Marine Genomics 23 (2015) 99–108

Contents lists available at ScienceDirect

Marine Genomics

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

Tunicate pregnane X receptor (PXR) orthologs: Transcript characterization and natural variation

Ingrid Richter [a,b,*], Andrew E. Fidler [a,c,d]

[a] Coastal and Freshwater Group, Cawthron Institute, Private Bag 2, Nelson 7012, New Zealand
[b] School of Biological Science, Victoria University of Wellington, PO Box 600, Wellington 6140, New Zealand
[c] Maurice Wilkins Centre for Molecular Biodiscovery, University of Auckland, Auckland 1142, New Zealand
[d] Institute of Marine Science, University of Auckland, Auckland 1142, New Zealand

ARTICLE INFO

Article history:
Received 16 February 2015
Received in revised form 6 May 2015
Accepted 6 May 2015
Available online 16 May 2015

Keywords:
Ciona intestinalis
Botryllus schlosseri
Nuclear receptor
SNPs
Xenobiotics
Marine Bioactives

ABSTRACT

Vertebrate pregnane X receptor (PXR, NR1I2), a ligand-activated nuclear receptor (NR), regulates expression of detoxification genes. Vertebrate PXR orthologs may adaptively evolve to bind deleterious/toxic xenobiotics typically encountered by organisms from their diet. Tunicates (phylum Chordata) are marine filter-feeders that form a sister clade to the Vertebrata. Genomes of two tunicate taxa, *Ciona intestinalis* and *Botryllus schlosseri*, encode at least two PXR orthologs (abbreviated VDR/PXRα and β). Here we report characterization of the transcript structures and sequence variation of three tunicate PXR orthologs: *C. intestinalis* VDR/PXRα and β, and *B. schlosseri* VDR/PXRα. The three predicted proteins consist of both DNA-binding (DBD) and ligand-binding (LBD) domains typical of NRs. The *C. intestinalis* VDR/PXRβ LBD may be significantly larger than that of the VDR/PXRα orthologs. In both tunicate taxa, the mRNAs were characterized by high frequencies of single nucleotide polymorphisms (SNPs, ca. 3 SNPs/100 base pairs). The majority of SNPs were synonymous and standard tests (Tajima's *D*, d_N/d_S ratios) indicated strong purifying selection. However, one base pair frameshift allelic variants were found in the *C. intestinalis* VDR/PXRα and β genes. The predicted proteins consisted of a DBD but lacked an LBD. The persistence of these variants may possibly reflect constitutive expression of detoxification genes as a selective advantage in the marine environment. These results provide a foundation for further investigations into the molecular evolution, population genetics and functioning of tunicate receptors involved in detection of marine bioactive compounds.

1. Introduction

Animals, particularly herbivores and omnivores, may ingest dietary chemicals that can alter their physiology and, ultimately, their selective fitness (Targett and Arnold, 2001; Raubenheimer and Simpson, 2009; Forbey et al., 2013). Metazoan genomes encode a range of sensors to detect such "xenobiotic" chemicals (from the Greek *xenos*: foreigner; *bios*: life); for example, G protein-coupled receptors that mediate taste and olfaction (Palmer, 2007; Kinnamon, 2012). Xenobiotic detection pathways operating independently of the nervous system are coordinated through ligand-activated nuclear receptors (NRs) which influence transcription of genes involved in xenobiotic detoxification (Nakata et al., 2006; Chai et al., 2013; Wallace and Redinbo, 2013). Thus, transcription of some detoxification genes is induced by those xenobiotic(s) that the pathway ultimately detoxifies (Xu et al., 2005; Yang et al., 2010; Testa et al., 2012; James and Ambadapadi, 2013). Xenobiotic-induced control of gene transcription may evolve adaptively to provide appropriate responses to the dietary chemical profiles associated with different animal diets (Mao et al., 2009; Whalen et al., 2010; Johnson et al., 2012).

Vertebrate genomes encode two NRs that modulate transcription of detoxification genes, constitutive androstane receptor (CAR; NR notation: NR1I3) and pregnane X receptor (PXR; NR notation: NR1I2; Kliewer et al., 2002; Kachaylo et al., 2011; Chai et al., 2013; Wallace and Redinbo, 2013). Of these two NRs, PXR is better understood with regards to how its ligand-binding domain (LBD) structure relates to ligand-binding and the associated permissiveness with respect to ligand

