General and Comparative Endocrinology 174 (2011) 60–69

Contents lists available at SciVerse ScienceDirect

General and Comparative Endocrinology

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

Cloning and expression pattern of the ecdysone receptor and retinoid X receptor from the centipede *Lithobius peregrinus* (Chilopoda, Lithobiomorpha)

Francesca Bortolin [a,*], Maria-Dolors Piulachs [b], Leonardo Congiu [a], Giuseppe Fusco [a]

[a] Department of Biology, University of Padova, via U. Bassi 58/B, I-35131 Padova, Italy
[b] Institut de Biología Evolutiva (CSIC-UPF), Passeig Marítim 37, 08003 Barcelona, Spain

ARTICLE INFO

Article history:
Received 17 May 2011
Revised 8 August 2011
Accepted 10 August 2011
Available online 19 August 2011

Keywords:
EcR
RXR
Molting
Myriapoda
Gene expression

ABSTRACT

In arthropods, molting events are mediated by the binding of the ecdysone hormone to a heterodimer of two nuclear receptors: the ecdysone receptor (EcR) and the retinoid X receptor (RXR), a homolog of ultraspiracle (USP). We have cloned partial sequences of several isoforms for *EcR* and *RXR* genes from the centipede *Lithobius peregrinus*, and studied their expression profile during the second post-embryonic stage. *LpEcR* and *LpRXR* inferred amino acid sequences are very similar to other arthropod orthologs, especially to those of chelicerates and hemimetabolous insects, and their expression levels are significantly higher during the 48 h that precede the molt. Results obtained in this study represent the first data on the genetic basis of the ecdysone signal pathway for a myriapod, and in particular for an animal that, through a stereotyped developmental schedule paced by the molt cycle, completes trunk segmentation during post-embryonic life.

1. Introduction

Arthropod ecdysteroids (zooecdysteroids) are steroid hormones responsible for regulating processes associated with developmental events (such as metamorphosis), reproduction and diapause. A prominent example is the α-ecdysone released by the prothoracic glands of insects, whose titer oscillations, or those of its biologically active form, the 20-hydroxyecdysone (20E, hereafter referred to as ecdysone), drives molting processes [27].

The rise in concentration of ecdysone in specific developmental phases initiates changes in tissue-specific gene expression through a hierarchy of ecdysone-responsive genes. In *Drosophila*, these events are mediated by the binding of the hormone to a heterodimer of nuclear receptors consisting of the ecdysone receptor (EcR) and ultraspiracle (USP), a retinoid X receptor (RXR) homolog [39]. Insect EcR is a distant relative of the vertebrate farnesoid X receptor (FXR) or liver receptor (LXR). It has been identified in several insects, crustaceans and chelicerates, but outside the arthropods, EcR orthologs have been reported only in some parasitic nematodes, as *Dirofilaria immitis* [30], *Brugia malayi* [36] and *Haemonchus contortus* [14]. The *USP* gene was originally identified in *Drosophila*, and presently the use of this name tends to be restricted to homologs from highly derived holometabolous insects clades, while the name *RXR* is more frequently used for those of other arthropod groups [16,27]. Ultraspiracle protein is an orphan receptor, i.e. a receptor that operates without a ligand-binding activity. However, the heterodimerization of EcR with USP is necessary to increase the binding affinity of ecdysteroids to EcR and for transcriptional activity [39]. RXR orthologs have been reported from other arthropods, as well as for other metazoans, including the nematode *Brugia malayi* [36] and the cubozoan jellyfish *Tripedalia cystophora* [23].

EcR and RXR belong to the superfamily of nuclear receptor (NR) proteins and share NR-typical domain structures and gene regulatory mechanisms. NR proteins are characterized by five distinct domains [3,10]: (i) domain A/B, a highly variable N-terminal domain involved in transcriptional activation; (ii) domain C, a highly conserved DNA-binding domain (DBD); (iii) domain D, a flexible and variable hinge region involved in the recognition and heterodimerization of ecdysone response elements; (iv) domain E, a rather complex ligand-binding domain (LBD) that is involved in hormone binding, heterodimerization and interaction with other transcription factors; and finally (v) domain F, a C-terminal highly variable domain, found only in EcR hortologues, whose function is not well understood. In both proteins, the amino acid sequences of the C (DNA-binding) and E (ligand-binding) domains are highly conserved, allowing the cloning of EcR and USP/RXR orthologs in several insects (e.g., *Locusta migratoria* [16,17]), crustaceans (e.g., *Marsupenaeus japonicus* [1]) and chelicerates (e.g., *Liocheles australasiae* [25]).

In insects, the 20E-receptor complex directly activates a small group of so-called 'early genes', among which *Broad-Complex*

* Corresponding author. Fax: +39 049 8276230.
E-mail addresses: francesca.bortolin@unipd.it (F. Bortolin), mdolors.piulachs@ibe.upf-csic.es (M.-D. Piulachs), leonardo.congiu@unipd.it (L. Congiu), giuseppe.fusco@unipd.it (G. Fusco).

0016-6480/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygcen.2011.08.008

(*Br-C*), *E74* and *E75*, each encoding a set of distinct transcription factor isoforms. The main role of these genes directly regulated by the hormone is to coordinate the temporal activation (in cascade) of appropriate sets of 'late genes'. In *Drosophila,* these late genes encode tissue-specific effector proteins, necessary for the developmental events that drive metamorphosis at the end of the last larval stage [2,22,27,35].

In myriapods, the endocrine system has been investigated in a few species through extirpation/reimplantation experiments, immuno- and radio-assays and histological and ultrastructural analysis (e.g., [21]), but biochemical and molecular approaches have never been pursued. Thus, neither the specific hormones, nor the genes involved in molting processes have been identified. In the stone centipede *Lithobius forficatus*, lymphatic strands surrounding the salivary glands are known to function as molting glands. This hypothesis is supported by both the ultrastructural similarity of this organ with the prothoracic glands of insects, and the biochemical affinity of its secretions to the ecdysteroids of other arthropods [29]. Injections of exogenous ecdysone increase the number of molts in adult specimens [20], but hormone titer has never carried out during juvenile (anamorphic, see below) stages. A fragment of a RXR ortholog was isolated in *L. forficatus* for a molecular phylogeny study in arthropods [4], but the authors did not provide any information about sequence characterization.

Here we take a molecular approach, consisting in the sequencing of candidate gene transcripts, in association with quantitative measures of their expression during development. This well-validated study design, together with other recent molecular- or genomic-based protocols, allows extending endocrinology studies to non-model species, where traditional phenotype-based approaches cannot be easily performed, thus significantly contributing to widening the comparative context of endocrine trait analyses [28].

We have cloned partial sequences of *EcR* and *RXR* gene homologs from the stone centipede *Lithobius peregrinus*, a close relative of *L. forficatus*, and analyzed their expression profiles during the second post-embryonic stage. The first five post-embryonic stages are called *anamorphic* stages (sometimes also referred to as 'larval' stages, and consequently labeled L0-L4), because during this developmental phase the number of trunk segments increases progressively at each molt through a stereotyped scheme of segment addition, until the adult trunk composition, with fifteen leg-bearing segments, is obtained (*hemianamorphic* development, see [12]).

2. Materials and methods

2.1. Centipede husbandry

Eggs were obtained from adult *L. peregrinus* collected in San Stino di Livenza (NE Italy) and reared in the laboratory (see [5]). Following hatching, juveniles were bred separately until the selected stage in Petri dishes with hardened poured plaster of Paris floor to maintain an adequate level of humidity at 21 ± 1 °C under natural photoperiod. They were checked daily for molting. The first two postembryonic stages (L0 and L1) do not feed, whereas the subsequent stages were fed with live fruit flies.

Despite controlled environmental parameters, stage duration varies considerably among individuals. The first post-embryonic stage (L0) lasts about one day, the second stage (L1) lasts about three days, whereas the following anamorphic stages (L2 to L4) last on average 20–25 days each.

