



SYMPOSIUM

Coral Comparative Genomics Reveal Expanded *Hox* Cluster in the Cnidarian–Bilaterian Ancestor

Timothy Q. DuBuc,* Joseph F. Ryan,[†] Chuya Shinzato,[‡] Nori Satoh[‡] and Mark Q. Martindale^{1,*}

*Kewalo Marine Laboratory, University of Hawaii, 41 Ahui Street, Honolulu, HI 96813, USA; [†]Sars International Centre for Marine Molecular Biology, University of Bergen, Thormøhlensgate 55, 5008 Bergen, Norway; [‡]Okinawa Institute of Science and Technology Graduate University, 1919-1 Tancha, Onna, Okinawa, 904-0495, Japan

The first two authors contributed equally to this work.

From the symposium “Evo-Devo Rides the Genomics Express” presented at the annual meeting of the Society for Integrative and Comparative Biology, January 3–7, 2012 at Charleston, South Carolina.

¹E-mail: mqmartin@hawaii.edu

Synopsis The key developmental role of the *Hox* cluster of genes was established prior to the last common ancestor of protostomes and deuterostomes and the subsequent evolution of this cluster has played a major role in the morphological diversity exhibited in extant bilaterians. Despite 20 years of research into cnidarian *Hox* genes, the nature of the cnidarian–bilaterian ancestral *Hox* cluster remains unclear. In an attempt to further elucidate this critical phylogenetic node, we have characterized the *Hox* cluster of the recently sequenced *Acropora digitifera* genome. The *A. digitifera* genome contains two anterior *Hox* genes (PG1 and PG2) linked to an *Eve* homeobox gene and an *Anthox1A* gene, which is thought to be either a posterior or posterior/central *Hox* gene. These data show that the *Hox* cluster of the cnidarian–bilaterian ancestor was more extensive than previously thought. The results are congruent with the existence of an ancient set of constraints on the *Hox* cluster and reinforce the importance of incorporating a wide range of animal species to reconstruct critical ancestral nodes.

Introduction

Hox genes are homeobox transcription factors that play a critical role in developmental patterning (McGinnis et al. 1984) and have been identified in every extant phylum outside of Porifera, Ctenophora, and Placozoa (placozoans have *ParaHox* but not *Hox* genes (Jakob et al. 2004; Ryan et al. 2010). The last common ancestor of protostomes and deuterostomes, which gave rise to 99% of all described animal species (Zhang 2011), is thought to have had an extensive cluster of *Hox* genes. Furthermore, since the expression of these genes in the body of many protostomes and deuterostomes is correlated with their position within the cluster, this ancestral condition was likely important for regulation of transcription (reviewed by Akam 1989).

Cnidarians (e.g., corals, sea anemones, hydroids, and medusae) are the only nonbilaterian phyla with

Hox genes, and therefore critical to our understanding of the early evolution of the *Hox* cluster. The exact relationship of cnidarian *Hox* genes with those of the protostomes and deuterostomes has been difficult to establish and has been debated (Finnerty and Martindale 1997; Gauchat et al. 2000; Yanze et al. 2001; Chourrout et al. 2006; Kamm et al. 2006; Jakob and Schierwater 2007; Ryan et al. 2007; Chiori et al. 2009). In addition, the genomic arrangement of *Hox* genes from sequenced cnidarian genomes have only been analyzed in *Nematostella vectensis* and *Hydra magnipapillata*. The *H. magnipapillata* genome shows no *Hox* cluster (Chapman et al. 2010) and the clustering in *N. vectensis* is limited to anterior *Hox* genes (Chourrout et al. 2006; Ryan et al. 2007; Putnam et al. 2007).

Most studies agree that cnidarians possess representatives of phylogenetically anterior *Hox* genes

including paralogous group 1 (PG1) and paralogous group 2 (PG2) (Finnerty and Martindale 1997; Chourrout et al. 2006; Ryan et al. 2007; Chiori et al. 2009). For example, the anthozoan *N. vectensis* has PG1 (*Anthox6* and *Anthox6a*) and PG2 genes (*Anthox7*, *Anthox8a*, and *Anthox8b*) (Finnerty and Martindale 1997; Chourrout et al. 2006; Ryan et al. 2007) and orthologs of these genes have been identified in the medusozoan *Clytia hemisphaerica* (Chiori et al. 2009).

The relationship of the other *Hox* genes in cnidarians is more controversial. For example, *N. vectensis* has a set of paralogs (*Anthox1* and *Anthox1a*) that have often been affiliated with posterior *Hox* genes (PG9–14) (Fig. 1) (Finnerty and Martindale 1997; Chourrout et al. 2006; Ryan et al. 2007). Three similar genes are found in *C. hemisphaerica* (Chiori et al. 2009). Our analyses, along with those of others (e.g., Chourrout et al. 2006; Chiori et al. 2009), show that these genes share almost equal phylogenetic affinity with central *Hox* genes (PG4–8) as they do with posterior *Hox* genes (Fig. 1). Therefore, either cnidarians have lost the ancestral central *Hox* gene, or the cnidarian–bilaterian ancestor possessed a single gene that gave rise to the bilaterian central and posterior genes as well as the cnidarian genes related to *Anthox1* and *Anthox1a* (Ryan et al. 2006).