Abbreviations: AF-1, transcription activation function 1; AR, androgen receptor; *Bs*VDR/PXRα, *Botryllus schlosseri* VDR/PXR ortholog α; BSA, bovine serum albumin; bp, base pair; CAR, constitutive androstane receptor; *Ci*VDR/PXRα, *Ciona intestinalis* VDR/PXR ortholog α; *Ci*VDR/PXRβ, *Ciona intestinalis* VDR/PXR ortholog β; DBD, DNA-binding domain; d_N/d_S, ratio of non-synonymous substitutions per non-synonymous site (d_N) to synonymous substitutions per synonymous site (d_S); ER, estrogen receptor; FXR, farnesoid X receptor; LBD, ligand-binding domain; LXR, liver X receptor; NCBI CD, National Centre for Biotechnology Information Conserved Domain database; NR, nuclear receptor; NR1I, nuclear receptor subfamily 1 class I; PAR, pregnenolone-activated receptor; PCR, polymerase chain reaction; Pfam, Protein Families database; π, nucleotide diversity; PXR, pregnane X receptor; qPCR, quantitative polymerase chain reaction; SMART, Simple Modular Architecture Research Tool; SNPs, single nucleotide polymorphisms; SXR, steroid and xenobiotic receptor; D_T, Tajima's *D*; Ti/Tv, transition/transversion ratio; VDR, vitamin D receptor; θ_W, Watterson's Theta.

* Corresponding author at: Coastal and Freshwater Group, Cawthron Institute, Private Bag 2, Nelson 7012, New Zealand. Tel.: + 64 3 548 2319; fax: + 64 3 546 9464.
E-mail address: Ingrid.Richter@cawthron.org.nz (I. Richter).

http://dx.doi.org/10.1016/j.margen.2015.05.006
1874-7787/

chemical structure (Chai et al., 2013). Vertebrate PXR ligands include a structurally diverse range of endogenously produced molecules (e.g. bile acids, steroid hormones and vitamins) along with exogenously acquired chemicals including both synthetic drugs and natural bioactive compounds (Chang and Waxman, 2006; Staudinger et al., 2006; Manez, 2008; Biswas et al., 2009; Hernandez et al., 2009; Zhou et al., 2009). Vertebrate PXRs display greater inter-taxa differences in LBD sequences than is typical of NRs, along with some evidence of positive selection within the LBD (Moore et al., 2002; Zhang et al., 2004; Krasowski et al., 2005a, 2005b). These observations have led to speculations that PXR LBD inter-taxa sequence differences may reflect adaptive evolutionary changes that enhance binding of xenobiotics typically encountered by an organism (Moore et al., 2002; Zhang et al., 2004).

Tunicates (Phylum Chordata) are marine filter-feeding invertebrates that form the sister clade to the Vertebrata (Delsuc et al., 2006, 2008; Putnam et al., 2008; Shenkar and Swalla, 2011). Filter-feeding marine invertebrates ingest small organisms filtered from seawater and therefore are exposed to myriads of xenobiotics associated with their diet (Sekiguchi et al., 2001; Echevarria et al., 2012; Roje-Busatto and Ujević, 2014). Those dietary xenobiotics that are toxic to tunicates (i.e. increase mortality/morbidity and/or reduce reproductive output) are likely to act as agents of natural selection — with xenobiotic receptor LBDs expected to evolve to include these natural toxins in their ligand repertoire. The genomes of two ascidian tunicates, *Ciona intestinalis* (*Ci*) and *Botryllus schlosseri* (*Bs*), encode at least two genes considered orthologous to the vertebrate PXR, CAR and vitamin D receptor (VDR), abbreviated as VDR/PXR (Dehal et al., 2002; Yagi et al., 2003; Voskoboynik et al., 2013). The tunicate VDR/PXRα LBDs, expressed in both mammalian cells and yeast, displayed ligand-binding characteristics consistent with a role in detecting marine xenobiotics, most notably microalgal biotoxins (Fidler et al., 2012; Richter and Fidler, 2014, 2015).

Previously, suggestions that positive selection acts on PXR LBD coding sequences were based on sequence comparisons between vertebrate PXR orthologs (Zhang et al., 2004; Krasowski et al., 2005b). Furthermore, physiological differences between humans and rodents in xenobiotic responses may be explained by differing PXR ligand affinities (Blumberg et al., 1998; Kliewer et al., 1998; Jones et al., 2000; LeCluyse, 2001; Tirona et al., 2004). Here we report intra-species sequence variation in VDR/PXR LBD coding sequences from two tunicate species, *C. intestinalis* and *B. schlosseri*, and examine the data for evidence of adaptive evolution and functional significance.

