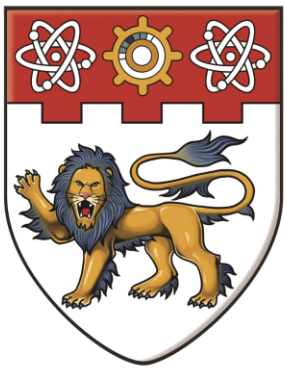

NANYANG TECHNOLOGICAL UNIVERSITY

# CHARACTERIZATION OF THE MURINE *HOXD4* AND HOX-ASSOCIATED *MIR-10B* TRANSCRIPTS

**PHUA SZE LYNN CALISTA**

**SCHOOL OF BIOLOGICAL SCIENCES**

**2012**

# CHARACTERIZATION OF THE MURINE *HOXD4* AND HOX-ASSOCIATED *MIR-10B* TRANSCRIPTS

**PHUA SZE LYNN CALISTA**

School of Biological Sciences

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Doctor of Philosophy

**2012**

## ACKNOWLEDGEMENTS

I am very much indebted to my supervisor Professor Mark Featherstone for his exceptionally strong support, keen guidance and the conceptualization of a most interesting and intellectually stimulating project. I am thankful to be one of the privileged few to have had Mark as a scientific mentor. It has indeed been a tremendously enriching and rewarding experience.

My heartfelt thanks go to the members of my Thesis Advisory Committee, Assoc Prof Chen Ken-Shiung and Assoc Prof Lin Chun Ling, Valerie, for their valuable input and time. I also thank and acknowledge the following people for their various contributions: Assoc Prof Peter Dröge and Feng Shu, for their help with the ribosome experiments; Asst Prof Zhang Li-Feng and Shao Yu for their assistance with the RNA FISH experiments; Dr Thomas Lufkin and V Sivakamasundari for performing the mouse *in situ* hybridizations; Cai Xiaohan for her work on *Hoxb4*/*miR-10a*; Asst Prof Eugene Makeyev for helpful discussions; and all staff and colleagues at the School of Biological Sciences, for help rendered in countless ways.

I also wish to express my warmest thanks to all past and present members of the Featherstone lab for their camaraderie and all the memorable times. I could not have asked for a more supportive, generous and amazing group of colleagues and friends. To Shen Hui, Ragini, Anusha, Eva, Yvonne, Fiona, Ser Yeng, Ravi, Lawrence, and the numerous friends and students who have helped one way or another, thank you all very much.

Above all, to my most precious family, Daddy, Mummy, Cheryl, Eugene, and my beloved fiancé, Sunny, I owe everything to you all. Thank you for your steadfast love and support and for all the sacrifices you have had to make; you make everything worthwhile and I am immensely blessed because of you. Dearest Sunny, thank you for your unwavering love and faith in me; you are my constant source of strength, friendship and motivation. To you all I dedicate this thesis.

*To my beloved family and fiancé*

## TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

## ABBREVIATIONS

| | |
|---|---|
| 5' RACE | 5’ Rapid Amplification of cDNA Ends |
| 5’ $m^7G$ | 5’ 7-methylguanosine |
| 5’ RLM-RACE | RNA-ligase mediated 5’ Rapid Amplification of cDNA Ends |
| A | anterior |
| a.a. | amino acid |
| ADARs | adenosine deaminases |
| AML | acute myeloid leukemia |
| Antp | Antennapedia |
| AP | anterior-posterior |
| ARE | AU-rich element |
| ATCC | American Type Culture Collection |
| ATRA | all-*trans* retinoic acid |
| ba | branchial arch |
| bp | basepairs |
| C | cervical vertebra |
| ChIP | chromatin immunoprecipitation |
| CIP | calf intestinal phosphatase |
| CML | chronic myeloid leukemia |
| CMV | cytomegalovirus |
| Cyto | cytoplasm |
| D | day |
| Dfd | Deformed |
| DR5 | direct repeat with a 5 base pair spacer |
| ds | double-stranded |
| E | Embryonic day |
| EC | embryonal carcinoma |

| | |
|---|---|
| eIF | eukaryotic initiation factor |
| EST | Expressed Sequence Tag |
| ex | exon |
| FISH | fluorescence *in situ* hybridization |
| flb | forelimb bud |
| FXR1 | fragile X mental retardation-related protein 1 |
| GW | glycine-tryptophan |
| HDAC | histone deacetylase |
| HMEC-1 | human microvascular endothelial cell |
| HMGA2 | high-mobility group A2 |
| IE | independent experiment |
| IRES | Internal Ribosomal Entry Site |
| kb | kilobases |
| KLF4 | Krüppel-like factor 4 |
| M | molecular weight marker |
| miRNA | microRNA |
| mRNA | messenger RNA |
| ncRNA | noncoding RNA |
| NE | neural enhancer |
| neuroepi | neuroepithelium |
| NPM | nucleophosmin 1 |
| NPMc+ | nucleophosmin 1 cytoplasmic positive |
| nt | nucleotides |
| NTC | no template control |
| Nuc | nucleus |
| opv | optic vesicle |
| P | posterior |
| P bodies | Processing bodies |
| PABPC | binding poly(A)-binding protein |

| | |
|---|---|
| PcG | Polycomb Group |
| POLR3D | RNA polymerase III (DNA directed) polypeptide D |
| pre-miRNA | precursor miRNA |
| pri-miRNA | primary microRNA |
| qPCR | quantitative PCR |
| qRT-PCR | quantitative RT-PCR |
| r | rhombomere |
| RA | retinoic acid |
| RAR | retinoic acid receptor |
| RARE | retinoic acid response element |
| RISC | RNA-Induced Silencing Complex |
| RP | ribosomal protein |
| RT-PCR | reverse transcription polymerase chain reaction |
| RXR | retinoid X receptor |
| T | thoracic segment |
| TAP | tobacco acid pyrophosphatase |
| Tiam1 | T lymphoma invasion and metastasis 1 |
| TNF-a | tumour necrosis factor, alpha |
| TOP | oligopyrimidine tract |
| U | untreated |
| Ubx | Ultrabithorax |
| UTR | untranslated region |
| v | ventricle (brain) |

## ABSTRACT

Patterning of the animal embryo's anterior-posterior axis is dependent on spatially and temporally regulated Hox gene expression. The murine *Hoxd4* gene has been proposed to harbour two promoters, an upstream promoter P2, and a downstream promoter P1, that lie 5.2 and 1.1 kilobase pairs (kb) upstream of the coding region respectively. The evolutionarily conserved *microRNA-10b* (*miR-10b*) gene lies in the *Hoxd4* genomic locus in the fourth intron separating exons 4 and 5 of the P2 transcript and is directly adjacent to the proposed P1 promoter. *Hoxd4* transcription is regulated by a 3' neural enhancer that harbours a retinoic acid response element (RARE). Here, we show that the expression profiles of *Hoxd4* and *miR-10b* transcripts during neural differentiation of mouse embryonal carcinoma P19 cells are co-ordinately regulated, suggesting that both *Hoxd4* and *miR-10b* expression are governed by the neural enhancer. Our observation that P1 transcripts are uncapped, together with the mapping of their 5' ends, strongly suggests that they are generated by Drosha cleavage of P2 transcripts rather than by transcriptional initiation. This is supported by the co-localization of P1 and P2 transcripts to the same posterior expression domain in the mouse embryo. These uncapped P1 transcripts do not appear to possess an Internal Ribosomal Entry Site (IRES), but accumulate within multiple punctate

bodies within the nucleus suggesting that they play a functional role. Finally, similar uncapped Drosha-cleaved P1-like transcripts originating from the paralogous *Hoxb4*/*miR-10a* locus were also identified. We propose that these transcripts may belong to a novel class of regulatory RNAs.

# CHAPTER I. INTRODUCTION

## 1.1 Hox genes

### 1.1.1 Introduction to Hox genes

The development of a multi-cellular organism from a single cell is an extraordinarily complex and highly regulated event. The established paradigm is that Hox genes act as “master regulators” in the specification of positional identity along the body axes of developing animals, determining the eventual body plan of the organism.

The Hox genes are a family of homeodomain-containing transcription factors involved in the regulation of segmental patterning and anatomical identity in the developing embryo. Hox genes are extraordinarily well conserved in the animal phyla and found to be essential in all animal species where they have been tested (McGinnis *et al*., 1984a). All Hox genes contain a 180 basepair (bp) DNA sequence termed the homeobox, which encodes a 60 amino acid (a.a.) homeodomain that is necessary for the DNA-binding and transcriptional activity of Hox proteins (McGinnis *et al*., 1984b; Scott and Weiner, 1984). The

homeodomain folds into a helix-turn-helix structure that is made up of three alpha helices connected by short loops. Two of the alpha helices are antiparallel and the third alpha helix is approximately perpendicular to the first two. This third alpha helix directly interacts with the major groove of DNA at a consensus sequence (TAAT/TANTNN) in cooperation with other cofactors to regulate transcription, while the N-terminal arm of the homeodomain makes base-specific contact to the minor groove (Gehring *et al*., 1994).

A hallmark of Hox genes in vertebrates and arthropods is that they are physically clustered together on the chromosome and display both spatial and temporal colinearity between their location on the chromosome and their expression along the anterior-posterior axis in the embryo, described as the "colinearity rule" (Duboule and Morata, 1994; Krumlauf, 1994). In general, the Hox genes located more 3' in each cluster, such as *HoxA1* and *HoxB1*, are expressed in the embryo earlier, and in a more anterior position (Duboule and Dolle, 1989; Graham *et al*., 1989). Hox genes are expressed in partially overlapping domains along the embryonic axes of the axial skeleton, central nervous system, branchial arches and limbs (Dolle *et al*., 1989; Giampaolo *et al*., 1989; Hunt and Krumlauf, 1992; Hunt *et al*., 1991a; Hunt *et al*., 1991b; Kessel and Gruss, 1991; Nelson *et al*., 1996; Wilkinson *et al*., 1989). A unique combination of the expressed Hox

proteins at a particular anterior-posterior level specifies its regional identity, which is also known as the “Hox code”.

Homeotic transformations and malformations arise when Hox gene expression in the developing embryo is dysregulated, and the precise spatio-temporal control of their expression is therefore critical to normal embryonic development. Changes in Hox gene expression due to mutations can lead to profound morphological defects.

Homeotic mutations in animals were first identified in the fruit fly *Drosophila melanogaster,* where one part of the embryo is transformed into another (Bridges and Morgan, 1923; Nusslein-Volhard and Wieschaus, 1980). One classical Hox mutant phenotype is the partial loss-of-function of *Ultrabithorax* (*Ubx*) in thoracic segment 3 (T3) of the fly which causes a homeotic transformation of T3 to T2 and consequently, two pairs of wings instead of one. In the partial gain-of-function *Antennapedia* (*Antp*) mutant, ectopic expression of *Antp* in the head segment causes a homeotic transformation of antennae to legs. In vertebrates such as the mouse, *Hoxc8*-null mice show a homeotic transformation of the first lumbar vertebra to a thoracic vertebra, resulting in an extra pair of ribs (Le Mouellic *et al*., 1992). In general, it was observed that loss-of-function Hox

mutations lead to anterior transformations of posterior structures while gain-of-function Hox mutations lead to posterior transformations of anterior structures. This was described as "posterior prevalence", which postulates that posterior Hox genes function by dominating over the activity of more anterior Hox genes. Mild variations in the domains of expression of Hox genes can therefore lead to changes in body plans, facilitating adaptation to natural evolutionary pressures which is of great interest in the field of evolutionary developmental biology.

Hox genes continue to be differentially expressed in differentiated cells in adult tissues such as fibroblasts, the gastrointestinal tract, adrenal glands, testes, kidneys and uterus in a colinear manner along body axes reminiscent of their embryonic expression pattern even though these tissues have different embryonic origins (Bomgardner *et al*., 2001; Kawazoe *et al*., 2002; Neville *et al*., 2002; Rinn *et al*., 2008; Takahashi *et al*., 2004; Taylor *et al*., 1997; Yamamoto *et al*., 2003). For example, only the 3' members of the Hox clusters (Hox1-5) are expressed in rostral tissues such as the hindbrain and trachea whereas almost all the Hox genes (Hox1-13) are expressed in caudal tissues such as the bladder and rectum (Morgan, 2006; Yamamoto *et al*., 2003). In addition, the same nested Hox gene expression is observed along the anterior-posterior axes of the adult gastrointestinal tract and the female reproductive tract (Du and Taylor, 2004;

Kawazoe *et al*., 2002; Morgan, 2006; Taylor *et al*., 1997; Yahagi *et al*., 2004).

While Hox expression in adult animals is less well-studied, they are known to play a role in the maintenance of differentiated cells and the dysregulation of Hox gene expression is implicated in various cancers, such as leukemia, lung, cervical, prostate and breast cancer (Abate-Shen, 2002; Abe *et al*., 2006; Gupta *et al*., 2010; Jung *et al*., 2004; Shah and Sukumar, 2010; Tan *et al*., 2009; van Scherpenzeel Thim *et al*., 2005; Zhai *et al*., 2007). For example, upregulation of *HOXC10* is associated with increased invasiveness of cervical cancer cell lines (Zhai *et al*., 2007) and *miR-10b*-mediated repression of *HOXD10* promotes tumour invasion and metastasis of breast cancer cell lines (Ma *et al*., 2007). A study of childhood acute lymphoblastic leukemia (ALL) has identified a mutation leading to a partial loss-of-function *HOXD4* protein as a susceptibility factor (van Scherpenzeel Thim *et al*., 2005). Transcriptional silencing of the *HOXD4* promoter mediated by *miR-10a* has also been shown in human breast cancer cells, suggesting a possible role for *HOXD4* in oncogenesis (Tan *et al*., 2009). These studies associate *HOXD4* expression with certain cancers but do not establish a causative effect between its gene expression and cancer progression.

### 1.1.2 Mammalian Hox gene cluster

In both mouse and human genomes, there are thirty nine Hox genes organized in four *Hox* clusters (A-D), each about 120 kilobases (kb) long on separate chromosomes, with a maximum of thirteen Hox paralogs within each cluster (Figure 2) (Duboule and Dolle, 1989; Duboule and Morata, 1994). It is hypothesized that the thirteen Hox paralogs arose by several gene duplication events and the four mammalian Hox clusters arose by a cluster quadruplication event, possibly via two genome duplications (Kappen *et al*., 1989). Following genome duplications, several Hox genes have been lost within each of the four clusters, resulting in an incomplete set of Hox genes in each vertebrate Hox cluster. Therefore, no single Hox cluster has retained all thirteen Hox paralogs and most paralog groups do not contain all four members (A-D), suggesting that some partial functional redundancy exists between paralogous Hox genes. In the course of evolution, some paralog clusters such as HoxA and HoxD have also acquired new functions in the specification of segment identity along the limb axes.

In the mouse, the four Hox clusters A, B, C and D are located on chromosomes 6, 11, 15 and 2 respectively. Just as in flies, there is a general correlation between

the physical position of the mammalian Hox genes in the Hox cluster and the time at which its expression is activated, as well as the domains of expression along embryonic axes. This phenomenon is termed temporal and spatial colinearity. While it is not an absolute hard and fast rule of Hox expression pattern, it is a remarkably well-conserved aspect of the Hox genes. For example, the Hox genes of paralog group 1 such as *Hoxa1* that are located at the 3' end of each Hox cluster are expressed earlier and have an anterior border of expression between rhombomeres 3 and 4 (r3/4) in the developing mouse hindbrain while Hox genes of paralog group 4, such as *Hoxa4*, *Hoxb4* and *Hoxd4* are expressed later and have an anterior border of expression between rhombomeres 6 and 7 (r6/7), posterior to that of *Hoxa1*. The exceptions to the colinearity rule in the mouse are the Hox genes of paralog group 2, *Hoxa2* and *Hoxb2*, which have a more anterior boundary of expression in the hindbrain compared to *Hoxa1* and *Hoxb1*.

During early development, the vertebrate hindbrain or rhombencephalon is transiently segmented from the neural tube into structures termed rhombomeres, with each rhombomere developing its own set of ganglia and neurons. In the chick hindbrain for example, the motor neurons of the trigeminal cranial nerve V originate from r2-r3, while those of facial cranial nerve VII originate from r4-r5

and those of glossopharyngeal cranial nerve IX originate from r6-r7 (Lumsden and Keynes, 1989). There are some subtle differences in the mouse hindbrain, with the cranial nerves V, VII and IX originating from r1-r3, r4-r5 and r6 respectively (Cordes, 2001; Fritzsch, 1998). Many Hox genes are expressed in the hindbrain with specific rhombomeric expression domains and their misexpression can cause a change in rhombomeric cell fate and affect cranial motor neuron development (Bell *et al*., 1999; Carpenter *et al*., 1993; Jungbluth *et al*., 1999).

The regulatory mechanisms directing temporal and spatial Hox gene expression appear to be separate, and Hox misexpression in either time or space is sufficient to induce homeotic transformations (Tschopp *et al*., 2009). In addition, some paralog clusters such as HoxA and HoxD have also acquired new functions in the specification of segment identity along the limb axes (Davis and Capecchi, 1996; Small and Potter, 1993). For example, the 5' HoxD genes (*Hoxd9-13*) have been shown to be critical for limb development, with surviving *Hoxa13*/*Hoxd13* double mutants showing a complete lack of digit formation (Davis and Capecchi, 1994; Davis *et al*., 1995; Dolle *et al*., 1993; Favier *et al*., 1995; Fromental-Ramain *et al*., 1996; Small and Potter, 1993; Zakany and Duboule, 1996).

A number of studies have shown that some Hox paralogs exhibit a certain degree of functional redundancy (Davis *et al*., 1995; Greer *et al*., 2000; Horan *et al*., 1995a). For example, a gene swapping experiment in which the *Hoxa3* coding sequence was replaced by that of *Hoxd3* resulted in viable mice, demonstrating that the HOXD3 protein, when expressed at the *Hoxa3* locus, is able to rescue the *Hoxa3*-null lethality and that HOXD3 is functionally equivalent to the HOXA3 protein, even though the two paralogs share less than 50% amino acid sequence identity and the respective knockout mice exhibit completely different phenotypes (Greer *et al*., 2000). This therefore establishes that Hox regulatory elements play a crucial role in establishing the precise space, time, and levels of Hox gene expression, all of which critically determines the function of the particular Hox gene product.

### 1.1.3 Hox genes of Paralog Group 4

There are four vertebrate Hox genes (*Hoxa4*, *Hoxb4*, *Hoxc4* and *Hoxd4*) within paralog group 4 of the Hox cluster that are orthologous to the *Deformed* (*Dfd*) gene of *Drosophila*.

Although *Hoxa4*, *Hoxb4* and *Hoxd4* have a similar anterior border of expression at r6/7 in the mouse hindbrain, *Hoxb4* and *Hoxd4* expression in the paraxial mesoderm extends to the first cervical vertebra (C1) while *Hoxa4* and *Hoxc4* have an anterior border of expression at the second (C2) and third (C3) cervical vertebra respectively, with each difference corresponding to a single somite (Gaunt *et al*., 1989; Horan *et al*., 1995b; Hunt and Krumlauf, 1992).

