THE HUMAN AHR: IDENTIFICATION OF SINGLE NUCLEOTIDE POLYMORPHISMS FROM SIX ETHNIC POPULATIONS

J. Craig Rowlands[1], Robert Budinsky[1], Bhaskar Gollapudi[1], Kenneth Drinkwater[2], Michael Storck[2]

[1]The Dow Chemical Company, Midland, MI; [2]Functional Biosciences, Inc., Madison, WI.

Abstract

Central to the mode of action (MOA) for 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) and related dioxin-like chemicals (DLCs) is binding to and activation of the cytosolic aryl hydrocarbon receptor (AHR). Although the functions of the various domains of the AHR are generally conserved across mammalian species, it is now evident that numerous molecular differences in receptors exist that affect receptor activity and sensitivity to TCDD and DLC's. For example, the ten-fold strain differences in sensitivity to TCDD between C57BL/6 sensitive mice and DBA2 insensitive mice is the result of a single amino acid polymorphism in the ligand-binding domain of the receptor that produces AHRs with high- and low-affinity for TCDD, respectively. Alternatively, differences in dioxin sensitive Long Evans and the TCDD-resistant Hans/Wistar rats does not correspond to differences in AHR ligand binding, but rather relates to differences in the trans-activation domain of their respective AHRs. The human AHR is also a low affinity receptor and reports indicate that over 100 individual AHR genes have been sequenced for polymorphism identification. These prior efforts have identified at least five single nucleotide polymorphisms (SNPs) in the human AHR coding sequence with no SNPs in the ligand-binding domain. The majority of these human AHR SNPs reside in the trans-activation domain, and one is in the DNA binding basic region. We sequenced up to 108 additional human AHR genes in an effort to further identify single SNPs within the open reading frames (ORFs) of the *AHR* locus. The DNA was sequenced from six ethnic populations that included Japanese, Chinese, European/Caucasian, African American, South East Asian and Hispanic. Eight exonic SNPs were identified; six had been described previously in the literature or public databases and two appear to be novel. Five of the SNPs lead to amino acid changes in the AHR protein and three of the SNPs lead to synonymous substitutions. With these new sequences, over 200 human AHR gene sequences have been analyzed for SNPs. The results indicate that core ligand binding region of the human AHR is not polymorphic, but other regions such as the trans-activation domain are polymorphic in the human population and some of these forms should be examined for altered receptor signaling.

Introduction

The AH receptor (AHR) is a central mediator of the toxicology of halogenated-dioxins, -dibenzofurans and -biphenyls. Upon binding these toxic ligands, the AHR partners with a related protein, ARNT and this dimeric transcription factor regulates the expression of target genes involved in xenobiotic metabolism, as well as genes involved in initiating a number of toxic endpoints. In humans, the AH receptor is encoded by the *AHR* locus, a ~ 50 kbp gene of at least 11 exons found on the short arm of chromosome 7. In both mouse and rats and other species such as birds, this *ahr* locus is highly polymorphic, and this polymorphism has been shown to influence the responsiveness of given strains to the toxic effects of dioxins. In an effort to understand if the human *AHR* locus is equally polymorphic and if polymorphic forms of the *AHR* give rise to human populations with differing sensitivities to dioxins and related pollutants, we have begun a survey of human populations for novel *AHR* alleles.

Experimental Design, Materials and Methods

Single nucleotide polymorphism (SNP) discovery within exons of human *AHR* locus. Approximately 100 human genomic DNA samples from lymphoblast cell lines were obtained from the Coriell Cell Repositories (Camden, NJ). The origins of these DNA samples were from 26 Caucasians, 26 African Americans, 13 Japanese, 10 Mexicans, 7 Southeast Asians and 18 Chinese. Each *AHR* exon from each DNA sample was amplified using PCR. The PCR reactions were cleaned up using an exo/SAP and sequenced using Big Dye v.

3.1. The SNPs were detected by analyzing the traces with Mutation Surveyor v. 3.10. Results were compared with the SNPs documented in the NCBI's dbSNP database. http://www.ncbi.nlm.nih.gov/SNP/.

