Diversity as Opportunity: Insights from 600 Million Years of AHR Evolution

Rebeka R. Merson  
*Rhode Island College, rmerson@ric.edu*

Mark E. Hahn  
mhahn@whoi.edu

Sibel I. Karchner  
*Woods Hole Oceanographic Institution*

Follow this and additional works at: [https://digitalcommons.ric.edu/facultypublications](https://digitalcommons.ric.edu/facultypublications)

Citation

Diversity as Opportunity: Insights from 600 Million Years of AHR Evolution

Mark E. Hahn\textsuperscript{1,3}, Sibel I. Karchner\textsuperscript{1}, Rebeka R. Merson\textsuperscript{2}

\textsuperscript{1}Biology Department, Woods Hole Oceanographic Institution, MS-32, Woods Hole, MA 02543, USA
\textsuperscript{2}Biology Department, Rhode Island College, 600 Mt. Pleasant Avenue, 251 Fogarty Life Sciences, Providence, RI 02908

\textsuperscript{3}Correspondence: mhahn@whoi.edu; 508-289-3242

Abbreviations: AHR, aryl hydrocarbon receptor; AHRR, AHR Repressor; AHRE, AHR response element; bHLH-PAS, basic helix-loop-helix Per-Arnt-Sim; BNF, beta-naphthoflavone; CYP, cytochrome P450; MYA, da, dendritic arborization; million years ago; TCDD, 2,3,7,8-tetrachlorodibenzo-p-dioxin.

Key words: Ah receptor, aryl hydrocarbon receptor, bHLH-PAS, dioxin, evolution, development, metazoan, vertebrate, fish, genome duplication, gene expression
Abstract

The aryl hydrocarbon receptor (AHR) was for many years of interest only to pharmacologists and toxicologists. However, this protein has fundamental roles in biology that are being revealed through studies in diverse animal species. The AHR is an ancient protein. AHR homologs exist in most major groups of modern bilaterian animals, including deuterostomes (chordates, hemichordates, echinoderms) and the two major clades of protostome invertebrates [ecdyzoans (e.g. arthropods and nematodes) and lophotrochozoans (e.g. molluscs and annelids)]. AHR homologs also have been identified in cnidarians such as the sea anemone *Nematostella* and in the genome of *Trichoplax*, a placozoan. Bilaterians, cnidarians, and placozoans form the clade *Eumetazoa*, whose last common ancestor lived approximately 600 million years ago (MYA). The presence of AHR homologs in modern representatives of all these groups indicates that the original eumetazoan animal possessed an AHR homolog. Studies in invertebrates and vertebrates reveal parallel functions of AHR in the development and function of sensory neural systems, suggesting that these may be ancestral roles. Vertebrate animals are characterized by the expansion and diversification of AHRs, via gene and genome duplications, from the ancestral protoAHR into at least five classes of AHR-like proteins: AHR, AHR1, AHR2, AHR3, and AHRR. The evolution of multiple AHRs in vertebrates coincided with the acquisition of high-affinity binding of halogenated and polynuclear aromatic hydrocarbons and the emergence of adaptive functions involving regulation of xenobiotic-metabolizing enzymes and roles in adaptive immunity. The existence of multiple AHRs may have facilitated subfunction partitioning and specialization of specific AHR types in some taxa. Additional research in diverse model and non-model species will continue to enrich our understanding of AHR and its pleiotropic roles in biology and toxicology.
1. Introduction

The aryl hydrocarbon receptor (AHR) was initially identified because of its role in regulating the induction of drug-metabolizing enzymes and in mediating the extreme toxic potency of chlorinated dibenzo-p-dioxins and related compounds [1,2]. For two decades after its discovery, this protein was of interest only to pharmacologists and toxicologists [3-5]. However, early on it was recognized by some investigators that study of this receptor and its high-affinity ligand 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) offered potential insights into fundamental biological processes, in the same way that other potent toxins and toxicants have been used to understand cellular functions [6]. Now, it is widely acknowledged that the AHR has multiple functions beyond toxicity [7-10], and the past decade has ushered in a new and exciting era in AHR biology as experts from a variety of fields in biomedicine have turned their attention to elucidate the roles of this pleiotropic protein in cell and developmental biology, immunology, and human disease.

