CMLS, Cell. Mol. Life Sci. 57 (2000) 809–827
1420-682X/00/010809-19 $ 1.50 + 0.20/0

CMLS Cellular and Molecular Life Sciences

Review

Origins and evolutionary diversification of the nuclear receptor superfamily

G. I. Owen and A. Zelent*

Leukaemia Research Fund Centre at the Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London SW3 6JB (UK), Fax + 44 171 352 3299; e-mail: arthur@icr.ac.uk

Received 6 December 1999; accepted 21 January 2000

Abstract. Nuclear receptors (NRs), which include those for steroid and thyroid hormones as well as retinoids, are encoded by a large gene superfamily that has evolved to regulate nearly every facet of metazoan life, from development to basic metabolism. This article reviews the conservation in structure and function of distinct receptors across different species and attempts to draw conclusions as to the evolution of this gene superfamily. Although sequences related to NRs can be found in plants and yeast, gene sequence analyses suggest that the NR ancestor(s) first appeared in the early metazoans and subsequently diversified into the six receptor sub-families, which were already recognisable at the time of the Arthropoda-Chordata split over 700 million years ago. At the time when a primitive NR emerged, the basic components of the transcription regulatory machinery, which are conserved from yeast to vertebrates, were already in place and the ancestral NR must have evolved with the ability to communicate with them. The first such NRs likely acted as monomers and in a ligand-independent fashion. As members of the NR superfamily acquired the ability to hetero- and homodimerise, and to bind and be regulated by ligands, the functional complexity of the NR superfamily increased. This exponentially increasing complexity subsequently provided a potential driving force for evolution of higher organisms by supplying a sophisticated regulatory gene network that could control complex physiological processes during development and in adult organisms.

Key words. Metazoa; arthopod; nematode; HOX genes; retinoic acid; steroid receptor; chromosome; genome duplication; phylogenetic tree; development.

Introduction and diverse biological roles of the superfamily

Regulation of gene expression at the transcriptional level is an essential component of important cellular and developmental processes such as growth, differentiation and lineage commitment. In this respect, the concerted action of cell-type-specific transcription factors, which bind to the DNA elements (response elements) located in the regulatory regions (such as promoters and/or enhancers) of specific genes and either inhibit or stimulate the rate of transcription initiation by RNA polymerase, is of extreme importance. Nuclear receptors (NRs), such as those for steroids, thyroid hormones and retinoic acid are soluble proteins that can bind as dimers to specific DNA regulatory elements (hormone response elements or HREs) and act as cell-type- and promoter-specific transcription factors. In contrast to other transcription factors, however, their activities can be modulated through binding of the corresponding hydrophobic ligands.

* Corresponding author.

In general, NRs act as homo- or heterodimers by binding two HREs consisting of two copies of a PuGGTCA core sequence arranged as inverted (palindromes), everted or direct repeats [1, 2]. A small number of NRs have been characterised, which bind as monomers to a single half-site. For receptors that act as heterodimers, the retinoid X receptor (RXR), that binds 9-*cis*-retinoic acid, appears to be a common partner with specificity of response element recognition for different heterodimers being dictated by differential spacing between the two directly repeated core sequences. Although most HREs for RXR heterodimers follow the so-called 1–5 rule [3–5], a number of exceptions have been noted in the past [6, 7]. Steroid receptors bind as homodimers to palindromic response elements with a spacing of three nucleotides. In addition to forming a variety of heterodimers, the RXR also appears able to act as a homodimer, but on direct repeats spaced by one nucleotide [8]. A few examples of thyroid hormone receptor (TR), and vitamin D_3 receptor (VDR) homodimers binding to everted repeats have been described [9, 10].

Recent studies on mechanisms of NR action suggest that these molecules can act both as transcriptional repressors and activators and that, in the case of most ligand-binding receptors, the ligand acts as a switch from repressing to activating activity. Constitutively active receptor (CARβ) is a good example where the unliganded factor activates transcription and binding of a ligand (adrostane metabolites) renders it a repressor [11]. Recently, a number of co-activators have been characterised [12, 13]. Proteins of the steroid receptor co-activator (SRC), or transcription intermediary factor (TIF) family and CREB-binding protein (CBP) have been shown to interact in a ligand-dependent manner with a number of NRs, including retinoic acid receptors (RARs), and to stimulate their activities [14–19]. SRC-1 and CBP proteins possess intrinsic histone acetyltransferase activities [20, 21]. Recent data have demonstrated that an RNA transcript, termed the steroid receptor RNA activator (SRA), can also act as a co-activator and is present in NR complexes along with SRC-1 [22]. In the absence of ligands, some NRs, such as TRs and RARs, remain associated with the nuclear receptor co-repressors, negative co-regulator (N-CoR) [23] or the silencing mediator for retinoid and thyroid hormone receptors (SMRT), [24], and repress basal transcription. Both N-CoR [25, 26] and SMRT [27] associate with the mammalian homologues (mSin3A and mSin3B) of the yeast global transcriptional repressor SIN3 [28–30] and histone deacetylase [31], and are thought to repress transcription through histone deacetylation, rendering the nearby chromatin inaccessible to transcriptional activators and/or basal transcription factors.

Thus far, accumulated sequencing data from a variety of organisms suggest that the NRs are exclusive to the metazoan kingdom. From early metazoans, throughout the course of evolution, NRs have been incorporated by different organisms to bind distinct ligands and regulate various physiological processes. Our knowledge of the molecular mechanisms of NR action is derived, to a large extent, from the early studies of the glucocorticoid and oestrogen receptors (GRs and ERs), which were the first members of this superfamily to be cloned nearly 15 years ago. Nevertheless, despite their important place in the history of discoveries relating to NRs, steroid hormone receptors such as GR and ER represent merely a well-studied specialised branch rather than an archetype for the superfamily [1]. Despite the fact that GR and other steroid receptors play a number of important physiological roles in mammals and other vertebrates, their homologues so far have not been detected in invertebrates. Although many non-steroid receptors, such as RAR, RXR, fushi tarazu factor 1 (FTZ-F1), chicken ovalbumin upstream promoter-transcription factor (COUP-TF) and hepatocyte nuclear factor 4 (HNF4), appear to have been highly conserved through evolution, their developmental roles have been adapted to different extents by distinct metazoans. For example, arthropods deficient in HNF4 show developmental defects in the midgut and Malpighian tubules, while in the mouse, the lack of HNF4 causes lethality with defects in gastrulation and mesoderm formation. These findings may suggest that the mammalian HNF4 has been utilised to perform an earlier, and/or perhaps a more critical, function in development than its arthropod homologue [reviewed in ref. 32].

Ecdysteroids and juvenile hormones, which are NR ligands, control both the moulting and metamorphic stages of the arthropod life cycle [33]. During the onset of *Drosophila* metamorphosis, ecdysone regulates expression of at least seven members of the NR superfamily [34, 35]. One of these nuclear receptors, DHR3, appears to function as the 'switch' that defines the transition from late larva to prepupa, while another NR called E75b modulates the timing of this process [34, 36]. On the basis of such observations, Truman and Riddiford [33] proposed that through their mediation of ecdysone and juvenile hormone action, NRs are responsible for the evolution of insect metamorphosis. Interestingly, in vertebrates where ecdysone is not a ligand, orthologues (direct descendants of an ancestral gene, after speciation) of DHR3 and E75 are represented by orphan receptors RORα and rev-erb, respectively, which have been incorporated into completely different signalling pathways [32, 37]. Likewise, the *Drosophila* ultraspiracle protein (USP), which is encoded by the RXR orthologue, has adapted to bind varying forms of juvenile hormone [38]. USP heterodimerises with the

ecdysone receptor (EcR) to control insect metamorphosis, while the RXR heterodimerises with RAR to mediate the regulatory effects of all-*trans* retinoic acid on vertebrate development.

Oestrogen and progesterone, whose activities are mediated via their respective receptors, ER and PR, are best known for their roles in mammalian reproduction, yet this may not have been the first role for these receptors. Baker [39] speculated that the origin of the ER was in the development and regulation of the nervous system. This notion may be supported by the role of oestrogen in the control and differentiation of species-specific behaviour and endocrine homeostasis in birds, with ER expression differing in the brain of songbirds and non-songbirds [40]. The origin of oestrogen synthesis is also interesting. Aromatase (the enzyme that catalyses the conversion of androgens to oestrogens) activity is present in the central neural tissues of birds, amphibians and teleosts (modern fish), but not in the hagfish, an intermediate on the way to vertebrates. Thus the failure to detect both ER and aromatase in invertebrates suggests that the ER is specific to vertebrates. The original role for PR is also believed to be different to the role it plays in mammals today. It has been suggested that the inflammatory and granulation tissue reaction to a foreign body has, under the influence of progesterone, been converted into an implantation response. Support for progesterone having a role in the immune system also comes with the finding that the immune-response-related colony-stimulating factors and interleukins are also involved in implantation [reviewed in refs 41, 42].

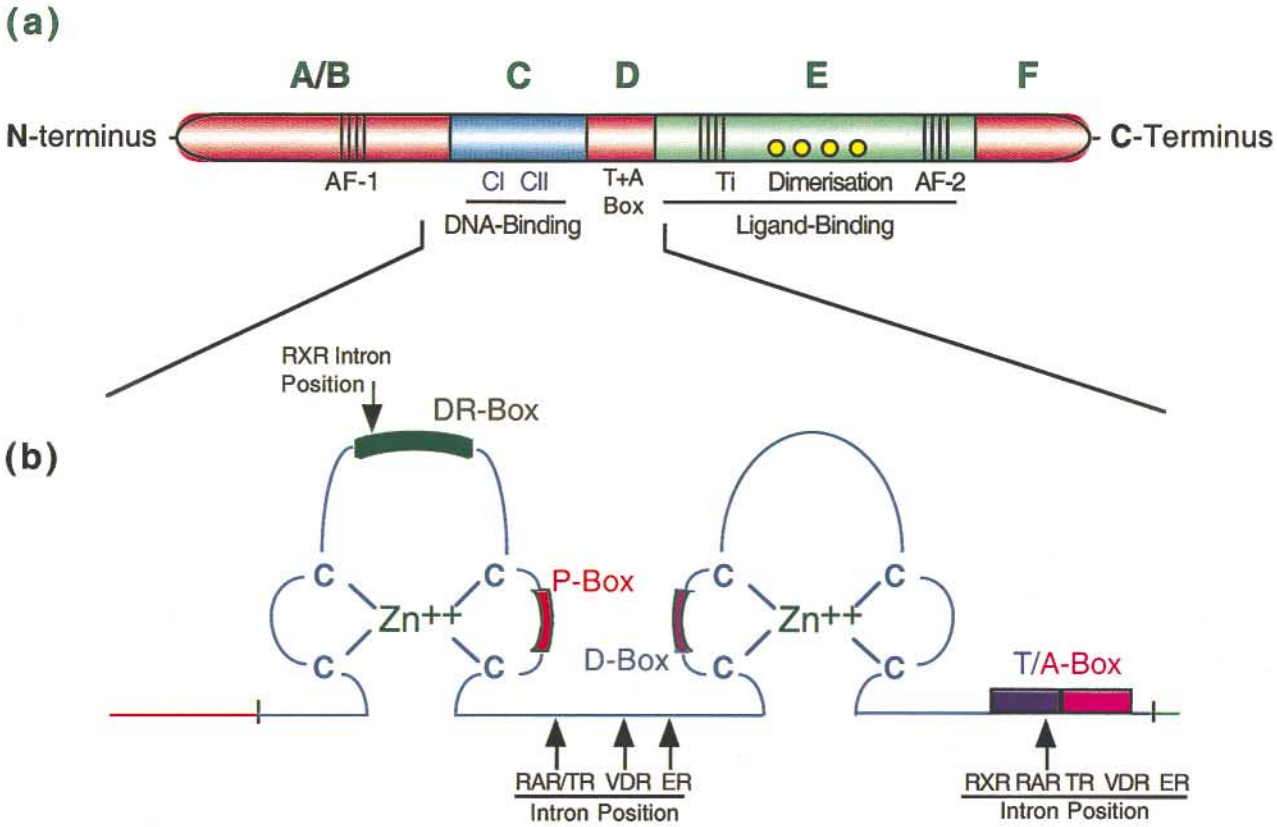
NR structure

The NRs constitute a superfamily of transcription factors, which have been classified on the basis of their conserved structural domains (fig. 1a). Recently, a standardised nomenclature that is based on evolutionary sequence conservation between different (or homologous) receptors in various species has been proposed [43]. This new nomenclature is used in figure 2 along with the trivial names of each receptor. The NR size varies considerably from 427 amino acids in the VDR to 1237 amino acids in the *Drosophila* E75 receptor [44]. Despite these size variations, NRs share common structural/functional domains. A canonical NR possesses five to six such functional regions named A through E/F (fig. 1a) [for comprehensive reviews on NR structure see refs 1, 44, 45]. The most highly conserved region among the various NRs is the DNA-binding domain (DBD) (region C), which facilitates sequence-specific interaction with the major groove of the double helix. The E region, with the ligand-binding domain (LBD), displays the next highest degree of conservation and the least conserved are domains A, B, D and F.

