

Early evolution of metazoan transcription factors

Bernard M Degnan¹, Michel Vervoort², Claire Larroux¹ and Gemma S Richards¹

Analyses of recently sequenced sponge, cnidarian, placozoan, and choanoflagellate genomes have revealed that most transcription factor (TF) classes and families expressed during bilaterian development originated at the dawn of the animal kingdom, before the divergence of contemporary animal lineages. The ancestral metazoan genome included members of the bHLH, Mef2, Fox, Sox, T-box, ETS, nuclear receptor, Rel/NF- κ B, bZIP, and Smad families, and a diversity of homeobox-containing classes, including ANTP, Prd-like, Pax, POU, LIM-HD, Six, and TALE. As many of these TF classes and families appear to be metazoan specific and not present in choanoflagellates, fungi and more distant eukaryotes, their genesis and expansion may have contributed to the evolution of animal multicellularity.

Addresses

¹ School of Biological Sciences, University of Queensland, Brisbane, QLD 4072, Australia

² Development and Neurobiology Program, Institut Jacques Monod - UMR7592 CNRS/Université Paris Diderot, 15 rue Hélène Brion, 75205 Paris Cedex, France

Corresponding author: Degnan, Bernard M (b.degnan@uq.edu.au)

Current Opinion in Genetics & Development 2009, 19:591–599

This review comes from a themed issue on
Genomes and evolution
Edited by Martha Bulyk and Michael Levine

Available online 31st October 2009

0959-437X/\$ – see front matter

© 2009 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.gde.2009.09.008

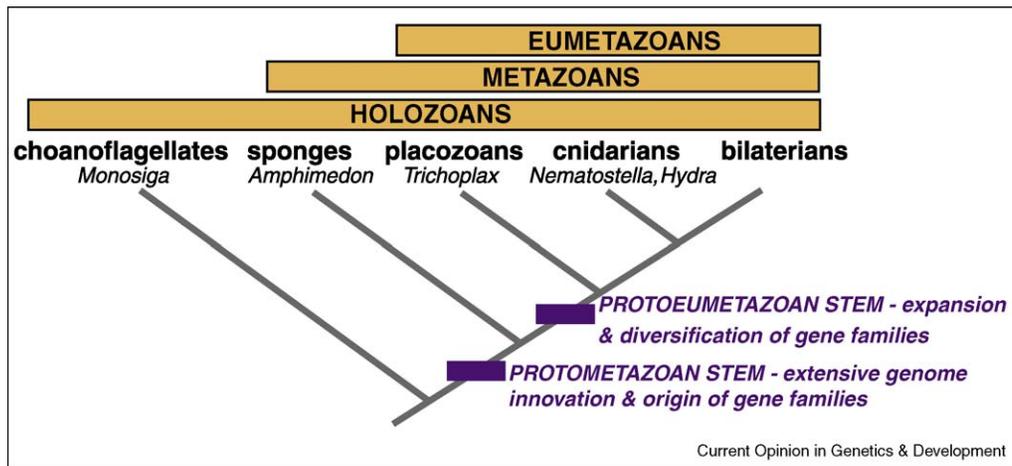
Introduction

Animal body plans vary significantly in their architecture and functional capabilities, with many phylum-level innovations lacking clear counterparts in other phyla. This disparity is particularly pronounced when comparing the markedly different body plans of the earliest known phyletic lineages — poriferans (sponges), ctenophores (comb jellies), cnidarians (anemones and jellyfish), and placozoans (Figure 1) [1,2]. The phylogenetic positions of these four metazoan phyla have not yet been completely resolved [3[•],4[•],5[•]], but there seems no doubt that all of them are basal to all other animals. Nonetheless, these comparatively simple animals share much with their more complex protostome and deuterostome cousins. Nowhere is this more clearly seen than in whole genome sequence comparisons among the anthozoan

cnidarian *Nematostella vectensis* (<http://genome.jgi-psf.org/Nemve1/Nemve1.home.html>) [6^{••}], the hydrozoan cnidarian *Hydra magnipapillata* (<http://hydrasome.metazome.net/cgi-bin/gbrowse/hydra/>; publication pending), the placozoan *Trichoplax adhaerens* (<http://genome.jgi-psf.org/Triad1/Triad1.home.html>) [7^{••}], the demosponge *Amphimedon queenslandica* (genome release and publication pending) [8], and the expanding list of bilaterian genomes (e.g. [9,10]). The genome organization and content of these early-branching phyla is remarkably similar to that of morphologically complex bilaterians, with the basic genomic underpinnings for key developmental, physiological, and immunological characteristics appearing before the divergence of any of these lineages (Box 1).

Comparison of these metazoan genomes with a representative of the closest living nonmetazoan lineage for which there is whole genome sequence — the choanoflagellate *Monosiga brevicollis* [11^{••}] — reveals that a vast number of these genomic features are unique to animals, evolving on the protometazoan stem after the divergence of choanoflagellate and metazoan lineages (Figure 1). Included among these animal-specific innovations are genes responsible for generating the metazoan multicellular condition through the process of development. A large majority of the signaling pathway components and transcription factors (TFs) that lie at the heart of bilaterian developmental gene regulatory networks (GRNs) also exist in sponges, placozoans, and cnidarians [6^{••},7^{••},12]. Indeed the expression of Wnt, TGF- β , Notch, and Hedgehog-like pathways in cnidarians and sponges provides compelling evidence that many of the developmental mechanisms for generating cell diversity and axial patterning during bilaterian development were already functioning in the last common ancestor to all living metazoans (LCAM) [12–14,15[•],16–20,21[•]]. This remarkable conservation of ligands, receptors, signal transducers, and TFs between lineages that have been evolving independently for over 600 million years reveals a deep and yet to be fully appreciated selective pressure acting on the evolution of these genes. The interaction between a TF and its target *cis*-regulatory elements is central to GRN outputs, and ultimately to the expression of cell lineage-specific and territory-specific differentiation products (reviewed in [22,23,24^{••}]). Thus, understanding the origin of the metazoan TF ‘toolkit’ is critical to deciphering the ancestral genomic architecture that underpinned the evolution of animal multicellularity.

Figure 1



Phylogeny of holozoans. Only holozoan and early-branching metazoan phyla for which there is a representative genome are depicted; genera listed. Tree topology is a consensus based on analyses of orthologous genes from complete animal genomes [6^{••},7^{••}]. Protometazoan and protoeumetazoan stems are highlighted as periods of extensive genome evolution.

