Independent losses of a xenobiotic receptor across teleost evolution

Marta Eide\textsuperscript{1}, Halfdan Rydbeck\textsuperscript{2}, Ole K. Terresen\textsuperscript{2}, Roger Lille-Langøy\textsuperscript{1}, Pål Puntervoll\textsuperscript{3}, Jared V. Goldstone\textsuperscript{4}, Kjetill S. Jakobsen\textsuperscript{2}, John Stegeman\textsuperscript{4}, Anders Goksøyr\textsuperscript{1} & Odd A. Karlsen\textsuperscript{1}

Sensitivity to environmental stressors largely depend on the genetic complement of the organism. Recent sequencing and assembly of teleost fish genomes enable us to trace the evolution of defense genes in the largest and most diverse group of vertebrates. Through genomic searches and in-depth analysis of gene loci in 76 teleost genomes, we show here that the xenosensor pregnane X receptor (Pxr, Nr1I2) is absent in more than half of these species. Notably, out of the 27 genome assemblies that belong to the Gadiformes order, the pxr gene was only retained in the Merluccidae family (hakes) and Pelagic cod (Melanonus zugmayeri). As an important receptor for a wide range of drugs and environmental pollutants, vertebrate PXR regulate the transcription of a number of genes involved in the biotransformation of xenobiotics, including cytochrome P450 enzymes (CYP). In the absence of Pxr, we suggest that the aryl hydrocarbon receptor (Ahr) have evolved an extended regulatory role by governing the expression of certain Pxr target genes, such as cyp3a, in Atlantic cod (Gadus morhua). However, as several independent losses of pxr have occurred during teleost evolution, other lineages and species may have adapted alternative compensating mechanisms for controlling crucial cellular defense mechanisms.

Teleost fishes represent the largest and most diverse vertebrate clade. By inhabiting a wide range of oceanic and freshwater habitats, ranging from tropical to arctic regions, the genetic composition and phenotype of teleost fishes have evolved accordingly\textsuperscript{1,2}. However, genes involved in the orchestrated cellular recognition, biotransformation and elimination of chemical compounds, denoted the chemical defensome, are largely conserved through evolution\textsuperscript{3–5}. Today, the increasing burden of man-made compounds poses a great risk to the aquatic environment. Since large pelagic and benthic teleost fish are the dominant group of predators in global oceans, they are particularly vulnerable to ecosystem-wide changes in energy resources\textsuperscript{6}, such as those caused by pollution. Thus, understanding the genomic defense mechanisms and their downstream effects in these species is paramount in conserving environmental diversity.

Atlantic cod (Gadus morhua) was the first commercially important, large teleost species of which the genome was sequenced and assembled\textsuperscript{7}. Furthermore, it is commonly used as an indicator species in environmental monitoring programs and is an emerging marine model organism in toxicological studies\textsuperscript{8–11}. Due to its habitats near offshore oil platforms, petroleum recovery facilities and coastal industries, the effects of contaminants in Atlantic cod are of great interest\textsuperscript{12–15}. We have recently mapped the full suite of cod cytochrome p450 (cyp) genes, providing further clues to the inherent capabilities for biotransformation of endogenous and xenobiotic compounds in this species\textsuperscript{16}. However, the proteins that regulate various chemical defense mechanisms have not been described in Atlantic cod.

Pregnane X receptor (PXR, NR1I2) and constitutive androstane receptor (CAR, NR1I3) are considered two of the main xenobiotic receptors in mammals that, in addition to binding endogenous steroid hormones, are activated by a structurally diverse range of drugs and environmental pollutants. Primarily expressed in the liver and small intestine, these xenosensors regulate the transcription of biotransformation enzymes, such as cytochrome P450 3A (CYP3A), and are thus central in orchestrating the chemical defensome. PXR and CAR belong to the large family of ligand-activated transcription factors known as nuclear receptors (NR), more specifically the subfamily NR1I, and show the common NR structural organization of a highly conserved DNA-binding domain

\textsuperscript{1}Department of Biological Sciences, University of Bergen, Bergen, Norway. \textsuperscript{2}Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, Oslo, Norway. \textsuperscript{3}Centre for Applied Biotechnology, Uni Research Environment, Bergen, Norway. \textsuperscript{4}Biology Department MS, Woods Hole Oceanographic Institution, Woods Hole, MA, USA. Correspondence and requests for materials should be addressed to O.A.K. (email: odd.karlsen@uib.no)
(DBD) and a less conserved ligand-binding domain (LBD)\(^1\). The early evolutionary history of this subfamily is not clear, but the NR1I genes are hypothesized to diverge from a common ancestral gene, such as that identified in the chordate invertebrate Ciona intestinalis\(^1\). A vitamin D receptor (VDR, NR1I1) has been cloned from sea lamprey (Petromyzon marinus), an ancient vertebrate lacking calcified skeleton and jaws, and proposed to function in part as a xenosensor\(^1\). However, early in, or before, vertebrate evolution, this ancestral gene diverged into VDR and PXR\(^8\).

Phylogenetic analyses suggest that the further divergence of CAR and PXR occurred in a whole genome duplication (WGD) in, or prior to, the formation of the vertebrate crown lineage, Sarcopterygii\(^2\). However, except for mammals, which have retained both xenosensors, several independent losses of CAR and PXR have occurred in vertebrate evolution, leaving birds and lizards with CAR (also known as CXR\(^2\)) and teosel species with PXR\(^2\)\^-\(^4\). The aryl hydrocarbon receptor (Ahr) is another important transcription factor in the chemical defensome of vertebrates. Also known as the "dioxin receptor", the ligands of this basic-Helix-Loop-Helix-PAS domain receptor include environmental pollutants such as 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and certain planar polycyclic aromatic hydrocarbons like benzo[\(a\)l]pyrene. The most studied piscine Ahr target gene is cyp1a\(^1\)\^-\(^5\), but recent studies have indicated its co-regulative role in mediating transcription of the Pxr target gene, cyp3a65, in zebrafish\(^6\)\^-\(^8\).