To investigate the expression pattern of genes that encode for the heterodimer EcR–RXR, we focused on the second post-embryonic stage (L1), because its duration (3.4 days, on average) is less variable (s.d. = 0.9 days) than other stages. From September 2009 to March 2010, 39 juveniles were collected at four different points in time during the L1 stage to be analyzed: 11 specimens immediately after the molt L0-L1 (group L1_0 h), 9 after 24 h (L1_24 h), 8 after 48 h (L1_48 h) and 11 after 72 h (L1_72 h).

2.2. RNA extraction and cDNA synthesis

Individuals were killed by freezing in N_2 and stored at −80 °C. Selected pools of individuals were then transferred to a ceramic mortar and ground to powder in liquid nitrogen.

Total RNA was isolated using the SV Total RNA Isolation kit (Promega, Madison, USA), according to the manufacturer's protocol, including a Dnase treatment. The concentrations and purity of RNA were determined by NanoDrop ND-1000 spectrophotometer (*NanoDrop* Technologies, Rockland, USA). For all RNA samples, A260/A280 and A260/A230 ratios were in the range 2.0–2.1 and 1.9–2.0 respectively.

Random primers and ImProm-II Reverse Transcriptase System (Promega) were used to perform first-strand cDNA. Synthesis started from 1 μg of total RNA, following manufacturer's instructions (see [5]). To minimize sample treatment disparity, all RNA samples were reverse-transcribed simultaneously.

2.3. Primer design

For both EcR and RXR, several degenerate primers were designed on multiple alignment of the DNA-binding domain (DBD) and Ligand-binding domain (LBD) of different arthropod homologs, to obtain the correspondent cDNA fragments from *L. peregrinus*. All primers were synthesized by MWG Biotech (Ebersberg, Germany).

In detail, to isolate the ecdysone receptor we aligned the EcR sequences of *Apis mellifera* (GenBank accession number AB267886), *Tribolium castaneum* (NM_001114178), *Bombyx mori* (NM_001043866), *Locusta migratoria* (AF049136), *Blattella germanica* (AM039690), *Daphnia magna* (AB274820), *Gecarcinus lateralis* (AY642975), *Ornithodoros moubata* (AB191193) and *Liocheles australasiae* (AB297929).

For *L. peregrinus* RXR isolation we used homologs in *Apis mellifera* (NM_001011634), *Tribolium castaneum* (NM_001114294), *Aedes aegypti* (AF305213), *Locusta migratoria* (AF136372), *Blattella germanica* (AJ854490), *Daphnia magna* (AB274819), *Celuca pugilator* (AF032983), *Ornithodoros moubata* (AB353290) and *Liocheles australasiae* (AB297930). The sequences of all primers used are available on request.

2.4. Sequence isolation

To isolate sequences of interest, 1 μl of cDNA (50 ng/μl) from a pool of juveniles of different stages was amplified using a standard PCR protocol.

The amplified cDNA fragments were cloned into the plasmid vector pGEM®-T Easy Vector (Promega), transforming *Escherichia coli* JM109 competent cells, and sequenced (BMR Genomics). In order to avoid cloning or PCR artifacts, several clones were screened for each fragment obtained.

2.5. Semi-quantitative PCR

PCR assays were performed to determine *RXR* and *EcR* temporal expression during L1 stage. RNA was isolated from each of the four L1 groups (L1_0 h, L1_24 h, L1_48 h, L1_72 h). cDNA synthesis was performed using 0.5 μg total RNA and ImProm-II Reverse Transcriptase (Promega).

Elongation factor 1-α (LpEF1-α; GenBank accession number FR714838) expression was used as an internal PCR control, by

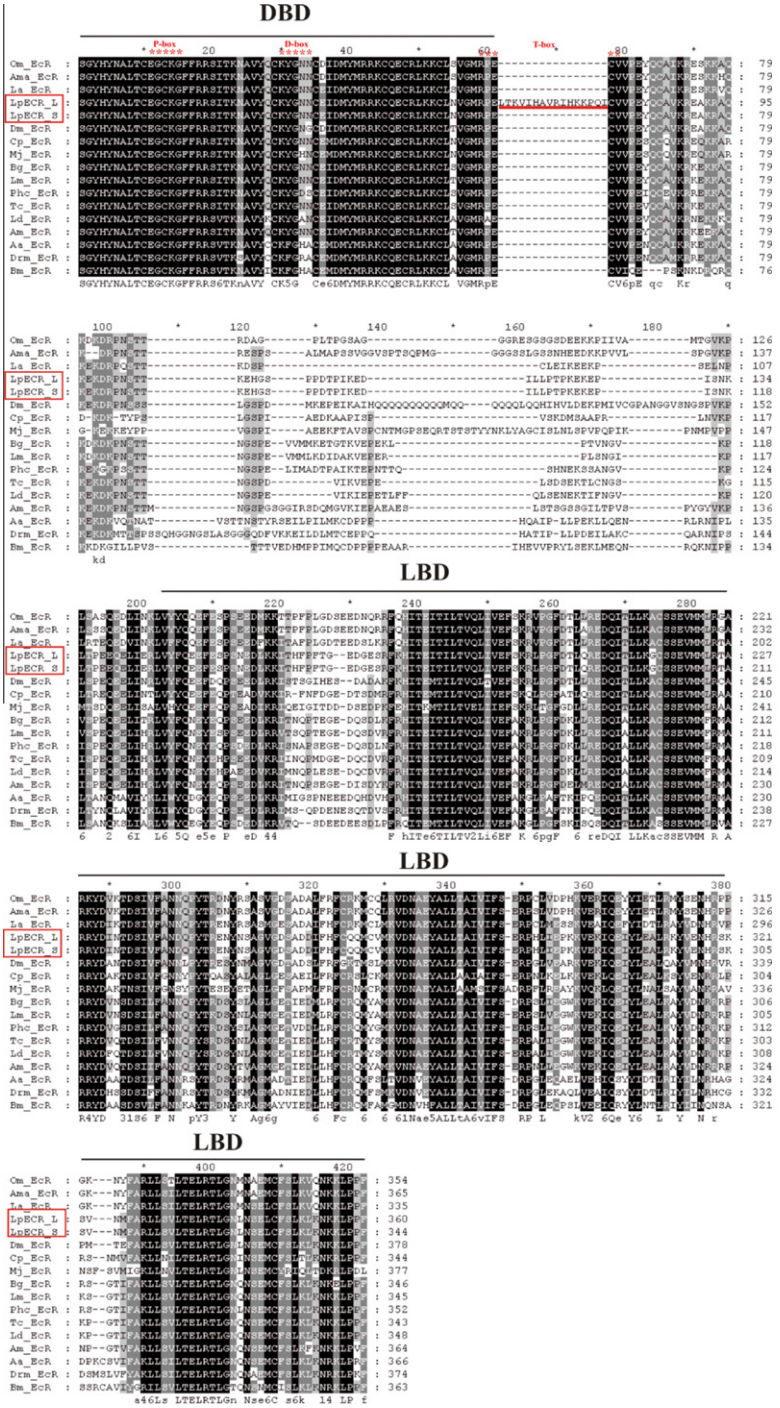

Fig. 1. Alignment of amino acid sequences of LpEcR and EcRs orthologs from species representative of the main groups of Chelicerata, Crustacea and Insecta: *Ornithodoros moubata* (Om; BAE45855), *Amblyomma americanum* (Ama; AAB94566), *Liocheles australasiae* (La; BAF85822), *Daphnia magna* (Dam; BAF49029), *Celuca pugilator* (Cp; AAC33432), *Marsupenaeus japonicus* (Mj; BAF75375), *Blattella germanica* (Bg; CAJ01677), *Locusta migratoria* (Lm; AAD19828), *Pediculus humanus corporis* (Phc; XP_002430228), *Tribolium castaneum* (Tc; NP_001107650), *Leptinotarsa decemlineata* (Ld; BAD99296), *Apis mellifera* (Am; BAF46356), *Aedes aegypti* (Aa; AAA87394), *Drosophila melanogaster* (Drm; Genbank Accession Number NP_724456), *Bombyx mori* (Bm; 001037331). Orthologous sequences from arthropod species were aligned with fragments from *L. peregrinus* (red boxed) using ClustalW program. Amino acids are shaded according to the degree of conservation using GeneDoc: black (similarity 100%); grey (similarity 80–90%); light grey (similarity 60–70%). Regions corresponding to DBD and LBD are marked with an upper bar, and the insertion in LpEcR_L is underlined. Amino acids of P-box, D-box and T-box are marked with stars.

amplification of a 502 bp fragment. *LpEF1-α* had already proved to be a suitable reference gene in *L. peregrinus* [5].