The regulatory zootype: a conservative estimate of transcription factor diversity in the metazoan ancestor

The phylogenetic relationship of the four early-branching metazoan phyla — Porifera, Ctenophora, Cnidaria, and Placozoa — is contentious [3[•],4[•],5[•]], although analyses of large nuclear gene sets consistently have Porifera, Placozoa, and Cnidaria branching sequentially from the stem that will give rise to the crown bilaterians, as depicted in Figure 1 [5[•],6^{••},7^{••}]. Given the uncertainty of the phylogenetic position of ctenophores (cf. [3[•],4[•],5[•]]) and the

lack of whole genome data from this phylum, they are not included in Figure 1 and discussed minimally here. Regardless of the true order of divergence, a minimal ancestral set of TFs can be inferred by comparing the TF repertoires in *Amphimedon*, *Trichoplax*, *Nematostella*, and *Hydra*. Comparison of this ancestral metazoan TF suite with the TFs in the choanoflagellate *Monosiga* [11^{••}] and other closely related opisthokont protist representatives [25,26,27[•],28[•]] uncovers the genomic innovations that occurred on the protometazoan stem before metazoan cladogenesis.

Box 1 Taxonomic terms

Opisthokonta — Cristidiscoidea (Nucleariida) plus Fungi, plus Ichthyosporea with Filasterea, corallochytrium and Choanoflagellata plus Metazoa. Cells possess a posterior cilium, flat cristae, narrow pseudopods, and possibly long thread-like (filose) projections.

Holozoa — Ichthyosporea plus Filasterea and Choanoflagellata plus Metazoa. Opisthokonts that possess a fused ubiquitin and ribosomal small subunit s30 protein. Filastereans, corallochytrium, choanoflagellates, and metazoan cells appear to share filose tentacles. Choanoflagellates and metazoans cells share a collar of aggregated microvilli tentacles that surrounds the cilium.

Metazoa — The animal kingdom; Porifera plus Eumetazoa. Multicellular heterotrophic holozoans with oogamy, and integrated (epithelial-like) and patterned layers and loose aggregates of cells interacting with an extracellular matrix.

Eumetazoa — Bilateria plus Cnidaria, Ctenophora (and Placozoa). Metazoans with integrated tissues, muscle, and nerve cells, and a digestive cavity.

Bilateria — Acoelomorpha plus Deuterostoma and Protostoma. Eumetazoans with bilateral symmetry and a nervous system with anterior condensation of interneurons.

(see [3[•],4[•],5[•],6^{••},7^{••},11^{••},26,27[•]] for more detailed descriptions)

Representatives of a wide range of classes and families of developmental TFs were present in the LCAM, as inferred by the analysis of sponge, placozoan, and cnidarian genomes. Basic helix–loop–helix (bHLH), Mef2, Fox, Sox, T-box, ETS, nuclear receptor, Rel/NF-κB, bZIP, Smad, and a diversity of homeobox-containing classes, including ANTP, Prd-like, Pax, POU, LIM-HD, Six, and TALE (Table 1), evolved before the divergence of sponges and eumetazoans (Figure 1) [7^{••},21[•],29–33,34[•],35^{••},36,37,38^{••},39–41]. Ctenophores also possess representatives of many of these TF families and classes [42–45,46[•]]; however, a comprehensive inventory of their TFs awaits whole genome sequence data.

Originally applied to the conserved expression and structure of Hox genes and other developmental regulators [47], the zootype concept can further be applied to derived aspects of genomes that are unique to animals, including TFs, signaling ligands, and other factors known to play critical roles in the regulation of development. Although some gene classes have a more ancient origin (e.g. Fox, bZIP, Rel/NK-κB, TALE, and typical (non-TALE) homeoboxes, and T-box [11^{••},21[•],28[•],35^{••},38^{••},48[•],49]), crown

Table 1

Metazoan TF gene families

Family	Expanded name	Representative genes
Homeobox		
ANTP	Antennapedia (<i>D. melanogaster</i>)	<i>hox, nk, tlx, msx, bar, dlx, emx, not</i>
LIM-HD	Lin11 (<i>C. elegans</i>), Isl-1 (<i>M. musculus</i>), Mec-3 (<i>C. elegans</i>) domain + homeodomain	<i>lim, lhx, lmx, isl, apt</i>
Prd-like	Paired-like homeobox	<i>rx, otx, arix, pitx, otp, gsc</i>
Pax	Paired domain + paired-like homeobox	<i>pax, pox, gsb</i>
POU	Pit (<i>M. musculus</i>), Oct (<i>M. musculus</i>), Unc (<i>C. elegans</i>) domain	<i>pou, pit, oct</i>
Six	Sine oculis (<i>D. melanogaster</i>) related homeobox	<i>six, so, arec</i>
TALE	Three amino acid loop extension (of the homeodomain)	<i>pbx, meis, tgif, iro</i>
bHLH	Basic helix–loop–helix	<i>achaete, scute, myoD, neuroD, ato, twist, max, myc, e12/e47</i>
bZIP	Basic leucine zipper	<i>fos, jun, atf, par, c/ebp, maf, nfe2, oasis, creb</i>
ETS	E26 transformation-specific	<i>ets, dets4, elf, tel, spi, gabp, erg, tcf, pea3, erf, pointed</i>
Fox	Forkhead	<i>fox, fkh</i>
Mef2	Myocyte enhancer factor 2/MADS box transcription enhancer factor	<i>mef2</i>
NR	Nuclear receptor	<i>hnf4, rxr, rar, ppar, ror, tlr, err, knirps</i>
Rel/NF- κ B	Rel homology domain/nuclear factor kappa-light-chain-enhancer of activated B cells	<i>NF-κB, relish, rel, dorsal</i>
Smad	Sma (<i>C. elegans</i>), Mothers against Decapentaplegic (<i>D. melanogaster</i>)	<i>smad, mad</i>
Sox	Sry-related HMG (high mobility group) box	<i>sox, sry</i>
T-box	–	<i>bra, tbr, tbx</i>

metazoans can be defined by numerous TF family synapomorphies (shared derived traits), including NK, Prd-like, Pax, POU, LIM-HD, and Six homeobox genes, groups A and C–F bHLHs, ETS, Runx, Mef2, Sox, and nuclear receptor families. Members of many of these gene families and classes are positioned within known bilaterian developmental GRNs, where they often have conserved roles and localized developmental expression patterns. Although the role and expression patterns of members of these gene families in sponges, placozoans, and cnidarians are largely unknown, their very presence in basal metazoan genomes lends support to the proposition that the origination of these gene families was a prerequisite for the evolution of metazoan multicellularity.

