

RESEARCH ARTICLE

Identification and evolution of nuclear receptors in Platyhelminths

Wenjie Wu[‡], Philip T. LoVerde^{‡*}

Departments of Biochemistry and Structural Biology and Pathology and Laboratory Medicine, University of Texas Health Sciences Center, San Antonio, Texas, United States of America

[‡] Current address: Canget BioTekpharma LLC, NY State Center of Excellence in Bioinformatics & Life Sciences, Buffalo, New York, United States of America* loverde@uthscsa.edu

OPEN ACCESS

Citation: Wu W, LoVerde PT (2021) Identification and evolution of nuclear receptors in Platyhelminths. PLoS ONE 16(8): e0250750. <https://doi.org/10.1371/journal.pone.0250750>

Editor: Mariya Y Pakharukova, Institute of Cytology and Genetics, RUSSIAN FEDERATION

Received: April 6, 2021

Accepted: July 18, 2021

Published: August 13, 2021

Copyright: © 2021 Wu, LoVerde. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting information](#) files.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: 2DBD-NR, nuclear receptor with two DBDs; AR, androgen receptor; CAR, constitutive androstane receptor; COUP-TF, chicken ovalbumin upstream promoter transcription factor; DBD, DNA binding domain;

Abstract

Since the first complete set of Platyhelminth nuclear receptors (NRs) from *Schistosoma mansoni* were identified a decade ago, more flatworm genome data is available to identify their NR complement and to analyze the evolutionary relationship of Platyhelminth NRs. NRs are important transcriptional modulators that regulate development, differentiation and reproduction of animals. In this study, NRs are identified in genome databases of thirty-three species including in all Platyhelminth classes (Rhabditophora, Monogenea, Cestoda and Trematoda). Phylogenetic analysis shows that NRs in Platyhelminths follow two different evolutionary lineages: 1) NRs in a free-living freshwater flatworm (*Schmidtea mediterranea*) and all parasitic flatworms share the same evolutionary lineage with extensive gene loss. 2) NRs in a free-living intertidal zone flatworm (*Macrostomum lignano*) follow a different evolutionary lineage with a feature of multiple gene duplication and gene divergence. The DNA binding domain (DBD) is the most conserved region in NRs which contains two C4-type zinc finger motifs. A novel zinc finger motif is identified in parasitic flatworm NRs: the second zinc finger of parasitic Platyhelminth HR96b possesses a CHC2 motif which is not found in NRs of all other animals studied to date. In this study, novel NRs (members of NR subfamily 3 and 6) are identified in flatworms, this result demonstrates that members of all six classical NR subfamilies are present in the Platyhelminth phylum. NR gene duplication, loss and divergence in Platyhelminths are analyzed along with the evolutionary relationship of Platyhelminth NRs.

1. Introduction

Platyhelminths (flatworms) are one of the largest animal phyla, which includes more than 20,000 species [1, 2]. Some flatworms are important disease-causing parasites of humans and livestock, e.g. Schistosomiasis, Paragonimiasis and Cestodiasis. Platyhelminths are bilaterally symmetrical, non-segmented acoelomates without an anus. Although they have an excretory system, they lack respiratory and circulatory systems. In addition, all flatworms are hermaphroditic, undergoing both asexual and sexual reproduction, with the exception of members in

DSF, *Drosophila* dissatisfaction gene; **E75**, *Drosophila* ecdysone-induced protein 75; **E78**, *Drosophila* ecdysone-induced protein 78; **EAR**, V-erbA-related gene; **EGON**, embryonic gonad protein; **ER**, estrogen receptor; **ERR**, estrogen related receptor; **FTZ-F1**, fushi tarazu-factor 1; **FXR**, farnesoid X receptor; **GCNF**, germ cell nuclear factor; **GR**, glucuronide receptor; **HNF-4**, hepatocyte nuclear factor 4; **HR96**, *Drosophila* hormone receptor 96; **LBD**, ligand binding domain; **LRH-1**, liver receptor homologue 1; **LXR**, liver X receptor; **MR**, mineralocorticoid receptor; **NGFI-B**, nerve growth factor-induced B; **NHR236**, nematode nuclear receptor 236; **NOR1**, neuron-derived orphan receptor 1; **NR**, Nuclear receptor; **Nurr1**, nuclear receptor related 1; **PNR**, photoreceptor-specific nuclear receptor; **PPAR**, peroxisome proliferators-activated receptor; **PR**, progesterone receptor; **PXR**, pregnane X receptor; **RAR**, retinoic acid receptor; **Rev-erb**, nuclear receptor-related protein coded on the opposite strand of the thyroid hormone receptor gene; **ROR**, retinoid-related orphan receptor; **RXR**, retinoid X receptor; **SF1**, Steroidogenic factor 1; **SR**, steroid receptor; **SupNRs**, supplementary nuclear receptors; **SVP**, seven up; **TLL**, *Drosophila* tailless gene; **TLX**, homologue of the *Drosophila* tailless gene; **TR**, thyroid hormone receptor; **TR2/4**, testicular receptor 2/4; **USP**, ultraspiracle; **VDR**, vitamin D receptor.

the Schistosomatoidea [3]. Platyhelminths are traditionally divided into four classes: Rhabditophora, Monogenea, Cestoda (tapeworms) and Trematoda (flukes). The class Rhabditophora includes all free-living flatworms, while all members in classes of Monogenea, Trematoda, and Cestoda are parasitic flatworms. From an evolutionary point of view, the parasitic classes arose from a primitive free-living flatworm [3].

Nuclear receptors (NRs) are important transcriptional modulators that regulate development, differentiation and reproduction of animals. Most of NRs share a common tertiary structure: A/B-C-D-E domains. The A/B domain is highly variable, the C domain is the DNA-binding domain (DBD) which is the most conserved region containing two zinc finger motifs, the D domain a flexible hinge between the C and E domains and is poorly conserved, the E domain contains the ligand binding domain (LBD) which is involved in transcriptional activation. Atypical NRs are found in some animals, e.g. NRs with a DBD but no LBD are found in arthropods and nematodes, members without a DBD but with a LBD are present in some vertebrates, and NRs with two DBDs and a single LBD (2DBD-NRs) are identified in protozoans. A phylogenetic analysis of the NRs divides them into six classical subfamilies by alignment of the conserved DBD [4]. The early study of the complete genome sequence of ecdysozoan Arthropods (*Drosophila melanogaster*) [5] the mosquito (*Anopheles gambiae*) [6], free-living nematodes (*Caenorhabditis elegans* and *C. briggsae*) [7, 8], tunicata (*Ciona intestinalis*) [9], mammals: rat (*Rattus norvegicus*), mouse (*Mus musculus*) [10] and human (*Homo sapiens*) [11] revealed that NRs in insects, tunicata and mammals share the same evolutionary lineage with extensive gene loss/duplication, while NRs in nematodes follow a different evolutionary lineage with a feature of multiple duplication of SupNRs and gene loss.

Since we identified the first complete set of Platyhelminths nuclear receptors (NRs) from *Schistosoma mansoni* a decade ago [12], more flatworm genome data has become available to identify their NR complement. Analysis of the NR complement of the blood fluke *Schistosoma mansoni* [12–14] and tapeworm *Echinococcus multilocularis* [13–15] shows that NRs in *S. mansoni* and *E. multilocularis* share the same evolutionary lineage as that of Deuterostomia and the arthropods in the Ecdysozoan clade of the Protostomia, but some divergent NRs in flatworms are not present in Deuterostomia or/and arthropods, for example 2DBD-NRs [16]. In this study, we analyzed genome databases of Platyhelminth Wormbase ParaSite including species from all of the four classes in Platyhelminths [17, 18]. Identification of the NR complement will contribute to a better understanding of the evolution of NRs in Platyhelminths.

A schematic that illustrates the different NR subfamilies and NRs in human, insects, Mollusca and Platyhelminths are shown in Fig 1.

2. Materials and methods

2.1. Data mining

Nuclear receptors in Platyhelminths were mined from the genome databases in Wormbase ParaSite (version: WBPS15 (WS276)) [17, 18]. Amino acid sequence of DBD of SmTRa (AY395038), SmHR96b (AY688259) and Sm2DBD α (AH013462) were used as the query to tblastn (with E-value threshold: 1e-1) against all available databases. Any sequence that contains a zinc finger structure of the NR DBD (Cys-X2-Cys-X13-Cys-X2-Cys or Cys-X5-Cys-X9-Cys-X2-Cys) was retained, and the deduced amino acid sequence of the DBD was obtained by the conserved junction position (JP) and GT-AG rule if there was an intron in the DBD coding region. The analyzed Platyhelminth genome databases are listed in Table 1. The classification of Platyhelminths and a short description of each Platyhelminth species can be found in Table 1.

Subfamily	Nomenclature	Name	Human	Insects	Mollusca	Platyhelminth	LB	DB
1	NR1A	TR	█	x	█	█	◆	RXR, H, M, H
	NR1B	RAR	█	x	█	x	◆	RXR
	NR1C	PPAR	█	x	█	x	◆	RXR
	NR1D	Rev-erb/E75	█	█	█	x		H
	NR1E	E78	x	█	█	█		M
	NR1F	ROR/HR3	█	█	█	x		RXR
	NR1H	LXR/FXR/EcR	x	█	█	x	◆	RXR
	NR1I	VDR/CAR/PXR	█	x	x	x	◆	RXR
	NR1J	HR96	x	█	█	█		RXR, M
	2	NR2A	HNF4	█	█	█	█	◆
NR2B		RXR/USP	█	█	█	█	◆	H, Heter, Heter, H, M
NR2C/D		TR2/TR4/HR78	█	█	█	█		H, H, M
NR2E1/2		TLX/TLL	█	█	█	█		M, H
NR2E3		PNR	█	█	█	█		H
NR2E4		DSF	█	█	█	█		
NR2E5		fax-1	█	█	█	▬		
*NR2E6		HR236	x	█	█	█		
NR2F	Coup-TF/SVP	█	█	█	█		H	
3	NR3A	ER	█	x	█	█	◆	H
	NR3B	ERR	█	█	█	▬		M, H
	NR3C	GR/MR/PR/AR	█	x	x	x	◆	H, M
4	NR4A	NGFI-B/HR38	█	█	█	█		RXR, M
5	NR5A	FTZ-F1	█	█	█	█	◆	M, M
	NR5B	DHR39/FTZ-F1b	x	█	█	█		M
6	NR6A	GCNF1/DHR4	█	█	█	▬		H
0	NR0A	KNI/KNRL/EGON	█	█	x	x		
	NR0B	DAX1/SHP	█	x	█	x		
2DBD			x	x	█	█		H
7/8			x	x	█	x		

Fig 1. A schematic illustrates the different NR subfamilies and NRs in human, insects, Mollusca and Platyhelminths. Solid color box shows NRs in different animals, black box shows the new NRs found in Platyhelminths in this study and X indicates NRs loss in different animals. LB: Ligand binding, DB: DNA binding, diamond on ligand binding column indicates the NR was known to bind its ligand. On a DNA binding column, RXR indicates the NR can form a heterodimer with RXR, H: Homodimer, Heter: Heterodimer with other NRs, M: Monomer. On DNA binding column, the red character indicates DNA binding of Platyhelminth NRs studied *in vitro* [5, 6, 11, 16, 19–35].

