# Hox, microRNAs and evolution : new insights into the patterning of the body axis

Woltering, J.M.

# *Chapter 1*

## The microRNA-*Hox* connection

**Joost M. Woltering**

## Abstract

The *Hox* genes play a key role in the determination of the anterior-posterior pattern of the metazoan body axes and belong with their clustered organization to the most intriguing loci in the genome. In the last 4 years advances in the non coding RNA field have resulted in the discovery of several microRNA genes within the *Hox* clusters and the *Hox* genes themselves have been shown to be subject to microRNA regulation. In this review I cover the literature with respect to the *Hox* - microRNA relations and present the current understandings of how this second layer of posttranscriptional control exerted by and on the *Hox* clusters fits within the already much described network of *Hox* gene regulation. In addition, I summarize the *Hox* related microRNAs present in the model vertebrates human, mouse, *Xenopus* and Zebrafish and correct some anomalies in the published annotations. I also describe a new divergent *miR-196* member from *Xenopus tropicalis* and present additional *miR-10* and *miR-196* expression data in avian embryos.

## *Hox* genes and regionalization of the body

The vertebrate trunk is highly regionalized along its anterior to posterior axis, with related structures having a different appearance at different positions in the body (1). This division of the trunk is part of the complex adaptations characterizing all higher metazoans where different parts of the body are dedicated to specialized functions.

During embryonic stages the basis for such regionalization is realized through the differential development of homologous elements, depending on where along the axis they form. Especially obvious is this regionalization in the axial skeleton and the central nervous system. In the paraxial mesoderm seemingly equal somites give rise to several different types of vertebrae. These are for instance rib carrying in the thoracic region, in the lumbar region they lack ribs and in the sacral region they contribute to the formation of the sacrum. The same concept of regionalization applies to the central nervous system where, as an example, motorneurons develop different characteristics and functions depending on their position along the anterior-posterior (AP) axis of the spinal cord and hindbrain (2). In this process of regionalization the *Hox* genes have been identified as the genetic key components (reviewed 3, 4). The *Hox* genes have differential expression along the trunk AP axis and are responsible for many of the choices between alternative pathways of development.

*Hox* genes have selector function within their expression domain and determine the anterior to posterior positional characteristics that a tissue will develop.

In both vertebrates and invertebrates there is much experimental evidence linking *Hox* genes to regional identity along the trunk axis. In the nervous system and the axial skeleton experimental interference with proper *Hox* functioning during embryonic stages, can induce cells to adopt a fate corresponding to an axial position different from the one that they would have acquired normally. Hereby it is for example possible to transform *lumbar*- (posterior) into *thoracic* (more anterior) vertebrae (5) or to change hindbrain rhombomere fate into either more posterior (6, 7) or more anterior (8). In Drosophila, the spectacular *Ultrabithorax* (*Ubx*) mutants develop four instead of two wings.

By now most *Hox* genes have been linked to the induction of a specific body segment and in this way a '*Hox* code' for the patterning of the primary body axis has been unraveled (9,10).

Comparison across a wide range of different taxa has shown that this code is generally conserved. Besides the patterning of the primary body axis, *Hox* genes are also expressed in and have a similar function in the patterning of the 'derived' body axes of the limbs and digestive system.

## *Hox* genes genomics

The *Hox* genes encode a family of closely related homeodomain transcription factors and presumably find their evolutionary origin in *cis*-duplications of a single ancestral gene (11). This way of creation has resulted in a clustered configuration of the *Hox* genes. In all vertebrates this clustering has been preserved till contemporary organisms and it is one of the most characteristic aspects of the *Hox* family. In non vertebrates, however, there are big deviations from the ancestral organization. In insects such as *Anopheles*, *Schistocera* and *Tribolium* the clustered organization is still intact but in other taxa the *Hox* clusters have undergone extensive rearrangements, contain additional gene duplications and are sometimes (partially) fragmented (e.g. *Drosophila spec.*, *Strongylocentrotus purpureus*, *Caenorhabditis elegans*, *Ciona spec.*, *Oikopleura dioica*) (12).

The two genome duplications in the vertebrate lineage have resulted in 4 *Hox* clusters in *Latimeria menadoniensis*, *Tetrapods* and Sharks (named *A*, *B*, *C* and *D*) (13). Afterwards redundancy among the duplicated gene groups has resulted in some degree of gene loss but overall the clusters have stayed relatively intact; each cluster still contains the majority of *Hox* genes and all genes are present in multiple paralogues in the genome. Members from one paralogous group have in general similar expression patterns and partially redundant functions. In the *Teleost* lineages an additional whole genome duplication occurred and these fish therefore originally possessed eight *Hox* clusters (named *Aa*, *Ab*, *Ba*, *Bb*, etc.) (13, 14). In today's species, usually one of each of two duplicated *Hox* clusters has partially degenerated and contains fewer genes. In all well characterized fish species this process has lead to the (virtual) disappearance of one *Hox* cluster; Fugu (*Tetraodon nigroviridis* and *Takifugu rupriceps*) and Medaka (*Oryza latipes*) both miss one *HoxC* cluster and in Zebrafish (*Danio rerio*) the *HoxDb* cluster doesn't posses *Hox* coding genes anymore (13).

## Colinearity

*Hox* genes are expressed in a highly structured sequential order along the primary body axis and in the limbs which is essential for the correct execution of their patterning functions. There is a remarkable relationship between the sequence of expression of the *Hox* genes and the sequence in which they are located in the clusters. The more 3' in the cluster a gene is located the more anterior its expression domain and the earlier its onset of expression during development. This transcriptional behavior is referred to as colinearity which can be both spatial (anterior-posterior) and temporal (timing).

## MicroRNAs

With the discovery of the *lin-14* founding member (15) of the microRNAs, a new paradigm in gene regulation was established and large scale microRNA cloning and sequencing efforts started to identify more microRNAs. Among the microRNAs cloned there were several located at conserved genomic positions inside the *Hox* clusters. In vertebrates the *miR-10*, *miR-196* (16, 17) and later the *miR-615* (18, 19) families have been identified and in *Drosophila* the *miR-10* and *IAB-4* microRNAs (20) are present.
MicroRNAs are ~22nt RNAs that function in posttranscriptional gene silencing and participate in a secondary layer of genetic control over the primary transcriptional regulation (reviewed 21, 22, 23). MicroRNAs are processed from hairpin structures in longer precursor transcripts by the RNAse III enzymes *Drosha* and *Dicer*. *Drosha* frees the ~ 100nt stemloop from a longer pri-miRNA in the nucleus (fig.1). This pre-miRNA is exported to the cytoplasm where it is processed by *Dicer* into the mature ~22nt single stranded miRNA and incorporated in a gene silencing complex. In general only one of the two strands of the pre-miRNAs is incorporated into the silencing complex and acts as mature microRNA. MicroRNAs function by inhibiting translation and promoting messengerRNA decay and thereby prevent the production and accumulation of protein. Most vertebrate microRNA targets are targeted through imperfect matching target sites in their UTRs. Target sites usually have only partial complementarity and the specificity of the interaction is determined by nucleotide 2-7 of the microRNA which needs a perfect match with the target site for recognition and silencing. Many microRNA families are present in multiple copies and isoforms in vertebrate genomes.