Preliminarily, amplification tests were performed in order to define the optimum cDNA quantity required to produce *LpEF1-α* band with similar intensity in each group. The amounts so set (17.5 ng cDNA for groups L1_0 h and L1_24 h, 12.5 ng cDNA for the group L1_36 h and 6.25 ng cDNA for group L1_72 h) were used to amplify both target and control gene.

PCR cycling was performed by standard protocol, with different annealing temperature for the three genes. PCR cycle number was optimized performing parallel amplifications (n = 22, 26, 30, 32, 34 and 36). After analyzing expression results from different cycles, 35

A

	C	D	E	LpEcR L	LpEcR S
L. australasiae	98	57	78	72	76
O. moubata	96	46	74	68	71
A. americanum	96	39	74	65	68
M. japonicus	96	48	61	54	57
C. pugilator	98	43	71	66	69
D. magna	94	56	76	65	68
B. germanica	98	55	72	69	73
L. migratoria	98	57	73	70	73
L.decemlineata	93	50	71	67	71
T. castaneum	98	57	72	70	73
P. humanus	94	48	73	67	71
A. mellifera	98	47	72	66	70
B. mori	88	28	57	52	55
A. aegypti	88	42	60	56	59
D. melanogaster	86	42	60	53	56

B

	C	D	E	LpRXR L	LpRXR M	LpRXR S
L. forficatus	98	95	90	95	96	92
L. australasiae	91	70	74	75	74	77
O. moubata	94	86	67	73	75	74
A. americanum	92	81	72	71	72	75
M. japonicus	94	52	70	69	69	72
C. pugilator	92	63	63	68	70	68
D. magna	94	72	74	73	75	78
B. germanica	97	81	76	76	77	81
L. migratoria	95	81	76	74	75	79
L.decemlineata	92	72	67	69	70	73
T. castaneum	94	72	66	69	70	72
P. humanus	91	76	74	72	74	77
A. mellifera	92	77	72	72	73	77
B. mori	92	45	43	52	53	53
A. aegypti	94	64	44	54	56	40
D. melanogaster	94	45	42	47	48	46

Fig. 2. Comparison of deduced amino acid sequences of EcR and RXR receptor homologs with LpEcR and LpRXR isoforms. The GeneBank accession numbers of the sequences are given in the legends for Figs. 1 and 3. (A) Percentage identities of amino acid sequences of EcR orthologs versus LpEcR. Identities values for C–E region were calculated only against LpEcR_S. (B) Percentage identities of amino acid sequences of USP/RXR orthologs versus LpRXR. Identities values for C–E region were calculated only against LpRXR_S. For *Amblyomma americanum*, *Blattella germanica* and *Locusta migratoria* we used only the short isoform of RXR (AmaRXR_S, BgRXR_S and LmRXR_S, respectively).

PCR cycles were selected for receptor analysis, and 22 PCR cycles were used for *LpEF1-α* expression analysis.

Five µl of each sample were added to 3 µl of loading buffer and they were run on 1.8% agarose gel in TAE 0.5X. One kilobyte ladder was used as molecular weight marker and bands were stained with GelRed™ (Biotium, Hayward, USA) and visualized on Gel Doc (Bio-Rad, Hercules, USA). The relative intensities of the amplified PCR products were determined using the program NIH ImageJ (http://rsb.info.nih.gov/ij) and expressed in arbitrary units (AU).

The intensities of the cDNA bands obtained for *LpRXR* and *LpEcR* in the different groups were normalized dividing the intensity of each band by the corresponding *LpEF1-α* specific PCR product density.

2.6. Sequence comparison and phylogenetic analysis

Sequence similarity search was performed using the program Blast (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Sequence alignment and identity calculation were carried out with the program ClustalX 1.83 (http://www.clustal.org [34]) and edited in GeneDoc software version 2.7.000 (www.psc.edu/biomed/genedoc).

In the alignments of the obtained sequences of LpEcR and LpRXR with the known arthropod EcR and RXR/USP homologs, poorly aligned positions and divergent regions were eliminated by using Gblocks 0.91 b [6]. The resulting alignments were analyzed by the program PhyML 3.0 [15] based on the maximum-likelihood principle with the amino acid substitution model. Four substitution rate categories with a gamma shape parameter of 1.444 were used. The data were bootstrapped for 100 replicates using PhyML.

Similarity between the phylogenies of the two receptors was measured on the basis of Pazos and Valencia's [26] method. This is based on the calculation of the linear correlation coefficient (Perason's correlation coefficient) between the matrices of pairwise distances among the amino acid sequences of each protein for the species ($N = 21$) shared by the two protein datasets. This measure is independent from the method of phylogeny reconstruction. Distances between amino acid sequences of domains A/B-F of EcR and RXR were estimated as number of site differences, using the software MEGA ver. 4 [33]. Missing data and gap sites in the multiple alignments were discarded from all species. When multiple sequences were available for the same species, we arbitrarily chose the longest, but alternative choices gave almost identical results.

2.7. Statistical analysis

All experiments were done at least in triplicates. Significance of the differences in means was calculated using ANOVA, and $p < 0.05$ were considered to be statistically significant. The program StatGraphics Centurion XV was used for all statistical analysis.

3. Results

3.1. Characterization of L. peregrinus EcR

We isolated two cDNA sequences, 1081 and 1033 bp long, encoding two polypeptides of 360 and 344 amino acids respectively. A database search with the program Blast indicated that both amino acid sequences encode domains C–E of *L. peregrinus* orthologs of EcR protein. They have been called LpEcR_L (long form, GenBank accession number FR846491) and LpEcR_S (short form, GenBank FR846492). The two proteins differ in an insertion/deletion (hereafter, insertion) of 16 amino acids in domain D. Deduced amino acid sequence has a structure typical of the nuclear receptor superfamily: a two-zinc-fingered DNA-binding domain, DBD (domain C, 59 aa), a hinge region (domain D, 73 aa for LpEcR_S, 89 aa for LpEcR_L), and a ligand-binding domain, LBD (domain E, 212 aa). Our cDNA sequences do not include either the ligand-independent activation domain A/B, nor the poorly conserved carboxyterminal domain F (Fig. 1).

LpEcR_L and LpEcR_S amino acid sequences were compared with EcR sequences from other arthropods (Figs. 1 and 2A). Domain C of LpEcR exhibits very high amino acid identity with those of other species' EcR (88–98%). In the DBD region, the P-box sequence of LpEcR (EGCKG) is 100% identical to that of other EcRs, whereas the D-box sequence of LpEcR (KYGNN) is much less conserved among arthropods.

Apart from the insertion present in the long isoform, domain D is quite similar to those of the chelicerate *L. australasiae*, the crustacean *D. magna*, and the insects *L. migratoria* and *B. germanica* (56–57%), while similarity is lower with respect to other arthropods (<48%). However, quite surprisingly, the insertion of LpEcr_L is located within the T-box (Fig.1), a 5-aa motif that is extremely conserved across arthropods, and that is thought to complement the DNA binding function of the flanking domain C.

Domain E of LpEcR is also highly similar to those of other EcRs (>60%), especially to those from Chelicerata (74–78%).

3.2. Characterization of L. peregrinus RXR

We isolated three cDNA sequences, 979, 961, and 915 bp long, encoding three polypeptides of 326, 320, and 305 amino acids respectively. A database search with the program Blast showed that these deduced sequences are highly homologous to domains C–E of other RXR/USP proteins, so they have been called LpRXR_L (long form, GenBank accession number FR846493), LpRXR_M (intermediate form, GenBank accession number FR846494), LpRXR_S (short form, GenBank accession number FR846495).