Despite the emerging understanding of the AHR, fundamental questions remain concerning its molecular mechanisms of action (including possible non-genomic mechanisms), target genes beyond the well-known genes encoding biotransformation enzymes, and networks of interactions with other signaling pathways. Similarly, the precise mechanisms underlying most AHR-dependent toxic effects of AHR ligands are not yet known. Complementing the studies of biomedical scientists, research by biologists in other fields has provided a comparative perspective that has yielded insights into the variety of biological functions carried out by AHRs in diverse species. The identification and characterization of AHRs in powerful model species such as Mus musculus [11,12], Caenorhabditis elegans [13], Drosophila melanogaster [14], Danio rerio [15], and Nematostella vectensis [16] has been especially valuable because the tools available for those species have facilitated manipulative experiments to assess AHR functions. Understanding the variety of AHR functions in biology may enable a better understanding of the mechanisms by which exposure to AHR ligands leads to toxicity.

Beyond the established model systems, a broader elucidation of the evolutionary history of the AHR, and of the bHLH-PAS (basic helix-loop-helix Per-Arnt-Sim) family to which it belongs, can provide a foundation for understanding shared and novel features of AHR biology, foster insight into toxic mechanisms, and support extrapolation and prediction of responses to chemicals among species [17,18]. To understand AHR evolution, we look at modern representatives of
early-diverging groups. In what organisms did AHR first appear? What were the original functions? Are there fundamental features of AHR action that are conserved throughout its evolutionary history? What new features and roles have evolved and how do they vary among taxa?

2. AHR origins: The AHR is an ancient protein

What were the first organisms to have an AHR? Early studies using radioligand binding assays with \[^3H\]TCDD or 2-azido-3-\[^{125}\]iodo-7,8-dibromodibenzo-p-dioxin suggested that AHR was a vertebrate protein [19,20]. Subsequently, the identification of AHR homologs in C. elegans [13,21] and D. melanogaster [14] revealed that AHR was more broadly distributed in multiple animal phyla. We now know that AHR is present throughout metazoans (Fig. 1). When did it first emerge? We cannot know for sure, but we can obtain clues by looking at genomes of modern descendants of some of the earliest diverging metazoans and their relatives.

Filastereans (represented by the genus Capsaspora) are not metazoans, but they are the closest unicellular relatives of metazoans and possess many of the transcription factors that are important for metazoan development [22,23]. Capsaspora has four bHLH-PAS proteins but no recognizable AHR homolog [22].

Choanoflagellates such as Monosiga and Salpingoeca also are single-celled but at certain stages can form aggregates held together by cell adhesion proteins, considered a primitive type of multicellularity [24-26]. Choanoflagellates have a few bHLH-PAS proteins but no clear AHR homolog [22].

Sponges (phylum Porifera; e.g. genus Amphimedon), which exhibit embryonic and larval stages and possess a large set of metazoan-specific developmental transcription factors, are considered the oldest extant metazoan lineage [27]. The Amphimedon genome encodes three bHLH-PAS proteins that resemble the ARNT/BMAL, HIF/SIM/TRH, and CLOCK proteins of more recently diverging animals, but again no AHR [28].

The most ancient metazoan lineage with a clearly recognizable AHR in its genome is the placozoan Trichoplax [29,30]. Placozoans have three cell layers and a variety of transcription factors involved in metazoan development and cell fate specification, but no recognizable
specialized sensory or nerve cells. The three bHLH-PAS proteins in *Trichoplax* include an AHR homolog that shows high amino acid sequence similarity to human AHR in its bHLH domain (84%), and substantial but lower similarity in its PAS-A (43%) and PAS-B (51%) domains. Nothing is known about the function of the placozoan AHR.

An AHR homolog is also found in a cnidarian, the starlet sea anemone *Nematostella vectensis*. This species, which is often studied because of its phylogenetic position basal to the bilaterian metazoans, has a nervous system, sensory organs, and a full toolkit of metazoan developmental regulatory proteins, including many shown to interact with AHR signaling in mammals (e.g. notch, hes, wnt, fgf). The *N. vectensis* genome encodes seven bHLH-PAS proteins, including homologs of HIF, SIM/TRH, ARNT, BMAL, and CLOCK in addition to the AHR and a second AHR-like protein [16,31]. Functional characterization of the AHR suggests differences compared to AHRs of vertebrate animals. For example, when expressed in vitro the *N. vectensis* AHR protein does not exhibit specific binding of [3H]TCDD or [3H]beta-naphthoflavone ([3H]BNF), prototypical ligands for vertebrate AHRs [16]. In addition, unlike vertebrate AHRs the *N. vectensis* AHR does not interact with ARNT or BMAL in vitro, suggesting that it may act independent of ARNT. *In situ* hybridization shows that AHR is expressed during larval development at the base of the apical tuft (a sensory structure) and later in the developing tentacles [16]. The expression patterns of AHR and ARNT are non-overlapping at most of these stages, providing additional evidence for ARNT-independent function of the *N. vectensis* AHR.