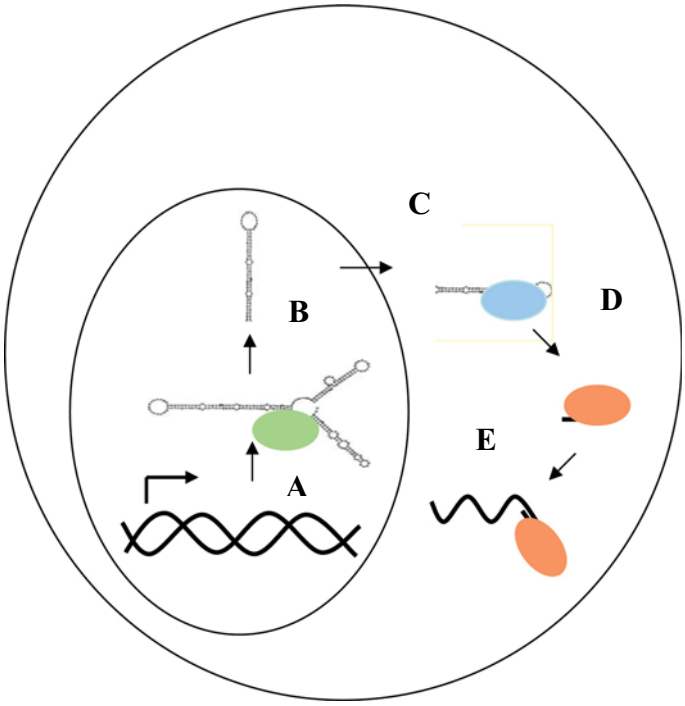


**Figure 1) Schematic representation of the microRNA metabolism and action within a cell.** The microRNA is produced in the nucleus (A), where it is further processed to the pre-miRNA by *Drosha* (green) (B), subsequently exported to the cytoplasm (C) where it is processed by *Dicer* (blue) to the final ~22nt mature single stranded miRNA that is incorporated in the RISC silencing complex (orange) (D). This complex can now recognize and silence target messengerRNAs (E).

**Box: microRNA nomenclature**
MicroRNA names are registered and assigned by miRBase (24) at the Sanger Centre (http://microrna.sanger.ac.uk/). The microRNA nomenclature for different isoforms and genomic paralogous copies in different species can be somewhat confusing. Due to the vertebrate genome duplications multiple copies and isoforms of a microRNA often exist within a single species. Isoforms are assigned different letters (eg miR-10**a**), if an additional isoform is discovered it will simply receive the next letter in the alphabet. Multiple genomic copies of one isoform within a species are distinguished by assigning a number suffix (eg miR-10b-**2**). The naming of the different microRNA isoforms and genomic copies is thus based on intra species sequence comparison and order of discovery but not on comparison of interspecies sequences or genomic homology. This means that when comparing different species, microRNAs carrying the same number suffix do not necessarily correspond to orthologous genes. In the same way, isoforms denoted by the same letter may also not correspond to identical sequences and/ or closest orthologues. Xenopus tropicalis *xtr-miR-10*c for instance differs in sequence from Zebrafish *dre-miR-10c* (Table 1). These two genes also do not represent closest orthologues since x*tr-miR-10c* is located in the *HoxC* cluster and is thus closest related to *dre-miR-10b-1* in the Zebrafish *HoxCa* cluster and *dre-miR-10c* is located in the *HoxBa* cluster and would thus correspond to *xtr-mir-10a* in the *HoxB* cluster. NB in this case the isoforms contain relatively recent lineage specific mutations and thus also differ in sequence from their true orthologues!! As a result miRNA names alone cannot be relied upon to convey complex inter-species relationships. When working with multiple isoforms and species this makes it necessary to carefully check miRbase sequences and locations within genomes to identify true othologues genes.

Isoforms usually differ in a few nucleotides outside the seed sequence and these mutations are not believed to result in any relevant change in target gene recognition. The homology between the microRNA and target sites makes it possible to predict putative target genes on basis of the seed sequence. Usually target RNAs are found to contain multiple target sites in close proximity of each other.

Since the discovery of microRNAs within the *Hox* clusters introduces new players into one of the best studied and most fascinating developmental systems, it has already excited many researchers. In this review, I discuss the current status of knowledge with respect to the connection between *Hox* genes and microRNAs. I give a detailed overview and annotation of the microRNAs present in the *Hox* clusters of the most important experimental model systems and correct some of the previously reported annotations. I discuss the literature with respect to the expression and developmental roles of these microRNAs.

## The microRNAs in the *Hox* clusters

Here I present an overview of the presence of microRNAs in the metazoan *Hox* clusters. In figure 2 a cladogram depicts the genomic configuration of the *Hox* cluster and shows the presence of the *Hox* microRNAs in several evolutionary or experimentally relevant taxa. In Table 1 the sequences of the *miR-10* and *miR-196* microRNAs in the most important vertebrate experimental model systems are listed per *Hox* cluster.

***miR-10***

*MiR-10* is one of the most conserved and ancient metazoan microRNAs. It is one of the 3 microRNAs identified in the cnidarian Sea Anemone (*Nematostella vectensis*) and therefore predates the microRNA radiation associated with the rise of the bilateralia (25). *MiR-10* has a significant homology with *miR-100/miR-99*, one of the other ancient microRNAs, and is thought to have evolved from it by a nucleotide insertion into the seed (fig2A). Since *miR-10* and *miR-100//miR-99* differ in their seed sequence they are predicted to have affinity for different target sites and most likely they will have different biological functions.

In vertebrates, amphioxus and arthropods the *miR-10* family is associated with the *Hox-4* paralogue *Hox* genes (26). In amniotes (mammals, birds) *miR-10* is present in two isoforms: *miR-10a* and *miR-10b* which are associated with the 5' *HoxB4* and *HoxD4* genomic region. In the anamniote Southern clawed frog (*Xenopus tropicalis*) an additional divergent *miR-10c* copy is present 5' of *HoxC4* (26). In the Teleost Zebrafish (*Danio rerio*) *miR-10* is present in 4 isoforms and 5 copies; *miR-10a* is located in the *HoxBb* cluster at a 4 paralogue position, *miR-10b-2* 5' of *HoxC4a*, *miR-10b-1* 5' of *HoxD4a*, *miR-10c* 5' of *HoxB4a* and *miR-10d* is located in the 8.1kb intergenic region between *lunapark* and *MTX2* which represents the remainders of the degenerated *HoxDb* cluster (27).

