Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Biologia

SCUOLA DI DOTTORATO DI RICERCA IN BIOSCIENZE
INDIRIZZO: BIOLOGIA EVOLUZIONISTICA
CICLO XXII

GENETIC CONTROL OF MOULTING AND SEGMENTATION DURING POST-EMBRYONIC DEVELOPMENT IN *Lithobius peregrinus* (CHILOPODA, LITHOBIOMORPHA)

Direttore della Scuola: Ch.mo Prof. Tullio Pozzan
Coordinatore d'indirizzo: Ch.mo Prof. Giorgio Casadoro
Supervisore: Dott. Giuseppe Fusco

Dottorando: Dott.ssa Francesca Bortolin

CONTENTS

ABSTRACT **1**
RIASSUNTO **3**
INTRODUCTION **5**
Postembryonic molting and segmentation in arthropods **5**
Molting genes **6**
Segmentation genes **7**
Post embryonic development in Lithobiomorpha **8**
Questions targeted by this thesis **9**
References **10**
MANUSCRIPT 1- Cloning and expression pattern of the ecdysone receptor and the retinoid X receptor from the centipede *Lithobius peregrines* **13**
Abstract **15**
Introduction **15**
Materials and methods **17**
Results **20**
Discussion **29**
References **31**
Appendix **34**
MANUSCRIPT 2- Post-embryonic expression of two segmentation genes, *engrailed* and *wingless*, during the anamorphosis of *Lithobius peregrinus* **39**
Abstract **41**
Introduction **41**
Materials and methods **43**
Results **47**
Discussion **55**
References **57**
Appendix **59**
GENERAL CONCLUSIONS **61**
References **62**

Abstract

The great diversity of arthropod body plans makes this group a remarkable taxon for studying the evolutionary diversification of developmental patterns. In the last decades, developmental genetic studies have been mainly focalized on embryonic development in several arthropod species, whereas post-embryonic development has been investigated to a considerably less extent, and within the limits of very few model species.

In many arthropods, post-embryonic development is characterized by an increase in the number of trunk segments. This process is called anamorphosis.

In lithobiomorph centipedes juveniles hatch with seven leg pairs and during the first five post-embryonic stages the definitive arrangement with fifteen leg-bearing segments is reached. In these species, despite conspicuous individual variation in growth rates and in temporal molting schedule, each anamorphic stage is characterized by a precise segmental composition of the body. Growth, molting and segmentation are thus intimately correlated developmental processes, and it is easy to hypothesize that variation (and evolvability) in each one of these processes is strictly dependent on the precise relationship it has with the others.

The aim of this research was to start an investigation of post-embryonic molting and segmentation process in the centipede *Lithobius peregrinus* through a genetic approach.

In arthropods, molting events driven by oscillations in the titer of the ecdysone, are mediated by the binding of the hormone to a heterodimer of nuclear receptors consisting of the ecdysone receptor (EcR) and the retinoid X receptor (RXR), a homologue of ultraspiracle (USP). In insects, the 20E-receptor complex directly activates a small group of so called 'early genes': *Broad-Complex* (*Br-C*), *E74* and *E75*, each of which encodes a set of transcriptional factors 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', involved in larval-pupal transformation.

In this study we have cloned partial sequences for *EcR* and *RXR* gene homologues from *L. peregrinus* (*LpEcR* and *LpRXR*) and we have analyzed their expression profiles during the second larval stage (L1). Sequence comparison of LpEcR and LpRXR with their orthologues in other arthropods shows a maximal degree of identity with chelicerates and hemimetabolous insects (68-81%), and a lesser degree with crustaceans and some holometabolous insects (65-70%). A distinctive character of *L. peregrinus* EcR and RXR is the presence of several variants: we identified two isoforms for EcR and three isoforms for the RXR receptor. Further studies are necessary to clarify the function of these different variants.

The expression profiles shown by LpEcR and LpRXR during the anamorphic stage L1 of *L. peregrinus* are similar. Expression levels of both receptors increase in the second part of the inter-molt period, suggesting that molting process start 48 hours after the previous molt. Similar expression profiles are found in other arthropods.

We failed to isolate an orthologues of *BR-C*. The lack of homologues in *L. peregrinus* and crustaceans (B. Konopova, personal communication) suggests that it could be an apomorphic trait of insects.

The genetic basis of the developmental mechanisms of body segmentation are well known in the fruit fly *Drosophila melanogaster*. So-called ‘segmentation genes’ discovered in the fly are classified into gap, pair-rule and segment-polarity genes. These genes occupy different levels in a hierarchical cascade leading from the early gap genes to the later-expressed pair-rule and segment-polarity genes. The latter (like, *engrailed*, *wingless*, *hedgehog* and *cubitus interruptus*) encode proteins that are expressed in the embryo as stripes with segmental periodicity, and that are required for the formation of the correct pattern of structures within each segment.

Homologs of several segment-polarity and pair-rule genes have been studied in other insects, as well as in chelicerates, myriapods, and crustaceans, but we lack basic information about expression of segmentation genes during post-embryonic development.

We investigated the expression of two segment-polarity genes, *engrailed* and *wingless*, already known to be involved in embryonic segmentation of centipedes, to ascertain their possible role in segmentation during anamorphic stages. We showed that in *L. peregrinus*, *en* and *wg*, are intensively expressed also during all anamorphic stages. Only *en* seems retain a segment-polarity role in post-embryonic development, although with different modality of expression, whereas *wg* function seems related to the development of nervous system and leg buds, but with no sign of periodic expression, thus suggesting that it does not retain the same role in segmental patterning that it has during embryonic development.

This research also marks a significant methodological advancement, consisting in the protocol set up, in a centipede, for two very powerful techniques of modern molecular biology: Real Time PCR and *in situ* hybridization on paraffin sections.

The results of this work on the anamorphosis of *Lithobius peregrinus* represent the first molecular data on post-embryonic development in a myriapod species, that add to the very few studies in non-insect arthropods. This work thus represents a basic starting point for the study of development and evolution of molting and segmentation in a poorly investigated phase of arthropod life cycle.

Riassunto

La grande diversificazione nell'organizzazione del piano corporeo degli artropodi fa di questo gruppo un taxon molto interessante per lo studio dell' evoluzione dei modelli di sviluppo. Negli ultimi decenni, gli studi di genetica dello sviluppo si sono focalizzati principalmente sulla fase embrionale in diverse specie di artropodi, mentre lo sviluppo post-embrionale è stato raramente studiato solo in alcune specie modello.

In molti artropodi lo sviluppo postembrionale è caratterizzato da un aumento del numero dei segmenti del tronco. Nei centopiedi litobiomorfi, l'animale emerge dall'uovo con soli sette segmenti del tronco perfettamente formati raggiungendo poi, al termine delle prime cinque mute, l'assetto definitivo con quindici segmenti pediferi. Questo tipo di sviluppo viene detto emianamorfico. In queste specie, nonostante ci siano ampie variazioni intraspecifiche nel tasso di crescita e nella scansione delle mute, ogni stadio anamorfico è caratterizzato da un precisa composizione segmentale del corpo. Crescita, muta e segmentazione sono quindi processi di sviluppo strettamente correlati ed è facilmente ipotizzabile che la variazione (e la capacità di evolvere) di ciascun processo sia strettamente dipendente dalla precisa relazione con gli altri.

L' obiettivo di questa ricerca è stato lo studio dei processi di muta e di segmentazione durante lo sviluppo postembrionale nel centopiedi *Lithobius peregrinus*, utilizzando un approccio molecolare.

Negli artropodi, la muta è controllata dall'oscillazione dei livelli dell'ecdisone. A livello molecolare, questo ormone svolge la sua funzione regolativa legandosi ad un complesso eterodimerico formato da due recettori nucleari: il recettore dell'ecdisone (EcR) e il recettore X retinoico (RXR), omologo del recettore ultraspiracle (USP). Negli insetti, il complesso EcR/USP regola direttamente i geni target primari dell'ecdisone, tra i quali *Broad-Complex* (*BR-C*), *E74* ed *E75*. Questi geni codificano fattori di trascrizione che mediano ed amplificano il segnale ormonale regolando un ampio assortimento di geni target secondari coinvolti nel passaggio larva-pupa.(review in Thummel, 1995; King-Jones & Thummel 2005).

In questo studio sono state clonate sequenze parziali dei geni omologhi EcR e RXR in *L. peregrinus* (*LpEcR* and *LpRXR*) ed è stato analizzato il loro profilo d'espressione durante il secondo stadio larvale (L1). Un confronto delle sequenze di *LpEcR* and *LpRXR* con sequenze ortologhe in altri artropodi ha evidenziato un'elevato livello d'identità con i chelicerati e gli insetti etero metaboli (68-81%), mentre con i crostacei e gli insetti olometaboli meno derivati (65-70%) l'identità è leggermente inferiore.

Una caratteristica di *L. peregrinus* è la presenza di più varianti di sequenza per entrambi i recettori: sono state identificate due isoforme per *EcR* e tre isoforme per *RXR*. Ulteriori analisi saranno necessarie per chiarire un eventuale diversa funzione di queste varie isoforme.

Il profilo d'espressione osservato per *LpEcR* e *LpRXR* durante lo stadio L1 è simile. I livelli d'espressione di entrambi i geni aumentano nella seconda parte dello stadio, e ciò suggerisce che il processo di muta in questo stadio inizi dopo circa 48 ore dalla precedente ecdisi. I risultati ottenuti sono in linea con quanto è stato osservato in altri artropodi.

Il mancato riscontro di un ortologo del gene *Broad-Complex* in *L. peregrinus* e nei crostacei (B. Konopova, comunicazione personale) fa supporre che si tratti di un'apomorfia degli insetti.

Le basi genetiche dei meccanismi di sviluppo della segmentazione sono stati ampiamente studiati nel moscerino della frutta *Drosophila melanogaster*. I cosiddetti 'geni della segmentazione' in Drosofila sono stati classificati in geni 'gap', 'pair-rule' e 'segment-polarity'. Questi geni occupano diversi livelli gerarchici nella cascata di regolazione genica della segmentazione: a monte ci sono i geni 'gap', successivamente i geni 'pair-rule' ed i geni 'segment-polarity'. Questi ultimi (come *engrailed*, *wingless*, *hedgehog* and *cubitus interruptus*) codificano proteine che vengono espresse nell'embrione con periodicità segmentale e che sono richieste per la corretta formazione del pattern di strutture entro ciascun segmento.

Geni omologhi dei geni 'pair-rule' e 'segment-polarity' sono stati studiati in diversi insetti, nei chelicerati, nei crostacei e nei miriapodi, ma mancano informazioni basilari sulla loro espressione durante la segmentazione postembrionale.

In questa ricerca è stata studiata l'espressione di due geni 'segment-polarity', *engrailed* e *wingless*, già coinvolti nella segmentazione embrionale dei centopiedi, per stabilire il loro ruolo nella segmentazione durante la fase anamorfica dello sviluppo. E' stato così dimostrato che in *L. peregrinus*, *en* e *wg* sono intensamente espressi durante l'anamorfosi. Solamente *en* sembra conservare un ruolo nella segmentazione postembrionale, sebbene con modalità d'espressione diverse, mentre *wg* sembra legato allo sviluppo del sistema nervoso e degli abbozzi degli arti, e non presenta espressione periodica. Questo suggerisce che questo gene non svolga più nel processo di segmentazione il ruolo che aveva durante lo sviluppo embronale.

Questa ricerca segna inoltre un avanzamento metodologico significativo: per la prima volta in un centopiedi sono stati messi a punto i protocolli di due tecniche di biologia molecolare molto importanti: la Real Time PCR e l'ibridazione in situ su sezioni di paraffina.

I risultati di questo studio sull'anamorfosi di *Lithobius peregrinus* rappresentano i primi dati molecolari disponibili sullo sviluppo postembrionale di un miriapode e rappresentano un importante punto di partenza per lo studio dello sviluppo e dell'evoluzione dei processi di muta e segmentazione in una fase dello sviluppo ancora poco conosciuta.

Introduction

The great diversity of arthropod body plans makes this group a remarkable taxon for studying the evolutionary diversification of developmental patterns. In the last decades, developmental genetic studies have been mainly focalized on embryonic development, and significant results have been obtained through the application of new molecular techniques in different arthropods species (e.g., Stollewerk, 2000; Janssen et al., 2004; Hatini et al., 2005; Brena et al., 2006). In contrast, developmental genetics of post-embryonic development has been investigated to a considerably lesser extent, and within the limits of very few model species.

Recent advances in the study of arthropods, e.g. about the origin and evolution of holometaboly in insects (Truman and Riddiford, 1999; Heming, 2003), have amply demonstrated that there is much more to late development than simply going on with processes already put in place during early embryogenesis. But holometabolous insects are not the only arthropods whose body architecture changes extensively during the life cycle.

The focus of this study is post-embryonic segmentation and its relationship with the molting cycle in a centipede.

Postembryonic molting and segmentation in arthropods

Arthropods growth is punctuated by the periodic shedding of the old exoskeleton that is replaced by a new and (usually) larger one. This process is called ecdysis, and the time between two successive ecdyses defines the molting cycle. In many species, post-embryonic growth is accompanied by a more or less extensive change in body organization, thus the sequence of developmental stages defined by the molting cycle may present a variably discontinuous ontogenetic change. This is the case of the segmental patterning of the main body axis.

In many arthropods, post-embryonic development is characterized by an increase in the number of trunk segments. This process is called anamorphosis. Among these groups there are sea spiders, proturans, most crustaceans and most myriapods.

In anamorphic species, juveniles hatch with an incomplete complement of segments, and the expected adult number of segments (when fixed) is reached later in ontogeny.

Different kinds of anamorphic development can be distinguished (Enghoff et al. 1993): *euanamorphosis*, when the addition of new segments continues until the last molt the animal undergoes, without any evidence of an expected fixed terminal number; *teloanamorphosis*, when the animal does not molt any more after it has reached the final number of segments; *hemianamorphosis*, when the final and fixed number of segments is reached after a number of molts, but growth continues through further molts without further increase in the number of body segments.

Post-embryonic segmentation schedule is invariant within many anamorphic species (as, for instance, in all anamorphic centipedes). In these species, despite conspicuous individual variation in growth rates and in temporal molting schedule, each anamorphic stage is characterized by a precise segmental composition of the body. Growth, molting and segmentation are thus intimately

correlated developmental processes, and it is easy to hypothesize that variation (and evolvability) in each one of these processes is strictly dependent on the precise relationship it has with the others.

Molting genes

Two hormones are the main actors in insect growth and development: the steroid α-ecdysone (and its metabolite 20-hydroxyecdysone, 20E) and the sesquiterpenoid juvenile hormone (JH) (Riddiford, 1996). Ecdysteroids and JH-like compounds were isolated also in crustaceans (Spinder et al., 1980) and chelicerates (Diehl et al. 1986). According to the classical tenets of insect endocrinology, the balance of the two hormones defines the outcome of each developmental transition.

Ecdysone synthesis occurs primarily in molting glands: prothoracic glands of insects and Y-organs of crustacean (Lachaise et al., 1993). The molecular mechanism of the 20E action has been extensively studied, especially in insects. The molecular target of the ecdysteroids is known as an ecdysone receptor (EcR), which belongs to the nuclear receptor family. EcR is a ligand-dependent transcription factor and it activates transcription of target genes by forming a heterodimer with another member of the nuclear receptor family, ultraspiracle protein (USP), which is the insect homologue of vertebrate retinoid X receptor (RXR) (Yao et al., 1992).

The 20E-receptor complex binds to an ecdysone response element within the promoters of target genes, altering their expression and thereby leading to changes in the genetic programs associated with molting (Li et al., 2003). Each target tissue propagates the 20E signal through a genetic regulatory hierarchy. In *Drosophila*, 20E bound to the receptor directly activates a small group of early genes: Broad-Complex (BR-C), E74 and E75, each of which encodes a set of protein isoforms of DNA binding transcriptional regulators (Riddiford, 2001). The protein products of early genes activate a much larger group of late genes that directly or indirectly perform distinct morphogenetic processes such as cell death, cell proliferation and tissues differentiation (Thummel, 2002).

EcR and *RXR*, the first genes activated by hormonal signal, show high levels of sequence conservation in arthropods. They have been isolated from several species (e.g. Hayward et al.,1999, 2003; Asazuma et al., 2007; Nakagawa et al., 2007), but further studies, especially in non-insect arthropods, are necessary to clarify their expression profile, identify their target genes and understand their precise role in molting regulation.

In myriapods, the endocrine system has been investigated through classical experiment of extirpation and/or reimplantation and biochemical, histological and ultrastructural analysis (e.g., Joly and Descamps, 1988), but a molecular approach have never been pursued. Endocrinological studies focussed on hormonal regulation of molting in *L. forficatus* identified lymphatic strands surrounding the salivary glands as a putative molting gland (Scheffel, 1969; Leubert, 1986; Joly and Descamps, 1988). The products of neurosecretory cells in the frontal lobes of the protocerebrum and the associated cerebral glands have inhibitory function on molt, whereas a role in stimulation by neurosecretory cells of the pars intercerebralis hypothesized in the adult (Joly, 1966) but not confirmed in the larvae (Scheffel, 1969). To date, neither specific molting hormones nor neuropeptides have been identified in myriapods, and nothing is known about the

genetic regulation pathways involved in molting process. A fragment of a RXR orthologues was isolated in *L. forficatus* in the context of a phylogenetic study on arthropods (Bonneton et al., 2003), but it was not characterized.

A study of *L. peregrinus* molting hormones is the subject of the first manuscript presented in this thesis.

Segmentation genes

Comparative analyses based on phylogeny and paleontological data suggest that anamorphosis is the primitive segmentation mode in arthropods (see Fusco 2005). However, beyond descriptive morphology data, very little is known about this important developmental process. We lack basic information about the expression of segmentation genes during post-embryonic development and about morphogenetic processes involved in anamorphosis (e.g. frequency and distribution of mitoses, orientation of the spindles, cell migration, etc.). Post-embryonic segmentation is traditionally considered to be the product of morphogenetic activity of a so-called subterminal 'generative' (or 'proliferative') zone, but this zone has never been suitably characterized.

