential therapeutic target for immune disorders (Pot, 2012) and for the treatment of different cancers, including breast cancer (Safe and Wormke, 2003).

In its nonliganded state, the AHR resides in the cytoplasm bound to a chaperone protein complex that includes heat shock protein 90, AHR-interacting protein (AIP), and p23 (Hankinson, 1995). Once activated by ligand, AHR translocates into the nucleus where it associates with its obligatory heterodimerization partner ARNT. The AHR/ARNT heterodimer binds to its cognate DNA sequence motif, referred to as an aryl hydrocarbon response element (AHRE; 5'-TNGCGTG-3') with a minimum core sequence of 5'-GCGTG-3' required for AHR/ARNT binding (Swanson *et al.*, 1995). Once bound to chromatin, the activated AHR/ARNT heterodimer induces the recruitment of coregulator proteins resulting in changes in target gene expression, including *cytochrome P450 1A1* (*CYP1A1*), *CYP1B1*, *nuclear factor (erythroid-derived 2)-like 2* (*NFE2L2*; *NRF2*), and *AHRR* (Baba *et al.*, 2001; Miao *et al.*, 2005; Whitlock, 1999; Zhang *et al.*, 1998). Recent chromatin immunoprecipitation assays coupled with high-density microarrays (ChIP-chip) have identified novel genomic sequences bound by AHR, providing new insight into the binding preferences of AHR across the genome (Ahmed *et al.*, 2009; De Abrew *et al.*, 2010; Dere *et al.*, 2011; Lo *et al.*, 2011; Pansoy *et al.*, 2010; Sartor *et al.*, 2009). However, compared with newer technologies such as chromatin immunoprecipitation coupled with next-generation high-throughput sequencing (ChIP-Seq), tiled array-based approaches have a lower signal-to-noise ratio, exhibit hybridization biases, and are limited to the sequences present on the array (Wold and Myers, 2008).

ARNT is a class II bHLH PAS transcription factor and the general heterodimerization partner for many bHLH proteins including AHR, hypoxia-inducible factor 1 and 2 (HIF-1 α and 2 α), and single-minded protein 1 and 2 (SIM1 and 2) (McIntosh *et al.*, 2010). Conditional knockout of ARNT prevented gene regulation by AHR and HIF-1 α , confirming a critical role of ARNT in AHR and HIF-1 α signaling (Tomita *et al.*, 2000). Given the requirement for ARNT in the transactivation of the AHR and HIF-1 α , identifying genomic binding sites recognized by both partners of the heterodimerization complex would provide a more accurate definition of the DNA recognition sequences after receptor activation. For example, Schödel *et al.* (2011) recently reported extensive overlap between HIF-1 α and ARNT (referred to as HIF-1 β in their study) in 356 distinct genomic regions using ChIP-Seq. The HIF-1 α and ARNT cobound regions reported in that study provided new insight into how the HIF-1 α /ARNT complex interacts with chromatin and regulates target gene expression across the genome. Despite numerous studies describing the genomic binding patterns of AHR, there have been no reports investigating the overlapping binding profiles of ligand-activated AHR and ARNT (Ahmed *et al.*, 2009; De Abrew *et al.*, 2010; Dere *et al.*, 2011; Lo *et al.*, 2011; Pansoy *et al.*, 2010; Sartor *et al.*, 2009). The identification of AHR/ARNT cobound genomic sequences would improve our understanding of AHR transactivation and AHR interactions with chromatin and provide a robust data set that can be used to re-evaluate the AHR/ARNT DNA recognition sequence. To this end, we performed ChIP-Seq and mapped AHR and ARNT cobound genomic regions in TCDD-treated MCF-7 human breast cancer cells.

MATERIALS AND METHODS

Chemicals and antibodies. Dimethyl sulfoxide (DMSO) was purchased from Sigma Aldrich (Oakville, ON, Canada), whereas TCDD was purchased from Wellington Laboratories, Inc. (Guelph, ON, Canada). Antibodies against AHR (H-211, sc-5579 Lot no. K1010) and ARNT (H-172, sc-5580, lot no. G3010) were purchased from Santa Cruz Biotechnology (Santa Cruz, CA). Validation of ARNT-bound regions was performed using an anti-ARNT antibody (NB100-110) from Novus Biologicals (Oakville, ON, Canada). All other reagents used were of the highest quality and scientific standards.

Cell culture. MCF-7 cells were incubated at 37°C and 5% CO₂ and maintained in Dulbecco's modified Eagle's medium (DMEM) 1.0 g/l glucose (Wisent, St-Bruno, QC, Canada) supplemented with 10% fetal bovine serum (Wisent) and 1% penicillin and streptomycin (PEST; Wisent). Cells were plated in phenol red-free DMEM 1.0 g/l glucose supplemented with 2.5% dextran-coated charcoal-treated fetal calf serum and 1% PEST for approximately 72 h before treatment.

ChIP-Seq. MCF-7 cells were plated at a density of 3 million per 10-cm dishes. Conventional ChIP assays were conducted as described previously (Matthews *et al.*, 2007). Immunoprecipitated DNA was eluted in a final volume of 50 μ l, and 20 μ l was separated by electrophoresis using a 1.0% agarose gel. DNA fragments in the 200–400 bp range were extracted from the gel using the QIAquick Gel Extraction Kit (Qiagen, MD), eluted in a final volume of 10 μ l, and amplified using SeqPlex according to manufacturer's protocol (Sigma Aldrich, St Louis, MO). Library preparation and high-throughput sequencing

were performed at BGI (Shenzhen, China). Briefly, isolated DNA was end repaired to create a 3'-dA overhang. Adapters were ligated to the end of DNA fragments. DNA fragments between 100 and 500 bp were selected after PCR amplification and sequenced using the HiSeq2000 (Illumina, Inc., San Diego, CA) for a final sequencing depth of 20 million reads per sample, 50 bp per read. Sequences were aligned to the human genome version 19 (hg19) using Short Oligo Analysis Package 2.21 (BGI) (Li *et al.*, 2009). Enriched peaks were normalized and analyzed using the two-sample peak calling function in CisGenome (Ji *et al.*, 2008) by comparing AHR and ARNT against duplicates of nonimmunoreactive IgG. Regions with a minimum length of 100 bp were extended by 150 bp and regions less than 50 bp apart were merged. Confirmation of AHR-bound and ARNT-bound regions was performed on extracts from DMSO or TCDD-treated MCF-7 cells that were immunoprecipitated with normal IgG, anti-AHR antibody (H-211, sc-5579 Lot no. K1010) from Santa Cruz Biotechnology, and anti-ARNT antibody (NB100-110) from Novus Biologicals.

Transcription factor binding site analysis. The presence of an AHRE was determined using a family of matrices for the AHR-ARNT heterodimers and AHR-related factors available at MatInspector (www.genomatix.de). Sequences with a matrix similarity score of at least 0.8 were considered AHRE-containing regions. Over-represented transcription factor binding site analysis was determined using the default parameter in RegionMiner (www.genomatix.de) based on the number of matches in immunoprecipitated DNA compared with the number of expected matches in genomic background. Module families and matrices were considered significant at z-score > 3. To identify transcription factors that might work in concert with AHR to modulate AHR transcription, we also determined the over-representation of transcription factor binding motifs located within 10–50 bp from putative AHREs in the AHR-ARNT-bound regions. Information on the definitions of the family matrices discussed herein is available on the Genomatix web page (www.genomatix.de). *De novo* motif discovery was conducted using the Gibbs motif sampler in CisGenome using a mean motif length of 12 bp and a maximum length of 30 bp. Significantly over-represented motifs extracted from Gibbs motif sampler were compared with existing binding motifs from the JASPAR (Wasserman and Sandelin, 2004) and TRANSFAC (Matys *et al.*, 2003) libraries using STAMP (Mahony and Benos, 2007).

RNA extraction and gene expression microarray analysis. MCF-7 cells were plated at a density of 350,000 per well. Total RNA was extracted using the Aurum Total RNA Mini Kit according to the manufacturer's recommendations (Bio-Rad, Toronto, ON, Canada). RNA was prepared using the Ambion WT kit and hybridized to the Affymetrix Human Exon 1.0 ST array at the Toronto Centre for Applied Genomics (Toronto, ON, Canada). Differentially regulated genes were identified with the Partek Genomics Suite (Partek, Inc., St Louis, MO) using a false detection rate of 5% (FDR5) and an absolute fold change of greater than 1.20 between TCDD- and DMSO-treated samples.

Functional analysis. Canonical and function pathways significantly enriched by TCDD through AHR/ARNT binding were determined using Fisher's exact test with the Ingenuity Pathway Analysis (Ingenuity Systems, Inc., Redwood, CA) software.

Statistical analyses. Statistical analysis was performed using ANOVA with Bonferroni's *post hoc* test at $p < 0.05$. The significance values of biological functions/diseases identified in the IPA studies were calculated using the right-tailed Fisher's exact test.

RESULTS

Identification of AHR/ARNT Binding Sites

To identify high-resolution genomic AHR/ARNT binding sites, we performed ChIP-Seq on chromatin isolated from MCF-7 breast cancer cells exposed to 10 nM TCDD for 45 min.