The DNA-binding domain

The DBD consists of two cysteine-rich zinc finger motifs (CI and CII), which appear to have arisen together as a single unit and not as the fusion of separate domains or a duplication of an ancestral gene [46]. This is despite the fact that the CI and CII zinc fingers are, in most NR genes, encoded by different exons [47, 48]. In addition to the zinc fingers, several other regions of the DBD show high degrees of sequence conservation. The P-box lies between the last two cysteines of the CI zinc finger and confers target DNA specificity (fig. 1b, c). NRs such as RAR, RXR, VDR, TR and ER have a P-box which confers binding to the AGGTCA sequence (referred to as the ER P-box group), while the GR, mineralocorticoid receptor (MR), androgen receptor (AR) and PR have adapted to recognise a AGAACA sequence (referred to as the GR P-box group) [49, 50]. Mutation of three P-box residues in the ER to the corresponding residues in the GR, switches its DNA-binding preference to that of the GR [51, 52]. It is interesting to speculate that mutations in an ancestral receptor ER P-box gave rise to a protein which recognised the sequence AGAACA, instead of AGGTCA, and that such a 'mutant' receptor gene became free to acquire further changes, resulting in the formation of the PR, GR, AR and MR. Analysis of the phylogenetic family tree shows that the ER P-box group is scattered throughout the superfamily, while the GR P-box group is tightly confined to sub-family III (fig. 2), suggesting it is the ER group which possesses the ancestral P-box.

Laudet [50] observed that the presence of certain P-box sequences correlated with the ability of an NR to heterodimerise with the RXR, and suggested that this ability arose in the early metazoans and has diversified during evolutionary history. Exceptions to this hypothesis are the germ cell nuclear factor (GCNF) and COUP-TF. GCNF, which has an identical P-box to RXR (CEGCKG), fails to heterodimerise, presumably due to variations in the sequences lying C-terminal to the P-box. COUP-TF on the other hand, heterodimerises despite having a diverse P-box (fig. 1c). While mutating both ER and GR P-box positions, Zilliacus and colleagues [49] observed that substitutions of certain amino acids resulted in receptors which either bound their respective response element with increased affinity or with increased fidelity, thus suggesting that through natural selection such mutations may be incorporated by members of the NR superfamily. Interestingly, one of the P-box mutations reported by the authors, which led to broad-specificity binding, has been found in the liver-enriched factor HNF4 and in the *Drosophila* tailless receptor (tll) [49]. To date, 76 distinct P-boxes have


(a)
A/B
C
D
E
F
N-terminus
C-Terminus
AF-1
CI CII
DNA-Binding
T+A Box
Ti
Dimerisation
AF-2
Ligand-Binding
(b)
RXR Intron Position
DR-Box
Zn++
P-Box
D-Box
T/A-Box
RAR/TR VDR ER
Intron Position
RXR RAR TR VDR ER
Intron Position

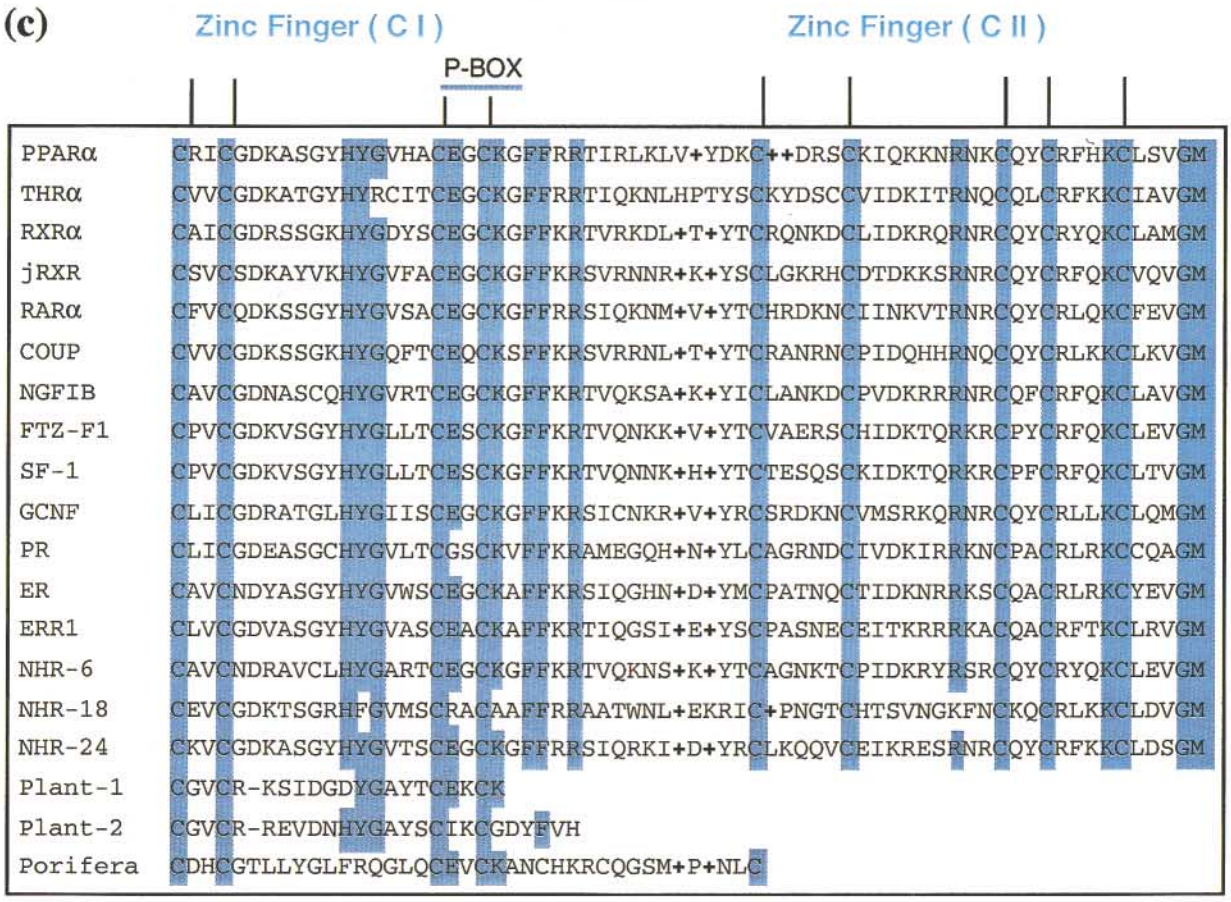

(c)
Zinc Finger (C I)
Zinc Finger (C II)
P-BOX
PPARα CRICGDKASGYHYGVHACEGCKGFFRRTIRLKLV+YDKC++DRSCKIQKKNRNKCQYCRFHKCLSVGM
THRα CVVCGDKATGYHYRCITCEGCKGFFRRTIQKNLHPTYSCKYDSCCVIDKITRNQCQLCRFKKCIAVGM
RXRα CAICGDRSSGKHYGDYSCEGCKGFFKRTVRKDL+T+YTCRQNKDCLIDKRQRNRCQYCRYQKCLAMGM
jRXR CSVCSDKAYVKHYGVFACEGCKGFFKRSVRNNR+K+YSCLGKRHCDTDKKSRNRCQYCRFQKCVQVGM
RARα CFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNM+V+YTCHRDKNCIINKVTRNRCQYCRLQKCFEVGM
COUP CVVCGDKSSGKHYGQFTCEQCKSFFKRSVRRNL+T+YTCRANRNCPIDQHHRNQCQYCRLKKCLKVGM
NGFIB CAVCGDNASCQHYGVRTCEGCKGFFKRTVQKSA+K+YICLANKDCPVDKRRRNRCQFCRFQKCLAVGM
FTZ-F1 CPVCGDKVSGYHYGLLTCESCKGFFKRTVQNKK+V+YTCVAERSCHIDKTQRKRCPYCRFQKCLEVGM
SF-1 CPVCGDKVSGYHYGLLTCESCKGFFKRTVQNNK+H+YTCTESQSCKIDKTQRKRCPFCRFQKCLTVGM
GCNF CLICGDRATGLHYGIISCEGCKGFFKRSICNKR+V+YRCSRDKNCVMSRKQRNRCQYCRLLKCLQMGM
PR CLICGDEASGCHYGVLTCGSCKVFFKRAMEGQH+N+YLCAGRNDCIVDKIRRKNCPACRLRKCCQAGM
ER CAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHN+D+YMCPATNQCTIDKNRRKSCQACRLRKCYEVGM
ERR1 CLVCGDVASGYHYGVASCEACKAFFKRTIQGSI+E+YSCPASNECEITKRRRKACQACRFTKCLRVGM
NHR-6 CAVCNDRAVCLHYGARTCEGCKGFFKRTVQKNS+K+YTCAGNKTCPIDKRYRSRCQYCRYQKCLEVGM
NHR-18 CEVCGDKTSGRHFGVMSCRACAAFFRRAATWNL+EKRIC+PNGTCHTSVNGKFNCKQCRLKKCLDVGM
NHR-24 CKVCGDKASGYHYGVTSCEGCKGFFRRSIQRKI+D+YRCLKQQVCEIKRESRNRCQYCRFKKCLDSGM
Plant-1 CGVCR-KSIDGDYGAYTCEKCK
Plant-2 CGVCR-REVDNHYGAYSCIKCGDYFVH
Porifera CDHCGTLLYGLFRQGLQCEVCKANCHKRCQGSM+P+NLC

Fig. 1.

been identified within the superfamily. Six of these are found throughout the metazoan kingdom, 7 are limited to vertebrates, with the remaining 63 found only in nematodes [53]. One of these novel P-box amino acid sequences (CRACAA; fig. 1c), which suggests novel binding specificity, is found in a third of predicted nematode NR genes.

The structure of the NR response element has also been highly conserved. A 19-nucleotide motif composed of two inverted RGGTCA sites with a seven-nucleotide spacing is present in NR promoters from echinoderms to mammals [54]. As with mammalian RXR, the arthropod homologue USP, when dimerised with EcR, can bind to both inverted and direct repeats with variable spacing [55]. Recent studies suggest that the RXR can form a tetramer in solution which is dissociated into a dimer upon ligand binding. These tetramers demonstrated the ability to bind both direct repeat and palindromic response elements [56]. As mentioned above, amino acids in the P-box confer half-site recognition while another region in the second zinc finger (the D-box) confers the selection pattern of half-site spacing (fig. 1b) [52].

The fact that RAR, RXR, TR and VDR bind as dimers to asymmetrical HREs suggests that their DBDs must possess structural features previously not described for steroid receptors. Indeed, a number of recent studies have identified novel functional determinants within the CI module (DR-box) and the very N-terminal sequences of the D region (A- and T-boxes) (see fig. 1b). The amino acids in the DR-box of the TR, RAR and VDR constitute an asymmetrical dimerisation interface and are critical for discrimination between different spacing of the DRs [57]. The amino acids in the T- and A-boxes mediate additional protein-protein and protein-DNA interactions probably necessary for better recognition of DRs with an appropriate spacing [58, 59] and regulation of differential orientations of the DBD on response elements with different symmetries, respectively. Interestingly, in the heterodimeric NR complexes which are bound to asymmetrical response elements, the RXR always occupies the 5′ half of the DR [7, 57, 60] (see fig. 1b).

The evolutionary acquisition by nuclear receptors of the ability to heterodimerise clearly increased the diversity of regulatory functions that these factors can mediate. This functional diversification through heterodimerisation of transcriptional regulators is not unique to NRs and can be observed among other families of transcription factors such as members of the Jun-Fos and ATF-CREB families [61, 62] as well as myogenic helix-loop-helix proteins [63, 64]. Coevolution of the DBDs of the receptors and the response elements appears to have given a very large number of combinatorial possibilities through which a plethora of genes or gene networks could be differentially regulated by a limited number of ligands and NRs.

The specificity of monomeric binding, which is prevalent among orphan receptors, is conferred by an A/T-rich region at the 5′ end of the response element [65]. Monomeric binding is neither restricted to orphan receptors nor to one branch of the NR family tree [37]. Monomeric sequence is utilised by sub-family I (rev-erb), sub-family III (oestrogen-like receptor, ERR1), sub-family IV (nerve growth factor, NGFIB) and sub-family V (steroidogenic factor-1, SF-1) (fig. 2). The orphan receptor and sub-family IV member, GCNF1, binds on a direct repeat sequence [66, 67]. Interestingly, the TR (sub-family I), which binds DNA as a heterodimer on everted repeats and as a homodimer and a heterodimer with RXR to direct repeats, can also bind as a monomer to A/T monomeric sequence [1] (fig. 2). Conversely the ER has been reported to bind as a monomer to the thyroid HRE consisting of an inverted palindrome without repeats [68] as well as binding as a monomer to half-site oestrogen-responsive elements (EREs) [69, reviewed in ref. 45]. Most sub-families possess monomeric binding to extended half-sites suggesting this may have been the ancestral mechanism of receptor binding conferring a biological response [70]. This hypothesis strengthened by the observation that the two earliest reported receptors FTZ-F1 and RXR are capable of monomeric binding [70, 71]. The jellyfish RXR (jRXR), which in vivo appears not to have a heterodimerisation partner, is reported to bind an extended half-site in the jellyfish crystallin gene and to

Figure 1. Conserved functional domains of the NRs. (*a*) The N-terminal (A/B) domain, the DNA-binding domain (C), a variable hinge domain (D), the ligand-binding domain (LBD) (E) and the C-terminal domain (F) are as shown. In the A/B domain, the ligand-independent activation function (AF-1) is indicated. Within the LBD, the conserved Ti domain, the residues utilised in dimerisation and the ligand-dependent activation function (AF-2) are shown. (*b*) Detail of the two zinc fingers (CI and CII) demonstrating the location of the DR-, P-, D- and T/A-boxes along with the location of exon/exon boundaries of selected NRs. (*c*) DNA-binding (C domain) sequence encompassing the DNA-binding zinc fingers CI and CII. Highly conserved residues are shaded and conserved cysteine residues marked by black lines. All sequences are human except for jRXR (accession number AAC80008) which is from the Cnidaria jellyfish, FTZ-F1 (accession number P33244) which is from *Drosophila*, NHR-6 (CNR-8/ceb-1 accession number AAD03682), NHR-18 (accession number AAD03690) and NHR-24 (CNR14/sex-1 accession number I45066) which are all nematode. Plant-1 and Plant-2 are putative zinc finger proteins from the plant *Arabidopsis thaliana* (accession numbers AAC28517 and AAD19774, respectively). The Porifera sequence is from the *Sycon raphanus* serine/threonine protein kinase (accession number CAA73557).