The results in *N. vectensis* strongly suggest that the common ancestor of cnidarians and bilaterian animals already possessed the modern animal set of familiar bHLH-PAS proteins, including AHR. The functions of this early AHR are unknown, but studies in modern bilaterians (protostomes and deuterostomes) have provided some clues.

### 3. AHRs in protostomes: Key roles in development of sensory systems

The protostomes include most of the major invertebrate phyla and two key model species that have provided important insights into possible ancestral functions of AHR [32,33].

An AHR homolog (called AHR-1) in the nematode *C. elegans* [13,21] resembles the cnidarian AHR in its inability to bind typical AHR ligands [13,34]. *AHR-1* is expressed during embryonic...
and larval development and primarily in developing neurons, including touch receptor neurons, GABAergic motor neurons, interneurons, and sensory neurons that contact the pseudocoelemic fluid [35]. Loss of AHR-1 function results in defective neuronal migration and axonal pathfinding, altered touch neuron fate, and changes in locomotor and social feeding behaviors [35-37]. The role of AHR-1 in neuronal development requires ARNT (AHA-1) [35,36] and may involve regulation of wnt signaling [38]. In addition to its roles in development, AHR-1 appears to have an ongoing role in regulating the expression of oxygen-sensing guanylate cyclases involved in the control of feeding behavior [39,40]. Together, these results support a role for AHR-1 in neuronal differentiation, migration, and cell fate determination [32] as well as post-embryonic neuronal functions.

Another powerful model, the fruit fly D. melanogaster, has also provided important insights into the pleiotropic developmental roles of AHR. The fly AHR homolog, the product of the spineless (ss) locus, is expressed in larval eye-antennal imaginal discs, the regions destined to becoming adult eyes and antennae [14]. Loss-of-function mutations demonstrate that ss specifies the identity of the distal segments of antennae (multi-sensory structures) and legs and the formation of mechanosensory bristles (the loss of which is reflected in the name “spineless”) [14]; in the antennae ss appears to have a specific role in development of olfactory sensillae [41]. The action of ss in controlling development of antennae and distal leg require the fly ARNT homolog tango (tgo) [42] and may in part involve the repression of gene expression by ss/tgo complexes [43].

Later in development, ss has a role in specifying photoreceptor cell fate in ommatidia of the developing compound eye. Stochastic expression of ss in specific photoreceptors determines the type of rhodopsin that is expressed, thus controlling color sensitivity in D. melanogaster [44,45] and other insects [46]. As seen for antenna and leg development, the role of ss in controlling photoreceptor cell fate requires tgo (ARNT) [47]. The expression of ss is maintained in these ss-specified photoreceptor subtypes in adults [45] and thus may be necessary to maintain the pattern of rhodopsin expression.

Yet another developmental role of the D. melanogaster AHR homolog is in controlling dendritic morphology on dendritic arborization (da) sensory neurons in the fly peripheral nervous system [48,49]. The effect of ss varies in different types of da neurons, with the end result of diversifying dendrite morphology. Although this role was originally suggested to be independent of
tgo/ARNT [48], more recent results indicate that co-expression of tgo is in fact necessary [47]. Because the C. elegans AHR has a similar role in controlling dendritic branching complexity, this has been suggested as an ancestral role of AHR [37].

The common theme of the research in C. elegans and D. melanogaster, with circumstantial support from studies in the cnidarian N. vectensis, is one of pleiotropic roles of AHR in controlling the development (cell fate and differentiation) and function of sensory structures and neural systems [18,32,33,49]. These functions appear to involve both activation and repression of gene expression by AHR [37,43].

What is the role of ligands in the functions of protostome and cnidarian AHR homologs? Although these AHRs do not appear to bind typical (i.e. vertebrate) AHR ligands [13,16,34] and there is some evidence for constitutive, ligand-independent activity [32,42,50], it is also possible that there are endogenous ligands or other regulatory mechanisms involved [51-53]. Nevertheless, we refer to these proteins as “protoAHRs” (Table 1) to highlight the apparently substantial differences in ligand specificity between these proteins and their vertebrate homologs, which function (at least in part) as true “aryl hydrocarbon receptors." It is important to note, however, that all of the evidence currently available is consistent with the idea that protoAHRs and vertebrate AHRs are true orthologs (i.e. descended from the same gene in the most recent common bilaterian ancestor [54]).