There are also some temporal and tissue-specific differences in their expression. For example, *Hoxd4* was shown to be highly expressed in the mesoderm and ectoderm earlier in time, between Embryonic day (E) E8.5 and E9.5, and its expression was absent from these tissues by E12.5. In contrast, *Hoxa4* and *Hoxb4* were weakly expressed in the same tissues between E8.5 and E9.5 but were highly abundant by E12.5. This difference in expression pattern points to a divergence in function between these paralog group 4 Hox genes.

Loss-of-function *Hoxa4* mutant mice showed a partial homeotic transformation of C3 to C2 while loss-of-function *Hoxb4* and *Hoxd4* mutant mice displayed a partial homeotic transformation of C2 to C1 in addition to other skeletal defects in the sternum and the basiooccipital bone (Horan *et al*., 1995a; Horan *et al*., 1995b; Horan *et al*., 1994; Kostic and Capecchi, 1994; Ramirez-Solis *et al*.,

1993). Interestingly, the homeotic transformations observed in each of these single mutants were at the anterior-most expression domain of each of the genes respectively. This is consistent with other observations on Hox-null mice and supports the posterior prevalence rule. However, in *Hoxa4/Hoxb4/Hoxd4* triple mutant mice, the homeotic transformations observed extended even more posteriorly, such that even C5 could be anteriorized to C1 identity, suggesting greater functional redundancy in more posterior regions of their expression domains. In the *Hoxb4* homozygous mutants and *Hoxb4/Hoxd4* double mutants, some embryonic lethality was observed, which could possibly be attributed to a neural complication. Curiously however, there are no other observable neuronal phenotypes of these paralog group 4 Hox mouse mutants. Consequently, the neural expression of *Hoxd4*, its regulation and function are of special interest to our laboratory.

## 1.2 MicroRNAs

### 1.2.1 Introduction to MicroRNAs

MicroRNAs are a class of highly conserved endogenous small noncoding RNAs (ncRNAs) about 22 nucleotides in length that were first discovered as the *lin-4* and *let-7* microRNAs controlling development in *Caenorhabditis elegans* (Lee *et al*., 1993; Reinhart *et al*., 2000). Its orthologs were later found to be expressed in a wide range of organisms such as plants, green algae, viruses, flies, fish and mammals (Griffiths-Jones, 2004; Griffiths-Jones *et al*., 2006). To-date, there are over eighteen thousand documented miRNAs expressing over twenty one thousand mature miRNA products in 168 species in the miRBase microRNA database (Griffiths-Jones *et al*., 2006), of which over 1500 are human miRNAs that are predicted to regulate some 60-90% of human genes (Friedman *et al*., 2009; Miranda *et al*., 2006). MicroRNAs have been observed to be expressed in a developmental and tissue-specific manner, indicative of a functional role in development (Chen *et al*., 2004; Houbaviy *et al*., 2003; Lagos-Quintana *et al*., 2002). They also have well established roles in various biological processes including cellular differentiation, proliferation, apoptosis and oncogenesis (Bushati *et al*., 2008; Debernardi *et al*., 2007; Giraldez *et al*., 2005).

### 1.2.2 MicroRNA nomenclature

MicroRNAs are classified based on the similarity of their mature sequence. In the zebrafish *Danio rerio*, dre-*miR-10b-1* and dre-*miR-10b-2* (with numbered suffixes) share an identical mature and seed sequence but are encoded by different genes in different loci. In mammals, the *miR-10* family members, *miR-10a* and *miR-10b* (with lettered suffixes), differ only by a single nucleotide in their mature sequence, and are located in a paralogous position upstream of *Hoxb4* and *Hoxd4* respectively. Orthologous microRNAs with the same name have an identical mature miRNA sequence and are denoted by a prefix identifying the species such as hsa-*miR-10b* and mmu-*miR-10b*. In addition, the mature microRNA strand can be derived from either arm of the pre-miRNA, and are named miR-17-5p (5′ arm) and miR-17-3p (3′ arm) respectively. Alternatively, the less common species can be denoted with an asterisk (*) in addition to the miRNA name, such as dme-*miR-10a**.

### 1.2.3 MicroRNA biogenesis and regulation

MicroRNAs are encoded within cellular genomes and may be located in either intergenic or intragenic regions. Intergenic microRNAs have their own promoters that control their transcription. When the miRNAs are positioned within the gene locus of other genes, its transcription may be controlled by the host gene. miRNAs can also be clustered together on the chromosome which can then be expressed as polycistronic transcripts (Lagos-Quintana *et al*., 2001; Lau *et al*., 2001; Lee *et al*., 2002).

miRNAs genes are transcribed by RNA Polymerase II to give a long primary microRNA (pri-miRNA) transcript that is 5' capped, spliced and 3' polyadenylated. In the canonical maturation pathway, the pri-miRNA forms a hairpin loop that acts as a signal for cleavage by an RNase III enzyme, Drosha and its cofactor DGCR8. Drosha cleaves the pri-miRNA near the base of the hairpin to give a 60-70 nt long precursor miRNA (pre-miRNA), which is then exported out of the nucleus into the cytoplasm by Exportin-5 (Lee *et al*., 2002; Lund *et al*., 2004; Yi *et al*., 2003). Another RNase III enzyme, Dicer, then cleaves the pre-miRNA at the terminal loop to give the miRNA:miRNA* duplex, which is formed by two single-stranded miRNA strands, one of which is the

predominant functional species. The strand with a lower internal stability at the 5' end is selected to be the mature miRNA strand, also known as the guide strand, while the other miRNA* strand, termed the passenger strand, may be cleaved and degraded by Ago2 (Khvorova *et al*., 2003; Schwarz *et al*., 2003).

The mature miRNA functions by base-pairing to target mRNAs and associates with Argonaute (AGO) family proteins to form ribonucleoprotein complexes such as the RNA-Induced Silencing Complex (RISC) to regulate gene expression, usually at the post-transcriptional level (Filipowicz *et al*., 2008). Recent evidence suggests that miRNAs are also able to epigenetically silence genes at the transcriptional level by recruiting silencing histone modifiers such as the Polycomb Group (PcG) protein EZH2, a histone methyltransferase, to promoters of target genes (Kim *et al*., 2008).

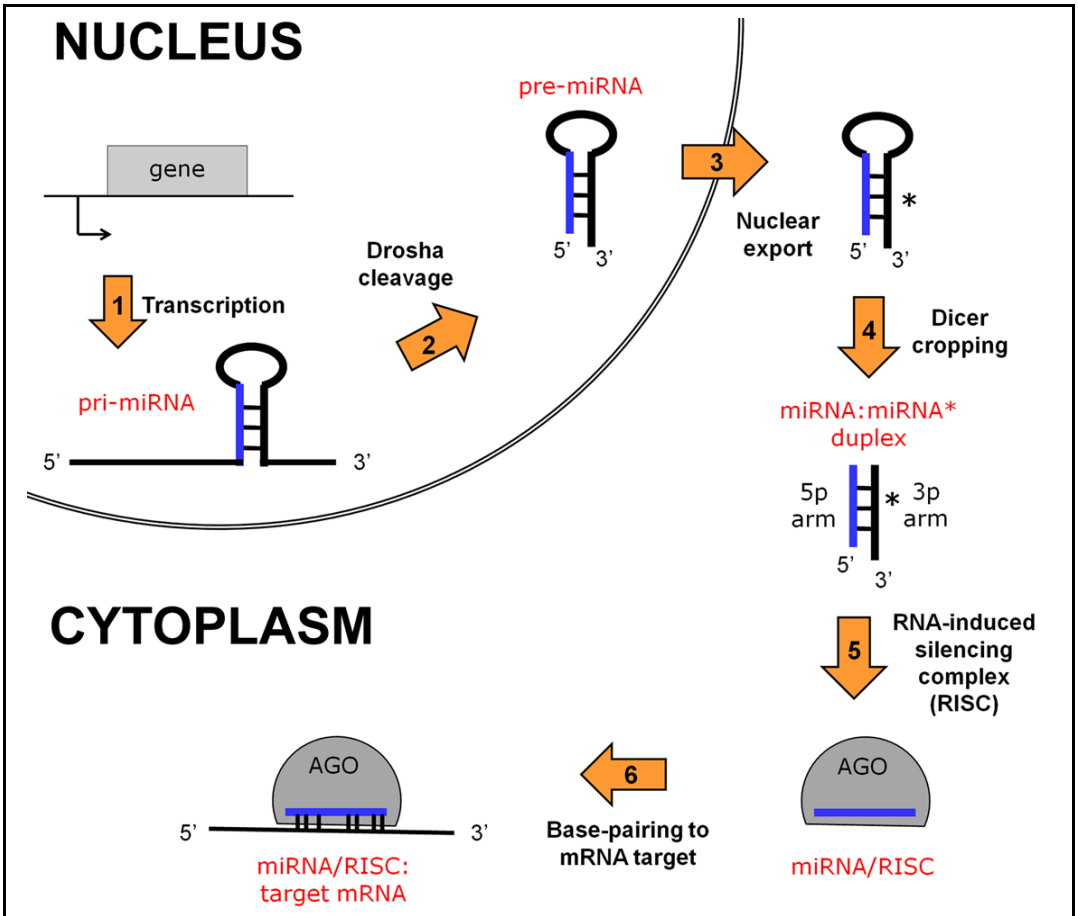


**Figure 1 The canonical microRNA processing pathway**

**(1)** microRNAs are transcribed from endogenous genes in the nucleus as long primary miRNA transcripts (pri-miRNA) that may be further spliced and polyadenylated. **(2)** The hairpin loop structure of the pri-miRNA acts as a signal for cleavage by a double-stranded (ds) ribonuclease, Drosha, to produce the precursor miRNA (pre-miRNA). **(3)** Exportin-5 mediates nuclear export of the 70-80 nt long pre-miRNA. **(4)** A cytoplasmic dsRNA nuclease, Dicer, crops the pre-miRNA at the hairpin loop leaving a miRNA:miRNA* duplex with 1–4 nt 3' overhangs. **(5)** The single-stranded mature miRNA (in blue) associates with AGO to form the RNA Induced Silencing Complex (RISC). **(6)** The

miRNA/RISC complex binds to partially complementary base pairs of specific target mRNAs to control its fate post-transcriptionally.

(Adapted from Ambion ® Technical Resources)

In addition, microRNAs can undergo post-transcriptional modifications such as RNA editing, 2'-O-methylation, adenylation and uridylation of miRNA 3' ends, which directly affects miRNA function and stability. Adenosine deaminases (ADARs) convert the RNA base adenosine to inosine in a process called A to I editing. A single base change was found to be sufficient to block either Drosha or Dicer cleavage for pri-*miR-142* and pri-*miR-151*, respectively (Kawahara *et al*., 2007a; Yang *et al*., 2006). Furthermore, it has been shown that a specific A to I editing of *miR-376* in its seed region affected its function by re-targeting the edited *miR-376* to repress an entirely different group of mRNAs (Kawahara *et al*., 2007b). The poly(A) polymerase GLD-2 was found to have a function in stabilizing mature *miR-122* levels in liver cells by the addition of a single adenosine base to its 3' end (Katoh *et al*., 2009). On the other hand, 3' uridylation of pre-*let-7* miRNA by terminal uridyltransferases TUT4/Zcchc11/PUP-2 blocked Dicer processing resulting in miRNA degradation in both mouse and human cell lines (Hagan *et al*., 2009; Heo *et al*., 2008; Kim *et al*., 2009). A number of other factors have been found to regulate miRNA turnover in various species, including an Argonaute protein, AGO2, that cleaves the passenger strand of the miRNA duplex, and the 5'-to-3' exonuclease XRN-2 that degrades unprotected cellular RNA (Chatterjee and Grosshans, 2009; Kai and Pasquinelli).

In one study, more than half (117 of 232) of all known mammalian microRNAs were found to be positioned within the introns of other coding genes (Rodriguez *et al*., 2004). Consequently, the transcription of the host gene and the miRNA may be co-regulated, and the unprocessed nascent RNA transcript can be either spliced to give the messenger RNA or cleaved by Drosha to give the pre-miRNA. In fact, it was shown that a single RNA transcript can give rise to both a mature miRNA and a protein-coding messenger RNA (Cai *et al*., 2004). Drosha cropping of the intronic miRNA can occur before splicing of the intron harbouring the miRNA and does not affect the production of correctly spliced messenger RNA (Kim and Kim, 2007).

A Drosha-independent microRNA biogenesis pathway was first described in Drosophila and later, in mammals (Berezikov *et al*., 2007; Okamura *et al*., 2007; Ruby *et al*., 2007a). Some microRNAs located in short introns of about 50-200 bp in length were found to be able to bypass Drosha processing in the nucleus. Instead, the splicing of the exons flanking the miRNA releases short hairpin introns in a lariat form that can be subsequently debranched to give the pre-miRNA for downstream Dicer processing. These short hairpin introns were termed mirtrons and their biogenesis was termed the mirtronic miRNA processing pathway.

### 1.2.4 MicroRNA function

The mature miRNA is about 22 nucleotides (nt) long and the second to seventh 5' nucleotide (5' 2-7 nt) is known as the 'seed' sequence critical for base-pairing with target transcripts. Imperfect base-pairing of the seed sequence to the mRNA targets are associated with translational repression while perfect base-pairing results in Ago2-directed mRNA cleavage (Zeng *et al*., 2003).

There are several miRNA-mediated gene regulatory mechanisms, such as by inhibition of translational initiation and elongation, premature termination of translation, inducing mRNA degradation via decapping and deadenylation and finally, by activating translation, which is a more novel recent discovery (Eulalio *et al*., 2008; Filipowicz *et al*., 2008; Orom *et al*., 2008; Vasudevan *et al*., 2007). More recently, miRNAs have also been shown to epigenetically control gene expression at the transcriptional level.

Inhibition of translational initiation

Cap-dependent initiation of translation begins with the binding of eukaryotic initiation complex eIF4F (which includes the cap-binding protein eIF4E, the

scaffolding protein eIF4G and the RNA helicase eIF4A) to the 5' 7-mG cap of an mRNA, and the poly(A)-binding protein PABPC1 to its poly(A) tail. The interaction between eIF4F and PABPC1 circularizes the mRNA to enhance translation.

Some studies have shown that miRNAs and their mRNA targets do not associate with the ribosomes (Pillai *et al*., 2005), and both the 5' cap and poly(A) tail of target mRNAs were required for let-7 miRNA-mediated gene-silencing (Humphreys *et al*., 2005). The addition of increasing amounts of purified eIF4F to an *in vitro* cell extract was also shown to relieve miRNA-mediated gene silencing (Mathonnet *et al*., 2007) while 5' cap-independent translation via an IRES was refractory to miRNA silencing, suggesting that cap-recognition is important (Humphreys *et al*., 2005; Kiriakidou *et al*., 2007; Mathonnet *et al*., 2007; Pillai *et al*., 2005; Wakiyama *et al*., 2007).

The Argonaute protein AGO2 has been proposed to bind to the 5' cap of target mRNAs to compete with eIF4E (Kiriakidou *et al*., 2007). Alternatively, AGO2 could repress translational initiation by binding and sequestering eIF6 and preventing 80S ribosome formation (Kiriakidou *et al*., 2007). Together,

these studies show that miRNAs are able to inhibit 5' cap-dependent translational initiation.

Inhibition of translational elongation and premature translational termination

Several independent groups had evidence on the contrary, showing that miRNAs and their target mRNAs were associated with ribosomes in sucrose sedimentation gradients even though there were no detectable levels of protein (Maroney *et al*., 2006; Nottrott *et al*., 2006; Petersen *et al*., 2006; Seggerson *et al*., 2002). In addition, inhibition of translational initiation caused an increase in ribosomal dissociation from the miRNA:mRNA transcripts (Petersen *et al*., 2006). This suggests that these miRNAs may inhibit protein synthesis after initiation of translation via cotranslational protein degradation and/or premature ribosome dissociation.

mRNA degradation *via* decapping and deadenylation

miRNA-mediated mRNA decay via deadenylation and decapping is mediated by Argonaute proteins, a glycine-tryptophan (GW) rich and RNA-binding protein GW182, the decapping complex DCP2:DCP2, additional decapping activators Ge-1 and EDC3, the RNA helicase RCK/p54 and the CAF1-CCR4-NOT deadenylase complex (Behm-Ansmant *et al*., 2006; Eulalio *et al*.,

2007b). miRNAs and their mRNA targets have been found to colocalize to cytoplasmic foci described as Processing bodies (P bodies) together with these proteins as a consequence, rather than cause of silencing (Eulalio *et al*., 2007a; Pillai *et al*., 2005).

Activating translation

Intriguingly, while most studies have found that miRNAs function mainly by recruitment of AGOs and RISC complexes to target mRNAs to repress translation, several new studies have shown that miRNAs are capable of activating translation of target mRNAs.

The seed sequence of *miR-369-3* is complementary to the AU-rich element (ARE) in the 3' untranslated region (3' UTR) of tumour necrosis factor-α (TNFα) mRNA and it was demonstrated to repress the translation of TNFα in synchronized proliferating HeLa cells via base-pairing to the ARE and the subsequent recruitment of AGO2. Serum-starvation of the HeLa cells, which led to cell-cycle arrest, permitted a switch from *miR-369-3*-mediated repression to translation activation through the additional recruitment of fragile X mental retardation-related protein 1 (FXR1) to the AGO2 complex (Vasudevan *et al*., 2007). Similarly, *let-7* was also shown to repress the

translation of the high-mobility group A2 (HMGA2) mRNA in proliferating cells but enhanced its translation upon cell-cycle arrest, demonstrating that miRNAs can induce translational activation under specific conditions (Vasudevan *et al*., 2007).

In a separate study, *miR-10a* has also been shown to induce translational activation of a subset of mRNAs encoding ribosomal proteins (RPs) during amino-acid starvation by its interaction with the 5' UTR region immediately downstream of a regulatory 5' oligopyrimidine tract, known as a 5' TOP motif (Orom *et al*., 2008).

More recently, AGO2 and FXR1-dependent miRNA-mediated translational activation was demonstrated in immature *Xenopus laevis* oocytes, where *xl-miR-16* was found to upregulate expression of Myt1 kinase to maintain the oocyte in an immature state (Mortensen *et al*., 2011).

Epigenetic control at the transcriptional level

In addition to the regulation of genes at the post-transcriptional level, there have been reports indicating that long ncRNAs and small RNAs are able to regulate genes at the transcriptional level by recruitment of chromatin

modifiers at promoter regions in plants and mammals (Khraiwesh *et al*., 2010; Kim *et al*., 2008; Morris *et al*., 2004; Reinhart and Bartel, 2002; Rinn *et al*., 2007; Wutz *et al*., 2002).

For example, *miR-320* was shown to direct EZH2, a histone-lysine N-methyltransferase and a member of the PcG family group of proteins, to the POLR3D promoter to inhibit its transcription in mammalian cells (Kim *et al*., 2008). *miR-10a* was also shown to repress *HOXD4* at the transcriptional level which was accompanied by *de novo* methylation of the *HOXD4* promoter in human breast cancer cell lines (Tan *et al*., 2009).