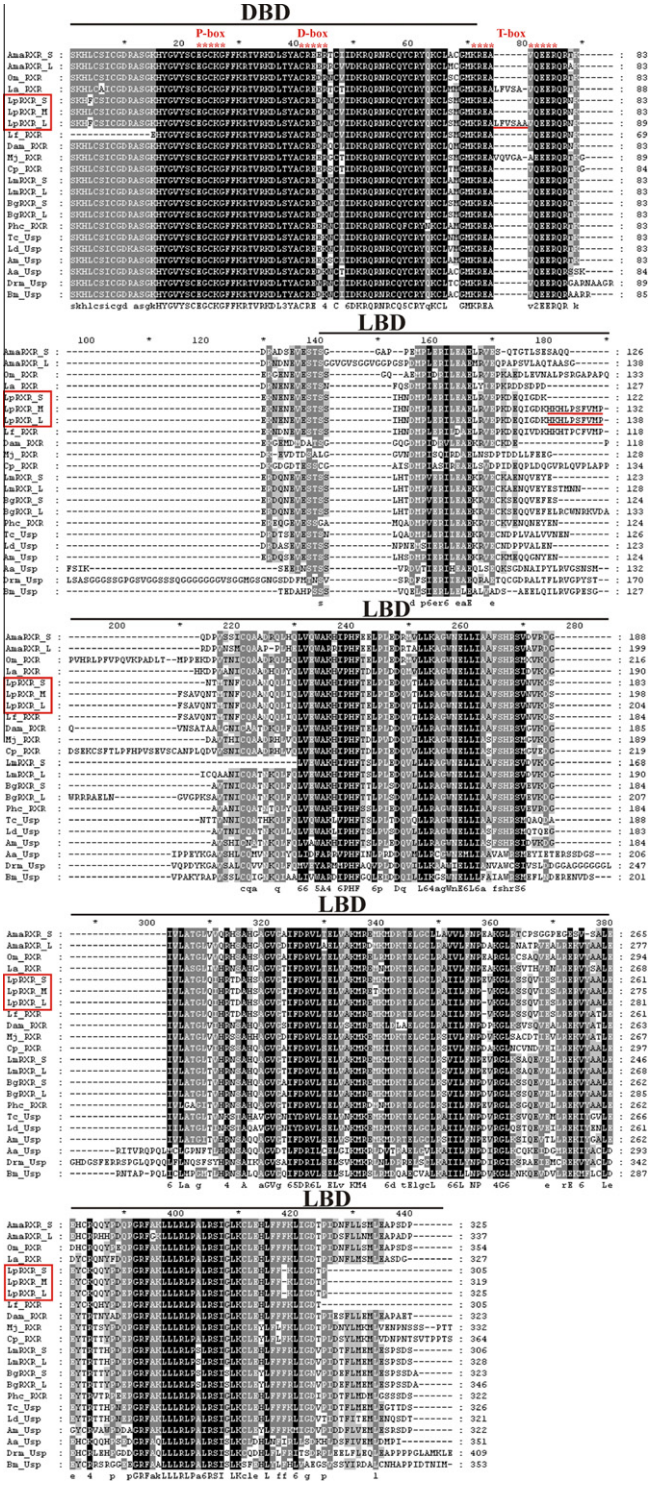

Fig. 3. Alignment of amino acid sequences of LpRXR and Usps/RXRs orthologs from species representative of the main groups of Chelicerata, Crustacea and Insecta, plus the only sequence available to date for Myriapoda: *Amblyomma americanum* (AmaRXR_L AAC15589; AmaRXR_S AAC15588), *Ornithodoros moubata* (Om; BAF91724), *Liocheles australasiae* (La; BAF85823), *Lithobius forficatus* (Lf; AAO18151), *Daphnia magna* (Dam; ABF74729), *Marsupenaeus japonicus* (Mj; BAF75376), *Celuca pugilator* (Cp; AAC32789), *Locusta migratoria* (LmRXR_L AAQ55293; LmRXR_S AAF00981), *Blattella germanica* (BgRXR_L CAH69898; BgRXR_S CAH69897), *Pediculus humanus corporis* (Phc; XP_002424949), *Tribolium castaneum* (Tc; NP_001107650), *Leptinotarsa decemlineata* (Ld; BAD99298), *Apis mellifera* (Am; NP_001011634), *Aedes aegypti* (Aa; AAG24886), *Drosophila melanogaster* (Drm; NP_476781), *Bombyx mori* (Bm; NP_001037470). Orthologous sequences from arthropod species were aligned with fragments from *L. peregrinus* (red boxed) using ClustalW program. Amino acids are shaded according to the degree of conservation using GeneDoc: black (similarity 100%); grey (similarity 80–90%); light grey (similarity 60–70%). Regions corresponding to DBD and LBD are marked with an upper bar, and short and long insertions in LpRXR isoforms are underlined. Amino acids of P-box, D-box and T-box are marked with stars.

These three proteins are identical except for two insertions/deletions (hereafter, insertions): a sequence of 6 amino acids in domain D (that discriminates LpRXR_L isoform from the other two) and a sequence of 15 amino acids in domain E (that discriminates LpRXR_S isoform from the other two). Amino acid sequence comparisons indicate that the proteins have a domain organization typical of a nuclear hormone receptor. Specifically, they include a DNA-binding domain, DBD (C domain, 71 aa), a hinge region (D domain, 22 aa for LpRXR_S and LpRXR_M, 28 aa for LpRXR_L) and a ligand-binding domain LBD (E domain, 211 aa for LpRXR_S, 226 aa for LpRXR_M and LpRXR_L). Our cDNA sequences do not include either the ligand-independent activation domain A/B, nor the poorly conserved carboxyterminal domain F (Fig. 3).

LpRXR_L, LpRXR_M and LpRXR_S amino acid sequences were compared with RXR/USP sequences from other arthropods, (Figs. 3, 2B). LpRXR_M sequence is very similar to *L. forficatus* ortholog (LfRXR) and both of them have the same insertion in domain E.

Similarly to LpEcR, amino acid identity of domain C of LpRXR to all other homologous sequences is very high (91–97%). The P-box of LpRXR (EGCKG) is 100% identical to that of other species RXR/USP, whereas the D-box (CREDR) is less conserved across arthropods.

Apart from for the insertion present in LpRXR_L, domain D is highly similar to those of the ticks *O. moubata* and *Amblyomma americanum*, and the insects *B. germanica* and *L. migratoria* (81–86%). The 6-aa insertion of LpRXR_L is located within the T-box, and shares several amino acids with corresponding sequences in domain D of the scorpion *L. australasiae* and the prawn *M. japonicus* (Fig. 4A).

Domain E of LpRXR_S is also highly similar to those of other arthropods (>63%), but for Lepidoptera and Diptera (42–44%). The loop between helices H1 and H3, where the 15-aa insertion of LpRXR_L and LpRXR_M is located, is quite divergent in arthropods. However, the insertion in *L. peregrinus* interestingly presents some similarity with the loop sequences of chelicerates and crustaceans (Fig. 4B).

3.3. Phylogenetic analyses

The phylogenetic trees of the two proteins (Fig. 5) have very similar topology, as measured by subtree prune and regraft distance ($d_{SPR} = 3$, on 21 shared terminal taxa, [18]). In particular, the non-insect subtrees of the two proteins have the same topology but for the position of the water flea *D. magna*. In both EcR and RXR phylogenies, *Lithobius* isoforms form an exclusive cluster that branches in between crustaceans and chelicerates.