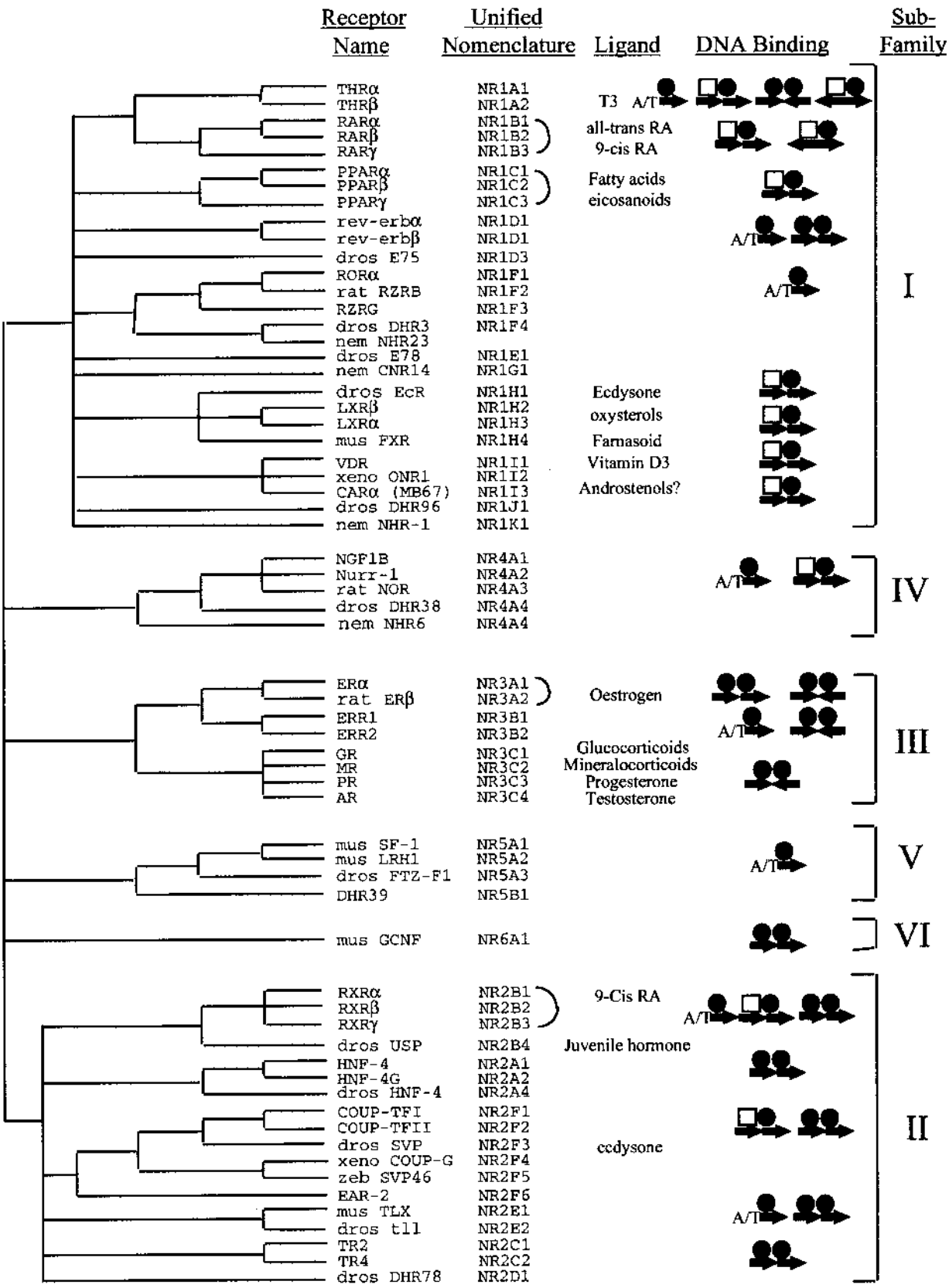


Figure 2. Consensus phylogenetic tree as reported by Laudet [50] (reproduced with permission). The black arrows represent the number and orientation of the DNA-binding core motif. The black circles represent the nuclear receptor, while the open square represents the ability to heterodimerise with the RXR. Identified receptor ligands are shown. All nuclear receptors used in this tree are human unless stated otherwise (dros, *Drosophila*; rat, rat; mus, mouse; zeb, zebrafish; nem, nematode). The unified nomenclature [43] of so-far classified receptors is shown. Sub-family classification is represented on the far right-hand column.

bind as a monomer to the direct repeat sequences with one-, four- or five-base pair spacing [71].

The ligand-binding domain

The LBD is functionally complex and possesses sub-regions implicated in ligand binding, dimerisation and transcriptional regulation. There are two highly conserved regions in the LBD. First, the activation function core motif (AF-2) which is involved in ligand-mediated transactivation by direct recruitment of a co-activator. This region is dependent on the presence of either one or two glutamic acid residues (E) in the motifs LLLEMLD in the ER or LIQEMLE in the RAR [72–76]. The ligand-dependent activity of the NRs involves the recruitment of co-activator binding with the incorporation of chromatin-remodelling factors [77]. The recruitment of co-activators such as SRC [19, 12] occurs through an LXXLL core consensus sequence referred to as LXDs. McInerney and colleagues [78] reported that only one LXD domain of the NCoA1/SRC-1 coactivator was required for the activation of the ER, while several LXD domains with the appropriate spacing were needed for the activation of the RAR, PR and TR. They suggest that LXXLL-containing motifs have evolved to enable the recruitment of co-activator complexes in both ligand- and receptor-specific manners. Another highly conserved sub-region within the LBD is the Ti domain (fig. 1a) with the central sequence WAKA or FAKK, as characterised in the ER or RAR, respectively [50, 79]. Mutations in the LBD might have given rise to receptors with either increased specificity for a given ligand or potential for binding of novel ligands. Thornton and Kelley [80] suggested that the presence of a unique threonine in the LBD of an ancestral AR sequence may have been the key event in the emergence of a receptor that binds testosterone specifically but not other steroids [81].

The earliest report of a ligand-binding NR was the jRXR [71]. The jRXR has higher homology with the human RXR (78% homology in the DBD and an astonishing 79% homology in the LBD) than with the *Drosophila* RXR homologue USP (figs 3, 4). The arthropod ixodid tick possesses two RXR isoforms which have a DBD closer to that of USP yet have LBDs closer to vertebrate RXRs suggesting that the LBD of the RXR has been well conserved from cnidarians to vertebrates and through the early branch that gives rise to the arthropod lineage [82] (figs 3, 4). Neither the USP, nor the tick RXRs, are thought to bind the ligand 9-*cis*-retinoic acid, while the jRXR is capable of binding this compound with higher affinity than the mammalian RXR [71]. Curiously, and despite being able to bind ligand, the AF-2 motif is not present in the jRXR and to date the mode of action of this receptor is unknown. This may either suggest that the important role of the RXR in mammals is as a heterodimerisation partner or, possibly, that 9-*cis*-retinoic acid is not the natural or unique ligand for the RXR. The presence of the jellyfish RXR receptor places the acquisition of ligand binding close to the base of the metazoan tree, which is a lot earlier than perhaps anticipated [50]. Garcia-Vallvé and Palau [83] observed that despite high variability between amino acid residues of the LBD there is strong conservation in the α-helical secondary structure. The authors proposed a hypothesis suggesting that mutations which alter secondary structure are not recognised and are thus discarded by molecular chaperones which are present in solution, while mutations which preserve the canonical secondary structure are accepted even if the ability of the receptor to perform its functional role is lost therefore allowing evolutionary changes to be gathered by the protein [83]. An important source of ligands (or their precursors) are dietary factors. Baker [84] reported on the endocrine activity of plant-derived compounds and that various flavinoids act in mammals by binding to ER, while Yamamoto [85] similarly speculated that metazoans evolved NRs to utilise an environmental compound or metabolite. ERR-1, currently regarded as an orphan receptor, has recently been shown to both bind and be transcriptionally activated by organochlorine pesticides [86]. Changes in environmental conditions, particularly leading to dietary changes of metazoans, may have lead throughout evolution to the acquisitions of new ligands and, hence, alterations in signalling pathways and development.

The less conserved regions A, B, D and F

The regions A and B are commonly referred to as the A/B domain and possess a ligand-independent activating function (AF-1). AF-1 functions in a promoter- and/or cell-context-specific manner and co-operates with AF-2 in regulation of transcription [87]. A number of recent papers have demonstrated that activity of AF-1 is regulated by phosphorylation [88–90]. In addition, the A/B domain of TRα and PPARγ regulate their DNA- and ligand-binding affinities, respectively, in a phosphorylation-dependent manner [91, 92]. Phosphorylation of the A/B domain facilitates ligand-independent DNA binding of the ER [93], while interaction between PPARγ2 AF-1 and the co-activator protein PGC-2 appears to be required for induction of fat cell differentiation through PPARγ agonists [94].

Although the A/B domain displays no sequence conservation between different NRs, for a given receptor it has been well conserved through evolution. Furthermore, the B (but not A) region has also been conserved among NRs that can be considered as paralogues

(paralogous genes arise by duplication within an ancestral species, e.g. RAR$\alpha\beta\gamma$). In the case of steroid receptors, however, the A/B domain of the AR is completely divergent from those of the GR and other steroid receptors, suggesting that this region in the AR might have been acquired more recently in evolution, perhaps as recently as the split between amphibians and mammals [80]. On these grounds, Garcia-Vallvé and Palau [83] speculated that the A/B domain was part of a single-copy ancient gene, which was incorporated into a locus encoding an NR. A possible example supporting such a hypothesis is seen in the plant *Arabidopsis thaliana*, where a predicted protein possesses similarity to the A/B domain of the RXRβ [95].

The D domain was originally thought to serve as a flexible hinge region between the adjacent C and E regions. In many receptors, this region also contains a nuclear localisation signal. Studies addressing structural requirements of DNA and ligand-binding domains revealed that the extremities of the D region are parts of these adjacent domains. For example, the N-terminal part of the D domain contains the T- and A-boxes which are involved in conferring dimerisation and half-site recognition respectively (fig. 1a), whereas the C-terminal part possesses the region responsible for ligand-regulated interactions between receptors and a co-repressor [96]. It is worth noting that both the C- and N-terminal regions of the D domain that carry important functional information have been conserved among various NRs, particularly well between paralogues, whereas the central region has not.

The very C-terminal F domain, which is absent in some NRs (RXRs, for example), is poorly conserved and its function is not well understood. In the steroid hormone receptors, the F domain has been shown to confer ligand specificity and influence transcriptional activation by the ER [97], as well as to modulate interaction between HNF-4 and its co-activator [98].

The NR phylogenetic tree

Several publications have reported phylogenetic trees of either the superfamily as a whole [1, 46, 49, 50, 83, 99–101] or the steroid hormone family [39]. This review reproduces, with some minor modifications, the consensus tree constructed by Laudet [50] (fig. 2), which along with the phylogenetic trees proposed by Detera-Wadleigh and Fanning [101] and Garcia-Vallvé and Palau [83] represents one of the most comprehensive analyses to date. It must be stressed that figure 2 is not an evolutionary tree, but represents the relationship between NRs based on sequence alignment of the DNA- and ligand-binding domains. Also of importance are the phylogenetic trees based upon sequence alignment of either the DBD or LBD alone [46, 83]. Figure 2 classifies the NR superfamily into six sub-families. This classification is in general agreement with that of other authors, although there are some areas of disagreement. Detera-Wadleigh and Fanning [101], in a comparison of trees generated using different algorithms, presented RAR and RXR grouped together, while Laudet [50] and Garcia-Vallvé and Palau [83] place the RAR and RXR in separate sub-families. The positioning of the PPAR also varies, from being placed into a separate family by Detera-Wadleigh and Fanning, to being grouped with the RAR and TR in other reports. Another major discrepancy involves the positioning of the tll receptor and HNF4, which are both placed in sub-family II by Laudet [50], but in distinct sub-families by Detera-Wadleigh and Fanning [101] who used a different algorithm for their analyses. Similarly, Garcia-Vallvé and Palau [83], using different algorithms, have placed the TR and RAR into separate sub-families than Laudet [50]. The arthropod orphan receptors THR4 [102] and GRF [Charles et al., unpublished accession AAD38900] cluster in phylogenetic trees with GCNF, which until recently was regarded as the sole member of sub-family VI (see fig. 4). GCNF, which has a restricted brain-specific expression during mouse development and then an exclusively germ cell expression in the adult, appears to possess unique DNA-binding and dimerisation properties, suggesting the use of novel mechanisms to regulate target gene expression [103]. Garcia-Vallvé and Palau [83], on the basis of amino acid sequence conservation in the DBD, placed this receptor in the vicinity of sub-family I. However, when sequence conservation in the LBD is examined, GCNF is closer to sub-families II and V than to I. The position in the phylogenetic tree is speculative for the orphan receptors such as the KNI, KNRL, EGON and the nematode ODR-7, which possess only a DBD, or the potentially paralogous receptors DAX-1 and SHP-1, which possess only a recognisable LBD [104–107]. It has been speculated that these domains may be the evolutionary building blocks of the NR ancestor [108]. Although this is an attractive hypothesis, due to their location within well-defined sub-families, these single-domain receptors are more likely to have arisen later in evolution through aberrant splicing events from an ancestral receptor [50]. These receptors have been conserved and incorporated into new regulatory pathways and in some instances their expression may be regulated by other members of a sub-family from which they diverged. For example, positive effects of SF-1 on expression of the murine DAX-1, which clusters within sub-family II on phylogenetic trees [50], is antagonised by sub-family II member COUP-TF1 [109].

