Contents lists available at [SciVerse ScienceDirect](#)

Peptides

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

An evolutionary comparison of leucine-rich repeat containing G protein-coupled receptors reveals a novel LGR subtype

Matthias B. Van Hiel^{a,b,1}, Hans Peter Vandersmissen^{a,1}, Tom Van Loy^a, Jozef Vanden Broeck^{a,*}

^a Animal Physiology and Neurobiology, Zoological Institute of the Katholieke Universiteit Leuven, Naamsestraat 59, P.O. Box 02465, B-3000 Leuven, Belgium

^b Department of Genetics, Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, 30 Flemington Road, Parkville, Melbourne, Victoria 3010, Australia

ARTICLE INFO

Article history:

Received 4 September 2011

Received in revised form 3 November 2011

Accepted 3 November 2011

Available online xxx

Keywords:

Leucine-rich repeat containing G
protein-coupled receptor
Glycoprotein hormone receptor
Evolution

ABSTRACT

Leucine-rich repeat containing G protein-coupled receptors or LGRs are receptors with important functions in development and reproduction. Belonging to this evolutionarily conserved group of receptors are the well-studied glycoprotein hormone receptors and relaxin receptors in mammals, as well as the bursicon receptor, which triggers cuticle hardening and tanning in freshly eclosed insects. In this study, the numerous LGR sequences in different animal phyla are analyzed and compared. Based on these data a phylogenetic tree was generated. This information sheds new light on structural and evolutionary aspects regarding this receptor group. Apart from vertebrates and insects, LGRs are also present in early chordates (Urochordata, Cephalochordata and Hyperoartia) and other arthropods (Arachnida and Branchiopoda) as well as in Mollusca, Echinodermata, Hemichordata, Nematoda, and even in ancient animal life forms, such as Cnidaria and Placozoa. Three distinct types of LGR exist, distinguishable by their number of leucine-rich repeats (LRRs), their type-specific hinge region and the presence or absence of an LDLa motif. Type C LGRs containing only one LDLa (C1 subtype) appear to be present in nearly all animal phyla. We here describe a second subtype, C2, containing multiple LDLa motifs, which was discovered in echinoderms, mollusks and in one insect species (*Pediculus humanis corporis*). In addition, eight putative LGRs can be predicted from the genome data of the placozoan species *Trichoplax adhaerens*. They may represent an ancient form of the LGRs, however, more genomic data will be required to confirm this hypothesis.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

G protein-coupled receptors (GPCRs) containing an N-terminal domain with leucine-rich repeats (LRRs), such as the mammalian glycoprotein hormone receptors, have long been thought to be restricted to vertebrates. However, based on molecular cloning efforts and genome analyses, the ancient origin of the leucine-rich repeat containing GPCRs (LGRs) was revealed. Three distinct types of LGRs have been defined based on their structural characteristics (type A, B and C). The differences between these three LGR types mainly lie in the structure of their ectodomain, while the serpentine

region is relatively well preserved between the three types. The three LGR types differ in the number of LRR motifs, the absence or presence of an LDLa motif (low density lipoprotein receptor-like cysteine-rich motif) and the type-specific hinge region (Fig. 1). An LRR motif consists of approximately 24 amino acids (AAs) and has a consensus sequence LxxLxxLxxNxLxxLPxxxFxx in which the leucines (L) can be readily replaced by an isoleucine (I) or a valine (V) and sometimes even with the weakly similar phenylalanine (F) or methionine (M) (Suppl. File 1A). Generally, type B LGRs have about twice the number of LRRs compared to the other two types. An exclusive feature of the type C LGRs is the presence of an LDLa motif in the ectodomain.

The sequence of the hinge region appears to be specific for every type of LGR. It is the sequence between the most C-terminal leucine-rich repeat and the start of the transmembrane domain. It functions, as the name implies, as a hinge between the anchored transmembrane domain and the extracellular horseshoe-shaped ectodomain.

Type A LGRs include the mammalian glycoprotein hormone receptors FSHR, LH/CGR and TSHR and the fruit fly LGR1. The type B LGRs contain the mammalian LGR4, 5 and 6, and the fruit fly LGR2 or bursicon receptor. The third type of LGRs (type C) contain the mammalian LGR7 and LGR8, also called RXFP1 and RXFP2 respectively, and the fruit fly LGR3 and LGR4 [67]. Some LGRs have already been

Abbreviations: FSH(R), Follicle stimulating hormone (receptor); GPCR, G protein-coupled receptor; ILP, Insulin-like peptide; LDLa, low density lipoprotein receptor-like motif; LGR, Leucine-rich repeat containing GPCR; LH/CG(R), Luteinizing hormone/choriogonadotropin (receptor); LRR, Leucine-rich repeat; TSHR, Thyroid stimulating hormone (receptor).

* Corresponding author. Tel.: +32 16324260; fax: +32 16323902.

E-mail addresses: mvanhiel@unimelb.edu.au (M.B. Van Hiel), HansPeter.Vandersmissen@bio.kuleuven.be (H.P. Vandersmissen), Tom.VanLoy@bio.kuleuven.be (T. Van Loy), Jozef.VandenBroeck@bio.kuleuven.be (J. Vanden Broeck).

¹ These authors contributed equally.

Fig. 1. Schematic representation of the LGR types. 7TM=seven transmembrane domain; CTE=C-terminal endodomain; CRR=cysteine-rich region; HR=hinge region; LDLa=low density lipoprotein receptor domain class A; LRR=leucine-rich repeat.

functionally characterized, such as the mammalian relaxin receptors [27,37] and glycoprotein hormone receptors, and fruit fly LGR2 [47].

Using genome databases available at the NCBI platform (National Center for Biotechnology Information) (www.ncbi.nlm.nih.gov) and Ensembl (www.ensembl.org) and a variety of genome project databases from individual species, such as the Californian purple sea urchin (www.spbase.org) or the waterflea (wfleabase.org), we predicted new LGR sequences in an attempt to generate a more comprehensive phylogenetic tree containing all presently known and predicted LGRs. This comparative analysis will contribute to a better understanding of the structural diversity and evolution of LGRs in Metazoa.