The genetic basis of the developmental mechanisms of body segmentation is well known in *Drosophila melanogaster* (Martinez-Arias, 1993). In the fly embryo, a cascade of maternal and zygotic gap genes is activated in the syncitial blastoderm that subdivides the ectoderm into smaller domains along the anterior-posterior (AP) body axis. A further subdivision of the main axis in seven transverse stripes, is accomplished by the activation of the pair-rule genes, whose expression represents the first sign of a periodic organization of the *Drosophila* body. Pair-rule genes expression defines embryo regions called parasegments. These are serial units of the same length as a segment, but shifted posteriorly about one third of a segment. In insects (Martinez-Arias and Lawrence, 1985), crustaceans (Dohle and Scholtz, 1988) and chelicerates (Damen, 2002), this initial segmental organization of the early embryo is subsequently replaced by the final segmental organization, observable in the later embryo, the larva and the adult. The last genes to act are the segment-polarity genes like *engrailed*, *wingless*, *hedgehog* and *cubitus interruptus*, which are required for the formation of the correct pattern of structures within each segment (review in Rivera-Pomar and Jäckle, 1996).

In *Drosophila* embryo all segments are thus originated almost simultaneously, while the blastoderm is still syncytial, but in other arthropods, more often only the most anterior segments originate synchronously, whereas the remaining segments appear sequentially from a posterior sub-terminal 'proliferative zone' (Davis and Patel, 2002).

In recent times, a small number of model organisms for the study of segmentation, other than *Drosophila* plus some other insect species (Davis and Patel, 2002) have been investigated (Damen, 2002; Hughes and Kaufman, 2002; Dearden et al., 2002; Chipman et al., 2004; Janssen et al., 2004) and in all these arthropods embryonic segmentation is sequential (i.e. develops in anterior-posterior direction, as during post-embryonic segmentation), rather than simultaneous as in *Drosophila*. At least for a significant posterior portion of the main body axis, sequential segmentation is generally considered the primitive condition in arthropods (Peel, 2004).

Until now, homologs of several segment-polarity and pair-rule genes have been studied in insects as well as in chelicerates, myriapods, and crustaceans (Abzhanov and Kaufman 2000; Damen, 2002; Hughes and Kaufman, 2002). However, most studies on segmentation in arthropods other than *Drosophila* is limited to the description of expression patterns of segmentation genes. Very few studies have been carried out on gene function in other insects (Copf et al., 2004; Shinmyo et al., 2005) and in the crustacean *Artemia franciscana* (Copf et al., 2004), using RNA interference or knockdown technique.

Current knowledge about developmental genetics of lithobiomorph centipedes is based on studies on two *Lithobius* species, *L. forficatus* and *L. atkinsoni*, and it is anyway limited to embryonic development. In particular, genes involved in segment formation (Hughes and Kaufman, 2002) and neurogenesis (Kadner and Stollewerk, 2004) have been identified. Hughes and Kaufman's (2002) work on the expression patterns of the genes *even-skipped* (eve), *engrailed* (*en*) and *wingless* (*wg*) in *L. atkinsoni* suggests that their basic roles in embryonic segmental patterning is largely conserved across the arthropods.

A study of *L. peregrinus* segmentation gene expression during post-embryonic development is the subject of the second manuscript presented in this thesis.

Post embryonic development in Lithobiomorpha

Lithobiomorph centipedes are hemianamorphic. During the first five post-embryonic stages, segment number increases through an invariant schedule of segment addition, from seven to fifteen leg-bearing segments (anamorphic phase of post-embryonic development). This phase is followed by additional stages in which the number of segments remains constant (epimorphic phase of post-embryonic development). Hemianamorphic development is typical of all pauropods, symphylans and basal millipedes (though not exclusively), but it is also found in basal chelicerates (Pycnogonida), basal hexapods (Protura), most crustaceans, and it was the typical mode of post-embryonic development in trilobites (Minelli et al. 2003). Among centipedes (Chilopoda), hemianamorphic are also the Scutigeromorpha and the Craterostigmomorpha, whereas the Scolopendromorpha and the Geophilomorpha are epimorphic, that is, begin their post-embryonic life with a full complement of segments.

Anamorphic stages are customary labeled as larval stages (from that the labeling L0-L4 for the five anamorphic stages), although the term 'juvenile' could be more appropriate. **Table 1** is a schematic representation of the anamorphic schedule of *L. peregrinus*. Note that although the anamorphic phase is defined on the basis of the formation of functional leg pairs, different segmental structures (tergites, sternites, the tracheal or the nervous system) could suggest a different point of transition between juvenile and adult segmental condition (see Minelli et al., 2006).

Tab. 1. Number of serial trunk structures in the first seven post-embryonic stages of *L. peregrinus*. L0-L4: anamorphic stages. PL1-PL2: first two post-anamorphic ('post-larval') stages. Figures in the column 'leg pairs' represent the number of pairs of 'completely formed legs' + 'incompletely formed legs' + 'leg buds'. Figures in parentheseis denote a partially incomplete formation of the last element of the segmental series. The gray background indicates the condition also found in the adult.

stage	leg pairs	tergites	sternites	ganglia pairs	spiracle pairs
L0	7 + 0 + 1	8	8	9	-
L1	7 + 1 + 2	8	8	10	2
L2	8 + 0 + 2	10	10	11	2
L3	10 +0 + 2	12	12	13	3
L4	12 + 0+ 3	14	(15)	15	4
PL1	15	(15)	15	15	5
PL2	15	15	15	15	6

Questions targeted by this thesis

The aim of this research was to start an investigation of post-embryonic molting and segmentation in an anamorphic arthropod through a genetic approach.

I tried to identify homologs of the genes involved in the ecdysone signaling cascade of *L. peregrinus*, focusing on the ecdysone receptor (EcR) and retinoid X receptor (RXR). Furthermore, I investigated the expression of two segment-polarity genes, *en* and *wg*, already known to be involved in embryonic segmentation, to ascertain their possible role in segmentation during anamorphic stages of this species.

The results of this work on the anamorphosis of *Lithobius peregrinus* represent the first molecular data on post-embryonic development in a myriapod species, that add to the very few studies performed thus far in non-insect arthropods. This work thus represent a basic starting point for the study of development and evolution of molting and segmentation in a poorly investigated phase of arthropod life cycle.

References

Abzhanov, A., and Kaufman, T. C. (2000). Homologs of Drosophila appendage genes in the patterning of arthropod limbs. Developmental Biology, **227**: 673–689.

Asazuma, H., Nagata, S., Kono, M. and Nagasawa, H. (2007). Molecular cloning and expression analysis of ecdysone receptor and retinoid X receptor from the kuruma prawn, *Marsupenaeus japonicus*. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, **148:**139-150.

Bonneton, F., Zelus, D., Iwema, T., Robinson-Rechavi, M., Laudet, V. (2003). Rapid divergence of the ecdysone receptor in Diptera and Lepidoptera suggests coevolution between ECR and USP-RXR. Molecular Biology and Evolution, **20:** 541–553**.**

Brena, C., Chipman, A.D., Minelli, A., Akam, M. (2006). Expression of trunk Hox genes in the centipede Strigamia marittima: sense and anti-sense transcripts. Evolution and Development, 8:252-265.

Chipman, A. D.; Arthur, W.; Akam, M. (2004): Early development and segment formation in the centipede, *Strigamia maritima* (Geophilomorpha). Evolution and Development **6**: 78-89

Copf T., Schröder R., Averof M. (2004). Ancestral role of caudal genes in axis elongation and segmentation. Proceedings of the National Academy of Science of the United States, **101**:17711-17715.

Damen, W.G.M. (2002). Parasegmental organization of the spider embryo implies that the parasegment is an evolutionary conserved entity in arthropod embryogenesis. Development, **129**:1239– 1250.

Davis, G. K., and Patel, N. H. (2002). Short, long and beyond: molecular and embryological approaches to insect segmentation. Annual Review in Entomology, **47**: 669–699.

Dearden, P.K, Donly, C., Grbic, M. (2002). Expression of pair-rule gene homologues in a chelicerate: early patterning of the two-spotted spider mite *Tetranychus urticae***.** *Development* , **129:**5461-5472.

Diehl, P. A., Connat, J.L., Dotson, E. (1986). Chemistry, function and metabolism of tick ecdysteroids. In *Morphology, Physiology, and Behavioral Biology of Ticks* (Sauer J. R. and. Hair J. A eds). John Wiley and Sons, New York, 165–192.

Dohle, W. and Scholtz, G. (1988). Clonal analysis of the crustacean segment: the discordance between genealogical and segmental borders. Development, **104**:147–160.

Enghoff H., Dohle W., Blower, J.G. (1993). Anamorphosis in millipedes (Diplopoda)—the present state of knowledge with some developmental and phylogenetic considerations. Zoological Journal of the Linnean Society, **109**:103–234

Fusco G. (2005).Trunk segment numbers and sequential segmentation in myriapods. Evolution & Development, **7**: 608-617.

Hatini, V., Green, R. B., Lengyel, J. A., Bray, S. J. and Dinardo, S. (2005). The Drumstick/Lines/Bowl regulatory pathway links antagonistic Hedgehog and Wingless signalling inputs to epidermal cell differentiation. *Genes and Development,* **19**:709-718.

Hayward D.C., Bastiani M.J., Trueman J.W.H., Truman J.W., Riddiford L.M. and Ball E.E. (1999) .The sequence of locusta RXR, homologous to *Drosophila* Ultraspiracle, and its evolutionary implications. Development Genes and Evolution, **209**:564-571.

Hayward, D. C., Dhadialla, T. S., Zhou, S., Kuiper, M. J., Ball, E. E., Wyatt, G. R. and Walker, V. K. (2003). Ligand specificity and developmental expression of RXR and ecdysone receptor in the migratory locust. Journal of Insect Physiology*,* **49:** 1135-1144.

Heming, B.S. (2003). Insect Development and Evolution. Ithaca: Cornell University Press.

Hughes C.L. and Kaufman T.C. (2002). Exploring myriapod segmentation: the expression patterns of *even-skipped*, *engrailed*, and *wingless* in a centipede. Developmental Biology, **247**:47-61.

Janssen, R. Prpic, N.-M., & Damen, W.G.M. (2004). Gene expression suggests decoupled dorsal and ventral segmentation in the millipede *Glomeris marginata* (Myriapoda: Diplopoda). *Developmental Bioogy,* 268: 89-104.

Jolv. R. (1966). Sur l'ultrastructure de la glande cerebrale de *Lithobius forficatus*, L. (Myriapode. Chilopode). Comptes Rendus hebdomadaires des Séances de l'Academie des Sciences de Paris, 263:374-377.

Joly. R., and Descamps, M. (1988) Endocrinology of Myriapods. In Endocrinology of selected invertebrate types. Laufer H. (ed.), New York: Alan R. Liss. lnc., 430-449.

Kadner, D., and Stollewerk, A. (2004). Neurogenesis in the chilopod *Lithobius forficatus* suggests more similarities to chelicerates than to insects. Development Genes and Evolution, **214:**8367–79.

Lachaise, F., Le Roux, A., Hubert, M., Lafont, R.(1993). The molting gland of crustaceans: localization, activity and endocrine control (a review). Journal of Crustacean Biology, **13**:198–234.

Leubert,. F. (1986). Untersuchungen zur Ecdysteroid—Biosynthese durch das Lymphstrang—Gewebe von *Lithobius forficatus* (L.) (Chilopoda). Zoologischer Jahresbücher für Physiologie, **83**:334–339.

Li, T.R. and White, K.P. (2003). Tissue-specific gene expression and ecdysone-regulated genomic networks in Drosophila. Developmental Cell, **5**:59–72.

Martinez Arias, A. (1993). Development and patterning of the larval epidermis of *Drosophila*. In The development of Drosophila melanogaster, (Bate and Martinez- Arias Eds.). Cold Spring Harbor, 747-841.

Martinez-Arias A., Lawrence P.A. (1985). Parasegments and compartments in the *Drosophila* embryo. Nature, **313**:639–642.

Minelli A., Fusco G. & Hughes N. C. (2003). Tagmata and segment specification in trilobites. In Proceedings of the Third International Symposium on trilobites and their relatives (Lane P. D., Fortey R. A. and Siveter D. J. eds). Special Papers in Palaeontology, **70**: 31-43.

Minelli A., Brena C., Deflorian G., Maruzzo D., Fusco G. (2006). From embryo to adult – beyond the conventional periodization of arthropod development. Development Genes & Evolution, **216**: 373-383.

Nakagawa, Y., Sakai, A., Magata, F., Ogura, T., Miyashita, M.,Miyagawa, H. (2007). Molecular cloning of the ecdysone receptor and the retinoid X receptor from the scorpion *Liocheles australasiae*. FEBS Journal, **274**:6191–6203.

Peel, A. (2004). The evolution of arthropod segmentation mechanisms. Bio-Essays, **26**:1108–1116.

Riddiford, L. (1996). Molecular aspects of juvenile hormone action in insect metamorphosis. *In Metamorphosis: Postembryonic reprogramming of gene expression in amphibian and insect cells* (L. I. Gilbert, J. R. Tata, and B. G. Atkinson eds.). Academic Press, San Diego, 223–253.

Riddiford, L. M. (2001). Ecdysone Receptors and their Biological Actions. Vitamins and hormones**, 60**:1-73.

Rivera-Pomar R. and Jäckle H. (1996). From gradients to stripes in *Drosophila* embryogenesis: Filling in the gaps. Trends in Genetics, **12**: 478–483.

Scheffel H. (1969). Untersuchungen uber die hormonale Regulation von Hautung und Anamorphose von *Lithobius forficatus* (L.) (Myriapoda, Chilopoda). Zoologische Jahrbücher Abteilung für allgemeine Zoologie und Physiologie der Tiere, **74**: 436-505.

Shinmyo Y., Mito T., Matsushita T., Sarashina I, Miyawaki K., Ohuchi H., Noji S. (2005). *caudal* is required for gnathal and thoracic patterning and for posterior elongation in the intermediate-germband cricket *Gryllus bimaculatus*. Mechanism of Development, **122**:231-239.

Spindler, K.D., Keller, R. , O'Connor, J. D. (1980). The role of ecdysteroids in the crustacean molting cycle. In Progress in Ecdysone Research (Hoffmann, J., ed.). Elsevier, Amsterdam, 247–280.

Stollewerk, A. (2000). Changes in cell shape in the ventral neuroectoderm of Drosophila melanogaster depend on the activity ofthe achaete-scute complex genes. Development Genes and Evolution, **210**:190–199.

Thummel, C.S. (2002). Ecdysone-regulated puff genes 2000. Insect Biochemistry and Molecular Biology, **32**:113–120.

Truman, J. W. and Riddiford, L. M. (1999). The origins of insect metamorphosis. Nature, **401**:447–452.

Yao, T. P. (1993). Functional ecdysone receptor is the product of *EcR* and *Ultraspiracle* genes. Nature, **366**:476-479.

Manuscript 1

Cloning and expression pattern of the ecdysone receptor and the retinoid X receptor from the centipede *Lithobius peregrinus*

Cloning and expression pattern of the ecdysone receptor and the retinoid X receptor from the centipede *Lithobius peregrinus*

Abstract

In arthropods, molting events driven by oscillations in the titer of the ecdysone, are mediated by the binding of the hormone to a heterodimer of nuclear receptors consisting of the ecdysone receptor (EcR) and the retinoid X receptor (RXR), a homologue of ultraspiracle (USP). In insects, crustacean and chelicerates, molecular basis of ecdysteroids action has been analyzed in great detail, whereas knowledge about endocrine process that regulate molting in myriapods are missing. In this study we have cloned partial sequences of several isoforms for *EcR* and *RXR* gene homologues from the centipede *Lithobius peregrinus* (*LpEcR* and *LpRXR*). Their amino acid sequences are very similar to other arthropod orthologues, especially to chelicerate and hemimetabolous insect ones. We investigate LpEcR and LpRXR expression patterns during the second post-embryonic stage, showing that expression levels of both receptors increase in the second part of the inter-molt period, suggesting that molting process start 48 hours after the previous molt. These results confirm some expression data obtained in other arthropods. Results obtained in this study represent the first data on the genes involved in the ecdysone signal pathway in a myriapod species, thus contributing to a more comprehensive understanding of the developmental processes mediated by these genes in arthropods.

1. Introduction

Arthropod ecdysteroids (zooecdysteroids) are steroid hormones responsible for regulating processes associated with development, metamorphosis, reproduction and diapause, a prominent example being the α-ecdysone released by the prothoracic gland of insects. In this group, molting is driven by oscillations in the titer of the α-ecdysone or its biologically active form, the 20-hydroxyecdysone, 20E (hereafter referred to as ecdysone; Riddiford et al., 2001).

The rise in concentration of ecdysone during development 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 the retinoid X receptor (RXR), a homologue of ultraspiracle (USP) (Yao et al., 1993). 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 orthologue has been reported only in some parasitic nematodes, as *Dirofilaria immitis* (Shea et al., 2010). The 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 for increasing binding affinity of ecdysteroids to EcR and for transcriptional activity (Yao et al., 1993). RXR orthologues have been reported from arthropods and other invertebrates, including the cubozoan jellyfish *Tripedalia cystophora* (Kostrouch et al., 1998) and the nematode *Brugia malayi* (Tzertzinis et al., 2010).

EcR and RXR belong to the superfamily of nuclear receptors (NR) and share NR-typical domain structures and gene regulatory mechanisms. NRs are

characterized by a structure comprising five distinct protein domains (Evans, 1988; Billas et al., 2009): (i) A/B domain, a highly variable N-terminal domain involved in transcriptional activation; (ii) C domain, a highly conserved DNA-binding domain (DBD); (iii) D domain, a flexible and variable hinge region involved in ecdysone-response elements recognition and heterodimerization; (iv) E domain, a rather complex ligand-binding domain (LBD) which is involved in hormone binding, heterodimerization and interaction with other transcription factors; and finally (v) a C-terminal F domain, a highly variable domain whose function is not well understood. For both proteins, the amino acid sequences of the C (DNA-binding) and E (ligand-binding) domains are the most highly conserved. As a consequence, EcR and USP/RXR orthologues have been cloned in several insects (e.g. *Locusta migratoria*; Hayward et al., 1999, 2003), crustaceans (e.g., *Marsupenaeus japonicus*; Asazuma et al., 2007) and chelicerates (e.g., *Liocheles australasiae*; Nakagawa et al., 2007).

