The orphan COUP-TF nuclear receptors are markers for neurogenesis from cnidarians to vertebrates

GAUCHAT-FEISS, Dominique Madeleine, et al.

Abstract

In bilaterians, COUP-TF nuclear receptors participate in neurogenesis and/or CNS patterning. In hydra, the nervous system is formed of sensory mechanoreceptor cells (nematocytes) and neuronal cells, both lineages deriving from a common stem cell. The hydra COUP-TF gene, hyCOUP-TF, which encodes highly conserved DNA-binding and ligand-binding domains, belongs to the monophyletic COUP-TFs orphan receptor family (NR2F). In adult polyps, hyCOUP-TF is expressed in nematoblasts and a subset of neuronal cells. Comparative BrDU labeling analyses performed on cells expressing either hyCOUP-TF or the paired-like gene prdl-b showed that prdl-b expression corresponded to early stages of proliferation, while hyCOUP-TF was detected slightly later. HyCOUP-TF and prdl-b expressing cells disappeared in sf-1 mutants becoming "nerve-free". Moreover hyCOUP-TF and prdl-b expression was excluded from regions undergoing developmental processes. These data suggest that hyCOUP-TF and prdl-b belong to a genetic network that appeared together with neurogenesis during early metazoan evolution. The hyCOUP-TF protein specifically bound onto the [...]
The Orphan COUP-TF Nuclear Receptors Are Markers For Neurogenesis From Cnidarians To Vertebrates

Dominique GAUCHAT¹,², Hector ESCRIVA¹, Marijana MILJKOVIC-LICINA¹, Simona CHERA¹, Marie-Claire LANGLOIS³, Agnès BEGUE², Vincent LAUDET²,³ and Brigitte GALLIOT¹*

¹Department of Zoology and Animal Biology, University of Geneva, Sciences III, 30 Quai Ernest Ansermet, CH-1211 Geneva 4, Switzerland.
²UMR 49 du CNRS, Ecole Normale Supérieure de Lyon, 46 allée d’Italie, F-69364 Lyon Cedex 07, France.
³UMR 319 du CNRS, Institut de Biologie de Lille, Institut Pasteur, 1 rue Calmette, F-59019 Lille Cedex, France. † Present address: Université de Montréal, Département de Microbiologie et d’Immunologie, Faculté de Médecine, Montréal, Qc. H3T 1J4. Canada
*corresponding author: brigitte.galliot@zoo.unige.ch

Keywords: Hydra; Cnidaria; orphan nuclear receptors; COUP-TF; paired-like homeobox gene; transactivation; mechanoreceptor cells; nematocytes; neurogenesis; stem cells; evolution; BrdU labeling

ABSTRACT
In bilaterians COUP-TF nuclear receptors participate in neurogenesis and/or CNS patterning. In hydra, the nervous system is formed of sensory mechanoreceptor cells (nematocytes) and neuronal cells, both lineages deriving from a common stem cell. The hydra COUP-TF gene, hyCOUP-TF, which encodes highly conserved DNA-binding and ligand-binding domains, belongs to the monophyletic COUP-TFs orphan receptor family (NR2F). In adult polyps, hyCOUP-TF is expressed in nematoblasts and a subset of neuronal cells. Comparative BrDU labeling analyses performed on cells expressing either hyCOUP-TF or the paired-like gene prdl-b showed that prdl-b expression corresponded to early stages of proliferation, while hyCOUP-TF was detected slightly later. HyCOUP-TF and prdl-b expressing cells disappeared in sf-1 mutants becoming “nerve-free”. Moreover hyCOUP-TF and prdl-b expression was excluded from regions undergoing developmental processes. These data suggest that hyCOUP-TF and prdl-b belong to a genetic network that appeared together with neurogenesis during early metazoan evolution. The hyCOUP-TF protein specifically bound onto the evolutionarily-conserved DR1 and DR5 response elements, and repressed transactivation induced by RAR:RXR nuclear receptors in a dose-dependent manner when expressed in mammalian cells. Hence a cnidarian transcription factor can be active in vertebrate cells, implying that functional interactions between COUP-TF and other nuclear receptors were evolutionarily conserved.

INTRODUCTION
Hydra is a freshwater bilayered animal, organised as a tube with at one extremity, a single opening surrounded by a ring of tentacles, defining the head region, and at the other, a basal disc used by the animal to attach to the substrate. Hydra belongs to the Cnidaria, which together with Ctenophora are the first metazoan phyla in the animal kingdom to differentiate highly specialized cell types such as myoepithelial cells, nematocytes and neurons. The emergence of active movements and feeding behavior, which relies on the presence of a nervous system, is a major hallmark of metazoan organization. Therefore hydra represents an excellent model to study molecular mechanisms implicated in the early differentiation of nervous systems in animal evolution. Moreover, cnidarians also provide a model system for developmental biology, since they display an oral-aboral patterning process, a process easily amenable to experimentation during budding, regeneration, transplantation and reaggregation (Holstein et al., 2003).

In cnidarians, active behaviors rely on a simple nervous system, composed of nerve cells, either sensory or ganglia, and nematocytes, which are mechanoreceptor cells, specific to the Cnidaria phylum. In hydra, nematocytes are abundant, representing 35% of all cell types (David, 1973), particularly in tentacles where they are responsible for the food capture (Tardent, 1995). The cnidocil apparatus of nematocytes displays similarities with cnidocils present in sensory cells of nematodes, insects and vertebrates and might thus represent a progenitor of metazoan mechanoreceptors (Holstein and Hausmann, 1988; Holtmann and Thurm, 2001b). Upon stimulation of their cnidocil by a prey, nematocytes discharge a highly osmotically
pressed capsule named nematocyst that contains anions and toxins. This discharge process occurs in few milliseconds (Holstein and Tardent, 1984) and can function independently of the nervous system (Aerne et al., 1991). Thus, nematocytes can be considered as autonomous receptor-effector units, even though synaptic connections to nerve cells were demonstrated ultrastructurally (Westfall, 1996; Westfall et al., 2002) and shown to regulate their activity (Holtmann and Thurm, 2001a; Kass-Simon and Scappaticci, 2002). The neuronal and nematocyte cell lineages derive from a common stem cell, a pluripotent interstitial cell, which can also differentiate towards the gland cell lineage and the gametes (Bode, 1996). In hydra, precursor cells to nerve cells are blocked in G2 and, upon appropriate signaling, undergo a limited number of mitotic events that will provide mature nerve cells (Schaller et al., 1989). In contrast, precursors to nematocytes (nematoblasts), which are located in the ectodermal layer along the body column, undergo a variable number of synchronous divisions, from 3 to 5, forming nests of cells communicating by cytoplasmic bridges, before differentiating the nematocyst, a sophisticated Golgi-derived intracellular structure (Lehn, 1951; Rich and Tardent, 1969; Tardent, 1995). Nematocysts can be of four different types, stenoteles, desmonemes, holotrichous and atrichous isorhiza (see in Holstein and Emschermann, 1995), according to a body axis position-dependent differentiation process (Fujisawa et al., 1986). Once harboring mature capsules, most nematocytes will migrate towards the tentacles according to a guiding process (Weber et al., 1978; Campbell and Marcum, 1980).

The underlying genetic mechanisms responsible for the emergence of mechanoreceptor and neuronal cell types along evolution are still poorly understood (Anderson, 1990; Mackie, 1990); however the cnidarian version of a dozen genes known to regulate neurogenesis in bilaterian species was characterized within the last ten years. Among those, four gene families were shown to be expressed in the nematocyte lineage in hydra: CnASH, the Achaetescute homolog (Grens et al., 1995), prdl-b, a paired-like homogene (Gauchat et al., 1998), hyzic, a Zic homolog (Lindgens et al., 2004) and hyDkk3, the Dickkopf-3 homolog (Fedders et al., 2004).

In this paper, we focused on the cellular and developmental regulation of the hydra COUP-TF homolog gene (Escriva et al., 1997) and used the prdl-b gene for comparative cell analyses. COUP-TFs genes bring a major contribution to neurogenesis and neurophysiology in vertebrates (Pereira et al., 2000; Cooney et al., 2001) and also affect the development of nervous system in Drosophila (Modzik et al., 1990) and amphioxus (Langlois et al., 2000). Nuclear receptors (NRs) are ligand-dependent transcription factors, activated by steroid hormones, and non-steroid molecules such as retinoic acid, thyroid hormone and vitamin D (Moras and Gronemeyer, 1998). However, no cognate ligand was identified for the COUP-TFs and EAR-2 NR, which are therefore considered as “orphan” NR (Giguere, 1999). Together with the HNF4 (NR2A), RXR (NR2B), TR2/4 (NR2C), DHR78 (NR2D) and TLL (NR2E) NRs, COUP-TFs (NR2F) define the class II of the NR superfamily (Nuclear Receptor Nomenclature Committee, 1999).

COUP-TF homologs were identified in a wide range of metazoans from cnidarians to protostomes and deuterostomes (Escriva et al., 1997; Langlois et al., 2000; Grasso et al., 2001; Devine et al., 2002), showing an extensive degree of sequence identity in the DNA-binding domain (C domain) and the ligand-binding domain (E domain). In cnidarians, a COUP-TF gene was detected in hydra, RXR and FTZ-F1 genes in the Anemonia sulcata (Escriva et al., 1997), a RKR homolog was characterized in the jellyfish Tripedalia cystophora (Kostrouch et al., 1998) and more recently, six distinct NR genes with clearly identifiable DNA-binding and ligand-binding domains were isolated in the coral Acropora millipora (Grasso et al., 2001), four of them being assessed to the class II NR. These results demonstrate that COUP-TF and RXR genes appeared very early during metazoan evolution, before the divergence between Cnidaria and Bilateria. Moreover, their high degree of conservation in the functional domains implies that they support essential functions among all metazoans. In this paper, we show that hyCOUP-TF like the paired-like gene prdl-b is expressed in both, nematocyte and neuronal cell lineages, as well as in a subset of interstitial stem cells. In the temperature-sensitive sf-1 mutant that eliminates its interstitial cell lineage at non-permissive temperature hyCOUP-TF and prdl-b expressing cells disappear. In addition measurements of the BrdU labeling indexes in hyCOUP-TF and prdl-b expressing cells clearly show that both genes are induced at early stages of the nematocyte pathway, although with different kinetics and distinct regulations along this pathway. Finally, when expressed in mammalian cells, hyCOUP-TF is able to repress transactivation of reporter constructs induced by RAR:RXR NR.