TFs associated with metazoan-specific signaling pathways also appear to have originated in the protometazoan stem and thus are part of the metazoan genomic zootype (reviewed in [12]). In the Notch pathway, the CBF1/Su(H)/Lag1 (CSL) DNA binding protein, along with a number of other nuclear cofactors including core binding factor β (CBF β) and CREB binding protein (CBP), appear to be pan-metazoan. Likewise, TFs and coregulators of the Wnt pathway, including β -catenin, T-cell specific transcription factor/lymphoid enhancer-binding factor 1 (Tcf/Lef), and the corepressor Groucho, are metazoan innovations. TGF- β signaling leads to the nuclear accumulation of Smads, which act with TFs Fos and Jun (bZIP family) and Myc (bHLH family), and transcriptional coactivators, CBP and CBF β . On the basis of *Amphimedon*, *Trichoplax*, and *Monosiga* genomes, these factors also originated after the divergence of metazoan and choanoflagellate lineages. Although the

canonical Hedgehog pathway may have first become operational along the cnidarian + bilaterian stem [12,18,20], associated TFs belonging to the Gli/Ci group of zinc finger binding proteins are found throughout the animal kingdom.

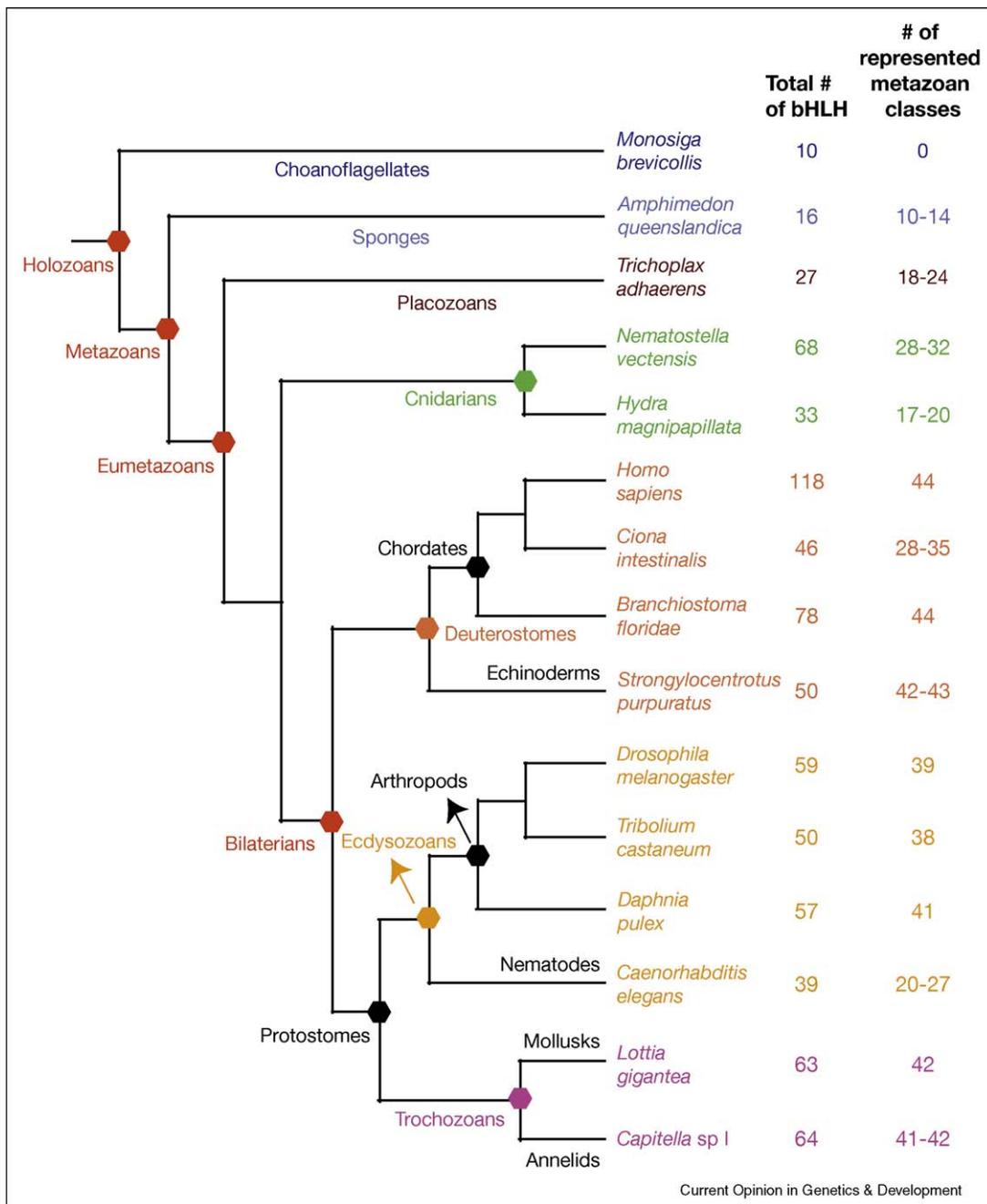
Genesis of metazoan transcription factor gene families

The genesis of metazoan-specific TFs can be classed into three groups. Using the parlance of Putnam *et al.* [6^{••}], these are type I, II, and III novelties. Animal TF genes that have no clear relatives outside the Metazoa are considered a type I novelty and currently include ETS and nuclear receptor families (the current sequencing of opisthokont protist genomes may reveal these to have a more ancient origin, e.g., the recent detection of T-box and Rel/NF- κ B outside the Metazoa [28[•]]). While there are a wide range of TF gene families and classes that are specific to animals, often a distant ancestral member can be found that predates the divergence of metazoan and choanoflagellate clades. For example, while the ANTP homeobox class appears to be a metazoan innovation, its origin can be traced to a more ancient typical (non-TALE) eukaryotic homeobox [38^{••},48[•]]. POU, Pax, and Six homeobox classes all can be classified as type II novelties, where the animal restricted POU, Six, and Pax domains are combined with the more ancient homeodomain to produce the metazoan novelty. In contrast, type III novelties are those where ancient premetazoan domains combine in novel ways to generate a metazoan-specific domain architecture; an example is the animal-specific way in which the ancient LIM and homeodomain combine in the LIM-HD class [38^{••}].