The experimental conditions were selected to best represent initial and maximal AHR/ARNT binding events that occurred prior to transcriptional changes in response to TCDD exposure (Pansoy *et al.*, 2010). The genomic coordinates of AHR- and ARNT-bound regions are provided in Supplementary tables S1 and S2. As expected, TCDD exposure resulted in the enrichment of multiple peaks in the previously characterized AHRE

clusters upstream of the bidirectional promoter for the AHR target genes, *CYP1A1* and *CYP1A2* (Fig. 1). We identified 2594 AHR-bound region at a false detection rate ~1% (FDR1) and 1352 ARNT-bound regions at FDR5. The FDRs were chosen to yield two data sets of similar size, while at the same time adjusting for differences in antibody specificity and affinity. AHR- and ARNT-bound regions were referred to as AHR_{number}

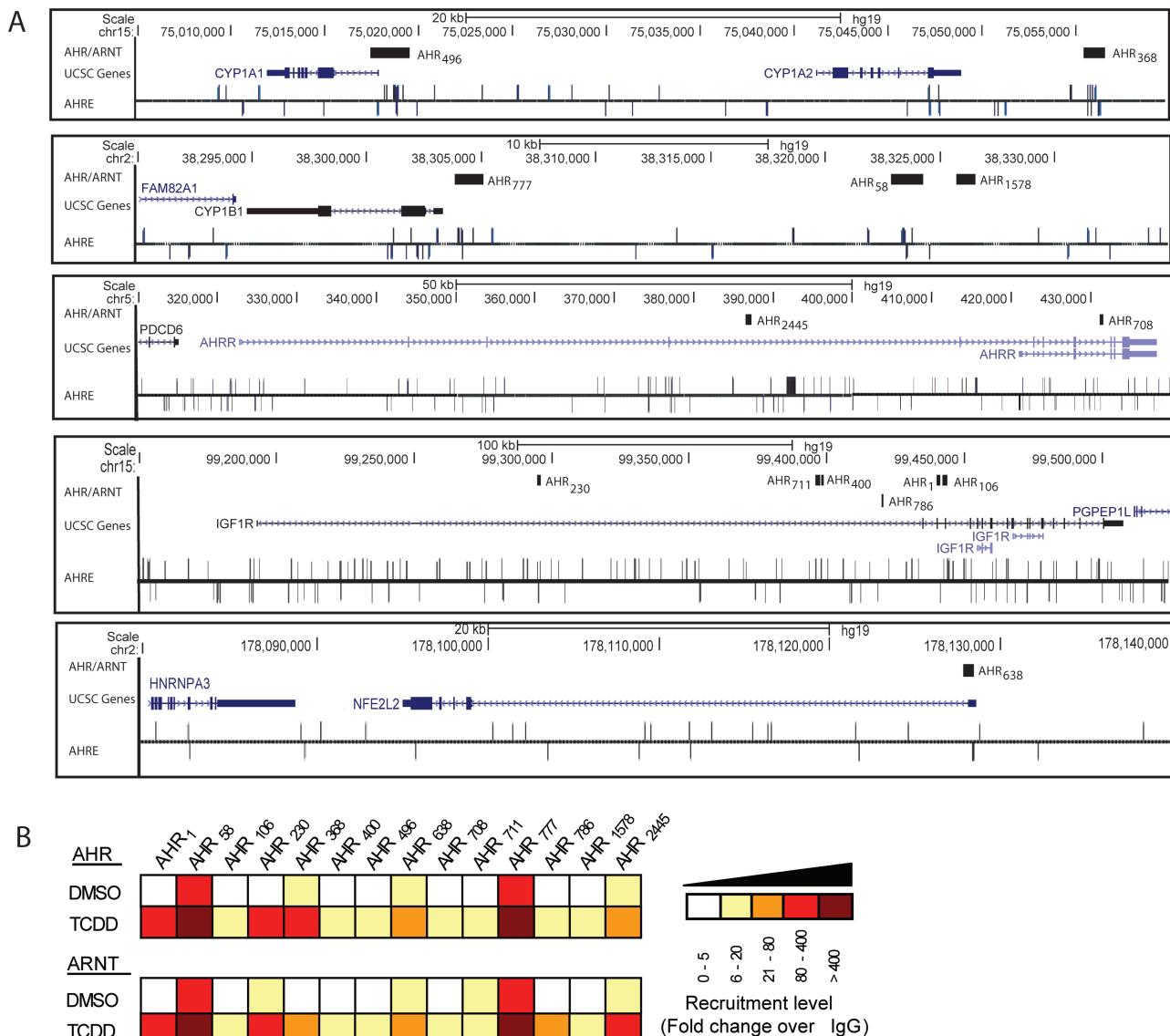


FIG. 1. Distribution of AHR- and ARNT-bound regions around six AHR-regulated genes—CYP1A1, CYP1A2, CYP1B1, AHRR, IGF1R, and NFE2L2. (A) Six TCDD-responsive genes from our microarray experiments were chosen to represent the distribution of AHR- and ARNT-bound regions around their target genes. In the top track, black blocks represent AHR/ARNT cobound regions identified by ChIP-Seq and were labeled according to their rank numbers in the AHR data set. Shown in the middle track are the coding regions and introns of UCSC genes, which are represented by blocks and connecting lines, respectively. The 5' and 3' untranslated regions (UTR) are indicated by thinner blocks, whereas arrows indicate the direction of transcription. In the bottom track, vertical lines indicate putative AHREs identified by MatInspector (Genomatix) at a core similarity of 0.75. AHR/ARNT cobound regions were distributed in both intragenic and intergenic regions across large genomic distances. (B) Confirmation of AHR- and ARNT-bound regions. All regions shown in Figure 1A were confirmed (ANOVA, $p < 0.05$) for AHR and ARNT recruitment by qPCR. Recruitment levels were shown as fold increase over IgG negative controls from at least two independent experiments and were represented by the color intensity on the heat map. Although some regions displayed constitutive AHR/ARNT occupancy, TCDD significantly increased AHR/ARNT recruitment to all regions.

TABLE 1
Genomic Distribution of AHR- and ARNT-Bound Regions

	Peak height	No. of peaks	Intergenic	Intragenic	Exon	Intron	CDS	Upstream from TSS		
								1 kb	10 kb	100 kb
AHR bound	93.4	2594	1508	1086	41	1051	15	58	261	1265
ARNT bound	75.5	1352	797	555	20	537	6	27	134	666
Overlap	153.2	882	501	381	9	373	3	15	85	446

and ARNT_{number}, respectively, where the number indicates the relative ranking within the analysis. The average peak lengths were 638 and 560 bp for AHR- and ARNT-bound regions, respectively. Locations of AHR- and ARNT-bound regions were divided into eight categories (intergenic, intragenic, exon, intron, coding DNA sequence [CDS], TSSup1k [1 kb upstream from TSS], TSSup10k, and TSSup100k) based on available RefSeq annotation (Table 1). There were no significant differences (Fisher's test, $p > 0.05$) between the genomic distribution of AHR and ARNT data sets. Peak density (i.e., number of peaks/bp) was highest immediately adjacent to TSS (Figs. 2C and D), and promoter regions (500 bp upstream/100 bp downstream as defined by www.genomatix.de) were enriched by 3.6-fold in the AHR data set and by 3.0-fold in the ARNT data set. In terms of absolute numbers, AHR and ARNT binding sites were distributed such that ~2% of the binding sites were immediately upstream (1 kb), ~10% within 10 kb upstream, and ~50% within 100 kb upstream from TSS. These findings support the notion that similar to members of the nuclear receptor family of transcription factors, AHR can regulate target gene expression through distal *cis*-acting elements.

AHR/ARNT binding sites were determined by the overlap of AHR- and ARNT-bound regions. Region overlap was determined based on a 100% sequence similarity and between regions that overlapped with > 50% of the width of the smallest region using the ViroBlast web server (Deng *et al.*, 2007). This resulted in 882 (65% overlap) AHR-/ARNT-bound regions (Fig. 2; the overlap list is provided in Supplementary table S3). AHR/ARNT cobound regions were labeled based on their corresponding rankings in the AHR data set (Supplementary table S1). The average peak height for the AHR/ARNT overlap data set was 153.2, which was significantly higher than those for AHR (93.4) and ARNT (75.5) (Table 1).

AHR/ARNT Binding Site Analysis

To evaluate the importance of the AHRE in the recruitment of the AHR complex, we interrogated the AHR, ARNT, and cobound data sets using a family of AHRE position weight matrices (PWMs) available in MatInspector (www.genomatix.de). Approximately 50.5 and 47.9% of the AHR and ARNT data sets contained at least one AHRE, whereas the percentage improved to 60% in the AHR/ARNT overlap data set (Fig. 2B). These results were in agreement with previous reports showing

that the AHRE was the major determinant, but not an absolute prerequisite for AHR/ARNT recruitment (Ahmed *et al.*, 2009; De Abrew *et al.*, 2010; Dere *et al.*, 2011; Lo *et al.*, 2011; Pansoy *et al.*, 2010; Sartor *et al.*, 2009).

