General and Comparative Endocrinology 295 (2020) 113527

Contents lists available at ScienceDirect

General and Comparative Endocrinology

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

Research paper

Cartilaginous fishes offer unique insights into the evolution of the nuclear receptor gene repertoire in gnathostomes

Elza Fonseca[a,b], André M. Machado[a], Nair Vilas-Arrondo[c,d], André Gomes-dos-Santos[a,b], Ana Veríssimo[b,e], Pedro Esteves[b,d], Tereza Almeida[b,e], Gonçalo Themudo[a], Raquel Ruivo[a], Montse Pérez[c], Rute da Fonseca[f], Miguel M. Santos[a,b], Elsa Froufe[a], Esther Román-Marcote[c], Byrappa Venkatesh[g,*], L. Filipe C. Castro[a,b,*]

[a] CIIMAR/CIMAR – Interdisciplinary Centre of Marine and Environmental Research, U.Porto, 4450-208 Matosinhos, Portugal
[b] FCUP – Faculty of Sciences, Department of Biology, U.Porto, 4169-007 Porto, Portugal
[c] AQUACOV, Instituto Español de Oceanografía, Centro Oceanográfico de Vigo, 36390 Vigo, Spain
[d] UVIGO, phD Program "Marine Science, Tehchology and Management" (Do *MAR), Faculty of Biology, University of Vigo, 36200 Vigo, Spain
[e] CIBIO – Research Center in Biodiversity and Genetic Resources, InBIO, Associate Laboratory, U.Porto, 4485-661 Vairão, Portugal
[f] Center for Macroecology, Evolution and Climate, GLOBE Institute, University of Copenhagen, Denmark
[g] Comparative Genomics Laboratory, Institute of Molecular and Cell Biology, A*STAR (Agency for Science, Technology and Research), Biopolis, Singapore 138673, Singapore

ARTICLE INFO

Keywords:
Nuclear receptors
Genome
Gene loss
Gene duplication

ABSTRACT

Nuclear receptors (NRs) are key transcription factors that originated in the common ancestor of metazoans. The vast majority of NRs are triggered by binding to either endogenous (e.g. retinoic acid) or exogenous (e.g. xenobiotics) ligands, and their evolution and expansion is tightly linked to the function of endocrine systems. Importantly, they represent classic targets of physiological exploitation by endocrine disrupting chemicals. The NR gene repertoire in different lineages has been shaped by gene loss, duplication and mutation, denoting a dynamic evolutionary route. As the earliest diverging class of gnathostomes (jawed vertebrates), cartilaginous fishes offer an exceptional opportunity to address the early diversification of NR gene families and the evolution of the endocrine system in jawed vertebrates. Here we provide an exhaustive analysis into the NR gene composition in five elasmobranch (sharks and rays) and two holocephalan (chimaeras) species. For this purpose, we generated also a low coverage draft genome assembly of the chimaera small-eyed rabbitfish, *Hydrolagus affinis*. We show that cartilaginous fish retain an archetypal NR gene repertoire, similar to that of mammals and coincident with the two rounds of whole genome duplication that occurred in the gnathostome ancestor. Furthermore, novel gene members of the non-canonical NR0B receptors were found in the genomes of this lineage. Our findings provide an essential view into the early diversification of NRs in gnathostomes, paving the way for functional studies.

1. Introduction

The homeostatic coordination of biological functions such as development or reproduction depends on the action of numerous transcription factors. Among these, Nuclear Receptors (NRs) are abundant and unique to metazoan genomes (Paps and Holland, 2018; Schmitz et al., 2016). NR monomers, homodimers or heterodimers typically triggered by ligand binding, selectively modulate transcription upon recognition of specific DNA responsive elements in the promoter region of target genes (Germain et al., 2006; Laudet and Gronemeyer, 2002). Their specific ligands comprise a vast array of small lipophilic molecules from endogenous or exogenous sources, such as hormones (e.g. thyroid hormones, steroids), morphogens (e.g. retinoic acid) or dietary components (e.g. fatty acids and vitamins) (Gronemeyer et al., 2004; Mangelsdorf et al., 1995; Mangelsdorf and Evans, 1995). Thus, NR-mediated gene expression is tightly controlled by the combinatorial effect of receptors, ligands and DNA responsive elements. The canonical NRs are divided into eight subfamilies, NR1 to NR8, and are

* Corresponding authors at: Comparative Genomics Laboratory, Institute of Molecular and Cell Biology (IBMC), 61 Biopolis Drive, Singapore 138673, Singapore (B. Venkatesh). CIIMAR/CIMAR – Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Terminal de Cruzeiros de Leixões, Av. General Norton De Matos s/n, 4450-208 Matosinhos, Portugal (L.F.C. Castro).
E-mail addresses: mcbbv@imcb.a-star.edu.sg (B. Venkatesh), filipe.castro@ciimar.up.pt (L.F.C. Castro).

https://doi.org/10.1016/j.ygcen.2020.113527
Received 7 April 2020; Received in revised form 15 May 2020; Accepted 3 June 2020
Available online 08 June 2020

characterized by a structural architecture that includes a DNA-binding domain (DBD) and a ligand-binding domain (LBD). The exception is the subfamily NR7 with two DBD (Kaur et al., 2015; Wu et al., 2007). The non-canonical NRs are grouped in the NR0 subfamily and lack the LBD in invertebrates and the DBD in vertebrates (Germain et al., 2006; Huang et al., 2015; Laudet and Gronemeyer, 2002). NRs have distinct binding specificities towards hormonal and non-hormonal compounds which is corroborated by the moderate conservation of their LBD amino acid sequences. Nevertheless, there are some ligand-orphan families of NRs (not exclusively ligand-activated) that exhibit unique structural adaptations to a redundant ligand-binding pocket (Kliewer et al., 1999).

A characteristic feature of NR biology is their activation or inhibition by the anthropogenic endocrine-disrupting chemicals (EDCs) (e.g. pharmaceutical agents, synthetic hormones, biocides, plastics, solvents and personal care products) (Capitão et al., 2017; Katsiadaki, 2019; le Maire et al., 2010; Lemaire et al., 2005). Some EDCs mimic or block the role of endogenous hormones and other signalling molecules, perturbing normal endocrine functions and causing physiological imbalances leading to diseases (e.g. Darbre, 2015). Thus, EDCs represent a major threat to animal ecosystem health, and understanding their mechanisms of action provides clues to anticipate deleterious effects in humans and wildlife.

Currently available genomes of bony vertebrates (sarcopterygians and actinopterygians) show a clear species-dependent pattern in NR gene content: 48 NRs in humans (Robinson-Rechavi et al., 2001), 49 NRs in mouse and rat, 52 NRs in the Western clawed frog, and 74 and 73 NRs in tilapia and zebrafish, respectively (Zhao et al., 2015). The number of NR genes is also extremely variable in invertebrate chordates: 17 NRs in sea squirt (urochordate) (Yagi et al., 2003) and 33 NRs in amphioxus (cephalochordate) (Lecroisey et al., 2012; Schubert et al., 2008). This variation in NR gene count in chordates has been shaped by events of gene loss and duplication (tandem or whole genome), denoting a flexible evolutionary configuration (Bertrand et al., 2011, 2004; Bridgham et al., 2010).

The extant jawed vertebrates are divided into groups, the bony vertebrates (Osteichthyes) and cartilaginous fishes (Chondrichthyes). To date, very few studies have addressed the diversity, function and ligand specificity of nuclear receptors in cartilaginous fishes (e.g.: Carroll et al., 2008; Filowitz et al., 2018; Inoue et al., 2010; Katsu et al., 2019). Cartilaginous fishes occupy an informative position in the evolution of vertebrates (Benton et al., 2009). They are composed of two subclasses: Holocephali (chimaeras) and Elasmobranchii (sharks and rays), that diverged approximately 400 million years ago (Ebert et al., 2013; Inoue et al., 2010). Cartilaginous fishes colonise a wide range of ecological habitats: while elasmobranchs are distributed from tropical to polar aquatic ecosystems and tolerate highly contaminated environments (Katsu et al., 2010), holocephalans are deep-water dwellers, living at depths of 500 m and deeper, populating all oceans with the exception of the Arctic and Antarctic oceans (Didier et al., 2012). Comparative genomic studies have shown that cartilaginous fish genomes are evolving at a slower rate compared to teleost fishes, sharing many similarities with mammalian genomes (Hara et al., 2018; Marra et al., 2019; Venkatesh et al., 2007, 2006). Thus, cartilaginous fishes are central to understand the evolution of vertebrate traits and lineage specific acquisitions.

Here we interrogated genomes and transcriptomes for the gene repertoire of NRs of five elasmobranch species – whale shark (*Rhincodon typus*), brownbanded bamboo shark (*Chiloscyllium punctatum*), small-spotted catshark (*Scyliorhinus canicula*), cloudy catshark (*S. torazame*) and little skate (*Leucoraja erinacea*) – and two Holocephali species, one with a publicly available genome – the elephant shark (*Callorhinchus milii*) – and another the small-eyed rabbitfish (*Hydrolagus affinis*) for which a draft genome assembly was generated in the present study.