<https://doi.org/10.1371/journal.pone.0250750.g001>

2.2. Phylogenetic analysis

Phylogenetic trees were constructed from deduced amino acid sequences of the DBD and/or ligand binding domain (LBD), the sequences are aligned with ClustalW [36], phylogenetic analysis of the data set is carried out using Bayesian inference (MrBAYES v3.1.1) [37] with a mix amino acid replacement model + gamma rates. The trees were started randomly; four simultaneous Markov chains were run for 5 million generations, the trees were sampled every 100 generations, the Bayesian posterior probabilities (BPPs) were calculated using a Markov chain Monte Carlo (MCMC) sampling approach implemented in MrBAYES v3.1.1, with a burn-in value setting at 12,500. The same data set was also tested by Maximum Likelihood method using PhyML 3.0 [38] with support values obtained by bootstrapping a 1000 replicates. The most optimal evolutionary models for phylogenetic reconstruction of NRs in each data set were determined by AIC criteria with Smart Model Selection in PhyML (SMS) [39] and the tested model for each data set was shown in each figure legend. For GenBank accession numbers of published NR sequences used in this study see S1 File.

Table 1. Platyhelminth species analyzed in Wormbase ParaSite (version: WBPS15 (WS276)).

CLASS	ORDER	FAMILY	SPECIES REFERENCES	SHORT DESCRIPTION
Rhabditophora	Rhabditophora	Rhabditophorae	<i>Macrostomum lignano</i> (Ml) [40, 41]	Free-living, intertidal flatworm that can regenerate most of its body parts.
	Tricladida	Dugesidae	<i>Schmidtea mediterranea</i> (Sme) [42–44]	Free-living, freshwater flatworm that can regenerate an entire organism.
Monogenea	Monopisthocotylea	Gyrodactylidae	<i>Gyrodactylus salaris</i> (Gs) [45]	Salmon fluke that lives on the body surface of freshwater fish.
	Polyopisthocotylea	Polystomatidae	<i>Protopolystoma xenopodis</i> (Px) [46]	Fluke of African clawed frogs where the adult worms live in the host's urinary bladder.
Cestoda	Cyclophyllidea	Hymenolepididae	<i>Hymenolepis diminuta</i> (Hd) [46]	Rat tapeworm. The intermediate hosts: arthropod; the definitive hosts: rodent.
			<i>H. microstoma</i> (Hm) [47]	Rodent tapeworm. The intermediate hosts: arthropod; the definitive hosts: rodent.
			<i>H. nana</i> (Hn) [46]	Rodent tapeworm. The intermediate hosts: arthropod; the definitive hosts: human and rodent.
		Taeniidae	<i>Echinococcus Canadensis</i> (Ec) [48]	Canid tapeworm. The intermediate hosts: cervid, camels, cattle etc.; the definitive host: dogs and other canids.
			<i>E. granulosus</i> (Eg) [47, 49]	Dog tapeworm. The intermediate hosts: human, horse; the definitive host: dogs.
			<i>E. multilocularis</i> (Em) [47]	Fox tapeworm. The intermediate hosts: wild rodents; the definitive hosts: foxes, coyotes, dogs and other canids.
			<i>Taenia asiatica</i> (Ta) [50]	Human tapeworm. The intermediate hosts: pigs; the definitive host: human.
			<i>T. multiceps</i> (Tm) [51]	Dog tapeworm. The intermediate hosts: sheep, goat and cattle; the definitive hosts: Canids.
			<i>T. saginata</i> (Ts) [50]	Beef tapeworm. The intermediate hosts: Cattle; the definitive hosts: human.
			<i>T. solium</i> (Tso) [47]	Pork tapeworm. The intermediate host: pigs; the definitive host: human.
			<i>Hydatigera taeniaeformis</i> (Ht) [46]	Cat tapeworm. The intermediate hosts: rodents; the definitive hosts: cats.
		Mesocestoididae	<i>Mesocestoides corti</i> (Mc) [46]	Mouse tapeworm. The intermediate host: birds and small mammals; the definitive host: cats and dogs.
	Diphyllobothridea	Diphyllobothriidae	<i>Diphyllobothrium latum</i> (Dl) [46]	Human tapeworm. The intermediate hosts: fish; the definitive hosts: human.
			<i>Schistocephalus solidus</i> (Ss) [46]	Bird tapeworm. The first intermediate hosts: cyclopoid copepod; the second intermediate hosts: fish; the definitive hosts: fish-eating water bird.
			<i>Spirometra erinaceieuropaei</i> (Se) [52]	Tapeworm. The first intermediate hosts: cyclopoid copepod; the second intermediate hosts: fish, reptiles, or other amphibians; the definitive hosts: cats and dogs.
Trematoda	Opisthorchida	Opistorchiidae	<i>Clonorchis sinensis</i> (Cs) [53]	Human liver fluke. The first intermediate hosts: freshwater snails; the second intermediate hosts: freshwater fish; the definitive hosts: human and other mammals.
			<i>Opisthorchis felinus</i> (Of) [54]	Cat liver fluke. The first intermediate hosts: freshwater snails; the second intermediate hosts: freshwater fish; the definitive hosts: human and other mammals.
			<i>O. viverrini</i> (Ov) [55]	Human liver fluke. The first intermediate hosts: freshwater snails; the second intermediate hosts: freshwater fish; the definitive hosts: human and other mammals.
	Echinostomida	Echinostomatidae	<i>Echinostoma caproni</i> (Eca) [46]	Intestinal fluke. The first intermediate hosts: freshwater snails; the second intermediate hosts: freshwater snails or frogs. Humans can be the definitive host.
		Fasciolidae	<i>Fasciola hepatica</i> (Fh) [56, 57]	Liver fluke. The intermediate hosts: freshwater snails; The definitive hosts: cattle, sheep and other mammals including human.
	Strigeidida	Schistosomatidae	<i>Schistosoma bovis</i> (Sb) [58]	Blood fluke. The intermediate hosts: freshwater snails; the definitive host: ruminants (cattle, goats, sheep, etc).

(Continued)

Table 1. (Continued)

CLASS	ORDER	FAMILY	SPECIES REFERENCES	SHORT DESCRIPTION
			<i>S. curassoni</i> (Sc) [46]	Blood fluke. The intermediate hosts: freshwater snails; the definitive hosts: ruminants (cattle, goat, sheep, etc).
			<i>S. haematobium</i> (Sh) [59]	Blood fluke. The intermediate hosts: freshwater snails; the definitive hosts: cattle, goat or sheep.
			<i>S. japonicum</i> (Sj) [60]	Zoonotic blood fluke. The intermediate hosts: freshwater snails; the definitive hosts: mammals including humans.
			<i>S. mansoni</i> (Sm) [61]	Human blood fluke. The intermediate hosts: freshwater snails; the definitive hosts: human.
			<i>S. margrebowiei</i> (Sma) [46]	Blood fluke. The intermediate hosts: freshwater snails; the definitive hosts: antelope, buffalo and waterbuck.
			<i>S. mattheei</i> (Sma) [46]	Blood fluke. The intermediate hosts: freshwater snails; the definitive hosts: bovid.
			<i>S. rodhaini</i> (Sr) [46]	Blood fluke. The intermediate hosts: freshwater snails; the definitive host: rodent.
			<i>Trichobilharzia regenti</i> (Tr) [46]	Neuropathogenic fluke. The intermediate host: freshwater snail; the definitive host: avian.

<https://doi.org/10.1371/journal.pone.0250750.t001>

3. Results

3.1. NR complements in Platyhelminths

NRs are identified in the genomes of 33 Platyhelminth species including 2 species (*Schistosoma mansoni* and *Echinococcus multilocularis*) that have had their NR complement reported [12, 15]. The analyzed flatworm species shown in Table 1, include 2 species from class Rhabditophora, 2 species from class Monogenea, 15 species from class Cestoda and 14 species from class Trematoda. The number of NRs varies from 15 to 61 in Platyhelminths, 18–23 NRs are present in Monogenea, 15–20 NRs are present in Cestoda, 21–22 NRs are found in Trematoda and 27–61 members are identified in Rhabditophora. For the amino acid sequences of the DBD identified in each Platyhelminth species see S2 File.

3.2. Phylogenetic analysis of NRs

Phylogenetic analysis of NRs using DBD sequences were carried out with Bayesian inference and Maximum Likelihood method. The DBD sequences of NRs from human, *Drosophila* and other animals were used as controls. Both methods place control NRs in mono groups and in correct families. Comparison of Bayesian inference and Maximum Likelihood method indicates Bayesian inference highly supports the analysis but Maximum Likelihood sometimes shows a lower support value (S3 File). This result indicates that Bayesian inference supports phylogenetic analysis of NRs using only DBD sequences.

The amino acid sequence of the P-box in the DBD is unique to NR members, groups or subfamilies. This sequence is helpful for identification of NRs when the phylogenetic analysis support value is low. For example, P-box sequence ESCKG is unique to NR subfamily 5 members (FTZ-F1 and DHR 39 orthologues). In this study, phylogenetic analysis of *G. salaris* NRs shows a lower support value for GsFTZ-F1 in subfamily 5 (S3 File), but the P-box sequence of GsFTZ-F1 (ESCKG) clearly demonstrated that it belongs to subfamily 5. The P-box sequence of all analyzed Platyhelminths NRs were checked to make sure that they share the same sequence with that of same group members. NRs in each Platyhelminth species are phylogenetic analyzed and the phylogenetic trees are reconstructed (S1–S17 Figs). NRs in Platyhelminths are summarized in Table 2.

Table 2. Nuclear receptors identified in Platyhelminths.

CLASS	ORDER	FAMILY	Genus/ Species	Platyhelminths				
				Rhabditophora		Monogenea		
				Macrostomida	Tricladida	Polyopisthocotylea	Monopisthocotylea	
			<i>H. sapiens</i>	<i>D. melanogaster</i>	Macrostomidae	Dugesidae	Polystomatidae	Gyrodactylidae
					Macrostomum	Schmidtea	Protopolystoma	Gyrodactylus
					<i>lignano</i>	<i>mediterranea</i>	<i>xenopodis</i>	<i>salaris</i>
NR1A	TR		2		2	3	1	1
NR1B	RAR		3					
NR1C	PPAR		3					
NR1D	Rev-erb/E75		2	1				
NR1E	E78			1	2		1	1
NR1F	ROR		3	1				
NR1H	ECR/LXR/ FXR		3	1				
NR1I	VDR/PXR/CAR		3					
NR1J	DHR96			1	16	3	3	2
NR1a					1	1	1	
NR1b						1	1	
NR2A	HNF4		2	1	1	1	1	1
NR2B	RXR/USP		3	1	2	2	2	1
NR2C/D	TR2/TR4/DHR78		2	1	2	1	1	1
NR2E1/2	TLX/TLL		1	1	1	1	1	1
NR2E3	PNR		1	1	1	1	1	1
NR2E4	DSF			1	1	1	1	1
NR2E5	fax-1			1	1			
NR2E new	NHR236				2	1	1	
NR2F	COUP-TF/SVP		3	1	2	2	2	2
NR3A	ER		2					
NR3B	ERR		3	1	4			
NR3C	MR/PR/AR/GR		4					
NR4A	NGFI-B/DHR38		3	1	2	2	1	1
NR5A	FTZ-F1		2	1		1	1	
NR5B	DHR39			1	1	1	1	1
NR6A	GCNF1		1	1	1			
NR0A/B	Knirps/KNRL/ EGON/DAX1/SHP		2	3				
2DBD-NR	2DBD-NR				2	4	3	3
Other					17	1		1
Total			48	21	61	27	23	18