In addition to disagreements as to the phylogenetic affinity of cnidarian *Hox* genes, there are differences in opinion as to the extent and nature of the cnidarian–bilaterian *Hox* cluster. Some authors have suggested, based on variations of expression domains in distant cnidarian species, that the cnidarian–bilaterian ancestor lacked *bona fide* *Hox* genes (Kamm et al. 2006; Schierwater and Kamm 2010). Others have asserted that the *Hox* cluster of the last common ancestor of cnidarians and bilaterians consisted of a maximum of two anterior-like *Hox* genes (Chourrout et al. 2006).

The criteria used by Kamm and coauthors to define what they call a “*Hox* system” are: (1) *Hox* genes are closely linked within the genome along the same chromosome; and (2) The *Hox* genes are primarily responsible for patterning structures along the primary body axis of a developing embryo (Kamm et al. 2006). In *N. vectensis*, the expression of *Hox* genes is restricted along the primary body axis (Finnerty et al. 2004; Ryan et al. 2007), but with only a partial cluster, it is perhaps difficult to assess these criteria *in totum*. However, the number of cnidarians examined to date and the extent of experiments involving cnidarian *Hox* genes are both far too few to rule out the existence of a functional *Hox* cluster in the cnidarian–bilaterian ancestor.

Recently, the genome of the staghorn coral *Acropora digitifera* was published (Shinzato et al. 2011). *Acropora digitifera* and *N. vectensis* are both members of the class Anthozoa; however, it is estimated that these two lineages diverged from each other some 500 million years ago (Shinzato et al. 2011). An analysis in another coral, *Acropora formosa*, was the first to show linkage between the *Hox*-related gene *Eve* and the PG1-related *Hox* gene *Anthox6* in a cnidarian (Miller and Miles 1993). In this article, we characterize the *Hox* cluster of *A. digitifera* and show that the *Hox* cluster in the last common ancestor of cnidarians and bilaterians was more elaborate than previously documented. This conserved synteny is consistent with an ancient functional constraint present in this ancestral *Hox* cluster.

Methods

Homeodomains from *N. vectensis*, *C. hemisphaerica*, *A. digitifera*, *Branchiostoma floridae*, *Homo sapiens*, *Drosophila melanogaster*, *Capitella telata*, and *Cupiennius salei* were aligned by eye. See Fig. 2 for accession numbers. We used ProtTest (AIC criteria) to determine that the LG + Gamma model best fit our multiple sequence alignment (Abascal et al. 2005). We ran RAXML version 7.2.8 with the following command line: `raxmlHPC-PTHREADS -T 6 -n hox -s hox.phy -m PROTGAMMALG -k -f a -N 100 -x 43241` (Stamatakis et al. 2008). The resulting tree is shown in Fig. 1. We used version 2.5.5 of Augustus (Stanke et al. 2006) to predict genes on the *A. digitifera* scaffold (DF093930) with the following command line: `augustus -species=human -AUGUSTUS_CONFIG_PATH=augustus.2.5.5/config DF093930.fa`.

Results

The *A. digitifera* genome (Shinzato et al. 2011) has the most extensive *Hox* cluster of all reported cnidarians. Phylogenetic analysis of 12 homeobox genes from the genome of *A. digitifera* revealed a total of six *Hox*, one *ParaHox*, three *Mox*, one *Eve*, and one *HlxB9* gene (Fig. 1 and Fig. 2). Two of the identified *Hox* genes (*Anthox6* and *Anthox7/8*) along with the homeobox genes *HlxB9* and *Eve* were found to be in the same 5′–3′ scaffold orientation as in *N. vectensis* (Fig. 3A and B). Contrary to *N. vectensis*, the 3′-end of the *A. digitifera* *Hox* cluster contains the central/posterior *Hox* gene *Anthox1a*. Interestingly, *Anthox1a* in *N. vectensis* is flanked by the pseudogene *Anthox9*, suggesting additional genes may have been present in the cnidarian–bilaterian cluster (Fig. 3).

