



HAL
open science

The evolution of the ligand/receptor couple: a long road from comparative endocrinology to comparative genomics

Gabriel V Markov, Mathilde Paris, Stephanie Bertrand, Vincent Laudet

► To cite this version:

Gabriel V Markov, Mathilde Paris, Stephanie Bertrand, Vincent Laudet. The evolution of the ligand/receptor couple: a long road from comparative endocrinology to comparative genomics. *Molecular and Cellular Endocrinology*, 2008, 293 (1-2), pp.5. 10.1016/j.mce.2008.06.011 . hal-00532041

HAL Id: hal-00532041

<https://hal.science/hal-00532041>

Submitted on 4 Nov 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

Title: The evolution of the ligand/receptor couple: a long road from comparative endocrinology to comparative genomics

Authors: Gabriel V. Markov, Mathilde Paris, Stephanie Bertrand, Vincent Laudet



PII: S0303-7207(08)00267-0
DOI: doi:10.1016/j.mce.2008.06.011
Reference: MCE 6896

To appear in: *Molecular and Cellular Endocrinology*

Received date: 11-2-2008
Revised date: 14-5-2008
Accepted date: 11-6-2008

Please cite this article as: Markov, G.V., Paris, M., Bertrand, S., Laudet, V., The evolution of the ligand/receptor couple: a long road from comparative endocrinology to comparative genomics, *Molecular and Cellular Endocrinology* (2007), doi:10.1016/j.mce.2008.06.011

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Title page

The evolution of the ligand / receptor couple: a long road from comparative endocrinology to comparative genomics

Gabriel V. MARKOV^{1,2}, Mathilde PARIS¹, Stéphanie BERTRAND³ and Vincent LAUDET^{1S}

1: Molecular Zoology team, Institut de Génomique Fonctionnelle de Lyon, Université de Lyon, Ecole Normale Supérieure de Lyon, Université Lyon 1, CNRS, INRA, Institut Fédératif 128 Biosciences Gerland Lyon Sud, France.

2: USM 501/UMR CNRS 5166 - Evolution des Régulations Endocriniennes. Muséum National d'Histoire Naturelle, Paris, France

3: Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645, edifici annex, 1a planta, 08028 Barcelona, Spain.

§ Corresponding author : Molecular Zoology, Institut de Génomique Fonctionnelle de Lyon; UMR 5242 du CNRS; INRA; IFR128 BioSciences Lyon-Gerland; Université de Lyon; Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07 – France. Vincent.Laudet@ens-lyon.fr; Tel 33 4 72 72 81 90; Fax 33 4 72 72 80 80 ; Mobile 33 6 16 41 73 34.

Keys words: nuclear receptors, comparative endocrinology, gene duplication, phylogeny, orthology

Running Title: Comparative endocrinology at the genomic scale

Abstract

Comparative endocrinology considers the evolution of bioregulatory systems and the anatomical structures and molecules that constitute the neuroendocrine and endocrine systems. One aim of comparative endocrinology is to trace the origins of the main endocrine systems. The understanding of the evolution of the ligand / receptor couple is central to this objective. One classical approach to tackle this question is the characterization of receptors and ligands in various types of non-model organisms using as a starting point the knowledge accumulated on classical models such as mammals (mainly human and mouse) and arthropods (with *Drosophila* among other insects). In this review we discuss the potential

caveats associated to this two-by-two comparison between a classical model and non-model organisms. We suggest that the use of an evolutionary approach involving comparisons of several organisms in a coherent framework permits reconstruction of the most probable scenarios. The use of the vast amount of genomic data now available, coupled to functional experiments, offers unprecedented possibilities to trace back the origins of the main ligand / receptor couples.

Contents

1. The ligand / receptor couple
 2. The phylogenetic framework: metazoan evolution
 3. Two by two comparison of receptors proceed at your own risks
 - 3.1. Gene duplication
 - 3.2. Gene loss
 - 3.3. Evolutionary shifts
 - 3.4. Recombination events
 4. The elusive ligands
 - 4.1. Effects
 - 4.2. Detection
 - 4.3. Other evidences
 5. Conclusion: Evolutionary vs. comparative approach: an evolutionary shift for comparative endocrinology.
- Acknowledgements
References

Comparative endocrinology, a very active branch of endocrinology, is mainly interested in the origin and diversification of hormonal systems in living organisms. Given the medically oriented knowledge that is a salient feature of modern endocrinology research, one basic focus of comparative endocrinologists is to trace back the origin of the human major endocrine systems and to understand the main events that have prompted the diversification of these systems (Norris, 2007). This is not an easy task, since this in fact tackles one of the major questions in comparative sciences, that is the origin of complex systems, here the cell-cell communication systems that are acting at the level of the whole organism. Comparative endocrinology, like any other field of life science with an evolutionary component relies on the comparison itself to infer the existence of a given process, anatomical structure, or molecule, back in time. Indeed, since it is usually impossible to have direct information on ancestor species (with the, still anecdotal, but promising exception of ancient DNA research, see Lalueza-Fox et al., 2007) it is the observation that a given feature is conserved in two distant species that allows to conclude that this feature was indeed present in the common ancestor of these two species. Thus a major aspect of research projects in comparative endocrinology is the characterization of endocrine systems in non-model organisms, using model organisms (mainly human but also mouse or arthropods such as *Drosophila*) as a starting point.

The purpose of this short review is double. First, we will emphasize that despite its apparent conceptual simplicity the comparative approach in endocrinology is paved with methodological difficulties. This clearly suggests that if artefactual inferences are to be avoided the traditional comparative approach should be replaced by multi-disciplinary evolutionary and functional studies. Second, we will detail how the impressive amount of data generated by recent genomic analysis offers unprecedented possibilities to carry out this type of research, thus placing comparative endocrinology in front of a major shift in its methods and approaches. Of course such a short review can only provide rapid glances on this burgeoning field. We will thus illustrate this paper by several examples taken from the recent literature without being exhaustive.

1. The ligand / receptor couple

The evolution of the ligand / receptor couple is a question that attracts considerable debate and theoretical discussion of complex experimental approaches. In fact hormones and receptors are central in the understanding of endocrine systems and, their origin as well as their parallel variation through co-evolution is a major evolutionary question. Indeed, divergence of proteins in different species requires ligand and receptor(s) coevolution to improve binding affinity and/or specificity. Coevolution is thus a ubiquitous process that is responsible for the parallel adaptative evolution of hormone / receptors couples in the broadest sense.

On this aspect of coevolution of ligand / receptor pairs the field is sharply cut into two parts given the chemical nature of the ligand. All the ligands that are peptides or proteins, i.e. that are encoded by genes provide conceptually relatively simple cases of ligand /receptor coevolution, with continuous adaptation across time. Protein-protein interaction in general is, from the coevolutionary point of view, not basically different from the interaction between a given receptor and its ligand (Waddell et al., 2007). In such cases, it is believed, and it has been demonstrated in several specific cases, that the genes encoding the ligand and the receptors are undergoing parallel evolution. One of the first examples of such coevolution is the one of the receptors for LH and FSH, which suggests that indeed the specificity of each ligand / receptor pair is maintained in divergent species (Moyle et al., 1994). More recently the case of prolactin receptors in mammals showed how episodes of adaptative evolution have modified the genes encoding the receptor and the ligand (Li et al., 2005, and references therein). In most mammals the prolactin gene evolved very slowly but this near-stasis was interrupted by bursts of rapid changes during the evolution of several mammals orders

such as artiodactyls, primates or rodents. Since prolactin has to bind its receptor to fulfill its function, it was anticipated that the gene encoding the prolactin receptor should be subjected to selective pressure in the same mammals. This has been shown to be the case and the correlation between the evolutionary rates of the ligand and the receptor is effectively indicative of such coevolution. Similar examples including G protein-coupled receptors (GPCRs) such as the receptors for PRXamides (Park et al., 2002) or the secretin (PACAP and VIP) and their receptors (Cardoso et al., 2007) support this theory of ligand-receptor coevolution. Thus, conceptually, the existence of evolutionary couples is relatively well understood and provides a coherent framework for functional evolution studies (Dean and Thornton, 2007).

This situation contrasts with the second, that of receptors for which the ligand is not a peptide or a protein but rather a small molecule. In such cases the ligand is not a gene product but is derived from a biochemical pathway that starts from an inactive precursor, sometimes derived from an external source such as food, which is transformed into the active molecule (see Simoes-Costa et al., 2008 for a recent illustration on retinoic acid metabolism). This is the case for some GPCRs but also for many nuclear receptors (NRs), for which the ligands are the products of complex biochemical pathways. In most cases these pathways contain a rate-limiting step, producing the active compound. This critical step is most often the one that is physiologically regulated. In addition it contains a catabolic part that is responsible for the degradation of the ligand and that is also subjected to precise regulation (see You, 2004, and Bélanger et al., 1998, for a review on steroids). In these cases of ligand / receptor pairs a simple coevolution mechanism obviously cannot operate. In the case of NRs, several models such as ligand exploitation (Thornton, 2001) or refinement of ligand-binding specificity by mutations (Escriva et al., 2006) have recently been proposed to explain how changes of specificity can take place during evolution. Nevertheless even if the situation for these ligand / receptor couples is more complex, the existence of coevolution is still possible. Indeed, a recent report on cannabinoid receptors suggests that the evolution of cannabinoid receptors is correlated with the evolution of diacylglycerol lipase, an enzyme implicated in the metabolism of anandamide and 2-arachidonyl glycerol (2-AG) the two endogenous ligands of cannabinoid receptors (McPartland et al., 2007). It remains to be investigated if such example can be generalized to other receptor systems.

It is clear nevertheless that any discussion on the evolution of a ligand / receptor couple should be confronted with experimental data, which consists in the characterization of ligands and receptors in various species, distantly related to the most classical models. It has also to be emphasized that receptors and ligands structures are often more conserved than their physiological function which depends on the expression patterns of receptors. This often blurs the recognition of orthology between related ligand/receptor pairs.

2. The phylogenetic framework: metazoan evolution

Any analysis of the evolution of a given ligand / receptor couple should take into account the phylogeny of the organisms from which these various pairs are coming from. Therefore, it is important at that level to rapidly present here the framework on which these comparative approaches are developed. Since most of these studies are done at the scale of animals we will limit this rapid presentation to the case of metazoans.