Platyhelminths							
Cestoda						Trematoda	
Cyclophyllidea			Diphylobothriidea			Opisthorchida	Strigeidida
						Echinostomida	
Hymenolepididae	Mesocestoididae	Taeniidae	Diphylobothriidae				
<i>Hymenolepis</i>	<i>Mesocestoides</i>	<i>Echinococcus</i>	<i>Diphylobothrium</i>	<i>Schistocephalus</i>	<i>Spirometra</i>		
	<i>corti</i>	<i>Hydatigera</i>	<i>latum</i>	<i>solidus</i>	<i>erinacei</i>		
		<i>Taenia</i>			<i>europaei</i>		
1	2	1	1	1	1	2	2
1	1	1	1	1	1	1	1
3	3	3	3	3	3	3	2

(Continued)

Table 2. (Continued)

	1			1	1	1	1
1	1	1	1	1	1		
1	1	1	1	1	1	1	1
			1	2	1	2	2
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
			1		1	1	1
						1	1
		1	1	1	1	2	2
1	1	1	1	1	1	1	1
		1	1	1	1	1	1
1	1	1	1	1	1	1	1
3	3	3	3	3	3	3	3
1	5	1	1	1	1		
15	21	17	19	20	20	22	21

<https://doi.org/10.1371/journal.pone.0250750.t002>

3.3. A novel zinc finger motif (CHC2) exists in DBD of parasitic Platyhelminth NRs

DBD is the most conserved region in NRs, it contains two C4-type zinc finger motifs. In each motif, four cysteine residues chelate one Zn^{2+} ion. The first zinc finger (CI) contains a sequence element, the P-box [62, 63] which is responsible for binding the target gene, and the second zinc finger (CII) contains a sequence element the D-box which is responsible for dimerization [62].

Previously, we isolated a partial cDNA of a trematode *Schistosoma mansoni* HR96b (SmNR96b) [12], unlike other NRs, SmNR96b contains a long amino acid sequence in the DBD with two introns located in this region. DBD of SmHR96b has a histidine residue replaced by a second cysteine residue in the D-Box in CII and it forms a novel zinc finger motif (CHC2). Recently, CHC2 zinc finger motif has been demonstrated to be present in HR96b of Cestoda and other Trematoda species [64, 65]. In this study, we determined that all SmHR96b orthologues in parasitic Platyhelminths contain this novel CHC2 zinc finger motif (Fig 2). Thus, CHC2 type zinc finger motif represents a novel NR CII which has diverged in ancient HR96b in a common ancestor of parasitic Platyhelminths. The function of this novel CHC2 type motif in NRs is unknown. Recent study shows that SmHR96b (named Vitellogenic Factor 1 in [64, 65]) is essential for female sexual development.

3.4. New NRs identified in Platyhelminths

Four members of NR subfamily 3 (MIERRs), one member of subfamily 6 (MINR6) and an orthologue of fax-1 (Mlfax-1) are identified in Rhabditophora *Macrostomum lignano* (S1 Fig), this result indicates that members of NR subfamily 3 and subfamily 6 are present in Platyhelminths. The identification of fax-1 adds a new member of NR2E super-group in Platyhelminths NRs. Phylogenetic analysis of the four *M. lignano* NR3 (MIERRs) shows that all of them are clustered together as a separate group within NR3 subfamily with a high ML/Bayesian support value (Fig 3). Though MIERRs clustered within ERR group with a low support value (BPP = 0.56), all of the MIERRs possess a P-box (EACKA) that is unique to ERR (Fig 3), thus the MIERRs may represent a new group related human/Drosophila ERRs.

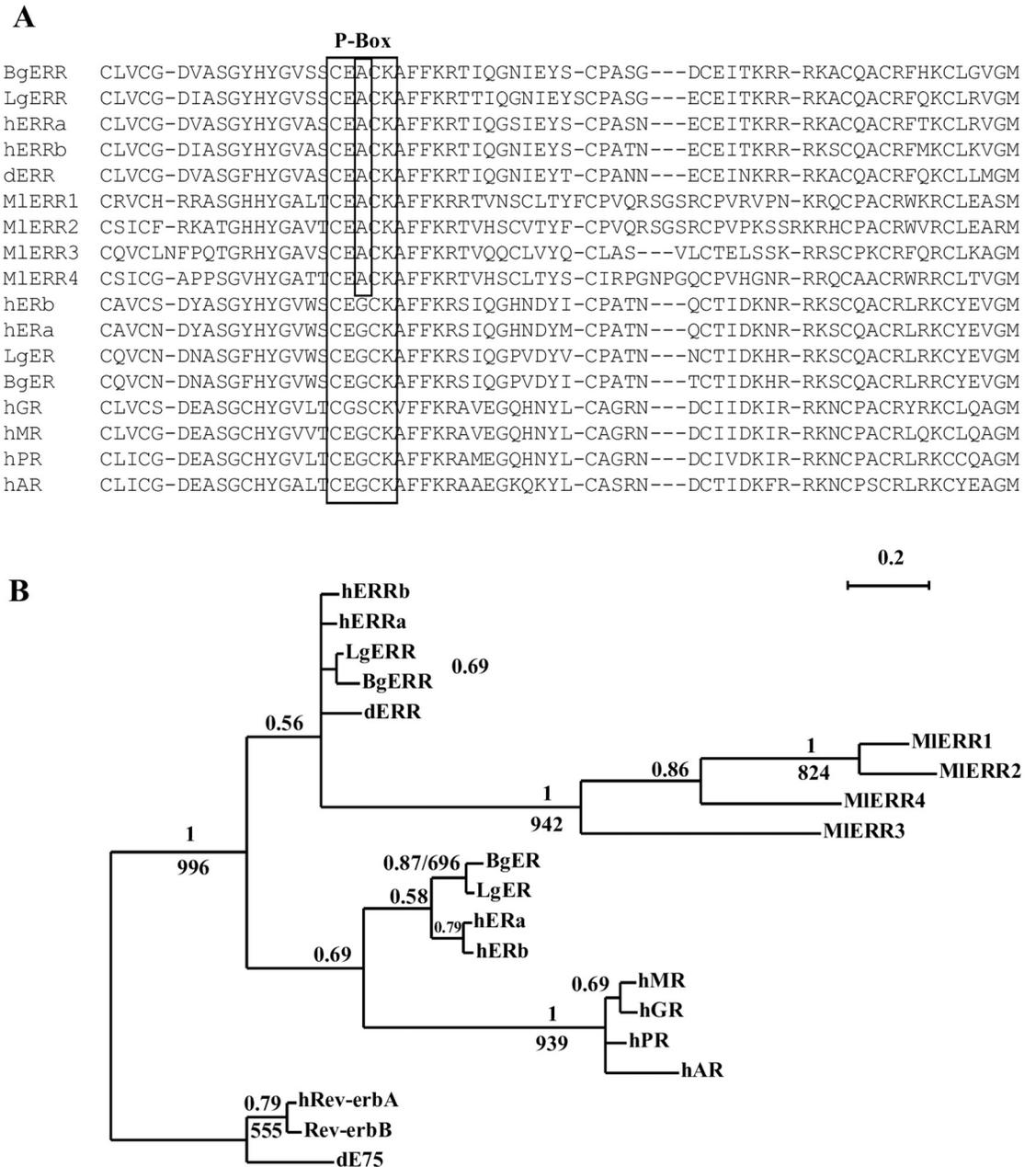


Fig 3. ERRs in Platyhelminths. A) Amino acid sequence alignment of DBD of MIERRs. **B)** Bayesian phylogenetic tree of MIERRs, the BPPs values are shown above each branch. Branches under the PPs 0.5 are shown as polytomies. The same data set was also tested by ML method using PHYML (v2.4.4) (under JTT +G model (equilibrium frequencies model, proportion of invariable sites was estimated as fixed 0.0, number of substitution rate categories: 4, gamma shape parameter was estimated as 0.547). Support values for the tree were obtained by bootstrapping a 1,000 replicates and bootstrap values above 500 are indicated below each branch (or after MrBAYES BPPs separated by Slash). Bg: *Biomphalaria glabrata*, d: *Drosophila melanogaster*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Ml: *Macrostomum lignano*.

<https://doi.org/10.1371/journal.pone.0250750.g003>

LBD sequences or only LBD sequence of NR1b orthologue (EgHR3 and SmeNR1b) were analyzed. Phylogenetic analysis shows that EgHR3 and SmeNR1b are clustered with the E75 group if both DBD and LBD sequences are analyzed, but they are clustered with E78 group if only LBD sequences are analyzed. These results suggest that the orthologues of Platyhelminths

NR1b are a group of divergent NRs that were present in a common Platyhelminth ancestor (S19 and S20 Figs).

3.5. NR genes lost in Platyhelminths

Comparison of NRs in Platyhelminths, orthologues of *RAR*, *PPAR*, *E75*, *ROR*, *ER*, *MR*, *PR*, *GR*, *AR*, *Knir* and *NR7/8* (a new identified subfamily [26, 29]) are missing in all analyzed Platyhelminth species. *Fax-1* is only identified in Rhabditophora *M. lignano*, it was lost in Rhabditophora *S. mediterranea* and all analyzed parasitic Platyhelminths and NHR236 orthologue is not identified in Cestoda and trematode species. In Cestoda, *RXR* and *PNR* are present in Diphylobothriidae Order but they are missing in Cyclophyllidea Order. NR gene lost in Platyhelminths is shown in Table 2.

3.6. Divergent NRs

Divergent NR refers to a NR which has a typical P-box sequence in the DBD but does not fall into any 'typical' NR groups, for example Platyhelminths NR1a and NR1b with a typical P-box of 'CEGCKGFFRR' belonging to NR subfamily 1 but they do not fit into any groups within the NR1 subfamily. It also refers to a NR which has an atypical P-box sequence in the DBD and does not fall into the present NR nomenclature [66].

Divergent NRs exist in various lineages of Platyhelminths. One divergent NR with a typical P-box (CEGCKGFFKR) which is the same as that of *RXR/TR4/NR4A* is found in Rhabditophora *S. mediterranea* and in Cestoda *Mesocestoides corti*. Most of the Platyhelminth divergent NRs possess an 'atypical' P-box sequence. For example, a NR with a P-box of 'CEPCKVFFKR' is identified in Monogenea *G. salaris*, a NR with a P-box of 'CEACKAFFQQ' is found in all analyzed species of Cestoda Hymenolepis family, a NR which has a P-box of 'CDSCRAFFEM' exists in Cestoda Taenia family, a NR with a P-box of 'CEACKSFFKR' is found in Cestoda Diphylobothriidea order and seventeen NRs with various 'atypical' P-box sequences were identified in Rhabditophora *M. lignano* (S21 Fig).

A phylogenetic tree of the divergent NRs including those of Platyhelminths and Mollusca was constructed (S21 Fig). Phylogenetic analysis shows no orthologues exist among Rhabditophora *M. lignano*, Cestoda and Mollusca. The result suggests that these divergent NRs diverged and duplicated independently in different animal lineages.