In insects, the 20E-receptor complex directly activates a small group of so called 'early genes', among which *Broad-Complex* (*Br-C*), *E74* and *E75*, each encoding a set of transcriptional factors 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' during the last larval instar encode tissue-specific effector proteins necessary for the developmental events that drive metamorphosis (Karim et al., 1993; Thummel, 1995; Riddiford et al., 2001).

Br-C encodes a family of four classes of protein isoforms (Z1–Z4), which share a common aminoterminal core domain, alternatively spliced to four distinct carboxy-terminal domains bearing pairs of zinc-finger DNA-binding domains (DiBello *et al.* 1991, Bayer *et al.* 1996). The common core region contains a highly conserved domain of 120 amino acids, called the BTB (for *bric-à-brac, tramtrack* and *Broad Complex,* three *Drosophila* genes where it was first identified) or POZ domain (for POx virus and Zinc finger domain). It appears to be involved in protein-protein interactions that affect binding to DNA (DiBello *et al.,* 1991). The *Br-C* gene has been identified as a key gene required for insect molting, metamorphosis and oogenesis.

In myriapods, the endocrine system has been investigated through extirpation/reimplantation experiments, immuno- and radio-essays and histological and ultrastructural analysis (e.g., Joly and Descamps,1988), but a biochemical and molecular approach have never been pursued. Thus, neither the specific hormones nor the genes involved in molting process have been identified.

In *Lithobius forficatus,* lymphatic strands surrounding the salivary glands are credited of functioning as ecdysteroidogenic glands. This hypothesis is supported by ultrastructural similarity with the prothoracic glands of insects, and biochemical affinity of their secretions to ecdysteroids of other arthropods (Seifert and Bidmon, 1988). Injections of exogenous ecdysone increases the number of molts in adult specimens (Joly 1964), but hormone titer has never carried out during anamorphic development. A fragment of a RXR orthologues was isolated in *L. forficatus* (GeneBank accession number)for a molecular phylogeny study in arthropods (Bonneton et al., 2003), but the authors did not provide any information about sequence characterization.

In this study we have cloned partial sequences for *EcR* and *RXR* gene homologues from *L. peregrinus* (*LpEcR* and *LpRXR*), and we have analyzed their

expression profiles during the second larval stage (L1). We also tried to isolate an orthologues of *BR-C* in a non-insect arthropod, but we only obtained the sequence of BTB domain.

2. Materials and methods

2.1 Centipede husbandry

Centipedes were collected during years 2008-2009 in a house garden in San Stino di Livenza (North-eastern Italy) and identified as *L. peregrinus* by Dott. Marzio Zapparoli (University of Tuscia, Italy). Adults were housed in plastic boxes with a hardened poured plaster-of-Paris floor to maintain elevate humidity and pieces of wood barks to let the animals hide underneath. Boxes were sprayed with water every few days and living crickets were provided weekly as food.

Eggs were collected periodically by rinsing out the woods and boxes with water and catching the eggs in a sieve. Eggs were kept in Petri dishes with a humid plaster floor, to hatch in 15-20 days since deposition. Since hatching, larvae were bred individually in Petri dishes with the same humid medium on the floor. They were checked daily for molt and fed with living fruit flies. All stages were kept at 21±1 °C under natural photoperiod.

Despite controlled environmental parameters, stage duration is quite variable between individuals. The first two molts (L0-L1 and L1-L2) occur within 1.5-3.5 days of post-embryonic life, whereas the duration of the following stages is much longer and variable: stages from L2 to L4 are on average 20-25 day long each.

2.2 Experimental animals

To investigate the expression pattern of genes that encode for the heterodimer EcR–RXR, we focused on the second larval stage (L1), because its length is less variable then other stages (3.5 days on average). From September 2009 to March 2010, 39 larvae were collected at four different points in time during the L1 stage to be scrutinized: 11 larvae immediately after the molt L0-L1 (group 0h), 9 after 24 hours (24h), 8 after 48 hours (48h) and 11 after 72 hours (72h).

2.3 RNA extraction and cDNA synthesis

The larvae were quickly frozen in N_2 and stored at -80 °C. Pools of frozen larvae (homogeneous for the time point) were transferred to a ceramic mortar and ground to powder in liquid nitrogen.

Total RNA was isolated using the SV Total RNA Isolation kit (Promega) according to the manufacturer's protocol, including a Dnase treatment. RNA was eluted using 50 µl RNase-free water and stored at -80 °C. The concentrations and purity of RNA were determined by NanoDrop ND-1000 spectrophotometer (*NanoDrop* Technologies). RNA samples used for the real time experiments met all of the following criteria: A260/A280 ratio >2.0, A260/A230 ratio >1.9.

To synthesize the first-strand cDNA, a mixture of total RNA (1 µg) and random primers (0.5 µg/reaction) in a volume of 10 µl was heated at 70 °C for 5 min and then chilled on ice for 5 min. To this mixture was added ImProm-II^{TM} 1× Reaction Buffer, $MgCl_2$ solution (2.25 mM), dNTP mix (0.5 mM each dNTP), RNasin Ribonuclease Inhibitor solution (10 U/µl), and ImProm-II Reverse Transcriptase (1 µl/reaction) to give a total volume of 10 µl. This mixture was

equilibrated at 25 °C for 5 min, extended at 42 °C for 60 min, and the enzyme was inactivated at 70 °C for 15 min. To minimize variations, all RNA samples were reverse-transcribed simultaneously. cDNA was stored at -20 °C until use.

2.4 Primer design

For both EcR and RXR, degenerate primers were designed on multiple alignment of the DNA-binding domain (DBD) and Ligand-binding domain (LBD) of different arthropod homologues, to obtain the correspondent cDNA fragments from *L. peregrinus*.

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 homologues 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).

To amplify a fragment of *Broad Complex* gene, initially we designed degenerated primers on the alignment of BTB sequence of different insect species: *Tribolium castaneum* (NM_001111264), *Apis mellifera* (NM_001040266), *Aedes aegypti* (AY499538), *Bombyx mori* (AB201230), *Blatella germanica*, *Manduca sexta* (AF032676), *Drosophila melanogaster* (X54665), *Acheta domesticus* (DQ176003) and *Oncopeltus fasciatus* (DQ176004). On the BTB fragment obtained, was designed a specific forward primer: BTBFor (5' GTGGATGTGACTGTAGCATGC 3'). It was used to amplify zinc-finger regions with degenerate reverse primers designed for each Z1-Z4 zinc-finger domains known in insects.

Primer information is listed in **Table 1**. All primers were synthesized by MWG Biotech (Ebersberg, Germany).

Table 1. Primer sequences of *EcR*, *RXR*, *Broad Complex* and *Elongation Factor 1-α*, PCR intermediate step of the amplified product. N means a mixture of A, T, G and C. In the same way, D (A, G, T), H (A, C, T), K (G, T), M (A, C), R (A, G), S (C, G), W (A, T) and Y (C, T) means a mixture of deoxynucleoside.

Gene	Forward primer sequence 5'-3'		Reverse primer sequence 5'-3'		PCR cycling
LpEcR	**EcRDFor2**	TCBGGNTACCACTAYAAYGC	**EcRLRev1**	TCTGARAADATKACDATGGC	35 cycles at 94 °C for 30 s, 54 °C for 30 s, 72 °C for 40 s
LpRXR	**RXRFor 1**	TCCAARCAYYTBTGYTCBATHTG	**RXRRev2**	GGNGTGTCGCCGATGAGTTTA	35 cycles at 94 °C for 30 s, 50 °C for 30 s, 72 °C for 40 s
LpBR-C	**BroadCore For**	TGCCTKCGNTGGAAYAAYTAYCA	**BroadCore Rev**	ACTTCRCCRTGGTAKATGAAYTC	35 cycles at 94 °C for 30 s, 55 °C for 30 s, 72 °C for 30 s
LpEF1-α	**ElFacFor**	GCTGGAATCTAGCCCCAAC	**ElFacRev**	CAATGTGAGCAGTGTGGCA	30 cycles at 94 °C for 30 s, 59 °C for 30 s, 72 °C for 30 s

2.5 Gene isolation and sequence analysis

To isolate genes of interest, 1 μl of cDNA (50 ng/μl) from a pool of different aged larvae was amplified using a standard PCR performed in 25 μl of reaction mix containing: 1X GoTaq® Reaction Buffer, 2 mM $MgCl_2$, 0.25 mM each dNTP, 0.4 μM primers, 1.25 units GoTaq® Polymerases (Promega). Between an initial denaturation step at 94 °C for 2 min and a final 5-min extension at 72 °C, the PCR conditions of the intermediate step changed according to primer pair (**Tab. 1**).

PCR products were separated by electrophoresis in a 1% agarose gel and visualized by staining with GelRed™. Predicted product sizes were verified with a 1 kb ladder of DNA markers (Promega) and the amplified cDNA fragments were cloned into the plasmid vector pGEM®-T Easy Vector (Promega), transforming *Escherichia coli* JM109 competent cells. Colonies were picked directly into a PCR mixture, and their inserts were then amplified with T7 forward and M13 reverse primers (MWG Biotech); inserts of the appropriate size were cleaned with ExoSAP-IT to remove an excess of primers and nucleotides (37 °C for 15 min, 80 °C for 15 min) and sequenced (BMR Genomics).

Sequence similarity search was performed using the program 'Blast' (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Sequence alignment and homology calculation were carried out with the program 'ClustalW' (http://www.ebi.ac.uk/Tools/clustalw/index.html) and edited in GeneDoc software version 2.7.000 (www.psc.edu/biomed/genedoc).

2.6 Semi-quantitative PCR

SQ-PCR was performed to determine *RXR* and *EcR* temporal mRNA expression during L1 stage. RNA was isolated from each of the four groups of larvae (0h, 24h, 48h, 72h). cDNA synthesis was performed using 0.5 μg total RNA and ImProm-II Reverse Transcriptase (Promega).

Primers listed in **Table 1** were used to amplify *LpRXR* and *LpEcR* cDNA fragments. *Elongation factor 1-α* expression was used as an internal PCR control, by amplification of the 502 bp fragment. All PCR reactions were carried out in a 12.5 μl reaction volume containing , 1X GoTaq® Reaction Buffer, 2 mM $MgCl_2$, 0.25 mM each dNTP, 0,4 μM primers, 1.25 units GoTaq® Polymerases (Promega).

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 0h and 24h larvae, 12.5 ng cDNA for 36h larvae and 6.25 ng cDNA for 72h larvae) were used to amplify both target and control gene.

The PCR cycle number was optimized performing parallel amplifications (n = 22, 26, 30, 32, 34 and 36). After analyzing expression results from different cycles, 35 PCR cycles were selected for receptors analysis, and 22 PCR cycles were used for *LpEF1-α* expression analysis.

The co-amplification of target and control gene in the same reaction tube was not possible because of the different annealing temperature of primers: 54 °C for *LpEcR*, 50 °C *LpRXR* and 59 °C for *LpEF1-α*. The cycling protocol was: one cycle at 94 °C for 2 min, 35/22 cycles at 94 °C for 30 s, 54/50/59 °C for 30 s, 72 °C for 30 s, and one cycle at 72 °C for 5 min for the final extension.

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. 1 kb ladder was used as molecular weight

marker and bands were stained with GelRed™ and visualized on Bio-Rad Gel Doc. The relative intensities of the amplified PCR products were determined using NIH ImageJ software 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.

2.7 Statistical analysis

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

3. Results

3.1 Cloning of *L. peregrinus* BTB domain

Using primers designed on the BTB region of *Broad Complex* in insects, we isolated a fragment of 258 bp that encodes a polypeptide of 86 amino acids. In Blast database this protein shows a high similarity (67-68% of amino acids identity) with BTB/POZ domains of two different insect gene: *Broad Complex* and *bric-à-brac*.

No fragments were obtained using a specific forward primer designed on the BTB domain with degenerate reverse primers designed for each of the Z1-Z4 zinc-finger domains known in *D. melanogaster.* Amplification of each single zinc-finger domain failed as well. The research did not proceed further because we hypothesized to have cloned the BTB domain of *bric à brac*.

3.2 Characterization of *L. peregrinus* EcR

Using an RT/PCR approach, we isolated two sequences, 1081 and 1033 bp long, encoding two polypeptides of 360 and 344 amino acids, respectively. A database search with Blast program indicated that both encode *L. peregrinus* orthologues of EcR protein. They have been called LpEcR_L (long form) and LpEcR_S (short form). These two proteins differ for an insertion/deletion (hereafter, insertion) of 16 amino acids in the 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). These cDNAs do not include the ligand-independent A/B activation domain and the poorly conserved carboxyterminal F domain (**Fig. 1**).

LpEcR_L and LpEcR_S amino acid sequences were compared with EcR sequences of other arthropods (see **Tab. 2**, **Fig. 1** and **Appendix**). The C domain of LpEcR shares a very high amino acid identity with that of other EcR sequences (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) retain less amino acids conserved with other arthropods. The E domain of LpEcR is also highly similar to those of other EcRs (>60%), especially EcRs from

Chelicerata (74-78%). Apart from the insertion present in LpEcr_L, the D domain is very similar to those of *L. australasiae*, *L. migratoria*, *D. magna* and *B. germanica* (56-57%), while similarity level is lower with respect to other arthropods (<48%).

Table 2. Identities of amino acid sequences of EcR orthologues versus LpEcR (%). Identities values for C, D and E region were calculated only against LpEcR_S. The GeneBank accession numbers of the sequences are listed in Appendix.

Species	**Identity against (%)**				
	LpEcR_L (total)	**LpEcR_S (total)**	**C region**	**D region**	**E region**
Liocheles australasiae	72	76	98	57	78
Ornithodoros moubata	68	71	96	46	74
Amblyomma americanum	65	68	96	39	74
Marsupenaeus japonicus	54	57	96	48	61
Celuca pugilator	66	69	98	43	71
Daphnia magna	65	68	94	56	76
Blattella germanica	69	73	98	55	72
Locusta migratoria	70	73	98	57	73
Leptinotarsa decemlineata	67	71	93	50	71
Tribolium castaneum	70	73	98	57	72
Pediculus humanus corporis	67	71	94	48	73
Apis mellifera	66	70	98	47	72
Bombyx mori	52	55	88	28	57
Aedes aegypti	56	59	88	42	60
Drosophila melanogaster	53	56	86	42	60

```
TCCGGATACCACTACAACGCCCTCACATGTGAAGGCTGCAAAGGTTTCTTT
S  G  Y  H  Y  N  A  L  T  C  E  G  C  K  G  F  F

AGGCGGAGCATTACAAAAAATGCTGTTTACCAGTGCAAATACGGAAACAAC
R  R  S  I  T  K  N  A  V  Y  Q  C  K  Y  G  N  N

TGTGAGATTGACATGTACATGCGCAGAAAGTGTCAGGAGTGCCGACTCAAG
C  E  I  D  M  Y  M  R  R  K  C  Q  E  C  R  L  K

AAGTGTCTCAACGTTGGGATGAGGCCAGAATTGACAAAGGTTATCCACGCT
K  C  L  N  V  G  M  R  P  E  L  T  K  V  I  H  A

GTCAGGATTCACAAAAAACCACAAATTTGTGTTGTCCCTGAATACCAATGT
V  R  I  H  K  K  P  Q  I  C  V  V  P  E  Y  Q  C

GCCGTGAAACGTGAGGCTAAAAGAGCTCAGAAGGAAAAAGACCGACCCAAC
A  V  K  R  E  A  K  R  A  Q  K  E  K  D  R  P  N

AGCACGACAAAAGAACATGGTTCACCACCGGACACACCTATCAAGGAAGAC
S  T  T  K  E  H  G  S  P  P  D  T  P  I  K  E  D

ATTTTGTTGCCCACTCCAAAAGAGAAAGAACCCATATCTAATAAGTTGACG
I  L  L  P  T  P  K  E  K  E  P  I  S  N  K  L  T

CCAGAAGAACAGGAATTAATCGAGAGGCTAGTCTATTTCCAAGAGGAGTTT
P  E  E  Q  E  L  I  E  R  L  V  Y  F  Q  E  E  F

GAATCTCCTTCAAATGAAGACCTCAAGAAAATCACGCACTTTCCCTTCACT
E  S  P  S  N  E  D  L  K  K  I  T  H  F  P  F  T

GGGGAAGATGGCGAGAGCCGCTTCAAGCACATCACCGAGATCACCATCCTC
G  E  D  G  E  S  R  F  K  H  I  T  E  I  T  I  L

ACTGTGCAACTCATAGTTGAGTTTTCAAAGCGCTTGCCTGGCTTTGATACC
T  V  Q  L  I  V  E  F  S  K  R  L  P  G  F  D  T

CTACAAAGGGAGGATCAAATCACACTGCTCAAGGGTTGTTCCAGTGAAGTG
L  Q  R  E  D  Q  I  T  L  L  K  G  C  S  S  E  V

ATGATGCTAAGGACTGCCAGGAGATATGACATAAACACTGATTCCATTGTA
M  M  L  R  T  A  R  R  Y  D  I  N  T  D  S  I  V

TTTGCCAATGACCAGCCGTACACACGAGAGAACTACAATAGTGCTGGCGTG
F  A  N  D  Q  P  Y  T  R  E  N  Y  N  S  A  G  V

GGAGACTCCGCTGACGACATCTTCCACTTCTGCCAGCAGATGTGCGTCATG
G  D  S  A  D  D  I  F  H  F  C  Q  Q  M  C  V  M

AAGGTGGATAATGCAGAATATGCACTCCTCACTGCCATTGTCATTTTCTCA
K  V  D  N  A  E  Y  A  L  L  T  A  I  V  I  F  S

GAACGGCCGCATTTGATTGAGCCGAAGAAGGTAGAGAAGATTCAAGAAATA
E  R  P  H  L  I  E  P  K  K  V  E  K  I  Q  E  I

TACCTTGAAGCACTCCGTAAATATGTGGAAAACCACCGCTCCAAATCCGTG
Y  L  E  A  L  R  K  Y  V  E  N  H  R  S  K  S  V

AACATGTTTGCGCGGCTGCTGTCCGTTTTAACAGAGCTACGTACCTTGGGT
N  M  F  A  R  L  L  S  V  L  T  E  L  R  T  L  G

AACCTGAATTCGGAGTTGTGCTTCTCTCTGAAGCTCAAGAACAAGAAGCTG
N  L  N  S  E  L  C  F  S  L  K  L  K  N  K  K  L

CCACCCTTCA
P  P  F
```

Fig. 1. Nucleotide and deduced amino acid sequence of LpEcR isoforms of *L. peregrinus*. The DNA-binding domain (DBD) is underlined and the ligand-binding domain (LBD) is underlined with dashes. Amino acid sequence of the insertion present in LpRXR_L isofrom is boxed.