The subfamilies of the metazoan bHLH TF superfamily are grouped into six higher order clades (named A–F), the genesis of which also features type II and III novelties. This diversity of bHLH families is unique to animals, but its ancestry can be traced to a more ancient group B bHLH shared with plants and fungi [35^{••}]. Among metazoans, most group E proteins contain a metazoan-specific orange domain in addition to the bHLH, thus

they are a type II novelty. Group C bHLH proteins appear to be a type III innovation, containing the ancient PAS domain in addition to the bHLH. Group F proteins (COE family) lack a basic domain and are characterized by the presence of the COE domain which is involved in both dimerization and DNA binding. As the COE domain is likely to have evolved from a more ancient IPT (TIG) domain, group F can also be considered a type III novelty.

Figure 2



bHLH gene family representation in holozoans. The total number of bHLH genes and the number of bHLH gene classes, as defined by the bilaterian gene set, are listed to the right of each species for which there is an assembled draft genome.

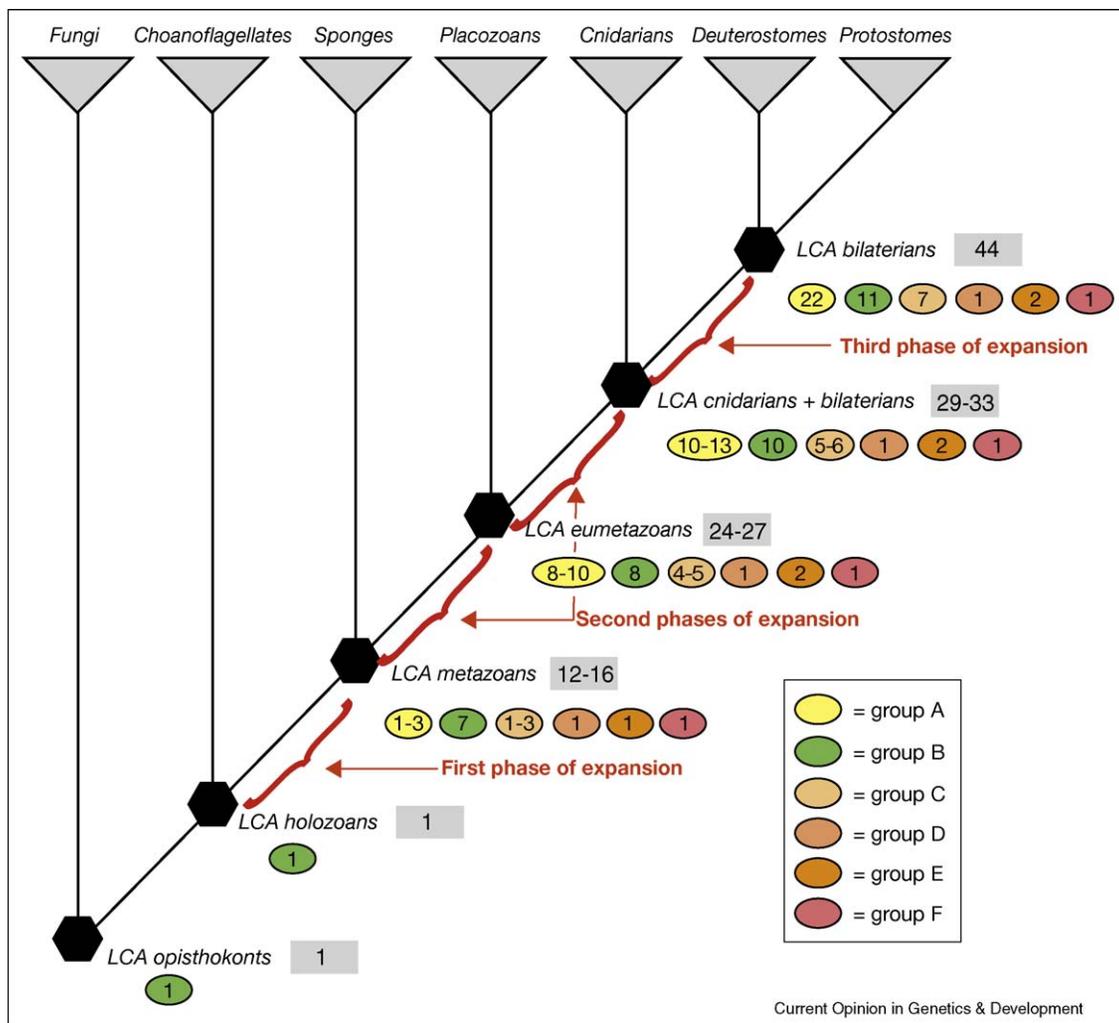
Transcription factor gene family expansion in early metazoans

The cnidarian *Nematostella* has a TF gene content that is remarkably similar to bilaterians, being in general two to three times larger than is observed in *Amphimedon* or *Trichoplax* [6^{**},15^{*},34^{*},35^{**},38^{**},50–52,53^{**},54^{**},55–60,61^{*}]. Given that cnidarians are classified consistently as the sister group to the bilaterians [3^{*},4^{*},5^{*},6^{**},7^{**}], it has been proposed that these differences are the result of gene family expansions on the stem leading to the last common ancestor of cnidarians and bilaterians (Figure 1). However, it is also likely that gene loss has occurred in both demosponge and placozoan lineages, as previously discussed [6^{**},7^{**},35^{**},38^{**},62,63]. Inferences of gene

duplication and loss during early metazoan evolution are often based on the most parsimonious explanation of a phylogenetic analysis. The reconstruction of the evolutionary history of a given gene family or class is possible through the integration of gene and organismal phylogenies. Although our knowledge of early-branching metazoan genomes is restricted to a handful of species, we can nonetheless use gene phylogenies to identify orthologs in distant taxa and uncover the minimal number of gene families and classes in a particular group of animals.