**MATERIAL AND METHODS**

**Hydra culture, regeneration experiments and production of “nerve free” animals**

*Hydra vulgaris* ( Hv) from the Zürich (gift of S. Hoffmeister) and Holland ( gift from D. Campbell) strains, *Hydra magnipapillata* (Hm) from the Hm 105 and sf-1 strains ( gift from T. Hosleitn) or AEP *Hydra*
**FIGURE 1**: (A) Scheme depicting the four distinct types of *hyCoup-TF* cDNAs isolated from *Hydra vulgaris*. These cDNAs encode either the C and D domains (type I), or the D, E and F domains (types II, III and IV) and differ by their unrelated distinct 5' UTR sequences. The position and the orientation of the C1, C2, C3, C4, D, E2 primers used in the RT-PCR experiment described in C are indicated. (B) Northern analysis of mRNA prepared from intact hydra showing two main bands indicated by arrows. (C) RT-PCR products obtained with each of the C1, C2, C3, C4, D, E2 primers used in the RT-PCR experiment described in C are indicated. Only cDNAs corresponding to D+E (lane D) or C+E (lane C1) domains were amplified. M: marker. (D) Scheme describing the structure of the *hyCoup-TF* isoforms whose deduced sequences are aligned in E, and the chimeric constructs used in functional assays. The *hyCoup-TF* I isoform, which was confirmed by RT-PCR in C (lane D), was produced by assembling overlapping type I and III cDNAs. Two distinct isoforms were characterized from the RT-PCR products shown in C (lane C1): *hyCoup-TF* 2 and *hyCoup-TF* 3, each of them encoding two distinct putative translation products. For the *hyCoup-chim1* construct, 25 residues including 21 residues of the *huCoup-TF1* C domain missing in the type I cDNA were inserted in front of *hyCoup-TF1*. For the *hyCoup-chim2* construct, the full *huCoup-TF1* N-terminus (A/B and partial C domains) was inserted in front of *hyCoup-TF1*. Initiation codons are indicated with horizontal arrows, stop codons with stars. (E) Alignment of the 5 putative *hyCoup-TF* isoforms sequences deduced from the different cDNAs (given in supplement 1) and the three distinct RT-PCR product sequences (given in supplement 3). Residues in italics correspond to the human N-terminal part of the C-domain (grey background). Accession numbers: *hyCoup-TF1*: Y69063; *hyCoup-TF2*: Y690638; *hyCoup-TF3*: Y690639.

*vulgaris* (gift from T. Bosch) were cultured in hydra medium (HM: 1 mM NaCl, 1 mM CaCl2, 0.1 mM KCl, 0.1 mM MgSO4, 1 mM Tris pH 7.6) and fed five days a week with hatched *Artemia nauplii*. After a two days starvation period, regeneration experiments were performed at 19°C on budless hydra, bisected at mid-gastric position. Sexual hydra were either collected among the *Hv* Holland strain or obtained from the *Hv*
AEP strain after induction (Martin et al., 1997). In order to obtain “nerve free” hydra, sf-1 hydra were induced at 26°C for two days (Marcum et al., 1980).

**Isolation of hyCOUP-TF cDNAs**
A 117 bp PCR fragment corresponding to the hyCOUP-TF gene (Escriva et al., 1997) was used as a probe to screen a lambda-gt11 Hv cDNA library (gift from C. Schaller). A single clone, named cDNA type I, 981 bp long, was isolated out of 106 phages. The region of this clone encompassing the C and D domains was used to screen a second cDNA library prepared from adult hydra polyps (gift from H. Bode). Hybridization were performed overnight at 37°C in 40% formamide, 7x SSC, 5X Denhardt’s solution, 0.5 % SDS, 0.1 mg/ml denatured salmon sperm DNA. Membranes were washed twice in 2x SSC, then 0.1x SSC at 37°C for 15 min. The inserts were subcloned in BlueScript vectors and sequenced on both strands.

**Phylogenetic analysis of the hydra COUP sequences**
NR sequences related to the deduced hyCOUP-TF protein were collected on databases using the Blast search on the Expasy server (www.expasy.org). 51 representative sequences were selected and aligned (see supplement 4) using ClustalW at the Pole Bioinformatique Lyonnais (http://pbi.univ-lyon1.fr) and the Genetic Data Environment 2.2 software (Larsen et al., 1993). Sequences and alignments can be found on Nurebase: http://ens-lyon/LBMC/laudet/nurebase.html (Duarte et al., 2002). Phylogenetic analyses were performed on sequences corresponding either to the C-F domains (318 residues after pairwise gap removal), or only the C domain (60 residues) following the procedure previously described (Galliot et al., 1999).

**Human-hydra hybrid constructs**
The hyCOUP-chim1 clone was constructed by the “overlap extension” mutagenesis technique (Yon and Fried, 1989) using the hyCOUP-TF type I and type III cDNAs as templates and adding 25 residues (MQQHIECVCGKSSGKHYGQFTCE) from the evolutionarily-conserved N-terminus of the huCOUP-TF1 C domain to the hyCOUP-TF N-terminus. Three consecutive PCR were performed using the following primers: 5C-hu: attatcagccacacccctcgctgctg, 3C-hu: tgcctaatcgtctcgtctgagaaacttt, 5D-hy: ctaaggtttctcatggttggtacg, 5C-chim: aaaaattctccaagaggaacattacg, 3D-hy: ttgatccattatatgttcgccc, 3tot-hy: cttggttccattatatgttcgccc. PCR1 amplified 3 distinct templates: huCOUP-TF1 (5C-hu, 3C-hu primers), hyCOUP-TF type I (5Chy, 3D-hy primers) and hyCOUP-TF type III (5D-hy, 3tot-hy primers). PCR2 was performed on 200 ng of the previously amplified and annealed hyCOUP-TF type I and type III fragments, and provided the full hyCOUP-TF1 (5C-hy, 3tot-hy primers). Finally, PCR3 was performed on 200 ng of the previously amplified huCOUP-TF1 (PCR1) and full hyCOUP-TF (PCR2) fragments using the 5C-hu and 3tot-hy primers. Profiles of PCR1, PCR2 and PCR3 were identical (94°C 10 min, [94°C 30 sec, 45°C 30 sec, 72°C 3 min] x5, [94°C 30 sec, 45°C 30 sec, 72°C 1 min] x20, 72°C 7 min). The PCR3 amplification product (hyCOUP-chim1, 1438 bp) was cloned into the PCR-TOPO vector (Invitrogen) and sequenced. The hyCOUP-chim2 clone was constructed similarly by adding to the 5’ end of hyCOUP-chim1 the region corresponding to the A/B domain of huCOUP-TF (94 residues). The GC-rich 271 bp fragment encompassing the huCOUP-TF1 5’ end was first amplified with the huCOUP-for1 (tagatagtggcaatgtagtagcagc) and huCOUP-rev2 (ccgccacacccacctcgtgctgctg) primers using the Advantage GC cDNA polymerase mix (Clontech) 1/5 (94°C 5 min, [94°C 30 sec, 49°C 1 min, 72°C 1 min] 30x, 72°C 10 min). The hyCOUP-chim1 fragment was amplified with the COUP-chim1.1 (cagcagcacatcg-aggttgctg) and COUP-chim1.2 (cattgtagcttaagaggagc) primers using the Qiagen Taq polymerase (94°C 5 min, [94°C 30 sec, 53°C 1 min, 72°C 2 min] 30x, 72°C 10 min). Both fragments were purified from gel, mixed, purified again (ROCHE, High Pure PCR Product Purification Kit) and then annealed in the absence of any primers but in the presence of Advantage GC cDNA polymerase mix, Taq polymerase (50X Taq mix used 1x), dNTP (0.2 mM each), GC cDNA PCR reaction buffer for 15 min at 45°C after denaturation (94°C 5 min) following a slow temperature downshift (1°C per minute). The huCOUP-for2 (cccatgatatgctgaatttagtgacagc) and COUP-chim.1.2 primers were subsequently added and the 1710 bp hybrid product was amplified (94°C 5 min, [94°C 30 sec, 53°C 1 min, 72°C 2 min] 30x, 72°C 10 min). After reamplification, the hyCOUP-chim2 PCR product was first sub-cloned into the pDrive vector (Qiagen), digested BamHI /SalI and then inserted into the pSG5 vector at the BamHI site for eukaryotic expression.

**In situ hybridization (ISH)**
Standard in situ hybridization was performed on whole mount animals, following the procedures described in (Greens et al., 1996). The digoxigenin (DIG)-labelled *hyCOUP-TF* riboprobe was prepared after *Pst*I digestion of the *hyCOUP-TF* type I cDNA (800 bp); the DIG-labelled *prdl-b* riboprobe, after *Nsi*I digestion of the *prdl-b* cDNA (clone 33, 700 bp). After NBT-BCIP staining, animals were stained with DAPI (0.01 µg/ml in PBS) for 2 to 5 min at RT, then washed 2x 5 min in PBS and once in water. Samples were mounted in DABCO (Sigma) and nests were counted on a Zeiss Axiosplan2 microscope. For cell-type determination, 10 µm cryosections were performed after NBT/BCIP or Fast Red staining. Double-labeling in situ hybridization was carried out in the InSituPro robot (INTAVIS Bioanalytical Instruments AG) as pre-
FIGURE 2: Alignment of various COUP-TF group members with the eukaryotic COUP-TF homologs, by COUP-TF1 and AmNR7. Identical amino acid positions are indicated by dashes, dots correspond to gaps. The A/B and C regions were not included since too variable. The C domain is shaded in grey and the D domain is boxed. The consensus shows residues conserved from eukaryotes to bilaterians that were derived, residues written in lowercase indicate conservative substitutions. Accession numbers of the various sequences: HyCOUP-TF1: U93425, AmNR7: trqQ5WLF7; COUP-TF1 Hs: sp1P10589; COUP-TF1 Xi: trqQ9W745; COUP-TF Xi trqQ91720; EAR2 Mm: sp1P43136; COUP Bf: AY211769; COUP Sp trqQ26641; Bm sp1Q69A9; NHR-Ae trqQ9BMU6; SVM Dm sp1P16375; RXR Mo X: M66223; RXR Tc: AF901121.
viously mentioned with following modifications: Animals were incubated in hybridization solution (50% formamide, 5x SSC, 0.1% Tween 20, 0.1% CHAPS, 1x Denhardt’s solution, 0.01% heparine, 0.02% tRNA in DEPC H2O) containing 1 to 2 µg/ml DIG-labeled prdl-b RNA probe and a fluorescein-labeled COUPI RNA probe for 16 hours at 55°C. For detection, samples were incubated with two antibodies simultaneously, the sheep anti-DIG Fab fragments coupled to alkaline phosphatase (1:2000, Roche) and the sheep anti-fluorescein Fab fragments coupled to horse-radish peroxidase (1:100, Roche) during 6 hours at RT in the robot. Unbound antibodies were removed by numerous washes in MAB (100 mM maleic acid, 150 mM Tris HCl, pH 7.5): 4x 12 min, 4x 30 min, 4x 1 hour, 2x 2 hours. Animals were then washed in TNT buffer (150 mM NaCl, 100 mM Tris HCl, pH 7.5, 0.05% Tween-20) 3x 5 min and incubated in biotinyl-tyramide mix diluted in the amplification buffer (1:50, PerkinElmer Life Sciences) for 30 min at RT, before being washed again in TNT buffer 3x 5 min. Alexa-488-conjugated streptavidin antibody (1:100, Molecular Probes) was added in the amplification buffer and incubation was carried out at RT for 30 min. Animals were subsequently washed in TNT buffer 3x 5 min and stained with Fast Red (DAKO Cytomation) for 15-90 min at RT in the dark. Staining was stopped by washing in TNT buffer 3x 5 min at RT. Samples were mounted as above and screened for fluorescent signals on a Zeiss Axiplan2 microscope equipped with the FITC, rhodamine and DAPI filters.

**Bromodeoxyuridine (BrdU) labeling**

Hydra were starved for one day and then exposed to 5 mM BrdU (Sigma) for variable periods of time as indicated. Animals were then fixed in PFA 4% and processed for in situ hybridization. After riboprobe detection, animals were washed in PBST (PBS, 0.1% Triton X100) 3x 5 min, treated with 2N HCl for DNA denaturation for 40 min at RT, washed in PBST at least 5x over 15 min and then incubated with the anti-BrdU antibody (1:20, BrdU labeling and detection Kit I Roche) for 1 hour at 37°C. After 3 washes in PBS animals were incubated for 3 hours at RT with the anti-mouse Alexa Fluor 488 antibody (1:500, Molecular Probes), washed again in PBS, mounted in Mowiol and screened for fluorescence as given above. Counting of nests displaying BrdU labeled nuclei was performed with the 100x objective on 40 to 100 nests per animal expressing either hyCOUP-TF or prdl-b and a minimum of 5 animals was counted for each condition.