2. Methods

Sequences were found either by keyword searches or by using the (PSI-)BLAST algorithm against all known LGRs. Protein sequences were checked for appropriate domains using the internet software program SMART [39,60]. Possible incorrectly in silico predicted sequences were manually checked at the genomic DNA level and, where possible, repredicted using softberry (<http://linux1.softberry.com/berry.phtml>) FGENESH+. Sequences were subsequently manually corrected using the “six-frame translation” software of Zbio.net (<http://molbiol.ru/index.php?langid=en>). By searching for protein domains and through comparison with the most closely related LGRs, aligned using ClustalX2 [25,38], exons were identified. New manually predicted sequences were added in a supplementary file (Suppl. File 1B). To check and further complement previous search results, novel sequences were used as additional queries.

Phylogenetic trees were generated using MUSCLE alignments [18] and PhyML [22] (WAG Substitution Model, 500 bootstraps) at the Phylogeny.fr server [15]. Trees were rendered with the online

TreeDyn application [11] and exported to MEGA [64]. Only complete or nearly complete sequences were used to avoid incorrect clustering as much as possible.

3. Results

3.1. Three distinctive types of LGRs exist

As with previously generated trees [68], three distinctly organized receptor types can be identified from our phylogenetic analysis. For a comprehensive overview, a phylogenetic tree generated based on a representative selection of LGR sequences from all phyla known to contain these receptors can be viewed in Suppl. File 2. A simplified version of this tree can be found in Fig. 2. When viewing this tree it can be concluded that, in addition to the vertebrates, arthropods (the water flea *Daphnia arenata* and *Drosophila melanogaster*) and echinoderms (the sea urchin *Strongylocentrotus purpuratus*) also possess one or more members of each LGR type. The most basal animal phylum in which LGRs have been described is *Cnidaria*. In *Hydra magnipapillata* and *Nematostella vectensis*, members of this phylum, one sequence of the type A and type B LGR groups have been identified. In molluscan species, up to now, a single type B (in the bivalve *Crassostrea gigas*) [23] and a single type C LGR (in the snail *Lymnaea stagnalis* [65]) were uncovered.

3.2. Type A LGRs

The LGRs belonging to type A typically have 7–9 LRRs in their ectodomain. Furthermore, they have a long hinge region, containing the consensus sequence LxxAxLTYPxHCCAF at the beginning of the hinge and the sequence VxCxPxPDAFNPCEDIMGYxFLRV at the end of the hinge (Suppl. File 3). In addition to the two cysteines in each of the consensus sequences, two more cysteines are present

Fig. 2. General phylogram of LGRs. The LGRs can be divided into three major, structurally distinct types (A, B and C). The type C LGRs consist of two distinct subtypes: C1 and C2. The LGRs found in Placozoa may represent an ancestral form of these receptors. This phylogram represents the simplified version of the tree displayed in [Suppl. File 2](#).

in the hinge region. Likely, those six cysteines form three disulfide bridges (S–S) stabilizing the entire structure. The hinge regions in the LGR sequences found in the sea anemones *Anthopleura* and *Nematostella* are further characterized by the occurrence of glycine- and proline-rich imperfect tandem repeats.

We generated a phylogenetic tree with all known type A LGRs and (nearly) complete predictions ([Suppl. File 4](#)). Incomplete sequences were excluded from the analysis if more than one full exon (as determined by comparison to sequences belonging to the same class of animals) was missing.

The analysis clearly shows the presence of type A LGRs throughout most animal phyla with the exception of Lophotrochozoa. In vertebrate genomes, multiple type A LGR sequences can be detected. In contrast, the analyzed invertebrates each seem to possess only one type A LGR, with the rare exception of *Tribolium castaneum* for which two were predicted from the genome.

Saccoglossus kowalewskii, a species of acorn worm belonging to the hemichordate phylum, has a type A LGR in its genome. This receptor's closest relative is the LGR1 of echinoderms, generally considered to be the sister group of the Hemichordata. The urochordate *Ciona* only has one type A LGR in its genome. Our general phylogenetic analysis ([Suppl. File 2](#)) places this receptor in a branch next to the cephalochordate sequences. In these animals, an incomplete type A LGR was found in the lancelet *Branchiostoma belcheri tsingtauense* (ACM66673) and a hypothetical LGR-like protein in *Branchiostoma floridae* (XM.002610196).

All arthropod sequences form a separate clade, with all closely related sequences clustering together. At present, the waterflea *D. arenata* is the only crustacean in which a type A LGR was found. Note that out of the sequences from twelve sequenced *Drosophila* species [13] only the sequence for *D. melanogaster* was added to our tree. Finally, the three nematode sequences (*Caenorhabditis elegans*, *Caenorhabditis briggsae* and *Brugia malayi*) and three cnidarian sequences (*H. magnipapillata*, *N. vectensis* and *Anthopleura elegantissima*) form separate clades.

3.3. Type B LGRs

Type B LGRs are characterized by the presence of 16–18 LRRs, roughly twice the number found in the other types of LGRs. They have medium length hinge regions compared to the type A and type C LGRs. At the beginning of the hinge, the sequence

IxxLxLYAYHCCxF is well conserved, while at the end of the hinge region the sequence IxCxPxPGPFxPCEDLFGSWxLRV can be found in most species. Unlike type A LGRs, LGR type B receptors do not have two extra cysteines in the hinge region. Therefore, only two disulfide bridges can potentially be formed ([Suppl. File 5](#)).

As can be concluded from [Supplementary Fig. 6](#), B-type LGRs are found in virtually all animal phyla. It is noteworthy that both predicted *Strongylocentrotus* and *Ciona* type B LGRs have shorter ectodomains and contain fewer LRRs than other type B representatives. Type B LGR sequences were also identified in the crustacean and tick species, *D. arenata* and *Ixodes scapularis*. These cluster together with the other arthropod type B LGRs. In fish, only complete sequences of LGR4 and LGR6 have been found. Regarding LGR5, a partial sequence was found only of *Takifugu rubripes*, containing the hinge region and part of the transmembrane domain (Ensembl database: ENSTRUP00000003360).

3.4. Type C LGRs

The total number of LRRs in type C LGRs is similar to this of the type A LGRs. One unique feature of this type is the presence of an LDLa motif, N-terminally of the LRRs. The arthropod and the chordate LGRs of type C and the sea urchin LGR7, only have a single such motif. Most interestingly however, the echinoderm (sea urchin LGR4, LGR5 and LGR8) and molluscan (*Lymnaea* LGR) LGRs of type C, on the other hand, harbor ten, six, five and twelve LDLa modules, respectively. A similar LGR, containing ten LDLa motifs, is present in the human body lice (*Pediculus humanis corporis*). From now on, we will refer to this type as type C2 LGRs. The more general type containing only one LDLa will be referred to as type C1.

