Toxicology and Applied Pharmacology 282 (2015) 30–41

Contents lists available at ScienceDirect

Toxicology and Applied Pharmacology

journal homepage: www.elsevier.com/locate/ytaap

# Analysis of the *AHR* gene proximal promoter GGGGC-repeat polymorphism in lung, breast, and colon cancer

Barbara C. Spink [a], Michael S. Bloom [b], Susan Wu [a], Stewart Sell [a,c], Erasmus Schneider [a,c], Xinxin Ding [a,b,c], David C. Spink [a,b,*]

[a] Wadsworth Center, New York State Department of Health, Albany, NY 12201, United States
[b] Department of Environmental Health Sciences, School of Public Health, University at Albany, State University of New York, Albany, NY 12201, United States
[c] Department of Biomedical Sciences, School of Public Health, University at Albany, State University of New York, Albany, NY 12201, United States

ARTICLE INFO

*Article history:*
Received 23 May 2014
Revised 26 September 2014
Accepted 27 October 2014
Available online 4 November 2014

*Keywords:*
Aryl hydrocarbon receptor
$(GGGGC)_n$ repeat polymorphism
Short tandem repeat
Lung cancer
Breast cancer
Colon cancer

ABSTRACT

The aryl hydrocarbon receptor (AhR) regulates expression of numerous genes, including those of the *CYP1* gene family. With the goal of determining factors that control *AHR* gene expression, our studies are focused on the role of the short tandem repeat polymorphism, $(GGGGC)_n$, located in the proximal promoter of the human *AHR* gene. When luciferase constructs containing varying GGGGC repeats were transfected into cancer cell lines derived from the lung, colon, and breast, the number of GGGGC repeats affected *AHR* promoter activity. The number of GGGGC repeats was determined in DNA from 327 humans and from 38 samples representing 5 species of non-human primates. In chimpanzees and 3 species of macaques, only $(GGGGC)_2$ alleles were observed; however, in western gorilla, $(GGGGC)_n$ alleles with $n = 2, 4, 5, 6, 7,$ and 8 were identified. In all human populations examined, the frequency of $(GGGGC)_n$ was $n = 4 > 5 \gg 2, 6$. When frequencies of the $(GGGGC)_n$ alleles in DNA from patients with lung, colon, or breast cancer were evaluated, the occurrence of $(GGGGC)_2$ was found to be 8-fold more frequent among lung cancer patients in comparison with its incidence in the general population, as represented by New York State neonates. Analysis of matched tumor and non-tumor DNA samples from the same individuals provided no evidence of microsatellite instability. These studies indicate that the $(GGGGC)_n$ short tandem repeats are inherited, and that the $(GGGGC)_2$ allele in the *AHR* proximal promoter region should be further investigated with regard to its potential association with lung cancer susceptibility.


## Introduction

Since its discovery as the receptor mediating the effects of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) and polycyclic aromatic hydrocarbons (PAHs) (Nebert and Gelboin, 1968; Poland et al., 1976; Conney, 1982, 2003; Okey, 2007; Xing et al., 2012), the aryl hydrocarbon receptor (AhR) has been identified as a key regulatory factor of carcinogenesis in the lung and other tissues. The AhR transcriptionally regulates the expression of the cytochrome P450 family 1 (CYP1) enzymes that metabolize PAHs, heterocyclic aromatic amines and other exogenous and endogenous compounds. The AhR-mediated induction of the CYP1 enzymes that catalyze the metabolic activation of PAHs to DNA adductive forms is integral to the carcinogenic process (Nebert et al., 2000, 2004), as is evidenced by the fact that AhR-null mice are refractory to PAH-induced carcinogenesis (Shimizu et al., 2000). However, induction of phase I enzymes by exposure to agents including TCDD can in some instances inhibit carcinogenicity (DiGiovanni et al., 1980). This presumably occurs by stimulating the metabolism of carcinogens through detoxification pathways to a greater extent than through toxification pathways (Conney, 1982).

Consistent with this longstanding hypothesis, Gelhaus et al. (2011) found that activation of AhR by pretreatment of H358 bronchial alveolar cells with TCDD caused a decrease in benzo[*a*]pyrene (BaP) adducts and an increase in BaP-glutathione adducts, the formation of which are catalyzed by GSTs (Kushman et al., 2007; Tang et al., 2010; Weng et al., 2005). Gelhaus et al. (2011) proposed that a detoxification pathway had been up-regulated as opposed to an activation pathway that had been down-regulated. The AhR thus appears to be involved in both toxification and detoxification events of several exogenous carcinogens. While there may be other pathways involved in lung carcinogenesis

*Abbreviations:* AhR, aryl hydrocarbon receptor; ATCC, American Type Culture Collection; BaP, benzo[*a*]pyrene; CHTN, Cooperative Human Tissue Network; CYP1, cytochrome P450 family 1; DMSO, dimethyl sulfoxide; $E_2$, 17β-estradiol; EROD, ethoxyresorufin-*O*-deethylase; FBS, fetal bovine serum; FFPE, formalin-fixed, paraffin-embedded; gDNA, genomic DNA; PAGE, polyacrylamide gel electrophoresis; PAH, polycyclic aromatic hydrocarbon; PCR, polymerase chain reaction; SNP, single nucleotide polymorphism; Sp, specificity protein; TCDD, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin; TSS, transcriptional start site.