2. Material and methods

2.1. Sequence retrieval and phylogenetic analysis

Amino acid sequences for putative NRs were retrieved via blast searches against publically available genome and transcriptomes obtained from GenBank, Skatebase (http://skatebase.org/), and Figshare (https://figshare.com/authors/Phyloinformatics_Lab_in_RIKEN_Kobe/4815111) using human (*Homo sapiens*) and zebrafish (*Danio rerio*) NR protein sequences as reference. Some Chondrichthyes NR sequences were retrieved via tblastn on Sequence Read Archive (SRA) using human NR sequences as reference and assembled with Geneious® (v.7.1.7). The accession numbers are listed in Supplementary Table S1.

The orthology of the retrieved sequences was inferred from phylogenetic analyses in the context of the NR superfamily. The sequences were aligned with the multiple alignment using fast Fourier transform (MAFFT) software (v.7) (Katoh and Toh, 2010) using the FFT-NS-2 model. A first sequence alignment combined all the retrieved NR amino acid sequences (edited with Geneious® (v.7.1.7), amino acid columns containing 90% of gaps were removed (alignment available on Figshare repository (https://figshare.com/s/a46a382a302531c78d9e). The final alignment containing 706 sequences and 894 positions was used in a Bayesian phylogenetic analysis with MrBayes (v.3.2.3) found at the CIPRES Science Gateway (v.3.3) (Miller et al., 2015). The following parameters were used: generation number = 27000000, rate matrix for aa = mixed (Jones), nruns = 2, nchains = 6, temp = 0.15, sampling set to 27,000 and burnin to 0.25. A second and more restricted Bayesian phylogenetic analysis with 334 sequences (3971 positions) from human, zebrafish, spotted gar (*Lepisosteus oculatus*), whale shark (*Rhincodon typus*), elephant shark (*Callorhinchus milii*) and Florida and European lancelets (*Branchiostoma floridae* and *B. lanceolatum*), was performed using the following parameters: generation number = 7000000, rate matrix for aa = mixed (Jones), nruns = 2, nchains = 4, temp = 0.2, sampling set to 5000 and burnin to 0.25. The phylogenetic trees were edited with iTOLv5 (https://itol.embl.de/).

2.2. Synteny analysis

To further investigate the orthology inferred from phylogenetics, the genomic location of individual NR genes was extracted from human (GRCh38), zebrafish (GRCz10 and GRCz11), spotted gar (LepOcu1), whale shark (ASM164234v2) and elephant shark (Callorhinchus_milii-6.1.3) genome assemblies. The NR-neighbouring genes were collected from GenBank using the human *loci* as reference to assemble the synteny maps of the remaining species. The orthology of NRb, one of the three novel cephalochordate NRs reported in previous studies (NRa, NRb and NRc) (Lecroisey et al., 2012; Schubert et al., 2008), could not be confirmed with our phylogenetic approach. Thus, the genomic location of this NR was retrieved from cephalochordate (*Branchiostoma belcheri*, Haploidv18h27 and *B. floridae* Version 2 GCA_000003815.1) genome assemblies; neighbouring genes were also collected and mapped in acorn worm (*Saccoglossus kowalevskii*, Skow_1.1) and purple sea urchin (*Strongylocentrotus purpuratus*, Spur_4.2) genomes to further investigate the gene orthology.

2.3. Sampling, DNA isolation and genome sequencing of the small-eyed rabbitfish

A small-eyed rabbitfish male specimen was collected from the EU-Spain Groundfish Survey (Gomes-dos-Santos et al., 2020; Román et al., 2019) in the NW Atlantic (NAFO Regulatory Area, Div. 3L from 47.3685, −46.6540 at a depth of 1159 m; to 47.3438, −46.6638 at a depth of 1157 m) (Supplementary Table S2). Small pieces of liver and gonad tissue were collected immediately after the capture of the specimen, preserved in RNAlater and stored at −20 °C, and transferred on arrival at the laboratory to −80 °C freezer, for future RNA extraction.

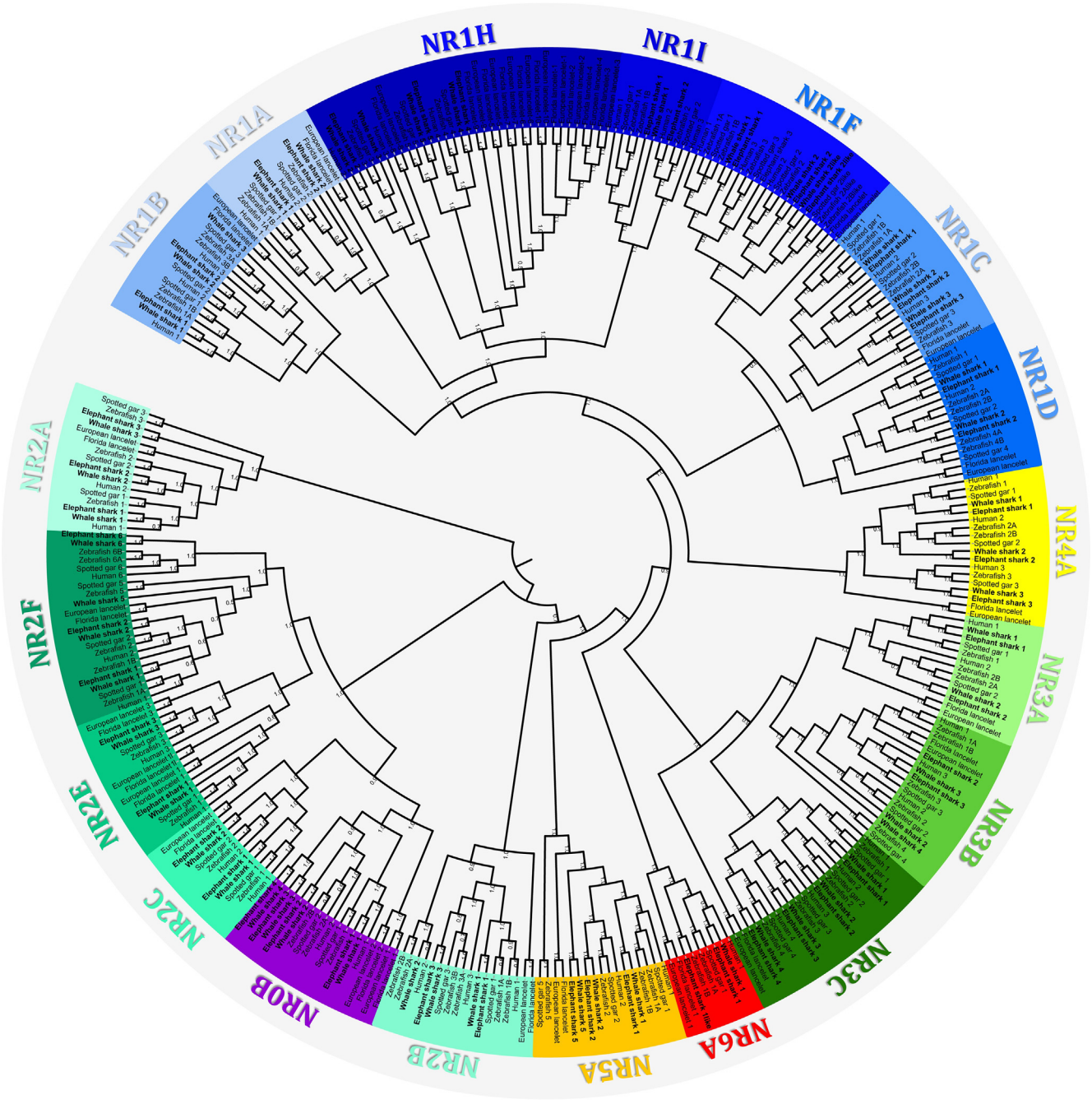


Fig. 1. Bayesian phylogenetic analysis of NR amino acid sequences of human, zebrafish, spotted gar, whale shark, elephant shark and Florida and European lancelets; numbers at nodes indicate posterior probabilities.

Additionally, a portion of muscle was also collected and preserved in ethanol (99%) and stored at −20 °C for DNA extraction. The integrity of the DNA was assessed on a 1% agarose TAE gel, stained with GelRed™ nucleic acid stain (Biotium, Hayward, CA, USA). The DNA sample was quantified using a Qubit Fluorometer (Invitrogen, USA). To sequence the paired-end (PE) short reads we used the Illumina HiSeq X Ten platform (Macrogen, Seoul, Korea) and generated a total of 76.2 Gbp of raw data.