The hinge region of the C-type LGRs is very short, compared to the other two types and has the consensus sequence LxxLxHlxFDR-FxYCxYAPHV at its beginning and CxPxxDGISxxEDLLSNxVLRV at its end. The type C1 LGR hinges contain only two cysteines, while these of type C2 LGRs contain four cysteines ([Suppl. File 7](#)). In addition, some conserved amino acids differ between the two subtypes. In consequence, LGRs of type A, B, C1 and C2 can be distinguished by their hinge region ([Suppl. File 1D](#)).

Whereas (partial) sequences of type C LGRs were found in *Aedes aegyptii* (XM.001650299), *Nasonia vitripennis* (XM.001600282) and *Acyrtosiphon pisum* (XM.001942974), no type C LGR was found in

the genome of, for instance, *Bombyx mori*. In *S. kowalevskii*, one type C1 LGR has been detected so far.

An exception to the “normal” types of LGRs, are those of *Trichoplax*, the sole genus of the Placozoa, basal multicellular organisms consistent of three cell layers. In their genome, we have predicted eight LGR-like sequences, which are placed near the type C LGRs by our alignments. They do have a short hinge region, but seem to lack an LDLa motif.

4. Discussion

As mentioned in Section 3, the Cnidarian genomes each hold one sequence of the type A and one of the type B LGR groups, which places the receptor's origin above 800–1000 million years ago [70]. Based on this observation, one may hypothesize that the third type of LGRs originated later in evolution, out of another type of LGR (by gene duplication) to which an LDLa encoding sequence was merged. This view, however, might be an oversimplification of the evolutionary process, since the hinge region is also different between all three types. Since the molluscan sequence analyses clearly show the presence of type C LGRs, this type had already originated before the divergence of protostomes and deuterostomes. Another factor which impedes a clear notion of the type C LGR origin are the sequences from *Trichoplax adhaerens* resembling this receptor type at the level of the hinge region. Generally, this placozoan organism is thought to be at the root of the animal kingdom, but its exact phylogenetic relationship is still a matter of debate [14].

4.1. Type A LGRs

Concerning the currently observed lack of a lophotrochozoan type A LGR, it is probably too early to draw final conclusions, given the disproportionately lower number of genome data that are currently available from this group of metazoans, although it is also quite plausible that one or more LGR types may have been lost in some branches of the evolutionary tree. This latter assumption can indeed be made in the case of the nematode *C. elegans*. This small worm has one of the first sequenced [6] and most studied genomes known, but only a single LGR has been found [36], which means that this nematode likely lost all other LGRs during evolution, as previously suggested [26]. The phylogenetic analysis considers this nematode LGR as most closely related to the type A receptors. However, detailed analysis of this receptor showed that although the hinge region clearly displays type A features, the membrane-spanning domain seems to be closer related to the type B receptors. These ambiguous properties may also account for the peculiar placement of the nematode receptors within the phylogenetic tree (Suppl. File 2).

Hitherto, in all genomes except for the vertebrate ones and this of *T. castaneum*, only one type A LGR can be identified. In vertebrates, three type A LGRs exist, namely the glycoprotein hormone receptors. These extensively studied receptors likely originate from two whole genome duplications within the vertebrate subphylum [53]. Smits and colleagues suggested that TSHR and FSHR evolved from a more recent common ancestor compared to LH/CGR, based on a mutagenesis study of TSHR and FSHR. A single substitution of the fifth amino acid of the first LRR by a leucine (L) or alanine (A) in TSHR, yielded slight sensitivity to FSH, while a single point mutation (L55Y, with Y being tyrosine) of the same residue in FSHR completely abolished sensitivity and affinity to FSH [61]. Earlier, the same conclusion was drawn from a study of several substitution mutants of the “determinant loop”, the sequence which is thought to confer binding specificity, of the ligands [7,21]. Based on the sequence similarity of the three glycoprotein hormone receptors

with respect to each other and to the (single) invertebrate LGR A types, it is difficult to account for this hypothesis. While the closest relative in non-chordates (*Strongylocentrotus*) most closely resembles TSHR, other non-chordate type A LGRs seem more similar to FSHR or LH/CGR. Nevertheless, in general, they resemble all three glycoprotein hormone receptors, one slightly more than the other.

While in the tetrapod branch the known taxonomy is respected (primates cluster with primates, etc.), the bony fish show a more remarkable branching. Whereas the sequences of closely related fish such as the Nile tilapia (*Oreochromis niloticus*) and the European seabass (*Dicentrarchus labrax*), both belonging to the Perciformes, cluster together in the FSHR-branch and the TSHR-branch, they do not group together in the LH/CGR-branch. The Medaka (*Oryzias latipes*) on the other hand, is in the same clade as the Nile tilapia in both the FSHR-branch as the LH/CGR-branch. However, it forms a different branch altogether for the TSHR sequence. This is in part explained by recent work of Chauvigne and colleagues, who investigated the synteny of these genes in vertebrates [9]. After the bony fish specific whole genome duplication (3R) [19], all fish would have retained the same copy of the *fshr* gene, with conserved synteny to the tetrapod lineage and would have retained either one or the other copy of the *tshr* paralogues. Due to the absence of duplicates in the genome and their synteny analyses, the authors suggested that all fish retained the same copy of *lhcgrr*. Therefore, it is hypothesized that the occurrence of two different *lhcgrr* genes in bony fish is the result of a gene conversion [9].

Low bootstrap values within clades in all generated trees are usually the result of extremely high sequence similarity of the LGRs between species within this clade, meaning that a sequence of a particular species can branch with the sequences of many of its close neighbors. This is especially true within the vertebrate clades, where differences can be as small as a few amino acid changes. This also implies that results can get slightly distorted when an exon or some amino acids are missing in a predicted sequence.