4. AHR in deuterostomes: Expansion through gene and genome duplications.

The other major group of bilaterian animals, the deuterostomes, includes echinoderms, hemichordates, and chordates (Fig. 1). Predicted AHR homologs are found in genomes of the echinoderm Strongylocentrotus (sea urchin) [55,56], hemichordate Saccoglossus (acorn worm) [57], and invertebrate chordates such as the cephalochordate Branchiostoma (amphioxus) [58,59] and urochordate Ciona (sea squirt) [60]. The Ciona AHR, like other invertebrate AHRs, does not bind TCDD (unpublished data), but nothing is known about the function of the other invertebrate deuterostome AHRs.

In contrast to the echinoderms, hemichordates, and invertebrate chordates, the vertebrate chordate lineage is notable for the diversification of AHRs (and other bHLH-PAS proteins [61]) (Table 1), a result of vertebrate- and teleost-specific whole genome duplications as well as an
early tandem duplication of AHR [61]. It is in the vertebrates where we first see AHRs that exhibit high-affinity binding of TCDD [18], AHR-dependent regulation of genes encoding xenobiotic-metabolizing CYP1 enzymes [62,63], and high sensitivity to toxic effects of dioxin-like compounds [64].

In the oldest extant vertebrate group, Agnatha (jawless fishes), represented by the lampreys Petromyzon marinus [65] and Lethenteron japonicum [66], we see a remarkable expansion, with five predicted AHR genes in each species ([67] and S.Karchner unpublished; Table 1). This increase in AHR genes was likely a result of the two whole genome duplications that occurred early in vertebrate evolution, more than 450 million years ago; current evidence suggests that both of these preceded the divergence of agnathan cyclostomes (jawless vertebrates) and gnathostomes (jawed vertebrates) [65,68] (but see also [69]).

In the jawed vertebrates (Chondrichthyes [cartilaginous fishes] and Osteichthyes [bony fishes] and their descendants, including tetrapods), our current understanding—supported by phylogenetic analyses and information from shared synteny—is that there are at least five classes of AHR-related genes (Table 1). We use AHR to refer to vertebrate orthologs of the AHR originally identified in mammals [11,12]. AHR genes are found in nearly all vertebrates, including sharks, gar, and sturgeon. However, teleosts (the largest group of ray-finned fishes) are notable in that most of those studied to date lack this AHR. The exception is zebrafish, where the enigmatic “AHR1a” [70-72] appears to be an AHR based on our recent analysis of shared synteny with human AHR and other AHR genes (Fig. 2A).

AHR1 and AHR2 are paralogs derived from a tandem gene duplication that occurred prior to the divergence of cartilaginous and bony fish lineages [21,61,71,73]. Although AHR1 was originally thought to be orthologous to mammalian AHR [73], more recent analysis of additional AHR sequences reveals that AHR and AHR1 represent distinct lineages (Table 1). AHR1 and AHR2 orthologs are found in tandem in cartilaginous fishes [74], bony fishes (which usually have duplicated AHR1-AHR2 pairs; see below) [61,71], coelacanth (a lobe-finned fish) [75], birds [76,77], reptiles [78], and an early diverging marsupial mammal, the opossum Monodelphis ([79]; C. Panti, S. Karchner, M. Hahn, unpublished) (Table 1). Based on Xenopus genomes and other analyses, the AHR1/AHR2 pair has been lost from at least some amphibians [80] (Table 1). Although orthologs of AHR1 and AHR2 are not found in rodents or humans, there are predicted AHR2 genes in several genomes in the mammalian orders Carnivora, Cetartiodactyla,
and Primates (both old world and new world monkeys, but not great apes; S. Karchner, R. Merson, and M. Hahn, unpublished) (Table 1). These occur without an adjacent AHR1, but the identity as AHR2 is supported by phylogenetic analyses (not shown) as well as by shared synteny with species possessing tandem AHR1-AHR2 pairs (Fig. 2B).

AHR1 and AHR2 genes are often found as duplicated pairs in teleosts. Thus, most teleost genomes include both an AHR1a-AHR2a tandem pair and an AHR1b-AHR2b tandem pair [61,71,81]; these duplicated pairs are thought to have arisen as part of the teleost-specific whole-genome duplication [82,83]. A prominent exception is the zebrafish (D. rerio), which has only one AHR1-AHR2 pair (orthologous to AHR1b-AHR2b in other teleosts) [61,71] and a separate AHR (currently designated AHR1a, but likely an “AHR”, as noted above) [71,72]. Interestingly, the recently sequenced genome of gar, representing the holostean lineage, which diverged prior to the teleost-specific genome duplication [84], does not contain an AHR1-AHR2 pair. However, it has—in addition to an AHR—an AHR2 that is orthologous to other fish AHR2 genes, based on phylogenetic analysis and shared synteny; thus, the tandem AHR1 appears to have been lost in this species. Similarly, sturgeon (a chondrosteian) has an AHR and an AHR2, but no AHR1 [85].