* Corresponding author at: Department of Environmental Health Sciences, School of Public Health, University at Albany, State University of New York, Albany, NY 12201, United States. Fax: +1 518 473 2895.
*E-mail address:* spink@wadsworth.org (D.C. Spink).

http://dx.doi.org/10.1016/j.taap.2014.10.017
0041-008X/

(Chiba et al., 2011), metabolic activation of smoke-borne carcinogens is almost universally accepted as a causative, initiating event.

A fully functional metabolic response to PAH exposure in airway epithelial cells involves the AhR-mediated upregulation of CYP1 enzymes and epoxide hydrolase, metabolism of PAHs to phenols, dihydrodiols and tetraols, phase II conjugation catalyzed by sulfotransferases and glucuronosyl transferases (Bernier et al., 1996; Nishimura and Naito, 2006; Vietri et al., 2003; Wiebel et al., 1986), and phase III transport of the metabolites out of the cell mediated by transporters such as AhR-regulated ABCG2 (Ebert et al., 2005, 2007; Leslie et al., 2005; Tan et al., 2010). It is the intermediates in the pathway, the dihydrodiol epoxides, and possibly other reactive metabolites (Xue and Warshawsky, 2005), that give rise to the mutagenic DNA adducts. Thus, while a complete loss of AhR expression and function would prevent PAH-induced lung carcinogenesis, a reduced, but not totally abolished, level of AhR could actually lead to incomplete metabolism and transport of PAH carcinogens, reduced levels of GSTs that inactivate mutagenic dihydrodiol epoxides, and consequently increased levels of the mutagenic intermediates that initiate cancers of the lung and other tissues.

Given the potentially pivotal role of the AhR in the development of lung and other cancers, there have been a number of studies focusing on genetic polymorphisms of the AhR and their potential roles in cancer susceptibility (Cauchi et al., 2001; Chen et al., 2009; Kawajiri et al., 1995; Kim et al., 2007; Li et al., 2013; Ng et al., 2010; Shin et al., 2008). Based on our previous studies (Englert et al., 2012), we contend that one of the most important genetic polymorphisms of the human *AHR* gene is the $(GGGGC)_n$ repeat (with $n = 2, 4, 5$ or 6) in the proximal promoter region, which has essentially been overlooked for technical reasons. The TATA-less promoter of the human *AHR* gene is dependent on specificity protein (Sp) transcription factors for promoter activation. We hypothesize that this $(GGGGC)_n$ repeat polymorphism, which is in the Sp transcription factor binding region of the *AHR* gene proximal promoter, influences inter-individual differences in susceptibility to PAH-induced carcinogens via its effect on AhR expression. As an initial step in determining the potential role of the $(GGGGC)_n$ repeat polymorphism in human cancers, this study was undertaken to investigate the incidence of alleles containing varying numbers of the $(GGGGC)_n$ repeats in DNA from patients with cancers of the lung, colon, or breast.

## Materials and methods

*Cell culture and media.* The NCI-H292 lung cancer cell line was obtained from the American Type Culture Collection (ATCC; Manassas, VA), and these cells were cultured in RPMI1640 medium (with phenol red) supplemented with 100 μM non-essential amino acids, 2 mM L-glutamine, 100 U/mL penicillin, 100 μg/mL streptomycin, and 10% fetal bovine serum (FBS; Sigma, Saint Louis, MO). MCF-7 cells were those used in our previous studies (Spink et al., 2003, 2012), and they were cultured in $DF_5$, which consists of DMEM (with phenol red) supplemented with 100 μM non-essential amino acids, 2 mM L-glutamine, 100 U/mL penicillin, 100 μg/mL streptomycin, 10 ng/mL human recombinant insulin (Gibco, Life Technologies, Grand Island, NY), and 5% FBS. The colon adenocarcinoma cell line, Caco-2 (from ATCC), was cultured in $DF_{10}$, which differed from $DF_5$ in that it contained 10% FBS and no added insulin. All cultures were maintained at 37 °C in humidified air containing 5% $CO_2$. Ethoxyresorufin-*O*-deethylase (EROD) assays for all cell lines were carried out in $DC_{10}$ medium, which differed from $DF_{10}$ in that it contained 10% Cosmic Calf Serum (Hyclone, Logan, UT) rather than FBS, and it did not contain phenol red.

*RNA isolation and real-time PCR.* For the determination of mRNA levels in NCI-H292, Caco-2, and MCF-7 cells, confluent cultures in 6-well plates were treated with 10 nM TCDD or 0.1% DMSO vehicle for 48 h. Total RNA was then isolated and reverse-transcribed, and levels of CYP1A1, CYP1B1, and 36B4 mRNA were quantified by real-time PCR using the primers and conditions previously described (Spink et al., 2003).