*MiR-10* has not been identified in the primitive proto-vertebrate *Ascidians Ciona savignii*, *Ciona intestinalis* and *Oikopleura dioica* by either bioinformatics means or wet lab experiments (25, 26) and can be assumed to have been lost. The *Ascidian* lineages also show severe *Hox* cluster fragmentation and the *Hox* genes have partially been lost and are present at dispersed locations in the genome (28, 29). In arthropods (*Drosophila sp.*, *Aedes Aegyptii* ) *miR-10* is located in the homologous region between the *Hox-4* homologue *deformed* (*dfd*) and the *Hox-5 sex combs reduced* (*scd*). *MiR-10* was originally not identified in *Caenorhabditis elegans* and *Caenorhabditis brigsae* (25) but *miR-57* in these species has high sequence homology (fig.2B) and an identical seed sequence and thus likely represents its ortholog (4).This is also supported by the absence of a 'true' *miR-57* microRNA gene outside the *Caenorhabditis* lineages. In *C. elegans* the organization of the *Hox* clusters is highly derived (30) and they lack the 2-4 paralogue group genes and *miR-57* is located on a different chromosome (4).

In the unusual sea urchin (*Strongylocentrotus purpureus*) *Hox* cluster *miR-10* is present 5' of the *Hox-3* paralogue (25, 31). *MiR-10* has recently also been cloned in the planarian (*Schmidtea mediterranen*) (32). No public genome assembly is available for this species so it was not possible to determine the genomic location of *miR-10*.

BLAST of the sea anemone *miR-10* precursor sequence locates it on a genomic contig also containing a *pax* homeodomain gene (NVHD074-paired class homeobox protein) but no true *Hox* gene. Although *Nemostella* is the most primitive species in which *Hox* genes have been identified the partially clustered state of the *Hox* genes is derived (33, 34) and it is not possible to tell whether the absence of linkage with *miR-10* represents an ancestral state in this species.

***miR-196***

*miR-196* is located between *Hox-9* and *Hox-10* genes or upstream of *Hox-9* in the *HoxA*, *HoxB* and *HoxC* cluster and no homolog has been detected in the *HoxD* cluster (25, 35). *MiR-196* is absent from ascidians, amphioxus and more distantly related organisms (4, 25) and appears to be a vertebrate specific microRNA. The most primitive species in which *miR-196* has been identified is the agnathan lamprey (*Petromyzon marinus*) (25). Its relationship to other microRNAs is unclear but it has been suggested to be distantly related to *let-7* (16). In human and mouse *miR-196* is present in 2 isoforms and 3 genomic copies; *miR-196a-1* is located in the *HoxB*, *miR-196a-2* in the *HoxC* and *miR-196b* in the *HoxA* cluster. In *Xenopus tropicalis miR-196* is present in the same *Hox* clusters but in this species there are 3 isoforms, with a previously undescribed *miR-196c* isoforms present in the *HoxC* cluster. This isoform is interesting since it carries a mutation in the seed compared to all other known *miR-196* sequences (fig. 2A, D and Table 1). As this mutation is located in the seed it is expected to alter target gene specificity. It is thus possible that this microRNA represents a functional evolutionary acquisition that contributed to the specification of the anuran lineages. In Zebrafish *miR-196* has been reported in two isoforms from the *HoxAa*, *HoxCa* and *HoxCb* clusters and was presumed to be absent from the *HoxB* clusters (25). I however locate the Zebrafish miRBase *miR-196b* sequence in the *HoxBa* cluster and cannot find a copy in the *HoxCb* cluster.

***miR-615***

*MiR-615* has sofar only been cloned in mouse and human (18, 19) and is located in the intron of *HoxC5*. I was not able to identify this microRNA in any species outside of the eutherian (placental) mammals. In *Xenopus tropicalis* where the *HoxC5* genomic sequence is well assembled its presence can be excluded. Although the genomic coverage from chicken, platypus (*Ornithorhynchus anatinus*) and opossum (*Monodelphis domestica*) is not complete its absence in all three species suggests that this microRNA is a very recent addition to the repertoire of regulatory sequences in the *Hox* clusters. It is noteworthy that the 5'and 3' ends of the *miR-615* precursor contain complementary low complexity repeat sequences and that the microRNA gene may have formed by an accidental localization of the two sequences in each others close vicinity (fig.2C).

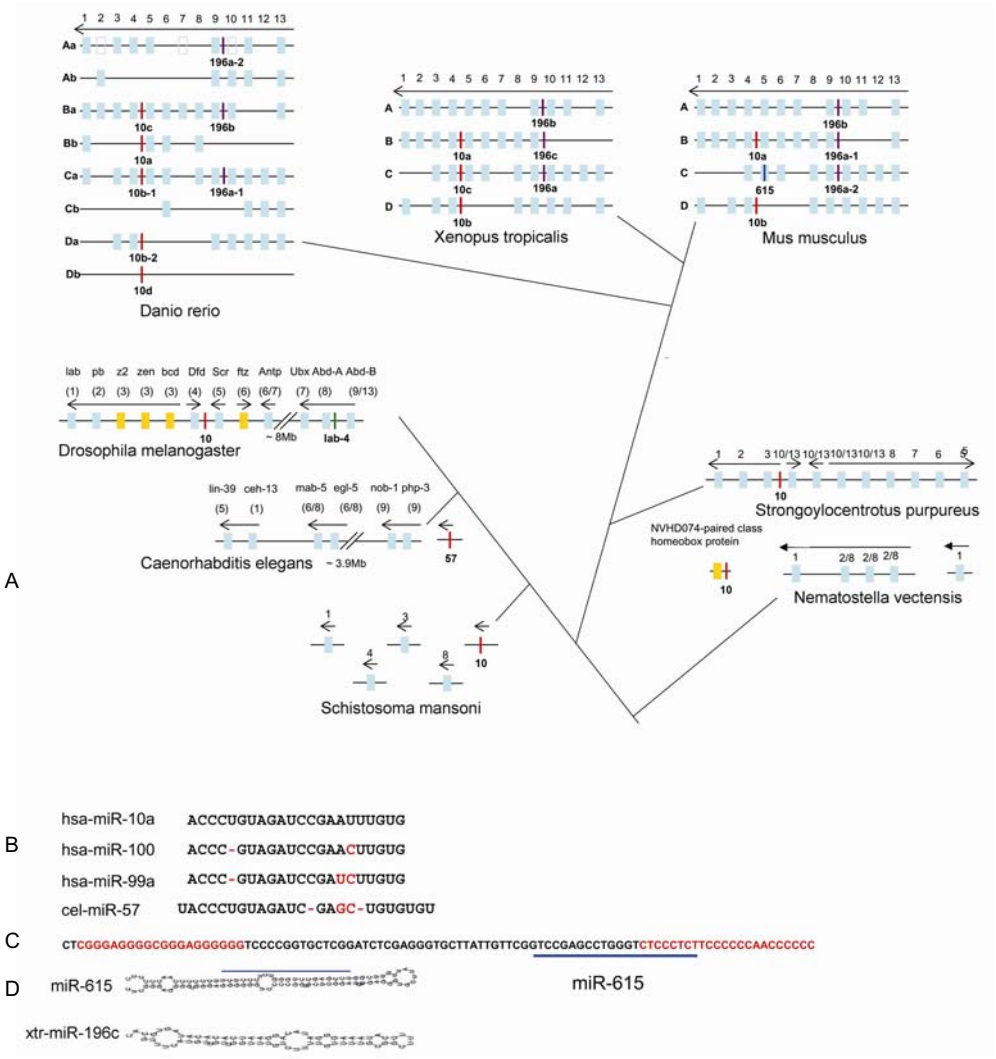


**Figure 2) Representation of the *Hox* related microRNAs throughout the metazoan kingdom.**
(A) Cladogram showing the presence of *miR-10*, *miR-196*, *miR-615*, *IAB-4* and *miR-57* in relation to the *Hox* clusters in Zebrafish (*Danio rerio*), *Xenopus tropicalis*, mouse (*Mus musculus*), fruit fly (*Drosophila melanogaster*), *Caenorhabditis elegans*, Sea urchin (*Strongylocentrotus purpureus*) and Sea anemone (*Nematostella vectensis*). (B) Relationship between *miR-10*, *miR-100/99* and *miR-57*. (C) Sequence and predicted RNA folding of *pre-miR-615*. The region of the mature *miR-615* is underlined blue. Low complexity repeat sequences are marked in red. (D) Predicted folding of the *miR-196c* pre-miRNA.