**Gel retardation assays**

The hyCOUP-chim1, hyCOUP-chim2 and amphiCOUP plasmids were used to produce the corresponding proteins in reticulocyte lysate according to the supplier procedure (TnT kit, Promega). The double-stranded oligonucleotides DR1: cgcgatttgactcaggtctgccagtta, DR5: cggatttgacccggtcctcaaggtcactcaggtctgccagtta and unr: cggagaaggtgcatcctccaaac-aac (Langlois et al., 2000) were labeled and incubated for binding as in (Galliot et al., 1995). For supershifts experiments, the polyclonal anti-huCOUP-TF1 T19 antibody (SantaCruz Biotechnology) was preincubated with 2.5 µl protein lysate for 15 min on ice. Then 10 fmole of the end-labelled DR5 were added and incubation was continued for 15 min at RT. Samples were loaded on 4% PAGE and run at 150 volts at RT with recycling buffer.

**Transactivation assays in mammalian cells**

Ros 17/2/8 (rat osteosarcoma) cells were maintained in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 5% fetal calf serum (FCS). A total of 1.5x10^5 cells in 24-well plates was transfected using 4 µl ExGen500 (Euromedex) with 1.0 µg total DNA including 0.1 µg reporter plasmid, 5 ng of AmphiRAR and AmphiRXR each and increasing amounts of AmphiCOUP (10 to 500 ng) and hyCOUP-chim2 (100 to 1000 ng). AmphiRAR, AmphiRXR, AmphiCOUP and hyCOUP-chim2 were cloned into the pSG5 expression vector (Stratagene). Three tandem repeats of oligonucleotides encompassing consensus DR5 sequence were inserted into the pGL2-promoter vector (Promega). The culture medium was changed 6 hours after transfection and all-trans retinoic acid (RA) diluted in DMSO was added (10 µM final). Cells were lysed 24 hours after transfection and assayed for luciferase activity.

**RESULTS**

**Isolation of multiple hyCOUP-TF isoforms in hydra**

A 117 bp probe corresponding to the hyCOUP-TF C domain was previously obtained from a PCR screen for NR in various metazoans (Escriva et al., 1997) and used for screening a Hydra vulgaris (Hy) cDNA library. A 981 bp long clone (cDNA type I) was isolated encoding a partial C-domain (60 residues), 81% identical to that of the human COUP-TFI (huCOUP-TFI) protein and a less conserved D domain (Fig. 1A and supplement 1). Upstream to the region encoding the C domain, a 686 bp divergent DNA stretch containing a stop codon at position 663 was found. In order to obtain full length cDNAs, the conserved regions of this cDNA were used to screen a second library. Ten distinct new clones were isolated, representing 3 distinct classes of cDNAs (types II, III and IV) all encoding ORFs showing similarities with the COUP-TF D and E domains, but lacking the C domain (Fig. 1A and supplements 1, 2). Thus these cDNAs only shared the D domain with the cDNA type I. These 3 types of cDNAs exhibited variable and divergent 5’ sequences, bearing no sequence identity.
to any sequence, but containing stop codons and thus likely representing non-coding sequences (described in supplemental data). Surprisingly, these 5' sequences diverged at the exact same position, at the boundary between the C and D domains (position 831 in cDNA type I, supplement 2B).

In order to confirm the presence of such transcripts, RT-PCRs on total RNA from hydra lower halves taken 3h after bisection were performed using specific primers located along the C (C1), D (D), and E (E2) domains and on the four divergent 5' sequences of cDNAs type I, II, III and IV (C0, C2, C3, C4 respectively, Fig. 1A). The RT-PCRs that used primers located at the 5' divergent regions failed to amplify any DNA fragment (Fig. 1C and not shown). Moreover, despite several attempts (inverted PCR, genomic library screening, PCR with conserved oligonucleotides), we were unable to isolate the corresponding genomic region. Since the divergent 5' regions of the various cDNAs were not detected by RT-PCR, we suspect that these cDNAs represented chimeric cDNA library artifacts. In contrast, the RT-PCRs that used the D+E2 primers detected one main PCR product while at least two distinct products were detected with the C1+E2 pair of primers (Fig. 1C). Interestingly, Northern analysis detected two main bands in the range of 3 and 3.5 kb, confirming thus the presence of at least two distinct hyCOUP-TF transcripts in hydra (Fig. 1B).

The molecular characterization of the C1+E2 RT-PCR products detected three distinct isoforms: hyCOUP-TF1, which encodes the C, D and E domains previously characterized in cDNAs type I-IV; hyCOUP-TF2, which contains a 131 bp deletion within the D region and finally, hyCOUP-TF3, which shows a 156 bp insertion also within the D region (Fig. 1D, 1E and supplement 3). Translation of hyCOUP-TF2 and hyCOUP-TF3 isoforms generates stop codons that produce truncated hyCOUP-TF proteins lacking the E domain, hyCOUP-TF2.21 and hyCOUP-TF3.2 respectively. Alternatively, these isoforms could also generate truncated hyCOUP-TF proteins that would contain only the E domain, hyCOUP-TF2.22 displaying an initiation codon in the D domain and hyCOUP-TF3.3 showing an initiation codon and a divergent sequence up to position 25 of the E domain. Therefore, given that RT-PCR experiments could detect transcripts where the regions encoding the C and E domains were linked, but in the absence of characterized hyCOUP-TF5's end, we decided to fuse the C-F domains obtained in type I and III cDNAs in a construct named hyCOUP-TF1. HyCOUP-TF1 was further used to construct chimeric human-hydra COUP-TF clones (Fig. 1D), which would contain either a complete C domain (hyCOUP-chim1) or a complete N-terminus with the addition of the human A/B domain (hyCOUP-chim2).

**Conservation of the COUP-TF functional domains from cnidarians to vertebrates**

The Blast search provided hyCOUP-TF related sequences that all belong to the class II NR: COUP, RXR, HNF4, TLL, TR2/4 (Nuclear Receptor Nomenclature Committee, 1999). When the COUP-TF C, D and E domains were aligned (Fig. 2), the overall identity of hyCOUP-TF was 47% with huCOUP-TFII, 45.7% with the human COUP-TFII, 43.8% with the coral AmNR1, 42.9% with the human EAR2 orphan receptor and 44.6% with the Drosophila SVP protein. These values reached 55% when conservative substitutions were taken into account. By comparison, the identity level was only 30.7% with the jellyfish RXR indicating that we isolated a hydra gene representative of the COUP-TF gene family.

Phylogenetic analyses performed either on the C-E domains (318 residues, Fig. 3) or on the C domain (60 residues, not shown) of 51 class II NR protein sequences confirmed that hyCOUP-TF is indeed a member of the COUP-TF group. However the coral COUP-TF homolog (AmNR7, (Grasso et al., 2001) exhibited a far more conserved sequence than the hyCOUP-TF, 80% identical to the huCOUP-TF in the C-E domains. In hyCOUP-TF, the C domain is the most conserved one, reaching 81 to 83% identity with vertebrate COUP-TF C domains. Among the different isolated cDNAs, the hyCOUP-TF type I was the only one to contained a C domain, although partial because lacking the 21 N-terminal amino acids when compared to the huCOUP-TF1, including the three first cysteines of the first zinc finger (Fig. 2). The similarity to NR actually started inside the recognition helix (CEGCKSFKRKSVR) a highly conserved motif, which mediates the direct interaction of the receptor to the major groove of the DNA double helix (Holmbeck et al., 1998; Zhao et al., 2000). Similarly, the second helix displays a very high level of conservation (positions 34-45 in Fig. 2). A very strong pressure was thus maintained along evolution on residues that are critical to the DNA-binding function. These observations strongly argue in favor of the fact that a regular COUP-TF protein, i.e. containing a complete and functional DNA-binding domain, is present in hydra.

The E domains from hyCOUP-TF and huCOUP-TFII harbor 44% sequence identity (Fig. 2). The homology is scattered over nine regions that roughly correspond to the α-helices determined by the 3D structure resolution of the RARα and RXR E domains (Bourguet et al., 1995; Renaud et al., 1995). This suggests that hyCOUP-TF, as its homologs in other metazoa, will fold in a common tertiary structure (Wurtz et al., 1996). The last remarkable feature of hyCOUP-TF is the presence of a long F domain, C-terminal to the E domain. This domain does not show identity to any other proteins. F domains are not
uncommon in NRs although their functions often remain elusive (reviewed in [Laudet and Gronemeyer, 2002]).

**hyCOP-TF and prdl-b are both expressed in the nematocyte and the neuronal cell lineages**

In situ hybridizations were performed with a **hyCOP-TF** riboprobe corresponding to the D and E domains and thus less prone to cross-hybridize with other nuclear receptor genes expressed in hydra. In adult polyps, **hyCOP-TF** was expressed as ectodermal multicellular spots located in the body column, but absent from the apical and basal regions (Fig. 4A-C). Rare scattered groups of **hyCOP-TF** expressing cells were detected in the peduncle and upper body column. This body column pattern was observed in budless polyps as well as in late-stage buds. **HyCOP-TF** expressing cells were clustered in nests of 2 to 16 cells, identified as dividing and differentiating nematoblasts (Fig. 4C, 5A-F). In heavily stained animals, we detected a second population of **hyCOP-TF** expressing cells located deeper than the nests of nematoblasts, along the mesoglea (Fig. 4D), appearing as small and single cells most often exhibiting long processes.

---

**FIGURE 3** : Phylogenetic relationships among 51 nuclear receptor sequences representing the COUP, RXR, HNF4, TLL families obtained by neighbour-joining method (Saitou and Nei, 1987) with a PAM-Dayhoff matrix (Dayhoff et al., 1978). Gnidarian sequences are written boldcase and boxed in grey. The tree was constructed with the C-E domains (318 residues after pairwise gap removal). Bootstrap values above 80% (500 replicates) are indicated. Species code: Aa: *Amblyomma americanum*; AaL: *Aedes albopictus* (forest day mosquito); Ae: *Aedes aegypti* (yellow fever mosquito); Ame: *Apis mellifera* (honeybee); Ami: *Acropora millipora* (coral); Bf: *Branchiostoma floridae* (amphioxus); Bg: *Biomphalaria glabrata* (blood fluke planorb); Bm: *Bombus mori* (silk moth); Br: *Brachydanio rerio* (zebrafish); Bt: *Bos taurus* (bovine); Fr: *Fugu rubripes* (pufferfish); Dm: *Drosophila melanogaster* (fly); Hs: *Homo sapiens* (human); Hv: *Hydra vulgaris*; Mm: *Mus musculus* (mouse); Oi: *Oryzia latipes* (medaka fish); Pm: *Petromyzon marinus* (sea lamprey); Po: *Paralichthys olivaceus* (flounder); Sp: *Strongylocentrotus purpuratus* (sea urchin); Tc: *Tripedalia cystophora* (jellyfish); Tm: *Tenebrio molitor* (yellow mealworm); Xi: *Xenopus laevis* (african clawed frog). The sequence alignment used for the construction of this tree is available in supplement 4.
characteristic of neuronal cells (Fig. 4E). These neuronal hyCOP-TF expressing cells were also observed in the hypostome region (Fig. 4D).