In summary, miRNAs are able to function as both activators and repressors of gene expression at either the transcriptional or post-transcriptional level by the common mechanism of base-pairing to its complementary mRNA or even DNA targets.

## 1.3 MicroRNAs in the Hox cluster

There are three known microRNAs or miRNA families embedded in vertebrate *Hox* clusters: *miR-10*, *miR-615* and *miR-196* (Figure 2). *miR-10* is located 5' of *Hoxb4* and *Hoxd4*, *miR-615* 5' of *Hoxc5* and *miR-196* 5' of *Hoxa9*, *Hoxb9* and *Hoxc9* (Figure 2). The position of these miRNAs within the *Hox* clusters is highly conserved during evolution. The *miR-10* family is not present in nematodes like *C. elegans* and tunicates like *C. intestinalis* where the Hox clusters have disintegrated into multiple pieces in both species. MicroRNAs located in Hox clusters have been shown to inhibit more anterior Hox genes, which is consistent with the posterior prevalence phenomenon associated with the Hox cluster (Lempradl and Ringrose, 2008).

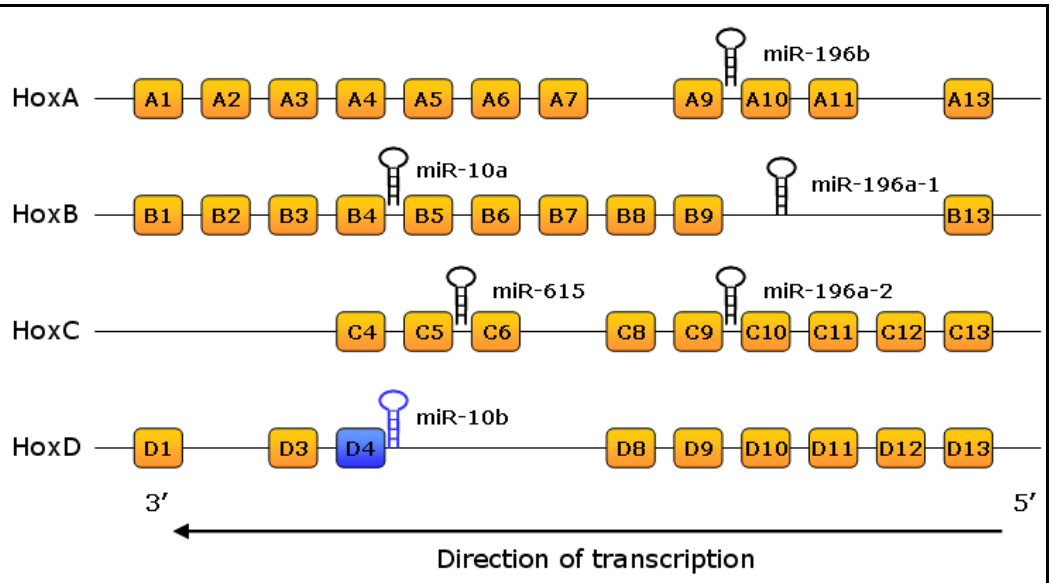


**Figure 2 Hox genes and their associated microRNAs**

A schematic of the four mammalian Hox clusters (HoxA through HoxD) and the position of the microRNAs embedded within them. The Hox genes are transcribed from the same DNA strand in a 5' to 3' direction and display temporal and spatial colinearity, i.e., genes located on the 3' most end of each Hox cluster are generally expressed earlier and with a more anterior border of expression along the developing embryonic axis.

(Adapted from Mansfield *et al*., 2004)

### 1.3.1 Known functions of the *miR-10* family

The *miR-10* family is one of the most well conserved microRNAs and has been detected in a diverse range of species, including the fruit fly (*Drosophila melanogaster*), zebrafish (*Danio rerio*), mouse (*Mus musculus*), chick (*Gallus gallus*) and human (*Homo sapiens*). The sequence of *miR-10b* is completely conserved in the mouse, chick, human and zebrafish and differs from *miR-10a* by a single central nucleotide. The mature sequences of the *miR-10* family members in five species are summarized in Table 1, with the differing nucleotides highlighted in colour.

**Table 1 Summary of the sequences of the *miR-10* family members**

| Accession | Species | Name | Sequence of 5p arm (5'-to-3') | Sequence of 3p arm (5'-to-3') |
|---|---|---|---|---|
| MI0000130 | *Drosophila melanogaster* | ***dme-mir-10*** | -acccuguagauccgaauuuguu | * caaauucgguucuagagagguuu |
| MI0007559 | *Gallus gallus* | ***gga-mir-10a*** | uacccuguagauccgaauuugu- | -aaauucguaucuaggggaaua |
| MI0001216 | *Gallus gallus* | ***gga-mir-10b*** | uacccuguagaaccgaauuugu- | N.A. |
| MI0000685 | *Mus musculus* | ***mmu-mir-10a*** | uacccuguagauccgaauuugug | caaauucguaucuaggggaaua |
| MI0000221 | *Mus musculus* | ***mmu-mir-10b*** | uacccuguagaaccgaauuugug | cagauucgauucuaggggaaua |
| MI0000266 | *Homo sapiens* | ***hsa-mir-10a*** | uacccuguagauccgaauuugug | caaauucguaucuaggggaaua |
| MI0000267 | *Homo sapiens* | ***hsa-mir-10b*** | uacccuguagaaccgaauuugug | acagauucgauucuaggggaau |
| MI0001363 | *Danio rerio* | ***dre-mir-10a*** | uacccuguagauccgaauuugu- | caaauucgugucuuggggaaua |
| MI0001364 | *Danio rerio* | ***dre-mir-10b-1*** | uacccuguagaaccgaauuugug | N.A. |
| MI0001887 | *Danio rerio* | ***dre-mir10b-2*** | uacccuguagaaccgaauuugug | N.A. |
| MI0001888 | *Danio rerio* | ***dre-mir-10c*** | uacccuguagauccggauuugu- | N.A. |
| MI0001889 | *Danio rerio* | ***dre-mir-10d*** | uacccuguagaaccgaaugugug | cagauucgguuuuaggggagua |

Sequences were obtained from miRBase at http://www.mirbase.org/index.shtml

*: the mature miRNA derived from the 3p arm of *dme-mir-10* is more abundant

N.A.: sequence not available

The Drosophila *miR-10* is located in the *Antennapedia* Hox cluster, upstream of *Deformed* (*Dfd*), the ortholog of mammalian *Hoxd4*, on the reverse strand and it is expressed in a Hox-like spatially and temporally specific manner, e.g., in the ventral nerve cord, posterior midgut and hindgut of a stage 11 embryo (Aboobaker *et al*., 2005). Both the *miR-10a* and *miR-10a** strands, derived from the 5p and 3p arms respectively, were shown to be present and active in Drosophila (Aravin *et al*., 2003; Lagos-Quintana *et al*., 2001; Ruby *et al*., 2007b; Sempere *et al*., 2003; Stark *et al*., 2007). The neighbouring Hox gene, *Sex combs reduced* (*Scr*), is one of the predicted targets of *miR-10* while *Abdominal-B* (*Abd-B*) and *Ultrabithorax* (*Ubx*) are predicted targets of *miR-10** (Stark *et al*., 2007). There have not been any experimental validations of these predicted Hox targets, nevertheless, it indicates that both arms of a pre-miRNA could be functional and that *miR-10* is a potential regulator of Hox genes.

In zebrafish, there are seven Hox clusters and five genes encoding *miR-10* (*miR-10a*, *miR-10b-1*, *miR-10b-2*, *miR-10c* and *miR-10d*) in the Hox *Bb*, *Da*, *Ca* and *Ba* respectively, with the last *miR-10d* gene located on what was the eighth vestigial Hox *Db* cluster present in both the *Takifugu* and *Tetraodon* pufferfish genomes, and that have since degenerated in the zebrafish lineage (Woltering and Durston, 2006). There are two genes encoding *miR-10b*, known as

*miR-10b-1* and *miR-10b-2* in the *HoxDa* and *HoxCa* cluster respectively, that gives rise to identical mature sequences (Table 1). In addition, *dre-miR-10c* is more similar to *dre-miR-10a*, while *dre-miR-10d* and *dre-miR-10b* are more similar, which is consistent with the genomic duplication in the zebrafish (Table 1). In zebrafish, *miR-10b* is expressed in the posterior trunk and spinal cord with an anterior boundary somewhat posterior to the r6/7 boundary (Woltering and Durston, 2008). The zebrafish *miR-10* family is also known to play a role in restricting the expression domain of anterior Hox genes by repressing *hoxb1a* and *hoxb3a* within the spinal cord in cooperation with *hoxb4a* (Woltering and Durston, 2008). It is therefore important for the correct development of the Xth cranial nerve and the migration of trunk motor neurons (Woltering and Durston, 2008). This is consistent with the general observation of posterior prevalence in the Hox gene cluster, with more posterior genes inhibiting more anterior genes. It is postulated that this may be an additional regulatory mechanism to reinforce and fine-tune the correct expression domains of Hox genes. Interestingly, the *miR-10* target sites in *Hoxb1* and *Hoxb3* are conserved in both mouse and human.

*miR-10a* was reported to bind to a 5' UTR region immediately downstream of a regulatory 5'TOP motif of mRNAs encoding ribosomal proteins (RP) to enhance their translation, thereby upregulating global protein synthesis and affecting

oncogenic transformation of mouse ES cells (Orom *et al*., 2008). Separately, *miR-10a* was also shown to regulate retinoic acid-induced smooth muscle cell differentiation from mouse ES cells in a NF-κB-dependent mechanism via repression of the histone deacetylase HDAC4 (Huang *et al*., 2010).

*miR-10a* has been implicated in several human cancers including pancreatic cancer, urothelial cancer, acute myeloid leukemia (AML) and chronic myeloid leukemia (CML) (Agirre *et al*., 2008; Ovcharenko *et al*., 2011; Veerla *et al*., 2009; Weiss *et al*., 2009). In fact, *miR-10a* and *HOXB4* are both upregulated in NPM1 cytoplasmic positive (NPMc+) AML (Garzon *et al*., 2008). Another study showed that *miR-10a* expression correlates with *HOXB4* expression in the K562 (a chronic myelogenous leukemia) cell line during megakaryocytic differentiation and that *miR-10a* represses *HOXA1* by binding to a region in its 3' UTR (Garzon *et al*., 2006). More interestingly, *miR-10a*, located upstream of *Hoxb4*, has been shown to repress *Hoxd4* transcription by targeting its promoter region in human breast cancer cells (Tan *et al*., 2009).

Like *miR-10a*, *miR-10b* has been implicated in many different cancers. Its expression has been shown to be highly upregulated in metastatic mouse and human breast cancer cell lines, human esophageal cancer, human pancreatic

cancer, AMLs, B-cell chronic lymphocytic leukemia and malignant gliomas (Bloomston *et al*., 2007; Ciafre *et al*., 2005; Debernardi *et al*., 2007; Garzon *et al*., 2008; Ma *et al*., 2007; Moriarty *et al*., 2010; Sasayama *et al*., 2009; Tian *et al*., 2010).

*miR-10b* was shown to promote cell migration and invasion by transcriptional inhibition of a known tumor suppressor KLF4 (Krüppel-like factor 4) in human esophageal cancer (Tian *et al*., 2010). In human metastatic breast cancer, *miR-10b* is known to target *HOXD10* via its 3' UTR region to repress its translation, resulting in the induction of the pro-metastatic gene, RHOC, thereby initiating tumour cell invasion and migration (Ma *et al*., 2007). In addition, *in vivo* silencing of *miR-10b* was shown to be effective in reducing the formation of lung metastases in the mouse mammary cancer model (Ma *et al*., 2010). An additional study confirming *HOXD10* as a target of *miR-10b* also showed that overexpression of *miR-10b* downregulated *HOXD10* expression and enhanced HMEC-1 (human microvascular endothelial cells) cell migration and tube formation, linking *miR-10b* to angiogenesis (Shen *et al*., 2011).

However, in a separate and contrasting study, *miR-10b* appeared to have the opposite effect of inhibiting the migration and invasion of breast carcinoma cells

via a novel target Tiam1 (T lymphoma invasion and metastasis 1), which is a guanidine exchange factor for Rac (Moriarty *et al*., 2010). Taken together, current research suggests that the function of *miR-10b* is possibly context-dependent and is more complex than simply being either pro- or anti-oncogenic.

### 1.3.2 *miR-10b* and *Hoxd4*

*miR-10a* and *miR-10b* are expressed in the central nervous system and trunk in a sub-domain of the *Hoxb4* and *Hoxd4* expression domains. This spatio-temporal restriction along the AP axis is reminiscent of Hox gene expression and is conserved in many animals such as *Drosophila*, zebrafish and mouse (Kloosterman *et al*., 2006; Kosman *et al*., 2004; Mansfield *et al*., 2004; Wienholds *et al*., 2005).

*Hoxd4* patterns the anterior cervical skeleton of the mouse, and has been implicated in acute lymphoblastic leukemia (Folberg *et al*., 1997; Horan *et al*., 1995a; Horan *et al*., 1995b; van Scherpenzeel Thim *et al*., 2005). *Hoxd4* knockout mice display homeotic transformations of the second cervical vertebrae (C2) to the first cervical vertebrae (C1) in addition to malformations of the neural arches of C1-C3 and of the basioccipital bone. However, there is no

known or characterized neuronal phenotype. *Hoxd4*-null mice appear to have normal motor skills, are responsive to sound and touch and also exhibit normal cranial ganglia staining (Horan *et al*., 1995a).

Two *Hoxd4* promoters, an upstream promoter termed P2 and a downstream promoter termed P1 (Figure 4) have been deduced using a combination of S1 nuclease and RNase protection assays, and 5' Rapid Amplification of cDNA Ends (5'RACE) (Folberg *et al*., 1997). While these are rigorous methods to map the 5' ends of RNA transcripts, they do not reveal whether these 5' ends are capped.

*Hoxd4* 3' neural enhancer

A 700 bp region termed the 3' neural enhancer (3' NE) (Figure 4) is essential for directing *Hoxd4* expression in the central nervous system of the developing mouse embryo to the correct r6/7 boundary in the hindbrain, as well as *Hoxd4* expression in neurally differentiating P19 cells (Morrison *et al*., 1997; Morrison *et al*., 1996; Nolte *et al*., 2003; Nolte *et al*., 2006; Rastegar *et al*., 2004; Zhang *et al*., 2000; Zhang *et al*., 1997). Using deletional mapping of the *Hoxd4* 3' NE, a non-consensus DR5 type RARE with imperfect repeats was identified and found

to be essential for the mediation of the response of *Hoxd4* neural expression to ATRA (Morrison *et al*., 1997; Nolte *et al*., 2003). Mutational analyses of the 3' DR5 RARE in transgenic mice showed that it is necessary for both the initiation and maintenance of normal *Hoxd4* neural expression in the developing CNS from E8.5 to E10.5 (Nolte *et al*., 2003). Alignment of the mouse, chicken, human *Hoxd4* and zebrafish *hoxd4a* sequence revealed seven blocks of highly conserved sequences in the 3' NE, termed sites A-F and including the RARE, as shown in Figure 3 (Nolte *et al*., 2003). Interestingly, the *Hoxd4* 3' DR5 RARE is also present in a paralogous position in the *Hoxb* cluster, 3' to *Hoxb4* (Gould *et al*., 1998; Zhang *et al*., 2000).

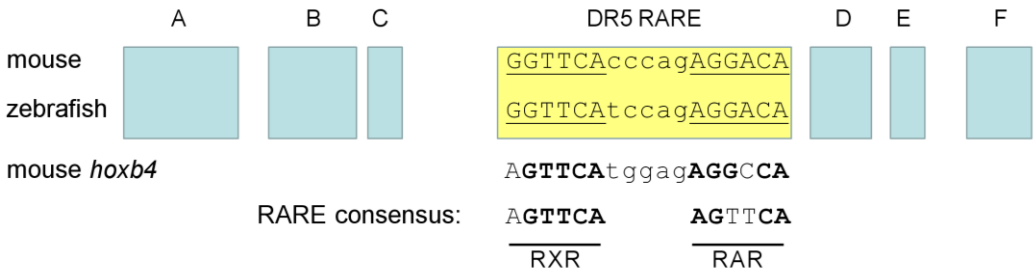


**Figure 3 Conserved blocks of homology in the mouse and zebrafish *Hoxd4* 3' Neural Enhancer**

This is a schematic showing the blocks of conserved sequences (A-F and DR5 RARE) between the mouse and zebrafish 3' neural enhancers. The sequence of the DR5 RARE is shown and the imperfect direct repeats underlined. The paralogous mouse *Hoxb4* 3' DR5 RARE sequence is shown with the conserved nucleotides highlighted in bold. The RARE consensus sequence consists of the RXR and RAR binding sites, and the conserved nucleotides are shown in bold.
DR5 RARE: retinoic-acid response element with a 5 bp spacer between two direct repeats.

(Adapted from Nolte *et al*., 2003)

In the mouse, the full *Hoxd4* expression domain in the developing central nervous system has an anterior limit at the boundary between rhombomeres 6 and 7 (r6/7) of the embryonic hindbrain. However, a subset of *Hoxd4* transcripts originating from an upstream P2 promoter is expressed more posteriorly at a level just above the forelimb bud. *miR-10b* is expressed in a similar temporal and spatial pattern as the *Hoxd4* P2 transcripts in the developing E9.5 mouse embryo, with an anterior expression border that is considerably posterior to the r6/7 anterior expression border (Folberg *et al*., 1997; Kloosterman *et al*., 2006). The anterior expression border of *miR-10a* is likewise posterior to the r6/7 expression border of the full *Hoxb4* domain (Mansfield *et al*., 2004).

The location of mouse *miR-10b* immediately adjacent to the presumptive P1 transcriptional start site and within the intron separating exons 4 and 5 of the P2 transcript (Figure 4) raised several questions regarding *Hoxd4* and the biogenesis of *miR-10b*. First, P2 may serve to encode both HOXD4 and *miR-10b*. In other words, the *Hoxd4* P2 transcript could also serve as the primary *miR-10b* transcript that is processed by Drosha. Second, expression of the *Hoxd4/miR-10b* transcript may be regulated by the same 3' neural enhancer. Third, the previously mapped 5' end of the *Hoxd4* P1 transcript may not be generated by

transcriptional initiation but by the cleavage of the *Hoxd4* P2 transcript by Drosha (Figure 4).