3.7. NR gene duplication in Platyhelminths

3.7.1. TR. One TR is identified in Monogenea; two are identified in Rhabditophora *M. lignano* and three are identified in *S. mediterranea*; and two are identified in Cestoda and Trematoda, respectively. Our previous study showed that two TR homologues are present in Platyhelminths [12–14, 34]. Phylogenetic analysis suggested that Platyhelminth TR gene duplicated after the split of the trematodes and the turbellarians [34], thus the paralogue of the trematode TR was not present in a turbellarian and in turn the paralogue of turbellarian TR was not present in trematode species. Phylogenetic analysis in this study shows that all TRs of parasitic Platyhelminths are clustered in a group, two *M. lignano* TRs are clustered in a group and three *S. mediterranea* TRs are clustered in another group. This result suggests that TRs duplicated independently in *M. lignano*, *S. mediterranea* and parasitic Platyhelminths. In parasitic Platyhelminth TR groups, trematode TRa (orthologues of *Schistosoma* TRa, SmTRa) are clustered together with those of Monogenea and Cestoda, but trematode TRb (orthologues of *Schistosoma* TRb, SmTRb) group only contain trematode TRs. This result suggested that one TR was present in a common ancestor of Platyhelminths and trematode TRa was an ancient TR gene. Phylogenetic analysis shows that the two Cestode *M. corti* TRs are clustered together

in Cestoda TR group, this result suggests that TRs in Cestoda and Trematoda underwent duplication independently (S22 Fig). Since TR duplicated independently in different Platyhelminths lineage, Platyhelminths TRb/TRc are paralogous genes.

3.7.2. HR96. Sixteen HR96s are identified in Rhabditophora *M. lignano* and two are found in *S. mediterranea*; three HR96s are identified in Monogenea, Cestoda and Trematoda, respectively. Phylogenetic analysis shows that Platyhelminths HR96s are clustered in four different groups: HR96a, HR96b, HR96c and HR96d. HR96a, HR96b and HR96c groups contain only Platyhelminth HR96s, but HR96d group contains both Platyhelminth and Mollusca members. The result suggests that Platyhelminth HR96d is an ancient NR gene (Fig 4). HR96d group only contains Platyhelminth *M. lignano* HR96s, but each group of HR96a, HR96b and HR96c contains NR96s from all species of the four classes of Platyhelminths.

Phylogenetic analysis shows that the BPP support values for HR96b and HR96c nodes are lower than 0.8 (Fig 4). The amino acid sequences and intron position of DBD region of HR96s were further analyzed, the result shows the amino acid length is the same within the members each of group of HR96a, HR96c and HR96d, but different in the members among groups. With the exception of *M. lignano* HR96a members have no intron located in their DBD region and the intron positions of the DBD region is the same in the members of each group but is different among groups. Amino acid length of members and intron position are unusual in Platyhelminth HR96b, most which have an extra exon inserted between the P-box and D-box region that resulted in 2 introns located in this region. Most important, DBD of all parasitic HR96b members have a histidine residue replaced by a second cysteine residue in the D-Box forming a novel zinc finger motif (CHC2). Phylogenetic analysis shows HR96d is an ancient gene which is supported by sequence analysis, with only members of Platyhelminth HR96d sharing the same DBD amino acid length and intron position with that of Mollusca HR96s (Fig 5). For all sequence alignment of HR96s and intron position see S4 File. This result suggests that there were four HR96s in a common ancestor of Platyhelminths and the ancient HR96 gene (HR96d) was lost in *Schmidtea mediterranea*, Monogenea, Cestoda and Trematoda (Fig 4). HR96 genes were amplified in *M. lignano* and each of the four HR96 genes gave birth to a total of sixteen NR96 genes.

3.7.3. RXR. Two RXRs were identified in Rhabditophora, Monogenea, Cestoda and Trematoda, respectively. Phylogenetic analysis shows that parasitic flatworm RXRs are clustered in two different groups: RXR1 group (*Schistosoma* RXR1 orthologues) and RXR2 group (*Schistosoma* RXR2 orthologues). RXR2 group is clustered with RXRs of free-living flatworms, *Drosophila*, Mollusca and human, which suggests that parasitic Platyhelminth RXR2 is an orthologue of free-living flatworms, *Drosophila*, Mollusca and human RXRs. RXR1 group contains only parasitic Platyhelminth RXRs, which suggests that parasitic Platyhelminth RXR2 duplicated after the split of their common ancestor with free-living Platyhelminths. In free-living flatworms, two *M. lignano* RXRs are clustered in a group while two *S. mediterranea* RXRs are scattered in the phylogenetic tree, this result suggests that RXR duplicated independently in free-living flatworms *M. lignano* and *S. mediterranea* (S23A Fig). Even though the support value of BPP for RXR1 and RXR2 group is 0.79 and 0.75, respectively, the intron position in the DBD of RXRs further supports the phylogenetic analysis (S23B Fig). RXR gene is lost in some of parasitic Platyhelminths after ancient RXR duplicated in a common ancestor of parasitic Platyhelminths: RXR1 was lost in Monogenea *G. salaris* and Cestoda *D. latum*; RXR2 was lost in Cestoda *S. erinaceieuropaei*; both RXR1 and RXR2 were lost in all the analyzed species of Cestoda Cyclophyllidea order.

3.7.4. Coup-TF. One Coup-TF was identified in Cestoda; two are identified in Rhabditophora, Monogenea and Trematoda, respectively. Phylogenetic analysis shows that parasitic flatworms Coup-TFs are clustered in two different groups: Coup-TFI group (*Schistosoma*



Fig 5. Sequence alignment and intron position of HR96s. The greater-than sign (>) indicates the intron position in DBD region.

<https://doi.org/10.1371/journal.pone.0250750.g005>

Coup-TFI orthologues) and Coup-TFII group (*Schistosoma* Coup-TFII orthologues). Coup-TFI group contains members from all three parasitic classes and they are clustered with Coup-TFs of free-living flatworms, *Drosophila*, Mollusca and humans. This result suggests that parasitic Platyhelminth Coup-TFI is an ancient Coup-TF gene and it is the orthologue of free-living flatworms, *Drosophila*, Mollusca and human Coup-TFs. Parasitic Platyhelminth Coup-TFII group contains Monogenea and Trematoda members, which suggests that Coup-TFI duplicated and gave rise to Coup-TFII in a common ancestor of parasitic Platyhelminths. In free-living flatworms, two *M. lignano* Coup-TFs and two *S. mediterranea* Coup-TFs are scattered in the phylogenetic tree, it suggests *M. lignano* and *S. mediterranea* Coup-TFs duplicated independently (S24A Fig); analysis of the intron position in DBD of Coup-TFs further supports this result (S24B Fig). Coup-TFI was missing in Cestoda Hymenolepididae family and in Mesocostoididae (*M. corti*); Coup-TFII was missing in all analyzed species of Cestoda.

3.7.5. NR3. Four ERRs were identified in Rhabditophora *M. lignano*. Phylogenetic analysis shows that all four *M. lignano* ERRs form a single group clustered with *Drosophila*, Mollusca and human ERRs. The result demonstrates that *M. lignano* ERRs duplicated after a split of *Drosophila*, Mollusca, Platyhelminths and humans (Fig 3B). ERR is missing in Rhabditophora *S. mediterranea* and in all analyzed parasitic Platyhelminths.

3.7.6. NR4. Two NR4As were identified in Rhabditophora *M. lignano* and *S. mediterranea*, respectively; one is found in all analyzed parasitic flatworms. Phylogenetic analysis shows that M1NR4Aa, SmeNR4Aa and parasitic flatworm NR4A are clustered with *Drosophila*, Mollusca and human NR4A; it suggests that M1NR4Aa, SmeNR4Aa and parasitic flatworm NR4A are orthologues of *Drosophila*, Mollusca and human NR4A. M1NR4Ab, SmeNR4Ab are scattered in the phylogenetic tree suggesting that they underwent a duplication after a split of *M. lignano* and *S. mediterranea* (S25 Fig).

3.7.7. 2DBD-NR. Two 2DBD-NRs were identified in Rhabditophora, *M. lignano* and four were identified in *S. mediterranea*; three 2DBD-NRs were found in parasitic Platyhelminths including Monogenea, Cestoda and Trematoda, respectively. A phylogenetic tree of 2DBD-NRs was constructed with the amino acid sequence of the second DBD, because the second DBD is more conserved than the first DBD. Phylogenetic analysis shows that parasitic Platyhelminth 2DBD-NRs clustered in three groups: 2DBDa (*Schistosoma mansoni* 2DBD α orthologues), 2DBDb (*S. mansoni* 2DBD β orthologues) and 2DBDg (*S. mansoni* 2DBD γ orthologues) groups. Parasitic Platyhelminth 2DBD-NRa and 2DBD-NRb groups contain members of three classes of parasitic flatworms, but 2DBD-NRg group contains 2DBD-NRs of all the four classes of Platyhelminths and members of the Mollusca. This result suggests that

Parasitic Platyhelminth 2DBD-NRg is an ancient gene, and 2DBD-NRa and 2DBD-NRb were formed by a second round of duplication. In free living Platyhelminths, both Rhabditophora *M. lignano* 2DBD-NRs are clustered in 2DBD-NRg group, but the four *S. mediterranea* 2DBD-NRs are clustered with different parasitic helminth 2DBD-NR groups. This result suggests that *M. lignano* 2DBD-NR gene underwent duplication after a split of *S. mediterranea* and parasitic Platyhelminths. For the four *S. mediterranea* 2DBD-NRs, one is clustered in parasitic Platyhelminth 2DBD-NRg group, one is clustered in 2DBD-NRb group and two are clustered with 2DBD-NRa group. Since the two *S. mediterranea* 2DBD-NRs (2DBD-NRa1 and 2DBD-NRa2) in 2DBDa group form a polytomy, it suggests that 2DBD-NRa underwent another round of duplication and formed two 2DBD-NRs as a common ancestor of *S. mediterranea* and parasitic Platyhelminths and then one 2DBD-NRa was lost in a common ancestor of parasitic Platyhelminths (S26 Fig). The DBD sequence length is different among the members of different 2DBD-NR groups, this further supports the phylogenetic analysis. The phylogenetic analysis of 2DBD-NRs in this study is consistent with the analysis of Platyhelminths and Mollusca 2DBD-NRs [67, 68].

The above results suggests that there were at least 24 ancient NRs in a common ancestor of Platyhelminths (Fig 4). There were 8 members from subfamily 1 (TR, E78, DHR96a, DHR96b, DHR96c, DHR96d, NR1a and NR1b); 10 members from subfamily 2 (HNF4, RXR2, TR4, TLL, PNR, DSF, dax1, NHR236, Coup-TFI and Coup-TFII); 1 member from subfamily 3 (ERR); 1 member from subfamily 4 (NR4A); 2 members from subfamily 5 (FTZ-F1 and HR39); 1 member from subfamily 6 (GRF) and 1 member from 2DBD-NRs (2DBD-NRg). Phylogenetic analysis shows that NR gene gain/loss occurred in different Platyhelminth lineages (Fig 6).

A scheme represents the phylogeny of Platyhelminths. The twenty-four ancient NR genes in a common ancestor of Platyhelminths are indicated in square box on the left of the figure. NR gene gained by gene duplication is shown above the branch of different flatworm lineages and the gene lost is shown under the branches (italic and strikethrough). The number behind flatworm species/families in parentheses indicates the number of NRs identified.