Similarly to hyCOP-TF, the prdl-b transcripts were detected in nests of cells spread along the body column, but absent from the apical and basal regions of the adult polyps (Fig. 4F, 4G) as previously described (Gauchat et al., 1998; Miljkovic-Licina et al., 2004). Interestingly, we also observed prdl-b expression in pairs of small dividing interstitial cells (Fig. 4H), and in small single cells, some of them displaying processes typical of neuronal cells (Fig. 4I). This prdl-b expression in the nerve cell lineage, which is more easily detected with fluorescent staining, was initially overlooked because of the massive expression in the nematoblasts. Therefore, in hydra hyCOP-TF and prdl-b are expressed in two distinct lineages that differentiate from a common interstitial stem cell (Bode, 1996). In fact, it was proven that apical and basal regions exhibit an inhibitory influence on nematocyte differentiation, which as a consequence, is excluded from these regions and restricted to the ectodermal layer of the body column (Fujisawa et al., 1986; Bode, 1996). In contrast, there is no restriction for neuronal differentiation.

**HyCOP-TF and prdl-b show partially overlapping cellular distributions in the nematocyte lineage**

To test whether prdl-b and hyCOP-TF were co-expressed at similar stages in the nematocyte lineage, we carried out double in situ hybridizations with the prdl-b and hyCOP-TF riboprobes (Fig. 4J-L). Colocalisation of the hyCOP-TF (green) and the prdl-b (red) signals was clearly observed in a large subset of cells (stars), reaching 85% ± 11.3 of the nests in case of prdl-b but only 63% ± 7.6 in case of hyCOP-TF. This difference reflects the fact that hyCOP-TF is expressed in a larger number of nests than prdl-b. In fact, 37% of the hyCOP-TF nests did not express prdl-b (Fig. 4J-L, arrowheads), while 15% of the prdl-b nests did not express hyCOP-TF (Fig. 4J-L, arrows). We furthermore identified more precisely the stages of the nematocyte pathway where hyCOP-TF and prdl-b were expressed, by counting the distribution of the hyCOP-TF and prdl-b expressing nests according to their size in animals fixed 24 hours after the last feeding (Fig. 5). HyCOP-TF and prdl-b were detected in nests of 2 to 16 cells (Fig. 5A-L), 4 and 8 cells clusters representing over 80% of the positive nests for both genes (Fig. 5M). Moreover, a comparative analysis of DIC views and DAPI staining showed that, beside the nucleus a typical differentiating capsule could be observed in about 30% of the hyCOP-TF and prdl-b expressing nests (Fig. 5D-F, 5I-L, arrowheads), proving that hyCOP-TF and prdl-b transcripts are still present at the differentiation stage. However, hyCOP-TF and prdl-b transcripts were never detected in nematocytes displaying a fully mature nematocyte, which displays a typical moon-shape nucleus (Fig. 5D, 5E, 5H, stars). Beside these similarities, we noted some clearcut differences between the hyCOP-TF and prdl-b cell populations: first, the most abundant nests were the 8 cells clusters in case of hyCOP-TF, representing about 55% of all nests, instead of the 4 cells clusters in case of prdl-b that represent about 50% of all nests (Fig. 5M). Second, 2 cells clusters were more frequently detected among prdl-b expressing nests, representing 8.5% of total nests versus 1.7% in case of hyCOP-TF (Fig. 5M). These results prove that hyCOP-TF and prdl-b are largely co-expressed in dividing nematoblasts; nevertheless hyCOP-TF expression is likely initiated in nematoblasts having undergone at least two rounds of cell divisions (4 cells clusters) while prdl-b expression is initiated at two different stages, first earlier than hyCOP-TF in pairs of interstitial cells and at a later stage in 4 cells clusters. This result is in agreement with our previous findings according to which, prdl-b expression was detected in the small interstitial cell fraction by elutriation (Gauchat et al., 1998). These pairs of interstitial cells represent dividing stem cells as well as precursors to nematoblasts and neuronal cells (Holstein and David, 1990; Bode, 1996). The 2 cells clusters that express hyCOP-TF and/or prdl-b could thus represent three distinct sub-populations, interstitial stem cells, precursors to nematoblasts and precursors to neuronal cells.

**HyCOP-TF and prdl-b expressing cells show distinct regulations upon starvation**

In the epithelial cell lineage, starvation does not lead to arrest of cell proliferation (Bosch and David, 1984) but rather induces on one hand an increase of the cell cycle length resulting in a 2 fold decrease of BrdU-labeled cells in the body column (Holstein et al., 1991) and on the other hand a phagocytotic process (Bosch and David, 1984), following a typical apoptotic process (Cikala et al., 1999). Hence, starvation results in dramatic cellular rearrangements. In order to compare the specific regulations of hyCOP-TF and prdl-b expressing nests in this context, we measured their respective distributions according to the size of the nests after 1, 2 or 3 days of starvation (Fig. 5M). The distribution of hyCOP-TF expressing nests showed a dramatic 4.5 fold increase from 1.7% ± 0.5 to 7.6% ± 1.5 in the number of 2 cells clusters after 3 days of starvation. Similarly, in case of the prdl-b gene, we recorded a 4.3 fold increase in the number of 2 cells clusters from 8.5% ± 2.9 to 36.8% ± 7.5. However, in case of hyCOP-TF, we did not register any significant variation in the number of 4 and 8 cells clusters. In contrast, in case of prdl-b together
**FIGURE 4:** *HyCoup-TF* (A-E) and *prdl-b* (F-I) are expressed in both the nematocyte and neuronal cell lineages. (A) *HyCoup-TF* expression pattern in hydra polyp. (B, D) DIC views of cryosections performed after whole mount ISH. (B) Expression of *HyCoup-TF* in nests of nematoblasts (arrowhead) restricted to the body column of adult and budding polyps. (C) Enlarged view of the body column showing nests of *HyCoup-TF* expressing cells. (D) Enlarged view of the apical region of the animal depicted in B showing scattered single small *HyCoup-TF* expressing cells in the head region (arrows). (E) Neuronal cell (nv) expressing *HyCoup-TF* (arrow). (F) *prdl-b* expression pattern in hydra polyp. (G) DIC view of a cryosection performed after double whole mount ISH where *prdl-b* expressing cells are red (arrow) and *HyCoup-TF* blue (arrowhead). (H) Confocal view of single cells and mitotic pairs of interstitial cells expressing *prdl-b* (arrows). (I) Confocal views of a single cell displaying processes typical of neuronal cells (arrowhead). (J-L) Co-localisation of nests expressing *HyCoup-TF* (green) and/or *prdl-b* (red). Arrowheads indicate nests expressing *HyCoup-TF* but not *prdl-b*, arrows point to nests expressing *prdl-b* but not *HyCoup-TF* and stars label nests that co-express *HyCoup-TF* and *prdl-b*. cy: nematocyte, ect, ectoderm; end, endoderm; ms, mesoglea; nb, nematoblasts; tae: taeniolaform. Bars: 400 μm in A, B; 50 μm in D; 20 μm in C, G; 7 μm in E; 64 μm in J-L; 3.6 μm in H; 1.8 μm in I.
with the increase in two cells clusters, we noted a symmetrical 4.1 fold decrease in the number of 8 cells clusters from 41.4% ± 3.9 to 10.5% ± 5.6 and a complete disappearance of the 16 cells clusters but no variation of the 4 cells clusters.

As previously mentioned the two cells clusters are composed of two distinct cell populations: the interstitial stem cells, characterized by a short cell cycle and the early precursors to several differentiation pathways, among which the precursors to nematocyte differentiation, which are medium-cycle cells (Holstein and David, 1990). At subsequent stages, the nematoblast nests are characterized by a tightly-tuned coordination between a variable number of synchronous divisions followed by the differentiation of the capsule, which is also a synchronous process within a nest (Bode, 1996). Consequently, the growth rate of the nematoblast population is the result of an equilibrium between the number of synchronous divisions, the stage where nematoblasts enter the differentiation process and possibly the magnitude of the apoptotic process (Fujisawa and David, 1984).

According to these parameters, we can propose the following explanations for the observed results: First, the dramatic increase in the ratio of 2 cells clusters expressing hyCOUP-TF and prdl-b can be explained by the fact that the division rate of stem cells, which are short-cycling cells (Holstein and David, 1990) is not affected upon starvation, while precursors to nematocyte differentiation, which are medium-cycling cells (Holstein and David, 1990) display a cell cycle lengthened upon starvation. The combination of these two mechanisms would lead to an increase in the pairs of interstitial cells. Second the absence of variation in the fraction of 4 cells clusters expressing hyCOUP-TF can simply be explained by the lengthening of the cell cycle of nematoblasts whatever their stage: the number of 4 cells clusters produced from 2 cells clusters is lower but assuming that the cell cycle length is also increased upon starvation in 4 cells clusters (the traverse of the 4 cell stage takes longer), then the number of 4 cells clusters expressing hyCOUP-TF will stay stable providing that initiation of hyCOUP-TF expression is maintained in 4 cells clusters, which seems to be the case.

Similar reasoning applies to the 8 cells clusters expressing hyCOUP-TF. Third the absence of variation in the fraction of 4 cells clusters expressing prdl-b could be explained the same way. However, the clearcut decrease in 8 cells clusters and the disappearance of 16 cells clusters expressing prdl-b, should be differently addressed. Either hyCOUP-TF and prdl-b are submitted to distinct regulations within a homogenous cell population, or their specific regulations reflect the regulation of two distinct sub-populations. As we found a partially overlapping cellular expression, both possibilities should be discussed. In the case of a homogenous cell population, the formation of 4 and 8 cells clusters is slowed down but not quantitatively altered upon starvation suggesting that either prdl-b expression is drastically and specifically down-regulated in the 8 cells clusters and absent in 16 cells clusters. In the case of distinct subpopulations, either as a consequence of the lengthening of the cell cycle the prdl-b expressing 4 cells clusters will differentiate earlier before reaching mitosis, thus preventing the formation of 8 and 16 cells, or 8 and 16 cells clusters upon starvation will dissociate providing transiently smaller nests (4 cells, 2 cells and single cells).

**BrdU labeling index of hyCOUP-TF and prdl-b expressing cells are different**

In order to better characterize the hyCOUP-TF and prdl-b expressing cell populations, we investigated their respective rates of proliferation by incubating hydra (Hm 105) with BrdU for either 1, or 24 or 48 hours to visualize cells that went through S phase. Animals were subsequently fixed and processed for in situ hybridization followed by anti-BrdU immunohistochemistry (Fig. 6A, 6B) and the percentage of labeled cells (BrdU labeling index) among the hyCOUP-TF and prdl-b expressing nests was counted (Fig. 6C). After 1 hour labeling, the BrdU labeling index was null for the 2 cell populations, probably because the animals were 24 hours starved. In the case of hyCOUP-TF, the BrdU labeling index reached 50% after 24 hours whatever the size of the nests and was over 80% in the 4 and 8 cells clusters after 48 hours.

Interestingly the BrdU labeling index of 2 cells clusters did not increase after 24 and 48 hours, suggesting that at least 50% of the hyCOUP-TF 2 cells clusters are either slower cycling cells or are not cycling at all, which corresponds with our assumption that these cells might represent precursors to neuronal cells. In case of prdl-b, only 35% of the 2 cells clusters were labeled after 24 hours whereas 95% of the 4 and 8 cells clusters were already BrdU labeled at that time. This result is in agreement with the heterogenous cell populations found among two cells clusters: on one side the stem cells that are fast-cycling and on the other the precursors of the neuronal and nematocyte lineage, which are cycling at a lower rate, specially when animals are starving.