NRs in the metazoan family tree

Metazoan life arose from mitochondrial-containing 'crown' Eukaryota, which through a multicellularisation step also gave rise to plants and fungi. Figure 3 represents a highly simplified evolutionary route taken between the rise of the first metazoan, around 1000–800 million years ago (mya) and the first appearance of mammals in the early Mesozoic era (200–160 mya) [110]. Screening of the nucleic acid sequence databases, which are representative of various organisms, has led to the detection of NRs only in the metazoans [50, 100]. The first characterised branch of the metazoan tree (the Porifera), which includes the living sponges, also failed to demonstrate the presence of NRs [100] (fig. 3). Although the skeleton of the NR first zinc finger (containing the four cysteines and the P-box) is present in some non-NR proteins throughout the plant and animal kingdoms, including a putative zinc finger protein from the plant *A. thaliana* and in the Porifera serine/threonine protein kinase (fig. 1c), there is insufficient evidence to speculate that genes encoding such proteins provided a building block in the construction of the NR ancestor. The most primitive organisms with genes encoding NRs (FTZ-F1, COUP-TF and RXR) are diploblastic Cnidaria, which incorporates the Coelenterata and includes the hydra, jellyfish and anemones. The FTZ-F1, COUP-TF and RXR may, therefore, be considered the most evolutionary conserved and perhaps closest, from among the known members of the superfamily, to the NR ancestor. The FTZ-F1 orthologues SF-1 (human), LRH (mouse) and DHR39 (arthropod) all contain the same unique P-box and a 30-amino-acid basic region abutting the C-terminal end of the zinc finger motif designated the FTZ-F1 box [111], as well as binding to the same sequence as monomers (fig. 2). SF-1 plays important roles in adrenal gland development and differentiation of male genitalia [37, 112]. The COUP-TF has been well conserved throughout metazoans with over 50% sequence identity between the arthropod, nematode and human (fig. 4). On the calculation of evolutionary rate, Laudet [50] observed that COUP-TFI has evolved extremely slowly [0.0073 Pauling units (PAU) as opposed to an average of 0.3 PAU], suggesting that this protein is strictly required for development. In this regard, it is noteworthy that mice lacking COUP-TFI die in utero while mice lacking COUP-TFII die shortly after birth [113, 114]. Although the jRXR does not cluster in phylogenetic trees with any of the vertebrate RXR paralogues (or any other RXR metazoan orthologue), RXR from Urochordata (figs 3, 4) shares 100% homology in the DBD and 92% homology in the LBD and clusters with the human RXRα [Kamimura et al., unpublished accession number BAA82618] (fig. 4). Interestingly, it is the knockout of RXRα, and not RXRβ or γ, which leads to embryonic lethality [114, 115] (fig. 4). With the knockout of the FTZ-F1-related murine SF-1 receptor leading to early post-natal lethality, it is tempting to suggest that the descendants of the earliest NRs tend to assume more important and non-redundant physiological roles. Since COUP-TF and RXR are members of family II, and FTZ-F1 is a member of family V (fig. 2), one can assume that the archaic NR was a precursor of these two families.

Metazoan phylogeny is currently undergoing a state of upheaval trying to reconcile the fossil records with the molecular genetic interpretations [116]. The next branch of the metazoan tree after the Cnidaria is where evolutionary relationships are most controversial. The established hypothesis saw the next major separation, before the rise of triploblastic organisms, being the large grouping of pseudocoelomates, which include the roundworm nematodes and the well-studied *Caenorhabditis elegans*. The following branch saw the protosomates, which contains the Arthropoda/Platyhelminthes/Mollusca lineage diverge, from the deuterosomates, which include the Echinodermata and Chordata. However, recent studies from Hox gene phylogeny propose the grouping of nematodes within the protosomates [116, 117]. This interpretation is presented in figure 3. Despite a substantial gene loss (as seen from Hox gene analysis [117]), the number of predicted NRs in the nematode roundworm is currently 228—fivefold higher than the 44 receptors identified so far in humans [53]. The majority of these NRs appear to have arisen through proliferation and diversification of one chromosome and many fall into phylogenetically unconserved groupings that are still in the process of identification [53]. Of the nematode receptors characterised to date, representatives from four of the six subfamilies appear to be present (figs 3, 4). The only absentees in the nematode are members of sub-families III and VI (figs 3, 4). In contrast to COUP-TF and FTZ-F1/SF-1, an orthologue of RXR was not identified in the nematode [53]. This is surprising, as RXR is present in the Cnidaria, arthropods and vertebrates (figs 3, 4). It is, nevertheless, possible that the nematode RXR orthologue has been lost along with some nematode Hox genes [117] during the evolution process.

As indicated by the star in figure 3, the path from the protodeuterosomate ancestor to Arthropoda indicates divergence of the Platyhelminthes (flatworms), Mollusca (snails, clams), Annelida (segmented worms) and Nematoda. A recent report demonstrated the presence of two constitutively expressed RXR homologues in the platyhelminth blood fluke *Schistosoma mansoni* (RXR-1 and 2) [118, 119]. Although classified as RXR family members on the basis of their high degree of amino acid sequence conservation in the DNA- and ligand-binding

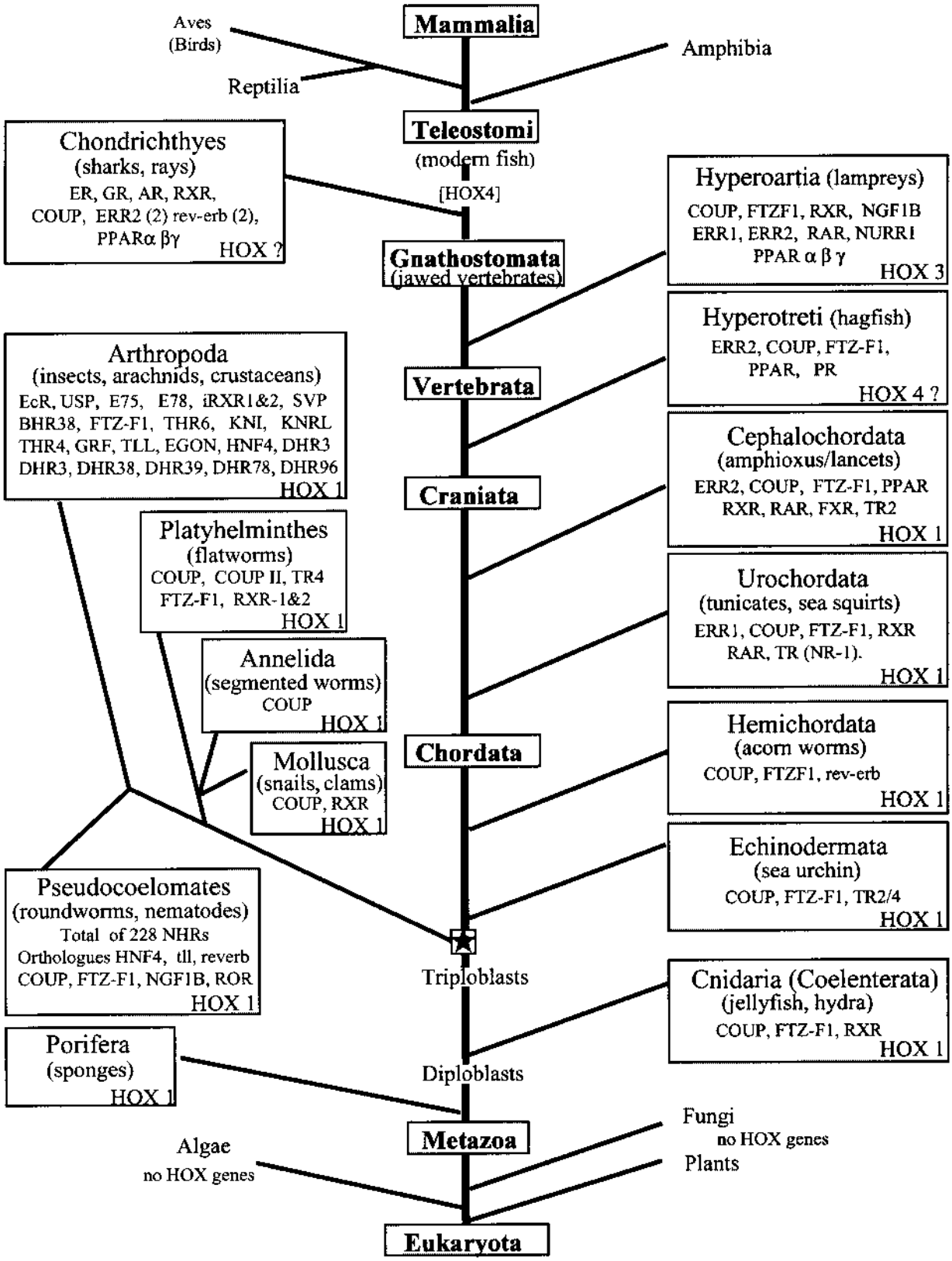


Figure 3. The Metazoan family tree. A simplified evolutionary family tree for the progression from early eukaryotes to mammals. This figure is not to scale and it should be noted that the NRs represent only those receptors that have been identified to date, and does not suggest the absence of any receptor in each class. The suspected HOX cluster number in each division is shown. The boxed star represents the diversification of the protosomes and the deuterosomes.

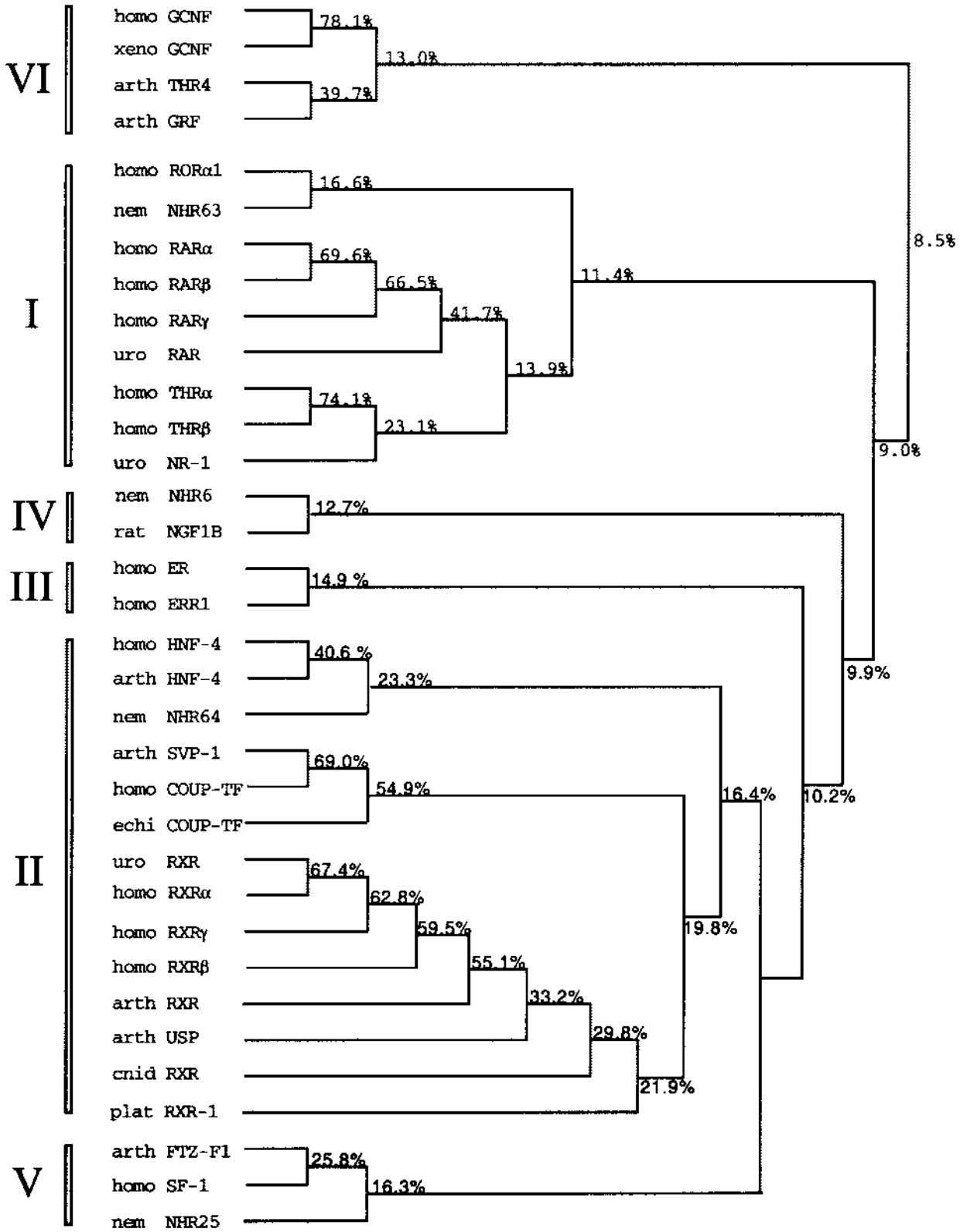


Figure 4. Phylogenetic tree indicating distribution of the six NR sub-families in metazoans. The tree was constructed with full-length NR sequences available in GenBank within the Mac DNASIS pro v3.6 program using the Higgins-Sharp algorithm (CLUSTAL4). The numbering on the branches corresponds to similarity scores which are calculated as the number of exactly matched residues. Abbreviations used are: homo, *Homo sapiens*; xeno, *Xenopus*; uro, Urochordata; arth, Arthropoda; nem, Nematodata; echi, Echinodermata; cnid, Cnidaria; plat, Platyhelminthes. Roman numerals represent the NR sub-family.