***IAB-4***

The *IAB-4* microRNA is present in a longer noncoding RNA transcribed from a cis-regulatory region between *AbdB* and *AbdA* in the insect (*Drosophlia sp.*, *Anopheles gambiae*, *Aedes aegyptii*, *Apios melifera*, *Tribolium castaneum*) *Bithorax* complex (20, 36). This region is analogous to the region where *miR-196* is present in the vertebrate *Hox* clusters. Sequence comparison however shows no significant homology between the two microRNAs and they are unlikely to represent orthologs (4). Two mature microRNAs are produced from opposite strands of the *IAB-4* precursor stemloop, *IAB-4-5p* and *IAB-4-3p*. Both of the strands are 100% conserved among a wide range of insects representing a 400 myrs evolutionary distance (36), which is very unusually for a precursor sequence outside of the functional microRNA. In this case it strongly suggests a function for the microRNAs produced from either strand of the stemloop.

## Expression of the *Hox* related microRNAs

Since no functional in situ technique was yet available in vertebrates the expression patterns for *miR-10* and *miR-196* in mice were initially inferred from transgenic sensor lines (37). These lines have an integrated constitutive active LacZ gene with multimerized microRNA target sites in its UTR (a 'sensor construct'). In regions where the microRNA is expressed the LacZ gene is repressed. In the transgenic embryos both microRNAs show a pattern essentially similar to their associated paralogue group *Hox-4* and *Hox-9* genes, implying that the microRNA and *Hox* genes in vertebrates are under shared transcriptional control. The presence of ESTs containing both *miR-196* and *HoxB9* supports this further (25). After the development of an efficient microRNA in situ hybridization technique based on locked nucleic acid (LNA) probes, the sensor data were confirmed in mouse (38), chicken (39) and Zebrafish (27). However, in vertebrates no precise somitic borders of expression have been reported for either microRNA so far. Our own expression data in Zebrafinch (*Taeniopygia guttata*) in situ hybridization show that in this species *miR-10a* is strongly expressed throughout the neural tube from a postotic hindbrain level and in the paraxial mesoderm much more weakly from the somite 7 level (fig. 3B ); *miR-196a* (fig.3A) is expressed in the neural tube with an anterior boundary at somite 23 and it has a stronger expression in the paraxial mesoderm reaching till somite 28.

| | | | | | |
|---|---|---|---|---|---|
| **HoxA** | | | | hs-miR-196b | UAGGUAGUUUCAAGUUGUUGGG |
| | | | | mm-miR-196b | UAGGUAGUUUCAAGUUGUUGGG |
| | | | | xtr-miR-196b | UAGGUAGUUUCAAGUUGUUGGG |
| | | | | dre-miR-196a-2 | UAGGUAGUUUCAUGUUGUUGGG |
| **HoxB** | hs-miR-10a | UACCCUGUAGAUCCGAAUUUGUG | | hs-miR-196a-1 | UAGGUAGUUUCAUGUUGUUGGG |
| | mm-miR-10a | UACCCUGUAGAUCCGAAUUUGUG | | mm-miR-196a-1 | UAGGUAGUUUCAUGUUGUUGGG |
| | xtr-miR-10a | UACCCUGUAGAUCCGAAUUUGUG | | xtr-miR-196c | UAGGCAGUUUCAUGUUGUUGG |
| | dre-miR-10c | UACCCUGUAGAUCCGGAUUUGU | | dre-miR-196b | UAGGUAGUUUCCUGUUGUUGG |
| | dre-miR-10a | UACCCUGUAGAUCCGAAUUUGU | | | |
| **HoxC** | xtr-miR-10c | CACCCUGUAGAAUCGAAUUUGU | | hs-miR-196a-2 | UAGGUAGUUUCAUGUUGUUGG |
| | dre-miR-10b-2 | UACCCUGUAGAACCGAAUUUGUG | | mm-miR-196a-2 | UAGGUAGUUUCAUGUUGUUGG |
| | | | | xtr-miR-196a | UAGGUAGUUUCAUGUUGUUGG |
| | | | | dre-miR-196a-1 | UAGGUAGUUUCAUGUUGUUGGG |
| **HoxD** | hs-miR-10b | UACCCUGUAGAACCGAAUUUGU | | | |
| | mm-miR-10b | CCCUGUAGAACCGAAUUUGUGU | | | |
| | xtr-miR-10b | UACCCUGUAGAACCGAAUUUGU | | | |
| | dre-miR-10b-1 | UACCCUGUAGAACCGAAUUUGUG | | | |
| | dre-miR-10d | UACCCUGUAGAACCGAAUGUGUG | | | |

**Table 1) MiRBase sequences of the micoRNAs in the *Hox* cluster in human, mouse, Xenopus tropicalis and Zebrafish organized by cluster. Mutations creating different isoforms between among closest orthologs are marked with color, red representing the minority isoforms.**

It is also expressed in the fore and hind limb. In the early stages *miR-196* is expressed in the hind but not fore limb field.

Depending on the place of the mismatch, LNA probes can exhibit single nucleotide resolution and are able to distinguish between the different isoforms. In Zebrafish the different *miR-10* isoforms have very similar expression patterns although there are also clear differences (27). I now also have been able to confirm this further by in situ hybridization using ~100nt precursor sequences only.

*MiR-615* has until now only been cloned from mouse colorectal tissue and the complete spatial embryonic expression pattern is unknown. If no additional level of posttranscriptional regulation is involved it is very likely to have an expression pattern identical to that of *HoxC5*, since it is located in its intron. In *Drosophila*, expression of *miR-10* and *IAB-4* has been described using precursor in situ hybridization in embryos (40, 41). The *miR-10* in situ data reveal a remarkable dynamic expression pattern in later embryos, which at later stages is quite different from the *dfd* expression pattern suggesting that the genes are under separate controls. In *Drosophila*, *IAB-4* is expressed at a more posterior level than *miR-10*, consistent with its association with more posterior *Hox* genes (40).