3.3 Characterization of *L. peregrinus RXR*

We isolated three sequences, 975, 961, and 918 bp long, encoding three polypeptides of 325, 319, and 305 amino acids, respectively. A database search with the Blast program showed that these deduced sequences are highly homologous to other RXR/USP proteins, so they have been called LpRXR_L (long form), LpRXR_M (intermediate form), LpRXR_S (short form). These three proteins are identical except for two insertions/deletions (hereafter, insertions): a short sequence of 6 amino acids in the D domain (that differentiates LpRXR_L isoform from the other two) and a longer sequence of 14 amino acids in the E domain (that differentiates 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, 212 aa for LpRXR_S, 226 aa for LpRXR_M and LpRXR_M). These cDNAs do not include the ligand-independent A/B activation domain and poorly conserved carboxyterminal F domain (**Fig. 2**).

Table 3. Identities of amino acid sequences of Usp/RXR orthologues versus LpRXR (%).Identities values for C, D and E region were calculated only against LpRXR_S. For *A. americanum*, *B. germanica*, *L. migratoria* we used only the short isoforms of RXR (AmaRXR_S, BgRXR_S, LmRXR_S). The GeneBank accession numbers of the sequences are listed in Appendix.

Species	Identity against (%)					
	LpRXR_L (total)	LpRXR_M (total)	LpRXR_S (total)	C region	D region	E region
Lithobius forficatus	95	96	92	98	95	90
Liocheles australasiae	75	74	77	91	70	74
Ornithodoros moubata	73	75	74	94	86	67
Amblyomma americanum	71	72	75	92	81	72
Marsupenaeus japonicus	69	69	72	94	52	70
Celuca pugilator	68	70	68	92	63	63
Daphnia magna	73	75	78	94	72	74
Blattella germanica	76	77	81	97	81	76
Locusta migratoria	74	75	79	95	81	76
Leptinotarsa decemlineata	69	70	73	92	72	67
Tribolium castaneum	69	70	72	94	72	66
Pediculus humanus corporis	72	74	77	91	76	74
Apis mellifera	72	73	77	92	77	72
Bombyx mori	52	53	53	92	45	43
Aedes aegypti	54	56	40	94	64	44
Drosophila melanogaster	47	48	46	94	45	42

```
TTCCAAGCATTTTTGTTCTATTTGTGGGGATAGGGCCTCAGGGAAGCATTA
 S  K  H  F  C  S  I  C  G  D  R  A  S  G  K  H  Y
TGGCGTCTACAGTTGTGAAGGCTGTAAAGGTTTTTTCAAGAGGACGGTACG
 G  V  Y  S  C  E  G  C  K  G  F  F  K  R  T  V  R
CAAGGACTTGTCATATGCCTGTAGAGAAGATAGAAATTGTGTGATTGACAA
 K  D  L  S  Y  A  C  R  E  D  R  N  C  V  I  D  K
GCGCCAGCGAAACAGATGTCAGTACTGCCGTTATCAGAAATGTCTTTCAAT
 R  Q  R  N  R  C  Q  Y  C  R  Y  Q  K  C  L  S  M
GGGCATGAAACGAGAAGCTCTGTTTGTTTCCGCAGCTGTGCAAGAGGAGCG
 G  M  K  R  E  A  L  F  V  S  A  A  V  Q  E  E  R
TCAGCGTAATAAGGAAAAGAATGAAAATGAGGTGGAGAGCACGAGCAGTAT
 Q  R  N  K  E  K  N  E  N  E  V  E  S  T  S  S  I
CCACAATGACATGCCCCTTGAACGCATTCTTGAGGCGGAGTTAAGGGTGGA
 H  N  D  M  P  L  E  R  I  L  E  A  E  L  R  V  E
GCCAAAAGATGAGCAAATAGGAGACAAGCACAAACACCTACCATCCTTTGT
 P  K  D  E  Q  I  G  D  K  H  K  H  L  P  S  F  V
CATGCCATTTTCTGCTGTCCAGAATACAATGACAAATTTCTGCCAAGCGGC
 M  P  F  S  A  V  Q  N  T  M  T  N  F  C  Q  A  A
CAATCAGCAGCTGATCCAGTTGGTGGAATGGGCCAAGCACATACCCCACTT
 N  Q  Q  L  I  Q  L  V  E  W  A  K  H  I  P  H  F
CACAGAACTGCCTATTGAAGATCAGGTGACCCTCTTGCGAGCAGGTTGGAA
 T  E  L  P  I  E  D  Q  V  T  L  L  R  A  G  W  N
CGAGTTGCTCATAGCAGCATTTTCTCACCGTTCAGTTAATGTCAAAGACAG
 E  L  L  I  A  A  F  S  H  R  S  V  N  V  K  D  S
CATTGTATTGGCAACGGGTCTTCAGATTCATCGGACTGACGCACACAGTGC
 I  V  L  A  T  G  L  Q  I  H  R  T  D  A  H  S  A
GGGAGTAGGCACAATTTTTGATCGAGTCCTAACAGAACTAGTAGCTAAAAT
 G  V  G  T  I  F  D  R  V  L  T  E  L  V  A  K  M
GAGGGAAATGAAGATGGATCGCACAGAACTTGGGTGCCTTAGGGCGATAAT
 R  E  M  K  M  D  R  T  E  L  G  C  L  R  A  I  I
CCTTTTCAATCCCGTTAAAGGTCTGCGATCCAGCCAAGTTATTGAGTCTCT
 L  F  N  P  V  K  G  L  R  S  S  Q  V  I  E  S  L
AAGGGAACGAGTATATGCCGCCCTAGAAGAATATTGCAAGCAACAGTATCC
 R  E  R  V  Y  A  A  L  E  E  Y  C  K  Q  Q  Y  P
CGATGAACCTGGCCGTTTTGCCAAGCTCTTGCTGCGTCTACCTGCCTTGAG
 D  E  P  G  R  F  A  K  L  L  L  R  L  P  A  L  R
AAGCATTGGACTTAAGTGTCTGGAGCACCTCTTCTTTTTAAACTCATCGG
 S  I  G  L  K  C  L  E  H  L  F  F  F  K  L  I  G
CGACACCCCA
 D  T  P
```

Fig. 2. Nucleotide and deduced amino acid sequence of LpRXR isoforms of *L. peregrinus*. The DNA-binding domain (DBD) is underlined and the ligand-binding domain (LBD) is underlined with dashes. Amino acid sequences of the short insertion (present in the LpRXR_L isofrom) and long insertion (present in the LpRXR L and RXR_M isoforms) are boxed.

LpRXR_L, LpRXR_M and LpRXR_S amino acid sequences were compared with RXR/USP sequences of other arthropods, (see **Tab. 3**, **Fig. 2** and **Appendix**). LpRXR_M sequence is very similar to *L. forficatus LfRXR* orthologues and both of them have the same insertion in LBD domain.

Similarly to DBD of LpEcR, amino acid identity of the C region of LpRXR is also very high among all sequences (91-97%). The P-box of LpRXR (EGCKG) is 100% identical to that of other RXR/USP, whereas the D-box (CREDR) retain less amino acids conserved with other arthropods as well as LpEcR. Apart from for the insertion present in LpRXR_L, D domain, is highly homologous to those of *O. moubata*, *A. americanum*, *B. germanica* and *L. migratoria* (81-86%), although they are less homologous to other USPs. The short insertion presents in LpRXR_L shares several amino acids with *L. australasiae* and *M. japonicus* D domain sequences (**Fig. 3A**).

Domain E, with the 14 amino acids insertion (in LpRXR_L and LpRXR_M), presents more similarity with LBD domain of other arthropods than the shorter isoform. Interestingly, the loop between helices H1 and H3 in LBD domain, where the longer insertion is located, is quite divergent in arthropods, but the insertion of *L. peregrinus* has high similarity to sequences of Chelicerata and Crustacea (**Fig. 3B**).

Fig.3. Comparison of the D and E domain sequences of LpRXR_S, LpRXR_M and LpRXR_L isoforms of *L. peregrinus* with other species. **A**) A portion of the D domain with the short insertion present in LpRXR_L is aligned with the homologue region of *M. japonicus* (Mj_RXR, Crustacea) and *L. australasiae* (La_RXR, Chelicerata). **B**) In the LBD domain (E) the region from helices H1 and H3 is aligned with the homologous region of *O. moubata* (Om_RXR, Chelicerata) and *C. pugilator* (Cp_RXR, Crustacea). Conserved sequences are shown in boxes. Regions corresponding to helices H1 and H3 are underlined with dashed line.

3.4 Gene expression profiles during the second larval stage (L1)

Under our rearing conditions (See Materials and Methods), the duration of the second larval stage (L1) in *L. peregrinus* is approximately 3.5 days. *EcR* and *RXR* gene expression were determined at four time points during this stage. The

housekeeping gene *Elongation factor 1-α* was used to normalize expression data of the two receptor genes.

Because of the small sequence differences, it is impossible to discriminate between the different isoforms of *LpEcR* and *LpRXR* in an agarose gel. Therefore, the intensity bands that we analyzed for the two genes represent the summation of the expression level in all isoforms of the same gene (**Fig. 4**).

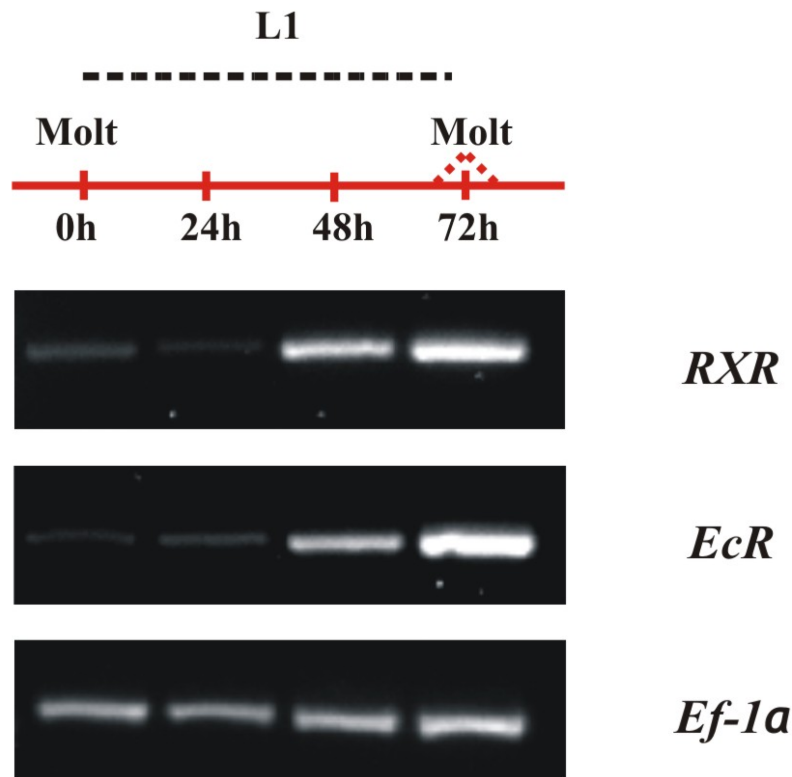

Fig.4. Expression patterns of *LpRXR* and *LpEcR* mRNA in second stage larvae (L1) of *L. peregrinus*. cDNA fragments of *LpEcR* and *LpRXR* were amplified from four groups of larvae collected immediately after the moult L0-L1 (group 0h), after 24 hours (24h), after 48 hours (48h) and after 72 hours (72h). These were separated by electrophoresis in a 1% agarose gel and visualized by staining with GelRed. *Elongation factor 1-α* levels were used as a reference.

The effects of experiment replicates and specimens grouping (0h-72h) on *LpEcr* and *LpRXR* relative expression level were tested with a two-way ANOVA. For both genes, expression levels in the four groups differ significantly ($p<0.0005$), whereas there are no significant differences between the three replicates of the experiment ($p>0.51$).

Concentration of *LpEcR* mRNA is low for the first 24 hours of the L1 stage, to significantly increase about 48 hours after the molt, with a significant peak of expression at 72 hours (Fischer's test LSD, $\alpha=0.01$; **Fig. 5A**).

LpRXR shows an expression profile very similar to *LpEcr*, with a significant peak of expression at 48 hours, that is then maintained almost unchanged until 72 hours after the molt (Fischer's test LSD, $\alpha=0.01$; **Fig. 5B**).

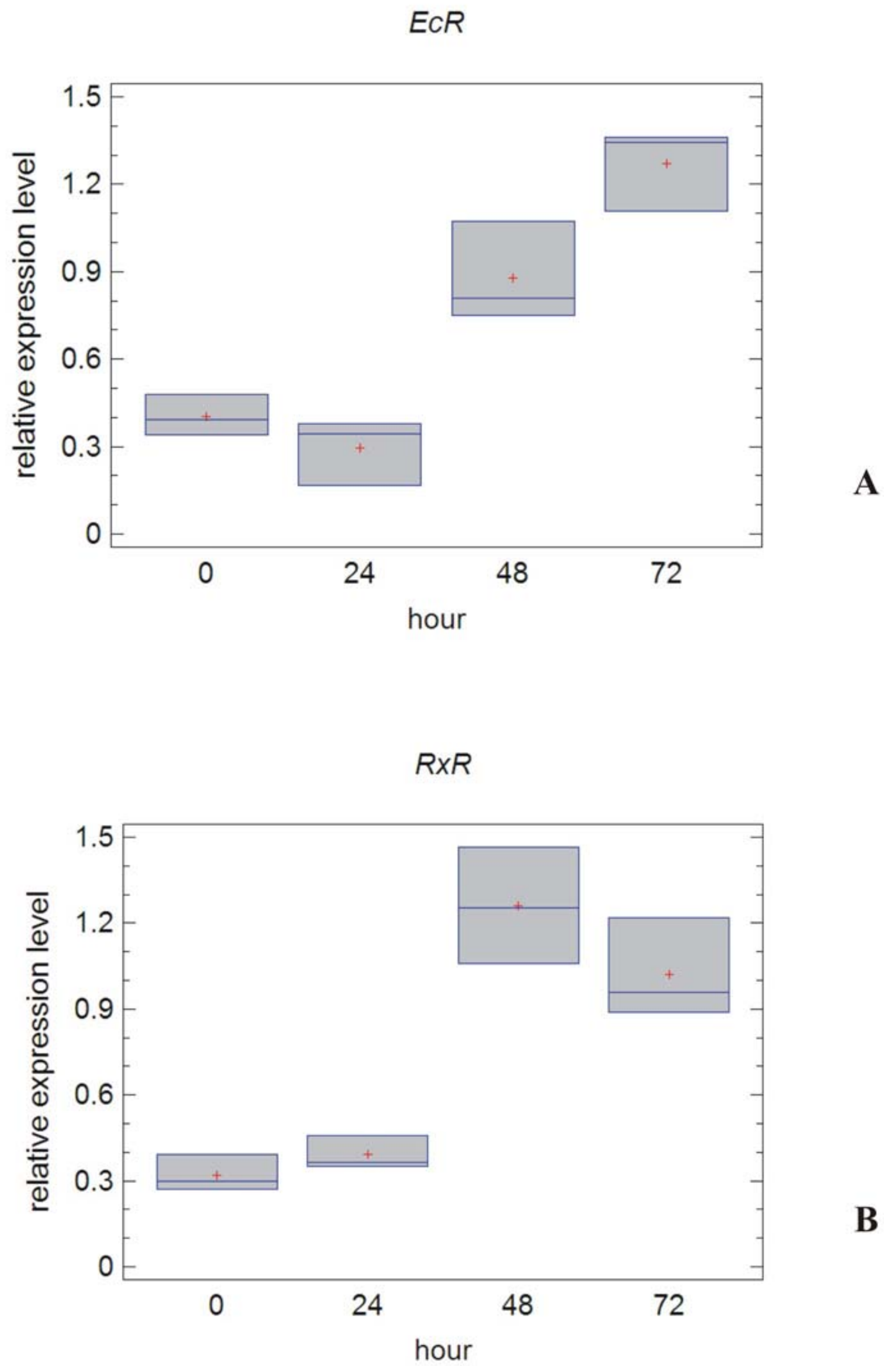

Fig 5. Relative expression level of *LpEcR* (**A**) *LpRXR* (**B**) during L1 stage of *L. peregrinus*. Data were normalized with respect to the expression level of *EF1-α*. Boxes represent the interval between lower and upper quartiles, with median (transverse line) and mean (small cross).

4. Discussion

Sequence comparison of the receptors LpEcR and LpRXR isolated in *L. peregrinus* with their orthologues 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 orthologues (40-55%). Comparative analyses have shown that EcR and RXR/USP from Diptera and Lepidoptera have co-evolved during the course of holometabolous insect evolutionary radiation, possibly leading to a functional divergence of the receptors (Bonneton et al., 2003). 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 (Hayward et al., 1999; Riddiford et al., 2001).

The C domain sequence (DBD) in both receptors is highly conserved across several species. In this region there are two zinc-finger domains containing, respectively, a proximal (P)-box and a distal (D)-box sequence that provide DNA-binding specificity (Umesono and Evans, 1989).

Aminoacids of the P-box, critical for DNA response element recognition, are identical in all EcR DBD, suggesting that they recognize similar response elements. D-box region forms a dimerization interface in several nuclear receptor (Umesono and Evans, 1989), and the substitution clustered in these sequence could reflect functional differences in protein-protein interaction among different receptors. For example, there is evidence that *Drosophila* EcR may bind nuclear receptor other than UPS (White et al., 1997).