domains, these receptors are longer than either their cnidarian or vertebrate relatives [118] (fig. 4). The RXR is extensively represented in the phylum Arthropoda having been identified in the silkworm, the crab (arthRXR in fig. 4), the tick (possessing two isoforms), and in the fruit fly (USP). Interestingly, the crab and tick RXR sequences resemble the vertebrate RXR more than that of the Platyhelminthes, suggesting that the modifications which occurred to the blood fluke RXRs did so after the divergence from the lineage resulting in arthropods (figs 3, 4). The presence of two RXR genes in both Arthropoda and Platyhelminthes suggests a duplication of the RXR gene shortly after the protodeuterosome split, with the *Droposphila* USP possibly arising from the acquisition of the ability to bind juvenile hormone. The recent finding of both sub-family III and VI members in arthropods demonstrates that NRs had already diversified into all six sub-families before the Arthropoda-Chordata (protodeuterosome) split (indicated by the star in fig. 3). Without complete sequence data from earlier organisms, the possibility that members of family I, III, IV and VI are also present in cnidarians can not be excluded at the present time. Likewise, finding NRs in earliest metazoans, such as Porifera (sponges), still remains a reasonable possibility.

The arthropod protosomate ancestor diverged from the deuterostomates, which include the Echinodermata (sea urchin, starfish), Hemichordata (acorn worms) and Chordata (fig. 3). The echinoderms, which possess COUP-TF, RXR and FTZ-F1, evolved from their ancestor by losing their ancestral locomotive tail and gill slit, while the chordates appear to have evolved with the development of a notochord [120]. The Hemichordata, which possess COUP-TF, FTZ-F1 and rev-erb, are represented in this family tree by a branch before the formation of the true Chordates. They share many common features with the chordates, such as a dorsal nervous system and pharyngeal arches, but lack the presence of a brain or notochord. The earliest Chordata, defined by the presence of a notochord at some stage in their life history, date from 525 mya and are currently placed within the sub-phylum Urochordata which include the tunicates, sea squirts and salps. It is in the sub-phylum Urochordata where a TR-related member of sub-family I, the tunicate receptor NR-1 [121], is found (figs 3, 4). Although NR-1 is 86% and 58% homologous in its DNA- and ligand-binding domains, respectively, to the human TR, it does not appear to bind thyroid hormone and does not possess the AF-2. The absence of the AF-2 in NR-1 is puzzling as this domain is present in the tunicate RAR [Hisata et al., unpublished accession BAA25569; see fig. 3] and in the *Drosophila* USP, which diverged nearly 200 mya prior to the emergence of tunicates. Therefore, NR-1 is either a divergent receptor and not the direct ancestor of the vertebrate TR or the TR AF-2 has arisen independently in the early vertebrate. As the authors discussed [121], it remains paradoxical that in the tunicate, which synthesises thyroid hormone involved in morphogenesis, a potential TR ancestor exists which does not have the ability to bind it [121]. The Urochordata also possess an RAR homologue [Hisata et al., unpublished accession BAA25569; fig. 4] which shows 94% and 77% homology in the DBD and LBD, respectively, with vertebrate RAR receptors and contains sequence suggesting that it binds ligand and possesses a functional AF-2.

It is noteworthy that the steroid hormone receptors, grouped in sub-family III, appear to lack homologues in lower metazoans. The earliest characterised member of sub-family III is the *Drosophila* oestrogen-like receptor (ERR) [53] (fig. 3). An ERR1 homologue has also been described in Urochordata [100]. The ERRs and steroid hormone receptors (ER, PR, GR, MR, AR) are believed to share a common ancestor, with the first steroid hormone receptor being a chordate, possibly craniate innovation (fig. 3) [39]. In accordance with the above, a further divergence of sub-family III, with ERR2 in the Cephalochordata amphioxus, is observed before the appearance of the PR in the Hyperotreti member hagfish [100]. The path taken between the early Craniata and the rise of modern fishes appears to have resulted in the appearance of the intermediate sub-phylum Hyperotreti and Hyperoartia as studied by work in the hagfish and the lamprey, respectively (fig. 3). The hagfish are eel-shaped jawless fishes differing from other vertebrates by the absence of extrinsic eye muscles, eye lens, cardiac innervation and radial muscles. The lamprey is also an eel-shaped jawless fish which possesses an eye lens, a thicker spinal cord and a true cartilaginous braincase, yet is devoid of the mineralised skeleton that is present in higher fish [122, 123]. Despite the detection of the PR, but not ER, in hagfish, there is a strong consensus from phylogenetic trees and P-box sequence analysis that the ER predates the PR and, thus, the absence of the ER is presumed to be merely a sampling artefact or the ER has been lost by the hagfish (fig. 3). The sub-family I member PPAR is also present in hagfish, with all three paralogues (α, β, γ) identified in the lamprey [100] (fig. 3).

Although the NR superfamily was diverse at the time of the Arthropoda-Chordata split and already possessed members from all six sub-families, the family went through a substantial increase in number during the Cambrian period ('Cambrian explosion of life') around 550–500 mya [110, 124].

Genome amplification during the cambrian explosion

Further advances in understanding developmental and evolutionary biology came with the discovery that vertebrate Hox genes are related to the homeotic selector (HOM) genes of *Drosophila* [for a review see Holland and Garcia-Fernandez [125]]. Hox genes, which are arranged in a well-conserved cluster, encode a family of helix-turn-helix transcription factors which operate in a regulatory, possibly self-regulatory cascade, determining embryonic orientation and development. Mammalian Hox genes are homologous to clustered Hox genes present in arthropods suggesting that the gene family arose through a series of cluster duplications to give the current higher vertebrate four-cluster state (termed Hox A–D). Further evidence suggesting genome duplication during evolution comes from analysis of gene numbers in different species. The arthropod *Drosophila* and the pseudocoelomate nematode possess approximately 12,000 and 16,000 genes respectively, while humans have an estimated 70,000 genes [126]. To date, a single Hox cluster has been identified in all branches of the metazoan family tree up to and including the Cephalochordata (designated HOX 1 in fig. 3). Higher vertebrates all possess four Hox clusters, suggesting that this four-cluster state arose during early vertebrate evolution during a period now referred to as the Cambrian explosion. If Hox genes operate in a combinatorial manner, then an increase in Hox gene number would be coupled with the formation of increasingly complex body plans. NRs, which increased in number in tandem with the Hox clusters, have been shown to control complex developmental processes in a variety of organisms, sometimes by directly targeting Hox gene expression [125, 127–130].

The mechanism by which one Hox gene cluster became four, and thus how the NR paralogues arose, is an issue of much debate. The major question concerns whether two genome duplications, occurred thus giving a quadruplication of the genome and potentially four copies of each gene or whether a three-step or multistep process took place. The scattered nature and unequal copy number of NRs, with two to four paralogues (fig. 3), is consistent with both theories on vertebrate genome duplications. In favour of the double-genome duplication hypothesis, Baker [39] proposed that in the steroid hormone family (sub-family III) an ancestral receptor of the GR-PR-AR-MR grouping present in Cephalochordata duplicated once to give GR/MR and AR/PR ancestors, and then again, followed by diversification within each group to give the four separate receptors. As previously mentioned, due to the absence of this receptor in fish, the MR is believed to be the last of the NRs to have been brought under tight evolutionary control [131]. It is also of interest that aldosterone, a ligand for the MR, is not known to be present in fish [80]. This suggests that despite the differing nomenclature of the four receptors, the GR, PR, MR and AR are paralogues located on separate chromosomes. This observation is noted in the new nomenclature system for the superfamily [43] where these four receptors are classified as NRC1–4 (fig. 2). This double-genome duplication hypothesis proposes a maximum of four paralogous receptors being present in vertebrates, and thus the recent cloning of a novel androgen receptor paralogue (ARβ) from the Japanese eel may seem unexpected [132]. Nevertheless, one cannot exclude the possible occurrence of an independent duplication and translocation of the eel AR gene giving rise to this paralogue. One also cannot exclude that, in analogy with zebrafish, the genomes of lineages leading to the Japanese eel might have undergone rapid expansions [133]. It remains to be seen whether the ARβ is present in mammals.

As an alternative to the double-genome duplication hypothesis, a four-cluster state may have been achieved through one round of genome duplication followed by duplication of partial blocks of the chromosomes [134, 135]—a so-called three-step or multistep hypothesis. In line with such a hypothesis, Pendleton and colleagues [136] and Sharman and Holland [137] observed the presence of three Hox clusters in the lamprey (designated HOX 3 in fig. 3). One cannot exclude, however, that the lamprey is an intermediate in the two-step genome duplication hypothesis, with one cluster duplicating to give a three-Hox cluster state, which is unique to this organism, or, alternatively, that the lamprey descended from a four-Hox-cluster-containing ancestor that subsequently suffered cluster loss. In support of this latter theory, Ruddle and colleagues [138] have speculated that the hagfish, which is believed to have a more distant evolutionary divergence date than the lamprey, possesses four Hox clusters (fig. 3). Although the mechanism is still a matter of debate, these data suggest that the increase in genome size, resulting in the Hox cluster number increase from one to four, and the emergence of three to four NR paralogues occurred in the relatively short period of time after the divergence of the Urochordata and before the specification of the Hyperotreti.

In support of NR paralogues arising through genome or block duplications, mapping studies have demonstrated the presence of extensive 'paralogy groups' which include NRs on different chromosomes. A paralogy group on human chromosomes 6, 9, 1, and 19 include the RXR, collagen, Notch and heat shock (HSP) genes [135]. Within these clusters, RXRα, β and γ are located on chromosomes 9, 6 and 1, respectively [135]. A paralogy group on human chromosomes 7, 17, 12, and 2 includes the Hox gene clusters, Evx home-

obox genes, glucose transport genes, wnt genes and the RAR and TR genes [139]. There are three paralogues of the RAR (α, β, γ) and two paralogues of the TR (α and β) [140, 141]. RARα and TRα are located together on human chromosome 17 along with the Hox B cluster. RARγ is located on chromosome 12 along with the Hox C cluster. This suggests that an invertebrate Hox cluster containing an ancient RAR gave rise to the vertebrate Hox clusters B and C by either genome or chromosome duplication. Possibly, a TR paralogue on chromosome 12 has either still to be identified or has been lost. RARβ and TRβ are also located together on human chromosome 3, which does not contain a Hox cluster grouping. This may suggest duplication of the RARα- and TRα-containing region of chromosome 17 followed by translocation to chromosome 3.

It is noteworthy that RARs are important regulators of Hox gene expression during vertebrate development. Control of the anterior-posterior axis linked to the development of a central nervous system is believed to be a chordate innovation [130, 142, 143]. Growing complexities of developmental systems, which occurred around the time of the Cambrian explosion, required co-evolution of more sophisticated regulatory gene networks, one of which was the regulation of the increasing number of Hox genes by the increasing number of RARs. The idea that the chordates incorporated the RAR into development is consistent with an absence of RARs before Urochordata (fig. 3) and a lack of effect of retinoic acid on Hox gene expression in Echinodermata and Arthropoda, but not Cephalochordata or the Urochordata [144, 145].

The first metazoan: setting the table for the arrival of the NR superfamily

The NR ancestor is likely to have appeared among the first metazoans around 1000–800 mya. The role of this NR ancestor was likely to activate and/or repress transcription of specific genes by binding to their DNA regulatory regions as a monomer and in a ligand-independent manner. For the early NR(s) to regulate transcription, it had to be compatible with the existing transcriptional apparatus. Observations that mammalian NRs are transcriptionally active in yeast [146–149] indicated that the factors required for NRs to act on chromatin structure and/or communicate with the basal transcriptional machinery must be highly conserved between yeast and mammalian cells, and must have existed long before the appearance of the first NRs on the evolutionary scene. Indeed, homologues of a large number of NR-associated transcriptional co-regulators have been described in yeast [150–152].