Prior to genome sequencing, the taxonomic status of the specimen was confirmed by PCR amplification and Sanger sequencing of the mtDNA COI gene. A fragment of ~599 bp of *CO1* gene was amplified by PCR using the universal primer pair FishF2 and FishR2 (Ward et al., 2005) (Supplementary Table S2) and later confirmed through the analysis of the complete mitochondrial genome (Gomes-dos-Santos et al., 2020).

2.4. Cleaning of raw dataset, de novo assembly and assessment of small-eyed rabbitfish genome

The small-eyed rabbitfish raw sequencing reads were inspected using the FastQC software (v.0.11.8) (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) (Andrews et al., 2015), trimmed with the Trimmomatic software (v.0.38) (Bolger et al., 2014) and corrected to sequencing errors with the Lighter tool (v.1.1.1) (Song et al., 2014) (See Supplementary Material 1 for detailed methods). Next, the clean raw reads were used to determine general genome features (e.g. genome size estimation, heterozygosity and homozygosity rates, among other), using several software such as GenomeScope2 (v.2.0) (Ranallo-

Benavidez et al., 2019), KmerGenie (v.1.7048) (Chikhi and Medvedev, 2014), and the Preqc module of SGA software (v.0.10.15) (Simpson, 2014). The *de novo* assembly of small-eyed rabbitfish genome was performed using multiple k-mers and two independent software's, W2RAP pipeline (v.0.1) (Clavijo et al., 2017) and the SOAPdenovo2 software (v.2.04) (Luo et al., 2012). Next, the QUAST (v.5.0.2) (Gurevich et al., 2013) and Benchmarking Universal Single-Copy Orthologs (BUSCO) (v.3.0.2) (Simão et al., 2015) tools were used to analyse the draft genome assemblies (Supplementary Table S2).

2.5. NRs in the genome of the small-eyed rabbitfish and phylogenetic analysis

To validate the absence of certain NR in the genome of the elephant shark and consequently in Holocephali, we performed blastn researches in the small-eyed rabbitfish genome. The retrieved nucleotide sequences were translated and the amino acid sequences were aligned with MAFFT software (v.7) (https://mafft.cbrc.jp/alignment/server/) (Katoh and Toh, 2010) using the L-INS-i model and the orthologies were assessed by maximum likelihood under the LG substitution model for amino-acid sequences using the PhyML (v.3.0) (Guindon et al., 2010) plugin (v.2.2.4) in Geneious Prime (v.2019.2.3) (https://www.geneious.com/). Branch support was calculated by bootstrapping using 500 replications; the proportion of invariable sites and the gamma distribution parameters were estimated.

3. Results and discussion

Previously, various studies indicate that the NR gene super family composition is rather variable among chordate lineages (e.g. Bertrand et al., 2004; Lecroisey et al., 2012; Campo-Paysaa et al., 2015; Supplementary Material 2). Yet, a detailed view of the NR repertoire in cartilaginous fishes is still missing. As the earliest diverging class of gnathostomes (jawed vertebrates), together with their slowly evolving genomes (Hara et al., 2018; Marra et al., 2019; Venkatesh et al., 2014, 2007), Chondrichthyes offer an exceptional opportunity to address the early diversification of NR gene families in jawed vertebrates and the evolution of the endocrine system (e.g. Fonseca et al., 2017). In this study, we interrogated publicly available genomes and transcriptomes of five elasmobranch species: whale shark, brownbanded bamboo shark (*Chiloscyllium punctatum*), small-spotted catshark (*Scyliorhinus canicula*), cloudy catshark (*S. torazame*) and little skate (*Leucoraja erinacea*); and one Holocephali species: elephant shark. Given the paucity of genomic data from Holocephali, we additionally generated a draft genome assembly from the small-eyed rabbitfish (see below). Moreover, we searched for NRs in the reptile green anole (*Anolis carolinensis*) and in the non-teleost ray-finned fish spotted gar and we reviewed the NR gene repertoires previously reported in chicken, sea squirt and amphioxus (Marlétaz et al., 2018; Putnam et al., 2008; Supplementary Material 2).

3.1. Detailed examination of NRs in Chondrichthyes

Our searches for NR DBDs and LBDs sequences in the genomes and transcriptomes of whale shark, brownbanded bamboo shark, small-spotted catshark, cloudy catshark, little skate and elephant shark retrieved full or partial open reading frames (ORF) sequences of a total of 52 NRs in elasmobranchs and the elephant shark (Fig. 1). All the NRs were classified into seven NR subfamilies (NR0-NR6), showing orthology with human and actinopterygian species (Fig. 1). For most NRs, we observed a conservative evolutionary relationship, consistent with traditional morphological and molecular systematics (Fig. 1; Supplementary Fig. S1) (Zhao et al., 2015). To support the phylogenetic analyses and to distinguish between gene loss or absence of sequencing data, we analysed the genomic location (synteny) of NR *loci* in human, zebrafish, spotted gar and elephant shark (Supplementary Fig. S2).

3.1.1. NR1 subfamily

In this study, we found 18 and 21 NRs in elasmobranch species and the elephant shark, respectively, belonging to the NR1 subfamily. In detail, two *NR1A* and three *NR1C* paralogs were found in all the examined Chondrichthyes; two *NR1D* paralogs were found in most Chondrichthyes; three *NR1B* paralogs were recognized in the whale shark and little skate, whereas only two were found in the elephant shark (Fig. 1; Supplementary Fig. S1). The genomic locations of the human *NR1A*, *NR1B*, *NR1C* and *NR1D* orthologs were next inspected. In humans, *NR1A1*, *NR1B1* and *NR1D1* genes are located on the same chromosome (Chr 17), and a cluster with *NR1A2*, *NR1B2* and *NR1D2* genes is located on Chr 3. The same pattern was observed for zebrafish and spotted gar (Supplementary Fig. S2). Our analysis of the elephant shark genome located the cluster *NR1A2-NR1B2-NR1D2* in the same scaffold (Sca NW_006890093.1), but *NR1A1*, *NR1B1* and *NR1D1* genes were distributed in three different scaffolds (Sca NW_006891573.1, Sca NW_006890316.1 and Sca NW_006891334.1, respectively). The human *NR1C* paralogs were located in different chromosomes. In the elephant shark, the three *NR1C* paralogs were also recovered at three different scaffolds displaying a conserved syntenic arrangement to humans. Regarding *NR1B3*, we identified orthologs in the whale shark and little skate but no sequence was retrieved for the elephant shark (Fig. 1 and Supplementary Figs. S1 and S2). Yet, orthologs of neighbouring genes of human *NR1B3* were found in three separate scaffolds in the elephant shark genome (Supplementary Fig. S2). Furthermore, in zebrafish, spotted gar and green anole, an additional NR1D paralog (*NR1D4s*), previously identified in teleosts (Zhao et al., 2015), was also retrieved (Supplementary Fig. S1). No *NR1D4* paralog was found in cartilaginous fish genomes. In zebrafish and spotted gar, *NR1B3* paralogs are followed by *NR1D4* paralogs (Supplementary Fig. S2). The *NR1B3* is found in elasmobranch genomes (Fig. 1) and we were able to identify neighbouring genes of *NR1B3* in elephant shark genome (Supplementary Fig. S2).

In the NR1F group, we detected four paralogs in the chimaera and three in elasmobranchs (Figs. 1, 2). Our blast searches failed to identify *NR1F3* paralogs in elasmobranchs and we were unable to further confirm the putative loss by synteny due to the poor genome assembly of the whale shark currently available. Moreover, we found a *NR1F2-like* receptor, previously identified in frog, turtle, birds and teleosts (Zhao et al., 2015), and now uncovered also in green anole, spotted gar and cartilaginous fish. Thus, it is reasonable to deduce that *NR1F2-like* was secondarily lost in the ancestor of mammals (Supplementary Figs. S1 and S2).

In a previous study, we addressed the evolution of NR1H and NR1I groups, both characterized by independent gene loss events (Fonseca et al., 2019, 2017). Briefly, we indicated that *NR1H2* and *NR1H3* are gnathostome-specific paralogs (Fonseca et al., 2017). Here, we were able to retrieve the full or near full sequence of these paralogs in all cartilaginous fish, with the exception of the very short sequence of the elephant shark *NR1H3* paralog (Fig. 1). Similarly, *NR1H4* (this study) and *NR1H5* (this study and Cai et al. (2007)) paralogs were identified in cartilaginous fish species, indicating that NR1H paralogs are gnathostome-specific (Fonseca et al., 2017). Contrasting with the ubiquitous occurrence of *NR1I1* in all studied gnathostome genomes, *NR1I3* displays a tetrapod-specific occurrence. Furthermore, independent losses of *NR1I2* during teleost evolution (Eide et al., 2018) and in Elasmobranchii (Fonseca et al., 2019) were suggested, paralleling the scenario described for the *NR1F3* gene.