As a case study, we applied this approach to bHLH evolution in animals [35^{**}], revising our previous analysis of this family by including new data. The bHLH inventory

Figure 3



Reconstruction of the evolution of the bHLH groups in the opisthokonts. The first phase of expansion occurred during the protometazoan stem. This includes the origination of bHLH groups A and C–F. Gene duplication and divergence within some of these groups has occurred before metazoan cladogenesis. Along the protoeumetazoan stem (two steps shown) there is an increase in gene number within groups A–C and D; groups A–C further expand on the protobilaterian stem. ‘LCA’ stands for last common ancestor. The numbers in gray boxes refer to the estimated number of bHLH genes in each LCA; numbers in colored ovals refer to the number of bHLH genes in each subgroup. Estimates of bHLH content in the various LCAs are based on phylogenetic inferences (see [35^{**},38^{**}] for details) and are larger than or equal to the estimates for taxa used to reconstruct the LCA (cf. Figure 2). A range in the estimate of bHLHs in some LCAs reflects uncertainty in assigning orthology to particular bHLH members in specific genomes.

in species from all the major phyletic and superphyletic lineages of animals identifies the stem in which a given family is likely to have originated (Figures 2 and 3). From this analysis we demonstrate general trends in TF gene family and class evolution [38**] that include: first, the origin of all the major metazoan-specific classes, often from a more ancestral prototype (as is the case in the bHLHs and homeoboxes), and second, an initial expansion phase before the divergence of sponge and eumetazoan lineages; third, a second period of expansion early in eumetazoan evolution, resulting often in a twofold to fourfold increase in gene number; and four, a third and typically less extensive phase of expansion in the bilaterian stem (Figure 3). Tracking the expansion of gene families in this way stimulates hypotheses and debates about the origin of particular TF subfamilies and classes of functional significance, for example, bHLH group A neurogenic factors [21*,64**] and the Hox and related homeobox genes [15*,34*,39,50,53**,54**,61*62,65*,66*,67*,68*].

Interestingly, the increase in the size of TF classes and families in eumetazoans correlates with morphological and cell type complexity, suggesting that an increase in TF gene number may have been an underpinning factor in expanding developmental potential and yielding new cell types. There are cases where the absences of specific TFs in *Amphimedon* correlate with the lack of a specific morphological complexity (e.g. sponges lack a true mesoderm and *Amphimedon* does not have TFs involved in bilaterian mesoderm development — Fkh, Gsc, Twist, Snail). In contrast, TFs involved in the determination or differentiation of eumetazoan-specific nerve and muscle cells are nonetheless present in *Amphimedon* (e.g. PaxB, Lhx, SoxB, Msx, Mef2, and group A bHLH neurogenic factors) [35**,38**,64**], suggesting related cell types may exist in sponges. For example, *Amphimedon* has a putative sensory cell type that appears to have affinity to eumetazoan neurons [69], and that expresses an Atonal-like bHLH gene that promotes neurogenesis in *Drasophila* and *Xenopus* [21*].

Conclusions

Comparative analyses of recently sequenced sponge, placozoan, and cnidarian genomes have revealed a rich and diverse TF gene repertoire. It can now be inferred that the large majority of TFs that control bilaterian development have their evolutionary origin at the base of the animal tree. This raises questions such as how this multifaceted TF resource was employed in the first metazoans, and how complex was the LCAM.

Given that the current representation of fully sequenced basal metazoan genomes is skeletal, reconstructions of last common metazoan and eumetazoan ancestors are likely to be conservative, taking minimal account of potential gene loss events in extant lineages. Phylogenetic analyses of TF families can help to identify poten-

tial points of gene loss [35**,38**,62], but the differential evolution of gene duplicates confounds inferences about the ancestral condition. As a consequence, it is only through the sequencing of additional sponge, ctenophore, cnidarian, and nonmetazoan opisthokont genomes that we can expand our understanding of the ancestral eumetazoan, metazoan, holozoan, and opisthokont TF inventories. Another future challenge will be to establish tools to study the functions of these TFs and the GRNs in early-branching metazoan phyla, for example, through the experimental reallocation of gene expression as recently demonstrated in cnidarians (e.g. [70*,71*,72*]). Comparative genomics, combined with gene knock-down studies and system-level network modeling [22,23], has the potential to reveal deeply conserved TFs at the core of the first metazoan GRNs. These approaches, along with defining *cis*-regulatory complexity in the genomes of early-branching metazoans relative to each other and bilaterians [73], may divulge how the evolution of regulatory architectures promote and constrain morphogenetic complexity and diversity.

Acknowledgements

We thank Sandie Degnan for critical comments on the manuscript. Research in the laboratories of BMD and MV is supported by the Australian Research Council and the Agence Nationale de la Recherche française (ANR), respectively.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
 - of outstanding interest
1. Valentine JW: *On the Origin of Phyla*. Chicago: The University of Chicago Press; 2004.
 2. Minelli A: *Perspectives in Animal Phylogeny & Evolution*. Oxford: Oxford University Press; 2009.
 3. Dunn CW, Hejnal A, Matus DQ, Pang K, Browne WE, Smith SA, Seaver E, Rouse GW, Obst M, Edgecombe GD *et al.*: **Broad phylogenomic sampling improves resolution of the animal tree of life**. *Nature* 2008, **452**:745-749.
See annotation to Ref. [5*].
 4. Schierwater B, Eitel M, Jakob W, Osigus HJ, Hadrys H, Dellaporta SL, Kolokotronis SO, Desalle R: **Concatenated analysis sheds light on early metazoan evolution and fuels a modern "urmetazoan" hypothesis**. *PLoS Biol* 2009, **7**:e20.
See annotation to Ref. [5*].
 5. Philippe H, Derelle R, Lopez P, Pick K, Borchellini C, Boury-Esnault N, Vacelet J, Renard E, Houliston E, Quéinnec E *et al.*: **Phylogenomics revives traditional views on deep animal relationships**. *Curr Biol* 2009, **19**:706-712.
Along with Refs. [3*,4*], this work reports metazoan phylogenies based on phylogenomic approaches with large data sets, but provide different early-branching metazoan phylogenies. Philippe *et al.* use the broadest sampling of basal metazoans and recover a monophyletic Porifera as the most basal extant metazoan lineage.
 6. Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, Salamov A, Terry A, Shapiro H, Lindquist E, Kapitonov VV *et al.*: **Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization**. *Science* 2007, **317**:86-94.
Description of the anthozoan cnidarian *Nematostella vectensis* genome reveals that the cnidarian plus bilaterian ancestor possessed a strikingly complex genome, showing remarkable similarity in organization and

content to modern bilaterians. The origin of the 'pan-eumetazoan toolkit' is discussed in relation to the type of metazoan novelty (I, II, or III).