The aim of the present study was to examine the evolution of Pxr orthologs in teleost fish. Importantly, based on global genomic searches and in-depth analysis of gene loci of 76 teleost genome assemblies\(^2\)\^-\(^4\), we show that several independent losses of pxr have occurred during piscine evolution. Furthermore, although we show that a pxr gene is absent in most species in the Gadiformes order, we detect its presence in the Merluccidae (hakes) family. Moreover, by mapping the complete NR complement in Atlantic cod, we show that the other nuclear receptors are conserved in this Gadiform species. Finally, lack of Pxr agonist-induced transcriptional response in Atlantic cod liver, as well as a highly increased number of putative Ahr response elements in the promoter region of Pxr target genes, suggest that Ahr has evolved a compensatory role in the absence of Pxr in this species. However, as several independent losses must have occurred during teleost evolution, other lineages may have adapted other compensating mechanisms. This is the first report with compelling evidence for multiple independent losses of the pxr gene in vertebrate evolution and accentuate the need for understanding how the chemical defensome system has evolved in teleost fish.

Results
In silico searches in 76 sequenced and assembled teleost genomes and using zebrafish Pxr as query sequence, showed that a pxr gene could not be identified in the majority of the fish species (Fig. 1). Additional searches using Pxr from medaka or tetraodon gave similar results (Supplementary Fig. 1). Surprisingly, pxr was only present in 33 out of the 76 genomes analyzed. Searches performed using zebrafish Vdros as query, however, showed that a vdr gene was present in all species (Supplementary Fig. 2).

As a positive control of the pipeline used for the gene mapping, searches using zebrafish Mhc II were included, showing an identical phylogenetic distribution as published previously (Fig. 1)\(^8\). Multiple hits on the NR DBD are the result of the high similarity between these domains in the different NR proteins within a genome.

Furthermore, we analyzed the genomic scaffolds containing MAATS1 and GSK3B, which flank PXR in all examined phyla that possess PXR. Figure 2 shows that the syntetic relationship of these genes is also maintained throughout the evolution of fishes. Importantly, in Atlantic cod and other species where pxr was not identified, no gene is present in the genomic region between maats1 and gsk3b. Our results show that the loss of pxr has occurred in several orders, including Zeiformes, Holocentriformes, and Beryciformes. Importantly, out of the 27 genome assemblies in the Gadiform order, we only identified sequences similar to Pxr in the Merluccidae family (hakes, represented by M. polli, M. merluccius, and M. capensis) and Melanonus zugmayeri (Pelagic cod). The syntetic relationship of maats1 and pxr was confirmed in European hake (M. merluccius), while gsk3b was assembled to another scaffold (Fig. 2).

In order to explore whether the presence of Merluccidae and Melanomonidae ppxr was due to retention or introgression, we aligned the translated LBDs of the identified Pxr sequences. Since Vdros (Nr1I1) are in the same NR subfamily as Pxr, we also included known teleost orthologs of this receptor, as well as orthologous retinoic acid receptor (Rar\(γ\), Nr1b3) and Rev-Era (Nr1d1) sequences as an outgroup. Although the families are distantly related, the sequences identified from Melanorus and Merluccius genomes are grouped together in the resulting phylogenetic tree, showing that their Pxr sequences are specific for the Gadiformes order (Fig. 3). The phylogeny also supported our initial bioinformatic searches, as none of the Pxr sequences grouped with the teleost Vdros sequences. However, some short sequences that were identified using medaka and tetraodon Pxr as query (Supplementary Fig. 1) were found to group with Rar\(γ\) and Rev-Erb\(α\) (Nr1d1), and not Pxr (Fig. 3). While loss of the third member of the NR1I subgroup in mammals, CAR, is demonstrated previously in teleost fish\(^2\)\^-\(^4\), we also performed searches using the unusual Xenopus laevis BXR\(^3\) without identifying any new Pxr-similar sequences.

Thus, our investigations support the hypothesis that Pxr is retained in Merluccidae and Melanomonidae even though several, possibly three, losses of pnxr have occurred in the evolution of the Gadiformes order alone (Fig. 4).

In the Atlantic cod genome (gadMor1 7), HMM-searches identified 72 nuclear receptor-encoding genes, which were further mapped to 50 unique members of the NR superfamily (Supplementary Table 2). To visualize their phylogenetic relationship and distribution into the NR families and subfamilies, the NR complement of Atlantic cod and zebrafish were compared by multiple sequence alignments and visualized as a cladogram (Fig. 5).

Importantly, whereas the zebrafish NR1I subfamily contains Pxr (Nr1i2) and two paralogous Vdros proteins (Nr1i1a and Nr1i1b, denoted Vdros and Vdr\(β\)), only Vdros and Vdr\(β\) orthologs were identified in the cod genome. Furthermore, the cod genome holds additional orthologs to defensome-related NRs that are not identified in the zebrafish genome, including retinoic acid receptors (i.e. two nr1b2 genes, and one nr1b3), hepatocyte nuclear factor 4a (nr2a1), gluocorticoid receptor (nr3c1), and androgen receptor (nr3c4).
Searches in the second version of the Atlantic cod genome assembly (gadMor2) supported the absence of a pxr gene. Moreover, we made several attempts to amplify a pxr transcript from cod liver cDNA through polymerase chain reactions (PCR) using CODEHOP primers (Supplementary Table 3). These primers have previously

Figure 1. In silico searches using zebrafish pregnane x receptor (Pxr, 430 amino acids) show the loss of the nr1i2 gene in many of the Gadiform species, as well as other teleost fish. Searches were iterated using zebrafish major histocompatibility class II (Mhc II, 234 amino acids) as a control. Each row represents the resulting coverage vectors of the pxr or mhcii BLAST hits identified in the query amino acid sequence in 76 different fish species. The DNA-binding domain (DBD) and ligand-binding domain (LBD) in zebrafish Pxr are indicated. The Cv values are average coverage of each nucleotide in the assemblies used from Malmstrøm, et al.29 and Malmstrøm et al.70, and asterisk indicate use of genomes available from ENSEMBL. The species are placed into the same phylogeny as described in Malmstrøm, et al.29 and species belonging to the Gadiformes order are denoted in red.