3.1.2. NR2 subfamily

In the NR2 subfamily a total of 14 and 12 NRs were identified in elasmobranchs and elephant shark, respectively. Cartilaginous fish display three NR2A paralogs (*NR2A1*, *NR2A2* and *NR2A3*) similarly to birds, reptiles, amphibians and teleosts (Figs. 1 and 2). Also, their genomes encode two paralogs in both NR2C (*NR2C1* and *NR2C2*) and NR2E (*NR2E1* and *NR2E3*) groups (Figs. 1 and 2). Concerning the NR2B

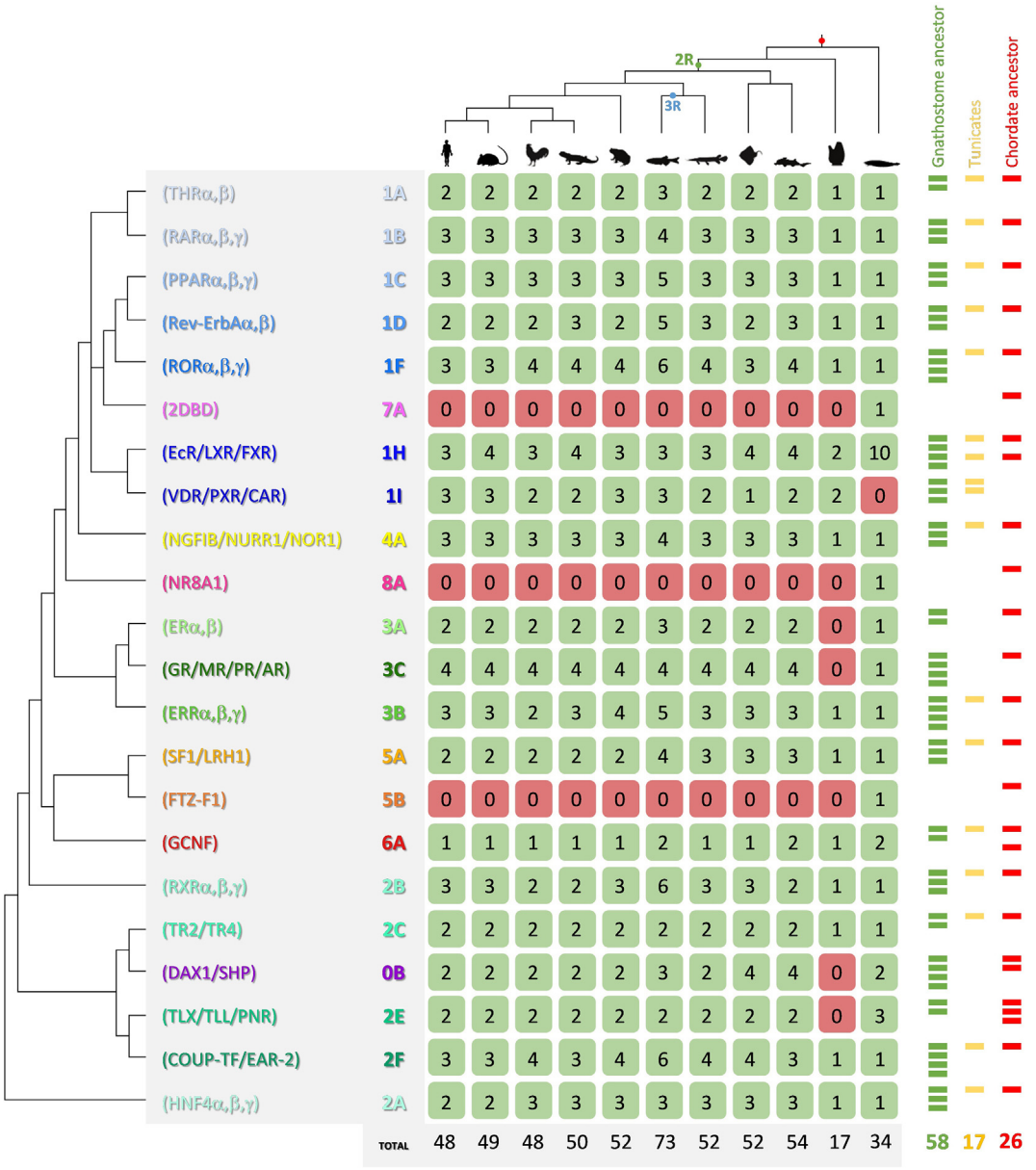


Fig. 2. Phylogenetic representation of NRs distribution across Chordata. The numbered green boxes refer to the number of NR orthologs, while red boxes indicate the absence of NR orthologs in the genome of the represented species. On the right, green, yellow and red boxes represent the estimated number of NRs copies in gnathostome ancestor, tunicates and chordate ancestor, respectively.

group, we found three paralogs (*NR2B1*, *NR2B2* and *NR2B3*) in elasmobranchs and two in elephant shark (*NR2B1* and *NR2B3*) (Figs. 1 and 2). Similarly, we recovered four NR2F paralogs (*NR2F1*, *NR2F2*, *NR2F5* and *NR2F6*) but only three paralogs (*NR2F1*, *NR2F2* and *NR2F6*) in elasmobranch and elephant shark genomes, respectively (Figs. 1 and 2). We next inspected the genomic regions of *NR2B2* and *NR2F5* in human, zebrafish and spotted gar to recover the *NR2B2* and *NR2F5* locations in elephant shark genome. However, the complete reconstruction of elephant shark *NR2B2* and *NR2F5 loci* was not possible due to fragmentation of this regions in the current assembly (Supplementary Fig. S2). Thus, the presence or absence of *NR2B2* and *NR2F5* in the elephant shark remains inconclusive. Yet, given that the full set of gene paralogs was retrieved in elasmobranch genomes, we put forward that the all three NR2B and four NR2F genes were encoded in the gnathostome ancestor genome (Fig. 2).

3.1.3. NR3 subfamily

Regarding the NR3 subfamily, we recognized 9 NRs in elasmobranchs and the elephant shark. The two NR3A paralogs (*NR3A1* and *NR3A2*) and the four NR3C paralogs (*NR3C1*, *NR3C2*, *NR3C3* and *NR3C4*) found in the studied gnathostomes were also retrieved in our analysis of cartilaginous fish genomes and transcriptomes (Figs. 1 and 2). Conversely, not all the NR3B gene paralogs were retrieved in our analysis. We found *NR3B2*, *NR3B3* and *NR3B4* gene orthologs. The later was secondarily lost in reptiles, birds and mammals (Fig. 1, Supplementary Fig. S1) (Zhao et al., 2015)). Similar to chicken and spotted gar, no *NR3B1* was found in the investigated cartilaginous fish genomes. The genomic location analysis for this gene paralog suggests the independent loss of *NR3B1* in this lineage and holosteans. However, the elephant shark *NR3B1* neighbouring gene orthologs were found in different scaffolds and the genomic location of spotted gar *NR3B1* is poorly assembled (Supplementary Fig. S2). Therefore, we suggest that the gnathostome ancestor displayed two NR3A paralogs, four NR3B

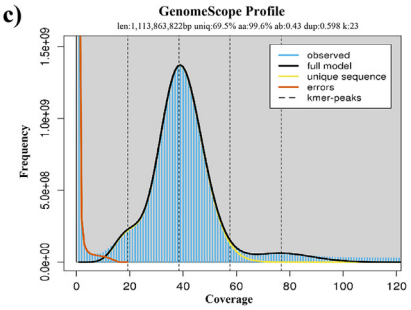
a)

b)

Raw data	
Raw sequencing reads (PE)	205663228
Clean reads	183667702
Assembly v. Haf_172k_RS_71 statistics (>500bp)	
Genome Size estimated, gb	0,941-1,268
Number of contigs [1,2,3]	252783
Contig N50 size, kb	15.875
Number of Scaffolds	245464
Scaffold N50 size, kb	19.535
Longest scaffold, kb	397.232
Total length Scaffolds, gb	1.114
Heterozygosity, %$_1$	0,41-0,38
BUSCO Completeness (%)	
BUSCO's library	**Vertebrata**
Complete	60.0
Complete and single copy	59.2
Complete and duplicated	0.8
Fragmented	26.1
Missing	13.9
Total BUSCO found	86.1

1) Kmergenie (1,268 Mb); **2)** Genomescope k19,21,23,25,27,29,31 (941-1,065 Mb); **3)** SgaPreQC (1,228 Mb)

c)



d)

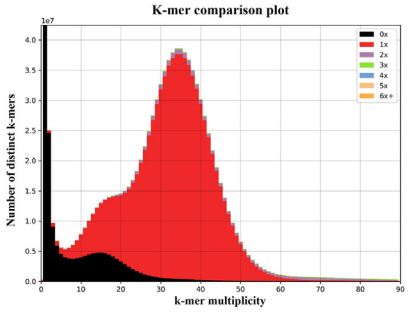


Fig. 3. A draft genome assembly of *Hydrolagus affinis*. a) Photograph of the sequenced specimen collected in Northwest Atlantic (photograph credits to Nair Arrondo); b) General statistics of the genome assembly; c) GenomeScope2 k-mer frequency distribution profile, showing genome size estimation, heterozygosity, unique content and repeat content (i.e. len, ab, uniq, dup respectively) for 23 k-mer; d) Stacked histogram showing read k-mer frequency spectrum versus *Hydrolagus affinis* genome assembly. Colours represent the frequency of read content present in the final assembly.

paralogs and four NR3C paralogs (Fig. 2).