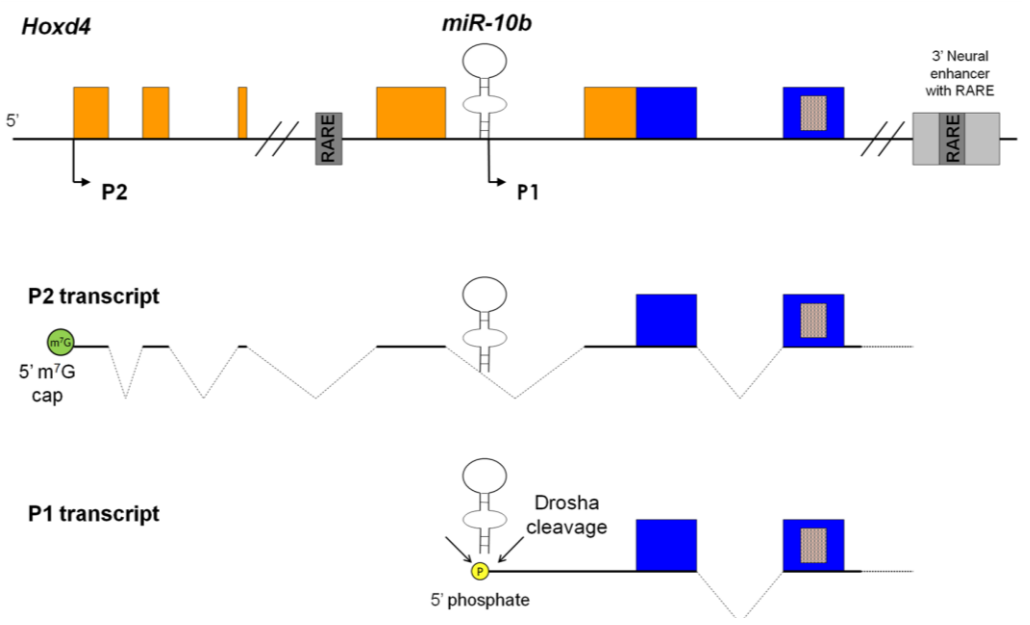


**Figure 4 Genomic organization of murine *Hoxd4***

This is a schematic showing the genomic organization of the mouse *Hoxd4* gene. Orange boxes indicate non-coding exons and blue boxes indicate coding exons. The nested shaded box indicates the homeobox within the second coding exon in blue. Regulatory elements are shaded grey. P2 denotes the upstream promoter and P1 denotes the presumptive downstream promoter. *miR-10b* is found directly upstream of the P1 promoter, in intron 4 of the P2 transcript. Dotted lines indicate spliced introns. The green circle at the 5' end of the P2 transcript denotes the 5' 7-methylguanosine ($m^7G$) cap and the circled "P" at the 5' end of the P1 transcript indicates a 5' phosphate that is generated by Drosha cleavage of a longer pri-*miR-10b* transcript.

RARE: retinoic-acid response element

(Adapted from Folberg *et al.*, 1997)

## 1.4 Research Objectives

The principal interest is in characterizing and understanding the regulation of *miR-10b* expression and its relationship, if any, to the regulation and processing of *Hoxd4* transcripts.

One aspect of this thesis aims to understand the regulatory mechanisms that direct the transcription and expression of *miR-10b*. We hypothesized that the *Hoxd4* neural enhancer is involved in this process and the mouse embryonal carcinoma P19 cell line was used as a system to study the appearance of *Hoxd4* and *miR-10b* transcripts during neural differentiation. The second aim is to determine whether the P1 transcript originates from a true transcriptional start site; the alternative being that it is generated by Drosha cleavage of the primary *miR-10b* (*pri-miR-10b*) transcript.

The expression profiles of *Hoxd4* and *miR-10b* transcripts during P19 differentiation are similar. This is consistent with the hypothesis that both *Hoxd4* and *miR-10b* is directed by the same regulatory elements. P1 transcripts were found to be uncapped and are therefore generated by Drosha cleavage rather than from a transcriptional start site. Strikingly, we also

detected a P1-like uncapped transcript that may encode *Hoxb4*, resulting from the Drosha cleavage of pri-*miR-10a*.

To further characterize these Drosha cleaved P1 transcripts, *in situ* hybridizations and nuclear-cytoplasmic fractionations in mouse embryos as well as neurally differentiating P19 cells were performed to determine its anatomical and subcellular localization. In addition, ribosome fractionation with high-speed sucrose gradient centrifugation and Internal Ribosomal Entry Site (IRES) assays were employed to determine if the P1 transcripts could be translated.

Finally, to investigate a potential relationship between Drosha cleavage and splicing of the pri-*miR-10b*/*Hoxd4* P2 transcript, morpholinos designed to block either Drosha cleavage or of splicing were used to transfect neurally differentiating P19 cells to determine its effect, if any, on *miR-10b* production and *Hoxd4* transcript levels. Drosha knockdown with siRNAs in neurally differentiating P19 cells was also performed in parallel.

Our specific aims are:

1. To investigate if the *Hoxd*4 and *miR-1*0b genes are co-ordinately regulated and expressed at the same time in neurally differentiating P19 cells
2. To investigate the possibility that *Hoxd4* P1 transcripts are generated by Drosha cleavage instead of a transcriptional start site
3. To investigate the paralogous *Hoxb4*/*miR-10a* genes
4. To characterize the pri-*miR-10b* transcript(s) and compare them with the known *Hoxd4* P2 transcript
5. To characterize the expression domains of these *Hoxd4*/*miR-10b* transcripts in the developing mouse embryo
6. To determine the subcellular localization of the Drosha-cleaved P1 transcripts
7. To investigate the possibility that the uncapped Drosha-cleaved P1 transcripts may be translated
8. To investigate possible interactions between the microRNA biogenesis machinery and the transcriptional and splicing machinery

# CHAPTER II. MATERIALS AND METHODS

## 2.1 Cell culture and differentiation

HEK293T and P19 cells were obtained from American Type Culture Collection (ATCC) and cultured in DMEM (Gibco) supplemented with 10% fetal bovine serum and 1% Penicillin/Streptomycin (Gibco). For differentiation, the P19 cells were seeded at a density of 1 x $10^5$ cells/ml in a 10 cm polystyrene bacterial petri dish (Greiner Bio-One), treated with 0.3 μM of all-*trans* retinoic acid (ATRA) and allowed to aggregate. Fresh media with 0.3 μM ATRA was replaced after two days. After four days, fresh media without ATRA was replaced and the cells plated onto tissue culture plates (Corning).

## 2.2 Quantitative Reverse Transcriptase PCR (qRT-PCR)

RNA was extracted using the PureLink™ Micro-to-Midi™ Total RNA Purification System (Invitrogen). Total RNA was pre-treated with DNase I (Fermentas). Reverse transcription of mRNA was carried out with

SuperScript® III First-Strand Synthesis (Invitrogen) and reverse transcription of miRNA was carried out with NCode™ miRNA First-Strand cDNA Synthesis (Invitrogen) as per manufacturer's instructions. The cDNA was then used for quantitative RT-PCR (qRT-PCR) together with SYBR GreenER™ qPCR SuperMix (Invitrogen) on a BioRad iCycler iQ5. 18S RNA and U6 snoRNA was used as an internal control for normalization of mRNAs and microRNAs respectively.

PCR cycling conditions for mRNA amplification:

| | | |
|---|---|---|
| Step 1 | 50 °C | 2 min |
| Step 2 | 95 °C | 10 min |
| Step 3 | 95 °C | 15 sec |
| Step 4 | 60 °C | 1 min |
| Step 5 | back to Step 3 for 40 cycles | |

followed by a melt curve analysis from 50-95 °C at 0.5 °C increments for 30 sec each.

PCR cycling conditions for miRNA amplification:

| | | |
|---|---|---|
| Step 1 | 50 °C | 2 min |
| Step 2 | 95 °C | 10 min |
| Step 3 | 95 °C | 15 sec |
| Step 4 | 57 °C | 1 min |
| Step 5 | back to Step 3 for 40 cycles | |

followed by a melt curve analysis from 50-95 °C at 0.5 °C increments for 30 sec each.

Sequences of qRT-PCR primers used are summarized in Table 2.

**Table 2 Sequences of qRT-PCR primer pairs**

| No. | Figure | Primer Name | Sequence (5'-to-3') |
|---|---|---|---|
| | **Figure 5B** | | |
| 1 | | P1+P2 | CTATGTCCATTCTGGGGGCT |
| 2 | | | CGGCAGGTAATCTCGGCTT |
| 3 | | P2 | CAC TCA CTT GGC TTA GGT TCT GG |
| 4 | | | AATCTCCCGCCCCCTTATCA |
| 5 | | nestin | CAGATGTGGGAGCTCAATCG |
| 6 | | | GCCTCCTCGATGGTCCGCTC |
| 7 | | miR-10b | CCCTGTAGAACCGAATTTGTG |
| 8 | | U6 | CTCGCTTCGGCAGCACA |
| 9 | | | AACGCTTCACGAATTTGCGT |
| 10 | | 18S | GTAACCCGTTGAACCCCATT |
| 11 | | | CCATCCAATCGGTAGTAGCG |
| | **Figure 8** | | |
| 12 | | P1 | CTATGTCCATTCTGGGGGCT |
| 13 | | | CGGCAGGTAATCTCGGCTT |
| 14 | | 18S | GTAACCCGTTGAACCCCATT |
| 15 | | | CCATCCAATCGGTAGTAGCG |
| 16 | | Hoxd4 ex5/6 | TGT GGT CTA CCC TTG GAT GAA |
| 17 | | | GAT CTG GCG CTC AGA CAG AC |
| 18 | | β-actin | TTCTTTGCAGCTCCTTCGTT |
| 19 | | | ATGGAGGGGAATACAGCCC |
| | **Figure 9** | | |
| 20 | | 47S rRNA | GAGAGAAGGAGGGGCAAGAC |
| 21 | | | GAAAGCCAGGCCTCTCAAAG |
| 22 | | snoU6 | CGCTTCGGCAGCACATATAC |
| 23 | | | CGAATTTGCGTGTCATCCTT |
| 24 | | β-actin | TTCTTTGCAGCTCCTTCGTT |
| 25 | | | ATGGAGGGGAATACAGCCC |
| 26 | | GR+P1 | CGACTGGAGCACGAGGACACTGA |
| 27 | | | TACAATTTCACCAGGCAAAGTCGATCAT |
| | **Figure 12** | | |

| | | | |
|---|---|---|---|
| 28 | | miR-10b | CCCTGTAGAACCGAATTTGTG |
| 29 | | Hoxd4 ex4/5 | GTCCTGGCTGCTCAGCTACT |
| 30 | | | TGCCCTCCTTACTCACCATC |
| 31 | | P1+P2 | CTATGTCCATTCTGGGGGCT |
| 32 | | | CGGCAGGTAATCTCGGCTT |
| 33 | | P2 | CAC TCA CTT GGC TTA GGT TCT GG |
| 34 | | | AATCTCCCGCCCCCTTATCA |
| 35 | | gapdh | AACGACCCCTTCATTGAC |
| 36 | | | TCCACGACATACTCAGCAC |
| | **Figure** 13 | | |
| 37 | | Drosha | GCCCTGAAAGGAGAAGACCT |
| 38 | | | CCCGGTGCCTGTGTCTCT |
| 39 | | Hoxd4 ex4/5 | GTCCTGGCTGCTCAGCTACT |
| 40 | | | TGCCCTCCTTACTCACCATC |
| 41 | | Hoxd4 ex5/6 | ACC CTT GGA TGA AGA AGG TG |
| 42 | | | GTT GGG CAG TTT GTG GTC TT |
| 43 | | P1+P2 | CTATGTCCATTCTGGGGGCT |
| 44 | | | CGGCAGGTAATCTCGGCTT |
| 45 | | P2 | CAC TCA CTT GGC TTA GGT TCT GG |
| 46 | | | AATCTCCCGCCCCCTTATCA |

### 2.3 GeneRacer® 5’ Rapid Amplification of cDNA Ends

RNA was extracted from P19 cells on day 3 of differentiation using the PureLink™ Micro-to-Midi™ Total RNA Purification System (Invitrogen) and 5’ RACE performed as per manufacturer’s instructions (GeneRacer® Kit, Invitrogen). Sequences of primers used for PCR are summarized in Table 3.

**Table 3 Sequences of primers used in GeneRacer PCR**

| No. | Primer Name | Sequence (5'-to-3') |
|---|---|---|
| 47 | GeneRacer 5’ primer | CGACTGGAGCACGAGGACACTGA |
| 48 | RACE P1-R | TACAATTTCACCAGGCAAAGTCGATCAT |
| 49 | GeneRacer Control Primer B.1 | GACCTGGCCGTCAGGCAGCTCG |
| 50 | RACE Hoxb4-R | TTGCTGAAGGCTGCAGTGTGC |

## 2.4 Ribosome extraction

Day 3 neurally differentiation P19 cells (~$10^6$ cells) were pelleted and washed twice in PBS. Cells were then resuspended in 0.5 ml lysis buffer (10% sucrose, 0.05 M Tris-HCl pH 7.5, 1 mM DTT, 0.1 mM EDTA, 150 mM KCl, 2 mM $MgCl_2$) and lysed by a homogenizer (Polytron PT 1600E, Kinematica) for 3 x 30 seconds. Alternatively, the cells can be lysed by passing through a syringe and a 19 gauge needle several times. The cell lysate was centrifuged at maximum speed in a microcentrifuge at 4ºC for 40 min. Supernatant was loaded onto a 30% sucrose cushion (30% sucrose, 0.05 M Tris-HCl pH 7.5, 1 mM DTT, 0.1 mM EDTA, 300 mM KCl, 2 mM $MgCl_2$) and the ribosomes pelleted after centrifugation at 100 000 rpm at 4ºC for 2 h with a T120 rotor in a TLX 120 Beckman OptimaMAX ultracentrifuge. The supernatant was

removed immediately, leaving the ribosome pellet at the bottom. The pellet was gently washed once with 0.5 ml lysis buffer, resuspended in 100 ul lysis buffer and kept on ice. Phenol-chloroform extraction and ethanol precipitation of the RNA in both the ribosome pellet and supernatant fractions were performed. The RNA obtained from the supernatant and the ribosome pellet was then individually ligated to the RNA oligonucleotide adaptor using T4 RNA ligase (Fermentas). Reverse transcription of mRNA was carried out with SuperScript® III First-Strand Synthesis (Invitrogen) and reverse transcription of miRNA was carried out with NCode™ miRNA First-Strand cDNA Synthesis (Invitrogen) as per manufacturer's instructions. The cDNA was then used for qRT-PCR with primers 12-19 in Table 2.

## 2.5 IRES assay

HEK293T cells were seeded at a density of 1 x $10^5$ cells/ml in a 12-well plate and allowed to attach overnight. Transfection was performed the next day with Lipofectamine2000 (Invitrogen) and 1.2 μg of IRES or inverted IRES plasmids together with 50 ng of Renilla reporter plasmid. Cell lysate was extracted 2 days after transfection. βgal activity was measured using the Dual-

Light® Combined Reporter Gene Assay System (Applied Biosystems) and normalized to Renilla activity which was measured using the Dual-Luciferase™ Reporter Assay System (Promega) on a Fluoroskan Ascent FL (Thermo). Experimental triplicates were performed.

## 2.6 Nuclear-Cytoplasmic fractionation

Day 3 neurally differentiation P19 cells (~$10^6$ cells) were pelleted and washed twice in PBS. Cells were then resuspended in 0.5 ml chilled lysis buffer (10 mM Tris pH 7.4, 3 mM $MgCl_2$, 10 mM NaCl, 150 mM sucrose, 0.5% NP-40) with 10-20 U of RNaseOUT (Invitrogen) and kept on ice for 5-10 min with gentle mixing every minute. The cell lysate was centrifuged at 250 g in a microcentrifuge at 4ºC for 5 min. Supernatant containing the cytoplasmic fraction was collected. The nuclear pellet was washed twice with 1 ml lysis buffer without NP-40 and resuspended in 100 ul of lysis buffer. RNA extraction of both nuclear and cytoplasmic fractions was performed with RNA mini kit (Invitrogen) followed by phenol-chloroform (pH 4.7) extraction and ethanol precipitation. RNA was then resuspended in equal volumes of DEPC-treated water. Equal volumes (1-2 ul) of the total RNA obtained from the

nuclear and cytoplasmic fractions was then individually ligated to the RNA oligonucleotide adaptor using T4 RNA ligase (Fermentas). Reverse transcription of mRNA was carried out with SuperScript® III First-Strand Synthesis (Invitrogen). The cDNA was then used for qRT-PCR with primers 20-27 in Table 2.

## 2.7 RNA *in situ* hybridization on mouse embryos

Staged mouse embryos were fixed overnight at 4ºC in 4% paraformaldehyde and subsequently processed for 10 um paraffin-embedded sections as described (Tribioli and Lufkin, 1999). RNA *in situ* hybridization with digoxigenin (DIG)-labeled probes was performed as previously described (Wang *et al*., 2001). The following plasmids were used as templates for synthesizing antisense DIG-labeled RNA probes: pGEMT *Hoxd4* uP1 and pCR4-TOPO *Hoxd4* P2 exon1-3. Following hybridization and washing, sections were stained with NBT/BCIP and exposed overnight at 4ºC in dark according to manufacturer's instructions (Roche). Sections were subsequently washed in PBS and mounted with glycerol gelatin. All sections were photographed using a Zeiss Axio Imager Z1.

## 2.8 RNA fluorescence *in situ* hybridization (FISH) on P19 cells

Day 0 and Day 3 neurally differentiated P19 cells were typsinized and diluted to a concentration of 7 x $10^5$ cells/ml before being cytospun onto glass slides using a cytocentrifuge, CytoSpin 4 (Thermo Scientific). Cells were then washed with ice-cold PBS for 5 min followed by fixation at room temperature in 4% paraformaldehyde (PFA) for 10 min. The slides were stored in 70% ethanol at 4°C. Before use, the slides were sequentially dehydrated through 80, 90 and 100% ethanol for 2 min each. The RNA FISH probes were labelled with the Nick translation kit and Cy3-dUTP (Roche) as per the manufacturer's instructions. Cot-1 mouse DNA was added to the Cy3-labelled probe and stored in hybridization buffer (50% formamide, 2x SSC pH 7.4, 2 mg/ml BSA, 10% Dextran Sulfate-500K) with a final concentration of 500 ng/ul and 50 ng/ul respectively. Hybridization was carried out at 42°C for 3 hours, followed by three washes of 5 min each in 50% (v/v) formamide in 2x SSC and another three washes of 5 min each in 2x SSC at 45°C. All cells were imaged using a Nikon Eclipse T*i*.

## 2.9 Morpholino-mediated knockdown in P19 cells

P19 cells were seeded at a density of 1 x $10^5$ cells/ml in a 35 mm polystyrene bacterial petri dish (Greiner Bio-One) in complete media and 0.3 μM of ATRA. Cells were transfected with Endo-Porter and 10 μM morpholino oligos (Gene Tools). Sequences of morpholinos used are summarized in Table 4. Fresh media with 0.3 μM ATRA was replaced after two days. Cells were assayed at 48 hpt and qRT-PCR was performed with primers 8-9 and 28-36 in Table 2.