In comparison 4 and 8 cells clusters expressing either hyCOUP-TF or prdl-b represent synchronous cell populations with almost identical BrdU labeling index at 24 and 48 hours. However these cell populations display significantly different BrdU labeling index at 24 and 48 hours: At 48 hours, while all prdl-b expressing...
nests went through S phase, this does not seem to be the case for all hyCOUP-TF expressing nests, as 15% to 20% of the 4 and 8 cells clusters did not. This percentage might represent the nests that escaped from proliferation and entered the differentiation process. Thus, these data indicate that these two cell populations are not identical, as a significant subset of hyCOUP-TF expressing nests are no longer cycling, which is not the case of prdl-b expressing nests.
In the sf-1 mutant, elimination of the interstitial cell lineage correlates with the disappearance of the hyCoup-TF and prdl-b expressing cells

In order to test whether expression of the hyCoup-TF and prdl-b genes was indeed linked with the differentiation of the interstitial cell lineage, we used the temperature sensitive sf-1 mutant that eliminates interstitial cells when animals are maintained above 23°C (Marcum et al., 1980; Terada et al., 1988). After 2 days at 26°C, ISH was performed and expressing nests were counted (Fig. 7). A normal hyCoup-TF and prdl-b expression pattern (Fig. 7A, 7C) was never observed in animals maintained at 26°C. Instead two distinct phenotypes were observed when compared to animals maintained at 18°C, a weak one where few expressing nests were recorded and a null one where we were unable to detect any expressing cells (Fig. 7B, 7D). At permissive temperature (Fig. 7E, 7G), nests of various sizes expressed hyCoup-TF and prdl-b while at non-permissive temperature where interstitial cells are eliminated, only few large size nests that strongly expressed hyCoup-TF and prdl-b could be detected in animals displaying a weak phenotype (Fig. 7F, 7H).

The number of animals displaying one or the other phenotype was very similar for the hyCoup-TF and prdl-b genes: around 45% for the weak phenotype and 55% for the null phenotype. When we screened for the presence of nematocytes in the expressing nests still present in the animals displaying the weak phenotype, we observed a dramatic increase in the percentage of nematocytes containing nests, over 85% in the case of prdl-b, around 50% in the case of hyCoup-TF compared to the 30% usually observed in wild type animals (Fig. 7I). It was actually shown that, while 2 cells clusters are quickly eliminated within the first 12 hours after the temperature switch, the number of nematoblasts and nematocytes will not be affected within the first two days (Terada et al., 1988). Thus our results are in agreement with this finding as we observed predominantly large nests of differentiating nematoblasts after 2 days. The difference observed between the prdl-b and hyCoup-TF expressing cells indicate that prdl-b transcripts longer than hyCoup-TF transcripts in cells clusters that entered differentiation. This might occur independently of the elimination process.

HyCoup-TF and prdl-b are no longer expressed in the nematocyte lineage during morphogenetic processes.

We recorded a similar regulation of the hyCoup-TF and prdl-b genes in regions undergoing budding, regeneration and sexual development, i.e. a clear exclusion of the hyCoup-TF and prdl-b transcripts from these regions (Fig. 8 and 9), analogous to the regulation we observed in apical and basal regions of the adult polyp. At early stages of budding (staging according to Otto and Campbell, 1977), hyCoup-TF and prdl-b were first repressed in a limited area of the parental body column, the bud spot from which the bud

**FIGURE 6:** (A, B) BrdU labeling of hyCoup-TF (A) and prdl-b (B) expressing cells after 16 hours exposure. Bar: 20 μm. (C) hyCoup-TF and prdl-b expressing cell clusters show distinct BrdU labeling index after 24 and 48 hours exposure.
FIGURE 7: *HyCoup-TF* and prdl-b expressing cells are eliminated in the sf-1 mutant hydra maintained at 26°C. (A, C) sf-1 mutant animals maintained at permissive temperature (18°C) and tested for *HyCoup-TF* (A) and prdl-b (C) expression. (B, D) SF-1 mutant animals maintained at non-permissive temperature (26°C) and tested for *HyCoup-TF* (B) and prdl-b (D) expression. (E-H) *HyCoup-TF* and prdl-b expressing nests observed in SF-1 mutant animals maintained either at permissive (E, G) or at non-permissive (F, H) temperatures for 2 days. Bars: 1 mm in A-D; 20 μm in E-H. (I) Elimination of interstitial cells leads to a dramatic increase of nematocysts-containing nests in residual prdl-b and *HyCoup-TF* expressing cells.

will emerge (**Fig. 8A, B, H**), then in the most distal part of the bud, which corresponded to the presumptive head region (**Fig. 8C-G, 8I-K**). From stage 4, we recorded a *HyCoup-TF* expression in the body column of the emerging bud at the same level as in the parental body column. From stage 8, we observed an exclusion of *HyCoup-TF* in the proximal region of the bud, which starts differentiating the basal disc (**Fig. 8G, 8K, arrowheads**). During regeneration, *HyCoup-TF* and prdl-b expressing cells immediately vanished from the healing region after amputation (**Fig. 8L and not shown**), and did not reappear at subsequent stages in apical and basal regenerating stumps (**Fig. 8M-Q**).

This observation is in agreement with the fact that cells of the nematocyte pathway rapidly disappear from regenerating tips without completing differentiation under the inhibitory influence of the *de novo* developing structure (Yaross and Bode, 1978). Cell death was actually recorded in the regenerating area during head and foot regeneration, a process which affects differentiating but not proliferating nematoblasts (Fujisawa and David, 1984).
In the remaining part of the body column, expression of **hyCoup-Tf** and **prdl-b** was not altered and the boundary between the expressing domain in the body column and the “empty” area in the regenerating stump, was rather sharp (**Fig. 8**, arrows). In addition, in the regenerating stump area, proliferation of epithelial cells is inhibited, forming an “empty” growth-arrested area during the first day of regeneration (Holstein et al., 1991). Thus this mitotic inhibition likely also affects proliferation of nematoblasts. During gametogenesis and early developmental stages, we could not detect any **hyCoup-TF** or **prdl-b** expression (**Fig. 9**): **hyCoup-TF** and **prdl-b** transcripts were detected neither in testes (**Fig. 9A-B, 9G-H**), nor in oocytes (**Fig. 9D-E, 9I-J**), nor in developing embryos (**Fig. 9C, 9F**). In female polyps, **hyCoup-TF** was also found transiently repressed in the region of the parental body column facing either the oocyte or the developing embryo. Thus, here **hyCoup-TF** and **prdl-b** display highly similar
the T19 antibody raised against the N-terminus of the huCoup-TF1 protein, a supershift was clearly observed (Fig. 10A, lanes 11, 12, 17, 18) with huCoup-chim2. Coup-TF from amphioxus (amphiCoup-TF) was used as a positive control in this experiment (Langlois et al., 2000). As expected, the complex formed by amphiCoup-TF onto DR5, was not supershifted in the presence of the T19 antibody (Fig. 10A, lanes 19-23).

In Coup-TFs transcriptional repressor activity was mapped within the E domain (Achatz et al., 1997). As this E domain is conserved in huCoup-TF, we tested whether huCoup-TF could repress the RAR:RXR transcriptional activity induced by retinoic acid (RA) in a heterologous context. We thus performed transactivation studies in mammalian cells transfected with a RARE-luc reporter plasmid, the AmphiRAR, AmphiRXR pSG5 expression plasmids (Escriva et al., 2002) and increasing amounts of pSG5-huCoup-chim2 in the presence of all-trans retinoic acid (10^{-8} M). AmphiCoup was used as a control (Fig. 10B). These assays showed that huCoup-TF significantly repressed the activation (about 2 fold) produced by RAR:RXR in the presence of all-trans RA upon the RARE and this repression was clearly dose-dependent.

**DISCUSSION**

The class II of nuclear receptors genes are well conserved from cnidarians to bilaterians

So far NR genes were detected exclusively in metazoans and our previous phylogenetic analyses had shown that the diversification of the NR superfamly occurred in two waves of gene duplication. An early one corresponded to the progressive emergence of the six classes of receptors shared by bilaterians (Laudet, 1997; Laudet and Gronemeyer, 2002) and a later one, vertebrate-specific, gave rise to the various paralogous copies of receptors such as COUP-TFI, COUP-TFII and EAR2 (Escriva et al., 1997; Escriva et al., 2000). According to concordant phylogenetic analyses (Grasso et al., 2001), the huCoup-TF gene belongs to the monophyletic COUP family. The more divergent position of the huCoup-TF gene compared to that of the coral COUP homolog possibly suggests that the two genes are not orthologous, implying that other members of the COUP-TF group remain to be discovered in cnidarians. The fact that representatives of class II (COUP, RXR) and class V (FTZ-F1) NR superfamly were isolated in cnidarians (Escriva et al., 1997; Kostouch et al., 1998; Grasso et al., 2001) supports the fact that the first wave of gene duplication occurred very early during metazoan evolution.
expressed at later stages in 8 and 16 cells clusters.

domains within the body column. It was actually differentiating nematoblasts, the precursor cells of index were characterized as dividing and morphology, organisation, location and BrdU labeling sized, round cells, which according to their morphology, organisation, location and BrdU labeling index were characterized as dividing and differentiating nematoblasts, the precursor cells of mature nematocytes. In most cells clusters prdl-b and hyCOUP-TF were found co-expressed while we recorded a clear restriction of their expression domains within the body column. It was actually demonstrated that nematocyte differentiation is inhibited in the head and basal regions (Fujiwasa et al., 1986) and other gene markers for the nematocyte lineage display similar restriction of their expression patterns (Grens et al., 1995; Gauchat et al., 1998; Engel et al., 2002; Lindgens et al., 2004; Fedders et al., 2004). Cellular comparative analyses performed onto the hyzic, CnASH and NOWA expressing cells showed that hyzic is expressed quite early in the nematocyte pathway, predominantly in 4 and 8 cells clusters, very similarly to prdl-b, while CnASH and NOWA were expressed at later stages in 8 and 16 cells clusters (Lindgens et al., 2004). The BrdU labeling index of hyzic expressing cells measured after one hour was already 40%, much higher than that measured for prdl-b, however after 24 hours hyzic and prdl-b BrdU labeling indexes reach similar values (over 90% of the cells). This difference noted after a short incubation time, beside a difference in the starvation conditions, might indicate that hyzic and prdl-b are expressed in distinct subpopulations of the 2 cells clusters. The comparative analyses we performed on the distribution of nest sizes and the presence of differentiating capsules in expressing nests, in the BrdU labeling index of starving animals and during elimination of the interstitial cell lineage indicate that hyCOUP-TF and prdl-b likely play distinct roles in the control of the nematocyte pathway. Two major differences were recorded in the regulation of these two genes: first prdl-b expression is activated at least at two distinct stages of the nematocyte pathway but significantly in 2 cells clusters, while that of hyCOUP-TF is induced one run of division later, its expression being predominant at the 8 cells stage; second, prdl-b is expressed in proliferating nests while a significant subset of hyCOUP-TF expressing cells clusters are no longer cycling, escaping from proliferation to differentiate. Consequently, hyCOUP-TF might promote entry into the differentiation process, while a possible function for prdl-b would rather be to keep the cells cycling. At differentiation stages, hyCOUP-TF expression did not seem to persist for long. However, the structural NOWa protein participates in nematocysts formation long after the NOWa gene expression is extinguished (Engel et al., 2002). Similarly we cannot exclude that the hyCOUP-TF transcription factor be still active in mature nematocytes.