A

Name	Scaffold GI	5' Scaff Pos.	Strand	Homeodomain Sequence
Ad_Ax1	344215709	303835	plus	GKRRKTAYTRKQELLEKEFFHNFHFLTERRAEMASQLNLTERQVKIWFQNRMMKWKKTN
Ad_MoxB	344215814	246320	plus	KRKDRTAFTKQIQELENEFERNNYLTLRRRYEIAVSLDLTERQVKVWFQNRMMKWKVKR
Ad_MoxD	344215814	253095	plus	LKRKERTVFTKSLQLEKEFFGRNNYLTLRRRYEAVSLDLTERQVKVWFQNRMMKWKVIR
Ad_MoxC	344215814	264146	plus	RRKRTAFTKHLQLENEFLRNNYLTLRRRYEIAVSLDLTERQVKIWFQNRMMKWKRIK
Ad_Hlxb9	344215894	27918	plus	QRRPRTAFSSQQLLTERQFQAQKYLTRPQRYELATSLMLTETQVKIWFQNRMMKWKRCG
Ad_Ax6	344215894	214976	minus	SNKKRFTTQRQLLEKEFFHSKYLTRRRIEIASNLDLTETQIKIWFQNRMMKWKREL
Ad_Evx	344215894	226289	minus	TRRYRTAFTREQLSRLKEFLRENYVSRTRRSELASMLNSETTKIWFQNRMMKAKRRR
Ad_Ax7/8	344215894	252838	minus	DKRRHVSYTNKQLEKEFFHNKYLCSRRSEIAKTLSLSERQVKIWFQNRMMKMKKDE
Ad_Ax1a	344215894	264704	minus	KHRRRVATRNQLEKEFFHTRYLTKRERSEMMLKLERQIKIWFQNRMMKWKND
AdGsx	344216069	40720	minus	SKRIRTAYSMQLLEKEFFSQNRYSRLRRRIQIAALLDLSEKQVKIWFQNRMMKWKDDK
Ad_Ax6a	344217215	33972	plus	ADKNRTIYSTRQLVELEKEFFHYNRYLCRPRRIEIAQTLGLTEKQVKIWFQNRMMKWKKEN
Ad_Ax6b	344217215	49094	plus	ADKNRTIYSTRQLVELEKEFFLYDKYLCRPRRIEIALSLGLTEKQVKIWFQNRMMKWKKEA