In studying the expression of any microRNA in vertebrate development it may be appropriate to take notice that there is recent evidence for an embryonic regulatory switch at the level of precursor processing or accumulation. Mature microRNAs could not be detected in early mouse embryos although their precursor sequences were abundantly present (42). How this affects the expression of the *Hox* related microRNAs is so far unclear but it implies that they may only be present after a certain stage in development, although they and their accompanying *Hox* genes are transcribed earlier on.

## The biological function of the *Hox* related microRNAs.

The conservation of microRNAs within the *Hox* clusters and their expression in *Hox* like patterns is strongly suggestive for a role in axial patterning. The biological functions of the *Hox* linked microRNAs are so far still largely unclear though. Until now most data are available for the *miR-196* and *IAB-4* microRNAs. The *miR-196* microRNAs have been shown to target *HoxA7*, *HoxB8*, *HoxC8* and *HoxD8* in mammals and the target sites in at least *HoxB8* and *HoxC8* are conserved from Teleosts to Tetrapods (35). The regulation of *HoxB8* is mediated through *miR-196* directed cleavage within the target site, which has only a single mismatch with the microRNA. Cleavage of target transcripts by microRNAs is common in Arabidopsis but until now this represents the only reported vertebrate example. How a target messenger is silenced depends on the degree of complementarity to the microRNA. An almost perfect match results in cleavage and a less perfect match in translational silencing. The sequence of the target site in *HoxB8* has been 100% conserved during the 500 myrs separating Teleost from Tetrapods and this suggest that there is a strong selective force favoring cleavage over translational silencing.
Hornstein et al.(43) discovered a role for *miR-196* in the patterning of the mouse and chicken limb buds where it is involved in the differential interpretation of retinoic acid signaling by fore and hind limb. In wildtype embryos retinoic acid induces *HoxB8* expression in the developing fore limb buds but not in hind limbs. However, in conditional *Dicer* knockout embryos that lack all mature microRNAs in the limbs, retinoic acid results in *HoxB8* induction in both fore and hind limbs. Using microarrays, *miR-196* was shown to be expressed in the hind limb field only (see also *miR-196a* expression in the Zebrafinch fig.3A).

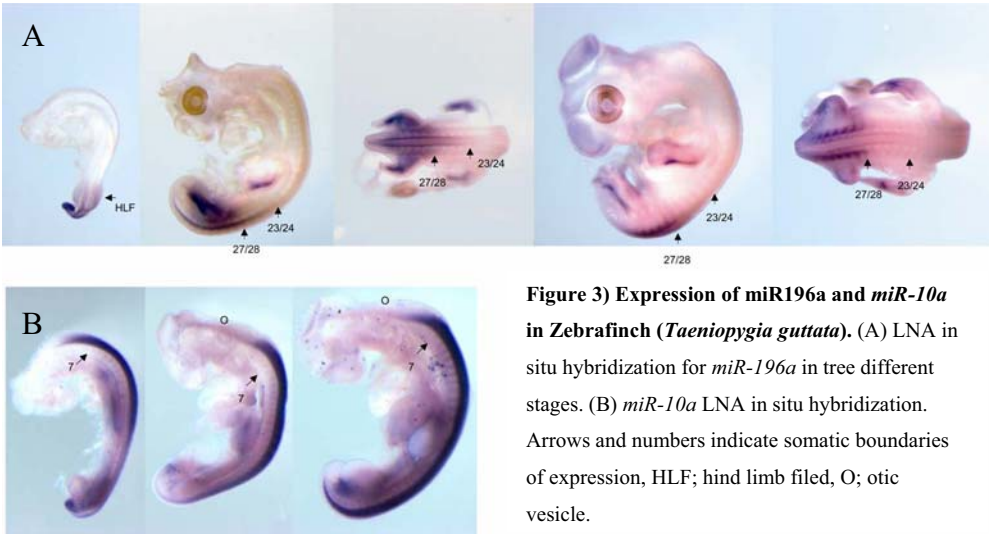


**Figure 3) Expression of miR196a and *miR-10a* in Zebrafinch (*Taeniopygia guttata*).** (A) LNA in situ hybridization for *miR-196a* in tree different stages. (B) *miR-10a* LNA in situ hybridization. Arrows and numbers indicate somatic boundaries of expression, HLF; hind limb filed, O; otic vesicle.

In chicken embryos infected with a *miR-196-1* containing RCAS virus, *HoxB8* expression is downregulated. In early chicken embryos *HoxB8* is expressed in the fore limb but not hind limb field and treatment with retinoic acid upregulates expression of *HoxB8* in fore limbs but not hind limbs. Interestingly, the absence of *miR-196* in the *Dicer* deficient mice does not, on its own, lead to ectopic expression of *HoxB8* in the hind limb but requires stimulation with retinoic acid. The primary level of *HoxB8* regulation thus seems to take place at the level of transcription. It is suggested by the authors (43) that the posttranscriptional repression of *HoxB8* in the hind limb may function as a failsafe system, protecting the developmental system from leaky gene transcription.

In the patterning of the axial mesoderm no putative role has yet been established for *miR-196* or its interaction with *Hox-8* genes. It seems plausible that *miR-196* downregulates *Hox-8* genes in the posterior part of the axis as well. Pre-dating the discovery of the microRNA, Bittner et al. (44) noticed a difference between *HoxC8* situ hybridization and immunolocalization patterns in *Xenopus laevis* and concluded that *HoxC8* is under translational control in the posterior part of its expression domain. Interesting and suggestive of multiple levels of control is that the identified mouse *HoxC8* enhancers are able to provide a correct anterior boundary of expression but result in an extended posterior

boundary (45). If indeed the absence of *Hox-8* genes from the posterior part of the embryos is (partially) due to posttranscriptional regulation, the phenotype of a complete *miR-196* loss of function in the paraxial mesoderm is likely to result in a (partial) lumbar to thoracic homeotic transformation, as observed in posterior overexpression of the *HoxC8* gene (46). The *IAB-4-5p* microRNA which is present in an analogues position in Drosophila has been shown to target the *Ultrabithorax* (*Ubx*) gene. *Ultrabithorax* plays a role in preventing the formation of wings from haltere structures. Mutations in *Ubx* result in the homeotic transformation where flies develop 2 instead of 1 pair of wings.
In *Drosophila* high levels of *IAB-4* expression are correlated to low levels of *Ubx* protein and ectopic expression prevents *Ubx* protein accumulation (36). When *IAB-4-5p* is ectopically expressed the halters develop the wing characteristic sensory hair rows, showing the effects of a classic homeotic transformation. Until now much less has been discovered about the biological role of *miR-10*. In mouse blood megakaryocyte differentiation, *miR-10* has been shown to be upregulated and to target *HoxA1* (47). The endogenous temporal expression patterns of *miR-10* and *HoxA1* in megakaryocytes show temporally mutually excluding patterns. I have found in Zebrafish that *miR-10* targets *Hox-1* and *Hox-3* genes and is necessary for the proper development of the posterior hindbrain (Woltering & Durston 2007, chapter this thesis). In *Drosophila miR-10* has been predicted to target *Sex Combs Reduced* (*Hox-5*) and in this case it also is a neighboring gene that is targeted (48).
*MiR-615* is the most recently discovered microRNA and no experimental data addressing its function have been published yet.