4.2. Type B LGRs

The best characterized type B receptor is certainly the fruit fly LGR2 or bursicon receptor [43,47]. This receptor is responsible for tanning, cuticle hardening and wing expansion after adult emergence. It is of interest to note that the type B LGRs of echinoderms (*Asterina* and *Strongylocentrotus*), mollusks (*Crassostrea*) and cnidarians (*Nematostella* and *Hydra*) are more similar to the arthropod type B LGRs than to chordate B-type LGRs (Suppl. File 6). Except for *H. magnipapillata*, which does not have a medusa stage (unlike some other members of the *Hydra* genus) these species all undergo metamorphosis from a larval stage to the adult stage. Whether B LGRs play a role during this process, analogous to the situation in insects, still remains an exciting hypothesis that requires more investigation. *Hydra* and *Nematostella*, belong to the Hydrozoa and Anthozoa respectively, two different classes of the Cnidaria that diverged from each other as much as protostomes diverged from deuterostomes [58]. This explains the relatively large distance between the sequences of these two species (Suppl. File 6).

Interestingly, the B-type receptors predicted for *Strongylocentrotus* and *Ciona* display more restricted ectodomains featuring a lower number of LRRs in comparison to other type B LGRs. The cause of this may lie in the preliminary nature of these predictions. Searches in the genome data indeed yielded additional LRRs, but assembly of the exons proved difficult, since the standard intron–exon junctions of “gt – ag” were not applicable to these putative exons. Further cDNA or EST data will be needed to overcome this obstacle.

As was discussed in the case of the type A LGRs, two genome duplication events after the branching off from the Urochordata probably tripled the sole type B LGR in invertebrates to three in

vertebrates. The similarity between the three type B LGRs is comparable to the similarity between the type A LGRs. It is of interest to note that in fish, only for *T. rubripes* an LGR5 orthologue could be identified. Perhaps, in future analyses of the genome or proteome of other fish this receptor too can be found.

Regarding the function of three B-type receptors in vertebrates, a number of studies have been published with a general tenor of their involvement in developmental processes. Research on LGR4 knockout mice revealed critical roles in embryonic growth [45], development of the male reproductive tract [46], renal development [31], keratinocyte development [32], hair formation [50], eye development [30,71] and bone formation [44]. Mice without LGR5 exhibit a full neonatal lethality, associated with ankyloglossia, the fusion of the tongue to the floor of the mouth cavity [51]. Furthermore, LGR5 and 6 have been demonstrated to be markers for epithelial stem cells [3,4]. LGR4 is localized in the small intestine where it is required for the final differentiation of Paneth cells [52].

4.3. Type C LGRs

According to our phylogenetic tree the two type C LGRs that exist in chordates (LGR7 and LGR8) resemble each other more than any other type C LGR (Suppl. File 8), with nearly 64% sequence similarity between human LGR7 and LGR8, also known as relaxin receptor 1 and 2 (RXFP1 and RXFP2). In contrast, the two fruit fly type C LGRs only display about 40% similarity, which is about the same as their similarity to the human type C LGRs. Although the three type C LGRs predicted in *B. floridae*, a very primitive chordate, do not fully match the two vertebrate types, they do share a mutual similarity of almost 70%, despite the fact that small parts of both genes have not been sequenced yet. This would be consistent with a recent gene duplication. Unfortunately, thus far no type C LGR has been identified in the genome of the sea squirt, a member of the Urochordata. Since the three LGR types have already been identified in arthropods, echinoderms and vertebrates, type C would have been expected to be present in the common ancestor of all chordates. While it is too early to draw definitive conclusions, this could mean that this type of receptors may have been lost in *Ciona*, as probably is the case for *C. elegans*.

The most closely related non-chordate in our tree, the sea urchin, a deuterostome, possesses both type C LGRs. At least three of these type C2 receptors, are present in the sea urchin genome (Suppl. File 8). This hinge region holds a motif that resembles the chordate motif and is also similar to the type A and B LGRs, albeit a lot shorter. In the body lice too, a type C2 is also present. These data might be interpreted as suggesting the possible existence of two kinds of type C LGRs (giving rise to type C1 and C2) before the divergence into proto- and deuterostomes. On the other hand, the current knowledge does not allow ruling out that this new receptor subtype has evolved independently in the respective phyla. More genomic data from other species are required to shift the discussion toward either option.

In addition, separate gene duplications have probably taken place in the chordate and arthropod lineage, since the two paralogues are more similar to each other than to their orthologues. Contrary to the type A and B LGRs, however, only two type C1 exist in the genome of vertebrates. Possibly, the paralogues of this gene were rapidly lost after the genome duplication events, retaining only two copies.

The eight LGR sequences identified from *T. adhaerens* are all very similar and, like C-type LGRs, they feature a short hinge region. However, these receptors lack an LDLA module, a C-type hallmark. Searches upstream of the predicted genes also revealed no such motifs. Therefore, the LGRs found in *Trichoplax* might be derived from the predecessors of the type C receptors, before their fusion with LDLA modules. Considering the organism's ancient lineage,

one can hypothesize that all three LGR types are derived from the genes encoding LGRs as they probably occurred in the common ancestor between Placozoa and Eumetazoa. However, extensive genomic information of Porifera and Ctenophora would be needed to investigate this hypothesis further. Alternatively, *Trichoplax* may have secondarily lost the LDLA motif [48]. It is interesting to note that previous in silico studies in *Trichoplax* also showed a very deviant gene organization for other genes, namely the pacifastins [5]. Although many annotated gene families seem to be present in *Trichoplax* [62], this organism just might prove to be an outcast with respect to most other species.