AHR has been reported to interact with an alternate response element termed AHRE-II (5'-CATGnnnnnnC[T/A]TG-3') (Boutros *et al.*, 2004; Sogawa *et al.*, 2004). Transcription factor binding sites analysis was performed to determine the number of AHR-/ARNT-bound regions that contained an AHRE-II site. We determined that 116/882 (13.1%) AHR-/ARNT-bound regions contained at least one AHRE-II. Of the 353 AHR-/ARNT-bound regions that did not contain an AHRE core, 38 of them contained an AHRE-II site. However, only one TCDD-regulated gene contained an AHRE-II site without a core AHRE in its upstream regulatory region. These findings confirmed the importance of the AHRE in comparison to the AHRE-II in AHR/ARNT interactions with DNA.

De novo Motif Discovery of a Symmetrical AHRE

To address the possibility of an alternate AHR/ARNT binding motif and to improve the existing AHRE binding motif, which was based on *in vitro* AHR-DNA interaction and mutation analysis of the CYP1A1 AHRE (Shen and Whitlock, 1992; Swanson *et al.*, 1995), we did *de novo* motif discovery on the top 500 regions in AHR-bound, ARNT-bound, and overlap data sets using Gibbs motif sampler in CisGenome. The unsupervised analysis identified the previously reported core AHRE (5'-GCGTG-3') in all three data sets, but not the extended AHRE (5'-TnGCGTG-3') (Swanson *et al.*, 1995) (Fig. 3). PWM indicated a strong preference for a guanine and a thymine flanking the 5' end of the core AHRE, resulting in an extended and symmetrical motif with the sequence 5'-GTGCGTG-3' (highlighted in dotted line box in Figure 3A). *De novo* motif discovery was repeated using a web-based application W-ChIPMotifs (Jin *et al.*, 2009), which confirmed the symmetrical motif 5'-GTGCGTG-3' (data not shown). Our motif matrices from the AHR/ARNT overlap data set were compared with the PWM library from JASPAR and TRANSFAC using STAMP (Fig. 3B). In addition to the symmetrical AHRE, Gibbs motif sampler identified additional binding motifs that aligned to MA0277.1_AZF1, MA0148.1_FOXA1, and MA0079.2_SP1.

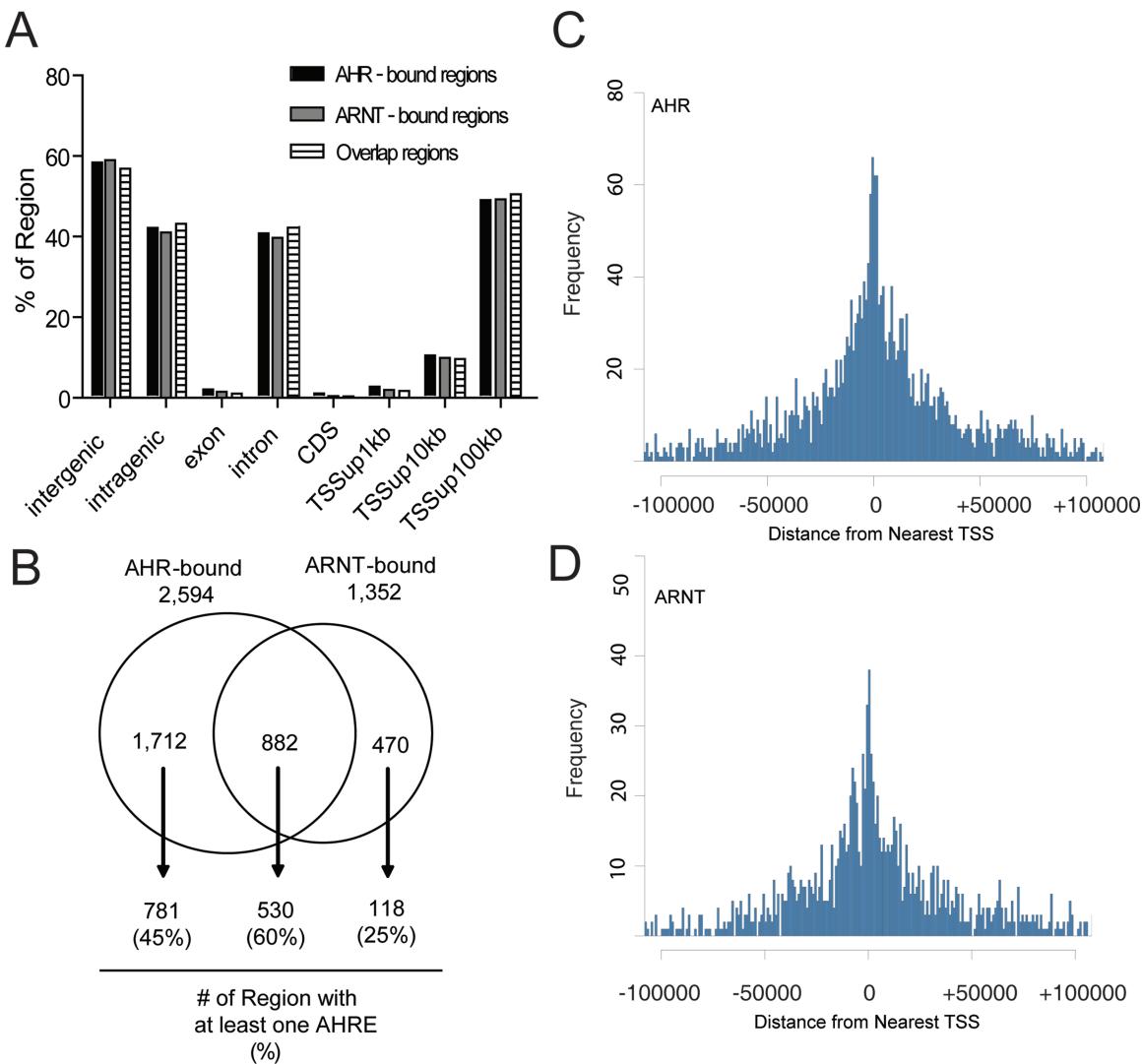


FIG. 2. Genomic distribution of AHR- and ARNT-bound regions and the percentage of regions containing an AHRE. (A) AHR- and ARNT-bound regions were defined as either intergenic (exon, intron, and CDS) or intragenic (TSSup1kb, TSSup10kb, and TSSup100kb). There were no significant differences (Fisher's exact test, $p > 0.05$) between the genomic distribution of AHR and that of ARNT. (B) AHR- and ARNT-bound regions containing at least one AHRE were identified by Genomatix (www.genomatix.de). Of the 882 AHR/ARNT cobound regions, approximately 60% contained at least one AHRE. Distribution of (C) AHR- and (D) ARNT-bound regions indicated a binding preference for regions proximal to TSS. Peak density was the highest immediately adjacent to TSS for both AHR and ARNT although the analysis also identified many distal AHR- and ARNT-bound regions.

Transcription Factor Binding Site Enrichment Analysis

To identify transcription factor binding motifs that were over-represented in the AHR-, ARNT-, and cobound regions, we did transcription factor binding site enrichment analysis using RegionMiner (www.genomatix.de; Table 2). Complete lists of over-represented transcription factor binding sites in all data sets are provided in Supplementary tables S4, S5, and S6. As expected, RegionMiner detected a significant over-representation of AHR_ARNT binding sites in the AHR-bound, the ARNT-bound, and the overlap data sets. Other notable over-represented binding motifs included early growth response 1 (EGR1), activator protein family (AP1 and AP2), SP1, estrogen receptor (ER), and HIF. The antioxidant response element

recognized by NRF2 was also over-represented in all three data sets, which was in agreement with noted interplay between AHR-NRF2 signaling (Yeager *et al.*, 2009). To identify other transcription factors that might work in concert with AHR to modulate AHR function, we searched the AHR/ARNT regions for enriched transcription factor motifs located within 10–50 bp of AHREs using RegionMiner (Genomatix). This analysis revealed enrichment of AHREs in close proximity to EGR, AP1, homologues of enhancer of split complex, and forkhead box (FKHD), suggesting possible physical interaction or coregulation of target genes between the candidate transcription factors and AHR (Table 3). A complete list is provided in Supplementary table S7. Interestingly, some module pairings