AHR3 is a novel AHR found originally in elasmobranchs (a subclass of cartilaginous fishes encompassing true sharks, skates, and rays; R. Merson & M. Hahn, unpublished; see also [61]) (Table 1). In the shark Squalus acantbias both AHR2 and AHR3 (but not AHR1) bind TCDD and are transcriptionally active in heterologous expression systems (Merson et al., in preparation). The genome of the elephant shark [86], a representative of the cartilaginous fish subclass Holocephali (chimaeras), contains two possible AHR3-like genes. AHR3-like genes are also found in some early diverging fishes such as lamprey, gar, and coelacanth (Table 1) but they form a distinct clade in phylogenetic analyses. The resolution of AHR3 relationships will require analysis of additional species and the completion of genome assemblies to assess the genomic context of this locus. AHR3 and AHR3-like genes do not appear elsewhere in the vertebrates.

AHRR, first identified in mouse [87], is distinct from AHR, AHR1, AHR2, and AHR3 and acts via multiple mechanisms to repress signaling through AHR and some other pathways [88]. The PAS-B region of AHRR, which in other AHR-related proteins forms the ligand-binding domain, is missing or highly divergent in AHRR [87]; consistent with that, AHRRs do not bind [3H]TCDD or
AHRR has been retained in nearly all vertebrate groups (Table 1), suggesting that it has an important physiological function. Recent findings regarding the possible roles of AHRR in the immune system [90,91], reproduction [92,93], and carcinogenesis [94,95] support that notion.

5. Functional divergence of a pleiotropic protein: Shared and divergent roles of metazoan AHRs

What are the ancestral roles of AHRs and how have they changed during metazoan evolution? We look for evidence in functions that may be shared among modern animals whose most recent common ancestor lived long ago, e.g. protostomes and deuterostomes, but these roles can be difficult to identify given the substantial developmental and physiological differences among long-diverged lineages.

Studies in protostomes provide evidence for pleiotropic roles of AHR in controlling the development and function of sensory structures and neural systems and these have been suggested as ancestral roles of AHR [18,32,33,37,49]. Possible roles of AHR in sensory/neural systems are less well understood in vertebrate animals, but results from AHR loss-of-function studies and effects of TCDD on neural development in fish and mammals [96-103] suggest that this is an area worth further exploration. For example, roles for AHR in developing GABAergic systems and in controlling dendrite growth may be shared by vertebrate and invertebrate species [36,37,96,99,102,104].

Other developmental and physiological roles of AHR have been identified in vertebrates, including some involving vascular development, reproductive function, immunological development, and stem cell biology [7,8,33,105-107]. Some of these developmental roles may explain the special sensitivity of vertebrate early life stages to disruption by AHR ligands [108,109]. However, the possible connection of these roles to those of invertebrate AHRs is not obvious. Conceivably, shared features of AHR function may be more readily identified at the level of molecular interactions such as those involving gene regulation or protein-protein interactions. For example, interaction with wnt signaling appears to be a shared, and thus possibly evolutionarily conserved, feature of vertebrate and invertebrate AHRs [38,110-113]. A comparative analysis of gene regulatory networks involving AHR in a variety of vertebrate and
invertebrate model systems could illuminate additional ancient molecular interactions [40,114-116].

What is the role of ARNT and AHR-ARNT interactions with AHR response elements (AHREs, also called DREs and XREs) in ancestral and modern functions of AHRs? At least some of the toxic effects of AHR ligands in mammals require both ARNT dimerization [117] and DNA binding [118]. Similarly, some of the developmental roles in invertebrate species are ARNT-dependent and/or involve interactions with AHRE sequences similar to those found in vertebrates [35,36,42,47]. Yet there also is evidence for ARNT-independent functions in Nematostella [16] and, increasingly, evidence for ARNT-independent or AHRE-independent roles of AHR in vertebrates [119-122].

The increased AHR diversity in vertebrate animals appears to have been accompanied by (and perhaps enabled) the emergence of new AHR adaptive functions, including regulation of the inducible expression of genes encoding xenobiotic-metabolizing enzymes such as cytochrome P450s (CYPs) in response to chemicals. Although CYPs and other biotransformation enzymes are inducible in *C. elegans*, the nematode AHR does not appear to be involved in this response [123]. In insects, there is evidence that AHR and ARNT homologs regulate the basal expression, but not the xenobiotic-inducible expression, of CYP6B1, which is involved in detoxification of dietary phytotoxins [124]. The first clear evidence for AHR-dependent regulation of inducible CYP1 genes is in jawed vertebrates [62,64]. How this association between AHR and CYP regulation evolved remains a mystery. However, one clue may be the intriguing tandem arrangement of AHR and CYP1-like genes in the urochordate *Ciona* [63], suggesting a possible mechanism whereby auto-regulation of AHR might have become co-opted by CYP1 through physical proximity on the chromosome.