4.4. Co-evolution of LGRs and their ligands

It is clear that this subgroup of rhodopsin-related G protein-coupled receptors displays a very strong conservation throughout the animal kingdom, exemplified by the presence of at least two types of LGRs in cnidarian species and a third one in Placozoa. Such extraordinary preservation is indicative for its importance. A couple of years ago, Sudo et al. found a ligand for *Drosophila* LGR1 that is related to the glycoprotein hormones [63]. While this is to date, with respect to LGRs, the only clear case upon which a discussion on ligand-receptor co-evolution crossing the bridge between proto-stomes and deuterostomes can be built, it does offer opportunities for further speculation. Apart from all vertebrate species, glycoprotein hormone related sequences can be found in the lancelet [17], the American dog tick [16], the owl limpet *Lottia gigantea* [69], the malaria mosquito, the red flour beetle, the nematode *C. briggsae*, the sea slug *Aplysia californica* and the sea urchin [68]. In addition, glycoprotein hormone sequences can be found in numerous other animals such as the desert locust *Schistocerca gregaria* [2], the hemichordate *S. kowalevskii* (LOC100370267, LOC100313754), the urochordate *Ciona intestinalis* (LOC100178300) and *Ciona savignyi* (CAR94706), the deer tick (CAR94694, CAR95342), the parasitic nematode *Trichinella spiralis* (CAR94700, CAR95342) *C. elegans* (ABK78775, CAR95352) and *B. malayi* (CAR94701, CAR95351). The mammalian glycoprotein hormone subunits and the invertebrate glycoprotein hormone subunits are probably derived from common ancestral sequences that already occurred before the divergences between ecdysozoan, lophotrochozoan and deuterostome lineages [55]. While type A LGRs have been found in the genome of cnidarian species, no glycoprotein hormone related sequences could so far be detected in these animals using pattern and BLAST searches.

The evolution of relaxin family genes in the vertebrate lineage has already been extensively studied [20,26,28,56,74]. Syntenic mapping of the relaxin family genes, showed that all mammalian relaxin genes likely evolved from three separate ancestral genes in the common ancestor of tetrapods and teleosts: (1) AncRFLA, giving rise to INSL5, (2) AncRFLB, the ancestral gene of INSL4, INSL6, RLN1 and RLN2, and (3) AncRFLC, which later diverged into RLN3 and INSL3 [68]. This last duplication took place even before the divergence of tetrapods and bony fish, indicating that the direct common ancestor already had four relaxin family genes [20]. Since RLN3 was the only gene which displayed strong evolutionary constraint, comparable to insulin, it has been suggested that RLN3 is the ancestral relaxin gene [20,74]. Apart from LGR7, RLN3 can also interact with GPCR135 and GPCR142, two non-LGR relaxin receptors [41,42]. It was proposed that RLN3 is the endogenous ligand of GPCR135, based on their similar tissue distribution. It was also suggested that RLN3 and GPCR135 co-evolved, since they both appear to have several gene copies in fish [10,73], and that the activation of LGR7 and LGR8 by relaxins was acquired later in mammalian evolution [73]. However, it must be noted that of the six copies found in puffer fish, only two are actual gene copies of RLN3 (which arose after 3R). The other four proved to be orthologues of INSL5, RLN2 and INSL3 [20]. Also, Park et al. showed that both relaxin 3 gene

(rln3a and rln3b, also called RFLC2 and RFLC1, respectively) products of zebrafish were able to activate both hLGR7 and hLGR8, suggesting that relaxins were already capable of activating type C LGRs prior to the appearance of mammals [56].

Relaxins belong to the insulin superfamily of peptides, which includes insulins, relaxins, insulin-like peptides (ILPs) and insulin-like growth factors (IGFs). No relaxins have unequivocally been detected in protostomian invertebrates, so far. However, it has already been suggested that ILPs might be the ligands of the invertebrate type C LGRs [75]. Many insulin-like peptides have already been detected in numerous insects belonging to different orders, such as the honey bee, locusts, mosquitoes, moths, beetles and flies [1,12,24,35,40,59,66,72,75]. ILPs can also be found in the jewel wasp (XM.001601680) and the pea aphid (XM.001950003, XM.001949218, XM.001949403). Besides insects, ILPs are found in other invertebrate phyla. For instance, three ILPs are present in *Ciona* [54], forty in *C. elegans* [29], four in the owl limpet [69] and one in the starfish *Asterina pectinifera* [49]. In the sea urchin, another echinoderm species, one ILP can be identified in the genome (XM.001175486). In the mollusks *A. californica* and *L. stagnalis*, at least three (AF364181, Q9NDE7, DQ479393) and four (CAA30043, AAB28954, AAA09966, AAB46831) ILPs can be detected, respectively. When performing a pattern search in the genome of *Trichoplax* one ILP can be found (XP.002111669), which resembles *Drosophila* ILP6. Since no orthologues of GPCR135 or GPCR142 are present in invertebrates and since several ILPs are present in invertebrate species, it is possible that one or more of the ILPs might bind invertebrate type C LGRs. Although they lack the characteristic relaxin motif, RXXXXXXI/V, some ILPs contain part of this motif. In addition, INSL3 (which also lacks the relaxin motif) is the endogenous ligand for LGR8 [33].

In every phylum containing a type A or a type C LGR, except for the Cnidaria, sequences related to their candidate ligands (glycoprotein hormones or ILPs, respectively) can be found. While this observation offers by no means full evidence for an existing co-evolution of the respective receptor/ligand pairs, it may constitute a good starting point for future research, which will likely reveal more ligand/LGR couples in a wide range of animal species.

Finally, assessing a path of co-evolution proves very difficult for the type B LGRs. In *Drosophila*, LGR2 was identified as the receptor for bursicon [47], a heterodimer of glycoprotein hormone-like cystine knot subunits which is well conserved in insects, mollusks and echinoderms, but not in chordates. In follow-up of this discovery, it was postulated that the ligand for this type of chordate receptor's could have been replaced by another class of ligands in the course of this evolutionary process [67]. Indeed, recent studies have provided evidence for the role of R-spondins as high affinity ligands for the LGR4 and 5 receptors in mouse and human [8]. Most remarkably, while the activation of these receptors increases Wnt-dependent LRP6 phosphorylation, the signal transduction does not seem to involve heterotrimeric G proteins or beta-arrestin. These results suggest there has been a change of ligand, as well as a change of signaling pathway, in the vertebrate LGR type B receptors when compared to the other receptors in this class. An interesting example of such a ligand change has recently been described for *Drosophila* SPR (sex peptide receptor). This GPCR is now thought to be activated ancestrally by the myoinhibiting peptides, but during the course of evolution, in *Drosophila*, has adopted an additional ligand, sex peptide [34,57].