3.1.4. NR4, NR5 and NR6 subfamilies

The three NR4A paralogs (*NR4A1*, *NR4A2* and *NR4A3*), two NR5A paralogs (*NR5A1* and *NR5A2*) and the *NR6A1* ortholog were recovered in all examined genomes (Figs. 1 and 2). Additionally, orthologs of zebrafish *NR5A5* were also identified in the spotted gar and cartilaginous fish genomes, having been secondarily lost in tetrapods (Fig. 1 and Supplementary Fig. S1). However, neither phylogenetics nor the synteny analyses were sufficiently robust to support the accurate orthology of *NR5A5* (Fig. 1; Supplementary Figs. S1 and S2). Importantly, we deduced the existence of a *NR6A1-like* gene ortholog located next to *NR5A2* in the elephant shark genome (Fig. 1; Supplementary Figs. S1 and S2). The analyses of the human, zebrafish and spotted gar *NR5A2 locus* allowed us to corroborate the loss of *NR6A1-like* gene in these species. The elasmobranch *NR5A2 locus* is dispersed into several scaffolds (data not shown), impeding a stronger support to the conclusion about *NR6A1-like* gene loss in these species.

3.1.5. NR0 subfamily

NR0B genes are non-canonical receptors which lack a DBD (Laudet and Gronemeyer, 2002). Until the present study only two gene paralogs (*NR0B1* and *NR0B2*) had been described in vertebrates (Zhao et al., 2015). Our searches into cartilaginous fish genomes and transcriptomes retrieved two novel NR0B paralogs, which we name *NR0B3* and *NR0B4* respectively (Figs. 1 and 2). Furthermore, by examining the genomic location of *NR0B3* and *NR0B4 loci* in human, zebrafish and spotted gar (Supplementary Fig. S2), we were able to deduce the secondary loss of these genes in the Euteleostomi lineage. Interestingly, we were able to show that the previously named *NR0B2* ortholog in frog is a *NR0B4* gene (Supplementary Fig. S2). In effect, the analyses of NR0B gene *loci* in this lineage (data not shown) supports that *NR0B2* and *NR0B3* paralogs were lost, while *NR0B1* and *NR0B4* paralogs were retained. Therefore, we propose that gnathostome ancestor had four NR0B paralogs (Fig. 2).

3.2. De novo genome assembly of the small-eyed rabbitfish

Our analysis of publicly available (or published) cartilaginous fish genomes allowed the identification of a total of 52 NR. Yet, in 5 cases we were unable to confirm the presence of a given NR in the available genome of the elephant shark – Holocephalii. To further clarify if these cases of NR gene absence in the elephant shark (or cartilaginous fish) are due to lineage-specific losses or result from missing sequencing data or assembly gaps, we generated a draft genome assembly of second chimaera species, the small-eyed rabbitfish (Fig. 3a). The sequencing produced approximately 205 M of paired end raw reads that, after trimming and quality control resulted in about 184 M reads to use for further analyses (Fig. 3b). We estimated a genome size between 1.08 and 1.27 Gb for the small-eyed rabbitfish (Fig. 3b). These values are smaller than genome size estimations for Chimaeriformes present in Animal Genome Size Database (AGSD – http://www.genomesize.com), ranging between 1.51 Gb (spotted ratfish, Hydrolagus colliei) and 2.01 Gb stipulated for these species. In addition, the GenomeScope2 analyses allowed the determination of the rate of heterozygosity of the small-eyed rabbitfish between 0.37 and 0.48 % (Fig. 3c and Supplementary Table S3). Interestingly, the only Chimaeriformes species with an available genome assembly, the elephant shark (NCBI – GCA_000165045.2) (Venkatesh et al., 2014), has a genome size of 0.937 Gb Mb, slightly smaller than the estimated for the small-eyed rabbitfish. Following the pre-established criteria (See Supplementary Material 1), we selected the Haf_172k version of the draft genome assembly of the small-eyed rabbitfish with an N50 contig and scaffold values of 15.857 and 19.535 kb, respectively, a total length of 1.117 Gb and a percentage of 76.3, 83.1, 86.1% of BUSCOS found in the three BUSCO libraries (Fig. 3b, Supplementary Table S4). Additionally, the KAT analysis showed that the major part of read k-mer content is present in the assembly (red colour of the histogram), mainly in the homozygotic peak (± 34x; k-mer multiplicity), while the k-mer content missing (black colour of the histogram), is mainly present in the heterozygotic peak (± 18x; k-mer multiplicity) (Fig. 3d).

We next investigated the NR gene catalogue in this draft genome assembly, focusing on genes missing in the elephant shark current assembly. In this species, *NR1D4*, *NR1B3*, *NR2B2*, *NR2F5* and *NR3B1* genes are not found (Fig. 1; Supplementary Fig. S2). Thus, we conducted blastn searches on the small-eyed rabbitfish genome, and assessed the orthology of the retrieved sequences with phylogenetic analyses (Fig. 4). Our analyses retrieved the three paralogs of both *NR1B* and *NR1D* subfamilies (Fig. 4a and b; Supplementary Table S5). Thus, we propose that the gnathostome ancestor genome encoded three *NR1B* and three *NR1D* paralogs (Fig. 2). Nevertheless, we failed to recovered *NR2B2*, *NR2F5* and *NR3B1* gene orthologs (Fig. 4c, 4d and 4e;

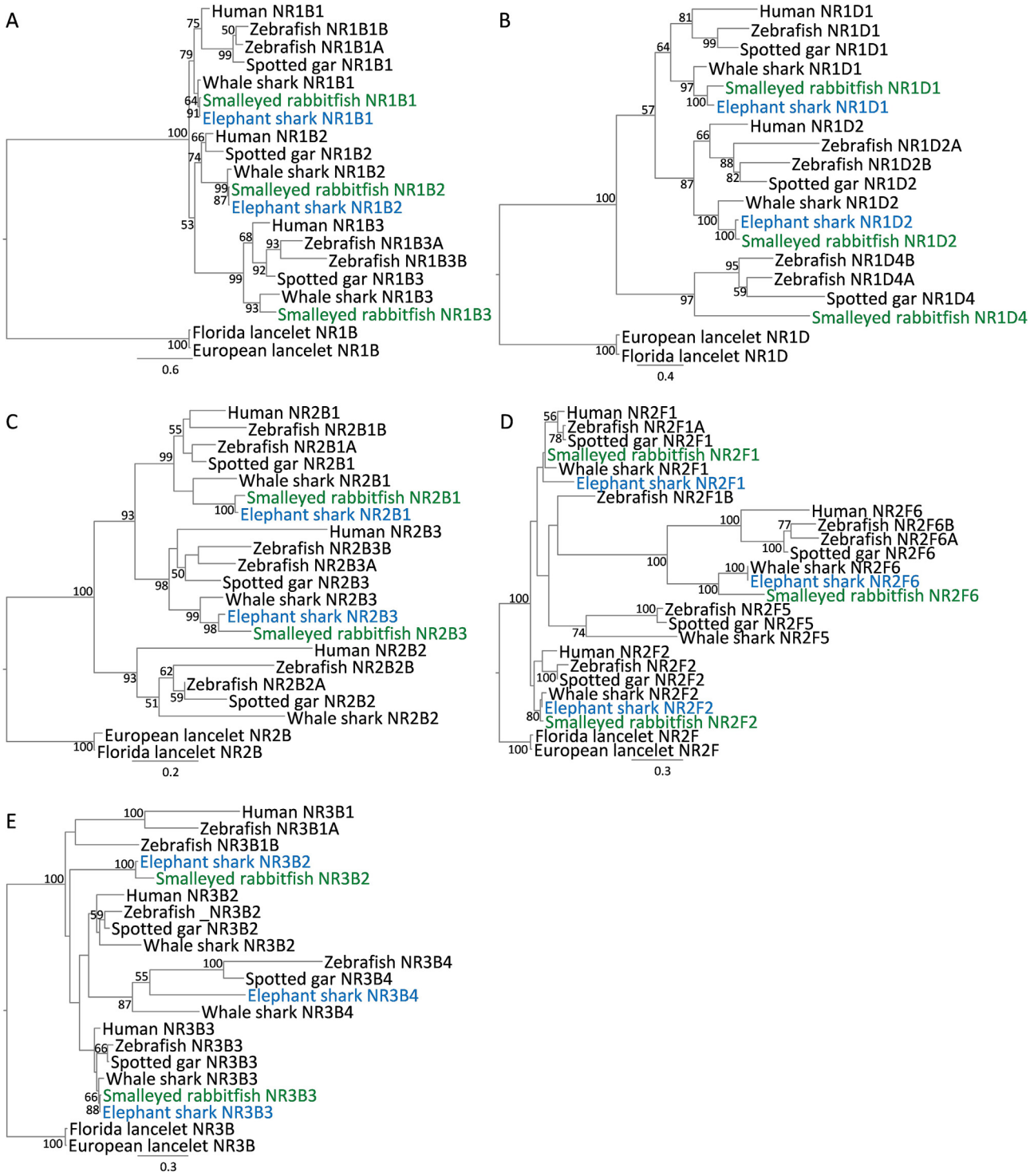


Fig. 4. Phylogenetic trees of NRs recovered from small-eyed rabbitfish genome. a) NR1B; b) NR1D; c) NR2B; d) NR2F; e) NR3B. Bootstrap values below 50% are not shown.