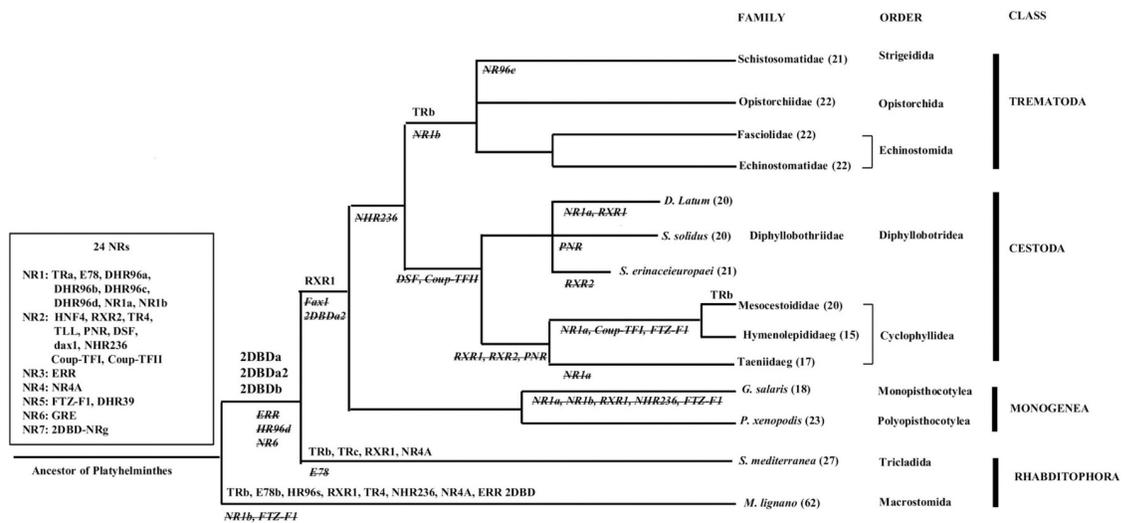


Fig 6. Summary of NR gene gain and loss in Platyhelminths.

<https://doi.org/10.1371/journal.pone.0250750.g006>

4. Discussion

In this study, we identified NRs in different species of Platyhelminths and performed sequence and phylogenetic analysis of them. Phylogenetic analysis of the DBD of NRs using DBD sequences was carried out with Bayesian inference and Maximum Likelihood method. Comparison of the two methods, Bayesian inference highly support phylogenetic analysis of NRs using only DBD sequences. (S3 File). This study shows that NRs in Platyhelminths have orthologues in Deuterostomes, arthropods or both, and the NRs in Platyhelminths diverged into two different evolutionary lineages: 1) Gene duplication and lost; 2) NR gene amplification and divergence.

NRs in Rhabditophora *S. mediterranea* and parasitic Platyhelminths follow the first evolutionary lineage: gene duplication and lost, the events occurred in different flatworm lineages. For example, fax-1 was lost in a common ancestor of *S. mediterranea* and parasitic Platyhelminths, NHR236 was lost in Cestoda and Trematoda, and DSF was lost in Cestoda. In parasitic Platyhelminths, extensive NR gene loss occurred in Cestoda Hymenolepididae and Taeniidae families, e. g. RXR, Coup-TF and FTZ-F1 were retained in most species of Platyhelminths, but they were missing in the cestode families Hymenolepididae and Taeniidae.

Comparison of NR complement in different Platyhelminth families shows that NR complement is most conserved in Trematoda. All species of Trematoda share the same NR complement except that a HR96 member was lost in Strigeidida (*Schistosoma* and *Trichobilharzia*). NR complement in Monogenea and in Cestoda exhibit more differences among different families, in Monogenea, there are 18 NRs in Gyrodactylidae (*G. salaris*), but 23 NRs are present in Polyopisthocotylea (*P. xenopodis*). In Cestoda Cyclophyllidea, there are 15 NRs in Hymenolepididae family, 17 in Taeniidae family and 20 in Mesocestoididae family. NR gene duplication occurred in different flatworm lineages. For example, 2DBD-NR underwent a second round of duplication in a common ancestor of *S. mediterranea* and parasitic Platyhelminths, RXR gene duplicated separately in *S. mediterranea* and a common ancestor of parasitic Platyhelminths, and TR duplicated independently in Rhabditophora (*S. mediterranea*), Cestoda and Trematoda.

NRs in Rhabditophora *M. lignano* follow another evolutionary lineage: Gene amplification and divergence. In *M. lignano*, 61 NRs are identified including 16 HR96s and 17 divergent members. This result suggests that NRs in *M. lignano* follows an evolutionary lineage with a feature of multiple gene duplication (amplification) and gene divergence. Multiple NR gene duplication was also found in other animals, for example, supplementary NRs (SupNRs) in nematodes and NR1H in Chordate *Cephalochordate amphioxus* [29, 69]. DHR96 (NR1J) plays a role in the response to xenobiotics. The regulation of lipid metabolism and xenobiotics is an ancestral function of members in NR1J group, HR96 links environmental conditions to physiology and development [70]. MIHR96 gene amplification in *M. lignano* may result from the adaption to its environment.

NRs in subfamily 3 and 6 were considered lost in Platyhelminths. In this study, we identified four ERRs and a NR6A in Rhabditophora *M. lignano*. This is the first known occurrence of NRs in subfamily 3 and 6 in Platyhelminths, but it is still not clear whether NRs from subfamily 3 and 6 are present in other Rhabditophora since the genome data is unavailable. This study also shows that divergent NRs are present in different flatworm lineages suggesting that novel NRs were acquired in different flatworms to adapt to the different living environments.

E75/Rev, DHR3/ROR and EcR/LXR, which are present in Deuterostomia, Arthropods and Lophotrochozoa Mollusca, are missing in Platyhelminths. In insects, E75, HR3 and EcR are directly involved in the control of the ecdysone pathway. E75 acts as a repressor of DHR3 and may through direct interaction [70].

RAR, PPAR, ER, which are missing in Arthropods but retained in Lophotrochozoa Mollusca, are also missing in Platyhelminths. In vertebrates, RARs are known to bind retinoic acid (RA). RA is a morphogen derived from vitamin A, it controls the patterning of the anteroposterior axis and the differentiation of various cell types [71]. Study of Mollusca *Nucella lapillus* RAR (NIRAR) showed that NIRAR binds to NR response elements as a heterodimer with RXR, but it does not bind all-trans retinoic acid or other retinoids [71]. In vertebrates, PPARs form a heterodimer with RXR binding their response elements, and their ligand includes free fatty acids, eicosanoids and Vitamin B3. In Mollusca *N. lapillus*, PPAR-responsive pathways is related to tributyltin (TBT) induced imposex [72]. Vertebrate ERs are activated by estradiol (E2) and have important roles in development of the nervous system and secondary sexual traits, but Mollusca ERs cannot bind estrogen, they are constitutively active and retained the ability to regulate their own gene transcription [73].

ERRs are orphan receptors in Chordata and Ecdysozoa, they play important roles in regulation of neurogenesis and metabolism in Chordata and are involved in control of larval growth in Ecdysozoa. The function of Lophotrochozoa ERRs is unknown [73]. In this study, four ERRs are identified in Platyhelminths (MIERRs), phylogenetic analysis shows that these members are closely related ERRs. The further characterization will reveal their role in development of Platyhelminths.

The NR2E6 gene was first identified in the genome of the honeybee *Apis mellifera* and then in other insects, but it was missing in *Drosophila*. The missing of NR2E6 in the major model organisms delayed the identification and functional analysis of this protein. Sometimes NR2E6 is named as PNR-like or PNR, but the true insect homolog of vertebrate PNR is HR51 (NR2E3) [70]. We demonstrated that a novel NR2E member (NHR236, orthologue of nematode *Caenorhabditis elegans* HR236) was present in Cnidaria, Arthropoda, Platyhelminths, Mollusca and Echinodermata [14]. Insect NR2E6 is an orthologue of NHR236 and all of NHR236 orthologues possess an unique P-box sequence (CDGCRG) in the DBD region.

Previously, we isolated a partial cDNA of *S. mansoni* HR96b (SmNR96b) [12], this member has a CHC2 zinc finger motif in the second zinc finger of DBD. In this study, we show that all parasitic Platyhelminth SmHR96b orthologues contain this novel motif. Whether the function of this new type of motif in NRs may change the DNA binding properties awaits further study.

5. Addendum

A paper on *Macrostomum lignano* NRs was published (Cheng, Y., et al., *Genome-Wide Characterization of the Nuclear Receptor Gene Family in Macrostomum lignano Imply Its Evolutionary Diversification*. *Frontiers in Marine Science*, 2021. 8(645)) while this paper was under review. The main differences of identification of *M. lignano* NRs between Cheng et al. and this contribution are as follows (for details see [S5 File](#)): (1). Cheng, Y., et al. identified 51 NRs in *M. lignano*, we identified 61 NRs in this species. (2). Cheng, Y., et al. identified 4 NR1Cs (PPARs) in *M. lignano*, we found no evidence for any PPARs in *M. lignano*. (3). Cheng, Y., et al. identified 2 NR1Fs (RORs). We identified the same sequences but our phylogenetical analysis showed that they belonged to NR1E (E78) group. (4). Cheng, Y., et al. identified 20 NR1Js (HR96s) in *M. lignano* including one sequence without a DBD sequence and 5 members each with an atypical P-box sequence which is different from the typical HR96 P-box sequence. We identified 16 HR96s in *M. lignano*, all of them possessed a typical HR96 P-box sequence. We also identified the same 5 members with an atypical P-box sequence that Cheng, Y., et al. identified, but put them into a divergent NR group. (5). Cheng, Y., et al. identified 10 NR3 members in *M. lignano* including four members that each possessed a typical ERR P-box sequences and 6 members that each possessed an atypical P-box sequence. We identified the same members

but put the six members that possessed an atypical P-box sequence into a divergent NR group. (6). Cheng, Y., et al. identified a NR8 member, we identified the same member but our phylogenetic analysis showed it was a divergent member. (7). Cheng, Y., et al. identified 2 NR0 members. We identified the same members but our phylogenetic analysis showed one of them belong to subfamily 6 and the other one is a divergent member because it possessed an atypical P-box sequence.

Supporting information

S1 Fig. Bayesian phylogenetic tree of *Macrostomum lignano* NRs. The Bayesian tree was constructed with the deduced amino sequences of the DNA binding domain (DBD) with a mix amino acid replacement model + invgamma rates. The PPs values are shown above each branch, branches under the PPs 0.5 are shown as polytomies. The same data set is also tested by ML method using PHYML (v2.4.4) under LG+G+I substitution model (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.158), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.956). Support values for the tree are obtained by bootstrapping a 1,000 replicates and bootstrap values above 500 and are indicated below each branch (or after MrBAYES BPPs separated by Slash). Star indicates the node obtained by Bayesian inference which is different from that obtained by ML method. Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, h: *Homo sapiens*, Lg: *Lottia gigantean*, ML: *Macrostomum lignano*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*. Red highlighted NRs show *M. lignano* NRs.
(TIF)

S2 Fig. Bayesian phylogenetic tree of *Schmidtea mediterranea* NRs. Methods for construction of phylogenetic trees see S1 Fig legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.165), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.751). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*, Sme: *Schmidtea mediterranea*. Red highlighted NRs show *S. mediterranea* NRs.
(TIF)

S3 Fig. Bayesian phylogenetic tree of *Protopolystoma xenopodis* NRs. Methods for construction of phylogenetic trees see S1 Fig legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.119), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.691). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*. Red highlighted NRs show *P. xenopodis* NRs.
(TIF)

S4 Fig. Bayesian phylogenetic tree *Gyrodactylus salaris* NRs. Methods for construction of phylogenetic trees see S1 Fig legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.119), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.691). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, Gs: *Gyrodactylus salaris*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*. Red highlighted NRs show *G. salaris* NRs.
(TIF)

S5 Fig. Bayesian phylogenetic tree of *Hymenolepis microstoma* NRs. The phylogenetic tree of *H. microstoma* NRs represents all analyzed *Hymenolepis* species because of the highly conserved DBD sequences in these species. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.092), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.709). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, h: *Homo sapiens*, Hm: *Hymenolepis microstoma*, Lg: *Lottia gigantean*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*. Red highlighted NRs show *H. microstoma* NRs.