B

Nv_Hlxb9	HRRPRTAFSSQQLLALERQQLHXYLTPQRYELATSLMLTETQVKIWFQNRMMKWKRCN	Dm_btn	NKRKRTAFSKIQLEKEAEFCYSNYLTLRRRYEIAVALELTERQVKVWFQNRMMKCKRIK
Hs_Hlxb9	CRRPRTAFSSQQLLELEHOFKNKYLSPKRFVAVATSLMLTETQVKIWFQNRMMKWKRSK	Dm_ro	QRRQRTFTSTQTLREVEFPHNRYISRRRPELAETLRLTETQIKIWFQNRMMKCKRIE
Dm_exex	TRRPRTAFSSQQLLELEHOFKONKYLSPKRFVAVATSLMLTETQVKIWFQNRMMKWKRSK	Ct_x10x	NKRRTAYTRAQLLELEKEFFPNRYTRPRRVELAHLNTEQIKIWFQNRMMKCKDV
Bf_Gbx	TRRRRTAFSSQQLLELEKEFFSKYLSLTERSQIAHALKLSVQVKIWFQNRMMKWKVKV	Nv_ax8a	SKRHRTSYTRQLLELEKEFFPNKYLCSRRRIEISKALQLTERQVKIWFQNRMMKWKDE
Bf_Cdx	KKRYRVVTSQQLLELEKEFFSNRYITIKREVQLANGLSSEQVKIWFQNRMMKQRMA	Nv_ax6	SQKRRTFTQMLVELEKEFFHSKYLTRRIELATLKLTEMQIKIWFQNRMMKWKRF
Bf_EvxA	VRRYRTAFTRGLARLEKEFFREYVSRPRCELAQMLPETTIKVFQNRMMKDKQR	Nv_gbx	RRKRTAFTKHLQLEKEFFHNNYLSLEERSVIATMLNTEQVKIWFQNRMMKSKRV
Bf_EvxB	VRRYRTAFSSQQLARLEKEFFRNDYLSRPRCELAALMLPETTIKVFQNRMMKDKQR	Nv_moxC	KKKRTAFSKHLQLEKEFFHNNYTLRRRYEIAVSLDLTERQVKVWFQNRMMKSKRVK
Dm_eve	TRRYRTAFTRDGLRLEKEFFENYVSRPRCELAQMLPETTIKVFQNRMMKDKQR	Nv_ax7	TKRYRTSYTRQLLELEKEFFHNNYLCGRRRELANAMKLERQVKVWFQNRMMKSKDE
Nv_Evx	TRRYRTAFTRDGLRLEKEFFENYVSRTRRCELANALMLPETTIKVFQNRMMKSKRRR	Nv_moxA	KKKRTAFTKHLQLEKEFFHNNYTLRRRYEIAVSLDLTERQVKVWFQNRMMKSKRVK
Bf_Gax	SRRRRTAFSSQQLLELEKEFFSNRYITIKREVQLANGLSSEQVKIWFQNRMMKQRMA	Nv_hd065	RSRRTAYTASQLEKEFFLYSRYTRRRELANTLQSEKIKIWFQNRMMKCKTD
Bf_Hox1	PNRRGTFTTQQLLELEKEFFHNNYLSRPRCELAALMLPETTIKVFQNRMMKDKQR	Nv_ax6a	SDKNRTIYSTRQLVELEKEFFHNNYLCRPRRIEIAQTLGLTEKQVKIWFQNRMMKCKEN
Bf_Hox2	SKRLRTVFTTQQLLELEKEFFHNNYLSRPRCELAALMLPETTIKVFQNRMMKDKQR	Nv_ax1a	KHRRRVATRNQLEKEFFHTRYLTKRERSEMMLKLERQIKIWFQNRMMKWKND
Bf_Hox3	GKRRRTAFSSQQLLELEKEFFHNNYLCRPRRVEAAMLMLETQIKIWFQNRMMKWKKEQ	Nv_ax1	GKRRRTAYTRKQLELEKEFFHNFHFLTERRAEMASQLNLTERQVKIWFQNRMMKWKCN
Bf_Hox4	TKRSRTAFTKQIQLELEKEFFHNNYTLRRRYEIAHSLGLTERQIKIWFQNRMMKWKVDN	Nv_moxB	KKERTAFTKHLQLEKEFFHNNYTLRRRYEIAVSLDLTERQVKVWFQNRMMKWKVKR
Bf_Hox5	VRRYRTAFTRDGLRLEKEFFENYVSRTRRCELANALMLPETTIKVFQNRMMKDKQR	Nv_ax9	NHAPRRTAFTKHLQLEKEFFHNNYLSRPRRIEIAVSLDLTERQVKIWFQNRMMKCKREA
Bf_Hox6	KKRGQTYTRYQLELEKEFFHNNYTLRRRYEIAHALGLTERQIKIWFQNRMMKWKKEN	Nv_moxD	KKKRTVFSKYLQLEKEFFHNNYTLRRRYEIAVSLDLTERQVKVWFQNRMMKSKRGR
Bf_Hox7	KKRGQTYTRYQLELEKEFFHNNYTLRRRYEIAHALGLTERQIKIWFQNRMMKWKKEN	Nv_ro	FRQRTFTTQQLLELEKEFFHNNYLSRPRRIEIAVSLDLTERQVKIWFQNRMMKCKREA
Bf_Hox8	KKRGQTYTRYQLELEKEFFHNNYTLRRRYEIAHALGLTERQIKIWFQNRMMKWKKEN	Nv_gbx	SKRIRTAYSMQLLELEKEFFSQNRYSRLRRRIQIAALLDLSEKQVKIWFQNRMMKWKDDK
Bf_Hox9	SKRRKCPYTRYQLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKWKKEN	Nv_ax8b	SKRHRTSYTRQLLELEKEFFHNNYLSRPRRIEISKALQLTERQVKIWFQNRMMKWKDE
Bf_Hox10	GKRRKCPYTRYQLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKWKKEN	Ch_Cdx	TRRSRCPSSHTRELEKEFFHNNYLSRPRRIEIAHALGLTERQVKIWFQNRMMKCKRQK
Bf_Hox11	SKRRKCPYTRYQLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKWKKEN	Ch_Eve	PHRTAFTKHLQLEKEFFHNNYLSRPRRIEIAHALGLTERQVKIWFQNRMMKCKRQK
Bf_Hox12	SKRRKCPYTRYQLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKWKKEN	Ch_Gax	SKRIRTAFTSILQLELEKEFFHNNYLSRPRRIEIAHALGLTERQVKIWFQNRMMKCKRQK
Bf_Hox13	GKRRKCPYTRYQLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKWKKEN	Ch_Hox1	QKRRKCPYTRYQLELEKEFFHNNYLSRPRRIEIAHALGLTERQVKIWFQNRMMKCKRQK
Bf_Hox14	VPRKRTAFSSQQLLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKWKKEN	Ch_Hox9_14A	GKRRKRTAFSSQQLLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKCKRQK
Bf_Mox	PKKRTAFTKQIQLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKWKKEN	Ch_Hox9_14B	HRKRTAYTASQLEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKCKRQK
Bf_Xlox	NKRRTAYTRAQLLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKCKRQK	Ch_Hox9_14C	TKRRKRTAYTASQLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKCKRQK
Dm_ind	SKRIRTAFTSILQLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKCKRQK	Ch_MoxC	SKRRTAFTKQIQLELEKEFFHNNYTLRRRYEIAVSLDLTERQVKIWFQNRMMKCKRQK
Dm_cad	KKRYRVVTSQQLLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKCKRQK	Ad_Ax1a	KHRRRVATRNQLELEKEFFHTRYLTKRERSEMMLKLERQIKIWFQNRMMKWKND
Dm_unp	SRRRRTAFSSQQLLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKCKRQK	Ad_Ax6a	ADKNRTIYSTRQLVELEKEFFHNNYLCRPRRIEIAQTLGLTEKQVKIWFQNRMMKCKEN
Dm_lap	PNRRGTFTTQQLLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKCKRQK	Ad_Ax6b	ADKNRTIYSTRQLVELEKEFFLYDKYLCRPRRIEIALSLGLTEKQVKIWFQNRMMKCKE
Dm_pb	PRRLRTAFTKHLQLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKCKRQK	Ad_Evx	TRRYRTAFTREQLSRLKEFLRENYVSRTRRSELASMLNSETTKIWFQNRMMKAKRRR
Cs_sen	AKRRRTAFTSILQLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKCKRQK	Ad_Gax	SKRIRTAFTSILQLELEKEFFHNNYLSRPRRIEIAHALGLTERQVKIWFQNRMMKCKRQK
Dm_dfd	FRQRTFTTQQLLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKCKRQK	Ad_Hlxb9	QRRPRTAFSSQQLLTERQFQAQKYLTRPQRYELATSLMLTETQVKIWFQNRMMKWKRCG
Dm_scr	TKRSRTAFTKQIQLELEKEFFHNNYTLRRRYEIAHSLGLTERQIKIWFQNRMMKWKVDN	Ad_MoxB	KKKRTAFTKHLQLEKEFFHNNYTLRRRYEIAVSLDLTERQVKVWFQNRMMKWKVKR
Dm_ftz	SKRRRTAYTRAQLLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKCKRQK	Ad_MoxC	RRKRTAFTKHLQLEKEFFHNNYTLRRRYEIAVSLDLTERQVKIWFQNRMMKWKRIK
Dm_antp	KKRGQTYTRYQLELEKEFFHNNYTLRRRYEIAHALGLTERQIKIWFQNRMMKWKKEN	Ad_MoxD	LKRKERTVFTKSLQLEKEFFHNNYTLRRRYEIAVSLDLTERQVKVWFQNRMMKWKVIR
Dm_ubx	RRRGQTYTRYQLELEKEFFHNNYTLRRRYEIAHALGLTERQIKIWFQNRMMKWKKEN	Ad_Ax1	GKRRRTAYTRKQLELEKEFFHNFHFLTERRAEMASQLNLTERQVKIWFQNRMMKWKCN
Dm_abdA	RRRGQTYTRYQLELEKEFFHNNYTLRRRYEIAHALGLTERQIKIWFQNRMMKWKKEN	Ad_Ax6	SNKKRFTTQRQLLELEKEFFHSKYLTRRRIEIASNLDLTETQIKIWFQNRMMKWKREL
Dm_abdE	VKKRRTAFSSQQLLELEKEFFHNNYLSRPRRIEIAHALGLTERQIKIWFQNRMMKCKRQK	Ad_Ax7/8	DKRRHVSYTNKQLELEKEFFHNKYLCSRRSEIAKTLSLSERQVKIWFQNRMMKMKKDE