A distinctive character of *L. peregrinus* EcR and RXR is the presence of several variants: we identified two isoforms for EcR and three variants for the RXR receptor. The sequences of EcR proteins, namely LpEcR_S and LpEcR_L, only differ by a 16 aa segment in the domain D. Two portions of this region are conserved among all arthropods: a T-box and an A-box motif that play a role in DNA recognition (Devarakonda et al., 2003). The insertion present in LpEcR_L is located inside the T-box, suggesting the possibility of a different response element recognition. The D region is also essential for a ligand-dependent heterodimerization with RXR. Some crustaceans species show multiple variants in the D domain. For instance, *C. pugilator* has four substutive variants (Chung et al., 1998), and *M. japonicus* has two. In *M. japonicus* (Asazuma et al., 2007) the longer variant is expressed more than shorter one in all tissue examined, but their different function is unknown. Further studies are required to ascertain whether these multiple variants of LpEcR have different properties in DNA binding and heterodimerization with respect to LpRXR.

In arthropods, the sequences of distinct variants in the A/B domain are produced by alternative splicing, and the expression of these variants is regulated by distinct promoters. In *Drosophila*, the different isoforms are expressed in tissue-specific and developmental stage-specific manner (Talbot et al., 1993). A complete full length cDNA sequence for LpEcR would allow to identify the A/B domain and possibly new isoforms of this gene.

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 D domain and the other one is in the LBD. As reported for LpEcR, the short insertion present in the D domain of LpRXR_L is located inside the T-box region. Interestingly, this fragment is almost identical to that found in *L. australasiae* and quite similar to that in *M. japonicus*. This region has a fundamental role in mediating hormone response element binding interactions with RXR homodimers or with heterodimers formed by RXR and other nuclear receptor (Zhao et al., 2000).

The 14 aminoacids insertion in LpRXR_L and LpRXR_M is located in the loop connecting helices H1 and H3 within the LBD, as in *L. forficatus* RXR. In invertebrates, RXR/USP isoforms differing for insertions/deletions of this type have been previously reported for *L. migratoria* (Hayward et al., 1999, 2003), *B. germanica* (Maestro et al., 2005) and the crab *C. pugilator* (Durica et al., 2002).

Sequence variation in this region could influence transactivation properties or ligand affinities, but further studies are necessary to clarify the function of the different isoforms.

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 start 48 hours after the previous molt. These results confirm some expression data obtained in other arthropods, as *C. pugilator,* where thoracic muscles show high level of EcR and RXR during a premolt phase (Chung et al., 1998), although both receptor genes can exhibit dissimilar expression profiles in other tissues.

In insects, expression profile of most of the ecdysteroid-regulated genes in the whole body directly correlates with the ecdysteroid titer (Sullivan and Thummel, 2003). However, the expression pattern of EcR and USP/RXR do not coincide with all peaks of ecdysteroid titer and these differences depend on developmental stage and tissue examined. These imply that the expression of these genes is not controlled by ecdysteroid exclusively.

The attempt to isolated *Broad Complex* in the centipede failed. To date, this gene has been studied only in insects, and the apparent lack of homologues in *L. peregrinus* and crustaceans (B. Konopova, personal communication) suggests that it could be an apomorphic trait of insects.

Results obtained in this study represent the first data on the genes involved in the ecdysone signal pathway in a myriapod species. Identification of target genes of LpEcR and LpRXR may help to understand their function in *Lithobius*, and further comparative study with EcR and RXR/USP homologues in other species, both within and outside myriapoda, would provide a more comprehensive understanding of the developmental processes mediated by these genes in arthropods.

References

Asazuma, H., Nagata, S., Kono, M. and Nagasawa, H. (2007). Molecular cloning and expression analysis of ecdysone receptor and retinoid X receptor from the kuruma prawn, *Marsupenaeus japonicus*. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, **148:**139-150.

Bayer, C. A., Holley, B. and Fristrom, J. W. (1996). A switch in *Broad-Complex* zinc-finger isoform expression is regulated post-transcriptionally during the metamorphosis of *Drosophila* imaginal discs. Developmental Biology, **177**, 1-14.

Billas, I.M.I., Browning, C., Lawrence, M.C., Graham, L.D., Moras, D., Hill, R.J. (2009). The structure and function of ecdysone receptors. In Ecdysone: Structures and Functions (Smagghe, G., ed.). Springer, Dordrecht, 335–360.

Bonneton, F., Zelus, D., Iwema, T., Robinson-Rechavi, M., Laudet, V. (2003). Rapid divergence of the ecdysone receptor in Diptera and Lepidoptera suggests coevolution between ECR and USP-RXR. Molecular Biology and Evolution, **20:** 541–553**.**

Chung, A.C., Durica, D.S., Clifton, S.W., Roe, B.A., Hopkins, P.M. (1998). Cloning of crustacean EcR and RXR gene homologs and elevation of RXR mRNA by retinoic acid. Molecular and Cellular Endocrinology, **139**:209–227.

Devarakonda, S., Harp, J.M., Kim, Y., Ozyhar, A., Rastinejad, F. (2003). Structure of the heterodimeric ecdysone receptor DNA-binding complex. EMBO Journal, **22**:5827-5840.

DiBello, P. R., Withers, D. A., Bayer, C. A., Fristrom, J. W. and Guild, G. M. (1991). The *Drosophila Broad-Complex* encodes a family of related proteins containing zinc fingers. *Genetics* **129,** 385-397.

Durica, D.S., Wu, X., Anilkumar, G., Hopkins, P.M., Chung, A.C. (2002). Characterization of crab EcR and RXR homologs and expression during limb regeneration and oocyte maturation. Molecular and Cellular Endocrinology, **189**:59–76.

Evans, R. M. (1988). The steroid and thyroid hormone receptor superfamily. Science, **240**:889-895.

Hayward D.C., Bastiani M.J., Trueman J.W.H., Truman J.W., Riddiford L.M., Ball E.E. (1999) .The sequence of locusta RXR, homologous to *Drosophila* Ultraspiracle, and its evolutionary implications. Development Genes and Evolution, **209**:564-571.

Hayward, D. C., Dhadialla, T. S., Zhou, S., Kuiper, M. J., Ball, E. E., Wyatt, G. R., Walker, V. K. (2003). Ligand specificity and developmental expression of RXR and ecdysone receptor in the migratory locust*.* Journal of Insect Physiology*,* **49:** 1135-1144.

Joly. R. (1964). Action de l'ecdysone sur le cycle de la mue de *Lithobius forficatus*. L . (Myriapode, Chilopode). Comptes Rendus des Seances de la Societe de Biologie et des ses filiales ,**158**:548-550.

Joly. R., and Descamps, M. (1988) Endocrinology of Myriapods. In Endocrinology of selected invertebrate types (Laufer H. ed.). New York: Alan R. Liss. lnc., 430-449.

Karim, F. D., Guild, G. M., Thummel, C. S. (1993). The *Drosophila Broad-Complex* plays a key role in controlling ecdysone-regulated gene expression at the onset of metamorphosis. Development, **118**:977-988.

Kostrouch, Z., Kostrouchova, M., Love, W., Jannini ,E., Piatigorsky, J., Rall, J. E. (1998). Retinoic acid X receptorin the diploblast, *Tripedalia cystophora*. Proceedings of the National Academy of Science of the United States, **95**:13442–13447.

Maestro, O., Cruz, J., Pascual, N., Martin, D., Bellés, X. (2005). Differential expression of two RXR/ultraspiracle isoforms during the life cycle of the hemimetabolous insect *Blattella germanica* (Dictyoptera, Blatellidae). Molecular and Cellular Endocrinology, **238**:27-37.

Nakagawa, Y., Sakai, A., Magata, F., Ogura, T., Miyashita, M.,Miyagawa, H. (2007). Molecular cloning of the ecdysone receptor and the retinoid X receptor from the scorpion *Liocheles australasiae*. FEBS Journal, **274**:6191–6203.

Riddiford, L. (2001). Ecdysone Receptors and their Biological Actions. Vitamins and hormones**, 60**:1-73.

Seifert G. and Bidmon H.J. (1988). Immunohistochemical evidence for ecdysteroid-like material in the putative molting glands of *Lithobius forficatus* (Chilopoda). Cell and Tissue Research, **253**:263-266.

Shea, C., Richer, J., Tzertzinis, G. and Maina, C. V. (2010). An EcR homolog from the filarial parasite, *Dirofilaria immitis* requires a ligand-activated partner for transactivation. Molecuar and Biochemical Parasitology, **171**: 55-63.

Sullivan, A.A., Thummel, C.S. (2003). Temporal profile of nuclear receptor gene expression reveal coordinate transcriptional responses during *Drosophila* development. Molecular Endocrinology, **17**:2125-2137.

Talbot, W.S., Swyryd, E.A., Hogness, D.S. (1993). Drosophila tissues with different metamorphic responses to ecdysone express different ecdysone receptor isoforms. Cell **73:**1323–1337.

Thummel, C.S. (1995). From embryogenesis to metamorphosis: the regulation and function of *Drosophila* nuclear receptor super family members. Cell, **83**:871-877.

Tzertzinis, G., Egana, A.L., Palli, S.R., Robinson-Rechavi, M., Gissendanner, C.R., Liu, C., Unnasch, T.R. and Maina, C.V. (2010). Molecular evidence for a functional ecdysone signaling system in Brugia malayi. PLoS Neglected Tropical Diseases , 4(3).

Umesono, K., and Evans, R.M. (1989). Determinants of target gene specificity for steroid / thyroid hormone receptors. Cell, **57:**1139–1146.

White, K.P., Hurban ,P., Watanabe, T., Hogness, D.S. (1997). Coordination of *Drosophila* metamorphosis by two ecdysone-induced nuclear receptors. Science, **276**: 114-117.

Yao, T. P. (1993). Functional ecdysone receptor is the product of *EcR* and *Ultraspiracle* genes. Nature, **366**:476-479.

Zhao, Q., Chasse, S.A., Devarakonda, S., Sierk, M.L., Ahvazi, B.,Rastinejad, F. (2000). Structural basis of RXR–DNA interactions. Journal of Molecular Biology, **296**:509–520.

Appendix I

Fig.1. Alignment of amino acid sequences of LpEcR and EcRs orthologues from species representative of the main groups of Insects, Crustacea and Chelicerata. *Drosophila melanogaster* (Drm; Genbank Accession Number NP_724456), *Leptinotarsa decemlineata* (Ld; BAD99296), *Pediculus humanus corporis* (Phc; XP_002430228), *Apis mellifera* (Am; BAF46356), *Bombyx mori* (Bm; 001037331), *Tribolium castaneum* (Tc; NP_001107650), *Aedes aegypti* (Aa; AAA87394), *Blattella germanica* (Bg; CAJ01677), *Locusta migratoria* (Lm; AAD19828), *Daphnia magna* (Dam; BAF49029), *Celuca pugilator* (Cp; AAC33432), *Marsupenaeus japonicus* (Mj; BAF75375), *Ornithodoros moubata* (Om; BAE45855), *Liocheles australasiae* (La; BAF85822), *Amblyomma americanum* (Ama; AAB94566).
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 underlined with a black bar, an insertion in LpEcR_L is underlined with a red bar.

DBD

LBD

LBD

LBD

Fig.2. Alignment of amino acid sequences of LpRXR and Usps/RXRs orthologues from species representative of the main groups of Insects, Crustacea, Chelicerata and the only sequence available to date for Myriapoda: *Drosophila melanogaster* (Drm; NP_476781), *Leptinotarsa decemlineata* (Ld; BAD99298), *Pediculus humanus corporis* (Phc; XP_002424949), *Apis mellifera* (Am; NP_001011634), *Bombyx mori* (Bm; NP_001037470), *Tribolium castaneum* (Tc; NP_001107650), *Aedes aegypti* (Aa; AAG24886), *Blattella germanica* (BgRXR_L CAH69898; BgRXR_S CAH69897), *Locusta migratoria* (LmRXR_L AAQ55293; LmRXR_S AAF00981), *Daphnia magna* (Dam; ABF74729), *Celuca pugilator* (Cp; AAC32789), *Marsupenaeus japonicus* (Mj; BAF75376), *Ornithodoros moubata* (Om; BAF91724), *Liocheles australasiae* (La; BAF85823), *Amblyomma americanum* (AmaRXR_L AAC15589; AmaRXR_S AAC15588), *Lithobius forficatus* (Lf; AAO18151).
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 underlined with a black bar, short and long insertions in LpEcR isoforms are underlined with a red bar.

DBD

LBD

LBD

LBD

LBD

Manuscript 2

Post-embryonic expression of two segmentation genes, *engrailed* and *wingless*, during the anamorphosis of *Lithobius peregrinus*

Post-embryonic expression of two segmentation genes, *engrailed* and *wingless*, during the anamorphosis of *Lithobius peregrinus*

Abstract

A segmental organization of the body plan is a defining characteristic of the arthropods, however the progressive increase in the number of trunk segments during post-embryonic development (anamorphosis), typical of several arthropod taxa, remains mostly unknown. In myriapods, knowledge about the developmental genetics of segmentation is limited to early embryonic development, whereas the genetic control of this process during anamorphosis has never been investigated. Using Real Time PCR and in situ hybridization on paraffin sections, we studied expression patterns of two segment polarity genes, *engrailed* and *wingless*, during the anamorphic stages of the centipede *Lithobius peregrinus*. Only en seems retain a segment-polarity role in post-embryonic development, whereas wg function seems related to the development of nervous system and leg buds, suggesting that it does not retain the same role in segmental patterning that it has during embryonic development. The results of this work are the first molecular data on post-embryonic development in a myriapod species, and thus represent a basic starting point for the study of anamorphosis, that is probably the primitive segmentation mode for arthropods.

1. Introduction

Arthropods are characterized by an obvious segmental body pattern. In most insects, in chelicerates and in some myriapods segmentation is completed before hatching, but in many other arthropods, among which the lithobiomorph centipedes, only a fraction of the segments is completed during embryogenesis. In these species, the final adult number of segments (when fixed) is reached later in ontogeny, as segments emerge sequentially, from anterior to posterior, through a series of molts. This post-embryonic developmental mode is called anamorphosis.

Several lines of evidence suggest that anamorphosis is the primitive segmentation mode for arthropods (see Fusco, 2005). However, beyond descriptive morphology, very little is known about this primitive developmental process. We lack significant information about the expression of segmentation genes during post-embryonic development and about the details of morphogenetic processes involved in anamorphosis. Post-embryonic segmentation is traditionally considered to be the product of the morphogenetic activity of an inadequately defined 'sub-terminal generative zone'.

The genetic basis of the developmental mechanisms of body segmentation is well known in the fruit fly *Drosophila melanogaster*. So-called 'segmentation genes' discovered in the fly are classified into gap, pair-rule and segment-polarity genes. These genes occupy different levels in a hierarchical cascade leading from the early gap genes to the later-expressed pair-rule and segment-polarity genes. The latter, like *engrailed*, *wingless*, *hedgehog* and *cubitus interruptus*, encode proteins that are expressed in the embryo as stripes with segmental periodicity, and are required for the formation of the correct pattern of structures within each segment (review in Rivera-Pomar and Jäckle, 1996). The mechanism of embryonic segmentation is extremely derived in the fly, and cannot be easily compared to that of most other arthropods.

Homologs of several segment-polarity and pair-rule genes have been studied in other insects as well as in chelicerates, myriapods, and crustaceans (Patel, 1994; Abzhanov and Kaufman 2000; Damen et al., 2000; Niwa et al., 2000; Damen, 2002; Hughes and Kaufman, 2002; Kettle et al., 2003). The available data suggest that at the level of the segment-polarity genes the segmentation process is very conserved among extant arthropods (Damen, 2002), whereas at the level of the pair-rule genes there seems to be more diversity and the role of maternal effect gene and gap gene homologs in non-insect arthropods is unclear.

Genetic regulation of segmentation presents high diversity among arthropods, but the role of the segment-polarity gene *engrailed* (*en*) and *wingless* (*wg*) in the establishment of a pattern within segments is on the whole highly conserved. In *Drosophila*, along the AP axis of the germ band, *en* is expressed in segmental bands in the anterior domain of each parasegment, a position corresponding to the prospective posterior margin of the segments, while *wg* is expressed in cell stripes adjacent and anterior to the bands expressing *en* (Sanson, 2001).

En, like several other developmental regulatory genes, encodes a homeodomain transcription factor (Desplan et al., 1985; Fjose et al., 1985) involved in specifying cell fate at segmental boundaries (Lawrence and Morata, 1976), but also in imaginal disc and wing development (DiNardo et al., 1985; Hidalgo, 1994), hindgut formation (Takashima et al., 2002) and neurogenesis (Siegler & Jia, 1999). Studies on crustacean larvae (review in Dohle et al. 2004) have shown segmental *en* expression, but any specificremark was made about the role of this gene during post embryonic development.

Wg is a member of the highly conserved wnt gene family of small signaling factors, characterized by a string of conserved cysteine residues (Rijsewijk et al., 1987). Its function is required throughout development in a wide range of patterning events at different times and in different tissues. In *Drosophila* early germ band, *wg* input consolidates parasegmental boundaries by maintaining *en* expression in adjacent epidermal cells (DiNardo et al., 1985; Martinez-Arias et al., 1988). Later in embryogenesis, *wg* input is no longer needed for *en* maintenance, becoming rather involved in specifying cell fate. It is implicated in the development of legs and nervous systems, but also of embryonic epidermis, head, midgut, heart, muscles and malpighian tubules (reviewed by Cadigan and Nusse, 1996).

Studies on developmental genetics of segmentation in Myriapoda are mostly limited to the germ band stage, during the early phases of embryonic development. In *Lithobius*, only two relatively early segmentation genes with segment-polarity function (*en* and *wg*) and a pair-rule gene (*even-skipped*) have been studied (Hughes and Kaufman, 2002), together with a few other later expressed genes (*ASH*, *Delta* and *Notch*) involved in segmental neurogenesis (Stollewerk and Kadner, 2004).

Thus, our work represents the first research on the expression of genes involved in segment formation during the anamorphic phase in a myriapod. The experimental protocols used for studying gene expression in this developmental period are innovative as well.

As all lithobiomorphs, the post-embryonic development of *L. peregrinus* is hemianamorphic. This centipede hatches with an incomplete number of trunk segments to reach the final segmental arrangement with fifteen leg-bearing segments within five developmental stages, following a precise schedule of per-

stage segment addition (anamorphic phase of post-embryonic development). Subsequently, the animal continues to growth and molt without any further variation to its segmental organization (epimorphic phase of post-embryonic development).

As a primary exploration of genetics of the process of myriapod anamorphic segmentation, we investigated the expression of two segment-polarity genes, *en* and *wg*, during the post embryonic development of *L. peregrinus* using Real Time PCR and in situ hybridization on paraffin sections, two techniques never applied in myriapods before. We found that *en* retains a role in segmentation also during anamorphosis, whereas *wg* seems to be involved in developmental processes less directly related to segmental patterning.