Supplementary Table S5). Overall, we found 54 NRs in the two Holocephali species genomes and our results suggest that the genome of the gnathostome ancestor encoded at least 58 NRs (Fig. 2).

3.3. NR gene repertoire in Chondrichthyes: functional considerations

The evolution of the Metazoa coincided with the appearance of numerous novelties at the genome level, including NRs (e.g. Paps and Holland, 2018). In effect, NRs are an important family of transcription factors found in extant metazoan genomes (Mangelsdorf et al., 1995; Schmitz et al., 2016). As transcription factors, NRs regulate the expression of genes involved in biological process, such as, reproduction, development, metabolism and immunity, being prime targets for hormones of the endocrine system (Evans and Mangelsdorf, 2014). Over the past decade, some studies have been conducted in cartilaginous fishes to elucidate the evolution of signalling pathways involving NRs. With the present work we took advantage of a new and several existing genomic datasets to gain insight into the emergence and evolution of

vertebrate signalling pathways involving NRs. Regarding cartilaginous fish, we were able to determine that the NR repertoire is relatively stable and the number of genes and their syntenic distribution are comparable to the ones of mammalian NR gene repertoire; despite the novelties described in the sections above (*NR0B3*, *NR0B4* and *NR6A1-like* genes). Yet, repertoire stability does not necessarily imply functional conservation. In fact, NR-mediated gene transcription operates within complex interaction networks including NRs and partner NRs, ligands, and DNA-binding sites, a network which is further entangled by the transcriptional modulation of NRs, their expression patterns, as well as downstream target genes and physiological processes: establishing the so called "*NR ring of physiology*" (Bookout et al., 2006; Siddiq et al., 2017). Although not exhaustive, some studies have provided clues on the functional evolution of cartilaginous fish NRs: highlighting cases of conserved responsiveness (estrogen receptors, (Katsu et al., 2010) (Filowitz et al., 2018); liver X receptors, (Fonseca et al., 2017)) or weak activity towards typical mammalian ligands (farnesoid X receptor β (Cai et al., 2007)); and also, lineage-specific specializations including interactions with elasmobranch-specific hormones (mineralocorticoid receptor and glucocorticoid receptor (Carroll et al., 2011, 2008)) or distinct activation and tissue expression profiles (pregnane X receptor, (Fonseca et al., 2019). Still, the current knowledge on NR function and evolution in this vertebrate lineage is still sparse and further studies are required to illuminate NR-dependent networks within this group. Importantly, and given their phylogenetic placement, the assessment of NR function in cartilaginous fish is crucial to decipher how mechanisms of receptor promiscuity, hormone exploitation or co-evolution of metabolic signalling pathways emerged in gnathostome history (Carroll et al., 2008; Thornton, 2001).

4. Conclusions

Overall, the investigation of cartilaginous fish genomes allowed the identification of the complete repertoire of NRs in this early branching lineage of jawed vertebrates offering a valuable resource for understanding the evolution of endocrine systems and physiology of gnathostomes.

CRediT authorship contribution statement

Elza Fonseca: Methodology, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing. **André M. Machado:** Methodology, Validation, Formal analysis, Investigation, Data curation, Writing - review & editing. **Nair Vilas-Arrondo:** Software, Validation, Investigation, Resources, Writing - review & editing, Visualization. **André Gomes-dos-Santos:** Validation, Investigation, Writing - review & editing. **Ana Veríssimo:** Validation, Investigation, Writing - review & editing, Funding acquisition. **Pedro Esteves:** Validation, Investigation, Writing - review & editing. **Tereza Almeida:** Validation, Investigation, Writing - review & editing. **Gonçalo Themudo:** Validation, Investigation, Writing - review & editing. **Raquel Ruivo:** Methodology, Validation, Investigation, Writing - review & editing. **Montse Pérez:** Validation, Investigation, Resources, Writing - review & editing, Funding acquisition. **Rute da Fonseca:** Methodology, Validation, Investigation, Writing - review & editing. **Miguel M. Santos:** Methodology, Validation, Investigation, Writing - review & editing. **Elsa Froufe:** Validation, Investigation, Writing - review & editing. **Esther Román-Marcote:** Validation, Investigation, Resources, Writing - review & editing, Funding acquisition. **Byrappa Venkatesh:** Validation, Investigation, Writing - review & editing, Supervision. **L. Filipe C. Castro:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgements

This research was funded by COMPETE 2020, Portugal 2020 and the European Union through the ERDF, grant number 031342, and by FCT through national funds (PTDC/CTA-AMB/31342/2017). The genome sequencing of Hydrolagus affinis was funded by the FCT through project PTDC/ASP-PES/28053/2017. A PhD fellowship awarded to Elza Fonseca (SFRH/BD/100262/2014) by FCT. R.R.F. thanks the Danish National Research Foundation for its support of the Center for Macroecology, Evolution, and Climate (grant DNRF96). Research stay fellowship Doctoral Program awarded to Nair Vilas Arrondo by Do*Mar (UVIGO). EU-Spain NAFO Groundfish survey has been co-funded by the European Union through the European Maritime and Fisheries Fund (EMFF) within the National Program of collection, management and use of data in the fisheries sector and support for scientific advice regarding the Common Fisheries Policy and IEO-BIOPESLE project. This research is part of the CIIMAR lead initiative *PortugalFishomics*.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygcen.2020.113527.

References

Andrews, S., Krueger, F., Seconds-Pichon, A., Biggins, F., Wingett, S., 2015. FastQC. A quality control tool for high throughput sequence data. Babraham Bioinformatics. Babraham Inst.

Benton, M., Donoghue, P., Asher, R., 2009. Calibrating and constraining molecular clocks. In: Hedges, S., Kumar, S. (Eds.), The Timetree of Life. Oxford University Press, pp. 35–86.

Bertrand, S., Belgacem, M., Escriva, H., 2011. Nuclear hormone receptors in chordates. Mol. Cell. Endocrinol. 334, 67–75. https://doi.org/10.1016/j.mce.2010.06.017.

Bertrand, S., Brunet, F., Escriva, H., Parmentier, G., Laudet, V., Robinson-Rechavi, M., 2004. Evolutionary genomics of nuclear receptors: from twenty-five ancestral genes to derived endocrine systems. Mol. Biol. Evol. 21, 1923–1937. https://doi.org/10.1093/molbev/msh200.

Bolger, A., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Bookout, A., Jeong, Y., Downes, M., Yu, R., Evans, R., Mangelsdorf, D., 2006. Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network. Cell 126, 789–799. https://doi.org/10.1016/j.cell.2006.06.049.

Bridgham, J., Eick, G., Larroux, C., Deshpande, K., Harms, M., Gauthier, M., Ortlund, E., Degnan, B., Thornton, J., 2010. Protein evolution by molecular tinkering: diversification of the nuclear receptor superfamily from a ligand-dependent ancestor. PLoSBiol 8, e1000497. https://doi.org/10.1371/journal.pbio.1000497.

Cai, S., Xiong, L., Wray, C., Ballatori, N., Boyer, J., 2007. The farnesoid X receptor FXRalpha/NR1H4 acquired ligand specificity for bile salts late in vertebrate evolution. Am. J. Physiol. Regul. Integr. Comp. Physiol. 293, 1400–1409. https://doi.org/10.1152/ajpregu.00781.2006.

Campo-Paysaa, F., Jandzik, D., Takio-Ogawa, Y., Cattell, M.V., Neef, H.C., Langeland, J.A., Kuratani, S., Medeiros, D.M., Mazan, S., Kuraku, S., Laudet, V., Schubert, M., 2015. Evolution of retinoic acid receptors in chordates: insights from three lamprey species, *Lampetra fluviatilis*, *Petromyzon marinus*, and *Lethenteron japonicum*. Evodevo 6, 18. https://doi.org/10.1186/s13227-015-0016-4.