*Western immunoblots of microsomal CYP1A1 and CYP1B1.* Confluent cultures of NCI-H292, Caco-2, and MCF-7 cells in 10-cm dishes were treated with 10 nM TCDD or 0.1% DMSO vehicle for 48 h. Microsomes were prepared as described (Spink et al., 1997) and were suspended in 10 mM Tris–HCl, pH 7.4, 150 mM KCl, and 20% glycerol. Microsomal proteins (30 μg/lane) were resolved with 10% Bis–Tris gels (NuPage; Invitrogen, Life Technologies) and blotted onto Invitrolon PVDF membranes (Invitrogen). Blots were probed with anti-CYP1A1 (H-70; Santa Cruz Biotechnology, Dallas, TX) or CYP1B1 (H-105; Santa Cruz Biotechnology) antibodies as described (Spink et al., 2003) and detected using the SuperSignal West Pico Chemiluminescent Substrate (Pierce, Thermo Fisher, Rockford, IL) and the ChemiDoc imaging system (BioRad, Hercules, CA).

*EROD assay.* EROD assays were carried out with cultures in 96-well plates as previously described (Spink et al., 2009). In brief, cells in their respective media were treated with 10 nM TCDD (Cambridge Isotope Laboratories, Andover, MA) or 0.1% dimethyl sulfoxide (DMSO) vehicle for 48 h. Conversion of the CYP1 substrate, ethoxyresorufin (Sigma), to resorufin (Sigma) was monitored after 30 min by measuring fluorescence at 590 nm elicited by excitation at 535 nm. EROD activities were normalized to total protein content as measured using the BCA protein assay (Thermo Fisher).

*Luciferase assays of the promoters of CYP1A1, CYP1B1, and polymorphic AHR.* The luciferase constructs containing approximately 1450 bp of the human CYP1A1 promoter, pHu-1A1-FL, or 1146 bp of the human *CYP1B1* promoter, p1B1Fluc, were those used in our previous studies (Spink et al., 2003, 2008). For the determination of in vitro promoter activities of the constructs, cells were seeded in 48-well plates in their respective medias and were transfected when the cultures were at approximately 60% confluence with 50 ng/well of the phRG-TK *Renilla* control vector, 150 ng/well of pHu-1A1-FL or p1B1Fluc, and 0.66 μL/well ExGen500 (Fermentas, Thermo Fisher, Rockford, IL) according to the manufacturer's protocol. After 24 h, cells were treated with 10 nM TCDD or the DMSO vehicle (0.1%) and assayed for luciferase activity 48 h after treatment using the Dual-Luciferase Reporter 1000 Assay System as recommended (Promega). Luminescent signals were recorded using a Lumat LB9501 luminometer (Berthold, Germany), and firefly luciferase activities were normalized to *Renilla* luciferase activities (Spink et al., 2009).

A deletion construct of the *AHR* promoter-luciferase reporter plasmid, pGL3-hAhRP, containing 5640 bp of the AhR promoter, $(GGGGC)_4$ in the proximal promoter region, and 159 bp corresponding to the untranslated region of the AhR mRNA, was prepared by restriction endonuclease digestion with *Kpn*I and *Apa*I (Fermentas) as described (Wolff et al., 2001). The resulting construct, referred to as $AhR\Delta(-120)_4$, in which the subscript refers to the number of GGGGC-repeats, encompasses 120 bp of the proximal *AHR* promoter region and 159 bp of the region corresponding to the mRNA. To prepare a construct containing $(GGGGC)_2$, two repeats, or 10 bp, were removed from the $AhR\Delta(-120)_4$ construct using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene, Agilent Technologies, Santa Clara, CA) as described (Englert et al., 2012), and is referred to as $AhR\Delta(-120)_2$. This strategy was also used to prepare a construct containing the single nucleotide polymorphism (SNP) C/A: $(GGGGC)_2GGGG\underline{\mathbf{A}}GGGGC$, in the $AhR\Delta(-120)_4$ construct, which is referred to as $Ah\underline{R}\Delta(-120)_{C/A}$. The oligonucleotides used for site-directed mutagenesis are summarized in Supplementary Table 1. After verification of the inserts by nucleotide sequencing, they were subcloned into the original vector using *Not*I/*Bgl*II (Fermentas) to ensure identical luciferase and surrounding sequences. To obtain constructs containing $(GGGGC)_5$ and $(GGGGC)_6$, genomic DNA (gDNA) from a neonate heterozygous for $(GGGGC)_{5/6}$ was amplified using the primers, AhRseq-F/R (Supplementary Table 1), and subsequent digestion with *Pau*I/*Apa*I (Fermentas) for cloning into pGL3-hAhRP. The $AhR\Delta(-120)_5$ and $AhR\Delta(-120)_6$ constructs containing $(GGGGC)_5$ and $(GGGGC)_6$,