## A role in posterior prevalence?

The emerging picture from functional studies is that the vertebrate *Hox* microRNAs have target interactions with more anterior genes located within the *Hox* clusters themselves (fig. 4). As discussed above the anterior boundaries of *Hox* genes in general follow the rules of colinearity. *Hox* genes however have nested expression patterns and the posterior boundaries of *Hox* genes are less well defined and tend to overlap with the expression of more posterior genes.

Within the *Hox* code a simple hierarchy exists that determines which of a number of overlapping *Hox* genes has selector function. As a rule: where genes are coexpressed the posterior genes are dominant over anterior genes, which is known as 'posterior prevalence' (49).

Although this is a well established phenomenon there are some indications that deviations from this rule exist. In the case of the *miR-196*/*Hox-8* interaction, the experiments done by Pollock and colleagues, where ectopic expression of *HoxC8* in the mouse paraxial mesoderm leads to an anterior transformation, indicate that the proposed dominance of lumbar *Hox* genes does not completely abolish *Hox-8* function (46). This would explain why an extra regulatory mechanism is necessary to prevent accumulation of *Hox-8* proteins in the lumbar paraxial mesoderm.

## Indications for additional microRNAs in the *Hox* clusters

The vertebrate *Hox* clusters contain many transcribed noncoding intergenic regions including ultra conserved elements (50). The possibility exists that there are more microRNAs located within these regions. Directed bioinformatics searches to identify these by analyzing secondary RNA structure surrounding conserved elements have so far however not identified likely candidates outside the *miR-10* and *miR-196* families (Mainguy, Woltering, Durston and others unpublished). The presence of more conserved microRNA families therefore does not seem likely. This in silico footprinting is however largely based on sequence homology and lineage specific micoRNAs like *miR-615* will therefore not be identified. Species specific microRNA cloning and sequencing efforts are therefore necessary to identify possible additional lineage specific *Hox* microRNAs.

## Other *Hox*-microRNA interactions

The *miR-181* microRNA which is strongly upregulated in regenerating muscles has been shown to target *HoxA11* during mammalian myoblast differentiation (51). *MiR-181* was knocked down by transfection of a *miR-181* antisense LNA oligo in differentiating C2C12 cells, which prevented the upregulation of muscle differentiation markers. In wildtype differentiating muscles *HoxA11* is downregulated and functions as an inhibitor of this

process. In the absence of *miR-181 HoxA11* downregulation is partially reduced. Neither upregulation of *miR-181* nor downregulation of *HoxA11* alone triggers muscle differentiation however, showing that the differentiation switch is operated through a intricate network involving multiple microRNA-target interactions.

Evidence for posttrancriptional regulation of additional *Hox* genes has been found in mouse for *HoxB4* and *HoxC6*, where expression of these genes has been characterized by both in situ hybridization and immunolocalization. *HoxB4* message can be located throughout the mouse neural tube. Howerer, antibody staining shows that the protein is absent from the posterior neural tube (52).

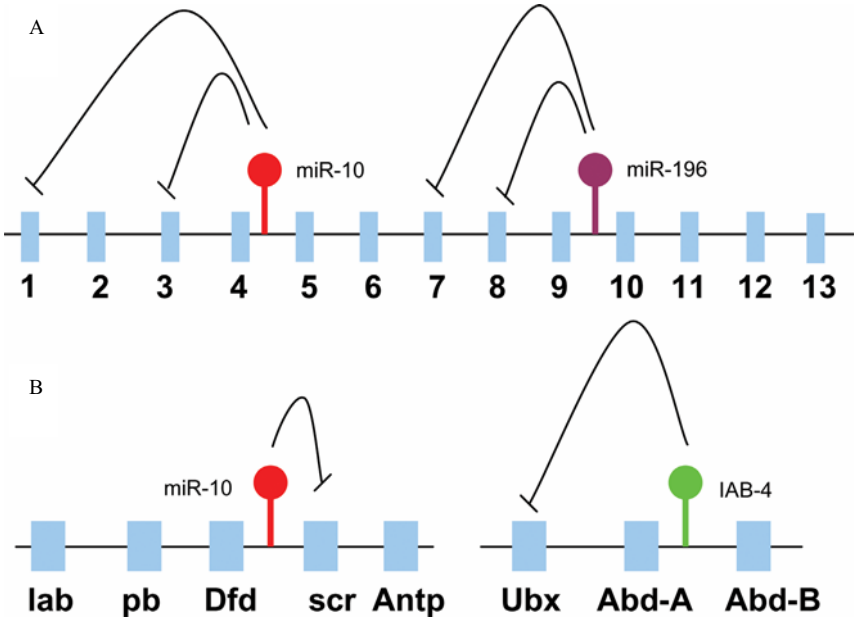


**Figure 4) Posttranscriptional interactions within the *Hox* clusters.** A) In the vertebrate *Hox* clusters there is evidence for interactions between *miR-10* and *Hox-1* and *Hox-3* paralogous genes. *MiR-196* has been conclusively shown to target *Hox-A7* and *Hox-8* genes. B) In Drosophila, *IAB-4* targets *Ubx* and the *scr* sequence has been shown to contain *miR-10* target sites.

*HoxC6* in chicken and mouse is expressed in both for and hind limb; *HoxC6* protein cannot be detected in the hind limb though (53). It is currently unknown whether these posttranscriptional effects are mediated through microRNAs in these two cases.

## Concluding remarks

With the discovery of the microRNAs in the *Hox* clusters, a new level of regulation has been introduced in one of the most exciting and best understood metazoan developmental loci. The functions of the endogenous microRNAs are still largely unresolved, mainly because of the absence of loss of function data. The presence of multiple, likely redundant copies in the vertebrate genomes probably necessitates the creation of double and triple mouse knockouts to reveal the full impact of the microRNAs on the patterning of the axis. Since oligo morpholinos can target and inhibit the function of complete microRNA families, knockdown in Zebrafish and amphibians may provide an elegant solution to circumvent this problem. In Zebrafish and *Xenopus* it is very well possible to study the patterning and development of the nervous system. Characterization of the development of the paraxial mesoderm, however, is hampered in these species as the axial skeleton differentiates very late. Knockdown in the Salamander *Triturus* or *Ambystoma* model systems, which have a much earlier development of the axial skeleton may therefore be used as an alternative.