Searches in the second version of the Atlantic cod genome assembly (gadMor2) supported the absence of a pxr gene. Moreover, we made several attempts to amplify a pxr transcript from cod liver cDNA through polymerase chain reactions (PCR) using CODEHOP primers (Supplementary Table 3). These primers have previously
been used to amplify \( pxr \) in other fish species, including Atlantic herring (\textit{Clupea harengus}), Atlantic wolfish (\textit{Anarhichas lupus}) and European eel (\textit{Anguilla anguilla}) (Supplementary Table 5). In line with the results from the bioinformatic searches, we were also able to successfully amplify \( pxr \) from the Gadiform species European hake (\textit{Merluccius merluccius}) (Supplementary Table 5). In contrast, none of the primers amplified sequences that matched \( pxr \) orthologs when using Atlantic cod cDNA as template.

To study the transcriptional effect on the biotransformation genes \( cyp3a \) and \( cyp1a \), we exposed precision-cut liver slices (PCLS) from Atlantic cod to a selected set of compounds known to activate mammalian and piscine PXR, and cod Ahr (BNF). The results show that the model zebrafish Pxr agonist, clotrimazole, had no significant effect on the transcription of \( cyp3a169 \) and \( cyp3a166 \), but it did induce transcription of \( cyp1a \) (\( p < 0.05 \)) (Fig. 6a). Transcription of all genes were slightly, but significantly (\( p < 0.05 \)) affected by the mouse PXR agonist, PCN, whereas the human PXR agonist, rifampicin, and the medaka Pxr agonist, B4HB, significantly (\( p < 0.05 \)) affected transcription of both \( cyp3a169 \) and \( cyp1a \) (Fig. 6a). Although slightly more induced than the \( cyp3a \) genes, the induction of \( cyp1a \) was much less following exposure to these compounds, compared to the exposure with the Ahr agonist, BNF (Fig. 6b). The highest dose of BNF also affected the transcription of both \( cyp3a \) genes, but the mean fold change remained low. Finally, we mapped the XRE and putative NR response elements (REs) in a 13,000 base-pair region upstream of the transcription start sites of human, mouse, zebrafish, and Atlantic cod \textit{CYP3A} and \textit{CYP1A} orthologs. These results show that the three active PXR-specific REs determined for human \textit{CYP3A4}\textsuperscript{34} are not conserved between these species, including mouse (Supplementary Fig. 2). The clusters of known XREs in human \textit{CYP1A1}, mouse \textit{Cyp1a1} and zebrafish \textit{cyp1a} orthologs\textsuperscript{35–37}, however, appear to be conserved and coincide position-wise. Importantly, whereas NR REs dominate the promoter regions of mammalian \textit{CYP3A} genes, the promoter regions of the fish \textit{cyp3a} orthologs have a higher number of potential XREs (Fig. 6c).

Figure 2. Schematic representation of the pregnane x receptor (\( pxr \)) syntenic regions in sequenced teleost genomes. The presence or absence of teleost \( pxr \) was confirmed in genomes where MYCBP-associated, testis expressed 1 (\textit{MAATS1}) and glycogen synthase kinase 3 beta (\textit{GSK3B}) were identified and assembled to the same scaffold. The \textit{Merluccius merluccius gsk3b} gene was located on a separate scaffold as \textit{maats1} and \textit{pxr}. Chr, chromosome; Sc, scaffold; LG, linkage group; Gr, group. The phylogeny is based on Malmstrøm, et al.\textsuperscript{29}.
Discussion

In the present work, we have demonstrated that multiple losses of pxr have occurred during teleost evolution. Through global genomic searches and in-depth analysis of gene loci of 76 teleost genome assemblies, we show that only 33 of these species have retained a pxr gene. Notably, our results show that within the Gadiformes order, only the Merluccidae family (hake) and Melanonus zugmayeri (Pelagic cod) have retained pxr, whereas the 23 other gadiform species represented have lost the gene. Based on current phylogeny of Malmstrøm, et al.29, Merluccidae and Melanonus zugmayeri are on separate branches deep within the Gadiformes order, and have arisen later in evolution than Bregmaceros cantori (striped codlet) but earlier than Macrouridae (grenadiers) and Muraenolepis marmorata (marbled moray cod), which all have lost the pxr gene. Thus, the presence of pxr in Merluccidae and Melanonus zugmayeri could be due to retention of an ancestral gene together with several independent losses in the other Gadiform species, or it could have been reintroduced through introgression. As we demonstrate through the conserved syntenic relationship of the genomic region, and the phylogenetic pattern of the LBD of Merluccidae and Melanonus zugmayeri Pxs, it is most likely that an ancestral gene is retained in these species. Moreover, based on current phylogeny, this means that multiple, independent losses have occurred in evolution of the Gadiformes order alone.

In other vertebrate species, PXR is an important xenobiotic receptor that governs much of the transcriptional response to chemical stressors. For instance, human PXR directly regulates expression of CYP3A isozymes, which are involved in the metabolism of more than 50% of all prescription medicines. The Cyp3a genes have been cloned...
and functionally characterized in several fish species\textsuperscript{39–45}, but, with the exception of zebrafish\textsuperscript{28}, the transcriptional regulation of teleost \textit{cyp3a} genes remains largely unexplored. However, exposure to prototypical Pxr agonists resulted in upregulated transcription of \textit{cyp3a} genes in Atlantic killifish (\textit{Fundulus heteroclitus})\textsuperscript{46–48}. Furthermore, expression patterns of the subfamilies \textit{cyp3b}, \textit{cyp3c}, and \textit{cyp3d} in ray-finned fishes (Actinopterygii) suggest involvement in xenobiotic metabolism, but this role has not yet been demonstrated\textsuperscript{11,49–51}. Thus, the widespread absence of \textit{pxr} orthologs in teleosts, and particularly among Gadiform species, raise interesting questions on how transcriptional regulation of biotransformation enzymes in these species has evolved.