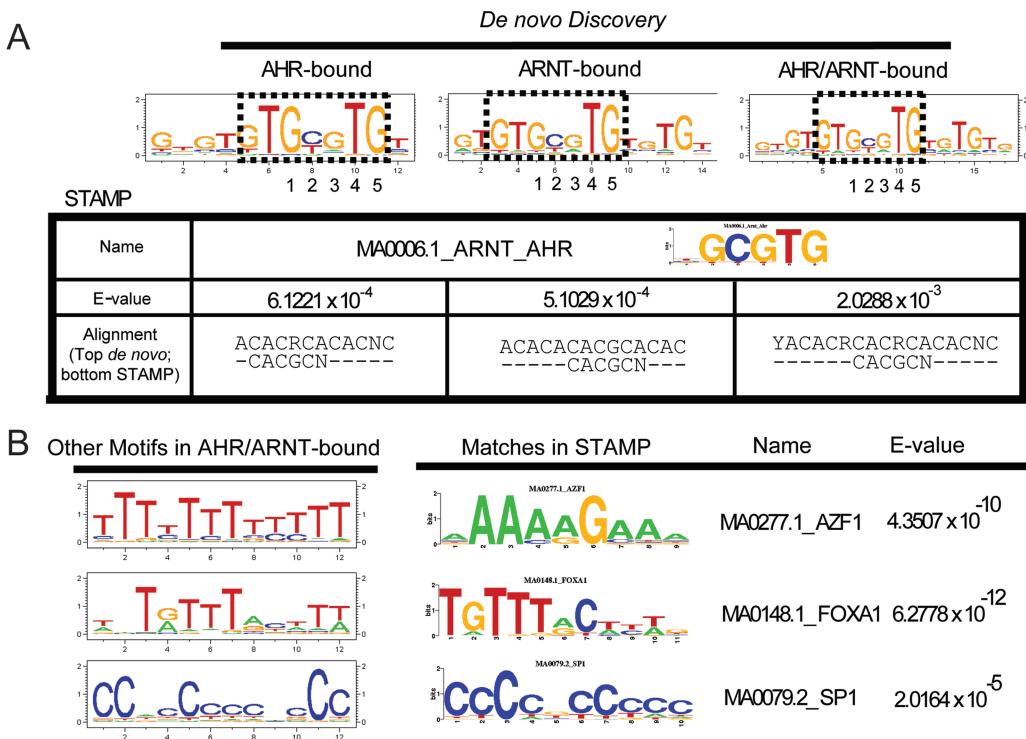


FIG. 3. Discovery of novel binding motifs for AHR- and ARNT-bound regions. *De novo* motif discovery was performed using Gibbs motif sampler in CisGenome (Ji *et al.*, 2008). (A) A symmetrical variation (highlighted in dotted box) of the AHRE was identified in all three data sets. Comparison to known STAMP and TRANSFAC motifs identified the AHRE as the most probable candidate binding motif. (B) *De novo* discovery yielded three motifs in addition to the AHRE. STAMP identified those three motifs to be AZF1, FOXA1, and SP1.

displayed spatial preferences for interaction. For example, the Maf_AP1-related factors (AP1R) binding motif and the ERE were uniformly distributed around an AHRE, whereas the pairing of AHREs and FKHD binding motifs appeared to have a spatial preference of > 25 bp apart in our AHR/ARNT cobound regions.

Confirmation of ChIP-Seq AHR/ARNT Binding Profiles Using qPCR ChIP

AHR/ARNT cobound regions and computationally determined AHREs were plotted using UCSC Genome Browser to investigate the relationship between the presence of an AHRE and AHR/ARNT recruitment to select AHR target genes. *CYP1A1* and *CYP1A2* are regulated by AHR/ARNT recruitment to their bidirectional promoter (Ueda *et al.*, 2006) (Fig. 1). In addition to the well-characterized AHRE cluster located approximately 1 kb upstream from *CYP1A1* TSS, we observed significant AHR/ARNT recruitment to a region 3' of *CYP1A2* that had been previously reported to be important for the AHR-dependent regulation of *CYP1A2* (Okino *et al.*, 2007). For *CYP1B1*, we confirmed AHR/ARNT recruitment to a region within 2 kb upstream of the TSS, which represents the proximal promoter (MacPherson and Matthews, 2010) and upstream AHR-rich enhancer (Matthews *et al.*, 2005). Our analysis also revealed AHR/ARNT binding to two novel

regions approximately 15 kb upstream of the TSS (Fig. 1). All AHR/ARNT cobound regions associated with *CYP1A1*, *CYP1A2*, and *CYP1B1* contained an AHRE core sequence with the exception of AHR₁₅₇₈.

Haarmann-Stemmann *et al.* identified a functional AHRE ~17 kb downstream of the AHRR TSS between exon 1 and exon 2 using reporter gene assays in HepG2 human hepatoma cells (Haarmann-Stemmann *et al.*, 2007). We did not detect AHR/ARNT recruitment to the AHRE-rich sequence between exon 1 and 2. We identified two AHR/ARNT cobound regions in the intronic regions of *AHRR*, regions AHR₇₀₈ and AHR₂₄₄₅ (Fig. 1). Neither contained an AHRE core sequence. Interestingly, we also identified an AHRE cluster at approximately chr15: 390,000–395,000 between *AHRR* exon 5 and 6, but this was not associated with an AHR-/ARNT-bound region.

Three AHREs have been identified in the 2 kb surrounding the murine *nrf2* TSS that might be involved in the AHR-dependent regulation of that gene (Miao *et al.*, 2005). Comparison between the mouse and the human genome revealed at least five potential AHREs in the regulatory region of the human *NRF2* (Miao *et al.*, 2005), but to our knowledge, AHR or ARNT recruitment to these regions has not been experimentally confirmed. We identified and confirmed one AHR-/ARNT-bound region (AHR₆₃₈) upstream of the *NRF2* start codon (Fig. 1), which contained two AHREs.

TABLE 2
Analysis of Enriched Transcription Binding Sites

TF families	TF description	Number of matches	Expected	z-score
Top 10 enriched transcription factor binding sites in AHR-/ARNT-bound regions				
AHR	AHR-arnt heterodimers and AHR-related factors	923	358.38	29.81
AP2F	Activator protein 2	1144	499.61	28.82
AP1F	Activator protein 1	1297	745.75	20.18
EGRF	EGR/nerve growth factor-related factors	1264	757.42	18.4
AP1R	MAF- and AP1-related factors(Antioxidant response element)	2567	1846.92	16.77
SP1F	GC-Box factors SP1/GC	1193	759.85	15.7
ZF02	C2H2 zinc-finger transcription factors 2	1499	1020.41	14.98
EREF	Estrogen response element	1059	730.53	12.14
HESF	Homologues of enhancer of split complex	892	598.79	11.97
MAZF	Myc associated zinc fingers	642	409.84	11.45
Top 10 enriched transcription factor binding sites in AHR-bound regions				
AP2F	Activator protein 2	2953	1296.35	46.01
AHR	AHR-arnt heterodimers and AHR-related factors	2086	929.92	37.9
AP1F	Activator protein 1	3415	1935.02	33.65
AP1R	MAF- and AP1-related factors(Antioxidant response element)	6669	4792.29	27.14
EGRF	EGR/nerve growth factor-related factors	3073	1965.32	24.99
SP1F	GC-Box factors SP1/GC	2811	1971.61	18.9
EREF	Estrogen response element	2710	1895.54	18.71
ZF02	C2H2 zinc-finger transcription factors 2	3497	2647.69	16.51
HIFF	Hypoxia-inducible factor, bHLH/PAS protein family	1363	888.32	15.91
CTCF	CCCTC-binding factor	1557	1054.79	15.45
Top 10 enriched transcription factor binding sites in ARNT-bound regions				
AP2F	Activator protein 2	1308	539.49	33.08
AHR	AHR-arnt heterodimers and AHR-related factors	1032	386.99	32.77
AP1F	Activator protein 1	1473	805.28	23.52
EGRF	EGR/nerve growth factor-related factors	1454	817.89	22.24
AP1R	MAF- and AP1-related factors(Antioxidant Response Element)	2887	1994.35	20
SP1F	GC-Box factors SP1/GC	1361	820.5	18.86
ZF02	C2H2 zinc-finger transcription factors 2	1679	1101.86	17.38
MAZF	Myc associated zinc fingers	755	442.56	14.83
KLFS	Krüppel like transcription factors	2698	2070.91	13.79
EREF	Estrogen response element	1154	788.85	12.99

TABLE 3
Analysis of Enriched Transcription Binding Sites Within 10–50 bp From an AHRE

Top 10 enriched modules in AHR-/ARNT-bound regions				
Modules with AHR	Modules associated with AHR	Number of matches	Expected	z-score
AHR_EGRF	EGR/nerve growth factor-related factors	308	69.96	28.4
AHR_AP1R	MAF- and AP1-related factors(Antioxidant response element)	317	74.34	28.09
AHR_AP1F	Activator protein 1	153	22.63	27.3
AHR_AP2F	Activator protein 2	196	38.8	25.16
AHR_HESF	Homologues of enhancer of split complex	198	45.79	22.42
AHR_KLFS	Krüppel like transcription factors	381	127.96	22.33
AHR_ETSF	Human and murine ETS1 factors	356	116.31	22.18
AHR_P53F	p53 tumor suppressor	181	40.75	21.89
AHR_ZF02	C2H2 zinc-finger transcription factors 2	271	77.83	21.84
AHR_FKHD	Forkhead transcription factors	344	112.6	21.76

Multiple AHR/ARNT cobound regions were identified and confirmed in the intragenic regions of *insulin-like growth factor 1 receptor (IGF1R)*. *IGF1R* is an estrogen-responsive gene that was positively regulated after 24 h of 0.1nM TCDD exposure *in vivo* or *in vitro* (Tanaka *et al.*, 2007). The AHR/ARNT

cobound regions were AHR₁, AHR₁₀₆, AHR₂₃₀, AHR₄₀₀, AHR₇₁₁, and AHR₇₈₆. Each of the regions contained at least one AHRE core sequence, with the exception of AHR₇₈₆.