2. Materials and method

2.1 Experimental animals

Eggs were obtained from adult centipedes collected in north-eastern Italy and reared in laboratory (for details, see manuscript 1). After hatching, larvae were isolated and kept individually until the selected stage.

For Real Time PCR experiments larvae were killed the day of molting and stored at -80 °C. Specimens of seven post-embryonic stages were collected: larvae from first (L0) to fifth (L4) stage, juveniles of the first post-larval stage (PL1) and adults of mixed age (**Tab. 1**). 60 embryos were also used to validate the protocols.

Larvae of stages L1 and L2, with a thinner cuticle, more permeable to fixative, were used for *in situ* hybridization.

Table 1. Numbers of specimens used in real time PCR assays.

Stage	Number of specimens
L0	36
L1	35
L2	24
L3	20
L4	11
PL1	7
Adult	2

2.2 RNA extraction and cDNA synthesis

Total RNA was isolated from larvae using the SV Total RNA Isolation kit (Promega, Madison, WI), according to the manufacturer's instructions. The RNA was treated with DNase I and cDNA synthesis was performed using 1 µg total RNA and ImProm-II Reverse Transcriptase (Promega). cDNA was stored at -20 °C until use.

2.3 Primer design

Specific primers for reference and target genes were designed to conserved regions on the basis of comparison of sequences from several arthropods, obtained from the GenBank database. Primer and amplicon lengths are listed in **Tab. 2**. All primers were synthesized by MWG Biotech (Ebersberg, Germany).

The only primer pair that worked for engrailed was designed on the sequence of *Lithobius atkinsoni* (GenBank accession number AF434998), whereas a portion of *wingless* gene was isolated using degenerate primers based on the *wg* sequence alignment of *Lithobius atkinsoni* (AF435006), *Cupiennius salei* (AJ315945), *Thermobia domestica* (AF214035) and *Gryllus bimaculatus* (AB044713).

For real time PCR experiments, two reference genes were selected: *actin* (*ACT*) and *elongation factor1-α* (*EF1-α*). *Actin* was isolated using a pair of primers based on the sequences of other arthropod species: *Plutella xylostella* (AB282645), *Manduca sexta* (L13764), *Ornithodoros moubata* (AY547732), *Homarus americanus* (AF399872), *Ceratitis capitata* (M76614). The amplicon sequence obtained were used to design specific primers. Specific primers (ActLp) were designed on the amplicon sequence obtained. For *EF1-α,* ElFacFor and ElFacRev primers were used, already tested on this species (see manuscript 1). A fragment of *tropomyosin* (*Tm*) was isolated as control for validation and optimization of *in situ* hybridization conditions. Degenerate primers were designed manually aligning nucleotide sequences of *tropomyosin* in *Tyrophagus putrescentiae* (AY623832), *Boophilus microplus* (AF124514), *Scolopendra* sp. (AY421743), *Jasus lalandii* (FJ169628), *Blattella germanica* (AF260897) and *Euphausia superba* (AB289603). Primers specific for *L. peregrinus* (TropomLp) were used to obtain the probe.

Table 2. Primer sequences of *engrailed*, *wingless* and three housekeeping genes, the amplification length and the melting temperature of the amplified product. N means a mixture of A, T, G and C. In the same way, D (A, G, T), H (A, C, T), K (G, T), M (A, C), R (A, G), S (C, G), W (A, T) and Y (C, T) means a mixture of deoxynucleoside.

Gene	Forward primer sequence 5'-3'		Reverse primer sequence 5'-3'		Length (bp)
LpEng	**EngFor4**	TATTCCGACCGGCCATCGT	**EngRev4**	CAATTCCCGAGCCAGATCCT	222
LpWg	**WgFor 2**	TGCACNTTCAAGACGTGCTGG	**WgRev3**	ACCCAACGAAGAATCGTTG C	290
LpEF1- α	**ElFacFor**	GCTGGAATCTAGCCCCAAC	**ElFacRev**	CAATGTGAGCAGTGTGGCA	504
LpACT	**ActFor**	AACTGGGATGACATGGAGAAG	**ActRev**	GGGTACATGGTGGTACC	689
	ActLpFor	TACAATGAGCTGCGTGTCG	**ActLpRev**	ATGGAGTTGAAGGTGGTCT	575
LpTm	**Tropom For3**	TGCAGGCSATGAAGYTGGA	**Tropm Rev2**	ACCTCCTTCTGGAGCTTCTG	736
	TropomLp For	TAAGGCCGAGGAGGAGGTTC	**TropmLp Rev**	TAGAGACTTCAGGTTGTTGC	511

2.4 Gene isolation and sequence analysis

Since no information is available about temporal expression for the genes of interest, different aged larvae were pooled for cDNA amplification. PCR cycling were performed in a Eppendorf thermocycler using 25 µl of reaction mix containing: 1X GoTaq® Reaction Buffer, 2 mM $MgCl_2$, 0,25 mM each dNTP, 0.4 µM primers, 1.25 units GoTaq® Polymerases (Promega). Between an initial denaturation step at 94°C for 2 min and a final 5-min extension at 72 °C, the PCR conditions of the intermediate step changed as follows, according to primer pair:

- ActFor/Rev: 30 cycles at 94 °C for 30 s, 56 °C for 30 s, and 72 °C for 40 s;
- ActLpFor/Rev: 30 cycles at 94 °C for 30 s, 59 °C for 30 s, and 72 °C for 40 s;
- TropomFor/Rev: 35 cycles at 94 °C for 30 s, 54 °C for 30 s, and 72 °C for 40 s;
- TropomLpFor/Rev: 35 cycles at 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 40 s;
- EngFor4/Rev4: 32 cycles at 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s;
- WgFor2/Rev3: 32 cycles at 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s.

PCR products were purified by agarose gel electrophoresis, cloned into the pGEM®-T Easy Vector (Promega) and sequenced (BMR Genomics.) In order to avoid cloning or PCR artifacts, several clones were screened for each fragment obtained.

Sequence similarity search was performed using the program 'Blast' (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Sequence alignment and homology calculation were carried out with the program 'ClustalW' (http://www.ebi.ac.uk/Tools/clustalw/index.html) and edited in GeneDoc software version 2.7.000 (www.psc.edu/biomed/genedoc).

2.5 Real-time PCR assays

Real-time PCR was performed on a Rotor Gene 3000 (Corbett Research) using the fluorescent marker SYBR Green (Applied Biosystems) to generate semiquantitative data. PCR premixes containing all reagents except for target cDNAs were prepared and aliquoted by a Robotic Liquid Handling System (CAS-1200, Corbett Robotics) into PCR tubes (Corbett Research). Primers used are listed in **Table 2**.

The PCR reaction was carried out following the program: 1 cycle at 95 °C for 10 min and 40 cycles consisting of 10 sec at 95 °C, 15 sec at 59 °C and 30 sec at 72 °C. This was followed by the measurement of fluorescence during a melting curve in which the temperature raised from 72 to 95 °C in sequential steps of 1 °C for 45 seconds. This insured the detection of one gene-specific peak and the absence of primer dimer peaks. Direct detection of the PCR product was measured by monitoring the increase in fluorescence caused by the binding of SYBR green dye to double-stranded DNA.

A fluorescence threshold was set manually to 0.05 on the log fluorescence scale to determine the cycle number (Ct value) at which the fluorescence passed the detection threshold. For each cDNA sample, relative expression levels of target gene were normalized by two reference genes: *actin* (*ACT*) and *elongation factor1-α* (*EF1-α*) .The amplification efficiencies ($E=10^{-1/\text{slope}}$) for each sample were calculated on the basis of the results of three amplification reactions, each with a different quantity of the template (6, 18, or 54 ng of total reverse transcribed RNA), according to Pfaffl (2001). Each reaction was performed in duplicate. The amplification efficiency, obtained for each sample, was used to

estimate the relative level of expression between L0 stage we chose as the 'calibrator' and the stage of interest (called 'Sample'):

$$R = \frac{E_{en}^{\Delta C_{en}}}{E_{ref}^{\Delta C_{ref}}}$$

where E_{en} is the amplification efficiency of Engrailed mRNA; E_{ref} is the amplification efficiency of reference mRNA; $\Delta C_{en}=Ct_c-Ct_s$ is the difference between cycle threshold (Ct) of the calibrator and the Ct of Sample for Engrailed mRNA; $\Delta C_{ref}=Ct_c-Ct_s$ is the difference between Ct of the calibrator and the Ct of Sample for reference mRNA.

2.6 Models

Two alternative models for engrailed expression pattern during anamorphosis were developed. The first model (model A) predicts that, at each stage, *en* is expressed both in the new forming segments at the rear of the trunk region and in the already formed (more anterior) trunk segments. The second model (model B) predicts that *engrailed* expression is localized only in the last segments added.

Each model provides for three variants that differ for the extension of the tissues that express *en* at each stage. The gene could be expressed in a row of cells that grows linearly with linear size (var. 1), in an epithelium that grows as a surface (i.e. with the square of linear size, var. 2) or in a three-dimensional portion of tissue that grows as a volume (i.e. with the cube of linear size, var. 3). These models are structurally identical and do not have adjustment parameters (see Appendix).

The two models share the same structural parameter values, computed on the basis of available information on anatomy, histology and external morphology of *L. peregrinus* (ontogenetic allometry, per-moult growth rate (weight), segmental composition of trunk and number of new segments for at each anamorphic stage). These parameter values had been previously estimated on samples of *L. peregrinus* bred in our laboratory in the same conditions of the animals investigated in this study.

2.7 Statistical analysis

StatGraphics Centurion XV software was used for statistical analysis. Significance of the differences in means were calculated using ANOVA tests, and values of $p<0.05$ were considered statistically significant.

2.8 Construction of *in situ* probes

RNA antisense probes for *wingless and tropomyosin* were prepared by cloning the fragments isolated as mentioned above, in pGEM®-T Easy Vector (Promega). After colony PCR, the plasmids from several positive colonies were purified using the Qiaprep SpinMiniprep kit (Qiagen) and sequenced. The plasmid, with a positively orientated insert of the cDNA, was linearized with an appropriate restriction enzyme (Sac II or Sall I, Promega) and used as the template for in vitro transcription of the antisense RNA probe using a T7 or SP6 RNA polymerase (Promega), according to the protocol supplied with the DIG RNA Labelling kit (Roche Molecular Biochemicals).

2.9 *In situ* hybridization

The following protocol was adapted from in situ hybridization studies on tunicates (Degasperi et al. 2009).
Specimens for ISH were killed by ethyl acetate to keep the muscle relaxed. Legs of dead animals were ablated in PBS pH 7.4 to assure a better permeability and fixed overnight in freshly prepared MOPS buffered (0.1 M MOPS (Sigma), 1 mM $MgSO_4$ (Sigma), 2 mM EGTA (Fluka), 0.5 M NaCl) 4% paraformaldehyde (Sigma). Fixative was removed by washing twice in PBS pH 7.4 (Sigma) and then samples were dehydrated through graded PBS/Ethanol to 100% then washed in xylene and embedded in Paraplast Plus (Sigma). Samples were serially sectioned (12 µm) and left to adhere to microscope slides, cleaned from the Paraplast with xylene (15 min), rehydrated in a graded series of ethanol to PBS, then used immediately.

Sections were incubated (6 min) in 10 µg/ml proteinase K (Promega) in PBS; the enzyme action was then stopped with a solution of 0.2% glycine in PBS, washed in PBS, postfixed in a 4% paraformaldehyde plus 0.2% glutaraldehyde solution in PBS and re-washed in PBS. Samples were then incubated in the hybridisation mix [50% formamide (Fluka), 1% Blocking Reagent (Roche), 5 mM EDTA (Fluka), 0.1% Tween-20 (Sigma), 0.1% CHAPS (Roche), 1 mg/ml heparin (Sigma) and 1 mg/ml tRNA (Roche), SSC 5×] 1 h at 65 °C followed by overnight incubation with 1-2 µg/ml DIG-labelled riboprobes. Specimens were washed twice in 2× SSC pH 4.5, twice in formamide 50% in 2× SSC pH 4.5 30 min at 42 °C, once in 2× SSC pH 4.5 30 min at 42 °C and twice in PBS-T (0.1% Tween-20 in PBS).

Subsequently slides were: i) incubated in a blocking solution [2% Blocking Reagent, 10% goat serum (Sigma) in PBS-T] 1 h at RT and then overnight with an alkaline phosphatase conjugate anti-DIG-antibody (Roche) for riboprobe detection, ii) treated with a NBT/BCIP solution (Roche) as alkaline phosphatase substrate, until the dye was detectable, iii) dehydrated in ethanol to a final step of xylene (15 min) and mounted in Eukitt (Electronic Microscopy Sciences). Sections were photographed with a Leica DM5000B light microscope accessorised with a Leica DFC 300 DX digital photo camera; images were then organised with CorelDRAW 11.
We tried to discriminate the specific signal of epidermis from the adjacent cuticle background using the fluorescent nuclear stain DAPI (4′,6-diamidino-2-phenylindole) after hybridization. Slides were stained with DAPI (5 µg/ml) in PBS 10 min and washed twice in PBS.

3. Results

3.1 Sequences of *engrailed* and *wingless*

L. peregrinus orthologues of *en* and *wg* were cloned from larval cDNA using standard methods. We recovered two different fragments with high sequence similarity to the *en* gene of other arthropods: *Lp_enL* (228 bp) and *Lp_enS* (180 bp). The two sequences differ for a short intervening sequence lacking stop

codons, between the EH2 and EH3 domains. This linker, also identified in the *Lithobius atkinsoni en* homolog (Hughes and Kaufman, 2002), has high similarity to the sequence of one of the *engrailed* paralogs of the spider *Cupiennius salei* (Damen, 2002; **Fig. 1**). Further DNA analysis are required to establish whether these sequences represented splicing variants.

No duplicates or transcriptional isoforms of *wg* were obtained from *L. peregrinus*. The fragment cloned, designed as *Lp_wg*, is 292 bp long, and it encodes a 97 aa protein. Lp_wg is most closely related to the wingless homologs of arthropods listed in **Figure 1**.

engrailed

wingless

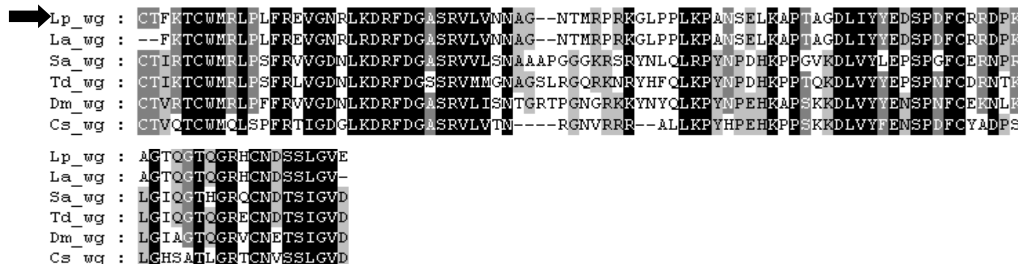

Fig.1 Sequence alignments of arthropod orthologs of *engrailed* and *wingless*. Arrows highlight the *L. peregrinus* sequences. A bar marks the insertion in Lp_enL, boxes mark similarity with the sequences of La_en and Cs_en1.
All sequences except those of *L. peregrinus* (Lp) were acquired from Genbank with the following Accession Number: *Lithobius atkinsoni* La_en (AF434998), *Strigamia maritima* Sm_en (AAL13136), *Artemia franciscana* Af_en (CAA50279), *Schistocerca americana* Sa_en (AAA29807), *Thermobia domestica* Td_en1 (AF104006) *Thermobia domestica* Td_en2 (AF104007), *Porcellio scaber* Ps_en1 (AF254262), *Porcellio scaber* Ps_en2 (AF254262), *Cupiennius salei* Cs_en1 (CAA07503), *Cupiennius salei* Cs_en2 (CAC87039); *Lithobius atkinsoni* La_wg (AAL36911), *Schistocerca americana* Sa_wg (AAD37798), *Thermobia domestica* Td_wg (AAF43000), *Daphnia magna* Dm_wg (BAJ05334), *Cupiennius salei* Cs_wg (CAC87040). Orthologous sequences from arthropod species were aligned with fragments from *L. peregrinus* using ClustalW program and and amino acids are shaded according to the degree of conservation using GeneDoc: black (100% similarity); grey (80–90% similarity); light grey (60–70% similarity).

3.2 engrailed expression pattern

In our experiments *wingless* was detected more than 15 PCR-cycles later than any of the standard gene candidates. Therefore, it proved difficult to obtain quantitative data for both *wingless* and the standard genes candidates for the same sample dilution series, as required. As consequence, only *en* PCR assays results are reported.

Using Real Time PCR it was impossible to discriminate between the different isoforms of *Lp_en,* so the expression data we analyzed for *en* represent the summation of the expression level in all isoforms of the gene.

Embryos express high levels of *en*, in conformity with what is reported by Hughes and Kaufman (2002) for *L. atkinsoni*. While validating our Real Time PCR protocol, this group of specimens was nevertheless excluded from subsequent analyses because we have no means to refer expression data to precise time points in embryonic development.

The effects of the reference housekeeping gene and the developmental stage on *en* relative expression values were tested with a two-way ANOVA on all the experimental groups to the exclusion of L0 (chosen as 'calibrator' of *en* expression, value set to 1). Expression levels in the six stages differ significantly ($p<0.0001$), whereas there is no significant difference between *EF1-α* and *ACT* data normalization ($p>0.26$). This suggest that both genes are appropriate as internal control of gene expression in this species. Accordingly, relative data for each stage obtained using the two reference genes were lumped together in subsequent analyses.

Engrailed RNA transcript is more abundant in L1 stage, to decrease toward a minimum value in adult stage. Variation in *en* level during anamorphosis suggests that gene expression is somehow modulated during this period of development.