A hypothetical evolutionary path that might have been taken by the first NR in the early metazoan and how it might have diverged into the six sub-families known today is represented in figure 5. The first NR arose as one unit consisting of the currently recognised DBD and LBD. This conclusion is based on the finding that the two domains exist together in lower metazoans and does not entirely dismiss a possibility that the two regions existed independently earlier in evolution, despite the fact that strong similarity to either domain has not yet been observed outside the metazoan kingdom. The ancestor DBD may have been similar to that of modern receptors with the LBD being present but not possessing a transactivation domain or the ability to dimerise. In constructing this evolutionary tree, several assumptions have been made concerning ancestral events. First, the earliest metazoans possess COUP, RXR and FTZ-F1, all of which are well conserved to the present day and thus are close to the NR ancestor, making it an ancestor of sub-family II and V. Second, the ability to heterodimerise with RXR arose once and has diversified through sub-families I, II and IV [50]. Third, as indicated by Laudet [50], there is no correlation between the type of ligand bound by a receptor and the position in the family tree. A reference is made to the presence of a ligand-binding RXR in Cnidaria, which may suggest that the archaic receptor assembled with the ability to bind ligand and many subsequent receptors have lost this feature. The presence of RXR in the earliest branches of the metazoans enables the extrapolation that an archaic family II member gave rise to the families I and IV which can also heterodimerise with the RXR (fig. 5). COUP-TF and RXR may have arisen from the same ancestor with the RXR heterodimerisation function being highly conserved. FTZ-F1 appears to be currently the earliest NR that does not possess the ability to heterodimerise with the RXR. The sub-family III receptors also fail to heterodimerise with the RXR, with ERR-1 and SF-1/FTZF1 acting independently on a common DNA target sequence [153, 154]. In this hypothesis, an FTZ-F1 ancestor originally descended from a common ancestor with RXR and COUP-TF, through changes in the DBD, with the LBD being conserved. This receptor was the ancestor of the ERR which, with divergence in the LBD, gave rise to the steroid hormone receptors of sub-family III (fig. 5). A major reasoning in reaching this hypothesis is the placing of the FTZ-F1 in the same region of the phylogenetic tree as the RXR and COUP-TF on comparison of the C-terminal domain, but with the ER family on comparison of DBDs [46]. It is also a possibility that the sub-family III and V ancestor was able to dimerise with the RXR and then subsequently lost this ability.

It could be assumed from recent data in Cnidaria that a direct RXR ancestor developed the ability to bind the

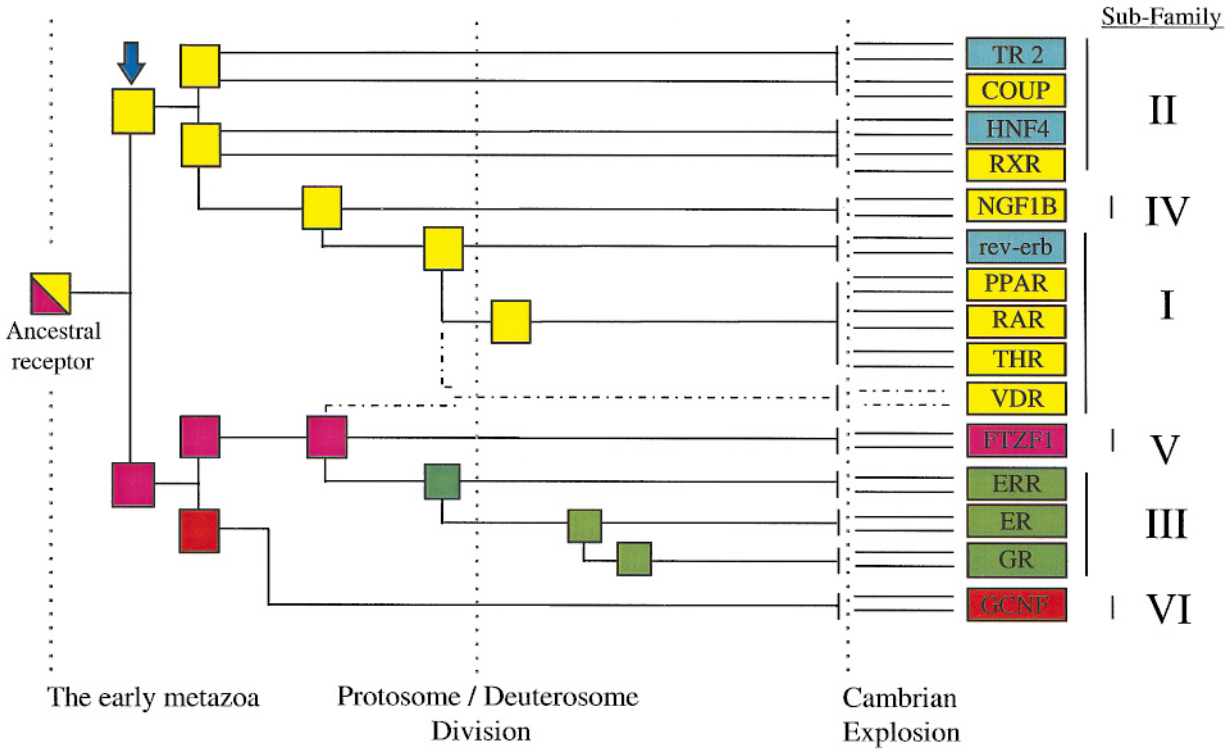


Figure 5. A hypothetical model for the evolution of the NR superfamily. The NR ancestor in the early metazoan was a common ancestor of sub-family II (yellow box) and V (pink box). Heterodimerisation was acquired early in the diversification of the family (blue arrow) and was either conserved (yellow box) or lost (blue box). The sub-family V ancestor gave rise to sub-family VI (red box) and to a receptor, which through further diversification of the ligand-binding domain gave rise to sub-family III (green box). Dashed lines represent the possible origin of a chimeric VDR. The vertical green dashed lines represent the time period of the first metazoan, the protosome-deuterosome split and the Cambrian explosion which saw the HOX gene cluster quadruple and the appearance of numerous NR paralogues that are present today in vertebrates.

ligand 9-*cis*-retinoic acid with a subset diverging into the RAR and obtaining the ability to bind all-*trans*-retinoic acid. Another possibility is that after acquiring the ability to heterodimerise, a common ancestor which had already diverged from RXR gave rise to both the NGF1B and RAR grouping (fig. 5). This approach is in line with the observation by Laudet [50] that sub-family I and IV are related, and leads to the assumption that the NGF1B grouping has diverged considerably in the ligand-binding region due to the initial loss of the ability to bind retinoids.

The possibility that the VDR is a chimera is in accordance with its clustering into different families based on the phylogenetic trees compiled from either DBD or LBD sequence [46]. The DBD of the VDR closely resembles that of FTZ-F1, while the LBD is closer to the RAR grouping. However, the more recently characterised FXR, CAR and LXR, along with several arthropod EcRs, cluster together with the VDR in phylogenetic trees based on both the DBD and LBD, thus throwing doubt on the VDR being a chimera. Convergent evolution may be in part responsible for this phenomenon which is observed in other receptors such as GCNF and NGF1B. The possibility that each of these receptors diverged from a common chimeric ancestor still remains and a mechanism for such an event, originally suggested by Laudet and colleagues [46], is given in figure 5. The sub-family VI member, GCNF, is the opposite to the VDR, with a DBD similar to sub-family I and a LBD closer to sub-family V [1, 103]. This receptor exhibits unique dimerisation and DNA-binding properties and acts as a repressor in the absence of ligand [155]. As mentioned earlier, the arthropod GRF and THR4 receptors cluster with *Xenopus* and mammalian GCNF (fig. 4) and are thus placed in the same grouping of the unique nomenclature system [43]. These receptors also show strong similarity with the FTZ-F1, possibly shedding light on the nature of a common ancestor (see fig. 5). The steroid-hormone-binding members of sub-family III, along with the RAR, PPAR, and TR, are not observed until the chordates, suggesting that these receptors were not brought under tight evolutionary conservation until a later date.

It is clear that without knowing the number and nature of NRs present today and in the early metazoans, the above discussion remains highly speculative. It is likely that a better understanding of evolution of the NR as well as other genes will emerge with the completion of the human genome project, scheduled to finish in the next decade, when the exact nature and number of NRs will become known. When the knowledge of all human genes is coupled with full genome sequences of other organisms, including lower metazoans, one will perhaps then be able to put all the pieces of the NR evolution and diversification puzzle together and in the correct order.

Acknowledgments. We thank Vincent Laudet for granting permission to reproduce figure 2 in this paper. We are grateful to Nicola Hawe and Leanne Wiedemann for comments on the manuscript and the Leukaemia Research Fund of Great Britain for support.

1 Mangelsdorf D. J., Thummel C., Beato M., Herrlich P., Schutz G., Umesono K. et al. (1995) The nuclear receptor superfamily: the second decade. Cell **83:** 835–839
2 Chambon P. (1996) A decade of molecular biology of retinoic acid receptors. FASEB J. **10:** 940–954
3 Yu V. C., Delsert C., Andersen B., Holloway J. M., Devary O. V., Naar A. M. et al. (1991) RXRβ: a coregulator that enhances binding of retinoic acid, thyroid hormone, and vitamin D receptors to their cognate response elements. Cell **67:** 1251–1266
4 Umesono K., Murakami K. K., Thompson C. C. and Evans R. M. (1991) Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D_3 receptors. Cell **65:** 1255–1266
5 Green S. (1993) Promiscuous liaisons. Nature **361:** 590–591
6 Smith W. C., Nakshatri H., Leroy P., Rees J. and Chambon P. (1991) A retinoic acid response element is present in the mouse cellular retinol binding protein I (mCRBPI) promoter. EMBO J. **10:** 2223–2230
7 Mader S., Chen J.-Y., Chen Z., White J., Chambon P. and Gronemeyer H. (1993) The patterns of binding of RAR, RXR and TR homo- and heterodimers to direct repeats are dictated by the binding specificities of the DNA binding domains. EMBO J. **12:** 5029–5041
8 Mangelsdorf D. J., Umesono K., Kliewer S. A., Borgmeyer U., Ong E. S. and Evans R. M. (1991) A direct repeat in the cellular retinol-binding protein type II gene confers differential regulation by RXR and RAR. Cell **66:** 555–561
9 Saatcioglu F., Deng T. and Karin M. (1993) A novel cis element mediating ligand-independent activation by c-ErbA: implications for hormonal regulation. Cell **75:** 1095–1105
10 Carlberg C., Bendik I., Wyss A., Meier E., Sturzenbecker L. J., Grippo J. F. et al. (1993) Two nuclear signalling pathways for vitamin D. Nature **361:** 657–660
11 Forman B. M., Tzameli I., Choi H. S., Chen J., Simha D., Seol W. et al. (1998) Androstane metabolites bind to and deactivate the nuclear receptor CAR-beta. Nature **395:** 612–615
12 Horwitz K. B., Jackson T. A., Rain D. L., Richer J. K., Takimoto G. S. and Tung L. (1996) Nuclear receptor coactivators and corepressors. Mol. Endocrinol. **10:** 1167–1177
13 Darimont B. D., Wagner R. L., Apriletti J. W., Stallcup M. R., Kushner P. J., Baxter J. D. et al. (1998) Structure and specificity of nuclear receptor-coactivator interactions. Genes Dev. **12:** 3343–3356
14 Baur E. V., Zechel C., Heery D., Heine M. J. S., Garnier J. M., Vivat V. et al. (1996) Differential ligand-dependent interactions between the Af-2 activating domain of nuclear receptors and the putative transcriptional intermediary factors Msug1 and Tif1. EMBO J. **15:** 110–124
15 Kamei Y., Xu L., Heinzel T., Torchia J., Kurokawa R., Gloss B. et al. (1996) A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. Cell **85:** 403–414
16 Le Douarin B., Zechel C., Garnier J. M., Lutz Y., Tora L., Pierrat B. et al. (1995) The N-terminal part of Tif1, a putative mediator of the ligand-dependent activation function (AF-2) of nuclear receptors, is fused to B-Raf in the oncogenic protein T18. EMBO J. **14:** 2020–2033
17 Halachmi S., Marden E., Martin G., MacKay H., Abbondanza C. and Brown M. (1994) Estrogen receptor-associated proteins: possible mediators of hormone induced transcription. Science **264:** 1455–1458
18 Voegel J. J., Heine M. J. S., Zechel C., Chambon P. and Gronemeyer H. (1996) TIF2, a 160 kDa transcriptional mediator for the ligand-dependent activation function AF-2 of nuclear receptors. EMBO J **15:** 3667–3675
19 Onate S. A., Tsai S. Y., Tsai M. J. and Omalley B. W. (1995) Sequence and characterization of a coactivator for the steroid-hormone receptor superfamily. Science **270:** 1354–1357
20 Spencer T. E., Jenster G., Burcin M. M., Allis C. D., Zhou J., Mizzen C. A. et al. (1997) Steroid receptor coactivator-1 is a histone acetyltransferase. Nature **389:** 194–198
21 Montminy M. (1997) Something new to hang your HAT on. Nature **387:** 654–655
22 Lanz R. B., McKenna N. J., Onate S. A., Albrecht U., Wong J., Tsai S. Y. et al. (1999) A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. Cell **97:** 17–27
23 Horlein A. J., Naar A. M., Heinzel T., Torchia J., Gloss B., Kurokawa R. et al. (1995) Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. Nature **377:** 397–404
24 Chen J. D. and Evans R. M. (1995) A transcriptional co-repressor that interacts with nuclear hormone receptors. Nature **377:** 454–457
25 Alland L., Muhle R., Hou H. Jr, Potes J., Chin L., Schreiber-Agus N. et al. (1997) Role for N-CoR and histone deacetylase in Sin3-mediated transcriptional repression. Nature **387:** 49–55
26 Heinzel T., Lavinsky R. M., Mullen T. M., Soderstrom M., Laherty C. D., Torchia J. et al. (1997) A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. Nature **387:** 43–48
27 Nagy L., Kao H. Y., Chakravarti D., Lin R. J., Hassig C. A., Ayer D. E. et al. (1997) Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase. Cell **89:** 373–380
28 Wang H. and Stillman D. J. (1990) In vitro regulation of a SIN3-dependent DNA-binding activity by stimulatory and inhibitory factors. Proc. Natl. Acad. Sci. USA **87:** 9761–9765
29 Wang H., Clark I., Nicholson P. R., Herskowitz I. and Stillman D. J. (1990) The *Saccharomyces cerevisiae* SIN3 gene, a negative regulator of HO, contains four paired amphipathic helix motifs. Mol. Cell. Biol. **10:** 5927–5936
30 Wang H. and Stillman D. J. (1993) Transcriptional repression in *Saccharomyces cerevisiae* by a SIN3-LexA fusion protein. Mol. Cell. Biol. **13:** 1805–1814
31 Taunton J., Hassig C. A. and Schreiber S. L. (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. Science **272:** 408–411
32 Thummel C. S. (1995) From embryogenesis to metamorphosis: the regulation and function of *Drosophila* nuclear receptor superfamily members. Cell **83:** 871–877
33 Truman J. W. and Riddiford L. M. (1999) The origins of insect metamorphosis. Nature **401:** 447–452