Independent ChIP assays confirmed TCDD-dependent recruitment of AHR and ARNT to all regions (Fig. 1B). We

did not observe any false positives during our qPCR validation of binding sites, consistent with our high stringency cutoffs of FDR1 and FDR5 for AHR and ARNT, respectively. In summary, AHR-/ARNT-bound regions were distributed in intergenic and intragenic regions across large genomic regions. Some AHR-/ARNT-bound regions were associated with a putative AHRE cluster, whereas others were not. These data supported a strong, but not an absolute relationship between AHR/ARNT binding and the presence of an AHRE cluster. They also revealed that AHR/ARNT bound to very small proportion of the computationally predicted AHRE across the genome.

Integration of AHR/ARNT Binding Events With Changes in Gene Expression

We next performed gene expression microarray experiments in MCF-7 exposed to 10nM TCDD for 6h. The resulting data set was then compared with genes associated with AHR- or ARNT-bound regions. We identified 104 unique differentially regulated genes at cutoff of $|fold\ change\ (FC)| \geq 1.2$ and FDR5 (the complete list is provided in [Supplementary table S8](#)). The primary effect of TCDD on MCF-7 cells was gene activation because the vast majority (98 or 94.2%) of the regulated genes were significantly induced at 6h. Of the 104 unique differentially regulated genes, 69 (66.3%) were associated with an AHR- or ARNT-bound region within 100kb of their corresponding TSS ([Fig. 4A](#)). The distribution of AHR- or ARNT-bound regions within 100kb of a TCDD-responsive gene was centered at 2102 bp downstream from TSS ([Fig. 4B](#)). Peak enrichment was located in the proximal promoter area (within 10kb) although a number of binding sites greater than 10kb from the TSS of TCDD responsive genes were also present. A list of the 69 TCDD-responsive and AHR/ARNT-bound genes is provided in [Table 4](#). Of the 69 genes, 59 (85.6%) contained at least one AHRE, representing TCDD-induced, AHR-mediated, and AHRE-dependent gene regulation.

Gene Function and Pathway Analysis

Function and pathway analysis was conducted on 69 TCDD-induced AHR or ARNT-bound genes to evaluate the direct impact of AHR activation at the cellular level. Pathway analysis of those 69 bound and regulated genes demonstrated significant changes in the AHR signaling pathway, xenobiotics metabolism by CYP450, xenobiotic metabolism signaling, fatty acid metabolism, and tryptophan metabolism. All of which correlated with the perturbation of cellular functions such as drug metabolism, small molecule biochemistry, nucleic acid metabolism, and lipid metabolism ([Table 5](#)). This was consistent with our previous genome-wide studies on mouse hepatic tissue ([Dere et al., 2011; Lo et al., 2011](#)) and with the role of AHR as a regulator of xenobiotic metabolism and lipid homeostasis. AHR-ARNT was among the top predicted transcription factors along with AIP, homeobox gene GSX2, and breast cancer 1, early onset (BRCA1).

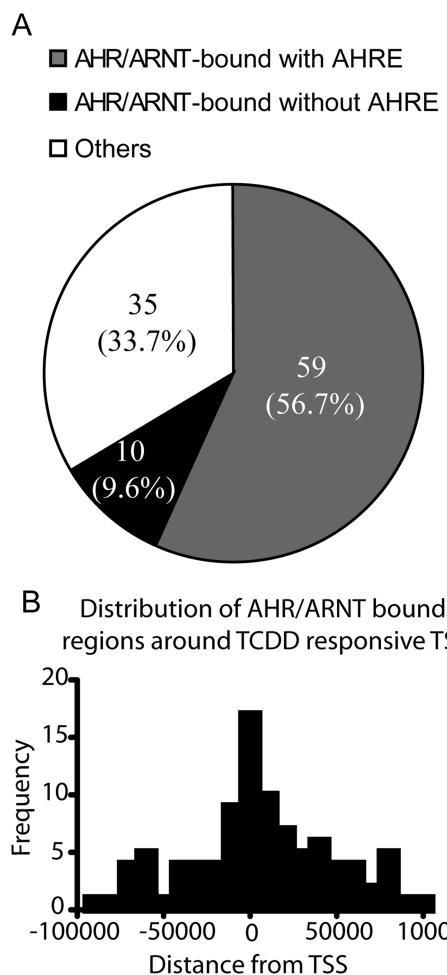


FIG. 4. (A) Percentage of TCDD-responsive genes that were associated with an AHR- or ARNT-bound region. Approximately 66.3% (69/104) of TCDD-responsive genes were associated with at least one AHR- or ARNT-bound region and 85.6% (59/69) of these regions contained at least one AHRE. (B) Distribution of AHR/ARNT-bound regions around TCDD-responsive TSS. Peak enrichment is located at the proximal promoter although there are many AHR/ARNT-bound regions greater than 10kb from TCDD-responsive TSS as well.

DISCUSSION

Our laboratory and others have described ligand-activated genomic binding of AHR using promoter-focused and genomic-wide microarrays ([Ahmed et al., 2009; De Abrew et al., 2010; Dere et al., 2011; Lo et al., 2011; Pansoy et al., 2010; Sartor et al., 2009](#)); however, similar studies have not been described for ARNT. We report here the genome-wide analysis of AHR and ARNT-binding sites in MCF-7 breast cancer cells. Our results reveal a high, albeit lower than expected overlap between AHR- and ARNT-bound regions. The AHRE sequence is over-represented in AHR-/ARNT-bound regions, but only 60% contain at least one AHRE core sequence. These findings support previous studies demonstrating that AHR/ARNT heterodimer is promiscuous in its binding to DNA

TABLE 4
Differentially Regulated Genes With an AHR- or ARNT-Bound Region Within 100 kb

Gene symbol		Genomic coordinate	Genomic location	Fold change	AHRE?
AHRR	chr5	431150	431649	Intron	3.70
	chr5	386600	387349	Intron	n
ALDH1A3	chr15	101389650	101390099	TSS_upstream	4.81
ALS2CL	chr3	46700600	46701649	TES_downstream	y
ATXN1	chr6	16766100	16766899	TSS_upstream	y
BATF	chr14	76032250	76033199	TES_downstream	y
	chr14	75919800	75920299	TSS_upstream	y
BCL3	chr19	45245100	45245699	TSS_upstream	y
BMF	chr15	40390900	40391549	Intron	2.49
	chr15	40385000	40385899	Intron	y
	chr15	40377300	40377849	TES_downstream	y
CACNA1D	chr3	53552600	53553899	Intron	1.54
CRISPLD2	chr16	84918050	84918599	Intron	y
CYP1A1	chr15	75055450	75056549	TSS_upstream	27.51
	chr15	75017450	75019499	TSS_upstream	y
CYP1A2	chr15	75055450	75056549	TES_downstream	y
	chr15	75017450	75019499	TSS_upstream	y
CYP1B1	chr2	38322850	38324249	TSS_upstream	4.86
	chr2	38303850	38305099	TSS_upstream	y
	chr2	38325700	38326549	TSS_upstream	y
DDIT4	chr10	74032050	74033149	TSS_upstream	1.81
DDX24	chr14	94425750	94426149	TES_downstream	1.19
DHX37	chr12	125405550	125406349	TES_downstream	y
DLL1	chr6	170501650	170502349	TES_downstream	1.52
	chr6	170698700	170699499	TSS_upstream	y
	chr6	170589100	170589699	TES_downstream	n
DLX1	chr2	172903100	172904099	TSS_upstream	1.90
DLX2	chr2	172903100	172904099	TES_downstream	y
DNMBP	chr10	101765300	101766449	5'UTR	y
	chr10	101767750	101768349	5'UTR	n
DPP9	chr19	4706600	4707599	Intron	1.32
EDC3	chr15	74934250	74936149	Intron	2.31
	chr15	75017450	75019499	TSS_upstream	y
ELF4	chrX	129243550	129244099	5'UTR	y
	chrX	129169200	129170049	TES_downstream	y
FAM32A	chr19	16283600	16284199	TSS_upstream	1.24
FAM65c	chr20	49294550	49295499	TSS_upstream	1.90
	chr20	49261950	49262999	TSS_upstream	y
FLVCR2	chr14	76032250	76033199	TSS_upstream	1.48
FOSL2	chr2	28617100	28618049	Intron	1.40
FREM2	chr13	39260100	39260999	TSS_upstream	y
IGF1R	chr15	99439950	99440999	Intron	1.25
	chr15	99442200	99443649	Intron	y
	chr15	99294900	99295949	Intron	y
	chr15	99397900	99399049	Intron	y
	chr15	99395750	99397549	Intron	y
	chr15	99419950	99420649	Intron	n
LAMA3	chr18	21209300	21209849	TSS_upstream	1.61
LMCD1	chr3	8497750	8499299	TSS_upstream	y
	chr3	8541950	8542799	TSS_upstream	y
MAPRE2	chr18	32559100	32559499	TSS_upstream	1.84
	chr18	32629450	32630099	5'UTR	y
MCOLN2	chr1	85422150	85423049	Intron	2.30
	chr1	85426000	85426599	Intron	y
NEDD9	chr6	11305000	11305649	TSS_upstream	n
NFE2L2	chr2	178127800	178128449	5'UTR	1.39
NIN	chr14	51377100	51378099	TSS_upstream	y
	chr14	51331750	51332699	TSS_upstream	y
	chr14	51370350	51371599	TSS_upstream	y
NPY1R	chr4	164225300	164225749	TES_downstream	-1.58
PITPNM2	chr12	123579100	123580599	5'UTR	1.60
	chr12	123560100	123560949	5'UTR	y
	chr12	123547100	123547699	5'UTR	n