By using Atlantic cod (\textit{Gadus morhua}) as a representative of a teleost without \textit{pxr}, we mapped the full complement of NR genes and show that Pxr is absent in the Nr1i subgroup. In line with the genome mining and synteny analysis, we were unable to amplify a \textit{pxr} gene from cod liver using CODEHOP primers. Notably, we were successful in using the primers to amplify a \textit{pxr} ortholog from European hake (\textit{Merluccius merluccius}), one of the few Gadiform species that have retained this gene.

Furthermore, we explored how transcription of cod \textit{cyp3a} and \textit{cyp1a} genes were affected by exposure to classical Pxr agonists. Exposure of cod PCLS to potent \textit{in vitro} activators of human, mouse, zebrafish, and medaka PXR/Pxr, including rifampicin, PCN, clotrimazole, and p-hydroxybenzoic acid, respectively\textsuperscript{52}, had little or no effect on \textit{cyp3a} transcript levels. Furthermore, although exposure to the established cod Ahr agonist BNF had a significant effect on \textit{cyp3a} transcription at the highest dose, the fold increase was very low compared to the BNF induction of \textit{cyp1a}. Similarly, in an early study of Atlantic cod, hepatic distribution of a CYP3A-like isozyme \textit{P450b} remained unaffected after \textit{in vivo} exposure to BNF\textsuperscript{53}. In mice, basal levels of \textit{Cyp3a11} expression was elevated approximately 4-fold in the livers of PXR knockouts relative to their wild-type counterparts\textsuperscript{54}, suggesting that PXR also has a repressing transcriptional role. We recently found a significant decrease in cod \textit{cyp3a} expression in PCLS following a 24-hour exposure to 10\textsubscript{\mu}M ethynylestradiol\textsuperscript{55}. This is in line with previous studies that indicate that xenoestrogens affect the expression and activity of Cyp3a and Cyp1a in Atlantic cod and Atlantic salmon (\textit{Salmo salar})\textsuperscript{56–57}, and suggests that the estrogen receptor (Er) signaling pathway also is involved in controlling \textit{cyp3a} expression in these species. Interestingly, the level of human ER proteins is partly regulated by activated AhR, which act as a ubiquitin ligase that selectively target sex steroid receptors for protein degradation\textsuperscript{58}.

In order to further decipher the transcriptional regulation of Atlantic cod \textit{cyp3a} and \textit{cyp1a} orthologs in the absence of Pxr, we mapped the response elements in the 13 000bp promoter region of these genes. The activated RXR/PXR heterodimer can bind several types of NR response elements, such as direct repeats (DR3, DR4,

\begin{figure}
\centering
\includegraphics[width=\textwidth]{Figure4.png}
\caption{Early, independent losses of pregnane x receptor (\textit{pxr}, \textit{nr1i2}) in the Gadiformes order. The hypothesized evolutionary losses of \textit{pxr} in Gadiform species are indicated by stars, leaving the gene retained only in the Merluccidae and Melanidae families. Stylephoriformes is used as an outgroup, showing that a loss of \textit{pxr} in this group precedes the multiple and independent losses in gadiformes lineage. The teleost illustrations by G. Holm and the phylogenetic relationship are both from Malmstrøm, \textit{et al.}\textsuperscript{29}.}
\end{figure}
Figure 5. Pregnane x receptor (Pxr, Nr1i2) is absent in the Nr1i subgroup of nuclear receptors in Atlantic cod. In contrast, zebrafish has both Pxr (Nr1i2) and vitamin D receptor (Vdr, Nr1i1) paralogs in the highlighted Nr1i subgroup. The nuclear receptor protein sequences identified by HMM and Pfam profiles in Atlantic cod (Gadus morhua, gm_ENSEMBL protein id) and zebrafish (Danio rerio, dr_ENSEMBL gene name) were aligned using ClustalX. Cloned sequences of cod Vdra and Vdrb were also included. The phylogenetic tree was made with MrBayes and drawn in FigTree. Nr0–6 indicate the seven different subfamilies of nuclear receptors.