In most contexts COUP-TF transcription factors behave as potent negative transcriptional regulators (Achatz et al., 1997). We could detect a significant hyCOUP-TF transcriptional repressive activity in mammalian cells. The fact that hyCOUP-TF, when expressed in mammalian cells, represses the activation mediated by other NRs suggests that this activity is conserved in all metazoans. Thus, the regulatory function of COUP-TF on other members of the superfamily observed in vertebrates or Drosophila is not the by-product of a receptor that has lost its ligand but is a fundamental ancestral activity of this gene product. Hence, interacting NRs likely played key functions in the genetic cascades that regulated differentiation early in metazoan evolution, similarly to what was observed with the bHLH family members in myogenesis (Muller et al., 2003). Although hyCOUP-TF transcriptional activity remains to be tested in hydra interstitial cells, this repressive activity might be required in nematoblasts to direct the transition from proliferation to differentiation stages. For example, a direct regulation of the hyCOUP-TF

10.1016/j.ydbio.2004.07.037
transcription factor upon the prdl-b gene can be envisaged similarly to the mouse Mix paired-like gene that displays COUP-TF binding sites in its promoter (Sahr et al., 2002).

**Highly-evolutionarily conserved regulatory genes as COUP-TF and paired-like genes trace back to the early steps of neurogenesis**

In vertebrates, COUP-TF1 disruption results in multiple defects of the central nervous system (Qiu et al., 1997). In mouse embryos, null mutation of mCOUP-TF1 induces defects in glossopharyngeal ganglion and abnormal axonal projections are observed in several regions. COUP-TF1 is involved in the segmentation of the diencephalon (Qiu et al., 1994) and is an early intrinsic factor for early regionalization of the neocortex (Zhou et al., 2001). In contrast, the COUP-TFII gene seems to be devoted to mesenchymal-epithelial interactions during organogenesis (Pereira et al., 1999). The counterpart of COUP-TF genes in Drosophila, Seven-up (svp) is implicated in photoreceptor determination (Mlodzik et al., 1990). Finally, in amphioxus, the COUP-TF gene is expressed in the nerve cord of late larvae (Langlois et al., 2000). Altogether these data show that COUP-TF related genes are implicated both in neurogenesis and/or CNS patterning during the embryonic life of a evolutionarily distant species.

Similarly, in hydra hyCOUP-TF is expressed in two lineages that define the rudimentary nervous system of cnidarians, mechanoreceptor cells and neuronal cells, and is thus likely involved in neurogenesis. Interestingly, hyCOUP-TF and prdl-b are two evolutionarily-conserved regulatory genes that are expressed in two cell lineages, which derive from a common stock of interstitial stem cells. We recently observed that the ParaHox Gsx-related cnox-2 gene (Gauchat et al., 2000) also displays this dual expression (M.M-L, unpublished). This “comparative molecular cell biology approach” implies that cells from the nematocyte and the neuronal cell lineages can be considered as “sister cell types that evolved from a common precursor by cell type diversification” (Arendt, 2003), suggesting that a yet unidentified fraction of interstitial cells are specific precursors for the nematocyte and neuronal cell lineages. We would expect that hyCOUP-TF, prdl-b and cnox-2 are markers of this common precursor.

We previously showed that the paired-like gene prdl-a is expressed in the neuronal cell lineage in hydra (Gauchat et al.,

---

**FIGURE 1B:** DNA-binding and repressor activities of the HyCOUP-TF transcription factor. (A) DNA-binding activity of the hyCOUP-chim1 (lanes 1-6), hyCOUP-chim2 (lanes 7-18) and amphiCOUP (lanes 19-23) proteins were assayed onto the DRS and DR1 response elements in gel retardation assays. Specific binding was not detected with hyCOUP-chim1, while hyCOUP-chim2 exhibited a specific binding on both DRS (lanes 7-10) and DR1 (lanes 13-16), supershifted in the presence of the T19 antiserum (star, lanes 11, 12, 17, 18). AmphiCOUP specific binding onto DRS was detected in the same experiment with a 5 times shorter exposure, but was not supershifted in the presence of T19 antiserum. (B) Transcriptional activities of the luciferase reporter construct pGL2-DRS-3x containing the DRS consensus RA-responsive element. Values are averages of three independent transfection experiments carried out in Ros 17.2/8 cells and are depicted as fold activation relative to control values obtained in the same experiment, i.e. AmphiRAR and AmphiRXR constructs expressed in the absence of all-trans-RA. The dose-dependent repression of transactivation induced by AmphiRAR/AmphiRXR after all-trans-RA treatment (10^(-8)M) on a DRS element was observed when either AmphiCOUP-TF (10 to 500 ng) or hyCOUP-TF-chim2 (100 to 1000 ng) were expressed. Error bars indicate standard deviations.
1998), the paired-like genes being mostly devoted to neurogenesis in Bilateria. Therefore we assume that neurogenesis corresponded to the ancestral cellular function of both COUP-TF and paired-like genes, similarly to what was proposed for the hyzic genes (Lindgens et al., 2004) and that neurogenesis evolved only once (Galli and Miller, 2000). According to the most parsimonious scenario (Miljkovic-Licina et al., 2004), a common ancestor to cnidarians and bilaterians would have differentiated “proto-neuronal” cells, possibly similar to mechanorceptor cells, under the control of regulatory genes like COUP-TF, paired-like and hyzic genes, whose function in neurogenesis remained conserved in most phyla.

Acknowledgements
We thank Sabina Hoffmeister, Dick Campbell, Thomas Bosch, Andreas Fröblius and Thomas Holstein for sending us various hydra strains, Chica Schaller and Hans Bode for their generous gifts of hydra cDNA libraries, Z. Kostrouch and J. Rall for the Tripedalia RXR clone, Daniel Roppolo and Yvan Rodriguez for helping us with the double in situ procedure, Volker Schmid for stimulating discussions. This work was supported by the Swiss National Foundation (FNS 31-59462.99), the Canton of Geneva, the Fonds Georges et Antoine Claraz, the Academic Society of Geneva, the Association in favor of the Cancer Research, the CNRS, the French National Education, Research and Technology Ministry and the Rhone-Alp District.

REFERENCES


The Orphan COUP-TF Nuclear Receptors Are Markers For Neurogenesis From Cnidarians To Vertebrates

Dominique GAUCHAT†, Hector ESCRIVA‡, Marijana MILJKOVIC-LICINA†, Simona CHERA†, Marie-Claire LANGLOIS§, Agnès BEGUE∥, Vincent LAUDET†‡∥ and Brigitte GALLIOT††


ONLINE SUPPLEMENT DATA LEGENDS:

Supplement 1: Sequences of type I and type III hyCOUP-TF cDNAs.
(A) In the hyCOUP-TF type I cDNA, the deduced partial C domain (47 Aas long) is underlined. In the DNA sequence, an upstream stop codon and a putative splicing acceptor site are underlined. Region upstream of the stop codon is in lower case letters. The codon where the sequences of type II, III and IV cDNAs start matching are indicated by two vertical arrows. (B) In the hyCOUP-TF type III cDNA, the deduced E domain (209 AAs long) is underlined. Note the long F domain (143 Aas long). The codon where the sequences of the type I, II and IV cDNAs start matching is indicated in bold and putative splicing acceptor sites are underlined.

Supplement 2: Structure of the hyCOUP-TF genes
(A) Scheme depicting the deduced structures of the various hyCOUP-TF cDNAs identified in Hydra vulgaris when compared to the human COUP-TF1 protein (huCOUP-TF1). In type I, the C and D domains (grey boxes) are linked to a seemingly non-coding 5′ UTR (hatched box). In type II, III and IV cDNAs, the D domain is linked to three distinct 5′ UTR sequences at an identical position, which corresponds to position 831 in type I cDNA (vertical arrows). In addition, some clones corresponding to these type II, III and IV cDNAs contain a 3 bp deletion at the 5′ extremity of the D domain depicted with a Δ sign. For each cDNA type the putative initiation (horizontal arrows) and upstream stop codons (*) are indicated. The position and orientation of the primers used in the RT-PCR experiment (C0, C1, C2, C3, C4, D, E2) are indicated. (B) Sequences of the respective 5′ UTR of the hyCOUP-TF type II, III and IV cDNAs. The different clones isolated for each type of cDNAs showed two possible points of junction with the D domain, which are indicated by vertical arrows. Putative coding sequences are uppercase and stop codons are underlined. Type I cDNA is aligned for comparison on top and domain C sequence is underlined.

Supplement 3: DNA and deduced protein sequences of the hyCOUP-TF isoforms described in Fig. 2.
Position of the primers used in the RT-PCR analysis or in the construction of the chimeric constructs are indicated. Accession numbers: hyCOUP-TF1: Y690637; hyCOUP-TF2: Y690638; hyCOUP-TF3: Y690639.

Supplement 4: Alignment of the 51 sequences used in the phylogenetic analysis corresponding to the C, D and E domains after pairwise gap removals (318 residues).
HYCoup-TF Isoforms Alignment

5′-by, 3′-hu, C1 primer

**hyCoup-TF1**

GTGCAGCAAGTTTTTCACAGACAGTCCAAAACATGCAATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF1**

GCRSSFRKRSIRRRNLAYTCTRAFQNC

**hyCoup-TF2**

.........GCATTAGACGAATTTAGCTTATACTTGCGCAGGCTTTCAAAACT

**hyCoup-TF2.21**

.................IRRRNLAYTCTRAFQNC

**hyCoup-TF3**

GCATTAGACGAATTTAGCTTATACTTGCGCAGGCTTTCAAAACT

**hyCoup-TF3.2**

.................IRRRNLAYTCTRAFQNC

**hyCoup-TF1**

GTTCATGAAATATTCACAGAAACCAATGCAATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF2**

GTTCATGAAATATTCACAGAAACCAATGCAATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF2.21**

SINDLNHRRNCQYCRLHKKWKKVGM

**hyCoup-TF3**

GTTCATGAAATATTCACAGAAACCAATGCAATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF3.2**

SINDLNHRRNCQYCRLHKKWKKVGM

**hyCoup-TF2**

GAGAAAGACGTCAAGGAGGATGCTCAATGCTAGAAGGTTCTATGTGATCAGTTTTAT

**hyCoup-TF1**

R KD TVRQKGRLP KPSLPESGSMVDQFI

**hyCoup-TF2**

GAGAAAGACGTCAAGGAGGATGCTCAATGCTAGAAGGTTCTATGTGATCAGTTTTAT

**hyCoup-TF2.21**

R KD TVRQKGRLP KPSLPESGSMVDQFI

**hyCoup-TF3**

GAGAAAGACGTCAAGGAGGATGCTCAATGCTAGAAGGTTCTATGTGATCAGTTTTAT

**hyCoup-TF3.2**

R KD TVRQKGRLP KPSLPESGSMVDQFI

**hyCoup-TF1**

GCTGGATATATATATCGTCCATATATGTGATACTTACTTACTTATTCCACCTACTTTATCG

**hyCoup-TF1**

CAAGGGTAAATATATACAGAAGGATGCTCAATGCTAGAAGGTTCTATGTGATCAGTTTTAT

**hyCoup-TF1**

S E H T T D E T

**hyCoup-TF3.2**

CAAGGGTAAATATATACAGAAGGATGCTCAATGCTAGAAGGTTCTATGTGATCAGTTTTAT

**hyCoup-TF3.2**

S E H T T D E T

**hyCoup-TF3.3**

CAAGGGTAAATATATACAGAAGGATGCTCAATGCTAGAAGGTTCTATGTGATCAGTTTTAT

**hyCoup-TF3.3**

M S A I I M D Y I R L Q I I L T F Y L

**hyCoup-TF1**

TTCAAATTGCACTATTAAAGTCTTCATGGAGTGAGCTTTATATT

**hyCoup-TF1**

GAGAAAAGACGTCAAGGAGGATGCTCAATGCTAGAAGGTTCTATGTGATCAGTTTTAT

**hyCoup-TF1**

ATGGCGAGCACAACAGAGGAAATTGCAATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF1**