TABLE 4—Continued

Gene symbol		Genomic coordinate	Genomic location	Fold change	AHRE?	
PLEC1	chr8	145040700	145041299	TSS_upstream	1.39	y
PRPS1	chrX	106920100	106920649	TES_downstream	1.89	y
	chrX	106869600	106870499	TSS_upstream		y
PSG5	chr19	43754550	43755099	TSS_upstream	-1.66	n
	chr19	43769050	43770049	TSS_upstream		n
PSPC1	chr13	20368400	20369399	TSS_upstream	1.34	y
PYGL	chr14	51377100	51378099	Intron	2.12	y
	chr14	51331750	51332699	TES_downstream		y
	chr14	51370350	51371599	TES_downstream		y
RND1	chr12	49275750	49276699	TSS_upstream	2.41	y
RRP12	chr10	99078600	99078999	TES_downstream	1.49	y
RUNX1	chr21	36250600	36251299	Intron	1.78	n
	chr21	36260050	36260949	5'UTR		y
RUNX2	chr6	45413500	45413949	Intron	2.11	n
SAMD12	chr8	119451700	119452449	Exon	1.95	n
SAT1	chrX	23815400	23815999	TES_downstream	1.44	n
SECTM1	chr17	80255800	80256299	TES_downstream	3.15	y
SIPA1L2	chr1	232736200	232736699	TSS_upstream	2.02	y
SLC27A2	chr15	50507600	50508449	Intron	1.39	y
SLC2A11	chr22	24198650	24199249	TSS_upstream	1.79	y
SLC3A2	chr11	62689250	62689649	TES_downstream	1.76	y
SLC7A5	chr16	87822300	87824849	TES_downstream	3.37	y
	chr16	87840300	87841199	TES_downstream		y
STRBP	chr9	126089750	126090599	TSS_upstream	1.31	y
	chr9	126097700	126098399	TSS_upstream		n
TFAP2A	chr6	10403200	10404399	Intron	1.60	y
TIPARP	chr3	156469350	156469899	TES_downstream	4.02	y
TMEM45B	chr11	129723600	129724899	Intron	1.55	y
	chr11	129716150	129717099	5'UTR		n
TMTC2	chr12	83079650	83080499	TSS_upstream	1.93	y
TPCN1	chr12	113683400	113684149	Intron	1.70	y
TRAFD1	chr12	112574650	112575299	Intron	1.25	y
TRUB2	chr9	131084700	131085199	TSS_upstream	1.19	y
TXNRD1	chr12	112555300	112555949	TSS_upstream	1.62	y
	chr12	104613750	104614499	Intron		y
	chr12	104571150	104572149	TSS_upstream		y
UBE2G2	chr21	46172650	46173399	TES_downstream	1.32	y
USP3	chr15	63788950	63790149	TSS_upstream	1.33	y
VDR	chr12	48297800	48298749	5'UTR	1.78	y
VIPR1	chr3	42538300	42538749	TSS_upstream	2.40	y
VTCN1	chr1	117811450	117812049	TSS_upstream	1.94	n
WDR63	chr1	85679900	85680799	TES_downstream	3.36	y

and not limited to AHREs (Ahmed *et al.*, 2009; Dere *et al.*, 2011; Lo *et al.*, 2011; Pansoy *et al.*, 2010). However, across the entire genome only a small proportion of these computationally predicted AHREs are bound by AHR/ARNT. Our analysis also identifies novel AHR-regulated bound regions in the regulatory regions of AHR target genes, including *AHRR* and *NRF2*. By integrating our ChIP-Seq data with gene expression microarray data, we provide a list of AHR-/ARNT-bound and TCDD-responsive genes with significant changes occurring in genes involved in lipid metabolism and drug metabolism.

Schödel *et al.* (2011) recently reported genome-wide mapping of HIF-1 α and ARNT (referred to as HIF-1 β in their study), revealing a ~89% overlap (356/400) between HIF-1 α and ARNT-binding sites using a different ARNT antibody, NB100–110 from Novus Biologicals. The degree of overlap between AHR and ARNT in our study is lower (65%; 882/1352) and

might be due to differences in antibody affinity as the average peak height for the ARNT data set is significantly lower than that for AHR. Differences in anti-ARNT antibody used might also account for this discrepancy. We are testing this hypothesis by repeating our ChIP-Seq studies with anti-ARNT antibody from Novus Biologicals (NB100–110). Comparison between the HIF-1 α -ARNT binding sites reported by Schödel *et al.* (2011) in hypoxic cells and the AHR-/ARNT-bound regions in our study reveals minimal overlap (data not shown), suggesting that the distribution of ARNT-binding sites is heavily determined by its heterodimerization partner. These findings are also consistent with the fact that AHR and HIF-1 α belong to two different molecular pathways with different sets of regulated genes and binding targets.

The lack of complete overlap between AHR-bound and ARNT-bound regions also suggests that both factors might

TABLE 5

Enrichment of Canonical Pathways, Functional Pathways, and Predicted Transcription Factors for the 69 Regulated and AHR-/ARNT-Bound Genes Using IPA

	<i>p</i> value
Top 5 canonical pathways	
Aryl hydrocarbon receptor signaling	1.2×10^{-5}
Fatty acid metabolism	8.86×10^{-5}
Xenobiotic metabolism signaling	6.35×10^{-4}
Metabolism of xenobiotics by CYP450	8.73×10^{-4}
Tryptophan metabolism	1.42×10^{-3}
Top 5 functions	
Gene expression	2.60×10^{-8}
Drug metabolism	5.00×10^{-8}
Small molecule biochemistry	5.00×10^{-8}
Lipid metabolism	5.23×10^{-7}
Nucleic acid metabolism	5.23×10^{-7}
Top 5 predicted transcription factors	
AHR-ARNT	5.59×10^{-10}
AHR-aryl hydrocarbon-ARNT	1.16×10^{-9}
AIP	1.18×10^{-6}
GSX2	3.29×10^{-5}
BRCA1	3.61×10^{-5}

interact independently with DNA. In support of this notion, AHR recruitment to a nonconsensus AHRE in the murine plasminogen activator inhibitor-1 (PAI-1) has been reported to be independent of ARNT (Huang and Elferink, 2012). PAI-1, however, is not associated with either AHR- or ARNT-bound regions in our present study, consistent with our previous ChIP-chip data in T-47D human breast cancer cells (Ahmed et al., 2009). This suggests that PAI-1 might not be regulated by AHR in human breast cancer cells, which is not surprising given the well-documented species differences in AHR signaling (Flavenvy et al., 2010). AHR- or ARNT-specific regions present in our data set may be influenced by the statistical cutoffs and the ability of the different antibodies to recognize their targets under our assay conditions, leading to potential false negatives. For example, in many cases AHR-bound regions at FDR1 not bound by ARNT at FDR5 are present in ARNT-bound regions at a more relaxed FDR10. We are currently using zinc-finger nuclease strategies to generate AHR and ARNT knockout cell lines to use as models to identify TCDD-induced and specific AHR- or ARNT-bound sequences and target genes.

In agreement with previous reports, our transcription factor binding site enrichment analysis reveals the overrepresentation of AHRE, ERE, SP1, and ARE motifs in AHR-/ARNT-bound regions (De Abrew et al., 2010; Dere et al., 2011; Lo et al., 2011; Sartor et al., 2009). Estrogen receptors and SP1 proteins have been previously reported to modulate AHR transactivation mechanisms (Ahmed et al., 2009; Kobayashi et al., 1996). NRF2, which regulates gene expression in response to oxidative stress through binding to AREs, is an AHR target gene but has also been shown to work in concert with AHR to regulate

the expression of numerous genes, including *NADPH quinone oxidoreductase 1* (*NQO1*) (Yeager et al., 2009). NRF2 knock-out prevents AHR-dependent regulation of *NQO1* (Yeager et al., 2009), suggesting that *NQO1* is directly regulated by NRF2, with its TCDD-dependent regulation likely a secondary effect of AHR-mediated increase in NRF2 expression (Yeager et al., 2009). We also report at least one AHR-bound region for NRF2 target genes such as *NQO1*, *GSTP1* and *GSTM1-5*, supporting a close functional relationship between AHR and NRF2 in the regulation of shared target genes. Although not studied in a recent ChIP-Seq analysis of NRF2 sites in lymphoblastoid cells, transcription factor binding site enrichment analysis of the NRF2 sites reported by Chorley et al. (2012) reveals that the AHRE is significantly overrepresented (data not shown). Taken together, these findings further support the functional interactions between the two transcription factors.