34 Thummel C. S. (1997) Dueling orphans – interacting nuclear receptors coordinate *Drosophila* metamorphosis. Bioessays **19:** 669–672

35 Crispi S., Giordano E., D'Avino P. P. and Furia M. (1998) Cross-talking among *Drosophila* nuclear receptors at the promiscuous response element of the ng-1 and ng-2 intermolt genes. J. Mol. Biol. **275:** 561–574

36 Carney G. E., Wade A. A., Sapra R., Goldstein E. S. and Bender M. (1997) DHR3, an ecdysone-inducible early-late gene encoding a *Drosophila* nuclear receptor, is required for embryogenesis. Proc. Natl. Acad. Sci. USA **94:** 12024–12029

37 Giguere V. (1999) Orphan nuclear receptors: from gene to function. Endocrinol. Rev. **20:** 689–725

38 Jones G. and Sharp P. A. (1997) Ultraspiracle: an invertebrate nuclear receptor for juvenile hormones. Proc. Natl. Acad. Sci. USA **94:** 13499–13503

39 Baker M. E. (1997) Steroid receptor phylogeny and vertebrate origins. Mol. Cell. Endocrinol. **135:** 101–107

40 Gahr M., Guttinger H. R. and Kroodsma D. E. (1993) Estrogen receptors in the avian brain: survey reveals general distribution and forebrain areas unique to songbirds. J. Comp. Neurol. **327:** 112–122

41 Finn C. A. (1996) Why do women menstruate? Historical and evolutionary review. Eur. J. Obstet. Gynecol. Reprod. Biol. **70:** 3–8

42 Finn C. A. (1998) Menstruation: a nonadaptive consequence of uterine evolution. Q. Rev. Biol. **73:** 163–173

43 Comitee N. R. (1999) A unified nomenclature system for the nuclear receptor superfamily. Cell **97:** 161–163

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

45 Truss M. and Beato M. (1993) Steroid hormone receptors: interaction with deoxyribonucleic acid and transcription factors. Endocrinol. Rev. **14:** 459–479

46 Laudet V., Hanni C., Coll J., Catzeflis F. and Stehelin D. (1992) Evolution of the nuclear receptor gene superfamily. EMBO J. **11:** 1003–1013

47 Ponglikitmongkol M., Green S. and Chambon P. (1988) Genomic organization of the human oestrogen receptor gene. EMBO J. **7:** 3385–3388

48 Zelent A. (1995) Molecular mechanisms of retinoid action. In: Retinoids in Oncology, pp. 3–25, Degos L. and Parkinson D. R. (eds), Springer, Heidelberg

49 Zilliacus J., Carlstedt-Duke J., Gustafsson J. A. and Wright A. P. (1994) Evolution of distinct DNA-binding specificities within the nuclear receptor family of transcription factors. Proc. Natl. Acad. Sci. USA **91:** 4175–4179

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

51 Green S., Kumar V., Theulaz I., Wahli W. and Chambon P. (1988) The N-terminal DNA-binding 'zinc finger' of the oestrogen and glucocorticoid receptors determines target gene specificity. EMBO J **7:** 3037–3044

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

53 Sluder A. E., Mathews S. W., Hough D., Yin V. P. and Maina C. V. (1999) The nuclear receptor superfamily has undergone extensive proliferation and diversification in nematodes. Genome Res. **9:** 103–120

54 Le Jossic C. and Michel D. (1998) Striking evolutionary conservation of a cis-element related to nuclear receptor target sites and present in TR2 orphan receptor genes. Biochem. Biophys. Res. Commun. **245:** 64–69

55 Wang S. F., Miura K., Miksicek R. J., Segraves W. A. and Raikhel A. S. (1998) DNA binding and transactivation characteristics of the mosquito ecdysone receptor-Ultraspiracle complex. J. Biol. Chem. **273:** 27531–27540

56 Kersten S., Reczek P. R. and Noy N. (1997) The tetramerization region of the retinoid X receptor is important for transcriptional activation by the receptor. J. Biol. Chem. **272:** 29759–29768

57 Perlmann T., Rangarajan P. N., Umesono K. and Evans R. M. (1993) Determinants for selective RAR and TR recognition of direct repeat HREs. Genes Dev. **7:** 1411–1422

58 Wilson T. E., Paulsen R. E., Padgett K. A. and Milbrandt J. (1992) Participation of non-zinc finger residues in DNA binding by two nuclear orphan receptors. Science **256:** 107–110

59 Towers T. L., Luisi B. F., Asianov A. and Freedman L. P. (1993) DNA target selectivity by the vitamin D3 receptor: mechanism of dimer binding to an asymmetric repeat element. Proc. Natl. Acad. Sci. USA **90:** 6310–6314

60 Kurokawa R., Yu V. C., Naar A., Kyakumoto S., Han Z., Silverman S. et al. (1993) Differential orientations of the DNA-binding domain and carboxy-terminal dimerization interface regulate binding site selection by nuclear receptor heterodimers. Genes Dev. **7:** 1423–1435

61 Abate C. and Curran T. (1990) Encounters with Fos and Jun on the road to AP-1. Semin. Cancer Biol. **1:** 19–26

62 Hai T. and Curran T. (1991) Cross-family dimerization of transcription factors Fos/Jun and ATF/CREB alters DNA binding specificity. Proc. Natl. Acad. Sci. USA **88:** 3720–3724

63 Lassar A. B., Davis R. L., Wright W. E., Kadesch T., Murre C., Voronova A. et al. (1991) Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo. Cell **66:** 305–315

64 Benezra R., Davis R. L., Lockshon D., Turner D. L. and Weintraub H. (1990) The protein id: a negative regulator of helix-loop-helix DNA binding proteins. Cell **61:** 49–59

65 Laudet V. and Adelmant G. (1995) Nuclear receptors: lonesome orphans. Curr. Biol. **5:** 124–127

66 Chen F., Cooney A. J., Wang Y., Law S. W. and O'Malley B. W. (1994) Cloning of a novel orphan receptor (GCNF) expressed during germ cell development. Mol. Endocrinol. **8:** 1434–1444

67 Cooney A. J., Hummelke G. C., Herman T., Chen F. and Jackson K. J. (1998) Germ cell nuclear factor is a response element-specific repressor of transcription. Biochem. Biophys. Res. Commun. **245:** 94–100

68 Hirst M. A., Hinck L., Danielsen M. and Ringold G. M. (1992) Discrimination of DNA response elements for thyroid hormone and estrogen is dependent on dimerization of receptor DNA binding domains. Proc. Natl. Acad. Sci. USA **89:** 5527–5531

69 Medici N., Nigro V., Abbondanza C., Moncharmont B., Molinari A. M. and Puca G. A. (1991) In vitro binding of the purified hormone-binding subunit of the estrogen receptor to oligonucleotides containing natural or modified sequences of an estrogen-responsive element. Mol. Endocrinol. **5:** 555–563

70 Wilson T. E., Fahrner T. J. and Milbrandt J. (1993) The orphan receptors NGFI-B and steroidogenic factor 1 establish monomer binding as a third paradigm of nuclear receptor-DNA interaction. Mol. Cell. Biol. **13:** 5794–5804

71 Kostrouch Z., Kostrouchova M., Love W., Jannini E., Piatigorsky J. and Rall J. E. (1998) Retinoic acid X receptor in the diploblast, *Tripedalia cystophora*. Proc. Natl. Acad. Sci. USA **95:** 13442–13447

72 Danielian P. S., White R., Lees J. A. and Parker M. G. (1992) Identification of a conserved region required for hormone dependent transcriptional activation by steroid hormone receptors. EMBO J. **11:** 1025–1033

73 Durand B., Saunders M., Gaudon C., Roy B., Losson R. and Chambon P. (1994) Activation function 2 (AF-2) of retinoic acid receptor and 9-cis retinoic acid receptor: presence of a conserved autonomous constitutive activating domain and influence of the nature of the response element on AF-2 activity. EMBO J **13:** 5370–5382

74 Webster N. J., Green S., Tasset D., Ponglikitmongkol M. and Chambon P. (1989) The transcriptional activation function located in the hormone-binding domain of the human oestrogen receptor is not encoded in a single exon. EMBO J. **8:** 1441–1446

75 Tora L., White J., Brou C., Tasset D., Webster N., Scheer E. et al. (1989) The human estrogen receptor has two independent non-acidic transcriptional activation functions. Cell **59:** 477–487

76 Tasset D., Tora L., Fromental C., Scheer E. and Chambon P. (1990) Distinct classes of transcriptional activating domains function by different mechanisms. Cell **62:** 1177–1187

77 Lemon B. D. and Freedman L. P. (1999) Nuclear receptor cofactors as chromatin remodelers. Curr. Opin. Genet. Dev. **9:** 499–504

78 McInerney E. M., Rose D. W., Flynn S. E., Westin S., Mullen T. M., Krones A. et al. (1998) Determinants of coactivator LXXLL motif specificity in nuclear receptor transcriptional activation. Genes Dev. **12:** 3357–3368

79 Bourguet W., Ruff M., Chambon P., Gronemeyer H. and Moras D. (1995) Crystal structure of the ligand-binding domain of the human nuclear receptor RXR-alpha. Nature **375:** 377–382

80 Thornton J. W. and Kelley D. B. (1998) Evolution of the androgen receptor: structure-function implications. Bioessays **20:** 860–869

81 Veldscholte J., Ris-Stalpers C., Kuiper G. G., Jenster G., Berrevoets C., Claassen E. et al. (1990) A mutation in the ligand binding domain of the androgen receptor of human LNCaP cells affects steroid binding characteristics and response to anti-androgens. Biochem. Biophys. Res. Commun. **173:** 534–540

82 Guo X., Xu Q., Harmon M. A., Jin X., Laudet V., Mangelsdorf D. J. et al. (1998) Isolation of two functional retinoid X receptor subtypes from the ixodid tick, *Amblyomma americanum* (L.). Mol. Cell. Endocrinol. **139:** 45–60

83 Garcia-Vallvé S. and Palau J. (1998) Nuclear receptors, nuclear-receptor factors, and nuclear-receptor-like orphans form a large paralog cluster in Homo sapiens. Mol. Biol. Evol. **15:** 665–682

84 Baker M. E. (1995) Endocrine activity of plant-derived compounds: an evolutionary perspective. Proc. Soc. Exp. Biol. Med. **208:** 131–138

85 Yamamoto K. R. (1997) Intracellular receptors: new instruments for a symphony of signals. In: Molecular Biology of Steroid and Nuclear Hormone Receptors, pp. 7–10, Freedman L. P. (ed.), Birkhauser, Boston

86 Yang C. and Chen S. (1999) Two organochlorine pesticides, toxaphene and chlordane, are antagonists for estrogen-related receptor alpha-1 orphan receptor. Cancer Res. **59:** 4519–4524

87 Leid M., Kastner P. and Chambon P. (1992) Multiplicity generates diversity in the retinoic acid signalling pathways. Trends Biochem. Sci. **17:** 427–433

88 Kato S., Endoh H., Masuhiro Y., Kitamoto T., Uchiyama S., Sasaki H. et al. (1995) Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase. Science **270:** 1491–1494

89 Taneja R., Rochette-Egly C., Plassat J., Penna L., Gaub M. and Chambon P. (1997) Phosphorylation of activation functions AF-1 and AF-2 of RAR alpha and RAR gamma is indispensable for differentiation of F9 cells upon retinoic acid and cAMP treatment. EMBO J. **16:** 6425–6465

90 Rochette-Egly C., Adam S., Rossignol M., Egly J. M. and Chambon P. (1997) Stimulation of RAR alpha activation function AF-1 through binding to the general transcription factor TFIIH and phosphorylation by CDK7. Cell **90:** 97–107

91 Tzagarakis-Foster C. and Privalsky M. L. (1998) Phosphorylation of thyroid hormone receptors by protein kinase A regulates DNA recognition by specific inhibition of receptor monomer binding. J. Biol. Chem. **273:** 10926–10932

92 Shao D., Rangwala S. M., Bailey S. T., Krakow S. L., Reginato M. J. and Lazar M. A. (1998) Interdomain communication regulating ligand binding by PPAR-gamma. Nature **396:** 377–380