L Q O Q O Q N Q Y L S Q N N L F G

**hyCoup-TF2**

cTaaagtttATGCACAAAATGGGCAACACACAGGAAATATATATACAGAAGGATGCTCAATGCTAGAAGGTTCTATGTGATCAGTTTTAT

**hyCoup-TF2.22**

H G S I Y E M A T R L LFNAVEWARNP

**hyCoup-TF2**

GGTGGGATATATATCGTCCATATATGTGATACTTACTTACTTATTCCACCTACTTTATCG

**hyCoup-TF3**

GDSIYEMATRLLFNAVEWARNP

**hyCoup-TF2**

GGTGGGATATATATCGTCCATATATGTGATACTTACTTACTTATTCCACCTACTTTATCG

**hyCoup-TF2.22**

GDSIYEMATRLLFNAVEWARNP

**hyCoup-TF3**

TTTCTTTTCAGCTTCTTCTCAATCATGCAATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF3**

 TGCTCTCTGCACCAAAATGCAATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF3**

TTTCTTTTCAGCTTCTTCTCAATCATGCAATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF3.22**

TGSALPTSDQIALKSSWSELYI

**hyCoup-TF3.3**

TTTCTTTTCAGCTTCTTCTCAATCATGCAATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF3.3**

TGCTCTCTGCACCAAAATGCAATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF3.2**

CTTTTCACACATGCAAAATGCAATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF3.3**

CTTTTCACACATGCAAAATGCAATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF3.2**

S S S S S *

**hyCoup-TF3.3**

LPTS QIALKSSWSELYI

**hyCoup-TF1**

CTCACATCTCCAAACTATCGCATTGGTATCTAAAATATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF1**

LSTS QH CIAFQINARALTSQNLS

**hyCoup-TF1**

CTCACATCTCCAAACTATCGCATTGGTATCTAAAATATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF2**

LSTS QH CIAFQINARALTSQNLS

**hyCoup-TF2**

CTCACATCTCCAAACTATCGCATTGGTATCTAAAATATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF2.22**

LSTS QH CIAFQINARALTSQNLS

**hyCoup-TF3**

CTCACATCTCCAAACTATCGCATTGGTATCTAAAATATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF3.3**

LSTS QH CIAFQINARALTSQNLS

**hyCoup-TF1**

GTAAGAAATAAAACACATCGCATTGGTATCTAAAATATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF1**

GTAAGAAATAAAACACATCGCATTGGTATCTAAAATATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF2**

GTAAGAAATAAAACACATCGCATTGGTATCTAAAATATATGCGATTGAAAAATTGTCGAAAGTTGAT

**hyCoup-TF2**

GTAAGAAATAAAACACATCGCATTGGTATCTAAAATATATGCGATTGAAAAATTGTCGAAAGTTGAT

10.1016/j.ydbio.2004.07.037
COUP-TF1 and neurogenesis in Hydra

**E I K K R S E G A N E A S V K M F E E L V E R**

dataset "chim1.2, 3tot" hy primer

**COUP-chim1.2, 3tot-hy primer**

P T T R H I K I E P P L R *
COPU-TF1 and neurogenesis in Hydra


13 RXXB_Mm DKQLTLTIWEV AKRIHPFSDL TLEDQVIIIL ARGWNELLIAS FSFHR5VSQVD G-ILLALTGLH
14 RXXR_Bg EKGEVESTIS G--ANNMDPV EQILEAELAV DKPIDT-YID AQ--------K DPVINVICQA
15 RXXR_Bf EKGEVESTTS N--FNECDFV EQIKEAEMAP EPFGDN-MVE Q--------N DPVINVICQA
16 RXXR_Aa DRNNVESTIS GGPGSDMPL ERILAEKRV EQPPBS-VLA QT--AASGR DPVNSMCQA
17 USP_Tm DRNNVESTIS N--MQAEEME RIKLAEAKRI EPCTAP-GSV GQV---------G DGVNNICQAT
18 Usp_Am ERSVESTIS S--LHDMPF ERILAEKRV ECKME----QQ ------OYNE NAVSHICNAT
19 Uspb_Ae KFSEEINITS S---VRDVTI ERINEAEQLS EQKGSDFYLR VGSNIFPEYK GAVSHLCVQW
20 Uspb_Aal KSFEENISTS S----VRDVTI ERITAAEQLS EQKGSDFYLR VGSNIFPEYK GAVSHLCVQW
21 Ultrasp RG-EDSESPS-S----VQELSI ERLIEIESLV APDPFEEQFLR VGPEPASYAKR AVFSLQCG
22 RXXR_Tc KRLSPAEVTIS EISSLATLVP DPGIQT-FAS AN--------T DPRHIVCLAA
23 HNF4a_Mm ---TRSSYYE D---SLLPSI NALLA奎ELS QOITSTPT--- ---NRAKIKI ANTIDVCESM
24 HNF4a_Ml ---TRSSYYE D---SLLPSI NVLQVQAEQ S0ITSSLT--- ---NRAKIKI AC1IDVCESM
25 HNF4a_Mm ---MRILSFAV D---SLLPSI NLQATQAMY PQI5SVPF--- ---NIYKIKI AS1IDVCESM
26 HNF4a_Ml ---MRILSFAV D---SLLPSI NLQATQAMY PQI5SVPF--- ---NIYKIKI AS1IDVCESM
27 HNF4a_Mm ---CRTSND D--DPNGDSLV GVISLKAENES RQSGKAE-- NI--NSMKF AS1IDVCESM
28 HNF4a_Ml ---CRTSND D--DPNGDSLV GVISLKAENES RQSGKAE-- NI--NSMKF AS1IDVCESM
29 HNF4a_Mm ---CRTSND D--DPNGDSLV GVISLKAENES RQSGKAE-- NI--NSMKF AS1IDVCESM
30 COUP_TF1_M Q-GQYALTIN GDPLNHSCYL GYSISLILLRA EYFPTSRYS GCMPQ---NNI MGIECIECA
31 Svpb_Ml Q-GQYALTIN GDPLNHSCYL GYSISLILLRA EYFPTSRYS GCMPQ---NNI MGIECIECA
32 COUP_TF1_X Q-GQYALTIN GDPLNHSCYL GYSISLILLRA EYFPTSRYS GCMPQ---NNI MGIECIECA
33 COUP_Bf QH-AQYALTIN GDPINGHSYL GYSISLILLRA EYFPTSRYS GCMPQ---NNI MGIECIECA
34 Svpb_Mr Q-GQFALTIN GDPLNHCSYL GYSISLILLRA EYFPTSRYS GCMPQ---NNI MGIECIECA
35 XCOUP_TFB Q-GQYALTIN GDPLNHCSYL GYSISLILLRA EYFPTSRYS GCMPQ---NNI MGIECIECA
36 Svpb_Mr Q-GQYALTIN GDPLNHCSYL GYSISLILLRA EYFPTSRYS GCMPQ---NNI MGIECIECA
37 NHR_Ae QPDPQYSTPM GTGPNHSYLS SYYISLILLRA EYFPTSRYS GCMPQ---NNI MGIECIECA
38 NHR_Aa QPDPQYSTPM GTGPNHSYLS SYYISLILLRA EYFPTSRYS GCMPQ---NNI MGIECIECA
39 Svp_Mm QPDPQYSTPM GTGPNHSYLS SYYISLILLRA EYFPTSRYS GCMPQ---NNI MGIECIECA
40 sev_Mm QPDPQYQIAT GAGFNGHSYL SYYISLLLRA EYFPTSRYS GCMPQ---NNI MGIECIECA
41 AmNR7x5 QVFQPHSAIN GVGNSHSHFL GSFISLILLRA EPFPTTFQQQ GNMMP---CCI MGIECIECA
42 EAR2_Mm 1LCPSGPQGAT GEPFTG-GPV SELIAQLLRA EYFPAARFG ---GG--GAV LGIDNVCLE
43 COUP_Hv PLEDQFIARM QNSHYNRQHI AYNSSNLSAR LQQQQQLYS Q--------NL NGIDSVIESM
44 AmNR5x4 QPEQCPQKRA SSSHNLPHQK QKQEPFLPTS SFRSSQVXL NLPTP---IST GSSMVEYIEA
45 AmNR4A QGZNSDMRQ GAGLNNHPYL SSYISLLLRA EPYPTSRFGS QCMQF---NNI MGIECIECA
46 NR2E1_Fr KQVFALYF VMNSGAAHPF SSLRGFRPPF TVQLEKYPH EGVSTYMLE VTESVCESCA
47 NR2E1b_Fr KQVFALYF VMNSGAAHPF SSLRGFRPPF TVQLEKYPH EGVSTYMLE VTESVCESCA
48 NR2E1_Mm ---KQVFALYF VMNSGAAHPF SSLRGFRPPF TVQLEKYPH EGVSTYMLE VTESVCESCA
49 Atm11x24 ---KQKMLKE AAATTAHAQ S---PSG--FI NTTLLAARFQS DLS-- ISMY SHPEAVCEAA
50 NR23_Mm ---DAMETS GVSAFPAK SGPPGTVS VSAMRVEENI DVTSN--ASP CSLDHIEHTS
51 FRAF_Bt ---DSVESET ELATTPALLG PSPRGPTPVS AARALGDENI DVTSNSSSSP CALDSIHETS