Fig. 2 Information on sequences of the *A. digitifera* homeodomain and the alignment with other taxa. (A) Sequences are organized by their linkage in the genome, whereby genes sharing the same color are found on the same scaffold. Genes in white were not linked to other *Hox* or *Hox*-related genes. (B) Alignment of homeodomain sequences.

N. vectensis PG2-related genes (*Anthox7*, *8*, and *8a*) are represented by only a single gene (*Anthox7/8*) in *A. digitifera*. This is congruent with the assertion that the *Anthox7*, *8*, and *8a* genes were most likely the result of recent duplications in the *N. vectensis* lineage (Ryan et al. 2007).

The *Hox* cluster of *A. digitifera* is situated on the end of a 270-kb scaffold. We predict 16 genes downstream of the *A. digitifera Hox* cluster (Fig. 4A). Of these, only *Rac3*, *Dars*, *Psm2*, and the *Hox*-related *Hlxb9* are linked to the cluster in both anthozoans (Fig. 4). *Dars* and *Psm2* are adjacent to each other in both genomes, but are on opposite sides of the cluster. Seven of the 16 downstream genes appear not to be linked to the *N. vectensis* cluster (Fig. 4C).

Discussion

For the first time, we show that a cnidarian *Hox* gene (i.e., the *A. digitifera Anthox1a* gene) related to the

central/posterior *Hox* genes of the Bilateria is linked to a cluster of genes that includes anterior *Hox* genes and an *Eve* gene (Fig. 3B). This new evidence and the genomic linkage between the *Anthox1a* and *Anthox9* genes of *N. vectensis* suggest that a larger cluster most certainly existed in the cnidarian–bilaterian ancestor (Fig. 3C). The *N. vectensis Anthox9* is thought to be a highly derived pseudogene and is unstable in phylogenies making it difficult to know its true identity (Kamm et al. 2006; Ryan et al. 2007). An ortholog of *Anthox9* was not identified in the *A. digitifera* genome. These data show that the *Hox* cluster of the cnidarian–bilaterian ancestor consisted of at least two anterior-related *Hox* genes, one central/posterior-related *Hox* gene, an *Eve homeobox*, and perhaps another gene related to *Anthox9* (Fig. 3C).