(TIF)

S6 Fig. Bayesian phylogenetic tree of *Mesocestoides corti* NRs. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.073), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.782). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Mc: *Mesocestoides corti*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*. Red highlighted NRs show *M. corti* NRs.

(TIF)

S7 Fig. Bayesian phylogenetic tree of *Echinococcus multilocularis* NRs. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.103), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.735). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Em: *Echinococcus multilocularis*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*. Red highlighted NRs show *E. multilocularis* NRs.

(TIF)

S8 Fig. Bayesian phylogenetic tree of *Hydatigera taeniaeformis* NRs. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.093), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.717). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, h: *Homo sapiens*, Ht: *Hydatigera taeniaeformis*, Lg: *Lottia gigantean*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*. Red highlighted NRs show *H. taeniaeformis* NRs.

(TIF)

S9 Fig. Bayesian phylogenetic tree of *Taenia saginata* NRs. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.123), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.898). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*, Ts: *Taenia saginata*. Red highlighted NRs show *T. saginata* NRs.

(TIF)

S10 Fig. Bayesian phylogenetic tree of *Schitocephalus solidus* NRs. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.129), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.744). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea*

gigas, d: *Drosophila melanogaster*, h: *Homo sapiens*, Ss: *Schitoccephalus solidus*, Lg: *Lottia gigantean*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*. Red highlighted NRs show *T. saginata* NRs.

(TIF)

S11 Fig. Bayesian phylogenetic tree of *Spirometra erinaceieuropaei* NRs. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.129), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.701). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Px: *Protopolystoma xenopodis*, Se: *Spirometra erinaceieuropaei*, Sm: *Schistosoma mansoni*. DBD of E78 sequence is partial, not used for construction of the tree. Red highlighted NRs show *S. erinaceieuropaei* NRs.

(TIF)

S12 Fig. Bayesian phylogenetic tree of *Clonorchis sinensis* NRs. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.127), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.747). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, Cs: *Clonorchis sinensis*, d: *Drosophila melanogaster*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*. Red highlighted NRs show *C. sinensis* NRs.

(TIF)

S13 Fig. Bayesian phylogenetic tree of *Opisthorchis viverrini* NRs. The phylogenetic tree of *O. viverrini* NRs represents *Opisthorchis viverrini* and *O. felineus* because of the highly conserved DBD sequences in these two species. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.123), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.830). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Ov: *Opisthorchis viverrini*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*. Red highlighted NRs show *O. viverrini* NRs.

(TIF)

S14 Fig. Bayesian phylogenetic tree of *Fasciola hepatica* NRs. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.133), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.834). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, Fh: *Fasciola hepatica*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*. Red highlighted NRs show *F. hepatica* NRs.

(TIF)

S15 Fig. Bayesian phylogenetic tree of *Echinostoma caproni* NRs. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.127), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.802). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, Eca: *Echinostoma caproni*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*. EcaCoup-TFII sequence is partial and not used for construction of the tree. Red highlighted NRs show *E. caproni* NRs.

(TIF)

S16 Fig. Bayesian phylogenetic tree of *Schistosoma haematobium* NRs. The phylogenetic tree of *S haematobium* NRs represents all analyzed *Schistosoma* species because of the highly conserved DBD sequences in these species. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.119), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.714). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Px: *Protopolystoma xenopodis*, Sh: *Schistosoma haematobium*, Sm: *Schistosoma mansoni*. Red highlighted NRs show *S. haematobium* NRs.

(TIF)

S17 Fig. Bayesian phylogenetic tree of *Trichobilharzia regent* NRs. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.093), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.709). Bg: *Biomphalaria glabrata*, Cg: *Crassostrea gigas*, d: *Drosophila melanogaster*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Px: *Protopolystoma xenopodis*, Sm: *Schistosoma mansoni*, Tr: *Trichobilharzia regent*. Red highlighted NRs show *T. regent* NRs.

(TIF)

S18 Fig. Bayesian phylogenetic tree of NR1 DBD sequences. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as JTT+G (Equilibrium frequencies: Model, Proportion of invariable sites: Fixed (0.0), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.378). Bg: *Biomphalaria glabrata*, d: *Drosophila melanogaster*, Dl: *Dibothriocephalus latus*, Eca: *Echinostoma caproni*, Eg: *Echinococcus granulosus*, Fh: *Fasciola hepatica*, h: *Homo sapiens*, Hd: *Hymenolepis diminuta*, Ht: *Hydatigera taeniaeformis*, Lg: *Lottia gigantean*, Mc: *Mesocestoides corti*, Ml: *Macrostomum lignano*, Ov: *Opisthorchis viverrini*, Px: *Protopolystoma xenopodis*, Se: *Spirometra erinaceieuropaei*, Sm: *Schistosoma mansoni*, Sme: *Schmidtea mediterranea*, Ss: *Schistocephalus solidus*, Tm: *Taenia multiceps*, Tr: *Trichobilharzia regent*. Red highlighted NRs show Platyhelminths NRs.

(TIF)

S19 Fig. Bayesian phylogenetic tree of NR1 DBD with LBD. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I+F (Equilibrium frequencies: Empirical, Proportion of invariable sites: Estimated (0.029), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (1.348). Bg: *Biomphalaria glabrata*, d: *Drosophila melanogaster*, Eg: *Echinococcus granulosus*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Sme: *Schmidtea mediterranea*. Red highlighted NRs show Platyhelminths NRs.

(TIF)

S20 Fig. Bayesian phylogenetic tree of NR1 LBD. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I+F (Equilibrium frequencies: Empirical, Proportion of invariable sites: Estimated (0.004), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (1.741). Bg: *Biomphalaria glabrata*, d: *Drosophila melanogaster*, Eg: *Echinococcus granulosus*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Sme: *Schmidtea mediterranea*. Red highlighted NRs show Platyhelminths NRs.

(TIF)

S21 Fig. Bayesian phylogenetic tree of Platyhelminth divergent NRs. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as JTT+G (Equilibrium frequencies: Model, Proportion of invariable sites: Fixed (0.0), Number of substitution rate categories:

4, Gamma shape parameter: Estimated (0.895). Bg: *Biomphalaria glabrata*, d: *Drosophila melanogaster*, Dl: *Dibothriocephalus latus*, Ec: *Echinococcus Canadensis*, Eg: *Echinococcus granulosus*, Em: *Echinococcus multilocularis*, Gs: *Gyrodactylus salaris*, h: *Homo sapiens*, Hd: *Hymenolepis diminuta*, Hm: Of *H. microstoma*, Ht: *Hydatigera taeniaeformis*, Mc: *Mesocostoides corti*, Ml: *Macrostomum lignano*, Sme: *Schmidtea mediterranea*, Ta: *Taenia asiatica*, Tm: *Taenia multiceps*, Ts: *Taenia saginata*, Tr: *Trichobilharzia regent*, Tso: *Taenia solium*. (TIF)

S22 Fig. Bayesian phylogenetic tree of Platyhelminth TRs. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G (Equilibrium frequencies: Model, Proportion of invariable sites: Fixed (0.0), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.411). Bg: *Biomphalaria glabrata*, Cs: *Clonorchis sinensis*, d: *Drosophila melanogaster*, Dl: *Dibothriocephalus latus*, Ec: *Echinococcus Canadensis*, Eca: *Echinostoma caproni*, Eg: *Echinococcus granulosus*, Em: *Echinococcus multilocularis*, Fh: *Fasciola hepatica*, Gs: *Gyrodactylus salaris*, h: *Homo sapiens*, Hd: *Hymenolepis diminuta*, Hm: Of *H. microstoma*, Hn: *Hymenolepis nana*, Ht: *Hydatigera taeniaeformis*, Lg: *Lottia gigantean*, Mc: *Mesocostoides corti*, Ml: *Macrostomum lignano*, Of: *Opisthorchis felineus*, Ov: *Opisthorchis viverrini*, Px: *Protopolystoma xenopodis*, Sb: *Schistosoma bovis*, Sc: *Schistosoma curassoni*, Se: *Spirometra erinaceieuropaei*, Sh: *Schistosoma haematobium*, Sj: *Schistosoma japonicum*, Sm: *Schistosoma mansoni*, Sma: *Schistosoma margrebowiei*, Smt: *Schistosoma mattheei*, Sme: *Schmidtea mediterranea*, Sr: *Schistosoma rodhaini*, Ss: *Schitoccephalus solidus*, Ta: *Taenia asiatica*, Tm: *Taenia multiceps*, Ts: *Taenia saginata*, Tr: *Trichobilharzia regent*, Tso: *Taenia solium*. (TIF)

S23 Fig. Bayesian phylogenetic tree of Platyhelminth RXR. A) Bayesian phylogenetic tree of Platyhelminth RXR. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as LG+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.339), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (1.457). Bg: *Biomphalaria glabrata*, Cs: *Clonorchis sinensis*, d: *Drosophila melanogaster*, Dl: *Dibothriocephalus latus*, Ec: *Echinococcus Canadensis*, Eca: *Echinostoma caproni*, Eg: *Echinococcus granulosus*, Em: *Echinococcus multilocularis*, Fh: *Fasciola hepatica*, Gs: *Gyrodactylus salaris*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Ml: *Macrostomum lignano*, Of: *Opisthorchis felineus*, Ov: *Opisthorchis viverrini*, Px: *Protopolystoma xenopodis*, Sb: *Schistosoma bovis*, Sc: *Schistosoma curassoni*, Se: *Spirometra erinaceieuropaei*, Sh: *Schistosoma haematobium*, Sj: *Schistosoma japonicum*, Sm: *Schistosoma mansoni*, Sma: *Schistosoma margrebowiei*, Smt: *Schistosoma mattheei*, Sme: *Schmidtea mediterranea*, Sr: *Schistosoma rodhaini*, Ss: *Schitoccephalus solidus*, Tr: *Trichobilharzia regent*. B) Sequence alignment shows the intron position (red color >) in DBD of RXRs. (TIF)

S24 Fig. Bayesian phylogenetic tree of Platyhelminth Coup-TF. A) Bayesian phylogenetic tree of Platyhelminth Coup-TF. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as FLU+G (Equilibrium frequencies: Model, Proportion of invariable sites: Fixed (0.0), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.320). Bg: *Biomphalaria glabrata*, Cs: *Clonorchis sinensis*, d: *Drosophila melanogaster*, Dl: *Dibothriocephalus latus*, Ec: *Echinococcus Canadensis*, Eca: *Echinostoma caproni*, Eg: *Echinococcus granulosus*, Em: *Echinococcus multilocularis*, Fh: *Fasciola hepatica*, Gs: *Gyrodactylus salaris*, h: *Homo sapiens*, Lg: *Lottia gigantean*, Ml: *Macrostomum lignano*, Of: *Opisthorchis felineus*, Ov: *Opisthorchis viverrini*, Px: *Protopolystoma xenopodis*, Sb: *Schistosoma bovis*, Sc: *Schistosoma curassoni*, Se: *Spirometra erinaceieuropaei*, Sh: *Schistosoma haematobium*, Sj:

Schistosoma japonicum, Sm: *Schistosoma mansoni*, Sma: *Schistosoma margrebowiei*, Smt: *Schistosoma mattheei*, Sme: *Schmidtea mediterranea*, Sr: *Schistosoma rodhaini*, Ss: *Schitoccephalus solidus*, Ta: *Taenia asiatica*, Tm: *Taenia multiceps*, Ts: *Taenia saginata*, Tr: *Trichobilharzia regent*, Tso: *Taenia solium*. B) Sequence alignment shows the intron position (red color >) in DBD of Coup-TFs.