and DR5), inverted repeats (IR6), and everted repeats (ER6 and ER8)\textsuperscript{59–62}. In humans, the active PXR response elements in the CYP3A4 promoter are known to include a distal enhancer module containing a DR3 motif, a xenobiotic-responsive enhancer module (XREM) containing a DR3 and an ER6 motif, and a proximal ER6\textsuperscript{63,64}. In addition, an HNF4α binding element (DR1/IR1) is important in transcriptional regulation of both human CYP3A4 and mouse Cyp3a11\textsuperscript{65,66}. Although the active PXR response elements are to some extent conserved between primates\textsuperscript{64}, we did not find these in mouse Cyp3a11 or fish cyp3a promoter regions. Notably, whereas the promoter regions of the mammalian CYP3A4 and Cyp3a11 were dominated by NR response elements, the promoter regions of the teleost orthologs also included Ahr-binding xenobiotic response elements (XRE). In fact, we found that the relative number of XREs in the promoter region of cod cyp3a orthologs are higher than in zebrafish cyp3a65, where it is demonstrated that Ahr are necessary to mediate effective transcription of this gene\textsuperscript{27}. 
**Figure 6.** Ex vivo liver exposure studies and promoter analysis suggest that in Atlantic cod (*Gadus morhua*), the aryl hydrocarbon receptor (Ahr) is involved in the transcriptional regulation of well-known pregnane x receptor (PXR) target genes. Precision-cut cod liver slices was exposed for 24 h to well-known ligands of PXR (a): clotrimazole (Clotr), pregnenolone 16α-carbonitrile (PCN), rifampicin (Rif), butyl 4-hydroxybenzoate (B4HB), and Ahr (b): β-naphthoflavone (BNF), before relative transcription of *cyp3a166*, *cyp3a169*, and *cyp1a* was measured using real-time quantitative PCR. Acidic ribosomal protein and ubiquitin was included as reference genes. The results are from three individual experiments, each with three technical replicates (n = 9). The different letters (a–c) indicate significant differences (p < 0.05) in gene transcription, as calculated on log2-transformed data in a Standard least square model combined with a Tukeys post hoc analysis. Relative composition of response elements (c) in 13 kb upstream regions of *CYP3A* and *CYP1A* orthologs. Species included are humans (hs), mouse (mm), zebrafish (*Danio rerio*, dr), and Atlantic cod (*Gadus morhua*, gm). Potential nuclear receptor response elements were revealed using NHR-Scan, and the AHR-binding xenobiotic response elements (XREs) were identified by searching for the consensus sequence ‘KNGCGTGM’.
Although the active XREs needed for transcription of human CYP1A1, mouse Cyp1a1, and zebrafish cyp1a are well described, we were not able to recognize the position of these sites in the cod ortholog. However, similar to the promoter regions of cod cyp3a166 and cyp3a169, we found that both the relative number of XREs, as well as the total number of response elements, in the cod cyp1a promoter region are higher than in the human, mouse, and zebrafish orthologs. Thus, transcriptional regulation of cod cyp3a genes appears more complex and may include extended coordinated signaling pathways and receptor crosstalk. This raises the question whether cod, and other fish lacking pXR, have a different response and sensitivity to environmental contaminants compared to fish that have retained the gene. When juvenile cod and rainbow trout (Onchorhyncus mykiss) were exposed in parallel to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) – a potent agonist for the aryl hydrocarbon receptor, cod liver accumulated higher concentrations compared to trout, whereas that hepatic Cyp1a1 activity was lower. Furthermore, exposing cod and turbot (Scophthalmus maximus) to two xenoestrogens, 4-nonylphenol and bisphenol A, showed indication of cod being the most sensitive and less specific in its response. The genomes of trout and turbot were not included in our study, but whereas pXR is previously identified in rainbow trout, we have not been able to clone this gene from turbot (unpublished results). Thus, it seems that there are inherent differences in how these fish species respond to chemicals that primarily interact with other receptor pathways, but, to our knowledge, a similar comparison aiming to activate the PXR-pathway has not yet been published.

In addition to its xenosensing role, PXR has frequently been linked to functions in the immune system of mammals, such as co-regulation of CYPs and cross-talk with the inflammation and immune responding transcription factor NF-κB. Interestingly, the initial sequencing of the Atlantic cod genome revealed the absence of several genes important for the adaptive immune system in other vertebrates, including the genes encoding the major histocompatibility complex (Mhc) II, Cd4 and invariant chain (Ii). Furthermore, low coverage genome sequencing of 66 teleost species demonstrated the absence of mhc ii across the entire Gadiformes order. However, whereas all Gadiformes, but no other fish genomes studied so far, lack mhc ii, our study show that the loss of pXR is not specific for the Gadiform order but has occurred several times during teleost evolution.

In conclusion, we present the first comprehensive evidence for evolutionary loss of the PXR in teleost fish. Importantly, the evolutionary pattern strongly suggests that multiple independent losses of this important xenobiatic receptor have occurred early in the Gadiformes order, as well as in several other teleost lineages. Furthermore, we suggest that the Ahr has evolved a regulatory role for cyp gene transcription in Atlantic cod. The loss of pXR, and possible evolved compensatory mechanisms for the lack of this gene, adds to the scientific interest in Atlantic cod as a toxicological model species. However, as several independent losses have occurred in the course of teleost evolution, other lineages may have adapted other compensating mechanisms, which accentuate the need for understanding how the chemical defenses have evolved in fish.

### Material and Methods

#### Searching 76 teleost species for a pXR gene.

The selection of teleost genomes was performed based on the most robust and recent multi-genome phylogeny available at the time of the study. This list includes 66 low-coverage genome assemblies, including 27 species from the Gadiformes order and a newly published cod genome assembly, and 10 other published teleost genome assemblies (listed in Supplementary Table 13). Performing tBLASTn searches, the nucleotide sequence of the fish genome assemblies was searched using the amino acid sequence of PXR from zebrafish (protein id. A5WYG8), medaka (Oryzias latipes, H2MUK7), tetraodon (Tetraodon nigroviridis, H3D8U6), and BXR (Bacillus subtilis) as query. To remove false positive BLAST hits that corresponded to the vitamin D receptor (Vdr, Nrl1), we removed any BLAST hits that also was listed as Vdr from the same fish species (i.e. zebrafish Vdr: Q9PTN2, medaka Vdr: H3CJ29) at a lower e-value. The output file of each search defined the coordinates (amino acid number for start and end) of the identified nucleotide segments (blast hits) in relation to the query PXR sequence used. Based on these coordinates, it was counted how many times every position in the query PXR sequence was overlapped by the identified segments, providing a coverage vector. The resulting coverage vectors of the BLAST hits identified from the PXR of each sequence were assembled and visualized as a heat map, guided by the phylogenetic tree published in Malmsfet et al. In these heat maps, one cell represents one amino acid in the PXR sequence (x-axis) identified in the 76 different fish species (y-axis).

As a positive control of this approach, we applied this pipeline to search for the mhc ii gene in the 76 fish genome assemblies, whose phylogenetic distribution in teleosts is recently characterized.

Based on their ligand binding domains, sequences identified in the bioinformatic searches using zebrafish, medaka, and tetraodon PXR as query were aligned with annotated or cloned teleost PXR (Nrl1), Vdr (Nrl2), retinoic acid receptor (RARγ, Nrlb) and Rev-ERBα (Nrl1d), using Clustal Omega v1.2.2 with default parameters. Due to incomplete sequences in the LBD, Cyp450, Cytochrome P450, and Halocentrus rupus was excluded from the alignment. The maximum likelihood phylogeny was analyzed in RAxML v.8.1.14 using the CATLG model, with 200 bootstrap replicates, and drawn in FigTree v1.4.2. The figure was prepared in Adobe Illustrator CS5.