93 Katzenellenbogen B. S. (1996) Estrogen receptors: bioactivities and interactions with cell signaling pathways. Biol. Reprod. **54:** 287–293

94 Castillo G., Brun R. P., Rosenfield J. K., Hauser S., Park C. W., Troy A. E. et al. (1999) An adipogenic cofactor bound by the differentiation domain of PPARgamma. EMBO J. **18:** 3676–3687

95 Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C. et al. (1998) Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. Nature **391:** 485–488

96 Heinzel T., Lavinsky R. M., Mullen T. M., Soderstrom M., Laherty C. D., Torchia J. et al. (1997) A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. Nature **387:** 43–48

97 Montano M. M., Muller V., Trobaugh A. and Katzenellenbogen B. S. (1995) The carboxy-terminal F domain of the human estrogen receptor: role in the transcriptional activity of the receptor and the effectiveness of antiestrogens as estrogen antagonists. Mol. Endocrinol. **9:** 814–825

98 Sladek F. M., Ruse M. D. Jr, Nepomuceno L., Huang S. M. and Stallcup M. R. (1999) Modulation of transcriptional activation and coactivator interaction by a splicing variation in the F domain of nuclear receptor hepatocyte nuclear factor 4alpha1. Mol. Cell. Biol. **19:** 6509–6522

99 Amero S. A., Kretsinger R. H., Moncrief N. D., Yamamoto K. R. and Pearson W. R. (1992) The origin of nuclear receptor proteins: a single precursor distinct from other transcription factors. Mol. Endocrinol. **6:** 3–7

100 Escriva H., Safi R., Hanni C., Langlois M. C., Saumitou-Laprade P., Stehelin D. et al. (1997) Ligand binding was acquired during evolution of nuclear receptors. Proc. Natl. Acad. Sci. USA **94:** 6803–6808

101 Detera-Wadleigh S. D. and Fanning T. G. (1994) Phylogeny of the steroid receptor superfamily. Mol. Phylogenet. Evol. **3:** 192–205

102 Mouillet J. F., Bousquet F., Sedano N., Alabouvette J., Nicola M., Zelus D. et al. (1999) Cloning and characterization of new orphan nuclear receptors and their developmental profiles during *Tenebrio* metamorphosis. Eur. J. Biochem. **265:** 972–981

103 Greschik H., Wurtz J. M., Hublitz P., Kohler F., Moras D. and Schule R. (1999) Characterization of the DNA-binding and dimerization properties of the nuclear orphan receptor germ cell nuclear factor. Mol. Cell. Biol. **19:** 690–703

104 Nauber U., Pankratz M. J., Kienlin A., Seifert E., Klemm U. and Jackle H. (1988) Abdominal segmentation of the *Drosophila* embryo requires a hormone receptor-like protein encoded by the gap gene knirps. Nature **336:** 489–492

105 Sengupta P., Colbert H. A. and Bargmann C. I. (1994) The *C. elegans* gene odr-7 encodes an olfactory-specific member of the nuclear receptor superfamily. Cell **79:** 971–980

106 Seol W., Choi H. S. and Moore D. D. (1995) Isolation of proteins that interact specifically with the retinoid X receptor: two novel orphan receptors. Mol. Endocrinol. **9:** 72–85

107 Burris T. P., Guo W. and McCabe E. R. (1996) The gene responsible for adrenal hypoplasia congenita, DAX-1, encodes a nuclear hormone receptor that defines a new class within the superfamily. Recent Prog. Horm. Res. **51:** 241–259

108 Mangelsdorf D. J. and Evans R. M. (1995) The RXR heterodimers and orphan receptors. Cell **83:** 841–850

109 Yu R. N., Ito M. and Jameson J. L. (1998) The murine Dax-1 promoter is stimulated by SF-1 (steroidogenic factor-1) and inhibited by COUP-TF (chicken ovalbumin upstream promoter-transcription factor) via a composite nuclear receptor-regulatory element. Mol. Endocrinol. **12:** 1010–1022

110 Kumar S. and Hedges S. B. (1998) A molecular timescale for vertebrate evolution. Nature **392:** 917–920

111 Ueda H., Sun G. C., Murata T. and Hirose S. (1992) A novel DNA-binding motif abuts the zinc finger domain of insect nuclear hormone receptor FTZ-F1 and mouse embryonal long terminal repeat-binding protein. Mol. Cell. Biol. **12:** 5667–5672

112 Ito M., Yu R. N. and Jameson J. L. (1998) Steroidogenic factor-1 contains a carboxy-terminal transcriptional activa-

tion domain that interacts with steroid receptor coactivator-1. Mol. Endocrinol. **12:** 290–301
113 Qiu Y., Krishnan V., Zeng Z., Gilbert D. J., Copeland N. G., Gibson L. et al. (1995) Isolation, characterization, and chromosomal localization of mouse and human COUP-TF I and II genes. Genomics **29:** 240–246
114 Kastner P., Mark M. and Chambon P. (1995) Nonsteroid nuclear receptors: what are genetic studies telling us about their role in real life? Cell **83:** 859–869
115 Sucov H. M., Dyson E., Gumeringer C. L., Price J., Chien K. R. and Evans R. M. (1994) RXR alpha mutant mice establish a genetic basis for vitamin A signaling in heart morphogenesis. Genes Dev **8:** 1007–1018
116 Morris S. C. (1998) Metazoan phylogenies: falling into place or falling to pieces? A palaeontological perspective. Curr. Opin. Genet. Dev. **8:** 662–667
117 Rosa R. de, Grenier J. K., Andreeva T., Cook C. E., Adoutte A., Akam M. et al. (1999) Hox genes in brachiopods and priapulids and protostome evolution. Nature **399:** 772–776
118 Freebern W. J., Osman A., Niles E. G., Christen L. and LoVerde P. T. (1999) Identification of a cDNA encoding a retinoid X receptor homologue from *Schistosoma mansoni*: evidence for a role in female-specific gene expression. J. Biol. Chem. **274:** 4577–4585
119 Freebern W. J., Niles E. G. and LoVerde P. T. (1999) RXR-2, a member of the retinoid X receptor family in *Schistosoma mansoni*. Gene **233:** 33–38
120 Jefferies R. P. (1991) Two types of bilateral symmetry in the Metazoa: chordate and bilaterian. Ciba Found. Symp. **162:** 94–120
121 Carosa E., Fanelli A., Ulisse S., Di Lauro R., Rall J. E. and Jannini E. A. (1998) *Ciona intestinalis* nuclear receptor 1: a member of steroid/thyroid hormone receptor family. Proc. Natl. Acad. Sci. USA **95:** 11152–11157
122 Janvier P. (1996) Early Vertebrates, Oxford University Press, Oxford
123 Braun C. B. (1996) The sensory biology of the living jawless fishes: a phylogenetic assessment. Brain Behav. Evol. **48:** 262–276
124 Ayala F. J. and Rzhetsky A. (1998) Origin of the metazoan phyla: molecular clocks confirm paleontological estimates. Proc. Natl. Acad. Sci. USA **95:** 606–611
125 Holland P. W. and Garcia-Fernandez J. (1996) Hox genes and chordate evolution. Dev. Biol. **173:** 382–395
126 Skrabanek L. and Wolfe K. H. (1998) Eukaryote genome duplication – where's the evidence? Curr. Opin. Genet. Dev. **8:** 694–700
127 Boncinelli E., Simeone A., Acampora D. and Mavilio F. (1991) HOX gene activation by retinoic acid. Trends Genet. **7:** 329–334
128 Krumlauf R. (1994) *Hox* genes in vertebrate development. Cell **78:** 191–201
129 Marshall H., Morrison A., Studer M., Popperl H. and Krumlauf R. (1996) Retinoids and Hox genes. FASEB J **10:** 969–978
130 Shimeld S. M. (1996) Retinoic acid, HOX genes and the anterior-posterior axis in chordates. BioEssays **18:** 613–616
131 Ducouret B., Tujague M., Ashraf J., Mouchel N., Servel N., Valotaire Y. et al. (1995) Cloning of a teleost fish glucocorticoid receptor shows that it contains a deoxyribonucleic acid-binding domain different from that of mammals. Endocrinology **136:** 3774–3783
132 Ikeuchi T., Todo T., Kobayashi T. and Nagahama Y. (1999) cDNA cloning of a novel androgen receptor subtype. J. Biol. Chem. **274:** 25205–25209
133 Amores A. and Postlethwait J. H. (1999) Banded chromosomes and the zebrafish karyotype. Methods Cell Biol. **60:** 323–338
134 Bailey W. J., Kim J., Wagner G. P. and Ruddle F. H. (1997) Phylogenetic reconstruction of vertebrate Hox cluster duplications. Mol. Biol. Evol. **14:** 843–853
135 Smith N. G., Knight R. and Hurst L. D. (1999) Vertebrate genome evolution: a slow shuffle or a big bang? Bioessays **21:** 697–703
136 Pendleton J. W., Nagai B. K., Murtha M. T. and Ruddle F. H. (1993) Expansion of the Hox gene family and the evolution of chordates. Proc. Natl. Acad. Sci. USA **90:** 6300–6304
137 Sharman A. C. and Holland P. W. (1998) Estimation of Hox gene cluster number in lampreys. Int. J. Dev. Biol. **42:** 617–620
138 Ruddle F. H., Amemiya C. T., Carr J. L., Kim C. B., Ledje C., Shashikant C. S. et al. (1999) Evolution of chordate hox gene clusters. Ann. NY Acad. Sci. **870:** 238–248
139 Bentley K. L., Bradshaw M. S. and Ruddle F. H. (1993) Physical linkage of the murine Hox-b cluster and nerve growth factor receptor on yeast artificial chromosomes. Genomics **18:** 43–53
140 Krust A., Kastner P., Petkovich M., Zelent A. and Chambon P. (1989) A third human retinoic acid receptor, hRAR-gamma. Proc. Natl. Acad. Sci. USA **86:** 5310–5314
141 Thompson C. C., Weinberger C., Lebo R. and Evans R. M. (1987) Identification of a novel thyroid hormone receptor expressed in the mammalian central nervous system. Science **237:** 1610–1614
142 Essner J. J., Johnson R. G. and Hackett P. B. Jr (1999) Overexpression of thyroid hormone receptor alpha 1 during zebrafish embryogenesis disrupts hindbrain patterning and implicates retinoic acid receptors in the control of hox gene expression. Differentiation **65:** 1–11
143 Maden M. (1999) Heads or tails? Retinoic acid will decide. Bioessays **21:** 809–812
144 Holland L. Z. and Holland N. D. (1996) Expression of AmphiHox-1 and AmphiPax-1 in amphioxus embryos treated with retinoic acid: insights into evolution and patterning of the chordate nerve cord and pharynx. Development **122:** 1829–1838
145 Katsuyama Y., Wada S., Yasugi S. and Saiga H. (1995) Expression of the labial group Hox gene HrHox-1 and its alteration induced by retinoic acid in development of the ascidian *Halocynthia roretzi*. Development **121:** 3197–3205
146 Heery D. M., Zacharewski T., Pierrat B., Gronemeyer H., Chambon P. and Losson R. (1993) Efficient transactivation by retinoic acid receptors in yeast requires retinoid X receptors. Proc. Natl. Acad. Sci. USA **90:** 4281–4285
147 Henriksson A., Almlof T., Ford J., McEwan I. J., Gustafsson J. A. and Wright A. P. (1997) Role of the Ada adaptor complex in gene activation by the glucocorticoid receptor. Mol. Cell. Biol. **17:** 3065–3073
148 Huang N., Baur E. vom, Garnier J. M., Lerouge T., Vonesch J. L., Lutz Y. et al. (1998) Two distinct nuclear receptor interaction domains in NSD1, a novel SET protein that exhibits characteristics of both corepressors and coactivators. EMBO J. **17:** 3398–3412
149 Gaudon C., Chambon P. and Losson R. (1999) Role of the essential yeast protein PSU1 in transcriptional enhancement by the ligand-dependent activation function AF-2 of nuclear receptors. EMBO J. **18:** 2229–2240
150 Emiliani S., Fischle W., VanLint C., AlAbed Y. and Verdin E. (1998) Characterization of a human RPD3 ortholog, HDAC3. Proc. Nat. Acad. Sci. USA **95:** 2795–2800
151 Morano K. A., Santoro N., Koch K. A. and Thiele D. J. (1999) A trans-activation domain in yeast heat shock transcription factor is essential for cell cycle progression during stress. Mol. Cell. Biol. **19:** 402–411
152 Baur E. vom, Harbers M., Um S. J., Benecke A., Chambon P. and Losson R. (1998) The yeast Ada complex mediates the ligand-dependent activation function AF-2 of retinoid X and estrogen receptors. Genes Dev. **12:** 1278–1289
153 Yang N., Shigeta H., Shi H. and Teng C. T. (1996) Estrogen-related receptor, hERR1, modulates estrogen receptor-mediated response of human lactoferrin gene promoter. J. Biol. Chem. **271:** 5795–5804
154 Bonnelye E., Vanacker J.M., Dittmar T., Begue A., Desbiens X., Denhardt D. T. et al. (1997) The ERR-1 orphan receptor is a transcriptional activator expressed during bone development. Mol. Endocrinol. **11:** 905–916
155 Greschik H. and Schule R. (1998) Germ cell nuclear factor: an orphan receptor with unexpected properties. J. Mol. Med. **76:** 800–810