(TIF)

S25 Fig. Bayesian phylogenetic tree of Platyhelminth NR4A. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as FLU+G (Equilibrium frequencies: Model, Proportion of invariable sites: Fixed (0.0), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.461). Bg: *Biomphalaria glabrata*, d: *Drosophila melanogaster*, Gs: *Gyrodactylus salaris*, h: *Homo sapiens*, Hd: *Hymenolepis diminuta*, Lg: *Lottia gigantean*, Mc: *Mesocostoides corti*, Ml: *Macrostomum lignano*, Sm: *Schistosoma mansoni*, Sme: *Schmidtea mediterranea*.

(TIF)

S26 Fig. Bayesian phylogenetic tree of Platyhelminth 2DBD. Methods for construction of phylogenetic trees see [S1 Fig](#) legend. ML model tested as JTT+G+I (Equilibrium frequencies: Model, Proportion of invariable sites: Estimated (0.131), Number of substitution rate categories: 4, Gamma shape parameter: Estimated (0.832). Bg: *Biomphalaria glabrata*, Cs: *Clonorchis sinensis*, d: *Drosophila melanogaster*, Dl: *Dibothriocephalus latus*, Em: *Echinococcus multilocularis*, Fh: *Fasciola hepatica*, Gs: *Gyrodactylus salaris*, h: *Homo sapiens*, Hd: *Hymenolepis diminuta*, Hm: *H. microstoma*, Lg: *Lottia gigantean*, Mc: *Mesocostoides corti*, Ml: *Macrostomum lignano*, Px: *Protopolystoma xenopodis*, Se: *Spirometra erinaceieuropaei*, Sm: *Schistosoma mansoni*, Sme: *Schmidtea mediterranea*, Ss: *Schitoccephalus solidus*, Tr: *Trichobilharzia regent*.

(TIF)

S1 File. Lists of GenBank accession number of published NR sequences used in this study.

(DOCX)

S2 File. DBD sequences of NRs in 33 species of Platyhelminthes.

(DOCX)

S3 File. Comparison of Bayesian inference and Maximum Likelihood method to analysis of NR DBD amino sequences.

(DOCX)

S4 File. Sequence alignment of DBD sequence of HR96s and intron in this regions.

(DOCX)

S5 File. Comparison of *Macrostomum lignano* NRs identified in Cheng, Y., et al 2021 and this study.

(XLSX)

Acknowledgments

We thank the Department of Biochemistry and Structural Biology for support.

Author Contributions

Conceptualization: Wenjie Wu.

Formal analysis: Wenjie Wu.

Funding acquisition: Philip T. LoVerde.

Investigation: Wenjie Wu, Philip T. LoVerde.

Methodology: Wenjie Wu.

Validation: Wenjie Wu, Philip T. LoVerde.

Writing – original draft: Wenjie Wu.

Writing – review & editing: Philip T. LoVerde.

References

1. Riutort M., et al., Evolutionary history of the Tricladida and the Platyhelminthes: an up-to-date phylogenetic and systematic account. *Int J Dev Biol*, 2012. 56(1–3): p. 5–17. <https://doi.org/10.1387/ijdb.113441mr> PMID: 22450992
2. Adell T., et al., Platyhelminthes, in *Evolutionary Developmental Biology of Invertebrates 2: Lophotrochozoa (Spiralia)*, Wanninger A., Editor. 2015, Springer Vienna: Vienna. p. 21–40.
3. Norena C., Damborenea C., and Brusa F., *Phylum Platyhelminthes*. Ecology and General Biology, Vol I: Thorp and Covich's Freshwater Invertebrates, 4th Edition, 2015: p. 181–203.
4. Laudet V., Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor. *J Mol Endocrinol*, 1997. 19(3): p. 207–26. <https://doi.org/10.1677/jme.0.0190207> PMID: 9460643
5. Adams M.D., et al., The genome sequence of *Drosophila melanogaster*. *Science*, 2000. 287(5461): p. 2185–95. <https://doi.org/10.1126/science.287.5461.2185> PMID: 10731132
6. Holt R.A., et al., The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science*, 2002. 298(5591): p. 129–49. <https://doi.org/10.1126/science.1076181> PMID: 12364791
7. Bertrand S., et al., Evolutionary genomics of nuclear receptors: from twenty-five ancestral genes to derived endocrine systems. *Mol Biol Evol*, 2004. 21(10): p. 1923–37. <https://doi.org/10.1093/molbev/msh200> PMID: 15229292
8. Sluder A.E. and Maina C.V., Nuclear receptors in nematodes: themes and variations. *Trends Genet*, 2001. 17(4): p. 206–13. [https://doi.org/10.1016/s0168-9525\(01\)02242-9](https://doi.org/10.1016/s0168-9525(01)02242-9) PMID: 11275326
9. Dehal P., et al., The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science*, 2002. 298(5601): p. 2157–67. <https://doi.org/10.1126/science.1080049> PMID: 12481130
10. Zhang Z., et al., Genomic analysis of the nuclear receptor family: new insights into structure, regulation, and evolution from the rat genome. *Genome Res*, 2004. 14(4): p. 580–90. <https://doi.org/10.1101/gr.2160004> PMID: 15059999
11. Maglich J.M., et al., Comparison of complete nuclear receptor sets from the human, *Caenorhabditis elegans* and *Drosophila* genomes. *Genome Biol*, 2001. 2(8): p. RESEARCH0029. <https://doi.org/10.1186/gb-2001-2-8-research0029> PMID: 11532213
12. Wu W., et al., *Schistosoma mansoni* (Platyhelminthes, Trematoda) nuclear receptors: sixteen new members and a novel subfamily. *Gene*, 2006. 366(2): p. 303–15. <https://doi.org/10.1016/j.gene.2005.09.013> PMID: 16406405
13. Wu W. and LoVerde P.T., Nuclear hormone receptors in parasitic helminths. *Mol Cell Endocrinol*, 2011. 334(1–2): p. 56–66. <https://doi.org/10.1016/j.mce.2010.06.011> PMID: 20600585
14. Wu W. and LoVerde P.T., Nuclear hormone receptors in parasitic Platyhelminths. *Mol Biochem Parasitol*, 2019. 233: p. 111218. <https://doi.org/10.1016/j.molbiopara.2019.111218> PMID: 31470045
15. Forster S., et al., Molecular characterisation of a serum-responsive, DAF-12-like nuclear hormone receptor of the fox-tapeworm *Echinococcus multilocularis*. *J Cell Biochem*, 2011. 112(6): p. 1630–42. <https://doi.org/10.1002/jcb.23073> PMID: 21328613
16. Wu W., et al., Evolution of a novel subfamily of nuclear receptors with members that each contain two DNA binding domains. *BMC Evol Biol*, 2007. 7: p. 27. <https://doi.org/10.1186/1471-2148-7-27> PMID: 17319953
17. Howe K.L., et al., WormBase 2016: expanding to enable helminth genomic research. *Nucleic Acids Res*, 2016. 44(D1): p. D774–80. <https://doi.org/10.1093/nar/gkv1217> PMID: 26578572
18. Howe K.L., et al., WormBase ParaSite—a comprehensive resource for helminth genomics. *Mol Biochem Parasitol*, 2017. 215: p. 2–10. <https://doi.org/10.1016/j.molbiopara.2016.11.005> PMID: 27899279

19. Bertin B., et al., The monomeric orphan nuclear receptor *Schistosoma mansoni* Ftz-F1 dimerizes specifically and functionally with the schistosome RXR homologue, SmRXR1. *Biochem Biophys Res Commun*, 2005. 327(4): p. 1072–82. <https://doi.org/10.1016/j.bbrc.2004.12.101> PMID: 15652506
20. Bonneton F., Chaumot A., and Laudet V., Annotation of *Tribolium* nuclear receptors reveals an increase in evolutionary rate of a network controlling the ecdysone cascade. *Insect Biochem Mol Biol*, 2008. 38(4): p. 416–29. <https://doi.org/10.1016/j.ibmb.2007.10.006> PMID: 18342247
21. De Mendonca R.L., et al., A functionally conserved member of the FTZ-F1 nuclear receptor family from *Schistosoma mansoni*. *Eur J Biochem*, 2002. 269(22): p. 5700–11. <https://doi.org/10.1046/j.1432-1033.2002.03287.x> PMID: 12423370
22. de Mendonca R.L., et al., Structural and functional divergence of a nuclear receptor of the RXR family from the trematode parasite *Schistosoma mansoni*. *Eur J Biochem*, 2000. 267(11): p. 3208–19. <https://doi.org/10.1046/j.1432-1327.2000.01344.x> PMID: 10824105
23. Freebern W.J., et al., Identification of a cDNA encoding a retinoid X receptor homologue from *Schistosoma mansoni*. Evidence for a role in female-specific gene expression. *J Biol Chem*, 1999. 274(8): p. 4577–85. <https://doi.org/10.1074/jbc.274.8.4577> PMID: 9988692
24. Hu R., Niles E.G., and LoVerde P.T., DNA binding and transactivation properties of the *Schistosoma mansoni* constitutive androstane receptor homologue. *Mol Biochem Parasitol*, 2006. 150(2): p. 174–85. <https://doi.org/10.1016/j.molbiopara.2006.07.011> PMID: 16962182
25. Hu R., et al., SmTR2/4, a *Schistosoma mansoni* homologue of TR2/TR4 orphan nuclear receptor. *Int J Parasitol*, 2006. 36(10–11): p. 1113–22. <https://doi.org/10.1016/j.ijpara.2006.06.003> PMID: 16839558
26. Huang W., et al., Evolution of a novel nuclear receptor subfamily with emphasis on the member from the Pacific oyster *Crassostrea gigas*. *Gene*, 2015. 567(2): p. 164–72. <https://doi.org/10.1016/j.gene.2015.04.082> PMID: 25956376
27. Kobayashi M., et al., Identification of a photoreceptor cell-specific nuclear receptor. *Proc Natl Acad Sci U S A*, 1999. 96(9): p. 4814–9. <https://doi.org/10.1073/pnas.96.9.4814> PMID: 10220376
28. Kumar V. and Chambon P., The estrogen receptor binds tightly to its responsive element as a ligand-induced homodimer. *Cell*, 1988. 55(1): p. 145–56. [https://doi.org/10.1016/0092-8674\(88\)90017-7](https://doi.org/10.1016/0092-8674(88)90017-7) PMID: 3167974
29. Lecroisey C., Laudet V., and Schubert M., The cephalochordate amphioxus: a key to reveal the secrets of nuclear receptor evolution. *Brief Funct Genomics*, 2012. 11(2): p. 156–66. <https://doi.org/10.1093/bfgp/els008> PMID: 22441553
30. Lu C., Niles E.G., and LoVerde P.T., Characterization of the DNA-binding properties and the transactivation activity of *Schistosoma mansoni* nuclear receptor fushi tarazu-factor 1alpha (SmFTZ-F1alpha). *Mol Biochem Parasitol*, 2006. 150(1): p. 72–82. <https://doi.org/10.1016/j.molbiopara.2006.06.011> PMID: 16890303
31. Lu C., et al., Identification and characterization of a novel fushi tarazu factor 1 (FTZ-F1) nuclear receptor in *Schistosoma mansoni*. *Mol Biochem Parasitol*, 2006. 150(1): p. 25–36. <https://doi.org/10.1016/j.molbiopara.2006.06.005> PMID: 16870276
32. Mangelsdorf D.J. and Evans R.M., The RXR heterodimers and orphan receptors. *Cell*, 1995. 83(6): p. 841–50. [https://doi.org/10.1016/0092-8674\(95\)90200-7](https://doi.org/10.1016/0092-8674(95)90200-7) PMID: 8521508
33. Vanacker J.M., et al., Transcriptional activities of the orphan nuclear receptor ERR alpha (estrogen receptor-related receptor-alpha). *Mol Endocrinol*, 1999. 13(5): p. 764–73. <https://doi.org/10.1210/mend.13.5.0281> PMID: 10319326
34. Wu W., Niles E.G., and LoVerde P.T., Thyroid hormone receptor orthologues from invertebrate species with emphasis on *Schistosoma mansoni*. *BMC Evol Biol*, 2007. 7: p. 150. <https://doi.org/10.1186/1471-2148-7-150> PMID: 17727708
35. Yu R.T., et al., Relationship between *Drosophila* gap gene *tailless* and a vertebrate nuclear receptor Tlx. *Nature*, 1994. 370(6488): p. 375–9. <https://doi.org/10.1038/370375a0> PMID: 8047143
36. Larkin M.A., et al., Clustal W and Clustal X version 2.0. *Bioinformatics*, 2007. 23(21): p. 2947–8. <https://doi.org/10.1093/bioinformatics/btm404> PMID: 17846036
37. Guindon S. and Gascuel O., A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*, 2003. 52(5): p. 696–704. <https://doi.org/10.1080/10635150390235520> PMID: 14530136
38. Guindon S., et al., New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*, 2010. 59(3): p. 307–21. <https://doi.org/10.1093/sysbio/syq010> PMID: 20525638
39. Lefort V., Longueville J.E., and Gascuel O., SMS: Smart Model Selection in PhyML. *Mol Biol Evol*, 2017. 34(9): p. 2422–2424. <https://doi.org/10.1093/molbev/msx149> PMID: 28472384

