

In Silico Identification of Dioxin Response Elements in Human, Mouse, and Rat Genomic Sequences

Sun, Y.V., Boverhof, D., Fielden, M., Zacharewski, T.R.

Department of Biochemistry and Molecular Biology, National Food Safety and Toxicology Center, and Institute for Environmental Toxicology, Michigan State University, East Lansing, MI, USA

Dioxin is a highly toxic compound produced by modern industry and causes a wide range of health problems. Most, if not all, of the biological effects elicited by dioxin and related compounds are mediated via ligand activation of the aryl hydrocarbon receptor (AhR), which forms a heterodimer with the AhR nuclear translocator (ARNT). This complex modulates gene expression through specific binding to dioxin response elements (DREs) in the regulatory regions of dioxin responsive genes. The purpose of this study is to identify conserved DREs using *in silico* approaches through the analysis of available human, mouse and rat genomic sequences corresponding to a nucleotide RefSeq accession (17,882, 11,697, and 3,896 promoters respectively) obtained from the UCSC Genome Browser (<http://genome.ucsc.edu/>). Thirteen *bona fide* DRE sequences, which included the 5bp core sequence (GCGTG) and the adjacent 7bp flanking sequence up- and downstream of the core sequence were used to establish a position weight matrix (PWM) with minimum similarity score of 0.8548 for the 19bp consensus sequence. Human, mouse and rat upstream regulatory regions and transcribed regions were scanned for DREs with matrix similarity scores greater than the 0.8548. DREs were disproportionately distributed in the upstream regulatory region close to transcription start site (TSS) in all three species. Putative functional response elements were identified by scanning regulatory elements -1500bp from the TSS for conserved DREs within orthologous human, mouse and rat genes identified using the HomoloGene database (build 31, 05/29/2003). Conserved DREs were identified in 365 out of 8872 orthologous human and mouse genes, 140 out of 3379 orthologous human and rat genes, 133 out of 2840 orthologous rat and mouse genes, and 48 out of 2437 orthologous genes across the three species. These findings corroborate many previously published reports, including the well characterized dioxin responsive genes *Cyp1a1* and *Cyp1b1*. Multiple sequence alignments identified six out of six conserved DREs in the *Cyp1a1* and three out of three conserved DREs in the *Cyp1b1* within the -1500bp region upstream from the TSS across the three genomes. For *Cyp1b1*, all three conserved DREs, which were the only three DREs, were suggested to be functional based on published thorough studies of human *CYP1B1* DREs. For *Cyp1a1* and *Cyp1b1*, the conserved DREs exhibit higher matrix similarity scores than their non-conserved counterparts. This *in silico* approach to identifying DREs in promoters sequences complements ongoing microarray studies of dioxin mediated alterations in gene expression as well as suggests that gene expression profiles across species may not be well conserved which could have significant implications in human risk assessment.

This work was supported by National Institutes of Health grant ES012245.