Tree similarity suggests that EcR and RXR have evolved in parallel, likely as a result of their functional interaction, as the two proteins are effective as an ecdysone receptor only when they

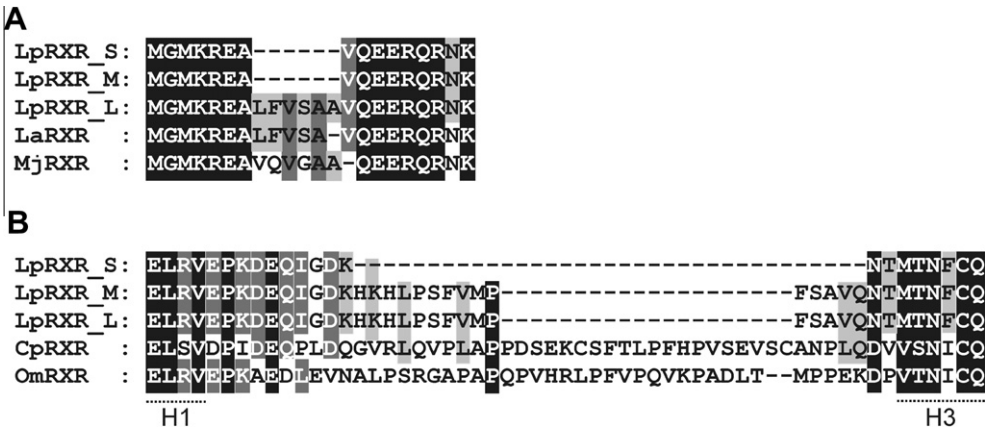

Fig. 4. Comparison of domain D and domain E sequences of LpRXR_S, LpRXR_M and LpRXR_L isoforms of *Lithobius peregrinus* with other species. (A) A portion of domain D with the 6-aa insertion present in LpRXR_L is aligned with the homolog region of *Marsupenaeus japonicus* (MjRXR, Crustacea) and *Liocheles australasiae* (LaRXR, Chelicerata). (B) In domain E, the region from helices H1 and H3 is aligned with the homologous region of *Ornithodoros moubata* (OmRXR, Chelicerata) and *Celuca pugilator* (CpRXR, Crustacea). Regions corresponding to helices H1 and H3 are underlined with dashed line. Amino acids are shaded according to the degree of conservation using GeneDoc: black (similarity 100%); grey (similarity 80–90%); light grey (similarity 60–70%).

are combined to form an heterodimer. To test further the interaction between the two receptors, we applied the Pazos and Valencia [26] approach to measure the level of coordinate evolution between the two proteins. For this analysis only the species shared by the two trees were considered. The correlation between the two matrices of pairwise distances among the sequences included in the analysis is quite high ($r = 0.75$, $N = 210$), and approximates the empirical threshold of 0.80, indicated by Pazos and Valencia [26] as a mark for true protein interaction. Correlation between EcR and RXR is remarkable in consideration of the fact that the 0.80 threshold has been determined on a large dataset that includes the interactions of distinct domains of the same protein. Thus, as pointed out by the same authors, correlation values for effectively interacting proteins that are independently coded and synthesized are expected to be smaller.

3.4. Receptor expression profiles during the second post-embryonic stage (L1)

Because of the little sequence differences, it was not possible to discriminate between the different isoforms of *LpEcR* and *LpRXR* in an agarose gel. Therefore, the intensity of bands that we analyzed for the two cDNAs represents the summation of the expression levels in all isoforms of the same gene (Fig. 6).

The effects of experiment replicates and groups (L1_0 h–L1_72 h) on *LpEcr* and *LpRXR* relative expression levels were tested with a two-way ANOVA. For both genes, expression levels in the four L1 groups differ significantly ($p < 0.0005$), whereas differences between the three replicates of the experiment are not significant ($p > 0.51$).

Concentration of *LpEcR* mRNA is low for the first 24 h of stage L1, to significantly increase about 48 h after the molt, with a significant peak of expression at 72 h (Fischer's test LSD, $\alpha = 0.01$; Fig. 7).

LpRXR shows an expression profile quite similar to that of *LpEcR*, with a significant increase in expression at 48 h, that however, at variance with *LpEcR*, does not change significantly in the next 24 h (Fisher's test LSD, $\alpha = 0.01$; Fig. 7).

4. Discussion

Sequence comparison of the receptors LpEcR and LpRXR isolated in *L. peregrinus* with their orthologs in other arthropods shows the highest degree of identity with chelicerates and hemimetabolous insects (68–81%), and a lesser degree with crustaceans and some holometabolous insects (65–70%). However, receptors from the more derived insect clades (e.g., Diptera and Lepidoptera) exhibit low sequence similarity with the centipede orthologs (40–55%).

Domain C (DBD) of both receptors is highly conserved across arthropods, and *L. peregrinus* is not an exception. In this region there are two zinc-finger domains containing respectively a proximal P-box sequence (5 aa) and a distal D-box sequence (5 aa) that provide DNA-binding specificity [37]. Amino acids of the P-box, critical for DNA response element recognition, are identical in all arthropods, suggesting that they recognize similar response elements. D-box region forms a dimerization interface in several nuclear receptors [37], and substitutions in this sequence can reflect functional differences in protein–protein interaction among different receptors. For instance, there is evidence that *Drosophila* EcR may bind nuclear receptor other than USP [38].

The ligand-binding domain of EcR (domain E) is highly conserved across arthropods, in agreement with the widespread occurrence of the 20E as molting hormone in this group. The high similarity of LpEcR to the orthologs of all other arthropods supports the hypothesis that the ecdysone has a role as molting hormone in *Lithobius* too.

A distinctive characteristic of *L. peregrinus* EcR and RXR is the presence of several variants: we identified two isoforms for EcR and three isoforms for RXR.

LpEcR isoforms, namely LpEcR_S and LpEcR_L, only differ by a 16-aa segment in domain D. Two portions of this domain, namely the T-box and A-box motifs, are highly conserved across arthropods and play a key role in DNA recognition [8]. Localization of LpEcR_L insertion within the T-box suggests the possibility of a different response element recognition between the two isoforms. Domain D is also essential for a ligand-dependent heterodimerization with RXR, but further studies are required to ascertain whether the multiple variants of LpEcR have different properties in DNA binding and heterodimerization with LpRXR. Multiple variants in domain D have been described in a few crustacean species, although their possible distinct function is unknown. For instance, *C. pugilator* exhibits four substitutive variants [7], while *M. japonicus* has two, which interestingly exhibit different level of expression [1]. In insect EcR, alternative sequences are observed in domain A/B. These variants are produced by alternative splicing, and their expression is regulated by distinct promoters. In *Drosophila*, the different isoforms are expressed in tissue-specific and developmental stage-specific manner [32]. A complete full length