Despite many examples of rearrangements and breakages, the persistence of a *Hox* cluster in disparate bilaterian lineages is attributed to constraints

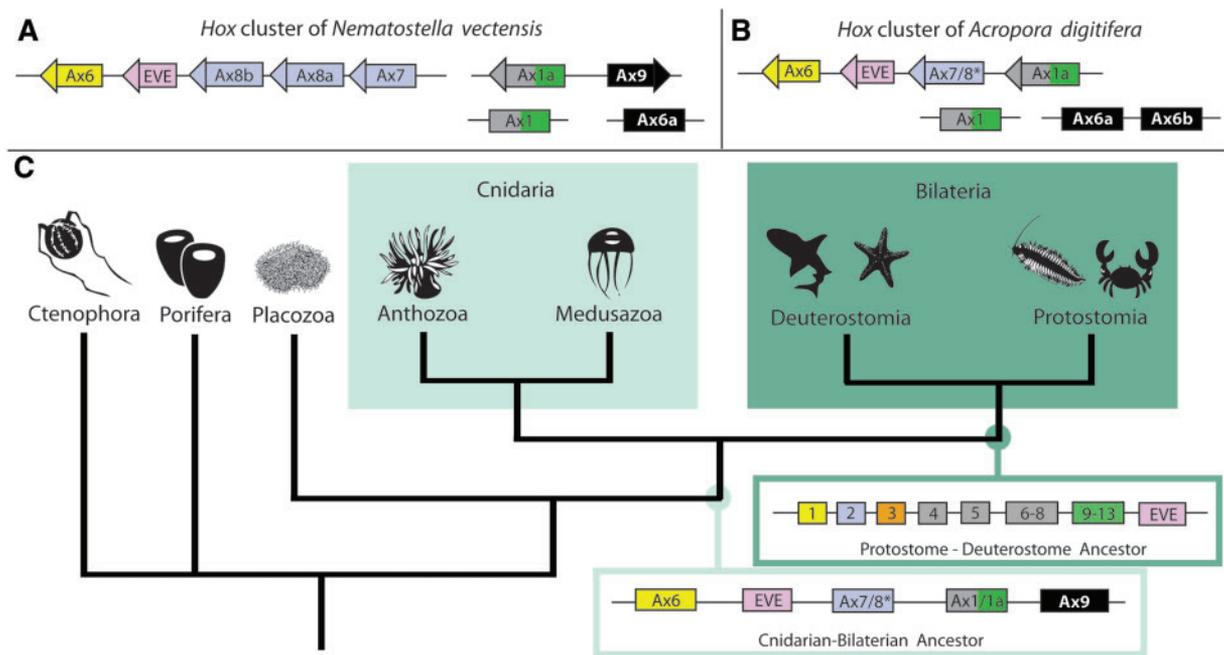


Fig. 3 The anthozoan complement of *Hox* genes and the implications of the evolution of the *Hox* cluster. Comparing the genomic linkage of *Hox* genes in the sea anemone *N. vectensis* and the staghorn coral *A. digitifera* confirms that cnidarians once had a *Hox* cluster that contained both anterior and posterior/central class *Hox* genes. (A) The *Hox* cluster of *N. vectensis* includes the anterior *Hox* genes *Anthox6* (PG1), *Anthox8b* (PG2), *Anthox8a* (PG2), and *Anthox7* (PG2) as well as the *Eve* homeobox gene. (B) The *Hox* cluster of *A. digitifera* includes the anterior *Hox* genes *Anthox6* (PG1) and *Anthox7/8* (PG2), and the posterior/central class *Hox* gene *Anthox1a* (PG4–14), as well as the *Eve* homeobox gene. Another gene *HlxB9* (also named *MNX*) is found upstream of *Anthox6* in the *Hox* cluster of both genomes (data not shown). (C) The metazoan tree of life with inferred ancestral *Hox* clusters. The ancestor to protostomes and deuterostomes is thought to have had two anterior class *Hox* genes (*Hox1* and *Hox2*), one paralogous group 3 gene (*Hox3*), three central class genes (*Hox4*, *Hox5*, and *Hox6–8*), one posterior class *Hox* gene (*Hox9–14*), and one *Eve* homeobox gene. Because of the extended cluster in *A. digitifera*, we can now say that the cnidarian–bilaterian ancestor had, at least, two anterior class *Hox* genes (*Anthox6* and *Anthox7/8*), a central/posterior class *Hox* gene (*Anthox1/1a*), and the *Eve* homeobox gene. It is unclear at what point the genomic rearrangement involving the *Eve* homeobox gene occurred. The origin of the PG3 *Hox* genes also is not clear. **Anthox7/8* has been categorized as a PG2 *Hox* gene in previous publications, but it is possible, based on our current phylogenetic analysis, that *Anthox7/8* descended from a *Hox* gene that was lost in bilaterians. Based on the genomic orientation of these genes, we also believe the ancestor likely had a fourth *Hox* gene potentially related to *Anthox9*. Abbreviations: PG = paralogous group, Ax = Anthox.

on the developmental regulation of these genes. The absence of *Hox* clustering in the the *H. magnipapillata* genome (Chapman et al. 2010) and the partial clustering in *N. vectensis*, a genome remarkable for its large-scale conserved synteny with vertebrate genomes (Putnam et al. 2007), presented the possibility that these regulatory constraints were established after cnidarians and bilaterians diverged. Nevertheless, the presence of an extensive *Hox* cluster in a third cnidarian lineage suggests that this regulatory constraint dates back prior to the last common cnidarian–bilaterian ancestor. The conservation of the clusters in the two anthozoan lineages despite the many genomic events that appear to have occurred in the region (Fig. 4) reinforces this view.