Results and Discussion

Our analysis identified 8 SNPs harbored within 11 exons of the *AHR* locus. These SNPs include those identified in the current study, as well as all SNPs documented by previous studies or found in the refSNP database. There were two previously documented SNPs that displayed no variation in the current study (i.e., Exon – 2 SNP2 and Exon 10 – SNP1). The majority of SNPs identified encoded non-synonymous base changes leading to predicted amino acid changes. Two of these SNPs do not appear to have been described previously (i.e., Exon 7 – SNP 1 and Exon 10 SNP – 4). Our data are consistent with a previous finding that the presence of Exon 10 - SNP 3 is linked to Exon 10 –SNP 2[3]. This finding is important because this double SNP has been suggested to lead to hypomorphic receptor activity[3]. Overall our data suggest that significant polymorphism exists at the human *AHR* locus and these differences may lead to altered receptor functional forms that could influence an individual's response to environmental dioxins and DLCs.

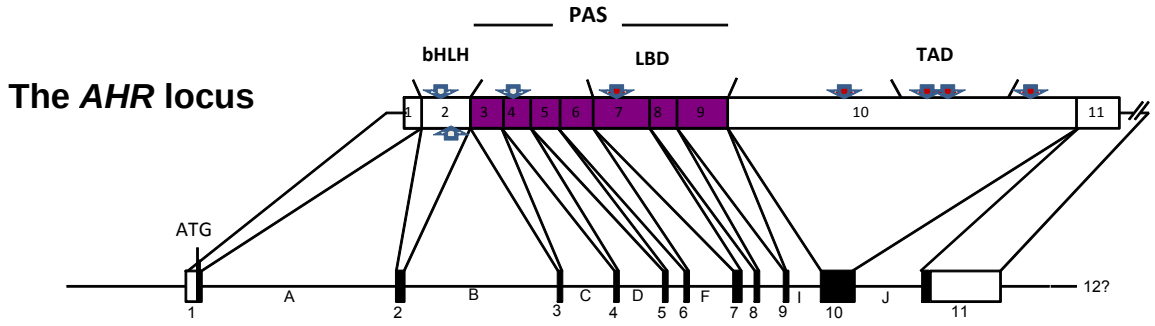

Figure 1: A representation of the human *AHR* locus as approximated from the human genome browser (http://genome.ucsc.edu). The exons are denoted as numbers, with open boxes representing untranslated sequence and darkened boxes representing translated sequence. Introns are denoted with letters. The deduced cDNA or mRNA is shown at the top. The domains are; basic-helix-loop helix (bHLH), ligand binding domain (LBD) and transcriptionally active domain (TAD). The PAS domain is shown colored and marked above. Domain structure is deduced from references 1 and 2. Red arrows are non-synonymous SNPS, open arrows are synonymous SNPS.

Exon 2 - SNP 1 : rs17779352 : N44N						
Samples		Genotype Frequency			Allele Frequency	
Group	People	T/T	T/C	C/C	T	C
Caucasian	26	0.923	0.077	-	0.962	0.038
African American	26	1.000	-	-	1.000	-
Japanese	13	0.923	0.077	-	0.962	0.038
Mexican	10	1.000	-	-	1.000	-
Southeast Asian	7	0.714	0.286	-	0.857	0.143
Chinese	18	0.944	0.056	-	0.972	0.028
Totals	100	0.940	0.060	-	0.970	0.030

Exon 2- SNP 2 : rs41273054 : L50L						
Samples		Genotype Frequency			Allele Frequency	
Group	People	G/G	G/A	A/A	G	A
Caucasian	26	1.000	-	-	1.000	-
African American	26	1.000	-	-	1.000	-
Japanese	13	1.000	-	-	1.000	-
Mexican	10	1.000	-	-	1.000	-
Southeast Asian	7	1.000	-	-	1.000	-
Chinese	18	1.000	-	-	1.000	-
Totals	100	1.000	-	-	1.000	-