**Table 4 Sequences of morpholinos**

| **Name** | **Sequence (5' → 3')** |
|---|---|
| Splice blocker (mmu *Hoxd4* intron4 SA) | CACAGCCACCAGGAACTCACCAAGC |
| *miR-10b* blocker (Pri-*miR-10b* Drosha*) | ATATTCCCCTAGAATCGAATCTGTG |
| Standard control oligo | CCTCTTACCTCAGTTACAATTTATA |

## 2.10 siRNA-mediated knockdown of Drosha in P19 cells

P19 cells were seeded at a density of 1 x $10^5$ cells/ml in a 35 mm polystyrene bacterial petri dish (Greiner Bio-One) in complete media and 0.3 μM of

ATRA. Cells were transfected with DharmaFect3 and 75 nM α-Drosha siRNApool or the non targeting control siRNA (Dharmacon). Cells were assayed at 48 hpt and qRT-PCR was performed with primers 10-11 and 37-46 in Table 2.

### 2.11 Plasmid constructs

Sequences of primers used for cloning of plasmid constructs are summarized in Table 5.

**Table 5 Sequences of primers for cloning of plasmid constructs**

| No. | Primer Name | Sequence (5'-to-3') |
|---|---|---|
| **51** | P2-2-F | GCTTGAGAGTTGACAAGCCAAA |
| **52** | P2-R | GTGGTTCTCCCTCGTTCTCC |
| **53** | hoxd4 P1-F | TGGTCGATGCAAAAACTTCA |
| **54** | hoxd4 P1-R | GGGTGCAAAATAAGGAAGAATG |
| **55** | Spe1-Hoxd4 ATG-F | AAAAGGACTAGTATGGCCATGAGTTCGTATA |
| **56** | PmeI-lacZ-R | TTGTCCCCGTTTAAACTTATTATTATTTTTGAC |
| **57** | SpeI-IRES(miR10b)-F | ACTAGTGCTGTGCTGAAGAGATCAGG |
| **58** | SpeI-IRES(miR10b)-R | ACTAGTTATTCCCCTAGAATCGAATCTGT |
| **59** | P1-start(4255)-F | TAT GGT CGA TGC AAA AAC TTC A |
| **60** | Hoxd4 ex5(5802)-R | ATC CAA GGG TAG ACC ACA GC |

The pGEMT *Hoxd4* uP1 plasmid used for mouse *in situ* hybridizations was made by T-tailed cloning of an approximately 700 bp PCR fragment amplified from pSNlacZpA (Bloomston *et al*., 2007; Zhang *et al*., 2000) starting from the P1 start site to the 3' splice acceptor of *Hoxd4* intron 4 into the pGEMT vector (Promega). Primers 53 and 54 (Table 5) were used. The pCR4-TOPO *Hoxd4* P2 exon 1-3 plasmid used for mouse *in situ* hybridizations was made by T-tailed cloning of an 287 bp PCR fragment amplified from cDNA of P19 day 4 differentiated cells, comprising of the first three *Hoxd4* exons into pCR4-TOPO vector using the TOPO® TA Cloning® Kit (Invitrogen). Primers 51 and 52 (Table 5) were used.

A non-IRES plasmid was first made by cloning a 3.2 kb fragment amplified from PSNlacZpA (Zhang *et al*., 2000) with sequence starting from the *Hoxd4* ATG start codon and extending throughout the entire *lacZ* coding sequence into SpeI and PmeI sites in the pMIR-REPORT Luciferase plasmid. Primers 55 and 56 (Table 5) were used. Next, a second PCR to amplify an approximately 150 bp fragment starting from approximately 140 bp upstream from the Drosha cleavage site of *miR-10b* to the sequence just 5' of the *Hoxd4* ATG start codon was performed on PSNlacZpA (Zhang *et al*., 2000) with primers 57 and 58 (Table 5). The 150 bp PCR product that consists of *miR-*

*10b* and the 5'UTR of P1 was then cloned via dephosphorylated SpeI sites into the non IRES Reporter plasmid.

The *Hoxd4* P1 plasmid used for RNA FISH was made by cloning a 1.5 kb PCR fragment with primers 59 and 60 (Table 5) into pCR4-TOPO T-tailed vector (Invitrogen).

## 2.12 Ethics statement

All animal procedures were performed according to the National Advisory Committee for Laboratory Animal Research of Singapore guiding principles and the Nanyang Technological University (NTU) Institutional Animal Care and Use Committee (IACUC) guidelines in the School of Biological Sciences Animal Research Facility.

# CHAPTER III. RESULTS

## 3.1 *Hoxd4* and *miR-10b* show similar expression profiles in neurally differentiating P19 cells

To determine if the expression of pri-*miR-10b* is controlled by the *Hoxd4* neural enhancer, the relative levels of *Hoxd4* P1, P2 and *miR-10b* transcripts were measured in differentiating P19 mouse EC cells with quantitative PCR. P19 cells can be induced to undergo neuronal differentiation by treatment with all-*trans* retinoic acid (ATRA) and cell aggregation (Rudnicki and McBurney, 1987). P19 neural differentiation recapitulates aspects of normal embryonic differentiation whereby Hox genes are expressed sequentially in a colinear manner (Baron *et al*., 1987; LaRosa and Gudas, 1988; Rastegar *et al*., 2004; Wei *et al*., 2002)

All *Hoxd4* P1, P2 and *miR-10b* transcripts were present at barely detectable levels in undifferentiated P19 cells (Figure 5B, P1, P2, *miR-10b* at 0 h). The accumulation of *nestin* transcripts, a neural lineage marker, served as a positive control for the P19 differentiation process. *Hoxd4* P1 and P2 transcripts were highly induced during P19 differentiation, peaking at maximum levels on day 3

of differentiation (Figure 5B, P1, P2 at D3). In a fashion similar to the *Hoxd4* transcripts, the *miR-10b* transcripts were first expressed at very low levels in undifferentiated cells and were then strongly induced upon RA treatment and aggregation, peaking at day 4 (Figure 5B, *miR-10b* at D4). When compared to the *Hoxd4* P1 and P2 transcripts, we observed that the *miR-10b* peaked a day later, on day 4. This delay in the expression peak may reflect differences in processing or stability; however, the overall expression profile is similar. This is in contrast to the expression profiles of other Hox genes such as *Hoxa1* which is induced by 6 h and peaks as early as 48 h of neural differentiation (LaRosa and Gudas, 1988).

In summary, *miR-10b* expression was shown to be induced together with *Hoxd4* P1 and P2 transcripts during P19 differentiation. This data is consistent with both *Hoxd4* and *miR-10b* transcripts being under the same regulatory control mechanism *via* the *Hoxd4* 3' neural enhancer.

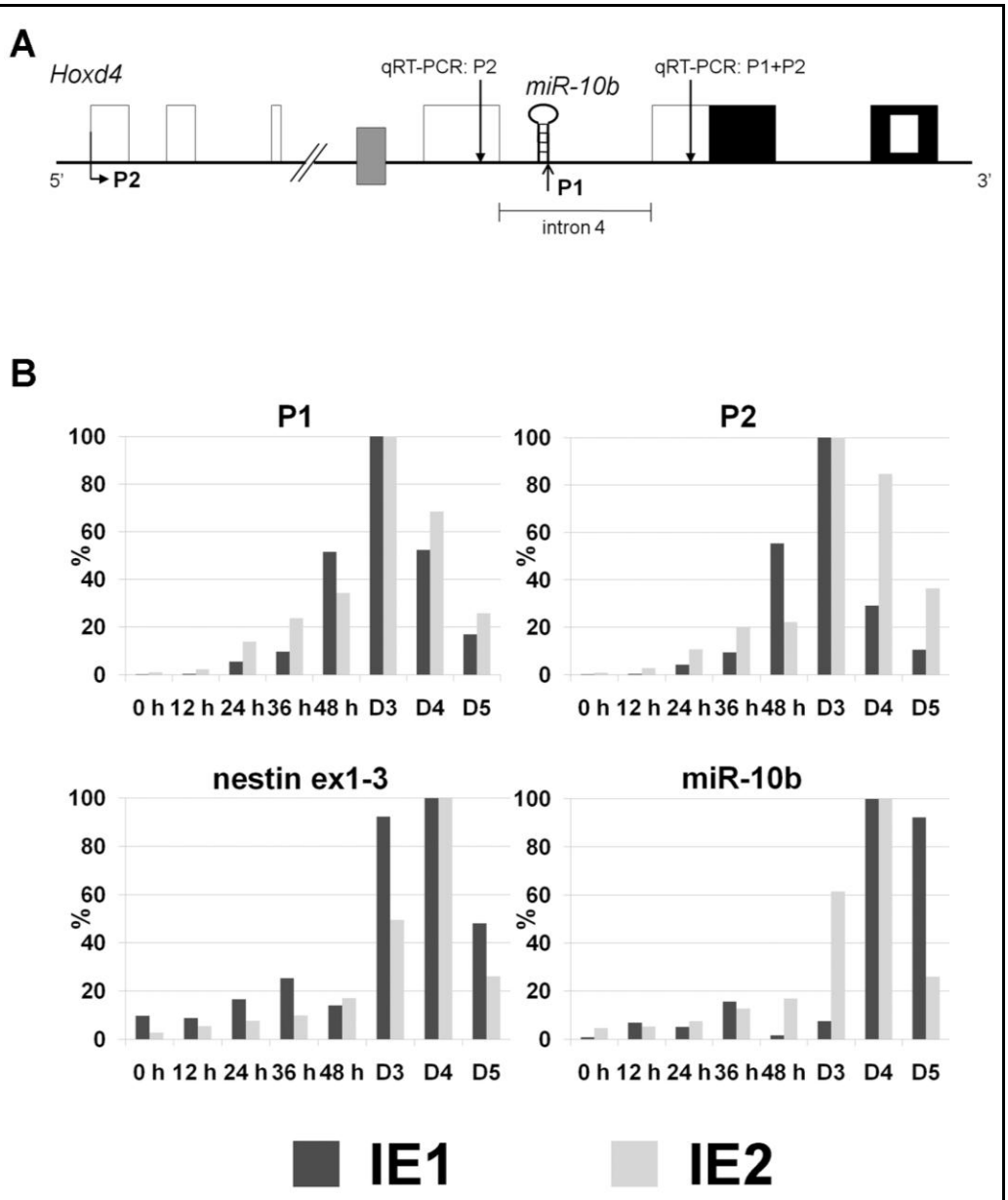


**Figure 5 Expression profiles of *Hoxd4* P1, P2 and *miR-10b* in differentiating P19 cells**

**(A) Schematic showing the position of the qRT-PCR primers used to detect P2 and P1+P2 transcripts respectively.** For specific detection of P2 transcripts only, or P1 plus P2 transcripts, primers were designed to amplify

regions of approximately 200 bp. **(B) qRT-PCR was carried out on cDNA derived from P19 cells undergoing neural differentiation.** The levels of the transcripts were arbitrarily assigned 100% at the time point where it peaks. Transcripts of *nestin*, a marker of neural differentiation, accumulate with time during P19 differentiation, peaking by day 4 (D4). *miR-10b* expression was induced together with *Hoxd4* P1 and P2 transcripts, peaking a day later on day 4 (D4). P1, P2 and nestin were normalized to 18S RNA; *miR-10b* was normalized to U6 snoRNA.

Data represent two independent experiments, each shown as the average of qRT-PCR technical duplicates. Dark grey bars: independent experiment 1 (IE 1); light grey bars: independent experiment 2 (IE 2).

## 3.2 *Hoxd4* transcripts originating from P1 are generated by Drosha cleavage

To determine if *Hoxd4* P1 transcripts are generated by Drosha cleavage or transcriptional initiation, an RNA-ligase mediated 5' Rapid Amplification of cDNA Ends (5' RLM-RACE) was performed. Transcripts whose 5' ends are generated by Drosha cleavage will bear a 5' phosphate and will be suitable substrates for RNA ligase, unlike intact transcripts synthesized by RNA pol II which will have a 5' 7-methylguanosine (5' $m^7G$) cap and therefore be unable to participate in RNA-ligase-mediated reactions.

Total RNA from P19 cells on day 3 of differentiation was collected. An aliquot was treated with calf intestinal phosphatase (CIP) to remove all free 5' phosphates in the total RNA. An aliquot of the CIP-treated RNA was subsequently treated with tobacco acid pyrophosphatase (TAP), which removes the 5' cap on mRNA transcripts to leave a free 5' phosphate. This was followed by ligation of all three RNA samples (untreated, CIP-treated and CIP/TAP-treated) to a synthetic RNA oligonucleotide and RT-PCR. Only transcripts bearing a 5' phosphate, such as produced by Drosha cleavage (Figure 4), can be amplified from untreated RNA samples. The untreated RNA sample gave a 188

bp band after PCR with a *Hoxd4*-sepecific reverse primer (RACE-P1-R), which corresponded to the distance from the primer to the predicted Drosha cleavage site (Figure 6A, white arrowhead). Moreover, no comparably sized products were amplified from the CIP/TAP-treated RNA which would have allowed detection of capped transcripts (Figure 6A). The mRNA for the β-actin control was amplified only following treatment with both CIP and TAP (Figure 6A, white arrow).

The 188 bp band was then cloned and sequenced. The 5' ends of all four clones began with 5'-TATGG-3', mapping precisely to the predicted Drosha cleavage site on the pri-*miR-10b* transcript, 11 nt from the base of the pri-miRNA stem junction (Han *et al*., 2006) (Figure 6D). Importantly, this site is within one nucleotide of the previously deduced P1 start site cluster (underlined in Figure 6D). Together, these observations suggest that P1 transcripts are generated by Drosha cleavage of the primary microRNA for *miR-10b*. The most likely origin for pri-*miR-10b* is a transcript initiating at P2 (Figure 4). In other words, pri-*miR-10b* and the *Hoxd4* P2 transcript are one and the same.

Next, we confirmed the presence of these Drosha-cleaved P1 transcripts in the mouse embryo. The otic vesicle is a morphological landmark that lies just

anterior to the *Hoxd4* expression domain (Figure 7D). Embryonic day (E) E9.5 mouse embryos were bisected just anterior to the developing otic vesicle and RNA extracted from anterior (*Hoxd4*-negative) and posterior (*Hoxd4*-expressing) tissues followed by 5' RLM-RACE. The 188 bp band indicating the presence of Drosha-cleaved P1 transcripts was amplified only from posterior tissue (Figure 6B). This demonstrates that the Drosha-cleaved P1 transcripts are generated during normal mouse embryonic development and only in embryonic regions expressing *Hoxd4*.

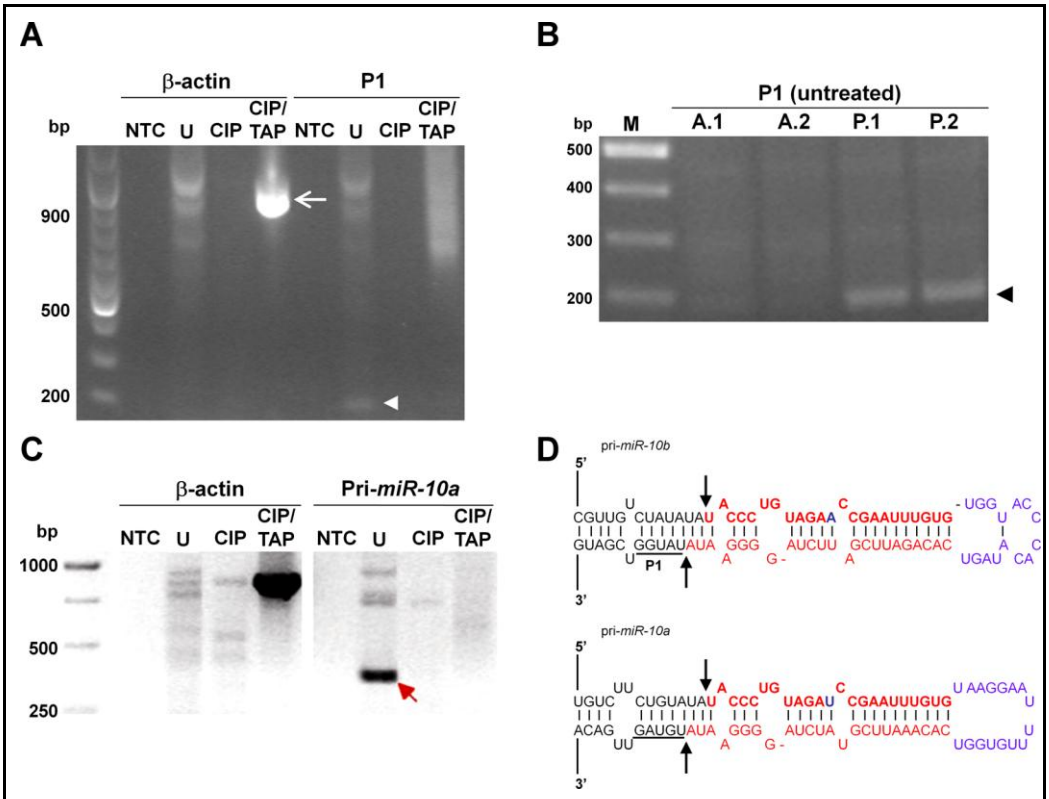


**Figure 6** ***Hoxd4*** **P1 transcripts are generated by Drosha cleavage**

**(A) Presence of uncapped P1 transcripts.** The untreated sample (U) gave an approximately 188 bp amplification product (arrowhead). As a positive control, an expected 900 bp product from β-actin is amplified only when the RNA has been treated with CIP and then TAP (arrow). NTC: no template control; U: untreated RNA; CIP: CIP-treated RNA; CIP/TAP: CIP and TAP-treated RNA. **(B) Uncapped P1 transcripts are present only in posterior tissues of the mouse embryo.** Untreated (U) RNA from the posterior tissues of two different mouse embryos (P.1, P.2) gave a 188 bp amplification product by RT-PCR that corresponds to the presence of uncapped P1 transcripts (black arrowhead). This band is at background levels in anterior tissues of two

mouse embryos (A.1, A.2) where *Hoxd4* is not expressed. M: molecular weight markers. **(C) Presence of uncapped *Hoxb4* transcripts**. The untreated sample (U) gave a 300 bp amplification product (red arrow), indicating the presence of an uncapped *Hoxb4* transcript. Image is the negative of the ethidium bromide stained gel. **(D) Drosha cleavage site of pri-*miR-10b and -10a*.** The nucleotides in red show the miRNA duplex formed by the mature *miR-10b/a* (top strand) and *miR-10b/a** (bottom strand) sequence. Black arrows show the Drosha cleavage sites on the pri-*miR-10b/a* hairpin. The cleavage site is exactly 11 bp from the bottom of both the pri-*miR-10b* and pri-*miR-10a* stem junction on the downstream side. This is within a single nucleotide of the previously mapped P1 start site (a cluster of 4 nt underlined and denoted "P1" on pri-*miR-10b*).