#### Syntenic relationship analysis.

The genomic region containing the mammalian PXR gene is conserved, with MAATS1 located immediately downstream and GSK3B immediately upstream of PXR. We used BLAST and Exonerate (EMBL-EMI) to search for these genes in 76 teleost genome assemblies. Since the searches were performed in relatively fragmented genome assemblies, synteny was only revealed in the assemblies where the genes were located on the same scaffold. The phylogenetic relationships between the different teleost species were derived from Malmsfet et al.
Identifying nuclear receptors in the Atlantic cod genome. We identified the cod NR proteins in the Atlantic cod protein database (ENSEMBL-Gadus_morhua.gadMor.1.73, based on gadMor1 annotated proteins) by hidden Markov model (HMM) searches using HMMER v3.1b1 (www.hmmer.org) and the Pam profiles ‘Hormone_recep’ Pf00104.25 and ‘sf-C4’ Pf00105.13. A Python script (available at https://doi.org/10.6084/m9.figshare.5752791) was used to perform the HMM searches, which were followed by reciprocal BLAST searches against the well-annotated zebrafish (Danio rerio) proteome (ENSEMBL-Danio_rerio.Zv9.73) to predict orthologous proteins. Based on the ENSEMBL protein identifiers, the corresponding gene identifiers were identified in the ENSEMBL databases and are presented in Supplementary Table 2.

Further, specific blastn searches were also performed in the second version of the cod genome assembly (gadMor2; available at http://cees-genomes.hpc.uio.no/blast/33), using zebrafish Pxr (protein id. A5WYG8-1) as query sequence. The resulting gene entries were extracted and orthologs identified by reciprocal blastp searches in the UniProtKB database.

Phylogenetic analyses of the NR-superfamily. A multiple sequence alignment of the Atlantic cod and zebrafish nuclear receptor proteins (available at https://doi.org/10.6084/m9.figshare.5752812) was constructed using ClustalX 2.1. Based on the resulting alignment, a phylogenetic tree was constructed using MrBayes v3.2.3 with the following settings: the prior for the amino acid model was set to mixed, and the number of generations used was 100 0000. FigTree v1.3.1 was used to visualize the tree, and the figure was prepared in Adobe Illustrator. For genes with predicted multiple protein sequences (splice variants), only the longest translation was included to ensure a non-redundant sequence set.

Amplifying pxr from cod liver with consensus-degenerate hybrid oligonucleotide primers. Total RNA was extracted from liver tissue from two different cods using phenol and guanidinium isothiocyanate solution (Trizol, Thermo Fisher Scientific), before Superscript II Reverse transcriptase was used to synthesize cDNAs from total RNA as described by the supplier (Invitrogen). Consensus-degenerate hybrid oligonucleotide (CODEHOP) primers (Supplementary Table 3) were designed based on known Pxr sequences from eight fish species and human PXR (Supplementary Table 4) using the CODEHOP web tool11. The PCRs were performed using GoTaq DNA polymerase according to the protocol from the supplier (Promega) as gradient PCRs (Tanneur 46–60 °C).

Ex vivo exposure to compounds known to activate mammalian and piscine PXR. Atlantic cod (Gadus morhua) was provided from Fjord Gadus AS (Fiskå, Norway), and maintained at The Industrial and Aquatic Research Laboratory (ILAB, Bergen, Norway). The fish were kept in 500 L tanks in natural seawater at 9 °C with a 12:12 hour light cycle regime. The fish were fed with a commercial diet ad libitum (EWOS, Bergen, Norway). Six immature cod of mixed sex, weighing 840–1549 g, was used in these experiments.

The fish were handled by FELASA C (Federation of European Laboratory Animal Science Association) approved personnel and killed according to the directions from the Norwegian Food Safety Authority (FOTS) § 16, second paragraph. Preparation of cod precision-cut liver slices (PCLS) was performed as described previously10. Freshly prepared cod PCLS were exposed for 24 hours to prototype PXR/Pxr-ligands described from other species, including clotrimazole, pregnenolone 16α-carbonitrile (PCN), rifampicin, and p-hydroxy benzoic acid (B4HB)12, as well as the well-known cod AhR-ligand and cyp1a-inducer β-naphthoflavone (BNF). Chemicals were purchased from Sigma-Aldrich (Oslo, Norway) and dissolved in Hybri-Max™ dimethyl sulfoxide (DMSO, Sigma-Aldrich) to final concentrations of 0.25, 2.5, and 25μM for PXR/Pxr-agonists, and 0.1, and 10μM for BNF. 0.01% DMSO was used as solvent control in all exposures. Cytotoxicity of the chemicals was monitored by measuring the release of lactate dehydrogenase (LDH; EC 1.1.1.27) into the culture medium, using the LDH Cytotoxicity Detection Kit (Roche Applied Sciences, Basel, Switzerland) according to the manufacturer’s protocol.

Total RNA was isolated from PCLS using the RNeasy Mini Kit (QIAGEN, Hilden, Germany). RNA concentration and RNA quality were assessed with a NanoDrop ND-1000 and agarose gel electrophoresis, and RNA was subsequently reverse transcribed to cDNA by either SuperScript III Reverse Transcriptase (Invitrogen, Thermo Fisher Scientific, Oslo, Norway) or qScript cDNA SuperMix (Quanta BioSciences, Gaithersburg, USA) following the producer’s protocol. The cDNA was diluted 1:30 in H2O prior to RT-qPCR using SYBR Green I Master (Roche) and the LightCycler® 480 Real-Time PCR System (Roche), according to the manufacturer’s procedure. Both a non-template control (H2O) and a positive control (calibrator cDNA) were included for each gene, and considering the efficiency of the reaction for each primer pair. Acidic ribosomal protein (arp and considering the efficiency of the reaction for each primer pair. Acidic ribosomal protein (arp and basal gene expression in individual fish is bypassed. Figures were prepared in Prism 6 (GraphPad), where the boxes represent the 25th to 75th percentile with the median indicated as a line and whiskers are drawn from the lowest to the highest value. Statistical significance (p < 0.05) between groups (control and two or three exposures groups) are marked by letters a, b, or c.