7. Srivastava M, Begovic E, Chapman J, Putnam NH, Hellsten U, Kawashima T, Kuo A, Mitros T, Salamov A, Carpenter ML *et al.*: **The *Trichoplax* genome and the nature of placozoans.** *Nature* 2008, **454**:955-960.
- The genome of one of the simplest free-living metazoans, the placozoan *Trichoplax adhaerens*. This genome encodes a large number of transcription factors despite the simplicity of this animal. From comparative genomics it is inferred that placozoans are eumetazoans.
8. Degnan BM, Adamska M, Craigie A, Degnan SM, Fahey B, Gauthier M, Hooper JNA, Larroux C, Leys SP, Lovas E *et al.*: **The demosponge *Amphimedon queenslandica*: reconstructing the ancestral metazoan genome and deciphering the origin of animal multicellularity.** *Emerging Model Organisms: A Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press; 2008: 139-166.
9. Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu JK *et al.*: **The amphioxus genome and the evolution of the chordate karyotype.** *Nature* 2008, **453**:1064-1071.
10. Tribolium Genome Sequencing Consortium: **The genome of the model beetle and pest *Tribolium castaneum*.** *Nature* 2008, **452**:949-955.
11. King N, Westbrook MJ, Young SL, Kuo A, Abedin M, Chapman J, Fairclough S, Hellsten U, Isogai Y, Letunic I *et al.*: **The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans.** *Nature* 2008, **451**:783-788.
- The first nonmetazoan holozoan genome to be sequenced and analyzed. Comparison of its genome content and organization with that of early-branching metazoans allows for the identification of metazoan-specific events that may have occurred along the protometazoan stem.
12. Richards GS, Degnan BM: **The dawn of developmental signaling in the Metazoa.** *Cold Spring Harb Symp Quant Biol* 2009, **74**: in press.
13. Kusserow A, Pang K, Sturm C, Hroudá M, Lentfer J, Schmidt HA, Technau U, van Haeseler A, Hobmayer B, Martindale MQ *et al.*: **Unexpected complexity of the *Wnt* gene family in a sea anemone.** *Nature* 2005, **433**:156-160.
14. Nichols SA, Dirks W, Pearse JS, King N: **Early evolution of animal cell signaling and adhesion genes.** *Proc Natl Acad Sci U S A* 2006, **103**:12451-12456.
15. Matus DQ, Pang K, Marlow H, Dunn CW, Thomsen GH, Martindale MQ: **Molecular evidence for deep evolutionary roots of bilaterality in animal development.** *Proc Natl Acad Sci U S A* 2006, **103**:11195-11200.
- Nematostella* orthologs of genes known to be expressed asymmetrically during bilaterian development were shown to be expressed in similar patterns during cnidarian development, demonstrating a deep and unexpected conservation between bilaterians and supposedly radial diploblasts — see also Ref. [50].
16. Rentzsch F, Anton R, Saina M, Hammerschmidt M, Holstein TW, Technau U: **Asymmetric expression of the BMP antagonists chordin and gremlin in the sea anemone *Nematostella vectensis*: implications for the evolution of axial patterning.** *Dev Biol* 2006, **296**:375-387.
17. Adamska M, Degnan SM, Green KM, Adamski M, Craigie A, Larroux C, Degnan BM: ***Wnt* and *TGF-β* expression in the sponge *Amphimedon queenslandica* and the origin of metazoan embryonic patterning.** *PLoS ONE* 2007, **2**:e1031.
18. Adamska M, Matus DQ, Adamski M, Green K, Rokhsar DS, Martindale MQ, Degnan BM: **The evolutionary origin of hedgehog proteins.** *Curr Biol* 2007, **17**:R836-837.
19. Käsbauer T, Towb P, Alexandrova O, David CN, Dall'armi E, Staudigl A, Stiening B, Böttger A: **The Notch signaling pathway in the cnidarian *Hydra*.** *Dev Biol* 2007, **303**:376-390.
20. Matus DQ, Magie CR, Pang K, Martindale MQ, Thomsen GH: **The Hedgehog gene family of the cnidarian, *Nematostella vectensis*, and implications for understanding metazoan Hedgehog pathway evolution.** *Dev Biol* 2008, **313**:501-518.
21. Richards GS, Simonato E, Perron M, Adamska M, Vervoort M, Degnan BM: **Sponge genes provide new insight into the evolutionary origin of the neurogenic circuit.** *Curr Biol* 2008, **18**:1156-1161.
- This work identifies a single *atonal/neurogenin* ortholog in the sponge *Amphimedon* that is expressed in putative sensory cells and has strong proneural activity in bilaterians, behaving like *neurogenin* in *Xenopus* and *Drosophila*, as predicted to exist in an ancestral gene before duplication and divergence.
22. Davidson EH: *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution* Oxford: Academic Press; 2006.
23. Busser BW, Bulyk ML, Michelson AM: **Toward a systems-level understanding of developmental regulatory networks.** *Curr Opin Genet Dev* 2008, **18**:521-529.
24. Carroll SB: **Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution.** *Cell* 2008, **134**:25-36.
- One of the best reviews in evolutionary developmental biology in the last few years. Proposes a general theory with several principles about how GRNs functions and alterations control morphological evolution in animals and emphasizes the importance of the evolution of *cis*-regulatory sequences.
25. Ruiz-Trillo I, Burger G, Holland PW, King N, Lang BF, Roger AJ, Gray MW: **The origins of multicellularity: a multi-taxon genome initiative.** *Trends Genet* 2007, **23**:113-118.
26. Ruiz-Trillo I, Roger AJ, Burger G, Gray MW, Lang BF: **A phylogenomic investigation into the origin of metazoa.** *Mol Biol Evol* 2008, **25**:664-672.
27. Shalchian-Tabrizi K, Minge MA, Espelund M, Orr R, Ruden T, Jakobsen KS, Cavalier-Smith T: **Multigene phylogeny of choanozoa and the origin of animals.** *PLoS ONE* 2008, **3**:e2098.
- See annotation to Ref. [28*].
28. Mikhailov KV, Konstantinova AV, Nikitin MA, Troshin PV, Rusin LY, Lyubetsky VA, Panchin YV, Mylnikov AP, Moroz LL, Kumar S *et al.*: **The origin of Metazoa: a transition from temporal to spatial cell differentiation.** *Bioessays* 2009, **31**:758-768.
- Along with Ref. [27*], this work found a number of metazoan-specific genes, including T-box and NF-κB genes, in nonchoanoflagellate holozoan protists, revealing these genes antedate the holozoan LCA.
29. Martinelli C, Spring J: **Distinct expression patterns of the two T-box homologues *Brachyury* and *Tbx2/3* in the placozoan *Trichoplax adhaerens*.** *Dev Genes Evol* 2003, **213**:492-499.
30. Hadrys T, DeSalle R, Sagasser S, Fischer N, Schierwater B: **The *Trichoplax PaxB* gene: a putative Proto-PaxA/B/C gene predating the origin of nerve and sensory cells.** *Mol Biol Evol* 2005, **22**:1569-1578.
31. Jager M, Quéinnec E, Houlston E, Manuel M: **Expansion of the SOX gene family predated the emergence of the Bilateria.** *Mol Phylogenet Evol* 2006, **39**:468-477.
32. Larroux C, Fahey B, Liubicich D, Hinman VF, Gauthier M, Gongora M, Green K, Wörheide G, Leys SP, Degnan BM: **Developmental expression of transcription factor genes in a demosponge: insights into the origin of metazoan multicellularity.** *Evol Dev* 2006, **8**:150-173.
33. Richelle-Maurer E, Boury-Esnault N, Itskovich VB, Manuel M, Pomponi SA, Van de Vyver G, Borchiellini C: **Conservation and phylogeny of a novel family of non-Hox genes of the Antp class in Demospongiae (Porifera).** *J Mol Evol* 2006, **63**:222-230.
34. Larroux C, Fahey B, Degnan SM, Adamski M, Rokhsar DS, Degnan BM: **The NK homeobox gene cluster predates the origin of Hox genes.** *Curr Biol* 2007, **17**:706-710.
- Survey of ANTP genes in *Amphimedon* reveals a cluster of NKL genes but no Hox or ParaHox genes, supporting the hypothesis that Hox is a eumetazoan synapomorphy.
35. Simonato E, Ledent V, Richards G, Thomas-Chollier M, Kerner P, Coornaert D, Degnan BM, Vervoort M: **Origin and diversification of the basic helix-loop-helix gene family in metazoans: insights from comparative genomics.** *BMC Evol Biol* 2007, **7**:33.
- One of the first detailed whole genome comparative analyses of the evolution of a TF family in metazoans. It revealed periods during early metazoan evolution where particular groups and subfamilies are likely to have originated and expanded.