## 3.3 Drosha cleavage of *pri-miR-10a* generates similar *Hoxb4* transcripts

As mentioned, *miR-10a* is located in a paralogous position as *miR-10b*, 5' to *Hoxb4*. qRT-PCR analysis showed that both *miR-10a* and *Hoxb4* transcripts are induced in a similar manner to *miR-10b* and *Hoxd4* transcripts during ATRA-induced P19 neural differentiation. To determine if there was a paralogous Drosha-cleaved *Hoxb4*/pri-*miR-10a* transcript, we attempted to amplify it with a *Hoxb4*-specific reverse primer (RACE-*Hoxb4*-R). Only the untreated RNA sample gave an amplification product of 300 bp, which corresponded to the distance between the reverse primer and the predicted Drosha cleavage site (Figure 6C, red arrow). Sequencing of the *Hoxb4* PCR product confirmed that it is indeed *Hoxb4*-specific, mapping it to the predicted Drosha cleavage site, also precisely 11 nt from the base of the pri-miRNA stem junction (Han *et al*., 2006) (Figure 6D). This indicated the presence of similar uncapped Drosha-cleaved *Hoxb4*/pri-*miR-10a* transcripts in differentiating P19 cells, suggesting that they may belong to a novel class of RNA species.

## 3.4 No capped *miR-10b* transcripts found to originate from the putative Twist-binding region

It was recently found using chromatin immunoprecipitation (ChIP) analysis that a transcriptional factor Twist, binds to a E-box region (E-box 1) that is 313 bp upstream of *miR-10b* (Ma *et al*., 2007). To detect possible capped pri-*miR-10b* transcripts originating from this putative promoter, we performed 5' RLM-RACE. However, after sequencing 12 possible clones, none of which were specific to the *Hoxd4* locus, we conclude that there is no evidence for a novel *miR-10b* promoter near the E-box 1 region. This is consistent with the failure to detect additional *Hoxd4* 5' ends despite previous extensive studies with S1 nuclease and RNase protection assays.

## 3.5 Drosha-cleaved *Hoxd4* P1 and P2/pri-*miR-10b* transcripts are all expressed posterior to the rhombomere 6/7 boundary in the mouse embryo

Using a probe against the *Hoxd4* 5' coding and non-coding region (P1+P2 ex5, Figure 7A), *Hoxd4* transcripts have been shown to have an anterior expression boundary between r6 and r7 within the developing hindbrain of mice (Figure 7D). This expression border is conserved in zebrafish (Folberg *et al*., 1997; Kloosterman *et al*., 2006). By contrast, a probe that specifically detected P2 transcripts (probe name = P2 ex4, Figure 7A) revealed a more posterior expression boundary in the anterior spinal cord above the forelimb bud (Folberg *et al*., 1997). Thus, the anterior part of the *Hoxd4* expression domain up to r 6/7 was attributed to the activity of a distinct P1 promoter. However, given our observation that P1 transcripts are likely generated by Drosha cleavage of P2 transcripts, and therefore do not originate from a transcriptional start site, we re-visited the spatial distribution of P1 and P2 transcripts in the mouse embryo.

We designed two new *in situ* hybridization probes to more precisely investigate the expression of these different transcripts in the embryo. First, a

P1 probe was designed to be just downstream of *miR-10b* and spanning the 5' untranslated region (5' UTR) of the Drosha-cleaved P1 transcripts (probe name = P1, Figure 7A). This region is spliced out in the P2 transcripts and thus the P1 probe will specifically detect P1 transcripts only. A second probe (probe name = P2 ex 1-3, Figure 7A), spans the first 3 exons of the P2 transcript and detects P2 transcripts only. *In situ* hybridizations with both the P1 and P2 ex 1-3 probes showed an expression domain that is posterior to the r 6/7 boundary, similar to the expression domain of mature *miR-10b* in the E9.5 embryo. Neither probe detected more anterior expression up to the r 6/7 boundary (Figure 7B, C, D). We conclude that P1 transcripts are derived by Drosha-cleavage of transcripts originating at the posteriorly active P2 promoter. The more anterior expression (up to r6/7) previously detected with probe P1+P2 ex 5 may be due to the action of a more anteriorly active promoter (transcript P3, Figure 7A). Alternatively, P2 may be the only *Hoxd4* promoter, but the presence or absence of destabilizing elements due to differential splicing determines whether P2-derived transcripts accumulate at anterior or posterior positions (hypothetical transcript P2.3, Figure 7A).

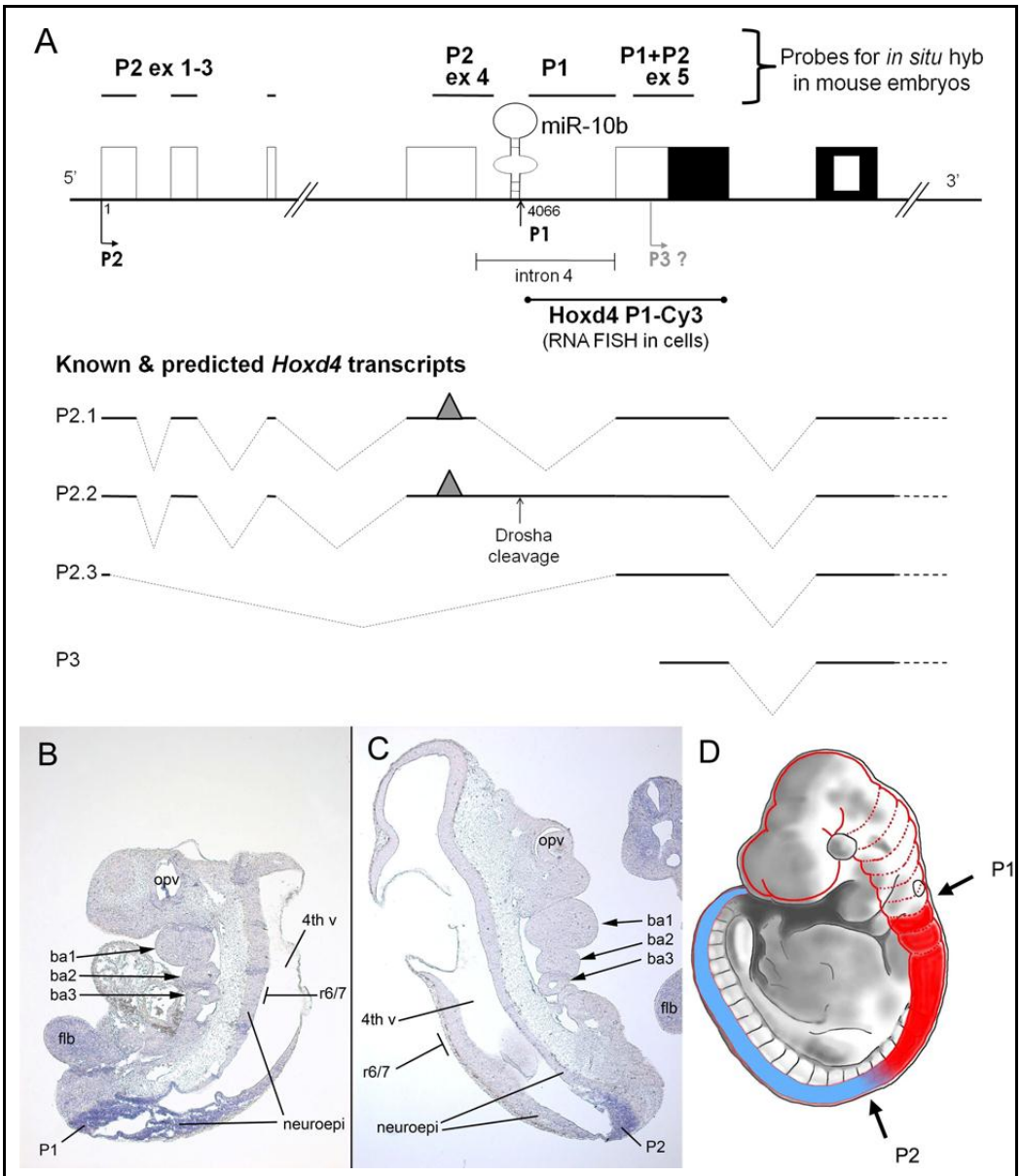


**Figure 7 Distribution of *Hoxd4*/*miR-10b* transcripts along the embryonic AP axis**

**(A) *Hoxd4* probes and transcripts.** The *in situ* hybridization probes are shown above a diagram of the *Hoxd4* genomic locus. Previously, the P2 ex 4

probe spanning exon 4 and extending partially into intron 4, detected P2 transcripts at a point posterior to the r6/7 boundary, while the P1+P2 ex 5 probe spanning exon 5 detected P1+P2 transcripts with a r6/7 boundary. The newly designed P2 ex 1-3 probe spanned the first 3 exons, and detected spliced P2 transcripts, while the P1 probe spans the 5' UTR of P1 transcripts, a region that is spliced out in mature P2 transcripts. The P1 5' end maps to nucleotide 4066 downstream of the major P2 start site. The grey triangle represents a hypothetical mRNA destabilizing element that is absent from hypothetical transcript P2.3. (**B,C) Expression of *Hoxd4* in E9.5 mouse embryos.** Both the Drosha-cleaved P1 transcript detected by the P1 probe **(B)** and the P2 transcript detected by the P2 ex 1-3 probe **(C)** show anterior expression boundaries that are posterior to the anterior limit of the full *Hoxd4* expression domain at r6/7. ba1,2,3: branchial arches 1, 2 and 3; $4^{th}$ v: fourth ventricle within the developing brain; neuroepi: neuroepithelium; flb: forelimb bud; opv: optic vesicle. The r6/7 boundary is approximately opposite branchial arch 3 as indicated. **(D) Schematic of distribution of *Hoxd4* transcripts in the E9.5 mouse embryo.** The solid red plus blue shading represents the cumulative distribution of all *Hoxd4* transcripts. The most anteriorly expressed transcripts (red) have an anterior expression border at the r6/7 boundary. Such an expression pattern is detected with the P1+P2 ex 5 probe (Folberg *et al*., 1997) as well as probes spanning the homeobox-containing exon and 3' UTR (Featherstone *et al*., 1988; Gaunt *et al*., 1989; Nolte *et al*., 2006). The region in blue represents the expression domain of those *Hoxd4* transcripts detected by the more upstream probes, P2 ex 1-3, P2 ex 4 and P1 (Folberg *et al*., 1997 and this thesis). The circle just anterior to the red region represents the otic vesicle.

### 3.6 Drosha-cleaved P1 transcripts are present in the ribosome pellet

We have determined that P1 transcripts are generated by Drosha cleavage. However, these uncapped P1 transcripts appear to be very abundant in differentiating P19 cells and the mouse embryo (Figure 6A and B) (Folberg *et al*., 1997). We therefore asked if these abundant, uncapped P1 transcripts are translated. As a first indication, we asked whether the uncapped P1 transcripts were present in the ribosome fraction, a site of active translation.

High-speed sucrose-gradient sedimentation was used to fractionate cell lysates and pellet the ribosome fraction, followed by ligation of recovered RNA to an RNA oligonucleotide adaptor and qRT-PCR. Our results indicated that the vast majority of β-actin transcripts and up to two-thirds of the *Hoxd4* P1 transcripts were associated with the ribosome pellet compared to the supernatant (Figure 8). This raised the possibility that Drosha-cleaved P1 transcripts are translated.

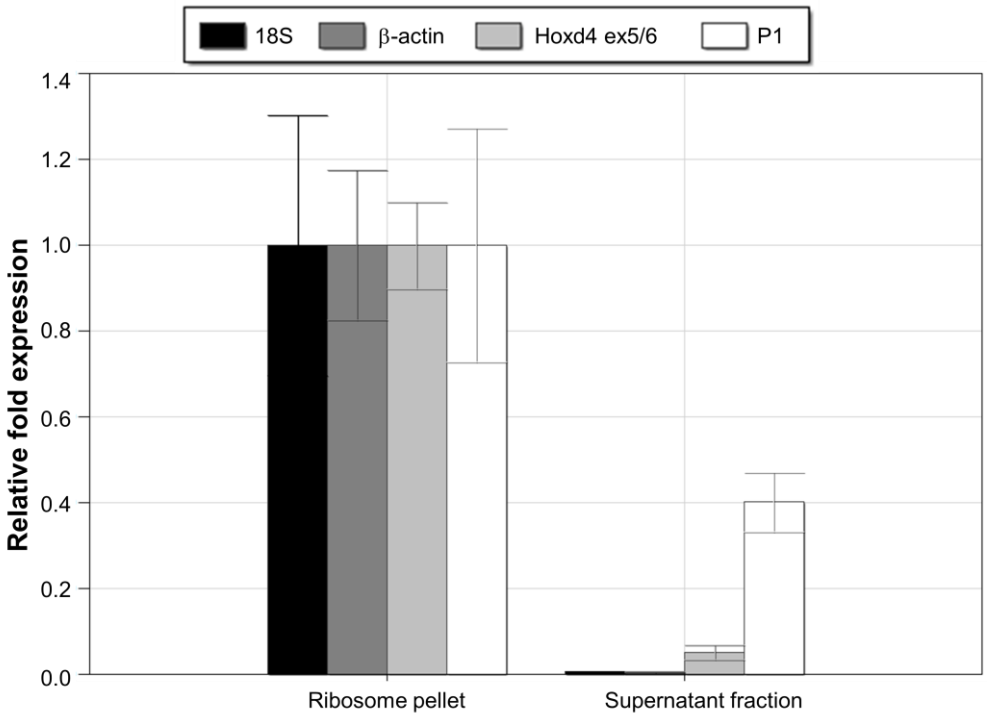


**Figure 8 Association of Drosha-cleaved P1 transcripts with ribosomes**

qRT-PCR of transcripts associated with ribosomes (high-speed pellet) and supernatant from P19 cells on day 3 of differentiation. Both 18S (black) and β-actin (dark grey) served as positive controls abundantly represented in the ribosome pellet. The majority of total *Hoxd4* transcripts (light grey) and P1-specific transcripts (white) were also found to be associated with ribosomes, though proportionately more P1 transcripts were present in the non-ribosome-associated supernatant. All values from the ribosome pellet were set to 1.

## 3.7 Drosha-cleaved P1 transcripts are detected in both the nucleus and cytoplasm

Concurrently, we wanted to determine whether the uncapped P1 transcripts were present in the cytoplasm. Nuclear-cytoplasmic fractionation of the differentiating P19 cells was performed, followed by ligation of recovered RNA to an RNA oligonucleotide adaptor and qRT-PCR. The controls, 47S pre-rRNA and snoU6 RNA were enriched in the nuclear fraction while β-actin mRNA was more abundant in the cytoplasmic fraction (Figure 9). Almost a quarter of the total uncapped Drosha-cleaved P1 transcripts were found in the cytoplasmic fractions, suggesting that these transcripts may be translated. However, the majority of the uncapped P1 transcripts were found in the nuclear fraction, thus raising the possibility that these transcripts may also possess a nuclear function.

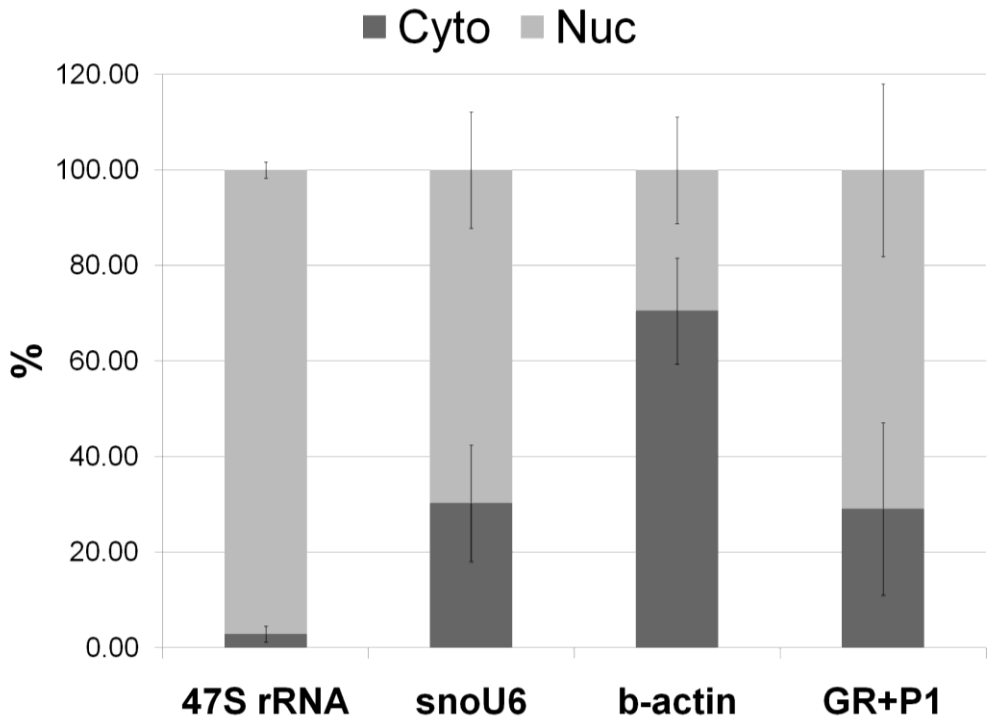


**Figure 9 Distribution of Drosha-cleaved P1 transcripts in nuclear and cytoplasmic fractions**

qRT-PCR of transcripts in nuclear (Nuc, light grey bars) and cytoplasmic (Cyto, dark grey bars) fractions of P19 cells on day 3 of differentiation. 47S pre-rRNA and snoU6 RNA served as positive controls which were enriched in the nuclear fraction while β-actin mRNA was more abundant in the cytoplasmic fraction. Almost three-quarters of the total uncapped Drosha-cleaved P1 transcripts (GR+P1) detected were present in the nuclear fraction. Data shown represent the mean of experimental triplicates (n=3), and error bars give standard deviation.

To further investigate the presence of *Hoxd4* P1 transcripts in the nucleus, we performed RNA fluorescence *in situ* hybridization (RNA FISH) to determine the sub-localization of *Hoxd4* P1 transcripts in day 3 differentiated P19 cells. A Cy3-labelled *Hoxd4* P1 probe spanning 1.5 kb from the Drosha cleavage site and extending to the *Hoxd4* coding region detected nuclear speckles (Figure 10, white arrows in bottom left panel). A control Cy3-labelled probe (pCR2.1 control) had no specific signal (Figure 10, top left panel). As an additional control, no signal was observed in undifferentiated P19 cells where *Hoxd4* and *miR-10b* are not expressed (data not shown).

Taken together, our results indicate the presence of Drosha-cleaved P1 transcripts in both cytoplasm and nucleus, with a majority of the transcripts abundantly localized to nuclear bodies. Therefore, the cytoplasmic Drosha-cleaved P1 transcripts may be translated via a 5' cap-independent mechanism, while the transcripts retained in the nucleus may carry out other functions.