Promoter analysis of CYP3A and CYP1A orthologs. Four orthologs to the zebrafish cyp3a65 gene were identified in the ENSEMBL cod genome (gadMor1) (Eide et al., in prep). These were confirmed in the
new cod genome assembly and provided official names by the Cytochrome P450 Nomenclature Committee: cyp3a166 (location in gadMor2: LG03, base pairs 18402807-18400860), cyp3a168P (LG03, base pairs 18384880-18386508), and cyp3a169 (LG01, base pairs 1472701-14723653). Genomic regions containing 13,000 base pairs of upstream sequences of cod cyp1a1 (LG14, base pairs 2303196-2307638), cyp3a166, and cyp3a169 were retrieved from the new genome assembly. The corresponding promoter regions of human CYP3A4 (ENSG00000160868), CYP1A1 (ENSG00000140465), and CYP1A2 (ENSG00000140505); mouse Cyp3a11 (ENSMUSG00000056035), Cyp1al (ENSMUSG0000032315), and Cyp1a2 (ENSMUSG00000032310); and zebrafish cyp3a65 (ENSDARG00000103295), and cyp1a (ENSDARG00000098315), were all derived from ENSEMBL.

Potential nuclear receptor response elements in the upstream regions were predicted using the nuclear hormone receptor (NHR)-scan available online. NHR-scan uses a HMM framework for describing binding behavior according to three 'match state chains', corresponding to site configurations of direct repeats, inverted repeats, and everted repeats, compared to a background state. We used the default settings with 0.01 probability of entering each of the match states. The extended Ahr-binding XREs were mapped by searching the sequences for the consensus sequence ‘KNGCGTGM’ in forward, reverse and complementary directions. The resulting response elements were visualized in Adobe Illustrator CS5 (v15.0.0), and horizontal slice diagrams of binding site distributions were prepared in Prism 6 (GraphPad).

**References**

1. Near, T. J. et al. Resolution of ray-finned fish phylogeny and timing of diversification. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 13698–13703 (2012).
2. Nelson, J. S. Fishes of the world, six, 601 p. (John Wiley, Hoboken, N.J., 2006).
3. Goldstone, J. V. et al. The chemical defenses: Environmental sensing and response genes in the Strongylocentrotus purpuratus genome. *Developmental Biology* **300**, 366–384 (2006).
4. Goldstone, J. V. Environmental sensing and response genes in cnidaria: the chemical defenses in the sea anemone *Nematostella vectensis*. *Cell Biology and Toxicology* **24**, 483–502 (2008).
5. Stegeman, J. J., Goldstone, J. V. & Hahn, M. E. Perspectives on zebrafish as a model in environmental toxicology. In *Fish physiology: Zebrafish* (Elsevier, 2010).
6. van Denderen, P. D., Lindegren, M., MacKenzie, B. R., Watson, R. A. & Andersen, K. H. Global patterns in marine predatory fish. *Nat Ecol Evol* **2**, 65–70 (2018).
7. Star, B. et al. The genome sequence of Atlantic cod reveals a unique immune system. *Nature* **477**, 207–210 (2011).
8. Hasselberg, L., Meier, S., Svardal, A., Hegelund, T. & Celander, M. C. Effects of alkylphenols on CYP1A and CYP3A expression in the developing zebrafish. *Arch Toxicol Suppl* **8**, 13698–13703 (2012).
9. Goksøyr, A., Solbakken, J. E. & Klungsøyr, J. Regioselective metabolism of phenanthrene in Atlantic cod (*Gadus morhua*). *Biochimica et Biophysica Acta* **1792**, 1359–1367 (2009).
10. Eide, M., Karlsen, O. A., Kryvi, H., Olsvik, P. & Goksøyr, A. Precision-cut liver slices of Atlantic cod (*Gadus morhua*): An in vitro system for studying the effects of environmental contaminants. *Aquatic Toxicology* (2014).
11. Olsvik, P. A. et al. Are Atlantic Cod in Store Lungegardsvann, a Seawater Recipient in Bergen, Affected by Environmental Contaminants? A qRT-PCR Survey. *Journal of Toxicology and Environmental Health Part A: Current Issues* **74**, 547–554 (2011).
12. Goksøyr, A. Purification of hepatic microsomal cytochromes P-450 from ß-naphthoflavone-treated Atlantic cod (*Gadus morhua*), a marine teleost fish. *Biochimica et Biophysica Acta* **840**, 409–417 (1985).
13. Goksøyr, A. et al. Species characteristics of the hepatic xenobiotic and steroid biotransformation systems of 2 teleost fish, Atlantic cod (*Gadus morhua*) and Rainbow trout (*Salmo-Gairdneri*). *Toxicology and Applied Pharmacology* **89**, 347–360 (1987).
14. Goksøyr, A., Solbakken, J. E. & Klyngstad, R. J. Regioselective metabolism of phenanthrene in Atlantic cod (*Gadus morhua*) - Studies on the effects of monooxygenase inducers and role of cytochrome P-450. *Chemico-Biological Interactions* **60**, 247–263 (1986).
15. Goksøyr, A., Serigstad, R., Solberg, T. S. & Stegeman, J. J. Response of cod (*Gadus morhua*), larvæ and juveniles to oil exposure detected with anti-cod cytochrome P-450 IgG and anti-scup cytochrome P-450 MAbs 1-12-3. *Aquatic Toxicology* **24**, 31–38 (1991).
16. Olsvik, P. A., Puntervoll, P. & Goksøyr, A. Mass spectrometric analyses of microsomal cytochrome P450 isoforms isolated from beta-naphthoflavone-treated Atlantic cod (*Gadus morhua*). *Molecular Endocrinology* **18**, 2–10 (2014).
17. Germain, P., Staels, B., Ducatteau, C., Spedding, M. & Laubet, V. Overview of nomenclature of nuclear receptors. *Pharmacological Reviews* **58**, 685–704 (2006).
18. Ekins, S., Reschly, E. J., Hagey, L. R. & Krasowski, M. D. Evaluation of pharmacologic specificity in the pregnant X receptor. *Bmc Evolutionary Biology* **9** (2008).
19. Whitfield, G. K. et al. Cloning of a functional vitamin D receptor from the lamprey (*Petromyzon marinus*), an ancient vertebrate lacking a calcified skeleton and teeth. *Endocrinology* **144**, 2704–2716 (2003).
20. Zhao, Y. B., Zhang, K., Giesy, J. P. & Hu, J. Y. Families of Nuclear Receptors in Vertebrate Models: Characteristic and Comparative *Toxicological Perspective*. *Scientific Reports* **5** (2015).
21. Handschin, C., Podvinec, M. & Meyer, U. A. CXR, a chicken xenobiotic-sensing orphan nuclear receptor, is related to both mammalian pregnane X receptor (PXR) and constitutive androstane receptor (CAR). *Proceedings of the National Academy of Sciences of the United States of America* **97**, 10769–10774 (2000).
22. Mathias, M. et al. Evolutionary History and Functional Characterization of the Amphibian XenoSensor CAR. *Molecular Endocrinology* **26**, 14–26 (2012).
23. Bertrand, S. et al. Unexpected novel relational links uncovered by extensive developmental profiling of nuclear receptor expression. *Plos Genetics* **3**, 2085–2100 (2007).
24. Moore, L. B. et al. Pregnan X Receptor (PXR), constitutive androstane receptor (CAR), and benzox A receptor (BXR) define three pharmacologically distinct classes of nuclear receptors. *Molecular Endocrinology* **16**, 977–986 (2002).
25. Goksøyr, A. Use of cytochrome P450 1A (CYP1A) in fish as a biomarker of aquatic pollution. *Arch Toxicol Suppl* **17**, 80–95 (1995).
26. Tseng, H. P., Hseu, T. H., Buhlker, D. R., Wang, W. D. & Hu, C. H. Constitutive and xenobiotics-induced expression of a novel CYP3A gene from zebrafish larva. *Toxicology and Applied Pharmacology* **205**, 247–258 (2005).
27. Chang, C. T. et al. Regulation of zebrasfish CYP3A65 transcription by AHR2. *Toxicology and Applied Pharmacology* **270**, 174–184 (2013).
28. Kubota, A. et al. Role of Pregnan X Receptor and Aryl Hydrocarbon Receptor in Transcriptional Regulation of pnr, CYP2, and CYP3 Genes in Developing Zebrafish. *Toxicological Sciences* **143**, 398–407 (2015).
29. Malmström, M. et al. Evolution of the immune system influences speciation rates in teleost fishes. *Nature Genetics* **48**, 1204–1210 (2016).
43. Ma, J. G., Bu, Y. Z., Li, Y., Niu, D. C. & Li, X. Y. Complementary DNA Cloning and Functional Characterization of Cytochrome P450