Exon 4 – SNP 1 : rs35966236 : D132D						
Samples		Genotype Frequency			Allele Frequency	
Group	People	T/T	T/C	C/C	T	C
Caucasian	26	1.000	-	-	1.000	-
African American	26	1.000	-	-	1.000	-
Japanese	13	0.923	0.077	-	0.962	0.038
Mexican	10	1.000	-	-	1.000	-
Southeast Asian	7	1.000	-	-	1.000	-
Chinese	18	1.000	-	-	1.000	-
Totals	100	0.990	0.010	-	0.995	0.005

Exon 7 – SNP 1 : novel SNP : I277V						
Samples		Genotype Frequency			Allele Frequency	
Group	People	A/A	A/G	G/G	A	G
Caucasian	26	1.000	-	-	1.000	-
African American	26	1.000	-	-	1.000	-
Japanese	13	1.000	-	-	1.000	-
Mexican	10	1.000	-	-	1.000	-
Southeast Asian	7	1.000	-	-	1.000	-
Chinese	18	0.944	0.056	-	0.972	0.028
Totals	100	0.990	0.010	-	0.995	0.005

Exon 10 – SNP 1 : described in reference 3 : P517S						
Samples		Genotype Frequency			Allele Frequency	
Group	People	C/C	C/T	T/T	C	T
Caucasian	26	1.000	-	-	1.000	-
African American	34	1.000	-	-	1.000	-
Japanese	13	1.000	-	-	1.000	-
Mexican	10	1.000	-	-	1.000	-
Southeast Asian	7	1.000	-	-	1.000	-
Chinese	18	1.000	-	-	1.000	-
Totals	108	1.000	-	-	1.000	-

Exon 10 – SNP 2 : rs2066853 : R554K						
Samples		Genotype Frequency			Allele Frequency	
Group	People	G/G	G/A	A/A	G	A
Caucasian	26	0.885	0.115	-	0.942	0.058
African American	34	0.324	0.441	0.235	0.544	0.456
Japanese	13	0.308	0.538	0.154	0.577	0.423
Mexican	10	0.800	0.200	-	0.900	0.100
Southeast Asian	7	0.571	0.286	0.143	0.714	0.286
Chinese	18	0.566	0.389	0.056	0.750	0.250
Totals	108	0.556	0.333	0.111	0.722	0.278

Exon 10 – SNP 3 : rs4986826 : V570I						
Samples		Genotype Frequency			Allele Frequency	
Group	People	G/G	G/A	A/A	G	A
Caucasian	26	1.000	-	-	1.000	-
African American	34	0.824	0.147	0.029	0.897	0.103
Japanese	13	1.000	-	-	1.000	-
Mexican	10	1.000	-	-	1.000	-
Southeast Asian	7	1.000	-	-	1.000	-
Chinese	18	1.000	-	-	1.000	-
Totals	108	0.944	0.046	0.009	0.968	0.032

Exon 10 – SNP 4 : novel SNP : Q666K						
Samples		Genotype Frequency			Allele Frequency	
Group	People	C/C	C/A	A/A	C	A
Caucasian	26	1.000	-	-	1.000	-
African American	34	0.971	0.029	-	0.985	0.015
Japanese	13	1.000	-	-	1.000	-
Mexican	10	1.000	-	-	1.000	-
Southeast Asian	7	1.000	-	-	1.000	-
Chinese	18	1.000	-	-	1.000	-
Totals	108	0.991	0.009	-	0.995	0.005

References

1. Burbach, K. M., Swanson, H. I. and Bradfield, C. A. *Proc Nat Acad Sci (USA)* 1993; 90: 8566.
2. Jain, S., Dolwick, K. M. and Bradfield, C. A. *J Biol Chem* 1994; 269: 31518.
3. Wong, J. M. Y., Okey, A. B. and Harper, P. A. *Bioch Bioph Res Comm* 2001; 288: 990.
4. Moffat I. D., Roblin S., Harper P. A., Okey, A. B. & Pohjanvirta, R. AH *Mol Pharmacology* 2007; 72: 956.