As we have noted earlier, the evolution of the ability of AHR to engage in high-affinity ligand binding associated with ligand-dependent adaptive functions was a vertebrate innovation [18], perhaps driven by a need to detoxify halogenated marine natural products [125-128]. Ironically, this new ligand-dependence of AHR also introduced a mechanism by which some persistent halogenated aromatic hydrocarbons could cause toxicity through high-affinity AHR binding and sustained AHR activation. Invertebrate animals, which have AHRs that lack the ability to bind dioxin-like compounds, are generally insensitive to the toxicity of these chemicals [18,64].
Additional AHR-mediated adaptive functions that may have first emerged in vertebrate animals are those involving the regulation of innate and adaptive immunity [7,129]. Some of these immunological roles of AHR appear to involve endogenous and microbiota-derived ligands [130-134], although it is not yet known whether AHR affinity for some of these ligands (many of them indole derivatives) evolved in parallel with these immune functions.

In addition to novel functions, the AHR expansion in vertebrates may have enabled AHR isoform specialization, through subfunction partitioning and subsequent functional refinement. AHRR may be one example of that, whereby loss of ligand-binding through degeneration of the PAS-B ligand-binding domain [89] led to a specialization for repression of gene expression.

Other examples of AHR specialization involve the apparent partitioning of tissue expression patterns, ligand specificity, and target gene specificity among multiple AHR paralogs in some non-mammalian vertebrates. Whereas most mammals have a single pleiotropic AHR, through which various classes of ligands must all act, there is evidence that the multiple AHRs of fishes, birds, reptiles, and amphibians have partitioned some of these functions. For example, the set of three AHR genes in zebrafish have evolved very different functional properties and expression patterns involving partitioning of ligand specificity and developmental versus adaptive roles [70-72,135-137]. Whereas zebrafish AHR2 appears to mediate most of the gene induction and developmental toxicity of TCDD, PCB-126, and some polycyclic aromatic hydrocarbons [138-143], AHR1a is preferentially involved in the response to some non-halogenated compounds such as leflunomide, pyrene, and oxygenated PAHs [70,136,144] and AHR1b may have a tissue-specific developmental role [137]. Similarly, AHR paralog-specific differences in ligand structure-activity relationships (e.g. for halogenated vs. non-halogenated ligands) or target gene specificity have been observed in chicken [76,145], alligator [78], and frog [146].

6. Conclusions

From the information summarized above (and other data that could not be covered in a brief review such as this) we offer some conclusions, some of which must necessarily be considered tentative.
Clearly, the AHR is an ancient protein, which has existed for more than 600 million years of animal evolution and should be considered part of the fundamental metazoan toolkit. In modern (living) invertebrates, AHR has roles in the development of sensory structures, including sensory neural systems; these may be some of the most ancient roles of metazoan AHRs. AHRs have undergone substantial diversification in the vertebrate chordates; this diversification was likely facilitated by the gene and genome duplications occurring prior to and after the vertebrate radiation. AHRs are pleiotropic, with multiple functions that vary by cell type and developmental stage within single species as well as among animal taxa. In some cases, those multiple functions are partitioned among AHR isoforms within a species. The emergence of the adaptive functions of AHR in the vertebrates is associated with the acquisition of high-affinity binding of planar aromatic hydrocarbons, which appears to be a vertebrate innovation. This broadened the capacity for inducible detoxification of xenobiotics but also introduced a mechanism by which some persistent, high-affinity ligands could cause toxicity.

It is worth noting that all of the information we have about AHR functions—including ligand-binding, protein-protein interactions, DNA binding, gene regulation, and developmental roles—is from studies in modern animals. All of these species—from *Trichoplax* to humans—can be considered “advanced” in that they are the result of a long evolutionary process; i.e. they are the “survivors.” These modern animals are the descendants of earlier species in which resided the ancestral functions of AHR that we seek to understand. Although we cannot turn the clock back to study these ancestors, it is now possible, through ancestral sequence reconstruction, to resurrect and study ancestral proteins [147,148]. The evolution of AHR ligand-dependence and ligand specificity, in particular, may be revealed by reconstruction and analysis of the ancestral AHR proteins that existed at key points in metazoan evolution, such as the emergence of bilaterians or the base of the vertebrate radiation.