45. Celander, M. & Stegeman, J. J. Isolation of a cytochrome P450 3A cDNA sequence (CYP3A30) from the marine teleost

59. Aranda, A. & Pascual, A. Nuclear hormone receptors and gene expression.

65. Liu, F. J., Song, X., Yang, D., Deng, R. & Yang, B. The far and distal enhancers in the CYP3A4 gene co-ordinate the proximal promoter

67. Hines, R. N., Mathis, J. M. & Jacob, C. S. Identification of Multiple Regulatory Elements on the Human Cytochrome P450ia1 Gene.

69. Larsen, B. K. Comparison of protein expression in plasma from nonylphenol and bisphenol A-exposed Atlantic cod (Gadus morhua) and turbot (Scophthalmus maximus) by use of SELDI-TOF. Aquatic Toxicology 78, S25–S33 (2006).

41. Vaccaro, E. et al. Cloning, tissue expression, and inducibility of CYP 3A79 from sea bass (Dicentrarchus labrax). Journal of Biochemical and Molecular Toxicology 21, 32–40 (2007).

42. Christen, V., Caminada, D., Arand, M. & Fent, K. Identification of a CYP3A form (CYP3A126) in fathead minnow (Pimephales promelas) and characterisation of putative CYP3A enzyme activity. Analytical and Bioanalytical Chemistry 396, 585–595 (2010).

54. Staudinger, J., Liu, Y. P., Madan, A., Habeebu, S. & Klaassen, C. D. Coordinate regulation of xenobiotic and bile acid homeostasis by...
Acknowledgements
The work was supported by the Research Council of Norway (RCN) as part of the iCOD and iCOD 2.0 (grant no. 192441/I30 and 244654/E40) to AG and OA, dCod 1.0 (248840 projects) to AG, and the National Institute of Health (USA) NIH P42ES007381 (Boston University Superfund Center), and NIH R21HD073805 grants to KSJ. The sequencing towards the new Atlantic cod reference genome was provided by the Norwegian Sequencing Centre (www.sequencing.uio.no), a national technology platform hosted by the University of Oslo and supported by the “Functional Genomics” and “Infrastructure” programs of the Research Council of Norway and the Southeastern Regional Health Authorities.

Author Contributions
M.E., O.A.K., A.G., and O.K.T. conceived the study. M.E. performed the multiple sequence alignment of Pxr sequences and built the Pxr phylogenetic tree, executed bioinformatical searches in ‘gadMor1’, planned and supported the ‘gadMor2’ and in the 75 other teleost genomes, retrieved the Pxr sequences, and performed synteny analysis. R.L.L. performed the cloning. P.P. designed and coded the Python script, performed the multiple sequence alignment of NRs, and built the NR phylogenetic tree. O.A.K., A.G., J.V.G. and J.S. participated in the conception of the study and drafting the manuscript. All authors contributed to, read, and approved the final manuscript.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-28498-4.

Competing Interests: The authors declare no competing interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018