Figure 1 depicts a simplified version of the currently endorsed evolutionary tree for metazoans. It is striking that even if the distribution of model organisms is very much biased towards two of the main clades of metazoans (namely ecdysozoans with *Drosophila* and *Caenorhabditis elegans* and deuterostomes with human, mouse, chicken, xenopus and zebrafish, among others) genomic data are now available for all the main metazoan clades. For example, looking at cnidarians and sponges, at the base of the metazoan tree, the genome of the sea anemone *Nematostella vectensis* (Putnam et al., 2007) is now available whereas the one of the sponge - *Amphimedon queenslandica* - is available as trace archives. These genomes are bringing crucial informations that now allow one to make inferences on early metazoan gene families (Simionato et al., 2007).

From the phylogenetic tree depicted in Figure 1, the bilaterians, that are the metazoans with three embryonic layers (ectoderm, endoderm and mesoderm) and with a clear antero-posterior axis, are divided into three main clades: on one hand, the lophotrochozoans and the ecdysozoans that together form the protostomes and, on the other hand, the deuterostomes. The monophyly of these three clades is relatively well accepted as is their branching order even if some analyses based on large genomic data sets tend to propose alternative schemes (for example the Coelomata hypothesis that groups arthropods and deuterostomes, excluding nematodes; see for example Rogozin et al., 2007).

The topology of the metazoan tree may have a major impact on our inferences regarding the origin and diversification of endocrine genes. This influence is well illustrated by a recent analysis of NR genes distribution and phylogeny in metazoans in which are compared the scenarios implied by the Ecdysozoan or Coelomata hypothesis on the evolutionary history of nuclear receptors (Bertrand et al., 2004). The use of alternative topologies of the metazoan tree (as the Coelomata hypothesis used for the analysis of the Forkhead family; Carlsson and Mahlapuu, 2002) may affect the conclusions drawn relative to the ancestry and evolution of specific genes. At the present stage of our phylogenetic knowledge, this should be regarded with caution.

Another important observation, developed below, is that the number of genes present in the genomes of several classical model organisms tends to be extremely variable because of large scale

events such as whole genome duplication, lineage specific expansion of specific genes or, alternatively gene loss (see Panopoulou and Poustka, 2005 for a review on vertebrates). It is now widely recognized that extremely important models such as *Drosophila*, *C. elegans*, or the urochordate *Ciona intestinalis* have experienced extensive gene loss (see Bertrand et al., 2004 for references). This may have important implications in terms of endocrine gene evolution. For example, the estrogen receptor has been found in vertebrates but not in invertebrate chordates such as *Ciona*, *Drosophila* or *C. elegans*. This has led to the proposal that this receptor was a key innovation of vertebrates (Laudet 1997, Escriva et al., 2000). In fact the observation that an estrogen receptor orthologous gene is present in several mollusks (Thornton et al., 2003, Keay et al., 2006) as well as in cephalochordates (Paris et al., 2008b) shows that it is in fact much more ancient than expected and that the gene was lost independently in ecdysozoans and urochordates (Bertrand et al., 2004; Escriva et al., 2004). A very similar situation was found for the thyroid hormone receptor (Bertrand et al., 2004., Wu et al., 2006, Wu et al., 2007). It is now widely accepted that no conclusion can be reached on the presence or absence of a given gene in the common ancestor of all bilaterians if data from the three main lineages of metazoans (that is lophotrochozoans, ecdysozoans and deuterostomes) are not available. Given that deuterostomes plus insects and nematodes contain the most dominant model organisms this conclusion should be strongly re-emphasized: no safe conclusion can be drawn on the ancestry of a given gene family without data from lophotrochozoans and/or cnidarians.

3. Two by two comparisons of receptors, proceed at your own risk!

A classical approach in comparative endocrinology is to identify in various organisms, for example the cephalochordate amphioxus, an orthologue of a given receptor known in a classical model organisms like human. This type of analysis may be important to clearly show if this receptor, and by extension, the corresponding signaling pathway is effectively present in the organism of interest. We will highlight below that, even if of course the conclusions reached from this type of analysis are interesting, this two-by-two comparative approach has several caveats that renders it quite risky. Thus, the conclusions and evolutionary models reached by these traditional comparative approaches, based on the assumption that one can extend to their zoological groups the knowledge accumulated in human or *Drosophila*, should be put into a larger perspective. It is only through their independent confirmation based on large-scale evolutionary analysis that these conclusions will be firmly assessed.

3.1. Gene duplication

An important result generated over the last 10 years of genomic analysis is that genes, and even genomes, duplications once believed to be relatively rare events are in fact a major evolutionary mechanism that has been instrumental in shaping the current biodiversity (Ohno, 1970, and see also Volff, 2005 for a recent review in teleosts). Gene duplication is an important mechanism of gene diversification that can be observed in nearly all organisms. Of course tandem duplication of individual genes, the simplest case of gene duplication, occurs quite often. But two more global processes that have broad implications for the functional anatomy of genomes should be emphasized. The first is the whole genome duplication, the importance of which was first highlighted in the 70's by Susumu Ohno (Ohno, 1970) and later revealed by the study of invertebrate chordates and early vertebrates (Garcia-Fernandez and Holland, 1994) as well as fishes (Wittbrodt et al., 1998). It is now well established that in several cases, such as at the base of the vertebrate tree (Dehal and Boore, 2005) or early on during actinopterygian fish evolution (Jaillon et al., 2004) whole genome duplication took place and impacted strongly on the appearance and diversification of complex features, including the endocrine systems (Holland et al., 2008). The fate of duplicated genes (non, sub or neo-functionalization, see Force et al., 1999 that is, respectively loss of one copy, sharing of the ancestral function between the two duplicates or acquisition of a new function by one of the copy), and its link with the origin of evolutionary novelties, is currently a major and fruitful research area and the study of endocrine genes should be placed in this context. A very convincing example of such an analysis in the context of actinopterygian fish whole genome duplication has been recently published for the MyoD family (Macqueen and Johnston, 2008).

The other extreme case of gene duplication is lineage-specific duplication events that can sometimes be extensive. This is the mechanism at the origin of the 270 nuclear receptor genes that are present in the genome of the nematode *Caenorhabditis elegans* (Sluder and Maina, 2001). It has been shown that most of these genes correspond in fact to orthologues of a unique gene encoding the orphan receptor HNF-4, which was massively duplicated in *C. elegans* and related species (Robinson-Rechavi et al., 2005). The functional significance of this burst of duplication is still under investigation, but it is clear that such an event should have consequences on the physiology of the animal and may completely modify the evolutionary scenarios constructed at the level of an individual gene family. GPCRs (Cardoso et al., 2006), but also receptors with tyrosine kinase activities (Rikke et al., 2000), the heterotrimeric G protein α -subunit (O'Halloran et al. 2006) and the *Hedgehog*-related genes (*Hog*) (Aspöck et al., 1999) also provide cases of lineage-specific expansion in nematodes.

The presence of these duplication events has a major consequence for the comparative endocrinologist, as it reinforces the need for an accurate phylogenetic study of any protein of

interest in order to avoid simplified, and often erroneous interpretations. Figure 2 shows the example of the thyroperoxidase (TPO) gene, which is implicated in the synthesis of vertebrate thyroid hormones. The vertebrate TPO is a member of a multigenic family, being the ortholog of a large group of peroxidases that experienced a complex series of duplication events within the vertebrate lineages. In mammals we find not only thyroid peroxidase but also lacto-peroxidase, myelo-peroxidase and eosinophil peroxidase, all these genes coming from a unique ancestor gene that underwent duplications at various levels within the vertebrate lineages (Heyland et al., 2006). Using primers corresponding to a various set of metazoan peroxidases, a gene encoding a peroxidase was cloned in the urochordates *Ciona intestinalis* and *Halocynthia roretzi*. Since in vertebrates four closely related types of peroxidases are found, one cannot conclude that this *Ciona* gene is indeed a TPO (see our recent review on the importance of gene nomenclature, Markov et al., 2008). It is only with additional data, here through the analysis of the expression pattern and the comparison to the previously reported biochemical activity that the conclusion that the gene encodes a TPO can be reached. Interestingly, a more extensive genomic analysis (Figure 2) shows that this protein is the product of only one of the two duplicated genes in the *Ciona* genome. The original “*Ciona* TPO” was shown to be expressed in a domain that does not overlap that of its classical regulator TTF1 in vertebrates, and that its expression domain was restricted to the endostyle zone 7, whereas previous histochemical studies reported a peroxidase activity also in other parts of the endostyle (zones 8 and 9) (Ogasawara et al., 1999). The presence of the second gene suggests that a more complete study is needed in order to test if it too has also a TPO-activity, if it is expressed in zones 8 and 9 of the endostyle and if the two genes have overlapping or complementary expression patterns and thus to reconstruct the detailed history of these genes. Since the duplication event that gave rise to the two *Ciona* genes is independent of the duplications that occurred in vertebrates, only functional analysis can determine whether the *Ciona* genes have activities related to thyroid, myelo-, lacto- or eosinophile peroxidases. The situation is even more complex in cephalochordates. Another gene encoding a TPO was characterised in the chinese amphioxus *Branchiostoma belcheri*, (Ogasawara, 2000) but the analysis of the *Branchiostoma floridae* genomes (Holland et al., 2008) shows that there are four orthologs of this gene, once again arising from an independent series of duplications from the ones that occurred in *Ciona* or vertebrates. Thus, starting from a unique gene with an unknown activity, the peroxidase gene family has been independently elaborated three times in vertebrates, urochordates and cephalochordates. Since proteins that undergo duplication are prone to subfunctionalisation or neofunctionalisation events (Force et al., 1999), this should be taken into account when comparing the functions of two proteins in two different organisms, because this can have major effects at the physiological level. This example illustrates that one should take into account the full set of genes and their complex

history in order to infer their ancestral functions.

3.2. Gene loss

Gene loss is an often neglected aspect that, due to the current interest in whole genome sequences, has recently been shown to be a frequent and important evolutionary mechanism that contributed significantly to the emergence of divergent animal lineages (Danchin et al., 2006). The analysis of the presence of NRs genes in complete genome sequences of metazoans shows that the NR complement of different animal models is extremely variable and that gene loss was effectively frequent, as discussed above in the case of *Drosophila*, nematodes and urochordates (Bertrand et al., 2004). The case of nematodes is particularly puzzling since the events of gene loss are hidden by the massive lineage specific expansion of the HNF4 gene discussed above (Robinson-Rechavi et al., 2005) thus illustrating the complexity of the individual gene family history (Figure 3).