5. Conclusions

This study describes and discusses the evolutionary aspects of all known and predicted LGRs. It is clear that LGRs are widely distributed and well preserved throughout the animal kingdom,

which is probably a testimony of their vital importance. LGRs have now been identified (and predicted from genome data) in Placozoa, Cnidaria, Nematoda, Arthropoda, Mollusca, Hemichordata, Echinodermata and Chordata. Based on their number of LRRs, their type-specific hinge region and the presence or absence of an LDLa motif, three major types of LGRs can be distinguished. Based on the number of LDLa motifs, type C LGRs can be further divided into two subtypes: a C1 type containing only one LDLa, which is present in nearly all phyla, and a C2 type containing multiple LDLa motifs and a hinge region distinct from those of the other LGR classes. So far, C2-type LGRs have only been found in echinoderms and mollusks, as well as in one insect species (*P.h. corporis*). In addition, eight putative LGRs were found in *Trichoplax*, which display similarity with type C LGRs, but lack the LDLa motif. Whether these receptors represent an ancient form of the type C LGRs, which occurred before the fusion of the LDLa motif to the rest of the receptor, or whether *Trichoplax* has lost this LDLa motif, remains unclear. While the evolutionary relationships of the type C LGRs seem very complex, the type A and B LGRs likely underwent two duplication events (and the loss of one of the resulting gene copies) giving rise to three members each in many species of the vertebrate lineage.

Authors' contributions

M.B. Van Hiel has performed the sequence predictions and phylogenetic analyses and drafted the manuscript. H.P. Vandersmissen performed phylogenetic analyses and co-drafted the manuscript. T. Van Loy checked the results of all predictions and analyses and contributed to the drafting of the manuscript. J. Vanden Broeck is the senior academic author who contributed to the design of the study, the discussion of the results and the correction of the manuscript. All authors have read and approved the manuscript.

Acknowledgements

This research was financially supported by the "Belgian program on Interuniversity Poles of Attraction (IUAP/PAI P6/14)", the Research Foundation of Flanders (FWO-Flanders) and the K.U. Leuven Research Foundation (GOA/11/002). M.B. Van Hiel obtained a Ph.D. Fellowship from the "Instituut voor de aanmoediging van Innovatie door Wetenschap en Technologie in Vlaanderen" (IWT) and H.P. Vandersmissen was supported by a FLOF grant (K.U.Leuven).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.peptides.2011.11.004.

References

- [1] Badisco L, Claeys I, Van Hiel M, Clynen E, Huybrechts J, Vandersmissen T, et al. Purification and characterization of an insulin-related peptide in the desert locust, *Schistocerca gregaria*: immunolocalization, cDNA cloning, transcript profiling and interaction with neuroparsin. *J Mol Endocrinol* 2008;40:137–50.
- [2] Badisco L, Huybrechts J, Simonet G, Verlinden H, Marchal E, Huybrechts R, et al. Transcriptome analysis of the desert locust central nervous system: production and annotation of a *Schistocerca gregaria* EST database. *PLoS One* 2011;6:e17274.
- [3] Barker N, Clevers H. Leucine-rich repeat-containing G-protein-coupled receptors as markers of adult stem cells. *Gastroenterology* 2010;138:1681–96.
- [4] Barker N, van Es JH, Kuipers J, Kujala P, van den BM, Cozijnsen M, et al. Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* 2007;449:1003–7.
- [5] Breugelmans B, Simonet G, van Hoef V, Van Soest S, Vanden Broeck J. Identification, distribution and molecular evolution of the pacifastin gene family in Metazoa. *BMC Evol Biol* 2009;9:97.
- [6] *C.elegans* genome consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 1998;282:2012–8.