Capitão, A., Lyssimachou, A., Castro, L., Santos, M., 2017. Obesogens in the aquatic environment: an evolutionary and toxicological perspective. Environ. Int. 106, 153–169. https://doi.org/10.1016/j.envint.2017.06.003.

Carroll, S., Bridgham, J., Thornton, J., 2008. Evolution of hormone signaling in elasmobranchs by exploitation of promiscuous receptors. Mol. Biol. Evol. 25, 2643–2652. https://doi.org/10.1093/molbev/msn204.

Carroll, S., Ortlund, E., Thornton, J., 2011. Mechanisms for the evolution of a derived function in the ancestral glucocorticoid receptor. PLoS Genet. 7, e1002117. https://doi.org/10.1371/journal.pgen.1002117.

Chikhi, R., Medvedev, P., 2014. Informed and automated k-mer size selection for genome assembly. Bioinformatics 30, 31–37. https://doi.org/10.1093/bioinformatics/btt310.

Clavijo, B., Accinelli, G., Wright, J., Heavens, D., Barr, K., Yanes, L., Federica, D.-P., 2017. W2RAP: a pipeline for high quality, robust assemblies of large complex genomes from short read data. bioRxiv. https://doi.org/10.1101/110999.

Darbre, P., 2015. Endocrine Disruption and Human Health. Overview of EDCs and Human Health which sets the Bigger Picture. Academic, New York.

Didier, D., Kemper, J., Ebert, D., 2012. Phylogeny, biology, and classification of extant holocephalans. In: Carrier, J., Musick, J., Heithaus, M. (Eds.), Biology of Sharks and Their Relatives. CRC Press, Boca Raton, FL, pp. 97–122.

Ebert, D., Ho, H., White, W., Carvalho, M., 2013. Introduction to the systematics and biodiversity of sharks, rays, and chimaeras (Chondrichthyes) of Taiwan. Zootaxa 3752, 5–19.

Eide, M., Rydbeck, H., Tørresen, O., Lille-Langøy, R., Puntervoll, P., Goldstone, J., Jakobsen, K., Stegeman, J., Goksøyr, A., Karlsen, O., 2018. Independent losses of a xenobiotics receptor across teleost evolution. Sci. Rep. 8, 10404. https://doi.org/10.1038/s41598-018-28498-4.

Evans, R., Mangelsdorf, D., 2014. Nuclear Receptors, RXR, and the Big Bang. Cell 157, 255–266. https://doi.org/10.1016/j.cell.2014.03.012.

Filowitz, G., Rajakumar, R., O'Shaughnessy, K., Cohn, M., 2018. Cartilaginous fishes provide insights into the origin, diversification, and sexually dimorphic expression of vertebrate estrogen receptor genes. Mol. Biol. Evol. 35, 2695–2701. https://doi.org/10.1093/molbev/msy165.

Fonseca, E., Ruivo, R., Machado, A.M., Conrado, F., Tay, B.H., Venkatesh, B., Santos, M.M., Castro, L.F.C., 2019. Evolutionary plasticity in detoxification gene modules: the preservation and loss of the pregnane X receptor in chondrichthyes lineages. Int. J. Mol. Sci. 10, E2331. https://doi.org/10.3390/ijms20092331.

Fonseca, E., Ruivo, R., Lopes-Marques, M., Zhang, H., Santos, M., Venkatesh, B., Castro, L., 2017. LXRα and LXRβ nuclear receptors evolved in the common ancestor of gnathostomes. Genome Biol. Evol. 9, 222–230. https://doi.org/10.1093/gbe/evw305.

Germain, P., Staels, B., Dacquet, C., Spedding, M., Laudet, V., 2006. Overview of nomenclature of nuclear receptors. Pharmacol. Rev. 58, 685–704. https://doi.org/10.1124/pr.58.4.2.

Gomes-dos-Santos, A., Arrondo, N., Machado, A., Veríssimo, A., Pérez, M., Román, E., Castro, L., Froufe, E., 2020. The complete mitochondrial genome of the deep-water cartilaginous fish Hydrolagus affinis Brito Capello, 1867 (Holocephali: Chimaeridae). Mitochondrial DNA Part B, 5:2, 1810 1812. DOI: 10.1080/23802359.2020.1749154.

Gronemeyer, H., Gustafsson, J., Laudet, V., 2004. Principles for modulation of the nuclear receptor superfamily. Nat. Rev. Drug Discov. 3, 950–964. https://doi.org/10.1038/nrd1551.

Guindon, S., Dufayard, J., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321. https://doi.org/10.1093/sysbio/syq010.

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29, 1072–1075. https://doi.org/10.1093/bioinformatics/btt086.

Hara, Y., Yamaguchi, K., Onimaru, K., Kadota, M., Koyanagi, M., Keeley, S., Tatsumi, K., Tanaka, K., Motone, F., Kageyama, Y., Nozu, R., Adachi, N., Nishimura, O., Nakagawa, R., Tanegashima, C., Kiyatake, I., Matsumoto, R., Murakumo, K., Nishida, K., Terakita, A., Kuratani, S., Sato, K., Hyodo, S., Kuraku, S., 2018. Shark genomes provide insights into elasmobranch evolution and the origin of vertebrates. Nat. Ecol. Evol. 2, 1761–1771. https://doi.org/10.1038/s41559-018-0673-5.

Huang, W., Xu, F., Li, J., Li, L., Que, H., Zhang, G., 2015. Evolution of a novel nuclear receptor subfamily with emphasis on the member from the Pacific oyster Crassostrea gigas. Gene 567, 164–172. https://doi.org/10.1016/j.gene.2015.04.082.

Inoue, J., Miya, M., Lam, K., Tay, B., Danks, J.A., Bell, J., Walker, T., Venkatesh, B., 2010. Evolutionary origin and phylogeny of the modern holocephalans (Chondrichthyes: Chimaeriformes): a mitogenomic perspective. Mol. Biol. Evol. 27, 2576–2586. https://doi.org/10.1093/molbev/msq147.

Katoh, K., Toh, H., 2010. Parallelization of the MAFFT multiple sequence alignment program. Bioinformatics 26, 1899–1900. https://doi.org/10.1093/bioinformatics/bty121.

Katsiadaki, I., 2019. Are marine invertebrates really at risk from endocrine-disrupting chemicals? Curr Opin Env. Sci Heal. 11, 37–42. https://doi.org/10.1016/j.coesh.2019.06.005.

Katsu, Y., Kohno, S., Narita, H., Urushitani, H., Yamane, K., Hara, A., Clauss, T., Walsh, M., Miyagawa, S., Guillette, L.J., Iguchi, T., 2010. Cloning and functional characterization of Chondrichthyes, cloudy catshark, Scyliorhinus torazame and whale shark, Rhincodon typus estrogen receptors. Gen. Comp. Endocrinol. 168, 496–504. https://doi.org/10.1016/j.ygcen.2010.06.010.

Katsu, Y., Kohno, S., Oka, K., Lin, X., Otake, S., Pillai, N., Takagi, W., Hyodo, S., Venkatesh, B., Baker, M., 2019. Transcriptional activation of elephant shark mineralocorticoid receptor by corticosteroids, progesterone, and spironolactone. Sci. Signal 12, eaar2668. https://doi.org/10.1126/scisignal.aar2668.

Kaur, S., Jobling, S., Jones, C., Noble, L., Routledge, E., Lockyer, A., 2015. The nuclear receptors of Biomphalaria glabrata and Lottia gigantea: implications for developing new model organisms. PLoS ONE 10, e0121259. https://doi.org/10.1371/journal.pone.0121259.

Kliewer, S., Lehmann, J., Willson, T., 1999. Orphan nuclear receptors: shifting endocrinology into reverse. Science (80-) 284, 757–760. https://doi.org/10.1126/science.284.5415.757.

Laudet, V., Gronemeyer, H., 2002. The Nuclear Receptors Factsbook. Academic Press, London.

le Maire, A., Bourguet, W., Balaguer, P., 2010. A structural view of nuclear hormone receptor: endocrine disruptor interactions. Cell. Mol. Life Sci. 67, 1219–1237. https://doi.org/10.1007/s00018-009-0249-2.

Lecroisey, C., Laudet, V., Schubert, M., 2012. The cephalochordate amphioxus: a key to reveal the secrets of nuclear receptor evolution. Br. Funct. Genomics 11, 156–166. https://doi.org/10.1093/bfgp/els008.

Lemaire, G., Balaguer, P., Michel, S., Rahmani, R., 2005. Activation of retinoic acid receptor-dependent transcription by organochlorine pesticides. Toxicol. Appl. Pharmacol. 202, 38–49. https://doi.org/10.1016/j.taap.2004.06.004.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C., Cheung, D., Yiu, S., Peng, S., Xiaoqian, Z., Liu, G., Liao, X., Li, Y., Yang, H., Wang, J., Lam, T., Wang, J., 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. GigaScience 1, 18. https://doi.org/10.1186/2047-217X-1-18.