But gene loss is not only revealed by comparison of distant evolutionary organisms. One example, taken from the evolution of NRs illustrates the importance of taking gene loss into account. When we compared, through careful phylogenetical analysis, the NR complement present in mammals and teleost fishes we were surprised to find cases that could not be explained by the classical “more genes in fish” scenario (Crollius and Weissenbach, 2005; Bertrand et al., 2004). This is well exemplified by the case of the Rev-erb genes, orphan nuclear receptors. In all known mammals, two paralogous Rev-erb genes, called Rev-erb α and Rev-erb β , corresponding to a unique orthologue in *Drosophila* (called E75) and amphioxus are known (Laudet, 1997). Given the actinopterygian fish whole genome duplication event discussed above, we were expecting to find a maximum of four Rev-erb genes in zebrafish, namely two Rev-erb α and two Rev-erb β . During our systematic analysis of NR genes expression patterns in zebrafish we were thus surprised to find in fact five Rev-erb genes in zebrafish whereas four are effectively present in pufferfishes (tetraodon and fugu) (Bertrand et al., 2007). By phylogenetical analysis (Figure 4) we were able to deduce that in fact three paralogous Rev-erb genes (Rev-erb α , β and a third gene called γ) were present in the ancestor of vertebrates. In tetrapods this third gene, Rev-erb γ was lost. In fishes, the three ancestral Rev-erb genes were duplicated giving rise to six genes. The differential loss of these fish-specific paralogues explains the different numbers of Rev-erb genes in various fish species: zebrafish lost one Rev-erb α copy whereas pufferfish lost the two Rev-erb α genes. The analysis of the expression patterns strongly suggests that if these genes share a similar function each of them has specific implications in distinct processes, probably linked to circadian rhythm regulation (Bertrand et al., 2007,

Kakizawa et al., 2007). This case is certainly not an isolated one and taking only the example of NR genes several other cases of specific gene loss in mammals were found (e.g. ERRs, COUP-TFs, SF-1) whereas, in contrast, one case of fish specific loss (CAR) was found (Bertrand et al., 2007). It is now known that these gene losses occurred in other gene families in mammals but also in other zoological groups (Danchin et al., 2006; Wyder et al., 2007).

All these examples illustrate that the evolutionary history of any given gene cannot be reconstructed without a careful phylogenetic analysis based on the use of adequate methods and that the importance of mechanisms such as gene duplication and gene loss should be adequately tested before concluding on the presence or absence of a given signaling pathway. Classical comparative work based on the two-by-two comparison of human and another animal will never match the quality of information reached by an in depth phylogenetical analysis.

3.3. Evolutionary shifts

Of course genes are not only duplicated or lost. Many genes are conserved as unambiguous orthologues in a wide variety of organisms, but even in such cases one should be careful when studying gene evolution since evolutionary shifts can have dramatic impacts in terms of gene function, as well as at the level of phylogeny. Long branch attraction is a well-known artefact in molecular phylogeny that can artificially group rapidly evolving clades together in a position clearly not compatible with the known phylogeny of species (Delsuc et al., 2005).

RXR evolution in insects provides an illustration of this phenomenon (Figure 5). The known orthologue of RXR in *Drosophila* is the USP gene that plays an important role as a common heterodimeric partner for several NRs, including the ecdysone receptor, EcR. When we did phylogenetic studies of USP in several insects we saw a very striking tree topology since all the USPs found in Diptera and Lepidoptera were not clustered with other arthropod sequences as expected, but at the base of all bilaterian RXRs (Bonneton et al., 2003). At the level of sequence identity this example is quite spectacular since, in the ligand-binding domain the RXR from a beetle (e.g. *Tribolium*) is more closely related to the human RXR than to the *Drosophila* USP. We found that in fact this corresponds to a very strong acceleration of evolutionary rate that occurred during the evolution of mecopterida, a monophyletic group with approximately 25% of extant insect species, including Trichoptera (caddisflies), Mecoptera (scorpionflies) and Siphonaptera (fleas) in addition to Diptera and Lepidoptera (Bonneton et al., 2006). Careful structural and functional analysis of RXR from *Tribolium* and *Drosophila* allowed us to reconstruct a complex and dynamic evolutionary history shaped by several functional shifts during which several events of loss and gain of ligand-binding occurred (Iwema et al., 2007). Interestingly, this event is not restricted to

RXR since several similar shifts were also recently observed in other nuclear receptors, all acting in the ecdysone cascade (Bonneton et al., 2008). This example illustrates that such evolutionary shifts are not just complicating factors for phylogenetic analysis, they also provide very fruitful cases in which the evolution of ligand / receptor couples can be scrutinized. In this case, this reveals an unsuspected evolutionary flexibility of NRs in terms of ligand binding.

The literature now contains several cases of evolutionary shifts that all provide illuminating examples of the power of natural selection acting at the level of endocrine pathways. Among these, the nuclear receptors provide two recent examples of receptors, ERs and TRs, that were believed to be recent but for which orthologues were recently found in non-model protostomes. Interestingly, in both cases these protostome receptors do not bind the *bona fide* vertebrate ligand suggesting a more complex history than anticipated. Indeed the ligand binding evolution of estrogen receptor is still far from understood (see contrasting views in Baker, 2003, Thornton et al., 2003 and Paris et al., 2008b) and similar questions are suggested by the recent characterization of thyroid hormone receptors in platyhelminthes, molluscs or even in a deuterostome, the cephalochordate amphioxus (Wu et al., 2007, Paris et al., 2008a). At a much smaller taxonomic level the careful analysis of the ligand-binding specificity of PXR in mammals also provides an example of shifts in the ligand-binding ability of a nuclear receptor (Krasowski et al., 2005 ; Reschly et al., 2007). For other receptor systems, the interested reader could, among a long potential list, refer to the cases of the Growth Hormone receptor in primates (Liu et al., 2001; Wallis et al., 2001), the TSH receptor which exhibits striking functional differences between mammals and actinopterygian fishes, but also among mammals (Farid and Szkudlinski, 2004) or the genes of the ectodysplasin pathway in vertebrates (Pantalacci et al., 2008).

3.4. Recombination events

Genomic data now available greatly facilitates assessment of gene orthology. In some multigenic families, the duplication rate is so high and variable for different genes that even a cautious phylogenetic analysis is not sufficient to establish orthology relationships. Such a problem occurs when examining the vertebrate CYP2 family that contains genes encoding xenobiotic metabolising enzymes, which have undergone many duplications events at various taxonomic levels (e.g. mammal-specific, rodent-specific, and mouse-specific duplications, with primate-specific or/and rat-specific duplications in parallel) (Nelson, 2005, Thomas, 2007). In such cases, data on the chromosomal localisation can provide very useful information since the analysis of neighboring genes allows one to assess the orthology of the whole syntenic region and, by extension, of the

genes of interest. This was indeed the case of the CYP2 genes since some of the genes are located in genomic clusters, where it is believed that recombination events occur quite often, giving rise to the wide diversity of CYP2 genes (Thomas 2007). Such events will erase the phylogenetic signal and make the functional comparisons between proteins that underwent different duplication events meaningless.

Another example of careful analyses of syntenic regions that led to large scale conclusions on complex evolutionary scenario are the evolutionary analyses of the Major Histocompatibility Cluster (MHC) by Pontarotti's group (Danchin et al., 2003). In the context of fish genome duplication the example of MyoD gene family evolution also illustrates how the use of synteny could help to infer orthology (Macqueen and Johnston, 2008). This type of analysis is particularly useful when studying genes encoding short peptide hormones, which are often difficult to unambiguously identify at a large evolutionary scale. The recent analyses of the POMC/AGPR/MCR gene repertoire in fugu (Klovins et al., 2004) and the study of the origin and coevolution between NPY receptor and prolactin-releasing hormone-receptors in vertebrates (Lagerström et al., 2005), also show the utility of complementary use of data from phylogenetic analysis coupled with careful synteny analysis. The construction of the “gene rosace” diagram that visualizes the respective relationships between genes located in different regions facilitates this type of analysis (see Jaillon et al., 2004, for an explanatory illustration). As for gene analyses, such studies take profit of in depth phylogenomic knowledge about the analysed genes, and also are greatly improved by comparaisons between more than two genomes (for detailed examples see Kasahara et al., 2007, Muffato and Crollius, 2008).

When possible, phylogenetic approach should integrate in a known phylogenetic species and gene trees all the information possibly available at the leave of the trees. Information are of course the sequences but also functional information such as binding properties and other physiological data. In some cases it is possible to calculate the probability to have this information a the node of the tree and from the information available at the node propagate dawn it to the non annotated leaves (see Thornton, 2004, for a detailed review). But even the finest probabilistic approaches are useless if the sampling is inappropriate.

4. The elusive ligands

The previous examples may give the impression that working on the evolution of receptor genes is full of obstacles that complicate analysis and interpretation. Thus, one solution to study evolution of a specific endocrine signaling pathway could be to test directly if the ligand of interest is present in

a wide variety of organisms. This attractive solution is nevertheless also paved with obstacles that are not easily solved and for which a comparative genomics approach does not always provide clues. The two main lines of evidence for the existence of a given ligand in a given animal are (i) the detection of an effect of this ligand in a given physiological or developmental process, most often considered from the known effect of this ligand in mammals and (ii) the direct detection of the ligand through genomic or analytical chemistry analysis.

4.1. Effects

The rationale behind these studies, which are commonly done in traditional comparative endocrinology is that if a given molecule has an effect in a given species this should be taken as a clear indication that there is a receptor for this molecule and thus a signaling pathway and a « physiological role ». Somehow mirroring the original receptor definition by Paul Ehrlich, for which a receptor was defined as any molecule that is able to bind exogenous elements in human cells (reviewed in Prüll, 2003), people tend to consider that any molecule of human origin acting on a living system is effectively proof that this molecule binds to another one, which can be called a “receptor”. But even if this assumption is true, it cannot be used as an indication that this “receptor” is related in anyway to the human receptor for this molecule. The problem is that many authors have applied this concept without taking into account this important caveat.