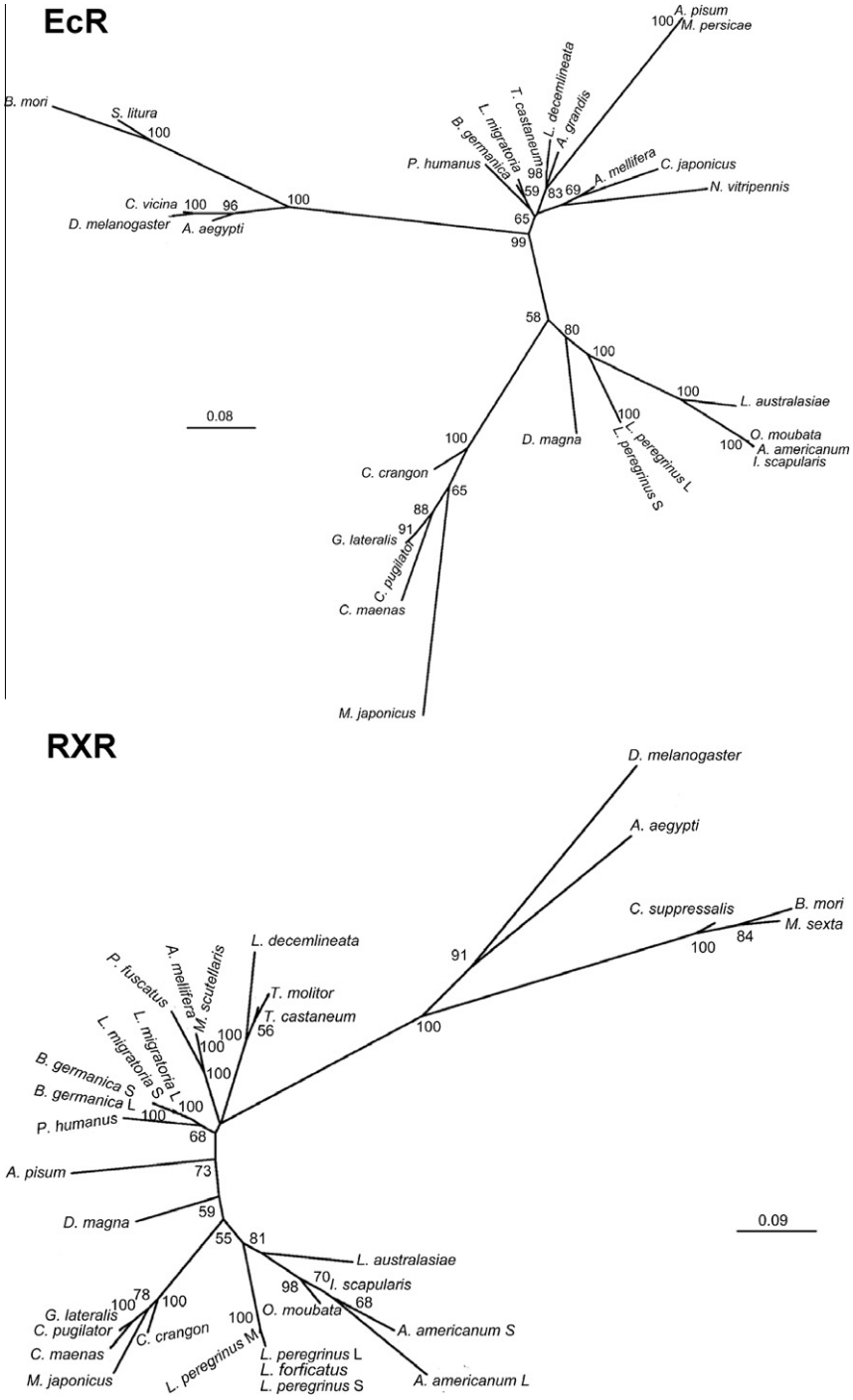

Fig. 5. Maximum likelihood phylogenetic trees of arthropod EcR and RXR. Branch lengths are proportional to sequence divergence per site (see bar). The GeneBank accession numbers of sequences are: *Lithobius forficatus* (LfRXR AAO18151); *Marsupenaeus japonicus* (MjEcR AB295492; MjRXR BAF75376); *Carcinus maenas* (CmEcR AAR89628; CmRXR ACG63787); *Celuca pugilator* (CpEcR AF034086; CpRXR AAC32789); *Gecarcinus lateralis* (GlEcR AAT77808; GlRXR AAZ20368); *Crangon crangon* (CcEcR ACO44665; CcRXR ACO44669); *Daphnia magna* (DamEcR AB274820; DamRXR ABF74729); *Liocheles australasiae* (LaEcR BAF85822; LaRXR BAF85823); *Ornithodoros moubata* (OmEcR AB191193; OmRXR BAF91724); *Amblyomma americanum* (AaEcR AF020187; AaRXR_S AAC15588; AaRXR_L AAC15589); *Ixodes scapularis* (IsEcR XP_002405625; IsRXR XP_002435070); *Acyrthosiphon pisum* (ApEcR NP_001152832; ApRXR ACR45970); *Myzus persicae* (MpEcR ABN11289); *Nasonia vitripennis* (NvEcR NP_001152829); *Camponotus japonicus* (CjEcR BAF79665); *Apis mellifera* (AmEcR AB267886; AmUSP NP_001011634); *Anthonomus grandis* (AgEcR ACK57879); *Leptinotarsa decemlineata* (LdEcR AB211191; LdUSP BAD99298); *Tribolium castaneum* (TcEcR NM_001114178; TcUSP NP_001107766); *Locusta migratoria* (LmEcR AF049136; LmRXR_S AAF00981; LmRXR_L AAQ55293); *Blattella germanica* (BgEcR AM039690; BgRXR_S CAH69897; BgRXR_L CAH69898); *Pediculus humanus corporis* (PhcEcR XM_002430183; PhcRXR XP_002424949); *Spodoptera litura* (SlEcR ABX79143); *Bombyx mori* (BmEcR NM_001043866; BmUSP NP_001037470); *Calliphora vicina* (CvEcR AF325360); *Drosophila melanogaster* (DmEcR NM_165461; DmUSP NP_476781); *Aedes aegypti* (AaEcR AAU02021; AaUSP AF305213); *Polistes fuscatus* (PfUSP AAX37292); *Melipona scutellaris* (MsUSP AAW02952); *Tenebrio molitor* (TmUSP CAB75361); *Manduca sexta* (MasUSP P54779); *Chilo suppressalis* (CsUSP BAC53670).

cDNA sequence for LpEcR would allow to identify domain A/B and possibly new isoforms of this protein.

Unlike *L. forficatus* RXR, which shows only one isoform, there are two deletion variant sites in the cDNAs of LpRXR: one is located in the domain D and the other one is in domain E. As in LpEcR, the 6-aa insertion in domain D of LpRXR_L is located inside the T-box region. Interestingly, this sequence is almost identical to that found in *L. australasiae* and quite similar to that in *M. japonicus*. This re-

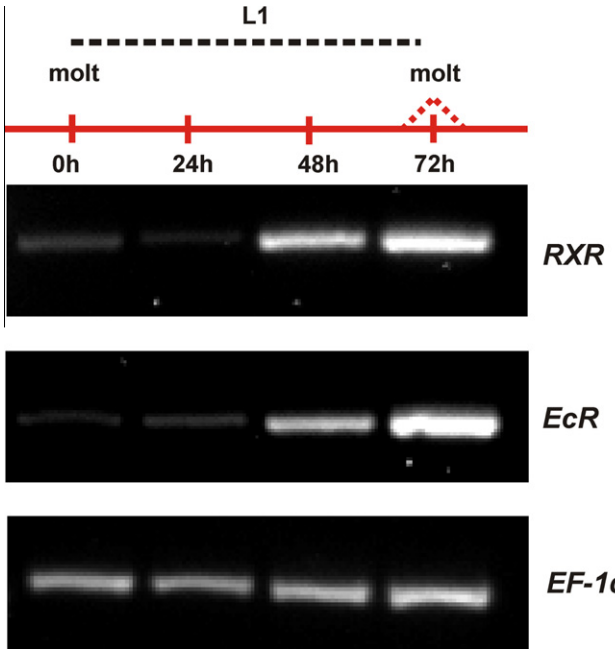

Fig. 6. Expression patterns of *LpRXR* and *LpEcR* mRNA during the second post-embryonic stage (L1) of *Lithobius peregrinus*. cDNA fragments of *LpEcR* and *LpRXR* were amplified from four groups of specimens collected immediately after the molt L0-L1 (group L1_0 h), after 24 h (L1_24 h), after 48 h (L1_48 h) and after 72 h (L1_72 h). Fragments were separated by electrophoresis in a 1% agarose gel, and visualized by staining with GelRed. *LpEF1-α* levels were used as a reference.