Adult *en* expression level differs significantly from 0 (T-test, $p<0.0011$) as it also differs significantly from the expression levels of all other stage but for L4 (Fischer's test LSD, $\alpha=0.01$, **Fig. 2A**). As the segmental pattern is completed in PL1, adult *engrailed* expression reveals the involvement of the gene in functions other than that related to the process of segmentation. Assuming that the during anamorphosis *en* is involved in the same functions (in addition to its role in segmentation), to get a better estimate of *engrailed* expression level in segmentation, we rescaled its expression values in each pre-adult stage subtracting a quantity equal to the transcript level in the adult. One-way ANOVA on this transformed dataset confirms the significance of the stage as a factor affecting *en* expression levels in L1-PL1 ($p<0.0001$, **Fig. 2B**).

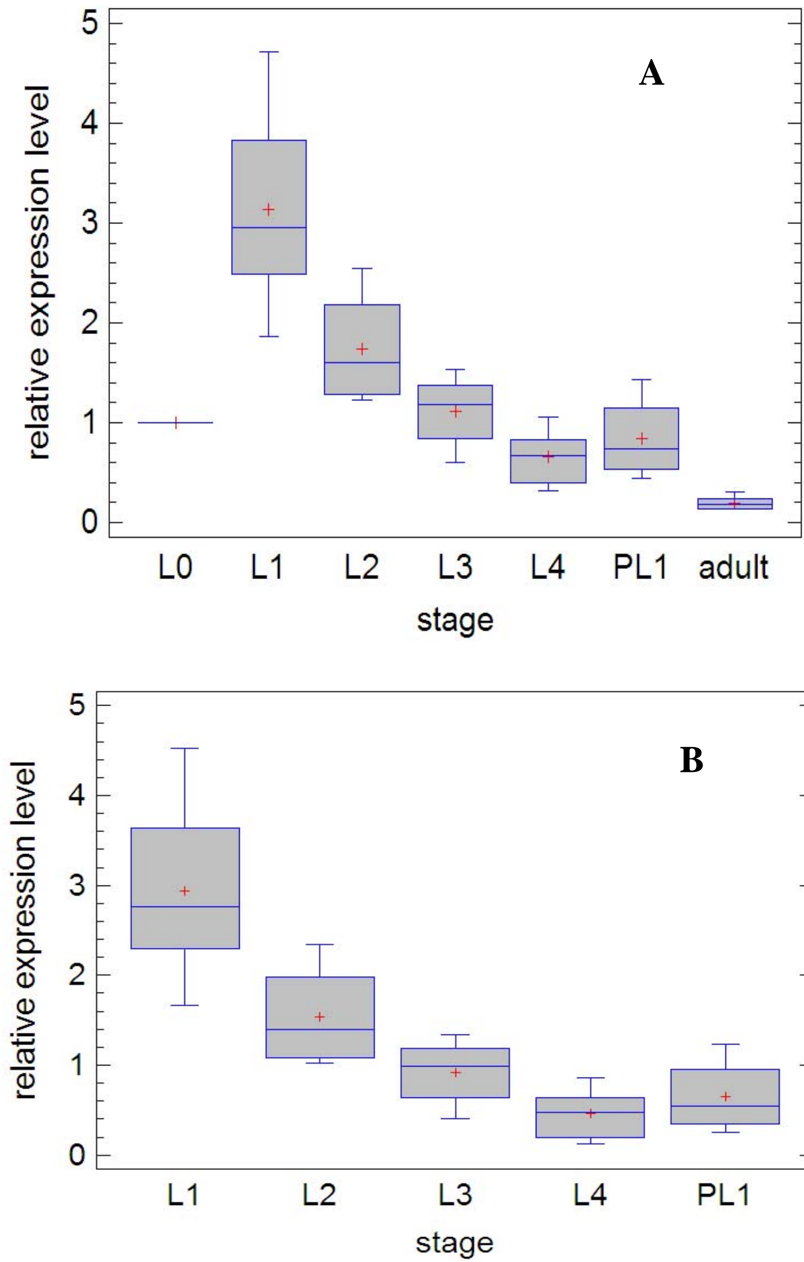

Fig. 2. Relative expression level of *engrailed* mRNA during anamorphosis. (**A**) Data normalized with respect to the expression level of housekeeping genes. (**B**) Same data as in (A), further rescaled by subtracting the adult level of expression from each pre-adult stage. Boxes represent the interval between lower and upper quartiles, with median (transverse line) and mean (small cross); vertical lines are ranges of variation.

For comparing the two alternative models of *en* expression pattern in *L. peregrinus*, we used this transformed dataset. The models were matched to observed data by calculating the Residuals Sum of Squares (RSS) between observed and expected values and the relative Coefficient of Determination (R^2), a measure of the variance explained by the model (**Tab. 3**).

Table 3. Residuals Sum of Squares (RSS) and Coefficient of Determination (R^2) for two models of *en* expression (coming in three variants each) when matched to observed data.

	model A			model B		
	variant A_1	variant A_2	variant A_3	variant B_1	variant B_2	variant B_3
RSS	82.54	71.89	64.64	37.33	28.18	27.42
R^2	0.39	0.47	0.52	0.72	0.79	0.80

Model B (expression restricted to the terminal trunk region) shows RSS values consistently lower than those of the alternative model (**Tab.3**, **Fig. 3**). The two variants of this model predicting an increase of the level of expression with the square and the cube of linear dimensions, respectively, produce similar R^2 values, both greater than 0.79.

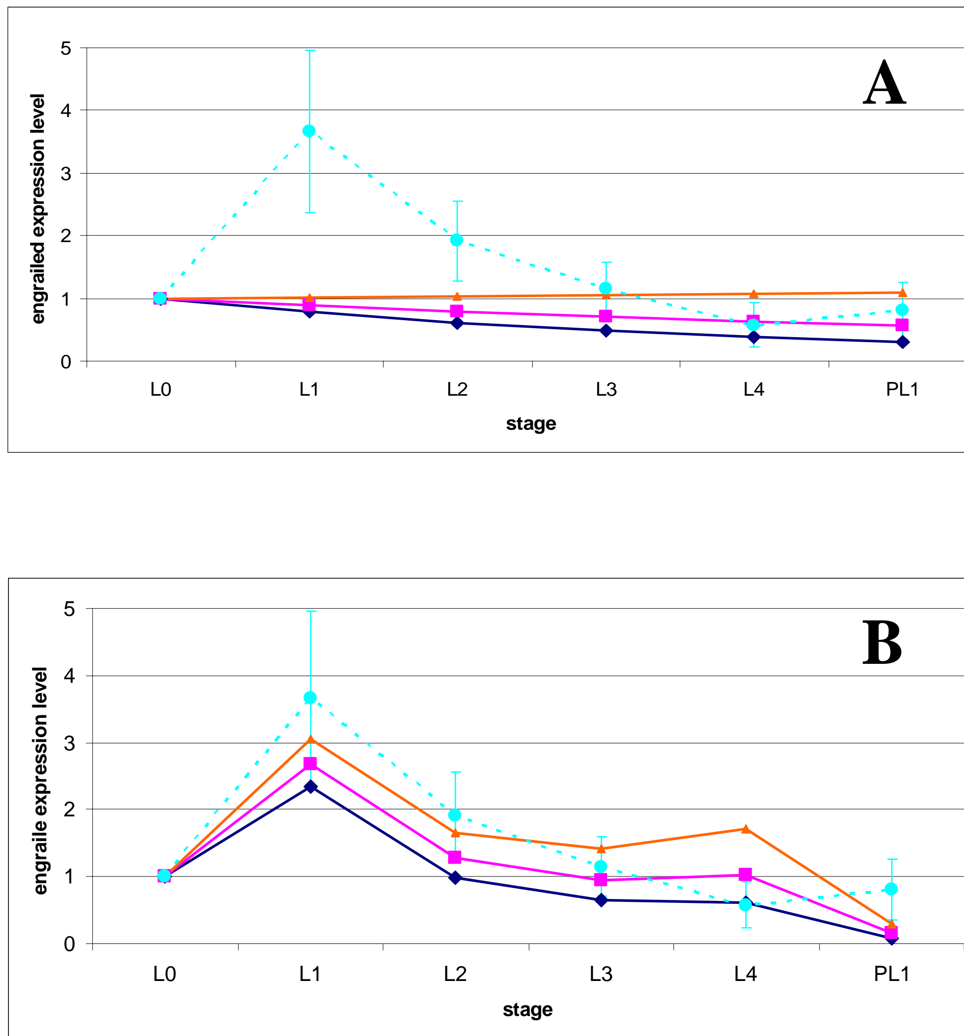

Fig. 3. Observed levels of *engrailed* expression during anamorphosis in *Lithobius peregrinus* and expected values under two models (with three variants each). (**A**) *En* expressed in all trunk segments. (**B**) *En* expressed only in the last segment formed at each stage. Dotted line: observed values and vertical lines are observed ranges of variation; blu line: *en* expressed in a row of cells (variant 1); purple line: *en* expressed in an epithelium (variant 2); red line: *en* expressed in a three-dimensional portion of a tissue (variant 3).

3.3 *wingless* expression pattern

In carrying out an in situ hybridization experiment it is very important to be confident that the hybridization reaction is specific and that the probe is in fact binding selectively to the target mRNA sequence rather than to other components of the cell or other closely related mRNA sequences. In addition, if no staining is observed with the probe, this could mean that there is actually no expression of that mRNA in the tissue, or that there may be problems with tissue preparation.

In order to avoid these problems we tested two different controls for in situ hybridization experiments: a probe for an housekeeping gene and a reaction with no riboprobe. Ignoring where and when *wg* or *en* could be expressed, the protocol was validated using a probe for *L. peregrinus* tropomyosin that specifically marks muscular tissues. The positive result obtained with tropomyosin probe confirmed RNA integrity and the signal specificity. Furthermore, a reaction control without riboprobe was used to test the specific binding of antibody used to visualized the probe. Cuticle, yolk (still present in early larvae) and gut tissues bind non-specifically all the probes that we tested, as it has also been noted in other arthropods (e.g., Chang et al., 1996).

In situ hybridization on sections of L1 and L2 larvae revealed that *wg* is expressed in the cortex (the tissue where neuronal cell bodies are located) of the encephalon, and all the ganglia of the ventral nerve cord (**Fig. 4A-C**). Well-formed neuromeres showed an intense signal in the cortex tissue, whereas no *wg* expression was detected in the network of axones, dendrites, and glial branches (neuropile). Specific signal was detectable also in other tissues of the more posterior segments still in formation. Here the undifferentiated tissues that will develop into the nervous system is in close connection with the ventral epithelium, forming a solid mass of cell equally stained by *wg* probe (**Fig. 4D-E**). Frontal sections of this region showed that *wg* is also expressed in undifferentiated tissues of the leg buds (**Fig. 4F**).

It was not possible to discriminate clearly the specific signal of epidermis from the adjacent cuticle background. Thus, further experiments are necessary to confirm the positive response of epidermis to *wg* probe.

In parallel an in situ hybridization with a probe for *engrailed* was performed, but, in opposition to Real Time PCR results, no specific signal was detected. The cause of the problem could be the short length of the probe that we used in *L. peregrinus* (see Erikson et al., 2009).

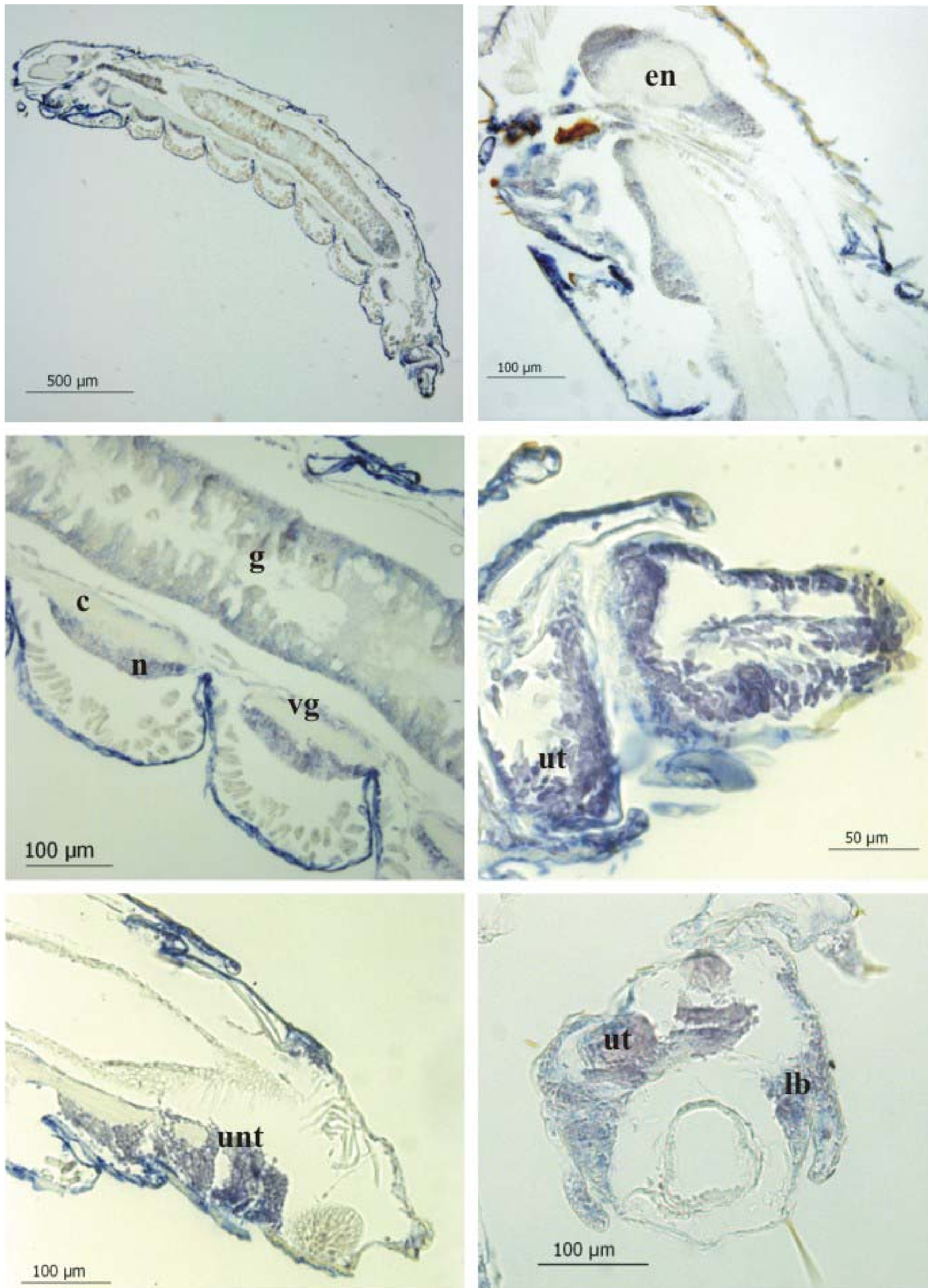

Fig.4 *wg* expression in paraffin sections *of L. peregrinus* larvae. (**A**) Sagittal section of L2 larvae. Specific signal for *wg* probe is detectable in the encephalon (**B,** sagittal section) and in all ganglia of ventral nervous system (**C**, sagittal section). *wg* is also expressed in the undifferentiated tissues of the more posterior segment in formation (**D** and **E**, sagittal sections) and in the leg buds (**F**, frontal section of L1larvae, dorsal view). **en** = encephalon; **g** = gut; **vg** = ventral ganglia; **c** = cortex; **n** = neuropile; **ut** = undifferentiated tissue; **unt** = undifferentiated nervous tissue; **lb** = leg buds.

4. Discussion

The post-embryonic development of *L. peregrinus* includes an anamorphic phase of sequential addition of trunk segments and the corresponding appendages. Although developmental genetic studies on post-embryonic segmentation had never been carried out before in myriapods, we hypothesized that the segment-polarity genes *engrailed* and *wingless*, involved in the establishing of an intra-segmental pattern of the trunk during embryogenesis in most arthropods investigated so far, could have a role in segmentation also during anamorphosis.

As previously found in *L. atkinsoni* embryos (Hughes and Kaufman, 2002), we have identified in *L. peregrinus* two sequences for the *en* homolog, *Lp_enL* and *Lp_enS*, differing for an insertion/deletion of 15 amino acids, and only an homologue for *wg*, *Lp_wg*. Using PCR it was impossible to discriminate between the different isoforms of *Lp_en,* so the expression data we analyzed for *en* represent the summation of the expression level in both isoforms of this gene.

The first result obtained in this work was the confirmation of *en* and *wg* expression during all anamorphic stages. The *en* transcript is expressed also in adults, but the level is significantly lower, compared to previous stages. This result supports the hypothesis that *en* has a specific role in post-embryonic segmentation.

In *L. atkinsoni* embryo, a stripe 3-4 cell wide expresses *en* at the posterior boundary of each segment. These stripes appear sequentially in the embryo, from anterior to posterior, reflecting the short germ-band mode of development of this arthropod (Hughes and Kaufman, 2002). From Real Time PCR it turns out that *en* expression is precisely modulated during anamorphosis of *L. peregrinus*, as its domain is limited to the new segments that are forming in each stage. Testing the three variants of this models of *en* expression during anamorphosis, it also turns out that within a single segment the portion of tissue expressing the gene growths most probably as an extending epithelium or a three-dimensional mass of cell, rather than as a stripe of cells of invariant width. This dissimilarity with respect to embryonic expression was somehow anticipated, considering the full three-dimensional architecture of the body of the larva, compared to the almost bi-dimensional constitution of the embryo at the stage of the germ-band. Therefore, the possibility that a segment-polarity role is retained, although with different mode of expression, during post-embryonic anamorphic stages is perfectly compatible with our data.

The high level of conservation of the *en* expression domain across the arthropods, on the posterior margin of embryo segments, probably reflects the gene's central role in segmental patterning. Actually, a different domain of expression has been recorded on the dorsal embryonic tissues of the pill millipede *G. marginata* (Janssen et al., 2004), possibly related to the dorso-ventral mismatch in segmental patterning typical of the Diplopoda. However, the different precise localization of *en* dorsal expression within the segment does not cast any doubts on its segment-polarity role.

En expression pattern observed in *L. peregrinus* confirms the role of this gene in segmentation also during post-embryonic development. A new series of in situ hybridization experiments, using a longer probe for *en*, will be necessary to

precisely localize the tissues expressing the gene and to investigate the possible different function of the two *en* isoforms, as it turned out to be the case in chelicerates (Damen et al., 2002) and crustaceans (Gibert et al., 2000).