A striking example is the case of the arboreal mycorrhizal fungi of the order Glomerales, such as *Glomus intraradices*, that live in symbiosis with many trees and play crucial roles at the root/soil interface. It has been shown that the roots of the plant have a positive effect in stimulating the growth of the fungal hyphae during pre-symbiosis steps (see Requena et al., 2007, for a review). The precise mechanism how plant signals are perceived by the fungi is still unknown but it has been shown that flavonoids have a strong chemotactic effect and that this effect can be mimicked by estrogens and blocked by anti-estrogens (Poulin et al., 1997; Scervino et al., 2005). This was taken as an indication that the genome of these fungi contained an estrogen receptor and that a specific receptor/interacting protein with a binding site for flavonoid or structurally related compounds (estrogens and antiestrogens) exists in these fungi (Requena et al., 2007; Poulin et al., 1997; Catford et al., 2006). However, our current knowledge of NR evolution indicates that the genome of *Glomus* (which is currently being sequenced) should be devoid of NR genes (Escriva et al., 1997). Among the possible candidates for mediating the effect of estrogens in *Glomus*, one of the most probable would be an orthologue of the estrogen binding protein already found in other fungi, including *Candida albicans* (Madani et al., 1994; Cheng et al., 2006). This protein, well known under the name of “Old yellow enzyme” is one of the most ancient enzymatic systems characterized, and is a

cytoplasmic oxydoreductase whose activity can be regulated by estrogen binding (Cheng et al., 2006 and references therein). This example demonstrates that whereas there is an effect of estrogen in this species, this could not be taken as an indication that the endogenous receptor is in any way evolutionary related to the nuclear estrogen receptor found in human.

RXRs provide here again an example of the difficulties associated with the identification of endogenous ligands. These receptors were first described in the early 90's as orphan receptors whose activity could be modulated by all-trans retinoic acid (see Laudet and Gronemeyer 2005, for references). It was later shown that 9-cis retinoic acid, an isomeric derivative of all-trans retinoic acid, is able to bind with a high affinity to RXR. Indeed several rexinoids, specific ligands of RXRs, were developed as pharmacological compounds for treatment of diseases such as diabetes and insulin resistance (Pinaire and Reifel-Miller, 2007). Ironically the detection of 9-cis retinoic acid *in vivo* has proven to be difficult, casting doubts on the *in vivo* relevance of this ligand (see Mic et al., 2003). This has been recently confirmed by genetic evidence in the mouse (Calleja et al., 2006). It is now thought that RXR acts rather as a fatty acid or fatty acid derivative sensor, although the exact identity and relevance of its ligand(s) remains a matter of debate (see references in Iwema et al., 2007). This example depicts that even in human, it is not because a molecule has an effect that it indicates its endogenous existence in a physiologically relevant manner!

These examples may appear artificial, but they illustrate a type of assumption that has been commonly used in some comparative endocrinology studies. The estrogen receptor of several molluscs such as *Aplysia*, *Octopus*, and others have been cloned and convincingly shown to be unable to bind to estrogens (Thornton et al., 2003; Keay et al., 2006; Kajiwara et al., 2006; Matsumoto et al., 2007, and Bannister et al. 2007). Nevertheless, effects of estrogens are reported in these animals (e.g. sex reversal in several mollusc species) and this is taken as evidence in favor of the existence of a classical nuclear estrogen receptor (see Lafont and Mathieu, 2007 for references). However, the actual data suggests that if such a receptor exists it is not the unique ER orthologue already characterized. Among the several possible explanations we propose that: (i) the molecule used for the treatment is metabolized to another compound that is the active one. Indeed we recently discovered such a situation in amphioxus, where thyroid hormones are metabolized to a compound that binds the TR (Paris et al., 2008a) ; (ii) a NR that binds estrogen exists but this is not an orthologue of ER ; (iii) the effect of estrogen is mediated by another receptor system, e.g. a transmembrane receptor (see Revankar et al., 2005 for an example) or a cytoplasmic protein such as in the case of *Glomus* discussed above. Thus one should treat claims of the existence of a signaling pathway based on treatment with human derived compound with much caution. Such data provide indications, but only indirect evidence that should be verified by more functional approaches including the cloning and characterisation of the relevant receptor.

4.2. Detection

Detecting a given ligand in one's favorite non-classical model organism is probably an excellent option for a comparative endocrinologist. One interesting case to illustrate the complexity of ligand evolution in different organisms is the one of the lamprey steroid hormones. For a long time, it was thought that as a vertebrate, lamprey should produce the classical steroids, estradiol, testosterone, and progesterone, and indeed, these were effectively identified by radioimmunoassays (RIA) and used for physiological (for an example see Bolduc and Sower, 1992) or functional (Thornton, 2001) studies. Nevertheless, in the 80's, chromatographic techniques and more recently, blood steroid profiles analyzed by high performance liquid chromatography (HPLC), contradicted this view and showed that the main circulating steroids are 15- α -hydroxylated (Kime and Callard, 1982, Lowartz et al., 2003). The actual nature of the active steroid hormones and their mechanism of action in lamprey are still a matter of debate (reviewed in Bryan, 2008). This example shows that the diversity of ligands in non-model species is often underestimated. In this case the presence of the classical steroids is not questioned, but they are not the physiologically relevant compounds.

Many reports have shown the presence of a typical human hormone in early vertebrates or even in invertebrates, steroids being the compounds for which the number of such reports is the highest. Most often the detection of these products is based on the use of antibodies, generated from human compounds and widely used on human tissue samples for medical application. These antibodies are mainly used for immunohistochemical staining and/or radioimmunoassay and two recent examples illustrate the caveats and uncertainties of such approaches. We have nevertheless to strongly insist on the fact that these examples were not selected because they are particularly dubious but rather because they illustrate numerous studies done in classical comparative endocrinology.

The first example concerns the detection of immunoreactivity for progesterone in the giant Rohde cells of the amphioxus *Branchisotoma belcheri* (Takeda et al., 2003). Using rabbit polyclonal antibodies against progesterone conjugated to bovine serum albumin from two commercial suppliers, the authors performed immunohistochemical staining on adult amphioxus and show staining in giant neurons known as the Rohde cells. The authors cautiously conclude that "progesterone-like" substances are likely to be present in these neurons, suggesting the existence of neurosteroids in this species. They performed a number of controls to avoid problems due to non-specific binding: they replaced the primary antibody by normal rabbit serum, they omitted the primary antibody from the staining reaction and they used PBS instead of the primary antibody. Furthermore, they also performed absorption tests with the antibodies, progesterone, and BSA, respectively, corroborating the fact that the staining is specific. The main problem with this type of

paper is that we have no information on the actual specificity of the antibodies. Are these reagents able to recognize only progesterone itself or do they show cross reactivity with other closely related molecules such as pregnenolone, deoxycorticosterone or 17β -hydroxyprogesterone? What if the amphioxus contains no progesterone but related steroids, such as those found in lamprey? The authors of this study were cautious and spoke of “progesterone-like substances”. It is interesting to note that the amphioxus genome contains several, but not all the enzymes implicated in steroidogenesis (e.g. a CYP21 orthologue is missing) and that several of these enzymes are duplicated (e.g. 3β -HSD for which six genes were found in amphioxus, compared to three in human and the SDR family that contains enzymes with 17β -HSD and 11β -HSD activities for which 31 genes are known in human for more than 100 in amphioxus) (Holland et al., 2008). However, only one steroid receptor, located at the base of the vertebrate GR, MR, PR and AR, is present in the amphioxus genome, and, to date, nothing is known about its specificity. Thus it would not be surprising if amphioxus steroidogenesis has been elaborated independently from vertebrate steroidogenesis, and searching for human compounds in amphioxus may simply be unfruitful.

The second example concerns the presence of estrogens in molluscs, and more precisely in the cephalopod *Octopus vulgaris*, an interesting case since it is much referred to in the debate on the ligand-binding evolution of estrogen receptors (Thornton et al., 2003; Keay et al., 2006; Paris et al., 2008b). It has been shown that, in this species, 17β -estradiol and progesterone are found in oviduct and ovarian tissues, and that the concentration of these hormones in females correlates with phases of the reproductive cycle (Di Cosmo et al., 2001; D’Aniello et al., 1996). The doses of hormones in these tissues are measured using a radioimmunoassay and we are faced with the same uncertainties about the specificity of the detection method in this study as in the previous one. In addition, the level of hormones detected both at the level of radioimmunoassay or by HPLC is very close to the lower limits of detection (D’Aniello et al., 1996). From all the data accumulated in *Octopus*, one can conclude that steroid hormone metabolism does exist in this species, but it is difficult to be more conclusive on the precise identity of the steroids that will be found (see below).

These two examples emphasize the difficulties associated with detection based on the use of antibodies. In fact, before the genome era the same situation was found for NR and many reports claimed for the detection of steroid receptors in a wide variety of non-metazoans organisms such as plants or fungi (see Agarwal et al., 1994, Milanesi et al., 2001, Milanesi and Boland, 2006). Despite the fact that, as in the above mentioned studies, all the controls were correct, all these studies displayed artefacts: there is no gene related to steroid receptors outside metazoans (Escriva et al., 1997, Escriva et al., 2004). We believe that basically the same situation holds for the immunological detection of NR ligands in invertebrates: perhaps some of these reports are correct but in the absence of more firm evidence this cannot be confirmed (Lafont and Mathieu, 2007).

Thus the use of antibodies as an evolutionary probe, even if this is attractive because these experiments are technically easy to perform, should be avoided.

4.3. Other lines of evidence

Other methods of analysis for the presence of specific compounds can be used. This is for example the ability of a given tissue to metabolize a given compound, i.e. to carry out a given reaction (e.g. hydroxylation at the position 3 β indicative of a 3 β -HSD activity see DiCosmo et al., 2001). Once again the problem in these experiments is the conclusions that can reasonably be drawn. That a given type of activity exists *in vivo* is interesting but given the diversity of steroidogenic enzymes (or enzymes using different substrates) in metazoans one cannot take for granted that because a reaction occurs with a given labelled intermediate it actually uses this very substrate *in vivo*.

Another approach often used is searching for the presence of the gene encoding a enzyme responsible for a critical step in the synthesis pathway of a given ligand. For example the existence of an aromatase in amphioxus (Castro et al., 2005) argues for the existence of estrogens in this species. This is convincing in this particular case since the phylogeny of the isolated clone was carefully assessed including synteny analysis. This is much less clear in other cases, such as the recent description of putative CYP11A and CYP17 in amphioxus (Mizuta and Kubokawa, 2007). The phylogeny of these clones, taking into account a more exhaustive dataset, with sequences from non-chordate species, shows that the phylogenetical inferences of this report were inexact and that in both cases, the cloned gene are orthologous to a new group of deuterostome CYPs, that was lost in vertebrates (Markov et al., in preparation).

Our critical arguments may well in turn be criticised as, we agree, it is easy to overstate previous publications without offering positive alternatives. It seems clear that the best studies are those done with careful analytical chemistry methods. For example HPLC coupled with mass spectrometry allows the clear detection and identification of compounds and offers a very strong and rigorous alternative (Bridgham et al., 2006 see also Lafont and Mathieu, 2007 for a similar conclusion). The problem with this method is of course that it is much more demanding and expensive than antibody detection. But they have the unique advantage of unambiguously assessing the presence and diversity of the compounds found in a given organism. The recent characterization of steroids present in the nematode *Caenorhabditis elegans* provides an interesting example of the variety of compounds that can be revealed through such a detailed analysis (Motola et al., 2006).