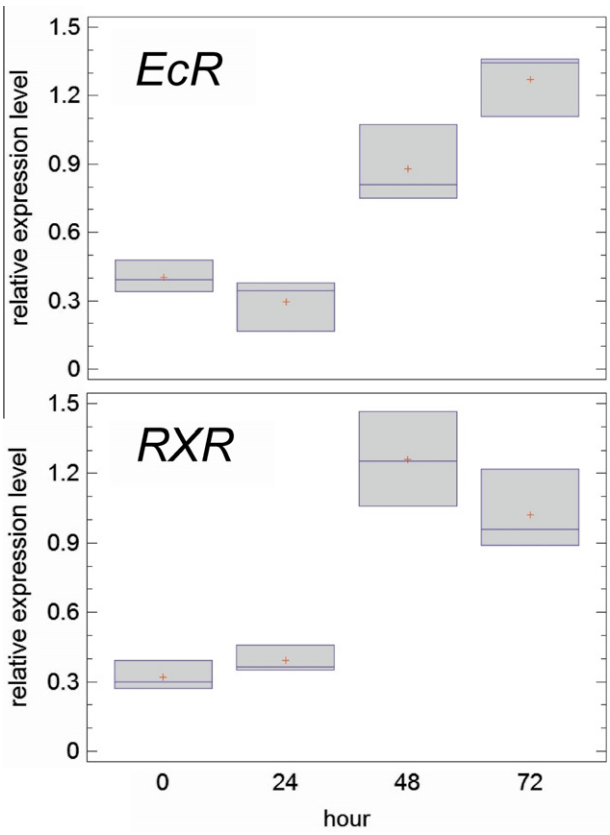

Fig. 7. Relative expression level of *LpEcR* and *LpRXR* during the second post-embryonic stage of *Lithobius peregrinus* (on average, 3.5 days). Data are normalized with respect to the expression level of *LpEF1-α*. Boxes represent the interval between lower and upper quartiles, with median (transverse line) and mean (small cross). For both genes, expression levels at 0 h and 24 h differ significantly from those at 48 h and 72 h, and in *LpEcR* there is also a significant difference between 48 h and 72 h expression levels.

gion has a key role in mediating interactions between the hormone response elements of downstream genes and either RXR homodimers, or RXR heterodimers formed by RXR binding to other nuclear receptors [41]. However, nothing is known about its function under these different variants.

The 15-aa insertion in LpRXR_L and LpRXR_M is located in the loop connecting helices H1 and H3 within the domain E, as in *L. forficatus* RXR. In arthropods, RXR/USP isoforms differing for insertions/deletions of this type have been reported for *L. migratoria* [16,17], *B. germanica* [24] and *C. pugilator* [9]. Sequence variation in this domain could influence transactivation properties or ligand affinities, but further studies are necessary to clarify the function of the different isoforms of *L. peregrinus*.

The phylogenetic trees of the two proteins have very similar topology, although there is a conspicuous difference in the branch length of the malacostracan cluster, indicating a more rapid rate of divergence of EcR sequences with respect to those of RXR in this clade.

The lack of suitable outgroup sequences for either EcR or RXR, precluded the possibility to build reliably radicated phylogenetic trees. The shared topology of the two receptor trees is equally compatible with the two currently competing major hypotheses for arthropod phylogeny: one version of the so called 'Mandibulata hypothesis' (Chelicerata + (Myriapoda + (Crustacea + Hexapoda)) [13]), and the so called 'Myriochelata hypothesis' ((Chelicerata + Myriapoda) + (Crustacea + Hexapoda) [11]), depending on the position of the root. However, it is not compatible with any hypotheses that encompass an Atelocerata clade (Myriapoda + Hexapoda).

We have provided evidence that EcR and RXR have evolved in parallel during the course of arthropod phylogeny, because of their functional interaction. This result confirms previous observations conducted on some arthropod taxa, as for instance the holometabolous insects [4].

The expression patterns shown by *LpEcR* and *LpRXR* during the anamorphic stage L1 of *L. peregrinus* are similar. Expression levels of both receptors are low during the first day after ecdysis, to increase in the second part of the inter-molt period, suggesting that molting process starts about 48 h after the previous molt. This observation is in agreement with expression profiles of the two receptors in other arthropods, as for instance in the crustaceans *C. pugilator* [7] and *M. japonicus* [1]. In these species, thoracic muscles show high level of EcR and RXR approaching the molt, although both receptor genes can exhibit dissimilar expression profiles in other tissues.

In insects, the expression profile of most of the ecdysteroid-regulated genes directly correlates with the ecdysteroid titer [31]. However, the expression profiles of *EcR* and *USP*/*RXR* do not coincide with all the peaks of ecdysteroid level, with these discrepancies depending on the developmental stage and the expressing tissue. This implies that the expression of these genes is not exclusively controlled by ecdysteroid.

Elucidation of the signal pathways associated to molting in *Lithobius* is of particular interest also from developmental and evolutionary perspectives, as these animals present anamorphic development, a very little known mode of segmentation [5]. In *Lithobius*, trunk segmental composition is completed during the first five stages of post-embryonic life, through a precise schedule of segment addition in the posterior of the trunk at each molt. Despite conspicuous variation in growth rates and temporal progression of the molts, there is no individual variation in the segmentation schedule, so that each anamorphic stage is characterized by a precise segmental composition of the body, that is almost invariant even at the level of the whole clade Lithobiomorpha

[12]. Growth, molting and segmentation are thus highly coordinated developmental processes in this clade, although the precise mechanism of interaction between these co-occurring processes is still to be understood. Clarifying this signaling system would have bearing on centipede development, but, more than that, this would be important for arthropod evolution and evolvability, as segmentation through anamorphosis is possibly the primitive condition within this major clade [19].

5. Conclusions

Results obtained in this study represent the first data on the genes involved in the ecdysone signal pathway in a myriapod species. Sequence similarity of the two *L. peregrinus* receptors to the orthologs in other arthropods, along with similarity of the expression profiles during inter-molt period, support the hypothesis that the molting process in *Lithobius* is regulated by the same hormonal signal cascade observed in all arthropods so far investigated. However, further studies will be necessary to identify LpEcR and LpRXR target genes. This will contribute to a better understanding of the functional role of the two receptors, as well as to clarify the whole hormonal signaling cascade involved in *Lithobius* molting processes.

The experimental approach taken here and the new molecular techniques in general focus on the important, but previously unapproachable, endocrine issue of the evolutionary change at the level of *localized* receptor expression, that adds to variation in *systemic* (titers) endocrine regulation. This will be a key topic for future research, as evolutionary change and adaptation in endocrinology traits are expected to occur at any level of the hormone regulation system [40]. For this reason, the study of receptor expression needs to be integrated with classical approaches, like measurement of hormone titers and experimental manipulations, to assess the functional significance of endocrine variation.

Viewed in the context of 'evolutionary endocrinology', an emerging field of studies that aims at integrating traditional comparative endocrine approaches with molecular studies of hormone receptors and intracellular signaling pathways [40], this study contributes new data to a little explored aspect of arthropod developmental evolution. Indeed, because of the close interactions between molting and segmentation during post-embryonic development in these animals, elucidation of structure and organization of their molting signaling system has implications for the study of development and body plan evolution of arthropods as a whole.

Acknowledgments

We thank C. Marcolin and R. Carrer for precious help with animal field collection, N. Franchi for assistance in analyzing electrophoresis gel images, E. Negrisolo for insightful suggestions about phylogenetic analysis, and S. Baldi and A. Minelli for their comments on a draft of this article. This research has been partly supported by a CaRiPaRo Foundation grant.