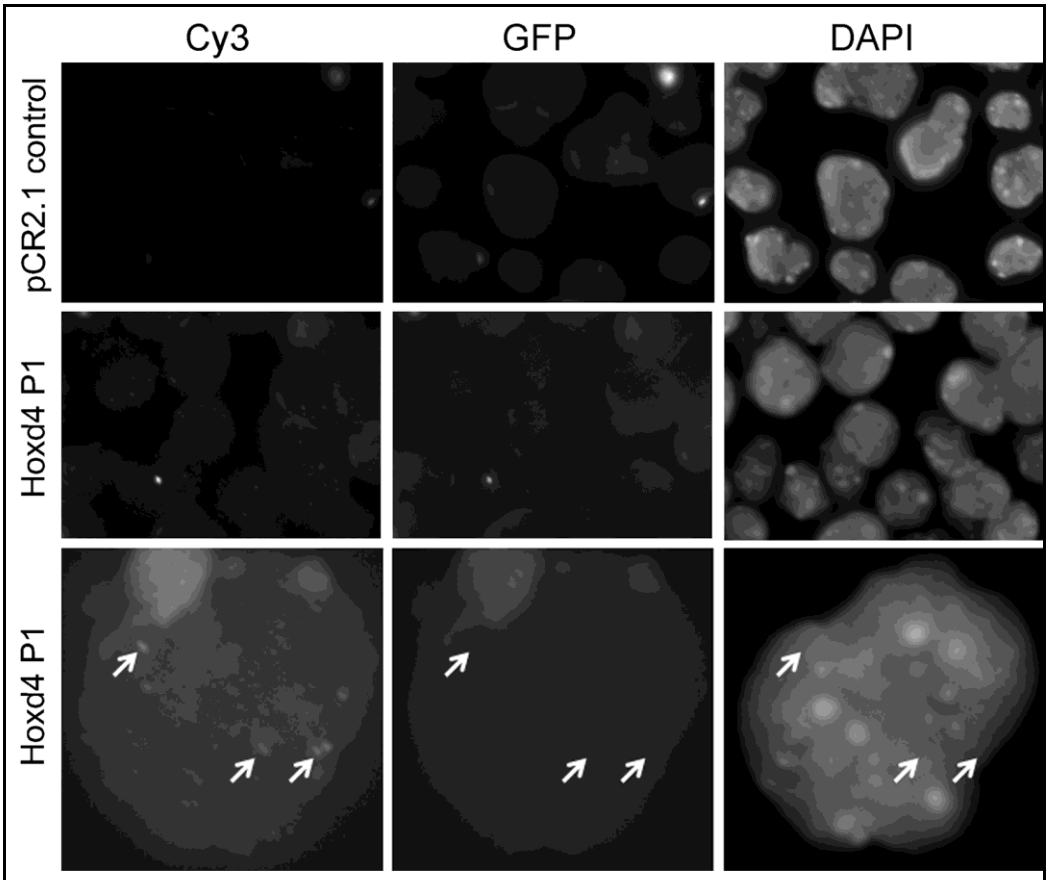


**Figure 10** ***Hoxd4*** **P1 transcripts localize to nuclear bodies**

**RNA FISH of P19 cells on day 3 of neural differentiation.** Vertical columns show images obtained under excitation for Cy3, GFP and DAPI. Images obtained with excitation for GFP control for autofluorescence. The top row shows a negative control Cy3-labelled probe (pCR2.1 control) and the middle and bottom rows show the localization of the Cy3-labelled *Hoxd4* P1 probe (*Hoxd4* P1). *Hoxd4* transcripts are specifically detected as speckles in the nucleus of neurally differentiated P19 cells. Images in the bottom row have been digitally enlarged in order to present the punctate bodies more clearly. The Adobe Photoshop application was used to enhance the contrast *simultaneously* to pairs of images in the Cy3 and GFP columns. Arrows

denote punctate bodies visible with the Cy3-labelled H*oxd4* P1 probe. These bodies are not visible when the sample is excited for GFP and do not correspond to chromatin structures detected by DAPI.

## 3.8 Absence of robust IRES activity within the 5' UTR of P1 transcripts

If the uncapped P1 transcripts are translated, then it should be through an IRES. To test for the ability of the 5'UTR of the uncapped Drosha-cleaved P1 transcript to initiate transcription of the downstream *Hoxd4* coding exons, an IRES reporter assay was designed. The IRES reporter consisted of a firefly luciferase cassette driven by a CMV promoter, followed by the putative IRES region driving *lacZ* gene expression and ending with an SV40 polyadenylation signal. The "Inverted IRES" construct had the putative IRES region inverted and cloned in the same position within the vector and served as a negative baseline control (Figure 11A). Any increase in β-galactosidase levels of the IRES construct compared to the Inverted IRES construct would thus reflect IRES activity.

The IRES test plasmids were co-transfected with control Renilla luciferase expression vectors in either HEK293T cells or RA-treated P19 cells. Cell lysates were assayed at 48 h post-transfection and β-galactosidase was normalized to Renilla luciferase activity to account for variations in transfection efficiency. No significant IRES activity was observed in the

HEK293T cells (Figure 11B). There was also no significant β-galactosidase activity in the transfected RA-treated P19 cells compared to untransfected cells (data not shown). This suggests that the Drosha-cleaved P1 transcripts are not translated through an IRES. This result also indicates the absence of any additional promoter activity in the putative IRES region that stretches for about 1 kilobase pairs upstream from the *Hoxd4* coding region. Efficient cleavage of the IRES reporter fusion transcript by Drosha was evident from the fact that firefly luciferase activity derived from this reporter was low by comparison to that obtained with the Inverted IRES reporter (data not shown).

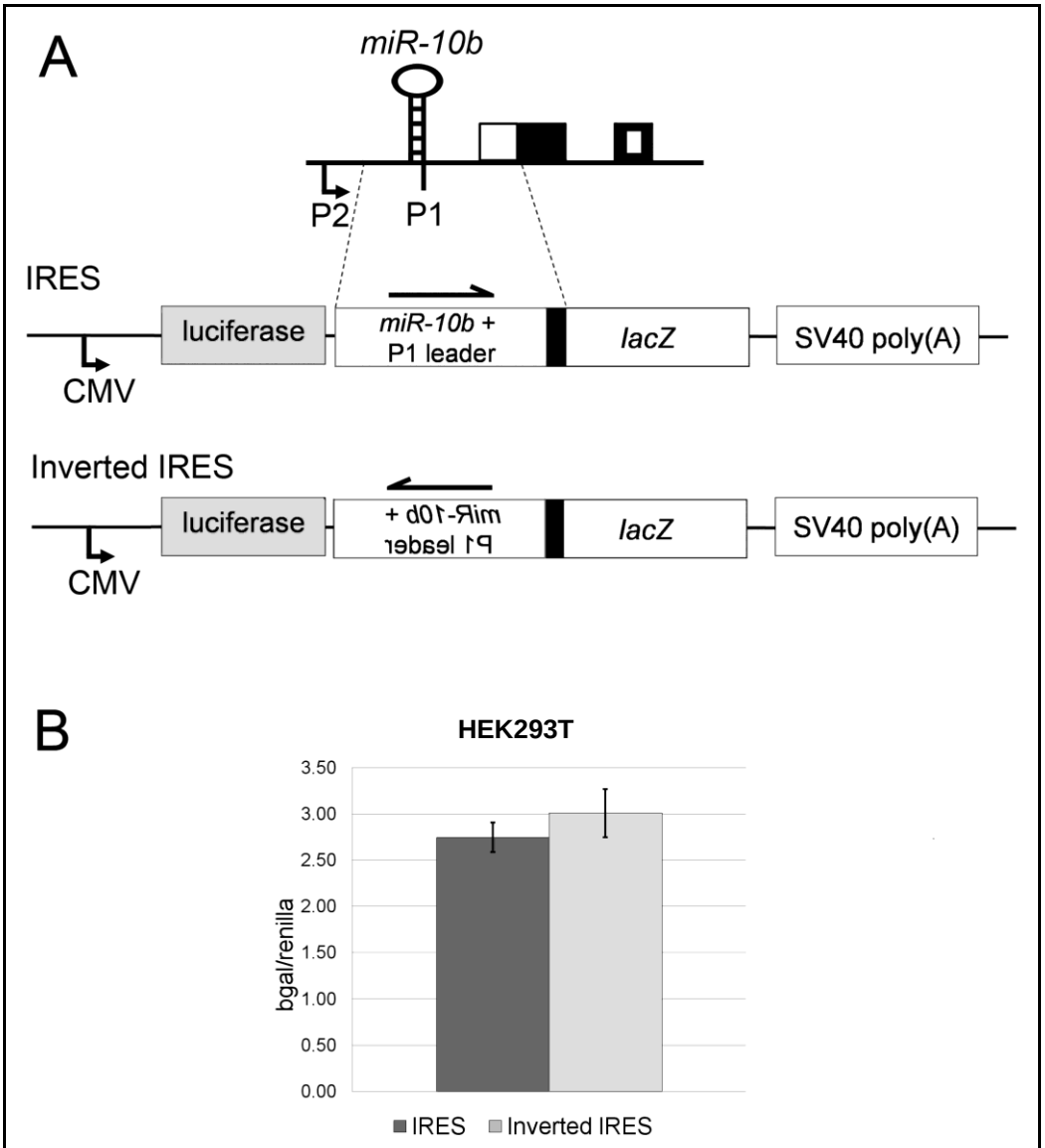


**Figure 11** **Absence of *Hoxd4* IRES activity within the 5' UTR of P1 transcripts**

**(A) Schematic of the IRES and control reporter constructs.** The IRES test construct consists of the luciferase cassette driven by a CMV promoter, followed by the P1 putative IRES region in front of the *lacZ* gene and SV40 polyadenylation signal. The inverted IRES construct consists of the same

components except that the putative IRES region, which is about 1 kb in length, is inverted. Boxes indicate exons as per Figure 1B. **(B) IRES reporter assay in HEK293T cells.** The activity of the IRES construct is not significantly higher than the control inverted IRES construct in either HEK293T cells. β-gal activity was first normalized to Renilla luciferase activity. Data shown represent the mean of experimental triplicates (n=3), and error bars give standard deviation.

### 3.9 Drosha cleavage and splicing of exons 4 and 5 of the pri-*miR-10b*/*Hoxd4* P2 transcript are independent events

Morpholinos are short oligonucleotides that have standard nucleic acid bases bound to 6-membered morpholine rings instead of deoxyribose rings and its bases are linked via non-ionic phosphorodiamidate groups instead of phosphates. Morpholinos act by binding to complementary RNA sequences as competitors to block binding of other proteins or RNA sterically. To determine the effect of blocking either Drosha cleavage or of splicing of pri-*miR-10b*/*Hoxd4* P2 transcript on mature *miR-10b* production and *Hoxd4* transcript levels, differentiating P19 cells were transfected with 10 uM of a *Hoxd4* splice blocker morpholino and a *miR-10b* maturation blocker respectively.

The levels of spliced *Hoxd4* exons 4/5 was heavily downregulated by more than 80% in the presence of the splice blocker, compared to the control morpholino (Figure 12, Student's *t* test, $p<0.05$). Concomitantly, a modest 50% decrease in mature *miR-10b*, *Hoxd4* P1 and P2 levels was observed (Figure 12). However, the change in the level of these transcripts compared to the control was not statistically significant (Student's *t* test, $p>0.05$).

Next, the *miR-10b* blocker resulted in an approximately 90% reduction in mature *miR-10b* levels (Figure 12, Student's *t* test, $p<0.05$). There was however, no significant effect on the splicing *of Hoxd4* intron 4, or on P1 and P2 levels (Figure 12). In other words, successfully blocking Drosha cleavage of pri-*miR-10b* did not influence *Hoxd4* transcript levels.

Therefore, our results suggest that Drosha cleavage of pri-*miR-10b* and splicing of *Hoxd4* P2 exons 4 and 5 are independent events.

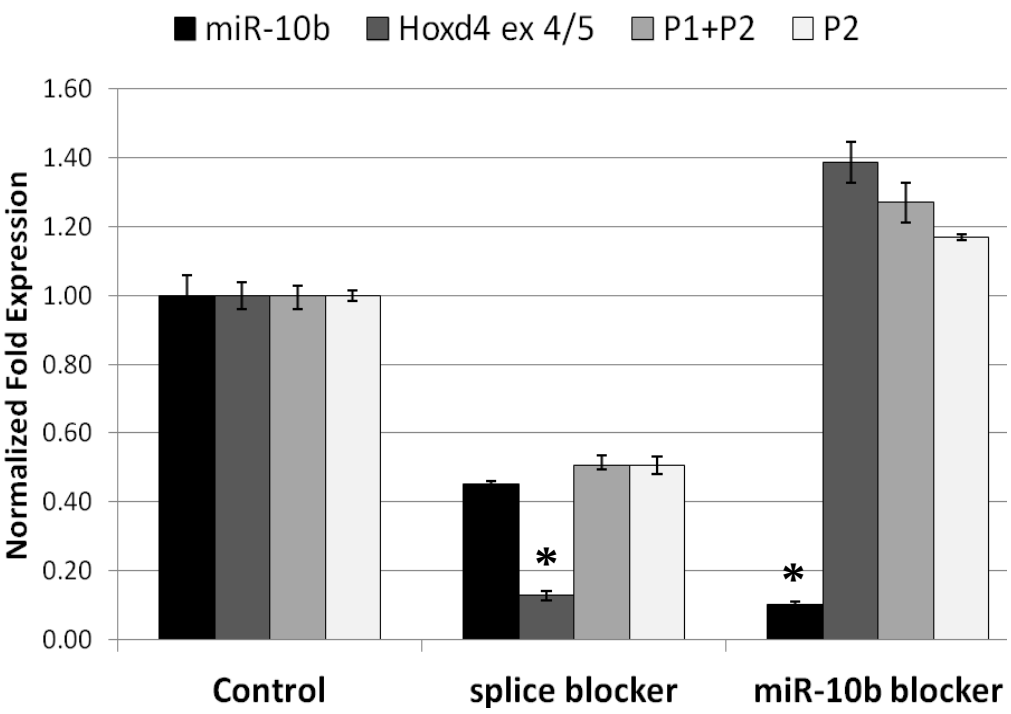


**Figure 12 Effect of blocking *miR-10b* maturation or splicing of *Hoxd4* exons 4 and 5 with morpholinos in differentiating P19 cells**

qRT-PCR of transcripts in day 3 of neurally differentiating P19 cells treated with a control, a *Hoxd4* splice blocking (splice blocker) and a *miR-10b* maturation blocking (miR-10b blocker) morpholino respectively. The splice blocker downregulated the levels of spliced *Hoxd4* exons 4 and 5 products (Student's *t* test, $p<0.05$), but had no statistically significant effect on mature *miR-10b* levels or *Hoxd4* P1 and P2 transcripts. The *miR-10b* blocker downregulated levels of mature *miR-10b* (Student's *t* test, $p<0.05$) but had no significant effect on either the levels of spliced *Hoxd4* or accumulation of *Hoxd4* transcripts.

mRNA (*Hoxd4* ex4/5, P1+P2, P2) were normalized to *gapdh* and *miR-10b* was normalized to U6 snoRNA. qRT-PCR primers used for detection of

*Hoxd4* P1+P2 and P2 transcripts are described in Figure 5A. Black bars: mature *miR-10b*; Dark grey bars: spliced *Hoxd4* exons 4 and 5; light grey bars: *Hoxd4* P1+P2 transcripts; white bars: *Hoxd4* P2 transcripts. Data represent three independent experiments (n=3), error bars represent standard deviation, * denotes a statistically significant difference of $p<0.05$.

### 3.10 Drosha knockdown has no significant effect on *Hoxd4* transcript levels

To investigate if Drosha processing of the pri-*miR-10b*/*Hoxd4* P2 transcript affects splicing and mRNA processing of the same *Hoxd4* transcript, knockdown of Drosha using siRNA was performed. If the pri-*miR-10b*/*Hoxd4* P2 transcripts were a substrate for Drosha processing, we would expect to see an accumulation of P2 transcripts upon Drosha knockdown.

qRT-PCR of transcripts in day 3 of neurally differentiating P19 cells treated with siRNA targeting Drosha (siDrosha) and a non-targeting siRNA control showed a significant 77% decrease in Drosha mRNA levels compared to the control siRNA (Figure 13, Student's *t* test, $p<0.05$). There was no change in the levels of both P1+P2 transcripts and spliced *Hoxd4* exons 5 and 6. While a 2-fold increase in levels of spliced *Hoxd4* exon 4 and 5 and a 1.6-fold increase in *Hoxd4* P2 transcripts were observed, it was not a statistically significant difference (Student's *t* test, $p>0.05$).

In conclusion, knockdown of Drosha did not appear to affect processing or splicing of the *Hoxd4* P2 transcript, which is consistent with the results from

blocking *miR-10b* maturation with morpolinos (Figure 12). It is possible however, that the sample size was not sufficiently large to detect a statistically significant difference between the control and siDrosha-treated samples using a two-tailed Student's *t* test. Alternatively, there is also the possibility that the Drosha protein levels and its activity were not drastically affected by the siRNA knockdown at the mRNA level. In support of that, western blots did not show a significant reduction in Drosha protein levels in siDrosha treated samples (data not shown).

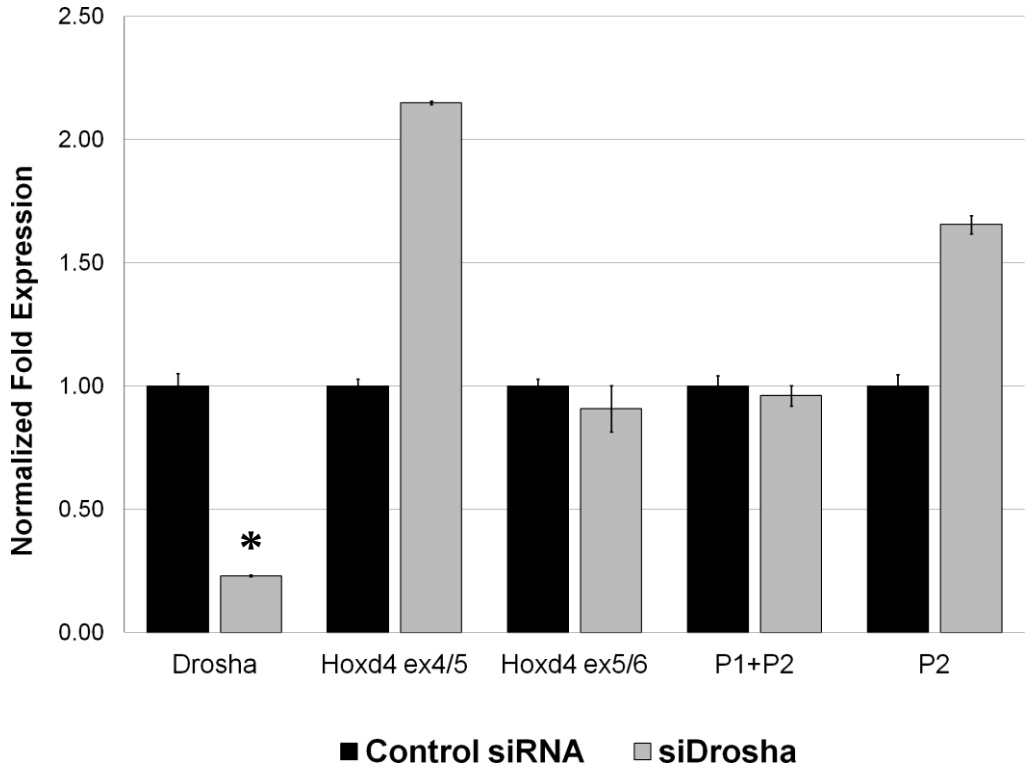


**Figure 13 Effect of Drosha knockdown in differentiating P19 cells**

qRT-PCR of transcripts in day 3 of neutrally differentiating P19 cells treated with siRNA targeting Drosha and a non-targeting siRNA control. There was a significant decrease of 77% in Drosha mRNA levels compared to the control (Student's *t* test, $p<0.05$). Drosha knockdown had no significant effect on *Hoxd4* transcript levels or its splicing. mRNA (*Hoxd4* ex4/5, *Hoxd4* ex5/6, P1, P2) were normalized to 18S. qRT-PCR primers used for detection of *Hoxd4* P1+P2 and P2 transcripts are described in Figure 5A. Black bars: control siRNA; light grey bars: Drosha siRNA. Data represent three independent experiments (n=3), error bars represent standard deviation, * denotes a statistically significant difference of $p<0.05$.