I predict that within the *Hox* clusters the *miR-10* and *miR-196* microRNAs play a role in the coordination of gene dominance effects that are not covered by the mechanisms of posterior prevalence. The posterior prevalence in the *Hox* code is believed to function through protein-protein interactions that establish a hierarchy in the DNA binding. This leads to different transctriptional responses of the genes. Several homeodomain proteins have however been reported to influence translation (53, 54), and based on conserved domain structures many more *Hox* genes are predicted to do so as well. Outside of the traditional function of *Hox* genes as transcription factors, different dominance relationships may exist. It is not unlikely that the posttranscriptional repression of microRNAs may play a role to influence these. The genes in the vertebrate *Hox* clusters and their functions are much conserved. The variation in the *Hox* clusters, like the acquisition of *miR-615* within mammals and the presence of an aberrant *miR-196c* in *Xenopus tropicalis,*

is therefore somewhat surprising and it is enticing to speculate that these differences may have played a role in evolution. This pattern of variation between closely related species may be a common theme in microRNA networks. Berezikov et al. (55) for instance reported the presence of human and chimp specific microRNAs expressed in the brain.
It is assumed that *miR-196* and *IAB-4-5p* are unrelated; still they have evolved at analogues positions in the *Hox* clusters with analogues functions and represent a striking case of convergent evolution of gene regulation within one locus. Such parallel developments suggest that these interactions add something essential to the coordination of the genetic network within the *Hox* cluster that is not easily achieved by other means of regulation.
It is very exciting that a *Hox* linked microRNA *miR-10* belongs to the group of 3 most ancient microRNAs present in the bilateralia lineage and study in *Nematostella* (as also suggested by Prochnik et al. (25)) is likely to provide insights, not only in the evolution of its own function, but also in the evolution of the intricate network of the current vertebrate microRNA target relationships.

*Acknowledgements: I would like to thank Tony Durston, Yanju Zhang and Flip Woltering for carefully reading the manuscript and pointing out mistakes; I am highly indebted to Merijn de Bakker and Michael Richardson for the Zebrafinch embryos.*

# References

1) Carroll, R. *Vertebrate Paleontology and Evolution* (Freeman, New York, (1988).

2) Dasen JS, Tice BC, Brenner-Morton S, Jessell TM.
A *Hox* regulatory network establishes motor neuron pool identity and target-muscle connectivity.
Cell. 2005 Nov 4;123(3):477-91.

3) Krumlauf R.
*Hox* genes in vertebrate development.
Cell. 1994 Jul 29;78(2):191-201.

4) Pearson JC, Lemons D, McGinnis W.
Modulating *Hox* gene functions during animal body patterning.
Nat Rev Genet. 2005 Dec;6(12):893-904.

5) Wellik DM, Capecchi MR.
*Hox*10 and *Hox*11 genes are required to globally pattern the mammalian skeleton.
Science. 2003 Jul 18;301(5631):363-7.

6) McClintock JM, Kheirbek MA, Prince VE.
Knockdown of duplicated zebrafish *Hox*b1 genes reveals distinct roles in hindbrain patterning and a novel mechanism of duplicate gene retention.
Development. 2002 May;129(10):2339-54.

7) Hooiveld MH, Morgan R, in der Rieden P, Houtzager E, Pannese M, Damen K, Boncinelli E, Durston AJ.
Novel interactions between vertebrate *Hox* genes.
Int J Dev Biol. 1999;43(7):665-74.

8) McNulty CL, Peres JN, Bardine N, van den Akker WM, Durston AJ.
Knockdown of the complete *Hox* paralogous group 1 leads to dramatic hindbrain and neural crest defects.
Development. 2005 Jun;132(12):2861-71.

9) Burke AC.
*Hox* genes and the global patterning of the somitic mesoderm.
Curr Top Dev Biol. 2000;47:155-81.

10) Burke AC, Nelson CE, Morgan BA, Tabin C.
*Hox* genes and the evolution of vertebrate axial morphology.
Development. 1995 Feb;121(2):333-46.

11) Garcia-Fernandez J.
The genesis and evolution of homeobox gene clusters.
Nat Rev Genet. 2005 Dec;6(12):881-92.

12) Lemons D, McGinnis W.
Genomic evolution of *Hox* gene clusters.
Science. 2006 Sep 29;313(5795):1918-22.

13) Hoegg S, Meyer A.
*Hox* clusters as models for vertebrate genome evolution.
Trends Genet. 2005 Aug;21(8):421-4.

14) Amores A, Force A, Yan YL, Joly L, Amemiya C, Fritz A, Ho RK, Langeland J, Prince V, Wang YL, Westerfield M, Ekker M, Postlethwait JH
Zebrafish *Hox* clusters and vertebrate genome evolution.
Science. 1998 Nov 27;282(5394):1711-

15) Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G.
The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans.
Nature. 2000 Feb 24;403(6772):901-6.

16) Lim LP, Glasner ME, Yekta S, Burge CB, Bartel DP.
Vertebrate microRNA genes.
Science. 2003 Mar 7;299(5612):1540.

17) Lagos-Quintana M, Rauhut R, Meyer J, Borkhardt A, Tuschl T.
New microRNAs from mouse and human.
RNA. 2003 Feb;9(2):175-9

18) Cummins JM, He Y, Leary RJ, Pagliarini R, Diaz LA Jr, Sjoblom T, Barad O, Bentwich Z, Szafranska AE, Labourier E, Raymond CK, Roberts BS, Juhl H, Kinzler KW, Vogelstein B, Velculescu VE.
The colorectal microRNAome.
Proc Natl Acad Sci U S A. 2006 Mar 7;103(10):3687-92.

19) Mineno J, Okamoto S, Ando T, Sato M, Chono H, Izu H, Takayama M, Asada K, Mirochnitchenko O, Inouye M, Kato I.
The expression profile of microRNAs in mouse embryos.
Nucleic Acids Res. 2006 Mar 31;34(6):1765-71.

20) Aravin AA, Lagos-Quintana M, Yalcin A, Zavolan M, Marks D, Snyder B, Gaasterland T, Meyer J, Tuschl T.
The small RNA profile during Drosophila melanogaster development.
Dev Cell. 2003 Aug;5(2):337-50..

21) He L, Hannon GJ.
MicroRNAs: small RNAs with a big role in gene regulation.
Nat Rev Genet. 2004 Jul;5(7):522-31.

22) Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R.
Control of translation and mRNA degradation by miRNAs and siRNAs.
Genes Dev. 2006 Mar 1;20(5):515-24.

23) Wienholds E, Plasterk RH.
MicroRNA function in animal development.
FEBS Lett. 2005 Oct 31;579(26):5911-22.

24) Griffiths-Jones S.
miRBase: the microRNA sequence database.
Methods Mol Biol. 2006;342:129-38.

25) Prochnik SE, Rokhsar DS, Aboobaker AA.
Evidence for a microRNA expansion in the bilaterian ancestor.
Dev Genes Evol. 2007 Jan;217(1):73-7.
26) Tanzer A, Amemiya CT, Kim CB, Stadler PF.
Evolution of microRNAs located within *Hox* gene clusters.
J Exp Zoolog B Mol Dev Evol. 2005 Jan 15;304(1):75-85.

27) Woltering JM, Durston AJ.
The zebrafish *Hox*Db cluster has been reduced to a single microRNA.
Nat Genet. 2006 Jun;38(6):601-2.

28) Seo HC, Edvardsen RB, Maeland AD, Bjordal M, Jensen MF, Hansen A, Flaat M, Weissenbach J, Lehrach H, Wincker P, Reinhardt R, Chourrout D.
*Hox* cluster disintegration with persistent anteroposterior order of expression in Oikopleura dioica.
Nature. 2004 Sep 2;431(7004):67-71.