Our current understanding of the early evolution of the *Hox* cluster is still at an early stage. As more cnidarian genomes are sequenced, and as experimental techniques are established for these new model systems (including *A. digitifera*), the structure and function of the cnidarian–bilaterian ancestor's *Hox* cluster will become even clearer. A better understanding of the similarities and differences between the *Hox* clusters of the cnidarian–bilaterian ancestor and the protostome–deuterostome ancestor will help explain the origin of bilaterian-specific complexities. Furthermore, more comprehensive surveys into the independent variation of *Hox* genes in cnidarian lineages will lead to a better understanding of the role these genes have played in establishing the vast diversity of body plans exhibited in the Cnidaria.

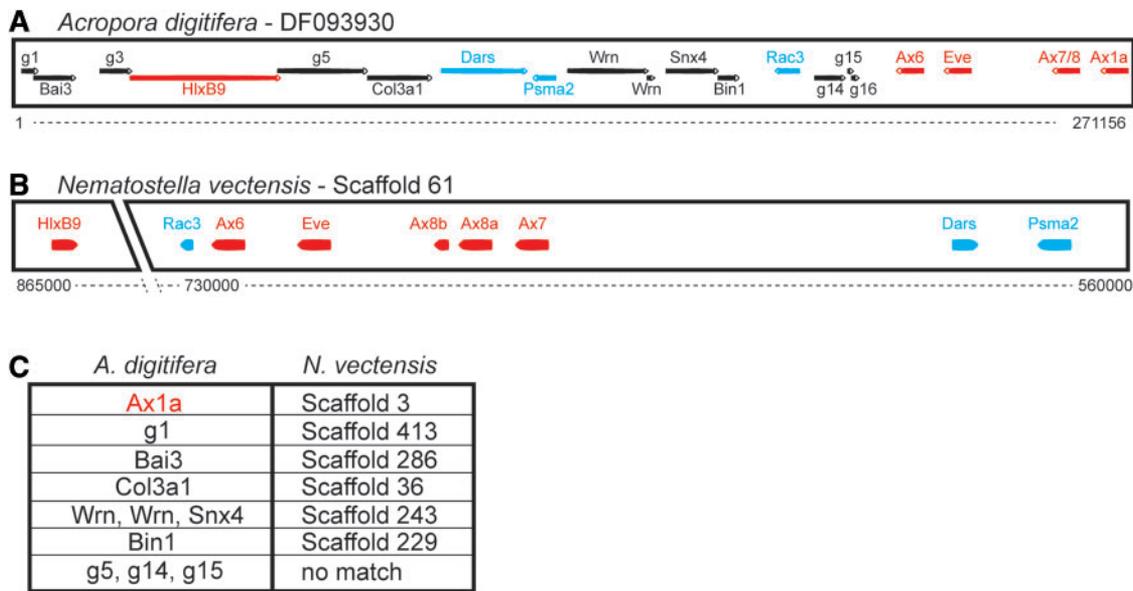


Fig. 4 Genes associated with the *Hox* cluster of *A. digitifera* and their location in the *N. vectensis* genome. **(A)** The *Hox* cluster of *A. digitifera* occurs on the scaffold with the NCBI accession DF093930 (total size = 271,156 bp). Genes on this scaffold were predicted with the Augustus gene finder (Stanke et al. 2006). If the resulting protein had a reciprocal best BLAST hit with a human RefSeq, the gene name associated with that RefSeq is used in the figure; if not, the genes were numbered from g1 to g16. Predictions and spatial relationships are roughly to scale. **(B)** The *Hox* cluster of *N. vectensis* occurs on the scaffold with the JGI ID# 61 (total size = 1,073,712 bp) in v1.0 of the JGI assembly. This scaffold includes three nonhomeobox genes (*DARS*, *PSMA2*, and *RAC3*) that are also associated with the *A. digitifera* *Hox* cluster. *Nematostella vectensis* genes not associated with the *A. digitifera* *Hox* cluster are not displayed. The broken line and broken box indicate that the distance between the *HlxB9* gene and the rest of the scaffold is not to scale. **(C)** Those genes in the *A. digitifera* cluster not on scaffold 61 in *N. vectensis* are listed along with the corresponding JGI scaffold number. The coordinates below the A and B panels indicate the scaffold coordinates of the region shown. Genes in red are *Hox* and *Hox*-related genes. Genes in blue are nonhomeobox genes that are associated with the *Hox* cluster in both *A. digitifera* and *N. vectensis*.

Acknowledgments

JFR would like to thank Andreas Hejnl for support and insightful conversations regarding this work. The authors would like to thank two anonymous referees for their insightful review of an earlier version of this manuscript, which led to significant improvements.

Funding

JFR was supported by the Sars Centre. This research was funded by the National Science Foundation.