There is an intuitive appeal to the hypothesis that the AHR had ancestral roles in the development of sensory structures and neurons that were later co-opted for novel roles in chemical sensing and adaptive responses. Studies of AHRs in new model and non-model systems will help to illuminate or refute that idea. One thing is indisputable: the AHR will continue to intrigue and surprise us over the next decade as its manifold roles are revealed.
Acknowledgments

This paper is based in part on an invited presentation at the AHR 2016 Symposium, “The Aryl Hydrocarbon Receptor as a Central Mediator of Health and Disease,” held at the University of Rochester Medical Center from August 3-6, 2016. The authors thank Drs. Ann Tarrant and Wade Powell for comments on an earlier version of the manuscript, and Diana Franks for assistance with AHR assays in a variety of species over many years. M.E.H. and S.I.K. are grateful for the long-term support of our AHR research from the National Institute of Environmental Health Sciences (NIEHS) through grants R01ES006272 and P42ES007381 (Superfund Research Program at Boston University). We also acknowledge support from a WHOI Independent Study Award funded by the Andrew W. Mellon Foundation Endowed Fund for Innovative Research. R.R.M. acknowledges support from the NIH National Center for Research Resources RI-INBRE (P20RR016457), National Science Foundation EPSCoR Cooperative Agreement #EPS-1004057, a MDIBL New Investigator Award funded by ME-INBRE (P20RR016463), and NIEHS grant P30ES003828. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. The funding agencies had no role in the preparation of the report or the decision to submit the manuscript for publication.
References cited:

Papers of particular interest are noted:
• of special interest
•• of outstanding interest

   • This insightful early paper drew parallels between TCDD and other highly toxic chemicals that have been used to probe biological systems. It predicted that understanding the mechanism of TCDD action through the AHR would lead to fundamental understanding in cell biology.


• This paper characterizes the AHR in an early-diverging metazoan, providing insight into possible ancestral roles, including possible ARNT-independent functions.


• This paper was the first report of multiple AHR genes in vertebrate animals.


   • This is an outstanding review of developmental roles of AHR in diverse animal models.

   • This paper demonstrated the inability of invertebrate AHRs (arthropods, nematodes, and molluscs) to bind typical AHR ligands TCDD and β-naphthoflavone.

   •• This is one of two elegant papers that first identified the neurodevelopmental roles of AHR in *C. elegans*.

   •• This is one of two elegant papers that first identified the neurodevelopmental roles of AHR in *C. elegans*.

   • This paper describes a role for AHR in specifying the cell fate of sensory neurons and identifies control of dendritic branching complexity as an evolutionarily conserved AHR function.


• This paper shows that *Drosophila* AHR function requires ARNT and that *Drosophila* AHR-ARNT complexes can recognize mammalian AHR response elements. They also present evidence that AHR function in *Drosophila* is ligand-independent.


•• This paper shows that Drosophila AHR acts as a switch to control the cell fate of photoreceptors, determining the type of rhodopsins expressed and thus regulating the development of color vision.


• This paper presents a detailed phylogenetic analysis of the CYP1 gene family and discusses the evolutionary history of AHR-dependent regulation of CYP1 expression.


• This paper describes the phenotype of AHR2-mutant zebrafish and provides new insight into the ligand-specific and regulatory diversity of the three zebrafish AHRs.


74. Merson RR, Mattingly CJ, Planchart AJ: Tandem duplication of aryl hydrocarbon receptor (AHR) genes in the genome of the spiny dogfish shark (Squalus acantthias). Bulletin of the Mount Desert Island Biological Laboratory 2009, 48:43-44.


• This paper was the first to identify AHRR as an AHR-related repressor of AHR signaling.


receptor repressor (AhRR) and AhR-mediated induction of CYP1A1, cytokines, and acute toxicity. *Environ Health Perspect* 2016, **124**:1071-1083.


• This paper is the first to systematically determine genes regulated by AHR in the presence and absence of a xenobiotic ligand.


125. Stegeman JJ, Hahn ME: Biochemistry and molecular biology of monooxygenases: Current perspectives on forms, functions, and regulation of cytochrome P450 in aquatic species. In Aquatic Toxicology: Molecular, Biochemical and Cellular Perspectives. Edited by Malins DC, Ostrander GK: CRC/Lewis; 1994:87-206.


141. Incardona JP, Linbo TL, Scholz NL: Cardiac toxicity of 5-ring polycyclic aromatic hydrocarbons is differentially dependent on the aryl hydrocarbon receptor 2 isoform during zebrafish development. Toxicology and applied pharmacology 2011, 257:242-249.