In situ staining of *wingless* shows a marked expression of the gene in the cortex of the encephalon and all the ganglia of the ventral nervous system, as well as in the undifferentiated tissues of the more posterior segments still in formation (leg buds included). This very extended domain of expression during postembryonic stages suggests that wingless does not retain in segmental patterning the same role it has during embryonic development.

Wg transcript localization in well-formed ganglia, as well as in undifferentiated ones, suggests an involvement of *wg* in both the formation and normal functioning of the nervous system, as demonstrated in other arthropods (reviewed in Cadigan and Nusse, 1996; Packard et al., 2002). The intense labeling in the leg buds suggests also a role for *wg* in the development of limbs. In the embryonic limb buds of several arthropods (e.g., Nagy and Carroll, 1994), as well as in the imaginal discs of wings and legs of some holometabolous insects (Cohen et al., 1993), *wg* contributes establishing dorso-ventral and/or antero-posterior identity of differentiating cells. Further investigations on the most posterior segments bearing not fully formed legs will clarify if *wg* function in leg development is conserved during anamorphosis.

Wg was the first segment polarity gene to be shown to have a conserved expression pattern not only in long- and short-germ insects (Nagy and Carroll, 1994), but also in other arthropods, as myriapods (Hughes and Kaufman, 2002), crustaceans (Prpic, 2008), chelicerates (Damen, 2002), and even in onychophorans (Eriksson et al., 2009). However, several lines of evidence suggest that not all *wg* functions are as much conserved. In the embryo of the pill millipede *G. marginata*, *wg* is expressed ventrally but not dorsally (Janssen et al., 2004), and dorso-ventral differences in gene expression have been also reported for the notostracan crustacean *Triops longicaudatus* (Nulsen andNagy, 1999) and the spider *Cupiennius salei* (Damen, 2002). Moreover, in insects, the involvement of *wg* in leg development seems to be an apomorphy of holometabolans. Studies on the cricket *Gryllus bimaculatus* and the milkweed bug *Oncopeltus fasciatus* showed that *wg* is necessary for normal body segmentation, but does not seem to have a role in leg development (Miyawaki et al, 2004; Angelini and Kaufman, 2005).

Beyond that, an important methodological goal reached by this work is the protocol setting up in a centipede for two very powerful techniques of modern molecular biology: Real Time PCR and *in situ* hybridization on paraffin sections. Using Real Time PCR, never applied on myriapods before, we demonstrated that the gene *elongation factor 1-α* (*LpEf1*-α) and *actin* (*LpAct*) are constitutively transcribed during the life cycle of *L. peregrinus*, validating, for the first time, their use as reference genes in this species. The protocol set up for in situ hybridization on paraffin sections is innovative for arthropods, although it is intensively applied in studies on tunicates (Degasperi et al., 2009).

Further studies on post-embryonic development of different anamorphic arthropods are needed to gain a better understanding of this segmentation process and the evolution of segmentation as a whole.

References

Abzhanov, A., and Kaufman, T. C. (2000). Homologs of *Drosophila* appendage genes in the patterning of arthropod limbs. Developmental Biology, **227**: 673–689.

Angelini, D.R., Kaufman, T.C. (2005). Functional analyses in the milkweed bug *Oncopeltus fasciatus* (Hemiptera) support a role for Wnt signaling in body segmentation but not appendage development. Developmental Biology, **283**:409–423

Cadigan, K. and Nusse, R. (1996). wingless signaling in the Drosophila eye and embryonic epidermis. Development, **122**: 2801–2812.

Chang, P.S., Lo, C.F., Wang, Y.C. and Kou, G.H. (1996). Identification of white spot syndrome associated baculovirus (WSBV) target organs in shrimp, *Penaeus monodon*, by in situ hybridization. Diseases of Aquatic Organisms,**27**:131–139.

Cohen, B., Simcox, A.A., Cohen, S.M. (1993). Allocation of the thoracic imaginal primordia in the *Drosophila* embryo. Development **117**:597– 608.

Damen, W., Weller, M., Tautz, D. (2000). Expression patterns of hairy, even-skipped, and runt in the spider *Cupiennius salei* imply that these genes were segmentation genes in a basal arthropod. Proceedings of the National Academy of Science of the United States, **97**:4515–19.

Damen, W.G.M. (2002). Parasegmental organization of the spider embryo implies that the parasegment is an evolutionary conserved entity in arthropod embryogenesis. Development, **129**:1239– 1250.

Degasperi V., Gasparini F., Shimeld S.M., Sinigaglia, C., Burighel, P., Manni, L. (2009). Muscle differentiation in a colonial ascidian: organisation, gene expression and evolutionary considerations. BMC Developmental Biology, 9:48.

Desplan C., Theis J., O'Farrell P.H. (1985). The *Drosophila* developmental gene, engrailed, encodes a sequence-specific DNA binding activity. Nature, **318**:630-635.

DiNardo, S., Kuner, J.M., Theis, J. and O'Farrel, P.H. (1985). Development of embryonic pattern in *D. melanogaster* as revealed by accumulation of the nuclear Engrailed protein. Cell, **43**: 59-69.

Dohle, W., Gerberding, M., Hejnol, A., Scholtz, G., (2004). Cell lineage, segment differentiation, and gene expression in crustaceans. In Evolutionary Developmental Biology of Crustacean, Crustacean Issue 15 (Scholtz, G. ed.). A.A.Balkema Publishers, Lisse, 95-133.

Eriksson, B.J., Tait, N.N., Budd G.E., Akam , M. (2009). The involvement of engrailed and wingless during segmentation in the onychophoran *Euperipatoides kanangrensis* (Peripatopsidae: Onychophora) (Reid 1996). Development Genes and Evolution , **219**:249–264.

Fjose A., McGinnis W.J., Gehring W.J. (1985). Isolation of a homeo box-containing gene from the engrailed region of *Drosophila* and the spatial distribution of its transcripts. Nature **313**:284-289.

Fusco G. (2005).Trunk segment numbers and sequential segmentation in myriapods. Evolution & Development, **7**: 608-617.

Gibert, J.M., Mouchel-Vielh, E., Quéinnec E. and Deutsch, J. S. (2000). Barnacle duplicate *engrailed* genes: divergent expression patterns and evidence for a vestigal abdomen. Evolution and Development, **2**: 194-202.

Hidalgo, A. (1994). Three distinct roles for the engrailed gene in *Drosophila* wing development. Current Biology, 4:1087–1098.

Hughes C.L. and Kaufman T.C. (2002). Exploring myriapod segmentation: the expression patterns of *even-skipped*, *engrailed*, and *wingless* in a centipede. Developmental Biology, **247**:47-61.

Janssen, R. Prpic, N.-M., & Damen, W.G.M. (2004). Gene expression suggests decoupled dorsal and ventral segmentation in the millipede *Glomeris marginata* (Myriapoda: Diplopoda). Developmental Biology*,* 268: 89-104.

Kadner D. and Stollewerk A. (2004). Neurogenesis in the chilopod *Lithobius forficatus* suggests more similarities to chelicerates than to insects. Development Genes and Evolution, **214**:367-379.

Kettle C., Johnstone J., Jowett T., Arthur H., Arthur W. (2003). The pattern of segment formation, as revealed by *engrailed* expression, in a centipede with a variable number of segments. Evolution and Development, **5**:198-207.

Lawrence P.A., Morata G. (1976). Compartments in the wing of *Drosophila*: a study of the *engrailed* gene. Developmental Biology, **50**:321-337.

Martinez-Arias, A., Baker, N. and Ingham, P.W. (1988). Role of segment polarity genes in the definition of cell states in the *Drosophila* embryo. Development, **103**: 157-170.

Minelli A., Brena C., Deflorian G., Maruzzo D., Fusco G. (2006). From embryo to adult – beyond the conventional periodization of arthropod development. Development Genes & Evolution, **216**: 373-383.

Miyawaki, K., Mito, T., Sarashina, I., Zhang, H., Shinmyo, Y., Ohuchi, H., Noji, S. (2004). Involvement of *Wingless*/*Armadillo* signaling in the posterior sequential segmentation in the cricket, *Gryllus bimaculatus* (Orthoptera), as revealed by RNAi analysis. Mechanism of Development, **121**:119–130

Nagy, L.M., Carroll, S., (1994). Conservation of *wingless* patterning functions in the short-germ embryos of *Tribolium castaneum*. Nature,**367**:460–463.

Niwa, N., Inoue, Y., Nozawa, A., Saitoh, M., Misumi, Y., Ohuchi, H., et al., (2000). Correlation of diversity of leg morphology in *Gryllus bimaculatus* (cricket) with divergence in *dpp* expression pattern during leg development. Development **127**:4373–4381.

Nulsen, C., Nagy, L.M. (1999). The role of *wingless* in the development of multibranched crustacean limbs. Development Genes and Evolution, **209**:340–348

Packard, M., Koo, E.S., Gorczyca, M., Sharpe, J., Cumberledge, S., Budnik, V. (2002). The *Drosophila wnt*, *wingless*, provides an essential signal for pre- and postsynaptic differentiation. Cell, **111**:319 –330.

Patel, NH. (1994). Developmental evolution: insights from studies of insect segmentation. Science, **266**:581–90.

Pfaffl, M.W.(2001). A new mathematical model for relative quantification in real-time RT-PCR. Nuceid Acids Research*,* **29**: e45.

Prpic, N.M. (2008). Parasegmental appendage allocation in annelids and arthropods and the homology of parapodia and arthropodia. Frontiers in Zoology, 5:17.

Rijsewijk, F., Schuermann, M., Wagenaar, E., Parren, P., Weigel, D. Nusse, R. (1987). The *Drosophila* homolog of the mouse mammary oncogene *int-1* is identical to the segment polarity gene *wingless*. Cell*,* **50**:649-57.

Rivera-Pomar R. and Jäckle H. (1996). From gradients to stripes in *Drosophila* embryogenesis: Filling in the gaps. Trends in Genetics, **12**: 478–483.

Sanson, B. (2001). Generating patterns from fields of cells. EMBO Reports, **2**:1083– 1088.

Siegler, M. V. S. and Jia, X. X. (1999). *Engrailed* negatively regulates the expression of cell adhesion molecules connection and neuroglian in embryonic *Drosophila* nervous system. Neuron, **22**:265-276.

Takashima, S., Yoshimori, H., Yamasaki, N., Matsuno, K., Murakami, R. (2002). Cell-fate choice and boundary formation by combined action of *Notch* and *engrailed* in the *Drosophila* hindgut. Development Genes and Evolution, **212**:534–541.

Appendix

Two alternative models for the expected level of expression of the gene *engrailed* have been devised. Model A assumes that *en* is constantly expressed in all trunk segments during anamorphosis, while model B assumes that *en* expression is restricted to the segments new forming at each anamorphic stage. Each model provides for three variants that differ for the way in which the tissues that express *en* grow across stages: in one, two, or three dimensions.

For each stage, the expected relative level of gene expression (e) is calculated as the ratio between the expected relative body mass (m_{Tot}, relative to the mass of L0) an the expected relative mass of the tissues expressing the gene (m_E, relative to the mass of the gene-expressing tissues in L0) under the specific variants of the models

$$e = \frac{m_E}{m_{Tot}}$$

m_{Tot} is directly computed from the observed weight growth rate, while for m_E, the relative volume of the trunk expressing the gene, computed form the observed linear-size growth rate, is used as a proxy of the measure. Thus, for each model, variant and stage the later is calculated as

$$m_E = \frac{N_E}{N_{Tot}} l^D$$

where N_E is the number of trunk segments expressing the gene, n_{Tot} is the total number of trunk segments, and l is trunk length at the specific stage, while D is the number of spatial dimensions of growth for the tissues expressing the gene in each segment. The two models differ for the n_E/n_{Tot} ratio, as for model A, $n_E=n_{Tot}$, while for model B, $n_E=n_F$, where n_F the number of new trunk segments in formation during the stage. The three variants for each model differ in the D value (1, 2, or 3).

Growth and segmentation parameter values had been previously estimated on samples of *L. peregrinus* bred in our laboratory. These are body weight per-moult growth rate (1.455), trunk length per-molt growth rate (1.140), and segmental composition of the trunk for each anamorphic stage, as inferred through morphological and anatomical observations. The following table lists the parameter values used in model calculations.

stage	n_{Tot}	n_F	l	m_{Tot}
L0	8	1	1.000	1.000
L1	8	3	1.140	1.455
L2	10	2	1.300	2.116
L3	12	2	1.482	3.079
L4	15	3	1.689	4.478
PL1	15	0.5*	1.925	6.515

* In the first post-larval stage (PL1), by definition anamorphosis should be already completed, at least on the basis of the count of the number of leg pairs, as traditionally reported. However, we assigned an n_F = 0.5 to stage PL1 to account for developmental processes of segmental pattern formation still active at this stage at the level of tergites, sternites, tracheal openings and nervous ganglia (Minelli et al., 2006).

General conclusions

This thesis represents the first study on post-embryonic gene expression in a myriapod, and also the applied techniques constitute a novelty for arthropod studies focused on this developmental period.

Anamorphic development is a common trait in many arthropod taxa, in Crustacea (e.g. Scholtz, 2000), Protura (Dallai, 1980), Pycnogonida (Hooper, 1980) and Mmyriapoda (e.g. Enghoff et al., 1993), but it has been seldom investigated through a molecular approach (Dohle, 2004).

We showed that in the centipede *L. peregrinus*, two segment-polarity genes, *engrailed* and *wingless*, involved in establishing the segmental pattern of the main body axis during embryogenesis, are intensively expressed also during all anamorphic stages. Only *en* seems to retain a segment-polarity role in post-embryonic development, although with different modality of expression, whereas *wg* function is seemingly related to the development of nervous system and leg buds, but with no sign of periodic expression, thus suggesting that it does not retain in segmental patterning the same role it has during embryonic development.

Comparative studies on expression pattern of segmentation genes in other anamorphic arthropods will be useful to confirm *en* role in segmentation in post-embryogenesis and to clarify which roles plays *wg* in these developmental stages.

The validation of two appropriate housekeeping genes in *L. peregrinus* guarantees that target genes expression detected in this study does not depend on experimental conditions, and provides at the same time a precious background for further investigations on gene expression in the centipede. Until now, mRNA transcript localization in arthropods was only carried out in embryos or in tissues isolated from larvae or adults. The possibility to apply in situ hybridization protocols on paraffin sections gives now the opportunity to study gene expression also in the whole body of arthropods during post embryonic development.

Results obtained in this study represent also the first data on the genes involved in the ecdysone signal pathway in a myriapod species. We isolated partial sequences of the ecdysone receptor (LpEcR) and the retinoid X receptor (LpRXR) in *L. peregrinus*. A comparison with EcRs and USP/RXRs orthologues of other arthropods shows high level of sequence identity with chelicerates and hemimetabolous insects, in particular in the DNA-binding domain and ligand-binding domain, that are very conserved across arthropods.

Different expression isoforms of both receptors have been investigated in insects (e.g. Maestro et al., 2005), chelicerates (e.g. Palmer et al., 1999) and crustaceans (e.g. Chung et al., 1998), but their function is unknown. Further studies are required to know whether the multiple variants of *LpEcR* and *LpRXR* are differently expressed in centipede and whether they have different properties in heterodimerization, DNA-binding and ligand-binding activity.

The expression patterns shown by LpEcR and LpRXR during the anamorphic stage L1 of *L. peregrinus* are similar. Expression levels of both receptors increase in the second part of the inter-moult period, as observed in other arthropods, where expression profile of most ecdysteroid-regulated genes directly correlates with the ecdysteroid titer (Sullivan and Thummel, 2003). Lacking specific hormonal titer measurements during post-embryonic development in this species, we could only hypothesize an ecdysone peak 48 hours after the previous moult, corresponding to the increase of LpEcR and LpRXR transcription.

Our research failed to isolate *Broad Complex* gene in centipede, and some lines of evidence suggest it is an insect apomorphy. However, there are several other target genes regulated by EcR and RXR, isolated in insects and other few arthropods as well, that could be investigated to get a better understanding molting process in *Lithobius*.

The results obtained in this work, while representing the first step towards the understanding of the molting and segmentation during post-embryonic life in myriapods, once put in a proper comparative context, will certainly contribute to shed light on the evolution of this developmental processes end their relationship in arthropods.

References

Chung, A.C., Durica, D.S., Clifton, S.W., Roe, B.A., Hopkins, P.M. (1998). Cloning of crustacean ecdysteroid receptor and retinoid-X receptor gene homologs and elevation of retinoid-X receptor mRNA by retinoic acid. Molecular and Cellular Endocrinology, **139:** 209–227.

Dallai, R. (1980). Considerations on Apterygota phylogeny. Bollettino di Zoologia, **47**:35–48.

Dohle, W., Gerberding, M., Hejnol, A., Scholtz, G., (2004). Cell lineage, segment differentiation, and gene expression in crustaceans. In Evolutionary Developmental Biology of Crustacean, Crustacean Issue 15 (Scholtz, G. ed.). A.A.Balkema Publishers, Lisse, 95-133.

Enghoff, H., Dohle, W., and Blower, J. G. (1993). Anamorphosis in millipedes (Diplopoda)-the present state of knowledge and phylogenetic considerations. Zoological Journal of the Linnean Society of London. **109**: 103–234.

Hooper, J.N.A. (1980). Some aspects of the reproductive biology of *Parapallene avida* Stock (Pycnogonida: Callipallenidae) from northern New South Wales. Australian Zoologist. **20**:473-484.

Maestro, O., Cruz, J., Pascual, N., Martin, D., Bellés, X. (2005). Differential expression of two RXR/ultraspiracle isoforms during the life cycle of the hemimetabolous insect *Blattella germanica* (Dictyoptera, Blatellidae). Molecular and Cellular Endocrinology, **238**:27-37.

Palmer, M.J., Harmon, M.A., Laudet, V. (1999). Characterization of EcR and RXR homologues in the ixodid tick, *Amblyomma americanum* (L.). American Zoologist, **39**: 747–757.

Scholtz, G. (2000). Evolution of the nauplius stage in malacostracan crustaceans. Journal of Zoological Systematics and Evolutionary Research, **38**:175–187.

Sullivan, A.A., and Thummel, C.S. (2003). Temporal profiles of nuclear receptor gene expression reveal coordinate transcriptional responses during *Drosophila* development. Molecular Endocrinology, **17**:2125–2137.