References

[1] H. Asazuma, S. Nagata, M. Kono, H. Nagasawa, Molecular cloning and expression analysis of ecdysone receptor and retinoid X receptor from the kuruma prawn, *Marsupenaeus japonicus*, Comp. Biochem. Physiol. B Biochem. Mol. Biol. 148 (2007) 139–150.
[2] M. Ashburner, Puffs, genes, and hormones revisited, Cell 61 (1990) 1–3.
[3] I.M.I. Billas, C. Browning, M.C. Lawrence, L.D. Graham, D. Moras, R.J. Hill, The structure and function of ecdysone receptors, in: G. Smagghe (Ed.), Ecdysone: structures and functions, Springer, Dordrecht, 2009, pp. 335–360.
[4] F. Bonneton, D. Zelus, T. Iwema, M. Robinson-Rechavi, V. Laudet, Rapid divergence of the ecdysone receptor in Diptera and Lepidoptera suggests coevolution between ECR and USP-RXR, Mol. Biol. Evol. 20 (2003) 541–553.
[5] F. Bortolin, C. Benna, G. Fusco, Gene expression during post-embryonic segmentation in the centipede *Lithobius peregrinus* (Chilopoda, Lithobiomorpha), Dev. Genes Evol. (2011), doi:10.1007/s00427-011-0359-3.
[6] J. Castresana, Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis, Mol. Biol. Evol. 17 (2000) 540–552.
[7] A.C. Chung, D.S. Durica, S.W. Clifton, B.A. Roe, P.M. Hopkins, Cloning of crustacean EcR and RXR gene homologs and elevation of RXR mRNA by retinoic acid, Mol. Cell. Endocrinol. 139 (1998) 209–227.
[8] S. Devarakonda, J.M. Harp, Y. Kim, A. Ozyhar, F. Rastinejad, Structure of the heterodimeric ecdysone receptor DNA-binding complex, EMBO J. 22 (2003) 5827–5840.
[9] D.S. Durica, X. Wu, G. Anilkumar, P.M. Hopkins, A.C. Chung, Characterization of crab EcR and RXR homologs and expression during limb regeneration and oocyte maturation, Mol. Cell. Endocrinol. 189 (2002) 59–76.
[10] R.M. Evans, The steroid and thyroid hormone receptor superfamily, Science 240 (1988) 889–895.
[11] M. Friedrich, D. Tautz, Ribosomal DNA phylogeny of the major extant arthropod classes and the evolution of myriapods, Nature 376 (1995) 165–167.
[12] G. Fusco, Trunk segment numbers and sequential segmentation in myriapods, Evol. Dev. 7 (2005) 608–617.
[13] G. Giribet, C. Ribera, The position of arthropods in the animal kingdom: a search for a reliable outgroup for internal arthropod phylogeny, Mol. Phyl. Evol. 9 (1998) 481–488.
[14] L.D. Graham, A.C. Kotze, R.T. Fernley, R.J. Hill, An ortholog of the ecdysone receptor protein (EcR) from the parasitic nematode *Haemonchus contortus*, Mol. Biochem. Parasit. 171 (2010) 104–107.
[15] S. Guindon, O. Gascuel, A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood, Syst. Biol. 52 (2003) 696–704.
[16] D.C. Hayward, M.J. Bastiani, J.W.H. Trueman, J.W. Truman, L.M. Riddiford, E.E. Ball, The sequence of locusta RXR homologous to *Drosophila* Ultraspiracle, and its evolutionary implications, Dev. Genes Evol. 209 (1999) 564–571.
[17] D.C. Hayward, T.S. Dhadialla, S. Zhou, M.J. Kuiper, E.E. Ball, G.R. Wyatt, V.K. Walker, Ligand specificity and developmental expression of RXR and ecdysone receptor in the migratory locust, J. Insect Physiol. 49 (2003) 1135–1144.
[18] G. Hickey, F. Dehne, A. Rau-Chaplin, C. Blouin, SPR distance Computation for *Unrooted* trees, Evol. Bioinf. 4 (2008) 17–27.
[19] N.C. Hughes, A. Minelli, G. Fusco, The ontogeny of trilobite segmentation: a comparative approach, Paleobiology 32 (2006) 602–627.
[20] R. Joly, Action de l'ecdysone sur le cycle de la mue de *Lithobius forficatus*. L. (Myriapode, Chilopode), Compt. Rend. Seances Soc. Biol. Et de ses Filiales 158 (1964) 548–550.
[21] R. Joly, M. Descamps, Endocrinology of Myriapods, in: H. Laufer (Ed.), Endocrinology of selected invertebrate types, Alan R. Liss. lnc, New York, 1988, pp. 430–449.
[22] F.D. Karim, G.M. Guild, C.S. Thummel, The *Drosophila Broad-Complex* plays a key role in controlling ecdysone-regulated gene expression at the onset of metamorphosis, Development 118 (1993) 977–988.
[23] Z. Kostrouch, M. Kostrouchova, W. Love, E. Jannini, J. Piatigorsky, J.E. Rall, Retinoic acid X receptorin the diploblast, *Tripedalia cystophora*, PNAS 95 (1998) 13442–13447.
[24] O. Maestro, J. Cruz, N. Pascual, D. Martin, X. Bellés, Differential expression of two RXR/ultraspiracle isoforms during the life cycle of the hemimetabolous insect *Blattella germanica* (Dictyoptera, Blatellidae), Mol. Cell. Endocrinol. 238 (2005) 27–37.
[25] Y. Nakagawa, A. Sakai, F. Magata, T. Ogura, M. Miyashita, H. Miyagawa, Molecular cloning of the ecdysone receptor and the retinoid X receptor from the scorpion *Liocheles australasiae*, FEBS J. 274 (2007) 6191–6203.
[26] F. Pazos, A. Valencia, Similarity of phylogenetic trees as indicator of protein-protein interaction, Protein Eng. 14 (2001) 609–614.
[27] L.M. Riddiford, P.T. Cherbas, J.W. Truman, Ecdysone receptors and their biological actions, Vitam. Horm. 60 (2001) 1–73.
[28] M. Robinson-Rechavi, Perspective: from molecular endocrinology of mouse to genomic endocrinology of animals, Mol. Cell. Endocrinol. 293 (2008) 2–4.
[29] G. Seifert, H.J. Bidmon, Immunohistochemical evidence for ecdysteroid-like material in the putative molting glands of Lithobius forficatus Chilopoda, Cell Tissue Res. 253 (1988) 263–266.
[30] C. Shea, J. Richer, G. Tzertzinis, C.V. Maina, An EcR homolog from the filarial parasite *Dirofilaria immitis* requires a ligand-activated partner for transactivation, Mol. Biochem. Parasitol. 171 (2010) 55–63.
[31] A.A. Sullivan, C.S. Thummel, Temporal profile of nuclear receptor gene expression reveal coordinate transcriptional responses during *Drosophila* development, Mol. Endocrinol. 17 (2003) 2125–2137.
[32] W.S. Talbot, E.A. Swyryd, D.S. Hogness, Drosophila tissues with different metamorphic responses to ecdysone express different ecdysone receptor isoforms, Cell 73 (1993) 1323–1337.
[33] K. Tamura, J. Dudley, M. Nei, S. Kumar, MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0, Mol. Biol. Evol. 24 (2007) 1596–1599.
[34] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The CLUSTAL_X Windows Interface. flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucl. Acids Res. 24 (1997) 4876–4882.
[35] C.S. Thummel, From embryogenesis to metamorphosis: the regulation and function of *Drosophila* nuclear receptor super family members, Cell 83 (1995) 871–877.

[36] G. Tzertzinis, A.L. Egaña, S.R. Palli, M. Robinson-Rechavi, C.R. Gissendanner, C. Liu, T.R. Unnasch, C.V. Maina, Molecular evidence for a functional ecdysone signaling system in Brugia malayi, PLoS Negl. Trop. Dis. 4 (3) (2010) e625.
[37] K. Umesono, R.M. Evans, Determinants of target gene specificity for steroid / thyroid hormone receptors, Cell 57 (1989) 1139–1146.
[38] K.P. White, P. Hurban, T. Watanabe, D.S. Hogness, Coordination of *Drosophila* metamorphosis by two ecdysone-induced nuclear receptors, Science 276 (1997) 114–117.
[39] T.P. Yao, B.M. Forman, Z. Jiang, L. Cherbas, J.D. Chen, M. McKeown, P. Cherbas, R.M. Evans, Functional ecdysone receptor is the product of *EcR* and *Ultraspiracle* genes, Nature 366 (1993) 476–479.
[40] A.J. Zera, L.G. Harshman, T.D. Williams, Evolutionary endocrinology: the developing synthesis between endocrinology and evolutionary genetics, Annu. Rev. Ecol. Evol. Syst. 38 (2007) 793–817.
[41] Q. Zhao, S.A. Chasse, S. Devarakonda, M.L. Sierk, B. Ahvazi, F. Rastinejad, Structural basis of RXR–DNA interactions, J. Mol. Biol. 296 (2000) 509–520.