Mangelsdorf, D., Evans, R., 1995. The RXR heterodimers and orphan receptors. Cell 83, 841–850. https://doi.org/10.1016/0092-8674(95)90200-7.

Mangelsdorf, D., Thummel, C., Beato, M., Herrlich, P., Schütz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., Evans, R., 1995. The nuclear receptor superfamily: the second decade. Cell 83, 835–839. https://doi.org/10.1016/0092-8674(95)90199-X.

Marlétaz, F., Firbas, P., Maeso, I., Tena, J., Bogdanovic, O., Perry, M., Wyatt, C., de la Calle-Mustienes, E., Bertrand, S., Burguera, D., Acemel, R., van Heeringen, S., Naranjo, S., Herrera-Ubeda, C., Skvortsova, K., Jimenez-Gancedo, S., Aldea, D., Marquez, Y., Buono, L., Kozmikova, I., Permanyer, J., Louis, A., Albuixech-Crespo, B., Le Petillon, Y., Leon, A., Subirana, L., Balwierz, P., Duckett, P., Farahani, E., Aury, J., Mangenot, S., Wincker, P., Albalat, R., Benito-Gutiérrez, È., Cañestro, C., Castro, F., D'Aniello, S., Ferrier, D., Huang, S., Laudet, V., Marais, G., Pontarotti, P., Schubert, M., Seitz, H., Somorjai, I., Takahashi, T., Mirabeau, O., Xu, A., Yu, J., Carninci, P., Martinez-Morales, J., Crollius, H., Kozmik, Z., Weirauch, M., Garcia-Fernàndez, J., Lister, R., Lenhard, B., Holland, P., Escriva, H., Gómez-Skarmeta, J., Irimia, M., 2018. Amphioxus functional genomics and the origins of vertebrate gene regulation. Nature 564, 64–70. https://doi.org/10.1038/s41586-018-0734-6.

Marra, N., Stanhope, M., Jue, N., Wang, M., Sun, Q., Pavinski Bitar, P., Richards, V., Komissarov, A., Rayko, M., Kliver, S., Stanhope, B., Winkler, C., O'Brien, S., Antunes, A., Jorgensen, S., Shivji, M., 2019. White shark genome reveals ancient elasmobranch adaptations associated with wound healing and the maintenance of genome stability. Proc. Natl. Acad. Sci. U.S.A. 116, 4446–4455. https://doi.org/10.1073/pnas.1819778116.

Miller, M., Schwartz, T., Pickett, B., He, S., Klem, E., Scheuermann, R.M., Kaufman, S., O'Leary, M., 2015. A RESTful API for access to phylogenetic tools via the CIPRES science gateway. Evol. Bioinform Online 11, 43–48. https://doi.org/10.4137/EBO.S21501.

Paps, J., Holland, P., 2018. Reconstruction of the ancestral metazoan genome reveals an increase in genomic novelty. Nat. Commun. 9, 1730. https://doi.org/10.1038/s41467-018-04136-5.

Putnam, N., Butts, T., Ferrier, D., Furlong, R., Hellsten, U., Kawashima, T., Robinson-Rechavi, M., Shoguchi, E., Terry, A., Yu, J., Benito-Gutiérrez, E., Dubchak, I., Garcia-Fernàndez, J., Gibson-Brown, J., Grigoriev, I., Horton, A., de Jong, P., Jurka, J., Kapitonov, V., Kohara, Y., Kuroki, Y., Lindquist, E., Lucas, S., Osoegawa, K., Pennacchio, L., Salamov, A., Satou, Y., Sauka-Spengler, T., Schmutz, J., Shin-I, T., Toyoda, A., Bronner-Fraser, M., Fujiyama, A., Holland, L., Holland, P., Satoh, N., Rokhsar, D., 2008. The amphioxus genome and the evolution of the chordate karyotype. Nature 453, 1064–1071.

Ranallo-Benavidez, T., Jaron, K., Schatz, M., 2019. GenomeScope 2.0 and Smudgeplots: Reference-free profiling of polyploid genomes. bioRxiv. https://doi.org/10.1101/747568.

Robinson-Rechavi, M., Carpentier, A., Duffraisse, M., Laudet, V., 2001. How many nuclear hormone receptors are there in the human genome? Trends Genet. 17, 554–556. https://doi.org/10.1016/S0168-9525(01)02417-9.

Román, E., González-Iglesias, C., González-Troncoso, D., 2019. Results for the Spanish Survey in the NAFO Regulatory Area of Division 3L for the period 2003-2018. NAFO SCR Doc No. 12, Se, 1–78.

Schmitz, J., Zimmer, F., Bornberg-Bauer, E., 2016. Mechanisms of transcription factor evolution in Metazoa. Nucleic Acids Res. 44, 6287–6297. https://doi.org/10.1093/nar/gkw492.

Schubert, M., Brunet, F., Paris, M., Bertrand, S., Benoît, G., Laudet, V., 2008. Nuclear hormone receptor signaling in amphioxus. Dev. Gene. Evol. 218, 651–665. https://doi.org/10.1007/s00427-008-0251-y.

Siddiq, M., Hochberg, G., Thornton, J., 2017. Evolution of protein specificity: Insights from ancestral protein reconstruction. Curr. Opin. Struct. Biol. 47, 113–122. https://doi.org/10.1016/j.sbi.2017.07.003.

Simão, F., Waterhouse, R., Ioannidis, P., Kriventseva, E., Zdobnov, E., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212. https://doi.org/10.1093/bioinformatics/btv351.

Simpson, J., 2014. Exploring genome characteristics and sequence quality without a reference. Bioinformatics 30, 1228–1235. https://doi.org/10.1093/bioinformatics/btu023.

Song, L., Florea, L., Langmead, B., 2014. Lighter: fast and memory-efficient sequencing error correction without counting. Genome Biol. 15, 509. https://doi.org/10.1186/s13059-014-0509-9.

Thornton, J., 2001. Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. Proc. Natl. Acad. Sci. U.S.A. 98, 5671–5676. https://doi.org/10.1073/pnas.091553298.

Venkatesh, B., Kirkness, E., Loh, Y.-H., Halpern, A., Lee, A., Johnson, J., Dandona, N., Viswanathan, L., Tay, A., Venter, J., Strausberg, R., Brenner, S., 2007. Survey sequencing and comparative analysis of the elephant shark (Callorhinchus milii) genome. PLoS Biol. 5, e101. https://doi.org/10.1371/journal.pbio.0050101.

Venkatesh, B., Kirkness, E., Loh, Y.-H., Halpern, A., Lee, A., Johnson, J., Dandona, N., Viswanathan, L., Tay, A., Venter, J., Strausberg, R., Brenner, S., 2006. Ancient noncoding elements conserved in the human genome. Science (80-). 314, 1892. https://doi.org/10.1126/science.1130708.

Venkatesh, B., Lee, A., Ravi, V., Maurya, A., Lian, M., Swann, J., Ohta, Y., Flajnik, M., Sutoh, Y., Kasahara, M., Hoon, S., Gangu, V., Roy, S., Irimia, M., Korzh, V., Kondrychyn, I., Lim, Z., Tay, B., Tohari, S., Kong, K., Ho, S., Lorente-Galdos, B., Quilez, J., Marques-Bonet, T., Raney, B., Ingham, P., Tay, A., Hillier, L., Minx, P., Boehm, T., Wilson, R., Brenner, S., Warren, W., 2014. Elephant shark genome provides unique insights into gnathostome evolution. Nature 505, 174–179. https://doi.org/10.1038/nature12826.

Ward, R., Zemlak, T., Innes, B., Last, P., Hebert, P., 2005. DNA barcoding Australia's fish species. Philos Trans R Soc L. B Biol Sci 360, 1847–1857. https://doi.org/10.1098/

rstb.2005.1716.
Wu, W., Niles, E., Hirai, H., LoVerde, P., 2007. Evolution of a novel subfamily of nuclear receptors with members that each contain two DNA binding domains. BMC Evol. Biol. 7, 27. https://doi.org/10.1186/1471-2148-7-27.
Yagi, K., Satou, Y., Mazet, F., Shimeld, S., Degnan, B., Rokhsar, D., Levine, M., Kohara, Y., Satoh, N., 2003. A genomewide survey of developmentally relevant genes in Ciona intestinalis: III. Genes for Fox, ETS, nuclear receptors and NFkappaB. Dev. Genes. Evol. 213, 235–244. https://doi.org/10.1007/s00427-003-0322-z.
Zhao, Y., Zhang, K., Giesy, J., Hu, J., 2015. Families of nuclear receptors in vertebrate models: characteristic and comparative toxicological perspective. Sci. Rep. 5, 8554. https://doi.org/10.1038/srep08554.