36. Gauthier M, Degnan BM: **The transcription factor NF-kappaB in the demosponge *Amphimedon queenslandica*: insights on the evolutionary origin of the Rel homology domain.** *Dev Genes Evol* 2008, **218**:23-32.
37. Gazave E, Lapébie P, Renard E, Bézac C, Boury-Esnault N, Vacelet J, Pérez T, Manuel M, Borchiellini C: **NK homeobox genes with choanocyte-specific expression in homoscleromorph sponges.** *Dev Genes Evol* 2008, **218**:479-489.
38. Larroux C, Luke GN, Koopman P, Rokhsar DS, Shimeld SM, ●● Degnan BM: **Genesis and expansion of metazoan transcription factor gene classes.** *Mol Biol Evol* 2008, **25**:980-996.
Using whole genome information from metazoans and choanoflagellates, this paper explores the origin and early evolution of homeobox, Sox, Fox, T-box gene families, revealing the protometazoan stem to be a key period in the origin of TF families and classes and the protoeumetazoan stem to be a time where these families expanded.
39. Schierwater B, Kamm K, Srivastava M, Rokhsar D, Rosengarten RD, Dellaporta SL: **The early ANTP gene repertoire: insights from the placozoan genome.** *PLoS ONE* 2008, **3**:e2457.
40. Sullivan JC, Sher D, Eisenstein M, Shigesada K, Reitzel AM, Marlow H, Levanon D, Groner Y, Finnerty JR, Gat U: **The evolutionary origin of the Runx/CBFBeta transcription factors — studies of the most basal metazoans.** *BMC Evol Biol* 2008, **8**:228.
41. Takahashi H, Kamiya A, Ishiguro A, Suzuki AC, Saitou N, Toyoda A, Aruga J: **Conservation and diversification of Msx protein in metazoan evolution.** *Mol Biol Evol* 2008, **25**:69-82.
42. Martinelli C, Spring J: **T-box and homeobox genes from the ctenophore *Pleurobrachia pileus*: comparison of *Brachyury*, *Tbx2/3* and *Tlx* in basal metazoans and bilaterians.** *FEBS Lett* 2005, **579**:5024-5028.
43. Derelle R, Manuel M: **Ancient connection between NKL genes and the mesoderm? Insights from *Tlx* expression in a ctenophore.** *Dev Genes Evol* 2007, **217**:253-261.
44. Yamada A, Pang K, Martindale MQ, Tochinai S: **Surprisingly complex T-box gene complement in diploblastic metazoans.** *Evol Dev* 2007, **9**:220-230.
45. Jager M, Quéinnec E, Chiori R, Le Guyader H, Manuel M: **Insights into the early evolution of SOX genes from expression analyses in a ctenophore.** *J Exp Zool B (Mol Dev Evol)* 2008, **310**:650-667.
46. Pang K, Martindale MQ: **Developmental expression of homeobox genes in the ctenophore *Mnemiopsis leidyi*.** *Dev Genes Evol* 2008, **218**:307-319.
An exhaustive RT-PCR survey of ANTP genes expressed in this ctenophore revealed a similar repertoire to that found in sponges, suggesting a possible close affinity.
47. Slack JM, Holland PW, Graham CF: **The zootype and the phylotypic stage.** *Nature* 1993, **361**:490-492.
48. Derelle R, Lopez P, Le Guyader H, Manuel M: **Homeodomain proteins belong to the ancestral molecular toolkit of eukaryotes.** *Evol Dev* 2007, **9**:212-219.
Detailed account of the early evolution of homeobox genes in eukaryotes. These and other ancient TF families may have contributed to the independent evolution of multicellularity in multiple lineages.
49. Mukherjee K, Bürglin TR: **Comprehensive analysis of animal TALE homeobox genes: new conserved motifs and cases of accelerated evolution.** *J Mol Evol* 2007, **65**:137-153.
50. Finnerty JR, Pang K, Burton P, Paulson D, Martindale MQ: **Origins of bilateral symmetry: Hox and dpp expression in a sea anemone.** *Science* 2004, **304**:1335-1337.
51. Martindale MQ, Pang K, Finnerty JR: **Investigating the origins of triploblasty: 'mesodermal' gene expression in a diploblastic animal, the sea anemone *Nematostella vectensis* (phylum, Cnidaria; class, Anthozoa).** *Development* 2004, **131**:2463-2474.
52. Magie CR, Pang K, Martindale MQ: **Genomic inventory and expression of Sox and Fox genes in the cnidarian *Nematostella vectensis*.** *Dev Genes Evol* 2005, **215**:618-630.
53. Chourrout D, Delsuc F, Chourrout P, Edvardsen RB, Rentzsch F, ●● Renfer E, Jensen MF, Zhu B, de Jong P, Steele RE et al.: **Minimal ProtoHox cluster inferred from bilaterian and cnidarian Hox complements.** *Nature* 2006, **442**:684-687.
See annotation to Ref. [54**].
54. Ryan JF, Burton PM, Mazza ME, Kwong GK, Mullikin JC, ●● Finnerty JR: **The cnidarian-bilaterian ancestor possessed at least 56 homeoboxes: evidence from the starlet sea anemone, *Nematostella vectensis*.** *Genome Biol* 2006, **7**:R64.
Along with Ref. [53**], this study reports the identification of the putative full complement of homeobox genes in a cnidarian. The large number and diversity of the retrieved homeobox genes indicate that an important diversification of the homeobox gene family has occurred before the split between cnidarians and bilaterians.
55. Bielen H, Oberleitner S, Marcellini S, Gee L, Lemaire P, Bode HR, Rupp R, Technau U: **Divergent functions of two ancient Hydra *Brachyury* paralogs suggest specific roles for their C-terminal domains in tissue fate induction.** *Development* 2007, **134**:4187-4197.
56. Kamm K, Schierwater B: **Ancient linkage of a POU class 6 and an anterior Hox-like gene in cnidaria: implications for the evolution of homeobox genes.** *J Exp Zool B (Mol Dev Evol)* 2007, **308**:777-784.
57. Matus DQ, Pang K, Daly M, Martindale MQ: **Expression of Pax gene family members in the anthozoan cnidarian, *Nematostella vectensis*.** *Evol Dev* 2007, **9**:25-38.
58. Mazza ME, Pang K, Martindale MQ, Finnerty JR: **Genomic organization, gene structure, and developmental expression of three clustered otx genes in the sea anemone *Nematostella vectensis*.** *J Exp Zool B (Mol Dev Evol)* 2007, **308**:494-506.
59. Sullivan JC, Kalaitzidis D, Gilmore TD, Finnerty JR: **Rel homology domain-containing transcription factors in the cnidarian *Nematostella vectensis*.** *Dev Genes Evol* 2007, **217**:63-72.
60. Shinzato C, Iguchi A, Hayward DC, Technau U, Ball EE, Miller DJ: **Sox genes in the coral *Acropora millepora*: divergent expression patterns reflect differences in developmental mechanisms within the Anthozoa.** *BMC Evol Biol* 2008, **8**:311.
61. Finnerty JR, Mazza ME, Jezewski PA: **Domain duplication, divergence, and loss events in vertebrate Msx paralogs reveal phylogenetically informed disease markers.** *BMC Evol Biol* 2009, **9**:18.
A phylogenomic analysis of the origin and evolution of the Msx gene family, revealing evolutionarily conserved domains that may inform clinical and functional studies.
62. Peterson KJ, Sperling EA: **Poriferan ANTP genes: primitively simple or secondarily reduced?** *Evol Dev* 2007, **9**:405-408.
63. Erwin DH: **Early origin of the bilaterian developmental toolkit.** *Philos Trans R Soc Lond B* 2009, **364**:2253-2261.
64. Galliot B, Quiquand M, Ghila L, de Rosa R, Miljkovic-Licina N, ●● Chera S: **Origins of neurogenesis, a cnidarian view.** *Dev Biol* 2009, **332**:2-24.
Recent and excellent review on the origin of neurogenesis, which integrates TF origins and gene duplications with the evolution of the nerve cell and nervous system.
65. Kamm K, Schierwater B, Jakob W, Dellaporta SL, Miller DJ: **Axial patterning and diversification in the Cnidaria predate the Hox system.** *Curr Biol* 2006, **16**:920-926.
See annotation to Ref. [68*].
66. Ryan JF, Mazza ME, Pang K, Matus DQ, Baxeavanis AD, ●● Martindale MQ, Finnerty JR: **Pre-bilaterian origins of the Hox cluster and the Hox code: evidence from the sea anemone, *Nematostella vectensis*.** *PLoS ONE* 2007, **2**:e153.
See annotation to Ref. [68*].
67. Hui JH, Holland PW, Ferrier DE: **Do cnidarians have a ParaHox cluster? Analysis of synteny around a *Nematostella* homeobox gene cluster.** *Evol Dev* 2008, **10**:725-730.
See annotation to Ref. [68*].