FKHD motifs are enriched adjacent to an AHRE, and a consensus FOXA1 binding motif is generated from our unsupervised *de novo* motif discovery using a set of AHR-/ARNT-bound regions. FOXA1 is the major determinant for ER binding where knockdown of FOXA1 significantly can reduce ER recruitment at most loci (Hurtado et al., 2011). We recently reported that FOXA1 is required for AHR-dependent regulation of cyclin G2, but not for CYP1A1, supporting a gene selective role for FOXA1 in AHR transactivation (Ahmed et al., 2012). The impact of FOXA1 on AHR transactivation at the genome-wide level has not been evaluated.

Our *de novo* motif analysis identifies a core AHRE, but not the extended AHRE or the AHRE-II. This implies that although AHR interaction with the other two variations of the AHREs is possible under cell type-specific or gene-specific context, the core AHRE remains one of the best predictor for AHR interactions with chromatin. The symmetrical AHRE (5'GTGCGTG3') we describe is identical to the SIM/ARNT motif and similar to the second ranking AHR/ARNT motif reported by Swanson et al. (1995). The extended AHRE (5'TNGCGTG3') is responsible for robust AHR/ARNT recruitment to the AHRE clusters for *CYP1A1*. However, the more flexible symmetrical AHRE might be a better predictor of AHR/ARNT binding sites throughout the genome.

We report 104 TCDD-responsive genes using microarray gene expression analysis with ~95% of the regulated genes being upregulated with only 66.3% associated with an AHR- or ARNT-bound region. Furthermore, AHR-/ARNT-bound regions identified by ChIP-Seq vastly outnumber TCDD-responsive genes in our expression microarray. These discrepancies between binding events and gene regulation may be due to temporal differences in mRNA expression (Boverhof et al., 2005), nongenomic effect of AHR activation, or the result of cell type-specific expression patterns through epigenetic modulation that are independent of AHR/ARNT occupancy. It also remains possible that the distal AHR-/ARNT-bound regions might be involved in the regulation of nonprotein coding transcripts, such as nonprotein coding RNA, small nucleolar RNAs, and

microRNA, that would have been missed by our microarray studies. Future experiments using RNA sequencing will provide valuable information on the bridging of multiple DNA elements with both coding and noncoding transcriptome.

AHR/ARNT binds preferentially to proximal promoter regions as peak density was the highest in regions within 1 kb from transcription start sites, indicating that many AHR target genes are directly regulated by AHR binding to the proximal promoter. On the other hand, there are also many AHR/ARNT cobound regions located distal (~100 kb) from annotated transcription start sites (Table 1; Figs. 2C and D). Although more data are needed, our findings suggest that AHR-mediated gene regulation might involve long-range interaction of distal *cis*-regulatory regions in a mechanism similar to that of nuclear receptors. This is supported by previous studies showing that nuclear receptor binding sites are also distal to regulated genes both *in vivo* upon acute ligand stimulation (Boergesen *et al.*, 2012; Hewitt *et al.*, 2012) and *in vitro* (Carroll *et al.*, 2006; Reddy *et al.*, 2009). Gene regulation by remote *cis*-acting regions has been experimentally shown to occur through chromatin remodeling, DNA looping (Fullwood *et al.*, 2009), or even inter- and intrachromosomal interaction (Hu *et al.*, 2008), making it challenging to correlate DNA interaction with downstream transcriptional outcomes.

Overall, our data are in agreement with previous studies supporting the important role of the AHRE in AHR transactivation. However, the lack of a core AHRE in all of the AHR-bound regions suggests that AHR exhibits more flexible DNA-binding preferences. This is supported by the less stringent symmetrical AHRE motif predicted from our study and the identification of cooperative transcription factors that might mediate indirect AHR-DNA interaction. Our results provide a comprehensive genome-wide AHR/ARNT binding site analysis and a robust data set that can be used to improve predictions for functional AHR/ARNT DNA interaction.

SUPPLEMENTARY DATA

Supplementary data are available online at <http://toxsci.oxfordjournals.org/>.

FUNDING

R.L. is the recipient of the Canadian Breast Cancer Foundation Doctoral Fellowship. J.M. is the recipient of the Canadian Institute of Health Research New Investigator Award. Canadian Institute of Health Research (MOP-82715); Canadian Breast Cancer Foundation—Ontario to J.M.

ACKNOWLEDGMENTS

We thank Sigma Aldrich for providing the SeqPlex amplification kits used in this study.

REFERENCES

- Ahmed, S., Al-Saigh, S., and Matthews, J. (2012). FOXA1 is essential for aryl hydrocarbon receptor-dependent regulation of cyclin G2. *Mol. Cancer Res.* **10**, 636–648.
- Ahmed, S., Valen, E., Sandelin, A., and Matthews, J. (2009). Dioxin increases the interaction between aryl hydrocarbon receptor and estrogen receptor alpha at human promoters. *Toxicol. Sci.* **111**, 254–266.
- Baba, T., Mimura, J., Gradin, K., Kuroiwa, A., Watanabe, T., Matsuda, Y., Inazawa, J., Sogawa, K., and Fujii-Kuriyama, Y. (2001). Structure and expression of the Ah receptor repressor gene. *J. Biol. Chem.* **276**, 33101–33110.
- Bock, K. W., and Köhle, C. (2006). Ah receptor: Dioxin-mediated toxic responses as hints to deregulated physiologic functions. *Biochem. Pharmacol.* **72**, 393–404.
- Boergesen, M., Pedersen, T. Å., Gross, B., van Heeringen, S. J., Hagenbeek, D., Bindesbøll, C., Caron, S., Laloyer, F., Steffensen, K. R., Nebb, H. I., *et al.* (2012). Genome-wide profiling of liver X receptor, retinoid X receptor, and peroxisome proliferator-activated receptor α in mouse liver reveals extensive sharing of binding sites. *Mol. Cell. Biol.* **32**, 852–867.
- Boutros, P. C., Moffat, I. D., Franc, M. A., Tijet, N., Tuomisto, J., Pohjanvirta, R., and Okey, A. B. (2004). Dioxin-responsive AHRE-II gene battery: Identification by phylogenetic footprinting. *Biochem. Biophys. Res. Commun.* **321**, 707–715.
- Boverhof, D. R., Burgoon, L. D., Tashiro, C., Chittim, B., Harkema, J. R., Jump, D. B., and Zacharewski, T. R. (2005). Temporal and dose-dependent hepatic gene expression patterns in mice provide new insights into TCDD-Mediated hepatotoxicity. *Toxicol. Sci.* **85**, 1048–1063.
- Carroll, J. S., Meyer, C. A., Song, J., Li, W., Geistlinger, T. R., Eeckhoute, J., Brodsky, A. S., Keeton, E. K., Fertuck, K. C., Hall, G. F., *et al.* (2006). Genome-wide analysis of estrogen receptor binding sites. *Nat. Genet.* **38**, 1289–1297.
- Chorley, B. N., Campbell, M. R., Wang, X., Karaca, M., Sambandan, D., Bangura, F., Xue, P., Pi, J., Kleeberger, S. R., and Bell, D. A. (2012). Identification of novel NRF2-regulated genes by ChIP-Seq: Influence on retinoid X receptor alpha. *Nucleic Acids Res.* **40**, 7416–7429.
- De Abrew, K. N., Kaminski, N. E., and Thomas, R. S. (2010). An integrated genomic analysis of aryl hydrocarbon receptor-mediated inhibition of B-cell differentiation. *Toxicol. Sci.* **118**, 454–469.
- Deng, W., Nickle, D. C., Learn, G. H., Maust, B., and Mullins, J. I. (2007). ViroBLAST: A stand-alone BLAST web server for flexible queries of multiple databases and user's datasets. *Bioinformatics* **23**, 2334–2336.
- Denison, M. S., and Nagy, S. R. (2003). Activation of the aryl hydrocarbon receptor by structurally diverse exogenous and endogenous chemicals. *Annu. Rev. Pharmacol. Toxicol.* **43**, 309–334.
- Dere, E., Lo, R., Celius, T., Matthews, J., and Zacharewski, T. R. (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.
- Flavenvy, C. A., Murray, I. A., and Perdew, G. H. (2010). Differential gene regulation by the human and mouse aryl hydrocarbon receptor. *Toxicol. Sci.* **114**, 217–225.
- Fullwood, M. J., Liu, M. H., Pan, Y. F., Liu, J., Xu, H., Mohamed, Y. B., Orlov, Y. L., Velkov, S., Ho, A., Mei, P. H., *et al.* (2009). An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature* **462**, 58–64.
- Haarmann-Stemmann, T., Bothe, H., Kohli, A., Sydlik, U., Abel, J., and Fritzsche, E. (2007). Analysis of the transcriptional regulation and molecular function of the aryl hydrocarbon receptor repressor in human cell lines. *Drug Metab. Dispos.* **35**, 2262–2269.
- Hankinson, O. (1995). The aryl hydrocarbon receptor complex. *Annu. Rev. Pharmacol. Toxicol.* **35**, 307–340.