2. Materials and methods

2.1. Tunicate collection and tissue sampling

C. intestinalis adults (>60 mm length, $n = 30$) were collected from four New Zealand South Island locations: (i) Nelson marina (41° 15′ 32.64″ S, 173° 16′ 55.53″ E, $n = 18$); (ii) Lyttelton Harbour (43° 36′ 38.63″ S, 172° 42′ 14.35″ E, $n = 3$); (iii) Pelorus Sound (41° 12′ 54.87″ S, 173° 52′ 46.91″ E, $n = 6$); and (iv) Queen Charlotte Sound (41° 12′ 31.73″ S, 174° 17′ 59.91″ E, $n = 3$). New Zealand *B. schlosseri* colonies ($n = 30$) were collected from two locations: Nelson marina ($n = 23$) and Lyttelton Harbour ($n = 7$). *B. schlosseri* colonies ($n = 10$) from the eastern Mediterranean Coast were collected from a single location: Michmoret beach, Ruppin Academic Centre, Israel (32° 24′ 31.85″ N, 34° 52′ 05.37″ E). *B. schlosseri* colonies were collected >1 m apart to reduce the chances of sampling clonally related colonies. *C. intestinalis* individuals were anesthetized overnight in seawater containing traces of menthol crystals to assist with the subsequent dissection of tissues (Hanashima et al., 2012). Dissected *C. intestinalis* tissues were frozen on dry ice before storage at −70 °C. *B. schlosseri* colonies from New Zealand waters were frozen on dry ice before storage at −70 °C. *B. schlosseri* colonies collected in the Mediterranean were placed in a RNA stabilization buffer (http://sfg.stanford.edu/RNAbuffer.pdf) and transported at ambient temperatures to New Zealand before long term storage at −70 °C.

2.2. Amplification and Sanger sequencing of cDNA sequences

Total RNA was isolated from either dissected *C. intestinalis* gut tissues or *B. schlosseri* colony fragments using Trizol (Life Technologies, Carlsbad, CA, U.S.A.) following the manufacturer's protocol. First strand cDNAs were synthesized from total RNA using random hexamer primers (Transcriptor First Strand cDNA Synthesis Kit, Roche Diagnostics, Penzberg, Germany) and used as templates in polymerase chain reactions (PCRs). Primers were designed to amplify three tunicate VDR/PXR partial coding sequences as follows. PCR primers flanking the predicted DBD and LBD coding sequences of *C. intestinalis* VDR/PXRα (*Ci*VDR/PXRα, GenBank acc. no. NM_001078379) and *C. intestinalis* VDR/PXRβ (*Ci*VDR/PXRβ, NM_001044366) were designed directly from the GenBank sequences (Table S1A). When this work was started, sequencing of the *B. schlosseri* genome was still in progress, thus complicating PCR primer design. Primers for amplifying the *B. schlosseri* VDR/PXRα coding sequence were developed using a combination of alignments of PXR sequences and *B. schlosseri* expressed sequence tag (EST) sequence data. The predicted DBD protein sequences of PXR orthologous from mouse (AF031814), rat (AF151377), rabbit (AF188476), human (AF061056), frog (AF305201), chicken (AF276753) and *C. intestinalis* (NM_001078379) were aligned. A conserved peptide sequence (CEGCKGFFR) was identified and a redundant forward primer designed from an alignment of the corresponding DNA sequences (Table S1A). This forward primer was paired with a reverse primer designed directly from two *B. schlosseri* EST sequences (contig-18161978, contig-18290615, A. Voskoboynik, pers. comm.) identified as encoding a VDR/PXRα ortholog (Table S1A). PCRs were carried out in 20.0 μl final volumes containing: 10.0 μl of high-fidelity *Taq* DNA polymerase (SAHARA™ DNA Polymerase Master Mix, Bioline, London, U.K.), forward and reverse primers (both 0.4 μM), 32.0 μg/ml bovine serum albumin (Sigma-Aldrich, St. Louis, MO, U.S.A.) and 2.0 μl of template cDNA. All three partial coding sequences (*Ci*VDR/PXRα, *Ci*VDR/PXRβ and *Bs*VDR/PXRα) were amplified using the following thermocycling conditions: 94 °C/10 min, 1 cycle; 94 °C/30 s, 45 °C/30 s ramping at 0.2 °C/s to 72 °C/2 min, 5 cycles; 94 °C/30 s, 55 °C/30 s, 72 °C/2 min, 35 cycles; 72 °C/7 min, 1 cycle; 15 °C/hold. The resulting PCR products were visualized on an ethidium bromide stained agarose gel and then extracted (Zymoclean™ Gel DNA Recovery Kit, Zymo Research, Irvine, CA, U.S.A.). Purified amplicons were ligated into pGEM-Teasy (Promega, Madison, WI, U.S.A.), transformed into DH5α *Escherichia coli* and plasmids were purified (High Pure Plasmid Isolation Kit, Roche Diagnostics). Plasmid inserts were bi-directionally sequenced by an external contractor (Massey Genome Service, Massey University, New Zealand). A representative sequence from a single cloned haplotype of each of the three transcripts was deposited on GenBank with the following accession numbers: *Ci*VDR/PXRα (1326 bp): KC561370, *Ci*VDR/PXRβ (1569 bp): KC561371 and *Bs*VDR/PXRα (1006 bp): KC561372.

2.3. Protein analysis software

Structural domains within protein sequences were identified using the Simple Modular Architecture Research Tool (SMART) software (Letunic et al., 2012; http://smart.embl-heidelberg.de), the National Center for Biotechnology Information (NCBI) Conserved Domain (CD) database (Marchler-Bauer et al., 2011; http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and the Protein Families (Pfam) database (Punta et al., 2012; http://pfam.xfam.org).