40. Wasik K., et al., Genome and transcriptome of the regeneration-competent flatworm, *Macrostomum lignano*. *Proc Natl Acad Sci U S A*, 2015. 112(40): p. 12462–7. <https://doi.org/10.1073/pnas.1516718112> PMID: 26392545
41. Wudarski J., et al., Efficient transgenesis and annotated genome sequence of the regenerative flatworm model *Macrostomum lignano*. *Nat Commun*, 2017. 8(1): p. 2120. <https://doi.org/10.1038/s41467-017-02214-8> PMID: 29242515
42. Grohme M.A., et al., The genome of *Schmidtea mediterranea* and the evolution of core cellular mechanisms. *Nature*, 2018. 554(7690): p. 56–61. <https://doi.org/10.1038/nature25473> PMID: 29364871
43. Robb S.M., et al., SmedGD 2.0: The *Schmidtea mediterranea* genome database. *Genesis*, 2015. 53(8): p. 535–46. <https://doi.org/10.1002/dvg.22872> PMID: 26138588
44. Robb S.M., Ross E., and Sanchez Alvarado A., SmedGD: the *Schmidtea mediterranea* genome database. *Nucleic Acids Res*, 2008. 36(Database issue): p. D599–606. <https://doi.org/10.1093/nar/gkm684> PMID: 17881371
45. Hahn C., Fromm B., and Bachmann L., Comparative genomics of flatworms (platyhelminthes) reveals shared genomic features of ecto- and endoparasitic neodermata. *Genome Biol Evol*, 2014. 6(5): p. 1105–17. <https://doi.org/10.1093/gbe/evu078> PMID: 24732282
46. International Helminth Genomes, C., Comparative genomics of the major parasitic worms. *Nat Genet*, 2019. 51(1): p. 163–174. <https://doi.org/10.1038/s41588-018-0262-1> PMID: 30397333
47. Tsai I.J., et al., The genomes of four tapeworm species reveal adaptations to parasitism. *Nature*, 2013. 496(7443): p. 57–63. <https://doi.org/10.1038/nature12031> PMID: 23485966
48. Maldonado L.L., et al., The *Echinococcus canadensis* (G7) genome: a key knowledge of parasitic platyhelminth human diseases. *BMC Genomics*, 2017. 18(1): p. 204. <https://doi.org/10.1186/s12864-017-3574-0> PMID: 28241794
49. Zheng H., et al., The genome of the hydatid tapeworm *Echinococcus granulosus*. *Nat Genet*, 2013. 45(10): p. 1168–75. <https://doi.org/10.1038/ng.2757> PMID: 24013640
50. Wang S., et al., Comparative genomics reveals adaptive evolution of Asian tapeworm in switching to a new intermediate host. *Nat Commun*, 2016. 7: p. 12845. <https://doi.org/10.1038/ncomms12845> PMID: 27653464
51. Li W., et al., The genome of tapeworm *Taenia multiceps* sheds light on understanding parasitic mechanism and control of coenurosis disease. *DNA Res*, 2018. 25(5): p. 499–510. <https://doi.org/10.1093/dnares/dsy020> PMID: 29947776
52. Bennett H.M., et al., The genome of the sparganosis tapeworm *Spirometra erinaceieuropaei* isolated from the biopsy of a migrating brain lesion. *Genome Biol*, 2014. 15(11): p. 510. <https://doi.org/10.1186/PREACCEPT-2413673241432389> PMID: 25413302
53. Wang X., et al., The draft genome of the carcinogenic human liver fluke *Clonorchis sinensis*. *Genome Biol*, 2011. 12(10): p. R107. <https://doi.org/10.1186/gb-2011-12-10-r107> PMID: 22023798
54. Ershov N.I., et al., New insights from *Opisthorchis felineus* genome: update on genomics of the epidemiologically important liver flukes. *BMC Genomics*, 2019. 20(1): p. 399. <https://doi.org/10.1186/s12864-019-5752-8> PMID: 31117933
55. Young N.D., et al., The *Opisthorchis viverrini* genome provides insights into life in the bile duct. *Nat Commun*, 2014. 5: p. 4378. <https://doi.org/10.1038/ncomms5378> PMID: 25007141
56. Cwiklinski K., et al., The *Fasciola hepatica* genome: gene duplication and polymorphism reveals adaptation to the host environment and the capacity for rapid evolution. *Genome Biol*, 2015. 16: p. 71. <https://doi.org/10.1186/s13059-015-0632-2> PMID: 25887684
57. McNulty S.N., et al., Genomes of *Fasciola hepatica* from the Americas Reveal Colonization with *Neorickettsia* Endobacteria Related to the Agents of Potomac Horse and Human Sennetsu Fevers. *PLoS Genet*, 2017. 13(1): p. e1006537. <https://doi.org/10.1371/journal.pgen.1006537> PMID: 28060841
58. Oey H., et al., Whole-genome sequence of the bovine blood fluke *Schistosoma bovis* supports interspecific hybridization with *S. haematobium*. *PLoS Pathog*, 2019. 15(1): p. e1007513. <https://doi.org/10.1371/journal.ppat.1007513> PMID: 30673782
59. Young N.D., et al., Whole-genome sequence of *Schistosoma haematobium*. *Nat Genet*, 2012. 44(2): p. 221–5. <https://doi.org/10.1038/ng.1065> PMID: 22246508
60. *Schistosoma japonicum* Genome, S. and C. Functional Analysis, The *Schistosoma japonicum* genome reveals features of host-parasite interplay. *Nature*, 2009. 460(7253): p. 345–51. <https://doi.org/10.1038/nature08140> PMID: 19606140
61. Berriman M., et al., The genome of the blood fluke *Schistosoma mansoni*. *Nature*, 2009. 460(7253): p. 352–8. <https://doi.org/10.1038/nature08160> PMID: 19606141

62. Luisi B.F., et al., Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature*, 1991. 352(6335): p. 497–505. <https://doi.org/10.1038/352497a0> PMID: 1865905
63. Lee M.S., et al., Structure of the retinoid X receptor alpha DNA binding domain: a helix required for homodimeric DNA binding. *Science*, 1993. 260(5111): p. 1117–21. <https://doi.org/10.1126/science.8388124> PMID: 8388124
64. Wang J., Chen R., and Collins J.J. 3rd, Systematically improved in vitro culture conditions reveal new insights into the reproductive biology of the human parasite *Schistosoma mansoni*. *PLoS Biol*, 2019. 17(5): p. e3000254. <https://doi.org/10.1371/journal.pbio.3000254> PMID: 31067225
65. Wang J. and Collins J.J. 3rd, Identification of new markers for the *Schistosoma mansoni* vitelline lineage. *Int J Parasitol*, 2016. 46(7): p. 405–10. <https://doi.org/10.1016/j.ijpara.2016.03.004> PMID: 27056273
66. Nuclear Receptors Nomenclature Committee, A unified nomenclature system for the nuclear receptor superfamily. *Cell*, 1999. 97(2): p. 161–3. [https://doi.org/10.1016/s0092-8674\(00\)80726-6](https://doi.org/10.1016/s0092-8674(00)80726-6) PMID: 10219237
67. Alvite G., et al., Bioinformatic analysis of a novel *Echinococcus granulosus* nuclear receptor with two DNA binding domains. *PLoS One*, 2019. 14(11): p. e0224703. <https://doi.org/10.1371/journal.pone.0224703> PMID: 31710619
68. Ladurner P., et al., A new model organism among the lower Bilateria and the use of digital microscopy in taxonomy of meiobenthic Platyhelminthes: *Macrostomum lignano*, n. sp (Rhabditophora, Macrostromorpha). *Journal of Zoological Systematics and Evolutionary Research*, 2005. 43(2): p. 114–126.
69. Gobert G.N., et al., Tissue specific profiling of females of *Schistosoma japonicum* by integrated laser microdissection microscopy and microarray analysis. *PLoS Negl Trop Dis*, 2009. 3(6): p. e469. <https://doi.org/10.1371/journal.pntd.0000469> PMID: 19564906
70. Bonneton F. and Laudet V., Evolution of Nuclear Receptors in Insects, in *Insect endocrinology* 2012. p. 219–252.
71. Gutierrez-Mazariegos J., et al., A mollusk retinoic acid receptor (RAR) ortholog sheds light on the evolution of ligand binding. *Endocrinology*, 2014. 155(11): p. 4275–86. <https://doi.org/10.1210/en.2014-1181> PMID: 25116705
72. Pascoal S., et al., Transcriptomics and in vivo tests reveal novel mechanisms underlying endocrine disruption in an ecological sentinel, *Nucella lapillus*. *Mol Ecol*, 2013. 22(6): p. 1589–608. <https://doi.org/10.1111/mec.12137> PMID: 23205577
73. Miglioli A., et al., Nuclear Receptors and Development of Marine Invertebrates. *Genes (Basel)*, 2021. 12(1). <https://doi.org/10.3390/genes12010083> PMID: 33440651