5. Conclusion: Evolutionary vs. comparative approaches: an evolutionary shift for comparative endocrinology.

Most often, endocrinology studies are realized with a long-term medical objective. This leads to a deformation of the evolutionary perspective, which may be too much "human-centered". Doing comparative endocrinology should therefore first be accompanied by an effort to go against this natural anthropocentric view and to consider equally the evidences coming from different taxa. This may seem obvious, but in fact "rampant" anthropocentrism and a graded view of evolution (e.g. that evolution is progressing toward complexity) is often difficult to avoid as nicely pointed out by Gould (Gould, 1996).

In fact a key solution to these methodological and interpretation difficulties relies in the comparative strategy itself that should incorporate evidences coming from various organisms and scientific approaches in an integrated manner. An interesting example of the power of such an approach is the case of the suiform aromatases (Gaucher et al., 2004). In this study, by combining bioinformatic, molecular evolution, paleontology, cladistics, structural biology and organic chemistry analysis, the authors propose that the conservation of three subfunctionalised aromatase paralogs in pigs is the result of a selection for *Suoidae* with larger litter than their ancestors. This selective event has allowed their survival during the global climatic shift that began in the Eocene. Bioinformatic analyses (estimation of divergence times, detection of positive selection by Ka/Ks analysis), were correlated with the presence of residues that were subject to positive selection in the substrate-binding site and with previous experimental data about different substrate specificities for these enzymes. Additionally, a detailed examination of the palaeontological record and of the number of pups in modern artiodactyls as well, correlated with data on global climate changes, led to this quite audacious hypothesis, that is consistent with data of many different research fields. It would be interesting – and possible, thanks to the availability of genome data - to check if such a correlation occurs in other vertebrates and could be statistically significant.

We argue that future comparative endocrinology studies should combine large-scale evolutionary analysis, with several standardised phylogenetic and chemical methods to ensure the robustness of the conclusions. The classical two-by-two comparison that is prone to artefactual conclusions based on partial analysis and the use of poorly refined detection methods, such as those based on antibodies should be replaced by such multidisciplinary studies, even if these approaches are experimentally much more difficult. Comparative endocrinology is thus now facing a real challenge: to perform multidisciplinary evolutionary approaches that will effectively offer solutions, using rigorous technical and conceptual basis, to the long-standing questions of the origin of the endocrine signalling pathways.

Acknowledgements

We are grateful to ARC, CNRS, EMBO, MENRT, and Region Rhone-Alpes for financial support. Work from our laboratory is also funded by the EU Cascade Network of Excellence and the integrated project Crescendo. We thank Barbara Demeneix, François Bonneton and Hector Escriva for critical reading of the manuscript.

References

- Agarwal, M.K., Mirshahi, M., Braq, S., Jullienne, A., Leblanc, N., Stibon, F., Guern, J., 1994. *Nicotiana tabacum* contains a putative mineralocorticoid receptor. *Biochem. Biophys. Res. Commun.* 200, 1230-1238.
- Aspöck, G., Kagoshima, H., Niklaus, G., Bürglin, T. R., 1999. *Caenorhabditis elegans* has scores of hedgehog-related genes: sequence and expression analysis. *Genome Res.* 9, 909-923.
- Baker, M.E., 2003. Evolution of adrenal and sex steroid action in vertebrates: a ligand-based mechanism for complexity. *Molecular and Cellular Endocrinology* 25, 396-400.
- Bannister, R., Beresford, N., May, D., Routledge, E.J., Jobling, S., Rand-Weaver, M., 2007. Novel estrogen receptor-related Transcripts in *Marisa cornuarietis*, a freshwater snail with reported sensitivity to estrogenic chemicals. *Environ. Sci. Technol.* 41, 2643-2650.
- Bélanger, A.; Hum, D.W., Beaulieu, M., Lévesque, E., Guillemette, C., Tchernof, A., Bélanger, G., Turgeon, D., Dubois, S., 1998. Characterization and regulation of UDP-glucuronosyltransferases in steroid target tissues. *J. Steroid Biochem. Mol. Biol.* 65, 301-310.
- Bertrand, S., Brunet, F.G., Escriva, H., Parmentier, G., Laudet, V., Robinson-Rechavi, M., 2004. Evolutionary genomics of nuclear receptors: from twenty-five ancestral genes to derived endocrine systems. *Mol. Biol. Evol.* 21, 1923-1937.
- Bertrand, S., Thisse, B., Tavares, R., Sachs, L., Chaumot, A., Bardet, P., Escriva, H., Duffraisse, M., Marchand, O., Safi, R., Thisse, C., Laudet, V., 2007. Unexpected Novel Relational Links Uncovered by Extensive Developmental Profiling of Nuclear Receptor Expression. *PLoS Genet.* 3, e188.
- Bolduc, T.G., Sower, S.A., 1992. Changes in brain gonadotropin-releasing hormone, plasma estradiol 17-beta, and progesterone during the final reproductive cycle of the female sea lamprey, *Petromyzon marinus*. *J. Exp. Zool.* 264, 55-63.
- 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. *Mol. Biol. Evol.* 20, 541-553.
- Bonneton, F., Brunet, F. G., Kathirithamby, J., Laudet, V., 2006. The rapid divergence of the ecdysone receptor is a synapomorphy for Mecoptera that clarifies the Strepsiptera problem. *Insect Mol. Biol.* 15, 351-362.
- Bonneton, F., Chaumot, A., Laudet, V. 2008. Annotation of *Tribolium* nuclear receptors reveals an increase in evolutionary rate of a network controlling the ecdysone cascade, *Insect Biochemistry and Molecular Biology*, *Insect Biochem. Mol. Biol.* 38, 416-429.

- Bridgham, J.T., Carroll, S.M., Thornton, J.W. 2006. Evolution of hormone-receptor complexity by molecular exploitation. *Science* 312, 97-101.
- Bryan, M.B., Scott, A.P., Li, W., 2008. Sex steroids and their receptors in lampreys. *Steroids* 73, 1-12.
- Calléja, C., Messaddeq, N., Chapellier, B., Yang, H., Krezel, W., Li, M., Metzger, D., Mascrez, B., Ohta, K., Kagechika, H., Endo, Y., Mark, M., Ghyselinck, N.B., Chambon, P., 2006. Genetic and pharmacological evidence that a retinoic acid cannot be the RXR-activating ligand in mouse epidermis keratinocytes. *Genes Dev.* 20, 1525-1538.
- Cardoso, J.C.R.; Pinto, V. C.; Vieira, F.A.; Clark, M.S., Power, D.M., 2006. Evolution of secretin family GPCR members in the metazoa. *BMC Evol. Biol.* 6, 108.
- Cardoso, J., de Vet, E., Louro, B., Elgar, G., Clark, M., Power, D., 2007. Persistence of duplicated PAC1 receptors in the teleost, *Sparus auratus*. *BMC Evol. Biol.* 7, 221.
- Carlsson, P., Mahlapuu, M., 2002. Forkhead transcription factors: key players in development and metabolism. *Dev. Biol.* 250, 1-23.
- Castro, L.F.C., Santos, M.M., Reis-Henriques, M. A., 2005. The genomic environment around the Aromatase gene: evolutionary insights. *BMC Evol. Biol.* 5, 43.
- Catford, J.G., Staehelin, C., Larose, G., Piché, Y., Vierheilig, H., 2006. Systemically suppressed isoflavonoids and their stimulating effects on nodulation and mycorrhization in alfalfa split-root systems. *Plant Soil* 285: 257-266.
- Cheng, G., Yeater, K.M., Hoyer, L.L., 2006. Cellular and molecular biology of *Candida albicans* estrogen response. *Eukaryot. Cell.* 5, 180-191.
- Crollius, H.R., Weissenbach, J., 2005. Fish genomics and biology. *Genome Res.* 15, 1675-1682.
- D'Aniello, A., Cosmo, A.D., Cristo, C.D., Assisi, L., Botte, V., Fiore, M.M.D., 1996. Occurrence of sex steroid hormones and their binding proteins in *Octopus vulgaris* lam. *Biochem. Biophys. Res. Commun.* 227, 782-788.
- Danchin, E.G.J., Abi-Rached, L., Gilles, A., Pontarotti, P., 2003. Conservation of the MHC-like region throughout evolution. *Immunogenetics* 55, 141-148.
- Danchin, E., Gouret, P., Pontarotti, P., 2006. Eleven ancestral gene families lost in mammals and vertebrates while otherwise universally conserved in animals. *BMC Evol. Biol.* 6, 5.
- Dean, A.M., Thornton, J.W., 2007. Mechanistic approaches to the study of evolution: the functional synthesis. *Nat. Rev. Genet.* 8, 675-688.
- Dehal, P., Boore, J.L., 2005. Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol.* 3, e314.
- Delsuc, F., Brinkmann, H., Philippe, H. 2005. Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet.* 6, 361-375.
- Di Cosmo, A., Cristo, C.D., Paolucci, M., 2001. Sex steroid hormone fluctuations and

morphological changes of the reproductive system of the female of *Octopus vulgaris* throughout the annual cycle. J. Exp. Zool. 289, 33-47.

Escriva, H., Safi, R., Hänni, C., Langlois, M.C., Saumitou-Laprade, P., Stehelin, D., Capron, A., Pierce, R., Laudet, V., 1997. Ligand binding was acquired during evolution of nuclear receptors. Proc. Natl. Acad. Sci. U. S. A., 94, 6803-6808.

Escriva, H., Delaunay, F., Laudet, V., 2000. Ligand binding and nuclear receptor evolution. Molecular and Cellular Endocrinology 22, 717-727.

Escriva, H., Bertrand, S., Laudet, V., 2004. The evolution of the nuclear receptor superfamily. Essays Biochem. 40, 11-26.

Escriva, H., Bertrand, S., Germain, P., Robinson-Rechavi, M., Umbhauer, M., Cartry, J., Duffraisse, M., Holland, L., Gronemeyer, H., Laudet, V., 2006. Neofunctionalization in vertebrates: the example of retinoic acid receptors. PLoS Genet. 2, e102.

Farid, N.R., Szkudlinski, M.W., 2004. Minireview: structural and functional evolution of the thyrotropin receptor. Endocrinology 145, 4048-4057.

Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L., Postlethwait, J., 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151, 1531-1545.

Garcia-Fernández, J., Holland, P.W., 1994. Archetypal organization of the amphioxus Hox gene cluster. Nature 370, 563-566.

Gaucher, E.A., Graddy, L.G., Li, T., Simmen, R.C.M., Simmen, F.A., Schreiber, D. R., Liberles, D. A., Janis, C.M., Benner, S.A., 2004. The planetary biology of cytochrome P450 aromatases. BMC Biol. 2, 19.

Gould, S.J. 1996. Full House. Harmony Books, New York.

Heyland, A., Price, D.A., Bodnarova-Buganova, M., Moroz, L.L., 2006. Thyroid hormone metabolism and peroxidase function in two non-chordate animals. J. Exp. Zoolog. B Mol. Dev. Evol. 306, 551-566.

Holland, L.Z., Albalat, R., Azumi, K., Benito-Gutierrez, E., Bronner-Fraser, M., Brunet, F., Butts, T., Candiani, S., Dishaw, L.J., Garcia-Fernandez, J., Ferrier, D.E.K., Gibson-Brown, J.J., Gissi, C., Godzik, A., Hallbook, F., Hirose, D., Hosomichi, K., Ikuta, T., Inoko, H., Kasahara, M., Kasamatsu, J., Kawashima, T., Kimura, A., Kobayashi, M., Kozmik, Z., Kubokawa, K., Laudet, V., Litman, G. W., McHardy, A. C., Meulemans, D., Nonaka, M., Olinski, R.P., Pancer, Z., Pestarino, M., Rast, J.P., Rigoutsos, I., Roch, G., Saiga, H., Sasakura, Y., Satake, M., Satou, Y., Schubert, M., Sherwood, N., Shiina, T., Takatori, N., Tello, J., Vopalensky, P., Wada, S., Xu, A., Ye, Y., Yoshida, K., Yoshizaki, F., Yu, J.K., Zhang, Q., Zmasek, C.M., Putnam, N.H., Rokhsar, D.S., Satoh N., Holland, P.W.H., 2008. The amphioxus genome illuminates vertebrate origins and cephalochordate biology. Genome Research, in revision.

Iwema, T., Billas, I. M. L., Beck, Y., Bonneton, F., Nierengarten, H., Chaumot, A., Richards, G., Laudet, V., Moras, D., 2007. Structural and functional characterization of a novel type of ligand-independent RXR-USP receptor. EMBO J. 26, 3770-3782.

Jaillon, O., Aury, J., Brunet, F., Petit, J., Stange-Thomann, N., et al., 2004. Genome duplication in

the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431, 946-57.

Kajiwara, M., Kuraku, S., Kurokawa, T., Kato, K., Toda, S., Hirose, H., Takahashi, S., Shibata, Y., Iguchi, T., Matsumoto, T., Miyata, T., Miura, T., Takahashi, Y., 2006. Tissue preferential expression of estrogen receptor gene in the marine snail, *Thais clavigera*. *Gen. Comp. Endocrinol.* 148, 315-326.

Kakizawa, T., Nishio, S., Triqueneaux, G., Bertrand, S., Rambaud, J., Laudet, V., 2007. Two differentially active alternative promoters control the expression of the zebrafish orphan nuclear receptor gene Rev-erbalpha. *J. Mol. Endocrinol.* 38, 555-568.

Kasahara, M., Naruse, K., Sasaki, S., Nakatani, Y., Qu, W., Ahsan, B., Yamada, T., Nagayasu, Y., Doi, K., Kasai, Y., Jindo, T., Kobayashi, D., Shimada, A., Toyoda, A., Kuroki, Y., Fujiyama, A., Sasaki, T., Shimizu, A., Asakawa, S., Shimizu, N., Hashimoto, S., Yang, J., Lee, Y., Matsushima, K., Sugano, S., Sakaizumi, M., Narita, T., Ohishi, K., Haga, S., Ohta, F., Nomoto, H., Nogata, K., Morishita, T., Endo, T., Shin-I, T., Takeda, H., Morishita, S., Kohara, Y., 2007. The medaka draft genome and insights into vertebrate genome evolution. *Nature* 447, 714-719.

Keay, J., Bridgham, J.T., Thornton, J.W., 2006. The *Octopus vulgaris* estrogen receptor is a constitutive transcriptional activator: evolutionary and functional implications. *Endocrinology* 147, 3861-3869.

Kime, D.E., Callard, G.V., 1982. Formation of 15 alpha-hydroxylated androgens by the testis and other tissues of the sea lamprey, *Petromyzon marinus*, in vitro. *Gen. Comp. Endocrinol.* 46, 267-270.

Klovins, J., Haitina, T., Fridmanis, D., Kilianova, Z., Kapa, I., Fredriksson, R., Gallo-Payet, N., Schiöth, H.B., 2004. The melanocortin system in Fugu: determination of POMC/AGRP/MCR gene repertoire and synteny, as well as pharmacology and anatomical distribution of the MCRs. *Mol. Biol. Evol.* 21, 563-579.

Krasowski, M. D., Yasuda, K., Hagey, L. R., Schuetz, E. G., 2005. Evolution of the pregnane x receptor: adaptation to cross-species differences in biliary bile salts. *Mol. Endocrinol.* 19, 1720-1739.

Lafont, R., Mathieu, M., 2007. Steroids in aquatic invertebrates. *Ecotoxicology* 16, 109-130.

Lagerström, M.C., Fredriksson, R., Bjarnadóttir, T.K., Fridmanis, D., Holmquist, T., Andersson, J., Yan, Y., Raudsepp, T., Zoorob, R., Kukkonen, J.P., Lundin, L., Klovins, J., Chowdhary, B.P., Postlethwait, J.H., Schiöth, H.B., 2005. Origin of the prolactin-releasing hormone (PRLH) receptors: evidence of coevolution between PRLH and a redundant neuropeptide Y receptor during vertebrate evolution. *Genomics* 85, 688-703.

Laudet, V., 1997. Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor. *J. Mol. Endocrinol.* 19, 207-226.

Laudet, V., Gronemeyer, H., 2005. *The Nuclear Receptor FactsBook*. Academic Press, London.

Lalueza-Fox, C., Römpler, H., Caramelli, D., Stäubert, C., Catalano, G. et al., 2007. Melanocortin 1 receptor allele suggests varying pigmentation among Neanderthals. *Science* 318, 1453-1455.

- Li, Y., Wallis, M., Zhang, Y., 2005. Episodic evolution of prolactin receptor gene in mammals: coevolution with its ligand. *J. Mol. Endocrinol.* 35, 411-419.
- Liu, J.C., Makova, K.D., Adkins, R.M., Gibson, S., Li, W.H., 2001. Episodic evolution of growth hormone in primates and emergence of the species specificity of human growth hormone receptor. *Mol. Biol. Evol.* 18, 945-953.
- Lowartz, S., Petkam, R., Renaud, R., Beamish, F.W.H., Kime, D.E., Raeside, J., Leatherland, J.F., 2003. Blood steroid profile and in vitro steroidogenesis by ovarian follicles and testis fragments of adult sea lamprey, *Petromyzon marinus*. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 134, 365-376.
- Macqueen, D.J., Johnston, I.A., 2008. An update on MyoD evolution in teleosts and a proposed consensus nomenclature to accommodate the tetraploidization of different vertebrate genomes. *PLoS ONE* 3, e1567.
- Madani, N.D., Malloy, P.J., Rodriguez-Pombo, P., Krishnan, A.V., Feldman, D., 1994. *Candida albicans* estrogen-binding protein gene encodes an oxidoreductase that is inhibited by estradiol. *Proc. Natl. Acad. Sci. U. S. A.* 91, 922-926.
- Markov, G., Lecointre, G., Demeneix, B., Laudet, V., 2008. The «street light syndrome», or how protein taxonomy can bias experimental manipulations. *Bioessays* 30, 349-357.
- Matsumoto, T., Nakamura, A.M., Mori, K., Akiyama, I., Hirose, H., Takahashi, Y., 2007. Oyster estrogen receptor: cDNA cloning and immunolocalization. *Gen. Comp. Endocrinol.* 151, 195-201.
- McPartland, J.M., Norris, R.W., Kilpatrick, C.W., 2007. Coevolution between cannabinoid receptors and endocannabinoid ligands. *Gene* 397, 126-135.
- Mic, F.A., Molotkov, A., Benbrook, D.M., Duester, G., 2003. Retinoid activation of retinoic acid receptor but not retinoid X receptor is sufficient to rescue lethal defect in retinoic acid synthesis. *Proc. Natl. Acad. Sci. U. S. A.* 100, 7135-7140.
- Milanesi, L., Monje, P., Boland, R., 2001. Presence of estrogens and estrogen receptor-like proteins in *Solanum glaucophyllum*. *Biochem. Biophys. Res. Commun.* 289, 1175-1179.
- Milanesi, L., Boland, R., 2006. Presence of vitamin D3 receptor (VDR)-like proteins in *Solanum glaucophyllum*. *Physiologia plantarum* 128, 341-350.
- Mizuta, T., Kubokawa, K., 2007. Presence of sex steroids and cytochrome P450 genes in amphioxus. *Endocrinology* 148, 3554-3565.
- Motola, D.L., Cummins, C.L., Rottiers, V., Sharma, K.K., Li, T., Li, Y., Suino-Powell, K., Xu, H. E., Auchus, R.J., Antebi, A., Mangelsdorf, D.J., 2006. Identification of ligands for DAF-12 that govern dauer formation and reproduction in *C. elegans*. *Cell* 124, 1209-1223.
- Moyle, W. R., Campbell, R.K., Myers, R.V., Bernard, M.P., Han, Y., Wang, X., 1994. Co-evolution of ligand-receptor pairs. *Nature* 368, 251-255.
- Muffato, M., Crollius, H.R., 2008. Paleogenomics in vertebrates, or the recovery of lost genomes from the mist of time. *Bioessays* 30, 122-134.