---

7

COUP-TF1 and neurogenesis in Hydra

1. RXRG_Mm: VRHSSAHASG-----VGSIF DR-----VL TELVSKMKDM QMDKSELGCL RAIVLFNPDA
2. RXRG_XL: VRHSSAHAG-----VGSIF DR-----VL TELVSKMKDM QMDKSELGCL RAIVLFNPDA
3. RXRA_Br: VRHSSAHASG-----VGSIF DR-----VL TELVSKMKDM QMDKTELGCL RAIVLFNPDA
4. RXRA_Po: VRHSSAHASG-----VGSIF DR-----VL TELVSKMKDM QMDKTELGCL RAIVLFNPDA
5. RXRA_XL: VRHNSAHASG-----VGAIF DR-----VL TELVSKMKDM QMDKTELGCL RAIVLFNPDA
6. RXRA_Mm: VRHNSAHASG-----VGAIF DR-----VL TELVSKMKDM QMDKTELGCL RAIVLFNPDA
7. RXR_Br: VRHNSAHAGAG-----VGAIF DR-----VL TELVSKMKDM QMDKTELGCL RAIVLFNPDA
8. RXRE_Br: VRHSSAHAGAG-----VGSIF DR-----VL TELVSKMKDM QMDKTELGCL RAIVLFNPDA
9. RXR_Bf: VRHSSAHAGAG-----VGAIF DR-----VL TELVSKMKDM QMDKTELGCL RAIVLFNPDA
10. RXR_Aa: VRQHSSAHAGAG-----VGAIF DR-----VL AELVAKMRDM QMDKTELGCL RAVLFNPDA
11. USP_Tm: VNKTSAAHAVG-----VGIY DR-----VL SELVNMKMKDM QMDDTELGCL RAIVLFNPDA
12. Usp_Am: VRHNSAAQAGG-----VGTIF DR-----VL SELVNMKMKDM QMDDTELGCL RAIVLFNPDA
13. Uspa_Ae: LHRNAAQAGG-----VDTLF DR-----VL CELGTMKMKDM QMVDAELGCL RAIVLFNPDA
14. Uspa_Aal: LHRNAAQAGG-----VDTLF DR-----VL CELGTMKMKDM QMVDAELGCL RAIVLFNPDA
15. Utrasp: LHRNAAQAGG-----VGTIF DR-----VL SELVNMKMKDM QMDDTELGCL RAVLFNPDA
16. RXR_Tc: LTRDNLKLAG-----VAGIY DR-----VL SEVFEKMQKI QMDRAEWGCL RAIVLFNPDA
17. HNF4A_Mm: VPHCPELA-----EMRSVL IR-----VL DELVLFPQEI QIDDNEYACL KAIIFFPDPA
18. HNF4G_Xl: IPNCPEL------EVGRVA VR-----VL DELVLFPQEI QIDDNEYACL KAIIFFPDPA
19. HNF4G_Mm: IHRNSCEY------EVSRA VR-----VL DELVLFPQEI QIDDNEYACL KAIIFFPDPA
20. HNF4b_Xl: MPPHCPEL------EIRAVP CR-----VL DELVKLRIE QIDDNYACL KAIIFFPDPA
21. HNF4b_Mm: MPPHCPEL------EIRAVP CR-----VL DELVKLRIE QIDDNYACL KAIIFFPDPA
22. CGG301_Dm: ITHCPCPDVL SPNLDSIRG AR-----II DELTVKMDV KIDDTESACI CALVFDPDPA
23. HNF4a_Ae: ITPKQPSDMKI APNLDSIRG AR-----II DELVSAIKDI KLDDSELACI CALVFDPDPA
24. COUP_TF1_M: ---------------- ASHPMA DRFMHDHIRIF EQQVEKLKAL HDVSAEYCSL KAIrafted
25. Svp4_Br: ---------------- ASHPMA DRFMHDHIRIF EQQVEKLKAL HDVSAEYCSL KAIrafted
26. COUP_TF1_X: ---------------- ASHPMA DRFMHDHIRIF EQQVEKLKAL HDVSAEYCSL KAIrafted
27. Svp4_Bf: ---------------- ASHPMA DRFMHDHIRIF EQQVEKLKAL HDVSAEYCSL KAIrafted
28. Svp4_Gr: ---------------- ASHPMA DRFMHDHIRIF EQQVEKLKAL HDVSAEYCSL KAIrafted
29. COUP_TF1B: ---------------- ASHPMA DRFMHDHIRIF EQQVEKLKAL HDVSAEYCSL KAIrafted
30. COUP_TF1F: ---------------- ASHPMA DRFMHDHIRIF EQQVEKLKAL HDVSAEYCSL KAIrafted
31. COUP_TF2F: ---------------- ASHPMA DRFMHDHIRIF EQQVEKLKAL HDVSAEYCSL KAIrafted
32. EAR2_Mm: ---------------- AAPPMA ERFMDQVRAF EQQVKGLGRL QVDAAEYCL KAIrafted

COPU-TF1 and neurogenesis in Hydra

1. RXRG_Mm  KGLSNPSEVE TLREKYATL EAYTQKYPE QGRFkALL LLPALRISIGL KCEILHFFFFK
2. RXRG_XL  KGLSNAEVE ALREKYATL ESYTKQ1PD QGRFkALL LLPALRISIGL KCEILHFFFFK
3. RXRA_Br  KGLSNPSEVE ALREKYAYL ESYTKQKYPD QGRFkALL LLPALRISIGL KCEILHFFFFK
4. RXRA_Po  KGLSNEVEVE ALREKYAYL ESYTKQKYPD QGRFkALL LLPALRISIGL KCEILHFFFFK
5. RXRA_XL  KGLSNPVEVE ALREKYAYL ESYTKQKYPD QGRFkALL LLPALRISIGL KCEILHFFFFK
6. RXRA_Mm  KGLSNPVEVE ALREKYAYL ESYTKQKYPD QGRFkALL LLPALRISIGL KCEILHFFFFK
7. RXRB_Br  KGLSNEVEVE ALREKYAYL ESYTKQKYPD QGRFkALL LLPALRISIGL KCEILHFFFFK
8. RXRB_O1  KGLSNEVEVE ALREKYAYL ESYTKQKYPD QGRFkALL LLPALRISIGL KCEILHFFFFK
9. RXRB_Br  KGLSNEVEVE ALREKYAYL ESYTKQKYPD QGRFkALL LLPALRISIGL KCEILHFFFFK
10. RXRG_Po KGLSNEVEVE ALREKYAYL ESYTKQKYPD QGRFkALL LLPALRISIGL KCEILHFFFFK
11. RXRB_XL  KGLSNPVEVE ALREKYAYL ESYTKQKYPD QGRFkALL LLPALRISIGL KCEILHFFFFK
12. RXRB_Mm  KGLSNPVEVE ALREKYAYL ESYTKQKYPD QGRFkALL LLPALRISIGL KCEILHFFFFK
13. RXRB_Br  KGLSNEVEVE ALREKYAYL ESYTKQKYPD QGRFkALL LLPALRISIGL KCEILHFFFFK
14. RXX_Bg  KGLTAVQVEE QRLEKYAYL EYETSKRFE EPRFkALL LLPALRISIGL KCEILHFFFFK
15. RXK_Br  KGLTNPSLVE ALREKYAYL ESYTCKQKYPD QGRFkALL LLPALRISIGL KCEILHFFFFK
16. RXA_Aa  KGLPNATVEE ALREKYAYL EYECRHHPHD QGRFkALL LLPALRISIGL KCEILHFFFFK
17. Usp_Tn  KGKSVQVEE MLREKYGVL EYETPTITHP EPRFkALL LLPALRISIGL KCSALEHFFFFK
18. Usp_Am  KGKLSQVEV TLREKYGVL EAYCRVAPD DGRFkALL LLPALRISIGL KCELALHFFFFK
19. Uspa_Ae  KGKQCXVEID GMREKYIACI DEHKQHQS EDGRFQQL LLPALRISIGL KCELHNFIR
20. Usp_Aal  KGKLSQNGGD GMREKYIACI DEHKQHQS EDGRFQQL LLPALRISIGL KCELHNFIR
21. Ultrasp KGKLNKPEV TLREKNSFIC DERYVSNSCA EEPRAASLL LLPALRHSGK KCEHFLLFFFFH
22. RXA_Tc  KGKTADQVEQ NYRELTSSTL EDHVKRKHPE QPERFTEKIL RLPALRSIGL KCEILHFFFFK
23. HNF4A_Mm KGKLDTPKXK RMYQVQVSL EDYINQYRD YRFGDELL LLPTQSITW QMEIQVFK
24. HNF4G_Xl KGKLDPTKXK RMYQVQVSL EDYINQYRD YRFGDELL LLPTQSITW QMEIQVFK
25. HNF4G_Mm KGKLDVQKXK NRMPQVQISL EDYINQYRD YRFGDELL LLPTQSITW QMEIQVFK
26. HNF4b_Xl KGKLDQVTKR SRMDFQVQISL EDYINQYRD YRFGDELL LLPTQSITW QMEIQVFK
27. HNF4_Dm KGKNEFPRH LSRQHICLNL EDYSISQYRE GFRSFEILL LLPSILQSVT QMEIQVFK
28. C9310_Dm KGKNEFPRH LSRQHIKNL EDYSISQYRE GFRSFEILL LPYVILQSVT QMEIQVFK
29. HNF4A_He KGKLDTPKXK RMYQVQVSL EDYINQYRD YRFGDELL LLPTQSITW QMEIQVFK
30. COUP_TF1_M KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
31. Svp44_Br KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
32. COUP_TF1_X KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
33. COUP_Bf  KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
34. Svp40_Br KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
35. XCOUP_TFB KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
36. Svp46_Br KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
37. COUP_TF1_XL KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
38. NHR_Ae  KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
39. Svp_Bm  KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
40. sev1_Dm  KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
41. AmnNRx5 KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
42. EAR2_Mm  KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
43. AmnNRx4 KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
44. AmnNR4A KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
45. NR2E1_Fr KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
46. NR2E1b_Fr KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
47. NR2E1_Mm KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
48. Am11x24 KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
49. NR2E2_Mm KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
50. NR2E3_Mm KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR
51. PRAP_Bt  KGCLSAHHSI SLQEKSGCAL EYESQVYNQ QPSFQKLL LRLPSLRTVS SVTEQLFVFR

10.1016/j.ydbio.2004.07.037 9
COUP-TF1 and neurogenesis in Hydra

Gauchat et al. (2004) *Dev Biol* 275, 104-123

RXRG_Mm  LIGDTPIDSF LMEMLETP
RXRG_XL  LIGDTPIDSF LMEMLETP
RXRA_Br  LIGDTPIDSF LMEMLEAP
RXRA_Po  LIGDTPIDSF LMEMLEAP
RXRA_XL  LIGDTPIDSF LMEMLEAP
RXRA_Mm  LIGDTPIDSF LMEMLEAP
RXR_Br  LIGDTPIDSF LMEMLEAP
RXRE_Br  LIGDTPIDSF LMEMLEAP
RXRD_Ol  LIGDTPIDSF LMEMLEAP
RXRD_Br  LIGDTPIDSF LMEMLEAP
RXRB_Xl  LIGDTPIDSF LMEMLEAP
RXRB_Mm  LIGDTPIDSF LMEMLEAP
RXR_Bg  LIGDTPIDSF LMEMLEAP
USP_Tm  LIGDTPIDSF LMEMLEAP
USP_Am  LIGDTPIDSF LMEMLEAP
Uspa_Ae  LLSKHLDSF YVEML---
Usp_Aal  LLSKHLDSF YVEML---
Ultrasp  LVADTSIASY IHDLRNN
RXR_Tc  LIGDVPMDDF LLDMLEVD
HNF4a_Mm  LGMAGKDNL LQBMMLGG
HNF4g_Xl  LGMAMGKDNL LQBMMLGG
HNF4g_Mm  LGMAMGKDNL LQBMMLGG
HNF4b_Xl  LGVARGDNL LQELMLGG
HNF4_Dm  LGVAIHDNL LQBMMLGG
CG9310_Dm  LGVAIHDSL LQBMMLGG
HNF4a_Ae  MGFVHDSL LQBMMLGG
Coup_TF1_M  LVGKTPIETL IREDMLSG
Svp40_Br  LVGKTPIETL IREDMLSG
Coup_TF1_X  LVGKTPIETL IREDMLSG
Coup_Bf  LVGKTPIETL IREDMLSG
Svp40_Br  LVGKTPIETL IREDMLSG
XCOUP_TFB  LVGKTPIETL IREDMLSG
Svp46_br  LVGKTPIETL IREDMLSG
Coup_TF_Xl  LVGKTPIETL IREDMLSG
NH2_Ae  LVGKTPIETL IREDMLSG
Svp_Bm  LVGKTPIETL IREDMLSG
Sev1_Dm  LVGKTPIETL IREDMLSG
AmNrr7x5  LVGKTPIETL IREDMLSG
Ear2_Mm  LVGKTPIETL IREDMLSG
Coup_Hv  MNSNIGYLS LSAMMLGG
AmNR8x4  MNSIISSS
AmNR4A  MNSIISSS
NR2E1_Fr  TIGNVPITRL LSYMKSS
NR2E1b_Fr  TIGNVPITRL LSYMKSS
NR2E1_Mm  TIGNVPITRL LSYMKSS
Am11x24  TIGNVPITRL LSYMKNND
NR23_Mm  TIGNTPMEKL LCMFPKN--
PRAP_Bt  TIGNTPMEKL LCMFPKN--