Figure legends

**Figure 1. Presence of AHR homologs in holozoans.** The tree shows the relationships of selected metazoan (animal) taxa and related unicellular eukaryotes and whether they possess AHR homologs. Ecdysozoa and Lophotrochozoa together comprise the protostomes. Protostomes and deuterostomes are bilaterian animals (green-shaded box). Solid boxes represent groups containing species from which AHR homologs have been confirmed by cloning. Dashed boxes occur around groups with AHR homologs predicted from sequenced genomes. The large yellow-shaded box encompasses the eumetazoans, a group that includes all the taxa in which AHR has been identified to date. See text for additional information. Phylogenetic relationships of choanoflagellates and filastereans after Torruella [149].

**Figure 2. Analysis of shared synteny supports AHR classification.** (A) Zebrafish AHR1a and related AHR genes in earlier-diverging fishes may be orthologous to human AHR and related AHR genes. (B) Predicted AHR2 genes found in several mammals exhibit shared synteny with AHR2 genes from fish and birds. Analysis of syntenic relationships was performed using Genomicus [150] and manual scanning of sequenced genomes.
Table 1. Phylogenetic distribution of multiple AHR forms in metazoans.

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Representative genus</th>
<th>common name</th>
<th>protoAHR</th>
<th>AHR</th>
<th>AHR1</th>
<th>AHR2</th>
<th>AHR3</th>
<th>AHRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filastereans</td>
<td>Capsaspora</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choanoflagellates</td>
<td>Monosiga Salpingoeca</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Porifera</td>
<td>Amphimedon sponge</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placozoans</td>
<td>Trichoplax</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cnidarians</td>
<td>Nemastella sea anemone</td>
<td>+ (2)</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arthropod</td>
<td>Drosophila fruit fly</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nematode</td>
<td>Caenorhabditis worm</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mollusc</td>
<td>Mya clam</td>
<td>+</td>
<td>+ (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Echinoderm</td>
<td>Strongylocentrotus sea urchin</td>
<td>+ (2)</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cephalochordate</td>
<td>Branchiostoma amphioxus</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urochordate</td>
<td>Ciona sea squirt</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jawless fishes</td>
<td>Petromyzon Lethenteron lamprey</td>
<td>+ (5)</td>
<td>+ (5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holocephalan</td>
<td>Callorhinchus elephant shark</td>
<td>+ (2)</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elasmobranch</td>
<td>Squalus spiny dogfish</td>
<td>+ (2)</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chondrosteian</td>
<td>Acipenser sturgeon</td>
<td>+ (2)</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holosteanian</td>
<td>Lepisosteus gar</td>
<td>+</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teleost</td>
<td>Fundulus, Oryzias killifish, medaka</td>
<td>+ (2)</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Danio zebrafish</td>
<td>+</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lobe-finned fishes</td>
<td>Latimeria Coelacanth</td>
<td>+</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amphibian</td>
<td>Xenopus frog</td>
<td>+</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reptile</td>
<td>Alligator alligator</td>
<td>+</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bird</td>
<td>Gallus chicken</td>
<td>+</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monotreme</td>
<td>Ornithorhynchus platypus</td>
<td>+</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marsupial</td>
<td>Monodelphis opossum</td>
<td>+</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eutherian mammals</td>
<td>Mus mouse</td>
<td>+</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Felis, Panthera cat, tiger</td>
<td>+</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bos cow</td>
<td>+</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Loxodonta elephant</td>
<td>+</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Equus horse</td>
<td>+</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chlorocebus, Callithrix green monkey, marmoset</td>
<td>+ (2)</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Homo, Gorilla human, gorilla</td>
<td>+ (2)</td>
<td>+ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table presents a summary of our current thinking regarding AHR multiplicity and the phylogenetic distribution of various AHR forms. Results are derived from published papers and from our analyses of numerous phylogenetic trees and syntenic relationships (shared synteny). The conclusions here should be considered tentative and will be subject to revision as additional sequences and species are examined. See sections 3 and 4 of text for details regarding the different types of AHRs.

Note: bold boxes indicate a demonstrated tandem arrangement of AHR1 and AHR2 genes.
*The classification of these as AHR3 is uncertain. Resolution will require additional sequences and synteny information.
Highlights

Diversity as Opportunity: Insights from 600 Million Years of AHR Evolution

Short title: 600 Million Years of AHR Evolution

Mark E. Hahn, Sibel I. Karchner, Rebeka R. Merson

- The AHR is an ancient protein, part of the fundamental eumetazoan genetic tool kit.
- Ancestral roles of AHR likely included control of sensory neural development.
- AHR has undergone expansion in vertebrates, generating at least 5 AHR types.
- Adaptive roles of AHR in xenobiotic sensing may have emerged in vertebrates.
- Modern AHRs exhibit both ancestral and recently evolved roles in cell biology.