29) Ikuta T, Yoshida N, Satoh N, Saiga H.
Ciona intestinalis *Hox* gene cluster: Its dispersed structure and residual colinear expression in development.
Proc Natl Acad Sci U S A. 2004 Oct 19;101(42):15118-23.

30) Aboobaker A, Blaxter M
*Hox* gene evolution in nematodes: novelty conserved.
Curr Opin Genet Dev. 2003 Dec;13(6):593-8.

31) Cameron RA, Rowen L, Nesbitt R, Bloom S, Rast JP, Berney K, Arenas-Mena C, Martinez P, Lucas S, Richardson PM, Davidson EH, Peterson KJ, Hood L.
Unusual gene order and organization of the sea urchin *Hox* cluster.
J Exp Zoolog B Mol Dev Evol. 2006 Jan 15;306(1):45-58.

32) Palakodeti D, Smielewska M, Graveley BR.
MicroRNAs from the Planarian Schmidtea mediterranea: a model system for stem cell biology.
RNA. 2006 Sep;12(9):1640-9.

33) Chourrout D, Delsuc F, Chourrout P, Edvardsen RB, Rentzsch F, Renfer E, Jensen MF, Zhu B, de Jong P, Steele RE, Technau U.
Minimal Proto*Hox* cluster inferred from bilaterian and cnidarian *Hox* complements.
Nature. 2006 Aug 10;442(7103):684-7.

34) Ryan JF, Mazza ME, Pang K, Matus DQ, Baxevanis 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 Jan 24;2:e153.

35) Yekta S, Shih IH, Bartel DP.
MicroRNA-directed cleavage of *HOX*B8 mRNA.
Science. 2004 Apr 23;304(5670):594-6

36) Ronshaugen M, Biemar F, Piel J, Levine M, Lai EC.
The Drosophila microRNA iab-4 causes a dominant homeotic transformation of halteres to wings.
Genes Dev. 2005 Dec 15;19(24):2947-52.

37) Mansfield JH, Harfe BD, Nissen R, Obenauer J, Srineel J, Chaudhuri A, Farzan-Kashani R, Zuker M, Pasquinelli AE, Ruvkun G, Sharp PA, Tabin CJ, McManus MT.
MicroRNA-responsive 'sensor' transgenes uncover *Hox*-like and other developmentally regulated patterns of vertebrate microRNA expression.
Nat Genet. 2004 Oct;36(10):1079-83.

38) Kloosterman WP, Wienholds E, de Bruijn E, Kauppinen S, Plasterk RH.
In situ detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes.
Nat Methods. 2006 Jan;3(1):27-9.

39) Darnell DK, Kaur S, Stanislaw S, Konieczka JK, Yatskievych TA, Antin PB.
MicroRNA expression during chick embryo development.
Dev Dyn. 2006 Nov;235(11):3156-65.

40) Aboobaker AA, Tomancak P, Patel N, Rubin GM, Lai EC.
Drosophila microRNAs exhibit diverse spatial expression patterns during embryonic development.
Proc Natl Acad Sci U S A. 2005 Dec 13;102(50):18017-22.

41) Kosman D, Mizutani CM, Lemons D, Cox WG, McGinnis W, Bier E.
Multiplex detection of RNA expression in Drosophila embryos.
Science. 2004 Aug 6;305(5685):846.

42) Thomson JM, Newman M, Parker JS, Morin-Kensicki EM, Wright T, Hammond SM.
Extensive post-transcriptional regulation of microRNAs and its implications for cancer.
Genes Dev. 2006 Aug 15;20(16):2202-7.

43) Hornstein E, Mansfield JH, Yekta S, Hu JK, Harfe BD, McManus MT, Baskerville S, Bartel DP, Tabin CJ.
The microRNA miR-196 acts upstream of *Hox*b8 and Shh in limb development.
Nature. 2005 Dec 1;438(7068):671-4.

44) Bittner D, De Robertis EM, Cho KW.
Characterization of the Xenopus *Hox* 2.4 gene and identification of control elements in its intron.
Dev Dyn. 1993 Jan;196(1):11-24.

45) Shashikant CS, Ruddle FH.
Combinations of closely situated cis-acting elements determine tissue-specific patterns and anterior extent of early *Hox*c8 expression.
Proc Natl Acad Sci U S A. 1996 Oct 29;93(22):12364-9.

46) Pollock RA, Jay G, Bieberich CJ.
Altering the boundaries of *Hox*3.1 expression: evidence for antipodal gene regulation.
Cell. 1992 Dec 11;71(6):911-23.

47) Garzon R, Pichiorri F, Palumbo T, Iuliano R, Cimmino A, Aqeilan R, Volinia S, Bhatt D, Alder H, Marcucci G, Calin GA, Liu CG, Bloomfield CD, Andreeff M, Croce CM.
MicroRNA fingerprints during human megakaryocytopoiesis.
Proc Natl Acad Sci U S A. 2006 Mar 28;103(13):5078-83.

48) Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS.
MicroRNA targets in Drosophila.
Genome Biol. 2003;5(1):R1.

49) Kmita M, Duboule D.
Organizing axes in time and space; 25 years of colinear tinkering.
Science. 2003 Jul 18;301(5631):331-3.

50) Sessa L, Breiling A, Lavorgna G, Silvestri L, Casari G, Orlando V.
Noncoding RNA synthesis and loss of Polycomb group repression accompanies the colinear activation of the human *HOX*A cluster.
RNA. 2007 Feb;13(2):223-39.

51) Naguibneva I, Ameyar-Zazoua M, Polesskaya A, Ait-Si-Ali S, Groisman R, Souidi M, Cuvellier S, Harel-Bellan A.
The microRNA miR-181 targets the homeobox protein *Hox*-A11 during mammalian myoblast differentiation.
Nat Cell Biol. 2006 Mar;8(3):278-84.

52) Brend T, Gilthorpe J, Summerbell D, Rigby PW.
Multiple levels of transcriptional and post-transcriptional regulation are required to define the domain of *Hox*b4 expression.
Development. 2003 Jun;130(12):2717-28.

53) Nelson CE, Morgan BA, Burke AC, Laufer E, DiMambro E, Murtaugh LC, Gonzales E, Tessarollo L, Parada LF, Tabin C.
Analysis of *Hox* gene expression in the chick limb bud.
Development. 1996 May;122(5):1449-66

54) Topisirovic I, Kentsis A, Perez JM, Guzman ML, Jordan CT, Borden KL.
Eukaryotic translation initiation factor 4E activity is modulated by *HOX*A9 at multiple levels.
Mol Cell Biol. 2005 Feb;25(3):1100-12.

55) Brunet I, Weinl C, Piper M, Trembleau A, Volovitch M, Harris W, Prochiantz A, Holt C.
The transcription factor Engrailed-2 guides retinal axons.
Nature. 2005 Nov 3;438(7064):94-8.

56) Berezikov E, Thuemmler F, van Laake LW, Kondova I, Bontrop R, Cuppen E, Plasterk RH.
Diversity of microRNAs in human and chimpanzee brain.
Nat Genet. 2006 Dec;38(12):1375-7.