- Nelson, D.R., 2005. Gene nomenclature by default, or BLASTing to Babel. *Hum. Genomics* 2, 196-201.
- Norris, D.O., 2007. *Vertebrate endocrinology*. Elsevier Academic Press, San Diego.
- O'Halloran, D.M., Fitzpatrick, D.A., McCormack, G.P., McInerney, J.O., Burnell, A.M., 2006. The Molecular Phylogeny and Functional Significance of a Nematode Specific Clade of Heterotrimeric G-protein α -Subunit Genes. *J. Mol. Evol.* 63, 87-94.
- Ogasawara, M., Lauro, R.D., Satoh, N., 1999. Ascidian homologs of mammalian thyroid peroxidase genes are expressed in the thyroid-equivalent region of the endostyle. *J. Exp. Zool.* 285, 158-169.
- Ogasawara, M., 2000. Overlapping expression of amphioxus homologs of the thyroid transcription factor-1 gene and thyroid peroxidase gene in the endostyle: insight into evolution of the thyroid gland. *Dev. Genes. Evol.* 210, 231-42.
- Ohno, S., 1970. *Evolution by gene duplication*. Springer-Verlag, Heidelberg.
- Panopoulou, G., Poustka, A.J., 2005. Timing and mechanism of ancient vertebrate genome duplications - the adventure of a hypothesis. *Trends Genet.* 21, 559-567.
- Pantalacci, S., Chaumont, C., Benoît, G., Sadier, A., Delsuc, F., Douzery, E., Laudet, V., 2008. Conserved features and evolutionary shifts of the EDA signaling pathway involved in vertebrate skin appendage development. *Mol. Biol. Evol.* 25, 912-928.
- Paris, M., H. Escriva, H., Schubert, M., Brunet, F., Brtko, J., Cieselski, F., Roecklin, D., Vivat-Hannah, V., Cravedi, J.-P., Scanlan, T.S., Renaud, J.-P., Holland, N.D., Laudet, V., 2008a. Amphioxus metamorphosis and the origin of the thyroid hormone signaling pathway. *Current Biology*, in press.
- Paris, M., Pettersson, K., Schubert, M., Bertrand, S., Pongratz, I., Escriva, H., Laudet, V., 2008b. An amphioxus orthologue of the estrogen receptor that does not bind estradiol : insight into estrogen receptor evolution. *BMC Evolutionary Biology*, in revision.
- Park, P.S., Palczewski, K., 2005. Diversifying the repertoire of G protein-coupled receptors through oligomerization. *Proc. Natl. Acad. Sci. U. S. A.* 102, 8793-8794.
- Pinaire, J.A., Reifel-Miller, A., 2007. Therapeutic potential of retinoid x receptor modulators for the treatment of the metabolic syndrome. *PPAR Res.* 941-56.
- Poulin, M.J., Simard, J., Catford, J.G., Labrie, F. and Piche, Y., 1997. Response of symbiotic endomycorrhizal fungi to estrogens and antiestrogens, *Mol Plant Microbe Interactions* 10.
- Prüll, C.-R., 2003. Part of a Scientific Master Plan? Paul Ehrlich and the Origins of his Receptor Concept. *Medical History* 47, 332-356.
- Putnam, N.H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shapiro, H., Lindquist, E., Kapitonov, V.V., Jurka, J., Genikhovich, G., Grigoriev, I.V., Lucas, S.M., Steele, R.E., Finnerty, J.R., Technau, U., Martindale, M.Q., Rokhsar, D.S., 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* 317, 86-94.

- Reschly, E.J., Bainy, A.C.D., Mattos, J.J., Hagey, L.R., Bahary, N., Mada, S.R., Ou, J., Venkataramanan, R., Krasowski, M.D., 2007. Functional evolution of the vitamin D and pregnane X receptors. *BMC Evol. Biol.* 7, 222.
- Requena, N., Serrano, E., Ocón, A., Breuninger, M., 2007. Plant signals and fungal perception during arbuscular mycorrhiza establishment. *Phytochemistry* 68, 33-40.
- Revankar, C.M., Cimino, D.F., Sklar, L.A., Arterburn, J.B., Prossnitz, E.R., 2005. A transmembrane intracellular estrogen receptor mediates rapid cell signaling. *Science* 307, 1625-1630.
- Rikke, B. A.; Murakami, S., Johnson, T. E., 2000. Paralogy and orthology of tyrosine kinases that can extend the life span of *Caenorhabditis elegans*. *Mol. Biol. Evol.* 17, 671-683.
- Robinson-Rechavi, M., Maina, C.V., Gissendanner, C.R., Laudet, V., Sluder, A., 2005. Explosive lineage-specific expansion of the orphan nuclear receptor HNF4 in nematodes. *J. Mol. Evol.* 60, 577-586.
- Rogozin, I.B., Wolf, Y.I., Carmel, L., Koonin, E.V., 2007. Analysis of rare amino Acid replacements supports the coelomata clade. *Mol. Biol. Evol.* 24, 2594-2597.
- Scervino, J.M., Ponce, M.A., Erra-Bassells, R., Vierheilig, H., Ocampo, J.A., Godeas, A., 2005. Arbuscular mycorrhizal colonization of tomato by Gigaspora and Glomus species in the presence of root flavonoids. *J. Plant. Physiol.* 162, 625-633.
- Simionato, E., Ledent, V., Richards, G., Thomas-Chollier, M., Kerner, P., Coornaert, D., Degnan, B. M., Vervoort, M., 2007. Origin and diversification of the basic helix-loop-helix gene family in metazoans: insights from comparative genomics. *BMC Evol. Biol.* 7, 33.
- Simões-Costa, M.S., Azambuja, A.P., Xavier-Neto, J., 2008. The search for non-chordate retinoic acid signaling: lessons from chordates. *J. Exp. Zool. B Mol. Dev. Evol.* 310, 54-72.
- Sluder, A.E., Maina, C.V., 2001. Nuclear receptors in nematodes: themes and variations. *Trends Genet.* 17, 206-213.
- Takeda, N., Kubokawa, K., Matsumoto, G., 2003. Immunoreactivity for progesterone in the giant Rohde cells of the amphioxus, *Branchiostoma belcheri*. *Gen. Comp. Endocrinol.* 132, 379-383.
- Thomas, J.H., 2007. Rapid birth-death evolution specific to xenobiotic cytochrome P450 genes in vertebrates. *PLoS Genet.* 3, e67.
- Thornton, J.W., 2001. Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. *Proc. Natl. Acad. Sci. U. S. A.* 98, 5671-5676.
- Thornton, J.W., Need, E., Crews, D., 2003. Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. *Science*, 301, 1714-1717.
- Thornton, J. W., 2004. Resurrecting ancient genes: experimental analysis of extinct molecules. *Nat. Rev. Genet.*, 5, 366-375.
- Waddell, P.J., Kishino, H., Ota, R., 2007. Phylogenetic Methodology for Detecting Protein Interactions. *Mol. Biol. Evol.* 24, 650-659.

Wallis, M., 2001. Episodic evolution of protein hormones in mammals. *J. Mol. Evol.* 53, 10-18.

Volff, J., 2005. Genome evolution and biodiversity in teleost fish. *Heredity* 94, 280-294.

Wittbrodt, J., Meyer, A., Scharl, M., 1998. More genes in fish? *Molecular and Cellular Endocrinology* 20, 511-512.

Wu, W., Niles, E.G., El-Sayed, N., Berriman, M., LoVerde, P.T., 2006. *Schistosoma mansoni* (Platyhelminthes, Trematoda) nuclear receptors: sixteen new members and a novel subfamily. *Gene* 366, 303-315.

Wu, W., Niles, E.G., LoVerde, P.T., 2007. Thyroid hormone receptor orthologues from invertebrate species with emphasis on *Schistosoma mansoni*. *BMC Evol. Biol.* 7, 150.

Wyder, S., Kriventseva, E., Schröder, R., Kadowaki, T., Zdobnov, E., 2007. Quantification of ortholog losses in insects and vertebrates. *Genome Biol.* 8, R242.

You, L., 2004. Steroid hormone biotransformation and xenobiotic induction of hepatic steroid metabolizing enzymes. *Chem. Biol. Interact.* 147, 233-246.

Figure 1: An updated view on metazoan phylogeny.

1R, 2R, and 3R indicate the whole genome duplication events that occurred in the vertebrate lineage.

Figure 2: Phylogenetic relationships in the chordate thyroperoxidase family.

Duplication events are indicated by green spots. Note that myeloperoxidase, lactoperoxidase and eosinophil peroxidases are the mammalian members of a gene family where duplications occurred at various levels in the vertebrate lineage. General tree topology is based on Heyland, 2006, and completed with data from *Ciona* complete genome (<http://www.treefam.org/cgi-bin/TFinfo.pl?ac=TF500225>) and from Holland et al., 2008 for *Branchiostoma floridae* sequences.

Figure 3: Gene duplications and losses in *C.elegans*: the example of the nuclear receptor superfamily.

The figure shows the loss of seven genes and duplications of two of them in *C. elegans*, in comparison with the *Drosophila* gene set. Data are from Bertrand et al., 2004. Note that the gene set of the “common ancestor” presented here is minimal, and coherent with the comparison of only these two species. A more complete analysis would show that some other genes were also lost independently in *Drosophila* and *C. elegans*.

Figure 4: Phylogenetical tree of the Rev-erb subfamily in zebrafish and human and the most

probable evolutionary scenario.

Branch-lengths are arbitrary. The two whole genome duplications that took place at the origin of vertebrates and the whole genome duplication at the origin of teleost fish are schematised by yellow and light-blue spots respectively. VertA means vertebrate ancestor. The most parsimonious scenario explaining the actual evolutionary relationships between human and zebrafish Rev-erb genes can be separated in two steps. The first one corresponds to the loss of one Rev-erb paralogue (here named Rev-erb δ) in the ancestor of vertebrates after the two whole genome duplications. The second step is the loss of Rev-erb γ paralogue in the ancestor of the vertebrates that diverged after fish divergence, and the loss of one Rev-erb α duplicate (here named Rev-erb α -B) in zebrafish after the whole genome duplication that took place at the base of teleost fish lineage.

Figure 5: Long branches attraction and functional shifts: the example of USP-RXR.

Adapted from Bonneton et al., 2003 and Iwema et al, 2007. A. USP-RXR protein tree, with an abnormal position of Mecoptera at the base of the bilaterians. Further analysis showed that this long branch, corresponding to an acceleration of the evolutionary rate in Mecoptera USP, correlates with changes in the ligand-binding abilities. B. Species tree based on classical neutral markers such as ribosomal RNA genes), showing the real position of Mecoptera within arthropods. The three different colours refer to the different binding abilities of the ligand-binding pocket (LBP).