- Hewitt, S. C., Li, L., Grimm, S. A., Chen, Y., Liu, L., Li, Y., Bushel, P. R., Fargo, D., and Korach, K. S. (2012). Research resource: Whole-genome estrogen receptor α binding in mouse uterine tissue revealed by ChIP-seq. *Mol. Endocrinol.* **26**, 887–898.
- Hu, Q., Kwon, Y. S., Nunez, E., Cardamone, M. D., Hutt, K. R., Ohgi, K. A., Garcia-Bassets, I., Rose, D. W., Glass, C. K., Rosenfeld, M. G., et al. (2008). Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 19199–19204.
- Huang, G., and Elferink, C. J. (2012). A novel nonconsensus xenobiotic response element capable of mediating aryl hydrocarbon receptor-dependent gene expression. *Mol. Pharmacol.* **81**, 338–347.
- Hurtado, A., Holmes, K. A., Ross-Innes, C. S., Schmidt, D., and Carroll, J. S. (2011). FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat. Genet.* **43**, 27–33.
- Ji, H., Jiang, H., Ma, W., Johnson, D. S., Myers, R. M., and Wong, W. H. (2008). An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nat. Biotechnol.* **26**, 1293–1300.
- Jin, V. X., Apostolos, J., Nagisetty, N. S., and Farnham, P. J. (2009). W-ChIPMotifs: A web application tool for de novo motif discovery from ChIP-based high-throughput data. *Bioinformatics* **25**, 3191–3193.
- Kobayashi, A., Sogawa, K., and Fujii-Kuriyama, Y. (1996). Cooperative interaction between AhR.Arnt and Sp1 for the drug-inducible expression of CYP1A1 gene. *J. Biol. Chem.* **271**, 12310–12316.
- Li, R., Yu, C., Li, Y., Lam, T. W., Yiu, S. M., Kristiansen, K., and Wang, J. (2009). SOAP2: An improved ultrafast tool for short read alignment. *Bioinformatics* **25**, 1966–1967.
- Lo, R., Celius, T., Forgaes, A. L., Dere, E., MacPherson, L., Harper, P., Zacharewski, T., and Matthews, J. (2011). Identification of aryl hydrocarbon receptor binding targets in mouse hepatic tissue treated with 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Toxicol. Appl. Pharmacol.* **257**, 38–47.
- Macpherson, L., and Matthews, J. (2010). Inhibition of aryl hydrocarbon receptor-dependent transcription by resveratrol or kaempferol is independent of estrogen receptor α expression in human breast cancer cells. *Cancer Lett.* **299**, 119–129.
- Mahony, S., and Benos, P. V. (2007). STAMP: A web tool for exploring DNA-binding motif similarities. *Nucleic Acids Res.* **35**, W253–W258.
- Matthews, J., Wihlén, B., Heldring, N., MacPherson, L., Helguero, L., Treuter, E., Haldosén, L. A., and Gustafsson, J. A. (2007). Co-planar 3,3',4,4',5-pentachlorinated biphenyl and non-co-planar 2,2',4,6,6'-pentachlorinated biphenyl differentially induce recruitment of oestrogen receptor alpha to aryl hydrocarbon receptor target genes. *Biochem. J.* **406**, 343–353.
- Matthews, J., Wihlén, B., Thomsen, J., and Gustafsson, J. A. (2005). Aryl hydrocarbon receptor-mediated transcription: Ligand-dependent recruitment of estrogen receptor alpha to 2,3,7,8-tetrachlorodibenzo-p-dioxin-responsive promoters. *Mol. Cell. Biol.* **25**, 5317–5328.
- Matys, V., Fricke, E., Geffers, R., Gössling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A. E., Kel-Margoulis, O. V., et al. (2003). TRANSFAC: Transcriptional regulation, from patterns to profiles. *Nucleic Acids Res.* **31**, 374–378.
- McIntosh, B. E., Hogenesch, J. B., and Bradfield, C. A. (2010). Mammalian Per-Arnt-Sim proteins in environmental adaptation. *Annu. Rev. Physiol.* **72**, 625–645.
- Miao, W., Hu, L., Scrivens, P. J., and Batist, G. (2005). Transcriptional regulation of NF-E2 p45-related factor (NRF2) expression by the aryl hydrocarbon receptor-xenobiotic response element signaling pathway: Direct cross-talk between phase I and II drug-metabolizing enzymes. *J. Biol. Chem.* **280**, 20340–20348.
- Okino, S. T., Quattrochi, L. C., Pookot, D., Iwahashi, M., and Dahiya, R. (2007). A dioxin-responsive enhancer 3' of the human CYP1A2 gene. *Mol. Pharmacol.* **72**, 1457–1465.
- Pansoy, A., Ahmed, S., Valen, E., Sandelin, A., and Matthews, J. (2010). 3-Methylcholanthrene induces differential recruitment of aryl hydrocarbon receptor to human promoters. *Toxicol. Sci.* **117**, 90–100.
- Pot, C. (2012). Aryl hydrocarbon receptor controls regulatory CD4+ T cell function. *Swiss Med. Wkly.* **142**, w13592.
- Reddy, T. E., Pauli, F., Sprouse, R. O., Neff, N. F., Newberry, K. M., Garabedian, M. J., and Myers, R. M. (2009). Genomic determination of the glucocorticoid response reveals unexpected mechanisms of gene regulation. *Genome Res.* **19**, 2163–2171.
- Safe, S., and Wormke, M. (2003). Inhibitory aryl hydrocarbon receptor-estrogen receptor alpha cross-talk and mechanisms of action. *Chem. Res. Toxicol.* **16**, 807–816.
- Sartor, M. A., Schnekenburger, M., Marlowe, J. L., Reichard, J. F., Wang, Y., Fan, Y., Ma, C., Karyala, S., Halbleib, D., Liu, X., et al. (2009). Genomewide analysis of aryl hydrocarbon receptor binding targets reveals an extensive array of gene clusters that control morphogenetic and developmental programs. *Environ. Health Perspect.* **117**, 1139–1146.
- Schödel, J., Oikonomopoulos, S., Ragoussis, J., Pugh, C. W., Ratcliffe, P. J., and Mole, D. R. (2011). High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq. *Blood* **117**, e207–e217.
- Shen, E. S., and Whitlock, J. P., Jr. (1992). Protein-DNA interactions at a dioxin-responsive enhancer. Mutational analysis of the DNA-binding site for the liganded Ah receptor. *J. Biol. Chem.* **267**, 6815–6819.
- Sogawa, K., Numayama-Tsuruta, K., Takahashi, T., Matsushita, N., Miura, C., Nikawa, J., Gotoh, O., Kikuchi, Y., and Fujii-Kuriyama, Y. (2004). A novel induction mechanism of the rat CYP1A2 gene mediated by Ah receptor-Arnt heterodimer. *Biochem. Biophys. Res. Commun.* **318**, 746–755.
- Swanson, H. I., Chan, W. K., and Bradfield, C. A. (1995). DNA binding specificities and pairing rules of the Ah receptor, ARNT, and SIM proteins. *J. Biol. Chem.* **270**, 26292–26302.
- Tanaka, J., Yonemoto, J., Zaha, H., Kiyama, R., and Sone, H. (2007). Estrogen-responsive genes newly found to be modified by TCDD exposure in human cell lines and mouse systems. *Mol. Cell. Endocrinol.* **272**, 38–49.
- Tomita, S., Sinal, C. J., Yim, S. H., and Gonzalez, F. J. (2000). Conditional disruption of the aryl hydrocarbon receptor nuclear translocator (Arnt) gene leads to loss of target gene induction by the aryl hydrocarbon receptor and hypoxia-inducible factor 1alpha. *Mol. Endocrinol.* **14**, 1674–1681.
- Ueda, R., Iketaki, H., Nagata, K., Kimura, S., Gonzalez, F. J., Kusano, K., Yoshimura, T., and Yamazoe, Y. (2006). A common regulatory region functions bidirectionally in transcriptional activation of the human CYP1A1 and CYP1A2 genes. *Mol. Pharmacol.* **69**, 1924–1930.
- Wasserman, W. W., and Sandelin, A. (2004). Applied bioinformatics for the identification of regulatory elements. *Nat. Rev. Genet.* **5**, 276–287.
- Whitlock, J. P. Jr. (1999). Induction of cytochrome P4501A1. *Annu. Rev. Pharmacol. Toxicol.* **39**, 103–125.
- Wold, B., and Myers, R. M. (2008). Sequence census methods for functional genomics. *Nat. Methods* **5**, 19–21.
- Yeager, R. L., Reisman, S. A., Aleksunes, L. M., and Klaassen, C. D. (2009). Introducing the “TCDD-inducible AhR-Nrf2 gene battery”. *Toxicol. Sci.* **111**, 238–246.
- Zhang, L., Savas, U., Alexander, D. L., and Jefcoate, C. R. (1998). Characterization of the mouse Cyp1B1 gene. Identification of an enhancer region that directs aryl hydrocarbon receptor-mediated constitutive and induced expression. *J. Biol. Chem.* **273**, 5174–5183.