- [7] Campbell RK, Bergert ER, Wang Y, Morris JC, Moyle WR. Chimeric proteins can exceed the sum of their parts: implications for evolution and protein design. *Nat Biotechnol* 1997;15: 439–43.
- [8] Carmon KS, Gong X, Lin Q, Thomas A, Liu Q. R-spondins function as ligands of the orphan receptors LGR4 and LGR5 to regulate Wnt/ β -catenin signaling. *Proc Natl Acad Sci U S A* 2011;108:11452–7.
- [9] Chauvigne F, Tingaud-Sequeira A, Agulleiro MJ, Calusinska M, Gomez A, Finn RN, et al. Functional and evolutionary analysis of flatfish gonadotropin receptors reveals cladal- and lineage-level divergence of the teleost glycoprotein receptor family. *Biol Reprod* 2010;82:1088–102.
- [10] Chen J, Kuei C, Sutton SW, Bonaventure P, Nepomuceno D, Eriste E, et al. Pharmacological characterization of relaxin-3/INSL7 receptors GPCR135 and GPCR142 from different mammalian species. *J Pharmacol Exp Ther* 2005;312:83–95.
- [11] Chevenet F, Brun C, Banuls AL, Jacq B, Christen R. TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* 2006;7:439.
- [12] Claeys I, Poels J, Simonet G, Franssens V, Van Loy T, Van Hiel MB, et al. Insect neuropeptide and peptide hormone receptors: current knowledge and future directions. *Vitam Horm* 2005;73:217–82.
- [13] Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, et al. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 2007;450:203–18.
- [14] Dellaportia SL, Xu A, Sagasser S, Jakob W, Moreno MA, Buss LW, et al. Mitochondrial genome of *Trichoplax adhaerens* supports placozoa as the basal lower metazoan phylum. *Proc Natl Acad Sci U S A* 2006;103:8751–6.
- [15] Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 2008;36:W465–9.
- [16] Donohue KV, Khalil SM, Ross E, Grozinger CM, Sonenshine DE, Michael RR. Neuropeptide signaling sequences identified by pyrosequencing of the American dog tick synganglion transcriptome during blood feeding and reproduction. *Insect Biochem Mol Biol* 2010;40:79–90.
- [17] Dos Santos S, Bardet C, Bertrand S, Escriva H, Habert D, Querat B. Distinct expression patterns of glycoprotein hormone- α 2 and - β 5 in a basal chordate suggest independent developmental functions. *Endocrinology* 2009;150:3815–22.
- [18] Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792–7.
- [19] Finn RN, Kristoffersen BA. Vertebrate vitellogenin gene duplication in relation to the “3R hypothesis”: correlation to the pelagic egg and the oceanic radiation of teleosts. *PLoS One* 2007;2:e169.
- [20] Good-Avila SV, Yegorov S, Harron S, Bogerd J, Glen P, Ozon J, et al. Relaxin gene family in teleosts: phylogeny, syntenic mapping, selective constraint, and expression analysis. *BMC Evol Biol* 2009;9:293.
- [21] Grossmann M, Szkudlinski MW, Wong R, Dias JA, Ji TH, Weintraub BD. Substitution of the seat-belt region of the thyroid-stimulating hormone (TSH) β -subunit with the corresponding regions of choriogonadotropin or folliculin confers luteotropic but not folliculin activity to chimeric TSH. *J Biol Chem* 1997;272:15532–40.
- [22] Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 2003;52:696–704.
- [23] Herpin A, Badaricotti F, Rodet F, Favre P. Molecular characterization of a new leucine-rich repeat-containing G protein-coupled receptor from a bivalve mollusc: evolutionary implications. *Biochim Biophys Acta* 2004;1680:137–44.
- [24] Hetru C, Li KW, Bulet P, Laguerre M, Hoffmann JA. Isolation and structural characterization of an insulin-related molecule, a predominant neuropeptide from *Locusta migratoria*. *Eur J Biochem* 1991;201:495–9.
- [25] Higgins DG, Sharp PM. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* 1988;73:237–44.
- [26] Hsu SY. New insights into the evolution of the relaxin-LGR signaling system. *Trends Endocrinol Metab* 2003;14:303–9.
- [27] Hsu SY, Nakabayashi K, Nishi S, Kumagai J, Kudo M, Sherwood OD, et al. Activation of orphan receptors by the hormone relaxin. *Science* 2002;295: 671–4.
- [28] Hsu SY, Semyonov J, Park JI, Chang CL. Evolution of the signaling system in relaxin-family peptides. *Ann N Y Acad Sci* 2005;1041:520–9.
- [29] Husson SJ, Mertens I, Janssen T, Lindemans M, Schoofs L. Neuropeptidergic signaling in the nematode *Caenorhabditis elegans*. *Prog Neurobiol* 2007;82:33–55.
- [30] Jin C, Yin F, Lin M, Li H, Wang Z, Weng J, et al. GPR48 regulates epithelial cell proliferation and migration by activating EGFR during eyelid development. *Invest Ophthalmol Vis Sci* 2008;49:4245–53.
- [31] Kato S, Matsubara M, Matsuo T, Mohri Y, Kazama I, Hatano R, et al. Leucine-rich repeat-containing G protein-coupled receptor-4 (LGR4, Gpr48) is essential for renal development in mice. *Nephron Exp Nephrol* 2006;104:e63–75.
- [32] Kato S, Mohri Y, Matsuo T, Ogawa E, Umezawa A, Okuyama R, et al. Eye-open at birth phenotype with reduced keratinocyte motility in LGR4 null mice. *FEBS Lett* 2007;581:4685–90.
- [33] Kawamura K, Kumagai J, Sudo S, Chun S-Y, Pisarska M, Morita H, et al. Paracrine regulation of mammalian oocyte maturation and male germ cell survival. *Proc Natl Acad Sci U S A* 2004;101:7323–8.
- [34] Kim YJ, Bartalska K, Audsley N, Yamanaka N, Yapici N, Lee JY, et al. MIPs are ancestral ligands for the sex peptide receptor. *Proc Natl Acad Sci U S A* 2010;107:6520–5.
- [35] Krieger MJ, Jahan N, Riehle MA, Cao C, Brown MR. Molecular characterization of insulin-like peptide genes and their expression in the African malaria mosquito, *Anopheles gambiae*. *Insect Mol Biol* 2004;13:305–15.
- [36] Kudo M, Chen T, Nakabayashi K, Hsu SY, Hsueh AJ. The nematode leucine-rich repeat-containing, G protein-coupled receptor (LGR) protein homologous to vertebrate gonadotropin and thyrotropin receptors is constitutively active in mammalian cells. *Mol Endocrinol* 2000;14:272–84.
- [37] Kumagai J, Hsu SY, Matsumi H, Roh JS, Fu P, Wade JD, et al. INSL3/Leydig insulin-like peptide activates the LGR8 receptor important in testis descent. *J Biol Chem* 2002;277: 31283–6.
- [38] Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics* 2007;23:2947–8.
- [39] Letunic I, Doerks T, Bork P. SMART 6: recent updates and new developments. *Nucleic Acids Res* 2009;37:D229–32.
- [40] Li B, Predel R, Neupert S, Hauser F, Tanaka Y, Cazzamali G, et al. Genomics, transcriptomics, and peptidomics of neuropeptides and protein hormones in the red flour beetle *Tribolium castaneum*. *Genome Res* 2008;18:113–22.
- [41] Liu C, Chen J, Sutton S, Roland B, Kuei C, Farmer N, et al. Identification of relaxin-3/INSL7 as a ligand for GPCR142. *J Biol Chem* 2003;278:50765–70.
- [42] Liu C, Eriste E, Sutton S, Chen J, Roland B, Kuei C, et al. Identification of relaxin-3/INSL7 as an endogenous ligand for the orphan G-protein-coupled receptor GPCR135. *J Biol Chem* 2003;278:50754–64.
- [43] Luo CW, Dewey EM, Sudo S, Ewer J, Hsu SY, Honegger HW, et al. Bursicon, the insect cuticle-hardening hormone, is a heterodimeric cystine knot protein that activates G protein-coupled receptor LGR2. *Proc Natl Acad Sci U S A* 2005;102:2820–5.
- [44] Luo J, Zhou W, Zhou X, Li D, Weng J, Yi Z, et al. Regulation of bone formation and remodeling by G-protein-coupled receptor 48. *Development* 2009;136:2747–56.
- [45] Mazerbourg S, Bouley DM, Sudo S, Klein CA, Zhang JV, Kawamura K, et al. Leucine-rich repeat-containing G protein-coupled receptor 4 null mice exhibit intrauterine growth retardation associated with embryonic and perinatal lethality. *Mol Endocrinol* 2004;18:2241–54.
- [46] Mendive F, Laurent P, Van SG, Skarnes W, Pochet R, Vassart G. Defective post-natal development of the male reproductive tract in LGR4 knockout mice. *Dev Biol* 2006;290:421–34.
- [47] Mendive FM, Van Loy T, Claeys S, Poels J, Williamson M, Hauser F, et al. *Drosophila* molting neurohormone bursicon is a heterodimer and the natural agonist of the orphan receptor DLGR2. *FEBS Lett* 2005;579:2171–6.
- [48] Miller DJ, Ball EE. Animal evolution: Trichoplax, trees, and taxonomic turmoil. *Curr Biol* 2008;18:R1003–5.
- [49] Mita M, Yoshikuni M, Ohno K, Shibata Y, Paul-Prasanth B, Pitchayawasin S, et al. A relaxin-like peptide purified from radial nerves induces oocyte maturation and ovulation in the starfish, *Asterina pectinifera*. *Proc Natl Acad Sci U S A* 2009;106:9507–12.
- [50] Mohri Y, Kato S, Umezawa A, Okuyama R, Nishimori K. Impaired hair placode formation with reduced expression of hair follicle-related genes in mice lacking Lgr4. *Dev Dyn* 2008;237:2235–42.
- [51] Morita H, Mazerbourg S, Bouley DM, Luo CW, Kawamura K, Kuwabara Y, et al. Neonatal lethality of LGR5 null mice is associated with ankyloglossia and gastrointestinal distension. *Mol Cell Biol* 2004;24:9736–43.
- [52] Mustata RC, Van Loy T, Lefort A, Libert F, Strollo S, Vassart G, et al. Lgr4 is required for Paneth cell differentiation and maintenance of intestinal stem cells ex vivo. *EMBO Rep* 2011;12:558–64.
- [53] Ohno S. The one-to-four rule and paralogues of sex-determining genes. *Cell Mol Life Sci* 1999;55:824–30.
- [54] Olinski RP, Dahlberg C, Thorndyke M, Hallbook F. Three insulin-relaxin-like genes in *Ciona intestinalis*. *Peptides* 2006;27:2535–46.
- [55] Park JI, Semyonov J, Chang CL, Hsu SY. Conservation of the heterodimeric glycoprotein hormone subunit family proteins and the LGR signaling system from nematodes to humans. *Endocrine* 2005;26:267–76.
- [56] Park JI, Semyonov J, Chang CL, Yi W, Warren W, Hsu SY. Origin of INSL3-mediated testicular descent in therian mammals. *Genome Res* 2008;18:974–85.
- [57] Poels J, Van Loy T, Vandersmissen HP, Van Hiel B, Van Soest S, Nachman RJ, et al. Myoinhibiting peptides are the ancestral ligands of the promiscuous *Drosophila* sex peptide receptor. *Cell Mol Life Sci* 2010;67:3511–22.
- [58] Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, Salamov A, et al. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* 2007;317:86–94.
- [59] Riehle MA, Fan Y, Cao C, Brown MR. Molecular characterization of insulin-like peptides in the yellow fever mosquito, *Aedes aegypti*: expression, cellular localization, and phylogeny. *Peptides* 2006;27:2547–60.
- [60] Schultz J, Milpetz F, Bork P, Ponting CP. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* 1998;95:5857–64.
- [61] Smits G, Campillo M, Govaerts C, Janssens V, Richter C, Vassart G, et al. Glycoprotein hormone receptors: determinants in leucine-rich repeats responsible for ligand specificity. *EMBO J* 2003;22:2692–703.
- [62] Srivastava M, Begovic E, Chapman J, Putnam NH, Hellsten U, Kawashim T, et al. The Trichoplax genome and the nature of placozoans. *Nature* 2008;454:955–60.
- [63] Sudo S, Kuwabara Y, Park JI, Hsu SY, Hsueh AJ. Heterodimeric fly glycoprotein hormone- α 2 (GPA2) and glycoprotein hormone- β 5 (GPB5) activate fly leucine-rich repeat-containing G protein-coupled receptor-1 (DLGR1) and stimulation of human thyrotropin receptors by chimeric fly GPA2 and human GPB5. *Endocrinology* 2005;146:3596–604.
- [64] Tamura K, Dudley J, Nei M, Kumar S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007;24:1596–9.