Figure 1

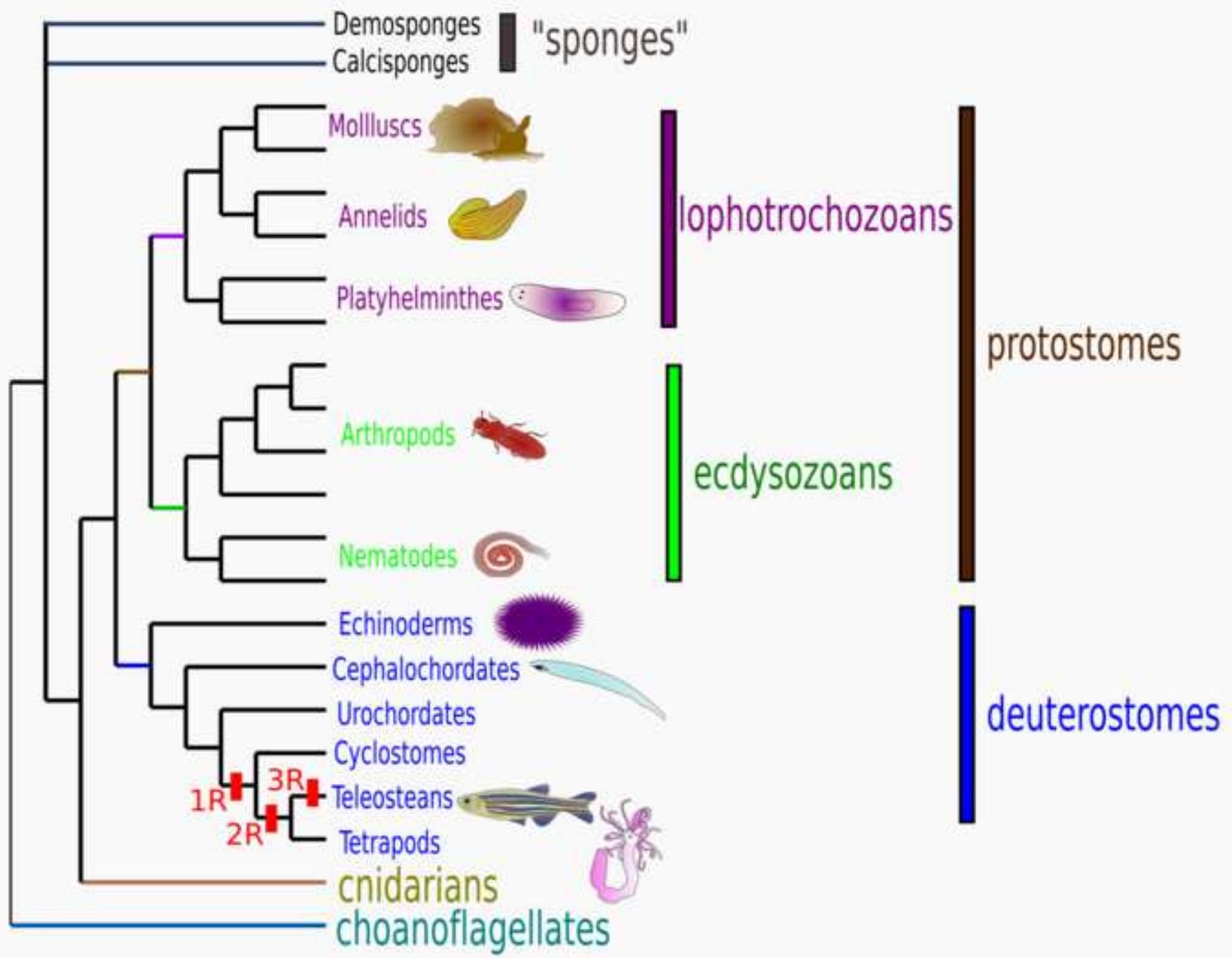


Figure 2

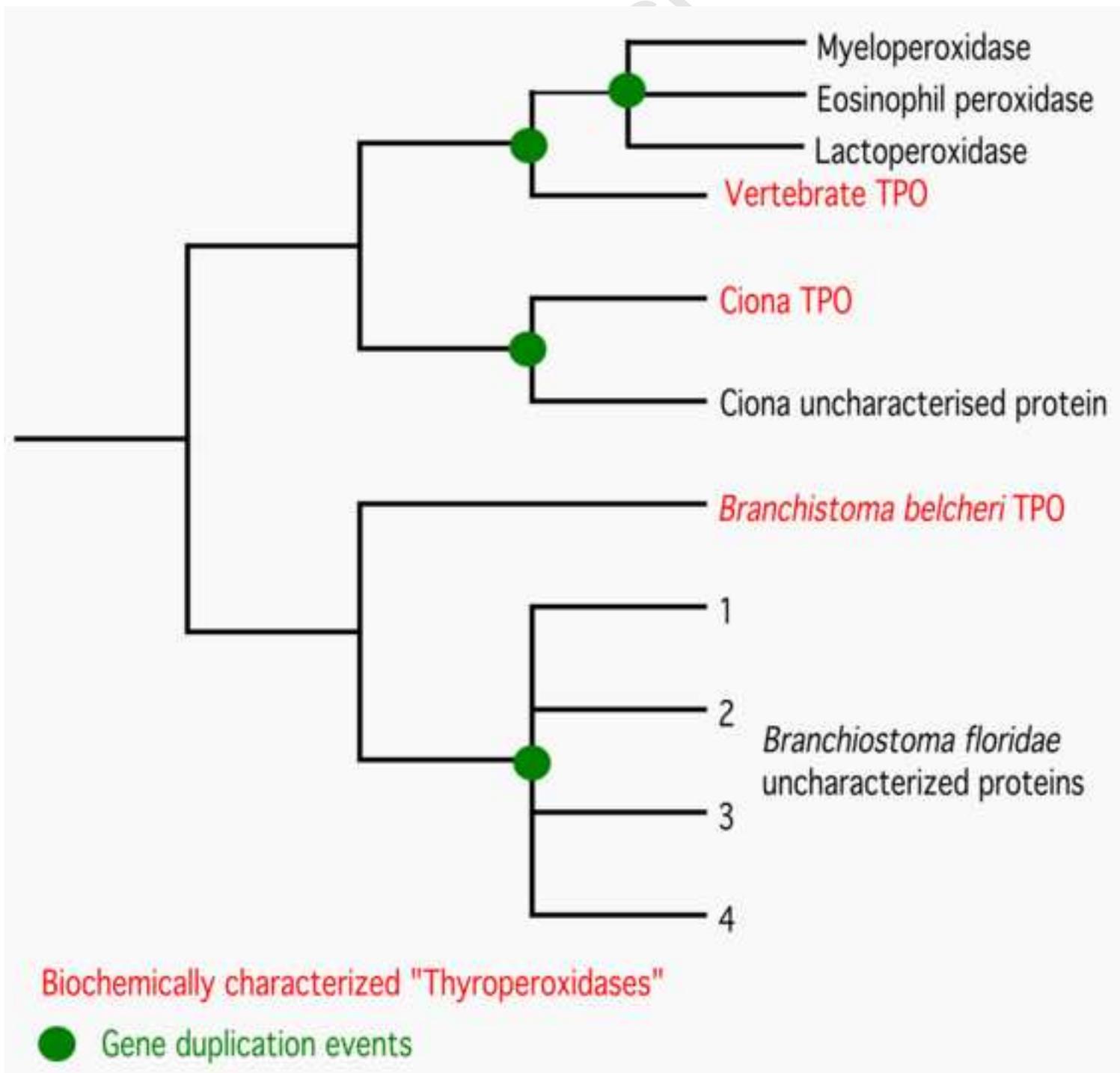


Figure 3

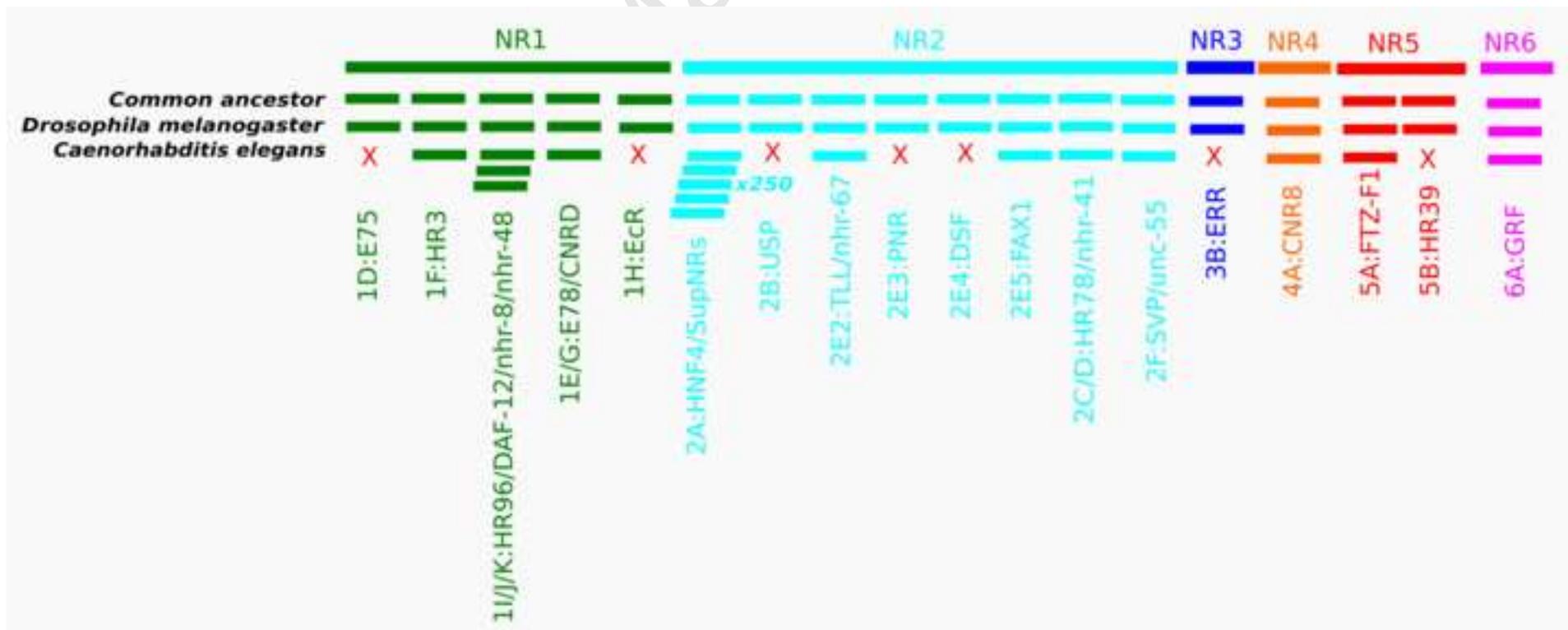


Figure 4