References

- Abascal F, Zardoya R, Posada D. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 21:2104–5.
- Akam M. 1989. *Hox* and *HOM*: homologous gene clusters in insects and vertebrates. *Cell* 57:347–9.
- Chapman JA, Kirkness EF, Simakov O, Hampson SE, Mitros T, Weinmaier T, Rattei T, Balasubramanian PG, Borman J, Busam D, et al. 2010. The dynamic genome of *Hydra*. *Nature* 464:592–6.
- Chiori R, Jager M, Denker E, Wincker P, Da Silva C, Le Guyader H, Manuel M, Quéinnec E. 2009. Are *Hox* genes ancestrally involved in axial patterning? Evidence from the

- hydrozoan *Clytia hemisphaerica* (Cnidaria). *PLoS One* 4:e4231.
- Chourrout D, Delsuc F, Chourrout P, Edvardsen RB, Rentsch F, Renfer E, Jensen MF, Zhu B, de Jong P, Steele RE, et al. 2006. Minimal Proto*Hox* cluster inferred from bilaterian and cnidarian *Hox* complements. *Nature* 442:684–7.
- Finnerty JR, Martindale MQ. 1997. Homeoboxes in sea anemones (Cnidaria:Anthozoa): a PCR-based survey of *Nematostella vectensis* and *Metridium senile*. *Biol Bull* 193:62–76.
- Finnerty JR, Pang K, Burton P, Paulson D, Martindale MQ. 2004. Origins of bilateral symmetry: *Hox* and *dpp* expression in a sea anemone. *Science* 304(5675):1335–7.
- Gauchat D, Mazet F, Berney C, Schummer M, Kreger S, Pawlowski J, Galliot B. 2000. Evolution of Antp-class genes and differential expression of *Hydra Hox/paraHox* genes in anterior patterning. *Proc Natl Acad Sci USA* 97:4493–8.
- Jakob W, Sagasser S, Dellaporta S, Holland P, Kuhn K, Schierwater B. 2004. The *Trox-2 Hox/ParaHox* gene of *Trichoplax* (Placozoa) marks an epithelial boundary. *Dev Genes Evol* 214:170–5.
- Jakob W, Schierwater B. 2007. Changing hydrozoan bauplans by silencing *Hox*-like genes. *PLoS One* 2:e694.
- Kamm K, Schierwater B, Jakob W, Dellaporta SL, Miller DJ. 2006. Axial patterning and diversification in the cnidaria predate the *Hox* system. *Curr Biol* 16:920–6.

- McGinnis W, Levine MS, Hafen E, Kuroiwa A, Gehring WJ. 1984. A conserved DNA sequence in homoeotic genes of the *Drosophila* Antennapedia and bithorax complexes. *Nature* 308:428–33.
- Miller DJ, Miles A. 1993. Homeobox genes and the zootype. *Nature* 365:215–6.
- Monteiro AS, Ferrier DE. 2006. Hox genes are not always Colinear. *Int J Biol Sci* 2:95–103.
- Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, Salamov A, Terry A, Shapiro H, Lindquist E, Kapitonov VV, et al. 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* 317:86–94.
- Ryan JF, Burton PM, Mazza ME, Kwong GK, Mullikin JC, Finnerty JR. 2006. The cnidarian-bilaterian ancestor possessed at least 56 homeoboxes: evidence from the starlet sea anemone, *Nematostella vectensis*. *Genome Biol* 7:R64.
- Ryan JF, Mazza ME, Pang K, Matus DQ, Baxevanis AD, Martindale MQ, Finnerty JR. 2007. Pre-bilaterian origins of the *Hox* cluster and the *Hox* code: evidence from the sea anemone, *Nematostella vectensis*. *PLoS One* 2:e153.
- Ryan JF, Pang K, NISC Comparative Sequencing Program, Mullikin JC, Martindale MQ, Baxevanis AD. 2010. The homeodomain complement of the ctenophore *Mnemiopsis leidyi* suggests that Ctenophora and Porifera diverged prior to the ParaHoxozoa. *Evodevo* 1:9.
- Schierwater B, Kamm K. 2010. The early evolution of Hox genes: a battle of belief? *Adv Exp Med Biol* 689:81–90.
- Shinzato C, Shoguchi E, Kawashima T, Hamada M, Hisata K, Tanaka M, Fujie M, Fujiwara M, Koyanagi R, Ikuta T, et al. 2011. Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature* 18:320–3.
- Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* 57:758–71.
- Stanke M, Tzvetkova A, Morgenstern B. 2006. AUGUSTUS at EGASP: using EST, protein and genomic alignments for improved gene prediction in the human genome. *Genome Biol* 7(Suppl 1):S11.1–8.
- Swalla BJ. 2006. Building divergent body plans with similar genetic pathways. *Heredity (Edinb)* 97:235–43.
- Yanze N, Spring J, Schmidli C, Schmid V. 2001. Conservation of Hox/ParaHox-related genes in the early development of a cnidarian. *Dev Biol* 236:89–98.
- Zhang Z. 2011. An introduction to higher-level classification and taxonomic richness. *Zootaxa* 3148:7–12.