- [65] Tensen CP, Van Kesteren ER, Planta RJ, Cox KJ, Burke JF, van HH, et al. A G protein-coupled receptor with low density lipoprotein-binding motifs suggests a role for lipoproteins in G-linked signal transduction. *Proc Natl Acad Sci U S A* 1994;91:4816–20.
- [66] Van de Velde S, Badisco L, Claeys I, Verleyen P, Chen X, Vanden Bosch L, et al. Insulin-like peptides in *Spodoptera littoralis* (Lepidoptera): detection, localization and identification. *Gen Comp Endocrinol* 2007;153:72–9.
- [67] Van Loy T, Van Hiel MB, Vandersmissen HP, Poels J, Mendive F, Vassart G, et al. Evolutionary conservation of bursicon in the animal kingdom. *Gen Comp Endocrinol* 2007;153:59–63.
- [68] Van Loy T, Vandersmissen HP, Van Hiel MB, Poels J, Verlinden H, Badisco L, et al. Comparative genomics of leucine-rich repeats containing G protein-coupled receptors and their ligands. *Gen Comp Endocrinol* 2008;155:14–21.
- [69] Veenstra JA. Neurohormones and neuropeptides encoded by the genome of *Lotia gigantea*, with reference to other mollusks and insects. *Gen Comp Endocrinol* 2010;167:86–103.
- [70] Waggoner BC, Collins AG. Reductio ad absurdum: testing the evolutionary relationships of Ediacaran and paleozoic problematic fossils using molecular divergence dates. *J Paleontol* 2004;78:51–61.
- [71] Weng J, Luo J, Cheng X, Jin C, Zhou X, Qu J, et al. Deletion of G protein-coupled receptor 48 leads to ocular anterior segment dysgenesis (ASD) through down-regulation of Pitx2. *Proc Natl Acad Sci U S A* 2008;105:6081–6.
- [72] Wheeler DE, Buck N, Evans JD. Expression of insulin pathway genes during the period of caste determination in the honey bee, *Apis mellifera*. *Insect Mol Biol* 2006;15:597–602.
- [73] Wilkinson TN, Speed TP, Tregear GW, Bathgate RA. Coevolution of the relaxin-like peptides and their receptors. *Ann N Y Acad Sci* 2005;1041:534–9.
- [74] Wilkinson TN, Speed TP, Tregear GW, Bathgate RA. Evolution of the relaxin-like peptide family. *BMC Evol Biol* 2005;5:14.
- [75] Wu Q, Brown MR. Signaling and function of insulin-like peptides in insects. *Annu Rev Entomol* 2006;51:1–24.