68. Chiori R, Jager M, Denker E, Wincker P, Da Silva C, Le Guyader H,
 • Manuel M, Quéinnec E: **Are Hox genes ancestrally involved in axial patterning? Evidence from the hydrozoan *Clytia hemisphaerica* (Cnidaria).** *PLoS ONE* 2009, **4**:e4231.

This paper along with Refs. [65*,66*,67*] as well as [34*,39,46*,50,53**,54**,62] contribute to the debate on the origin of ANTP homeobox genes, including NKL, Hox, and ParaHox gene clusters.

69. Sakarya O, Armstrong KA, Adamska M, Adamski M, Wang I-F, Sachdu D, Tidor B, Degnan BM, Oakley T, Kosik KS: **A post-synaptic scaffold at the origin of the animal kingdom.** *PLoS ONE* 2007, **2**:e506.
70. Khalturin K, Anton-Erxleben F, Milde S, Plötz C, Wittlieb J,
 • Hemmrich G, Bosch TC: **Transgenic stem cells in Hydra reveal an early evolutionary origin for key elements controlling self-renewal and differentiation.** *Dev Biol* 2007, **309**:32-44.
 See annotation to Ref. [72*].

71. Momose T, Derelle R, Houliston E: **A maternally localised Wnt ligand required for axial patterning in the cnidarian *Clytia hemisphaerica*.** *Development* 2008, **135**:2105-2113.
 See annotation to Ref. [72*].

72. Rentzsch F, Fritzenwanker JH, Scholz CB, Technau U: **FGF signalling controls formation of the apical sensory organ in the cnidarian *Nematostella vectensis*.** *Development* 2008, **135**:1761-1769.
 Along with Refs. [70*,71*], this paper reports transgenesis and functional analysis tools to study development in three cnidarian species and how these tools can provide significant insight in the genetic and molecular control of development in these species.

73. Fahey B, Larroux C, Woodcroft BJ, Degnan BM: **Does the high gene density in the sponge NK homeobox gene cluster reflect limited regulatory capacity?** *Biol Bull* 2008, **214**:205-217.