# CHAPTER IV. DISCUSSION

We have established that the 5' ends of *Hoxd4* P1 transcripts are not capped, bear a terminal phosphate and map to the predicted Drosha cleavage site at the base of the stem of the pri-*miR-10b* stem-loop. The 5' end of the *Hoxd4* P1 transcript is precisely 11 nucleotides from the base of the pri-miRNA stem junction. This result is consistent with a validated model in which Drosha cleaves at a position 11 nt from the stem-ssRNA junction (Han *et al*., 2006). We conclude that *Hoxd4* P1 transcripts are indeed the result of Drosha cleavage of the pri-*miR-10b* transcript, and are not generated by transcriptional initiation from a distinct promoter, but by the action of Drosha on transcripts initiated at the P2 promoter. These uncapped P1 transcripts do not possess an IRES and are thus not translated but accumulate within multiple punctate bodies within the nucleus. In addition, similar P1-like uncapped transcripts resulting from Drosha cleavage of pri-*miR-10a* could also be detected. We propose that these transcripts belong to a novel class of noncoding RNAs.

## 4.1 *Hoxd4* and *miR-10b* may share a common promoter and other regulatory elements

A significant majority of human miRNAs resides in intronic regions and in the same orientation as the host coding genes (Rodriguez *et al*., 2004). The expression pattern of these intronic miRNAs frequently coincides with the genes in which they are embedded, indicating that they could be regulated by common *cis*-regulatory elements (Baskerville and Bartel, 2005). Expression of *Hoxd4* transgenes in P19 cells is critically dependent on a 3' neural enhancer (Rastegar *et al*., 2004). The expression profiles of *Hoxd4* and *miR-10b* during P19 differentiation as measured by qRT-PCR are similar, with low basal expression levels in undifferentiated P19 cells and strong induction of their expression upon differentiation peaking at day 3 or 4 and declining thereafter. This is in contrast to the expression profiles of other Hox genes in P19 cells such as *Hoxa1* which has been shown to be induced as early as 6 hours of neural differentiation (LaRosa and Gudas, 1988). Our data are therefore consistent with both *Hoxd4* and *miR-10b* transcripts coming under the control of this same regulatory region.

Primary miRNAs are transcribed by RNA polymerase II and are 5' capped and 3' polyadenylated, making them structurally identical to messenger RNAs. It has been reported that human pri-miRNA transcripts can also simultaneously function as an mRNA coding for a protein (Cai *et al*., 2004). This is consistent with our observation that the P2 transcript codes for both *miR-10b* and HOXD4. Such an arrangement facilitates the co-regulated expression of both *Hoxd4* and *miR-10b* during development as activation of the P2 promoter would lead to production of both *miR-10b* and HOXD4. This is supported by *in situ* data that showed an extensive overlap of mature *miR-10b* and *Hoxd4* expression both spatially and temporally (Folberg *et al*., 1997; Kloosterman *et al*., 2006, and this paper). The coordinated regulation of both genes suggests that they may have shared functions during early development such as has been described for the shared repressive functions of the *miR-10* family and *hoxb4* in zebrafish (Woltering and Durston, 2008).

## 4.2 *Hoxd4* expression along the antero-posterior axis

The full *Hoxd4* expression domain extends anteriorly to the boundary between r6 and r7 in the developing hindbrain (Fig. 4D, combined blue plus red shading)

as revealed by *in situ* probes overlapping the 5' or 3' ends of the coding region (Featherstone *et al*., 1988; Folberg *et al*., 1997; Gaunt *et al*., 1989; Nolte *et al*., 2006). However, *miR-10b* and *Hoxd4* P2 and P1 transcripts are not detected in the anterior-most *Hoxd4* expression domain up to r6/7 (Folberg *et al*., 1997, and this paper). If a single P2 promoter drives expression of all transcripts derived from the *Hoxd4* locus, then there must be post-transcriptional controls which prevent some RNAs from accumulating in anterior tissues up to r6/7. In one possible mechanism, P2 transcripts that include a (hypothetical) destabilizing element located in the *Hoxd4* 5' UTR are unstable and degraded in the anterior-most part of the *Hoxd4* expression domain up to r6/7 (Figure 7, shaded red). In this scenario, only alternatively spliced transcripts that lack this destabilizing element, such as the hypothetical P2.3 transcript shown in Figure 7A, accumulate in the anterior-most domain.

Alternatively, an as-yet uncharacterized promoter (hypothetical promoter P3, Figure 7A) is active in anterior neural tissue and is responsible for expression in the anterior-most portion of the *Hoxd4* expression domain. This is supported by the presence of cDNA and Expressed Sequence Tag (EST) clones (FANTOM) whose 5' ends map immediately upstream of the *Hoxd4* coding region. In addition, a putative human *HOXD4* promoter has been mapped 21 bp 5' of the

ATG start codon (Cianetti *et al*., 1990). However, such *Hoxd4* transcripts may not be expressed in the neural lineage, and we have been unable to detect additional *Hoxd4* 5' ends despite extensive S1 nuclease and RNase protection assays (Folberg *et al*., 1997) and 5'RLM-RACE. In addition, the IRES assays (Figure 11) failed to reveal promoter activity in the 1 kb sequence upstream of the *Hoxd4* ATG start codon. The lack of a promoter at presumptive P3 is further substantiated by a low density of elongating RNA pol II, p300, TBP and H3K4me3 adjacent to the *Hoxd4* coding region. Supporting one or more far upstream promoters, the density of these four factors is high in a broad region spanning P2 (ChIP-seq data of CH12 cells from ENCODE/Stanford/Yale displayed on the UCSC Genome Browser) (Myers *et al*., 2011).

The relationship between the *miR-10* family and *Hox4* genes is surprisingly well conserved through evolution. In the zebrafish genome, three *miR-10* members, *miR-10b-1*, *miR-10b-2*, and *miR-10c*, are positioned upstream of the Hox group 4 paralogs *Hoxd4a*, *hoxc4a* and *hoxb4a*, respectively (Woltering and Durston, 2006). The remaining two *miR-10* family members, *miR-10c* and *miR-10d*, are located at homologous positions near sites from which 4$^{th}$ group paralogs have been lost in the *HoxBa* and (vestigial) *HoxDb* clusters (Woltering and Durston, 2006). The sequences of mature *miR-10b-1* and *miR-*

*10b-2* are identical and they are expressed slightly posterior to the r 6/7 boundary, reminiscent of the situation in the mouse embryo (Folberg *et al*., 1997; Kloosterman *et al*., 2006; Woltering and Durston, 2008).

Long intergenic primary Hox transcripts have been documented from the earliest days of the field (Simeone *et al*., 1988). More recently, high-resolution transcriptional profiling of the Hox clusters has revealed extensive polycistronism and a high degree of transcriptional complexity within the mammalian Hox clusters (Mainguy *et al*., 2007). This increases the possible sources of transcripts acting as primary microRNAs and/or Hox messenger RNAs. A long primary transcript (Genbank BK005082) originating downstream of zebrafish *hoxb5a* spans *miR-10c, hoxb4a* and *hoxb3a* and is expressed in a domain posterior to *hoxb4a* (Hadrys *et al*., 2004; Woltering and Durston, 2008). A similar situation exists in mouse where a long range *Hoxd3* transcript initiated from the *Hoxd4* P2 promoter has the potential to code for *miR-10b* as well (NM_010468). Thus, it may be generally true of Hox complexes that long intergenic transcripts governed by a variety of enhancers contribute to miRNA accumulation at different points along the antero-posterior axis.

## 4.3 Splicing as an important regulatory mechanism for Hox gene expression

No less than seven polycistronic clusters have been identified in the mammalian Hox genes, where different genes are co-transcribed on long isoforms and are alternatively spliced to give different gene products (Mainguy *et al*., 2007). Mutant mice that lack the spliceosomal protein Sf3b1 display severe Hox-like skeletal transformations, indicating that splicing is an essential regulatory component of Hox gene transcription (Isono *et al*., 2005). While *Hoxd4* expression in the paraxial mesoderm was not significantly altered, the authors noted a slight alteration in its expression in the second branchial arch of *Sf3b1* heterozygotic mice (Isono *et al*., 2005). Together, this supports our finding that *miR-10b* and *Hoxd4* share a single promoter and that the splicing machinery may determine whether the potentially bicistronic P2 transcript produces *miR-10b*, HOXD4 protein or even both.

Intronic miRNAs can be processed by Drosha before or after splicing. In fact, Drosha cleavage of these intronic miRNAs before splicing does not appear to affect the production of spliced mRNA and the integrity of the remains of the pri-miRNA (Kim & Kim, 2007). On the other hand, splicing of the *Hoxd4*

P2/pri-*miR-10b* transcript may occur before Drosha cleavage. Recently described in both invertebrates and mammals, mirtrons are short hairpin introns that use splicing to bypass Drosha cleavage, and are alternative precursors for miRNA biogenesis (Ruby *et al*., 2007, Berezikov *et al*., 2007). The spliced and debranched mirtrons do not require further processing by Drosha and can be cleaved by Dicer directly.

Our results from 5' RACE of P1 transcripts supports the former model, in which Drosha cleavage occurs before splicing. However, it does not exclude the latter possibility, in which the spliced out P2 intron 4 bypasses Drosha cleavage and can be directly processed by Dicer to give the mature miRNA.

## 4.4 Control of *miR-10b* versus HOXD4 protein production

HOXD4 protein was not detected in undifferentiated P19 cells but can be detected from day two to four of P19 neural differentiation, with protein levels generally peaking at day three (results not shown). Our current results support a single *Hoxd4* promoter, P2. Prior to splicing out of intron 4, the P2 transcript can be cleaved by Drosha to give pre-*miR-10b* and the uncapped P1 transcript

that cannot be translated. Alternatively, the P2 transcript can be spliced and translated to yield HOXD4 protein. Splicing of intron 4 may be compatible with concurrent processing by Drosha (Kim and Kim, 2007). Thus, *miR-10b* could be produced from both spliced and unspliced P2 transcripts, but HOXD4 protein would only be translated from capped and spliced P2 transcripts that have not been truncated by Drosha. Consequently, it is tempting to speculate that the relative timing of splicing *vs.* cleavage by Drosha could influence the balance between steady state levels of mature *miR-10b* and HOXD4. In this manner, the cell may calibrate the relative ratio of *miR-10b* to HOXD4 protein, a process that could be influenced by Drosha itself.

## 4.5 An alternative E-box-driven promoter

ChIP analysis has been used to implicate the human transcription factor Twist in binding to an E-box region (E-box 1) located 313 bp upstream of the human *miR-10b* hairpin. The study further showed that Twist upregulates *miR-10b* expression and the authors proposed that pri-*miR-10b* transcripts may initiate at a nearby promoter (Ma *et al*., 2007). To detect capped pri-*miR-10b* transcripts originating from this putative promoter, we employed a 5' RLM-RACE

approach with mouse E9.5 RNA. However, no specific clones were recovered, and we were therefore unable to provide evidence for a novel *miR-10b* promoter near the E-box 1 region. This is consistent with a failure to detect additional *Hoxd4* 5' ends despite previous extensive studies with S1 nuclease and RNase protection assays on P19 and whole embryonic RNA (Folberg *et al*., 1997), and further supports a model whereby *Hoxd4* P2 transcripts simultaneously code for both *miR-10b* and HOXD4. We note, however, that our studies in P19 cells have focused on neural-specific expression, while the E-box-directed promoter was implicated in studies on breast cancer cells derived from the mammary epithelium.

## 4.6 Presence of Drosha-cleaved P1 transcripts in the ribosome-associated fractions

Surprisingly, we observed that almost two-thirds of total Drosha-cleaved P1 transcripts were found to be associated with the ribosome fraction compared to the supernatant even though they do not appear to be translated through an IRES (Figure 8 and 11).

Processing bodies (P bodies) are punctate cytoplasmic foci containing aggregates of protein and sequestered RNA that are translation incompetent sites formed as a consequence of the mRNA degradation pathway (Bashkirov *et al*., 1997; Eulalio *et al*., 2007a; Leung *et al*., 2006; van Dijk *et al*., 2002). As all material denser than 30% sucrose will pellet together with the ribosomes during the centrifugation step, it is possible that the P1 transcripts could have been co-sedimented as complexes with the P bodies.

The Drosha-cleaved P1 transcripts could also have co-sedimented with other high molecular weight material. Both splicing and Drosha cleavage is known to occur co-transcriptionally in the nucleus (Morlando *et al*., 2008), and RNA transcripts are also shown to be retained at the transcription sites on the chromatin template (Pawlicki and Steitz, 2008, 2009; Steitz and Vasudevan, 2009). Thus, the Drosha-cleaved P1 transcripts could have co-sedimented together with chromatin and the transcriptional machinery, spliceosomes, protein complexes involved in Drosha cleavage and the tethering of these macromolecules together.

## 4.7 Putative functions of Drosha-cleaved P1 transcripts

Although uncapped transcripts are typically degraded rapidly (Franks and Lykke-Andersen, 2008), the P1 product of Drosha cleavage appears to be both stable and abundant in the mouse embryo and neurally differentiating P19 cells. P1 transcripts retain the potential to be translated and are localized to nuclear speckles in differentiating P19 cells. It is known that the Drosophila *Antp* and *Ubx* transcripts can be efficiently translated via an IRES in their 5' UTRs (Hart and Bienz, 1996; Oh *et al*., 1992). Therefore, the presence of IRES activity within the 5' UTR of the Drosha-cleaved P1 transcripts was examined. However, they do not appear to have an IRES for cap-independent translation (Figure 11).

There is now widespread evidence for the prevalence and importance of actively transcribed ncRNAs in mammalian cells (Birney *et al*., 2007; Carninci *et al*., 2005; Okazaki *et al*., 2002). The highly regulated expression of these ncRNAs in a temporal or spatial manner is suggestive of functional significance. Some Hox-associated ncRNAs which are 5' capped, spliced and polyadenylated, such as HOTAIR and HOTTIP were found to interact with Polycomb/PRC2 and Trithorax/MLL complexes respectively, controlling histone methylation and

chromatin remodeling and thereby silence or direct transcription (Rinn *et al*., 2007; Wang *et al*., 2011). Intriguingly, the HOTAIR ncRNA employs a *trans* mode of action by silencing transcription at a distant HOXD locus while the HOTTIP ncRNA activates gene expression in *cis*, on the proximal HOXA genes. It is thus conceivable that the *Hoxd4* P1 transcripts may perform a similar role by binding chromatin remodeling complexes to direct transcription of genes in *trans* or *cis*.

A search of the FANTOM and mouse ensembl database revealed several other transcripts spanning the *Hoxd4* gene locus, most notably *Hoxd3* (NM_010468) and a differentially spliced noncoding transcript (AK136751). There are, however, no annotated antisense transcripts or known miRNAs that target *Hoxd4* transcripts. Nevertheless, *Hoxd4* P1 transcripts could act as decoy transcripts to the full length *Hoxd4* P2 transcript to sequester either antisense RNA or RNA binding proteins and thereby regulate its own expression post-transcriptionally. This may account for the localization of *Hoxd4* P1 transcripts to nuclear bodies in differentiating P19 cells.

There are known precedents for RNA molecules with dual functions that are able to act as a functional noncoding RNA and a protein coding transcript. For

example, the bacterial small RNA *SgrS* and the *Drosophila* maternal effect gene *oskar* have been reported to function as both a regulatory RNA as well as a protein coding RNA (Jenny *et al*., 2006; Wadler and Vanderpool, 2007). It is conceivable, therefore, that the Drosha-cleaved P1 transcript may have a regulatory or structural role independent of its HOXD4 protein coding ability.

## 4.8 Current limitations and future directions

The P19 EC cell line is a useful model to study the regulation and initiation of transcription at the *Hoxd4* locus, as *Hoxd4* is expressed only upon retinoic acid induction and cell aggregation. It however, has its limitations. As the neural differentiation process in the P19 cells is asynchronous, coupled with the fact that cell aggregation is not a tightly controlled process, the reproducibility of quantitative data by qRT-PCR may be compromised. For example, while we know that *Hoxd4* expression is highly upregulated by day three of neural differentiation, its fold expression (compared to undifferentiated P19 cells) can vary greatly. In addition, *Hoxd4* expression has been found to peak at either day three or four of neural differentiation, but not consistently at a specific time point. As *Hoxd4* expression is induced over

time, its expression level is not maintained at a steady state. There is increased complexity with additional manipulation of transcript levels by siRNA or morpholino-mediated knockdown. The heterogeneity and variability of the data makes quantitative analysis especially challenging. A literature search and analysis of several cell lines was conducted but did not reveal a more amenable neural system in which *Hoxd4* transcripts were constitutively expressed.

The Drosha-cleaved P1 transcript was shown to possess a phosphate group at its 5' end, and cloning of a 188 bp band confirmed its sequence and mapped its position in the *Hoxd4* genomic locus. The P1 transcript was previously identified and mapped using a combination of S1 nuclease analysis, Northern blots and RNase protection assay (Folberg *et al*., 1997). However, it must be noted that the full-length P1 transcript, which includes the *Hoxd4* coding exons, were not experimentally confirmed in this thesis. The high GC content and length of the *Hoxd4* coding region hindered reverse transcription for 5' RACE as well as PCR amplification. Additionally, the 5' end of P2 transcripts originating from the *Hoxd4*/*miR-10b* promoter has yet to be characterized for the presence of 5' $m^7G$ cap structures.

Using the 1.5 kb Cy3-labelled *Hoxd4* P1 probe, *Hoxd4* transcripts were found to localize to nuclear bodies in differentiating P19 cells in a manner suggestive of function. To further explore its function, co-localization studies with known nuclear proteins such as coilin, a Cajal bodies marker, SC35, nuclear speckle markers, PSP1 and SR protein, a marker of paraspeckles, Pc2, a Polycomb body marker and PML antigen, a marker of PML bodies could be performed. Mutational analysis of the *Hoxd4* intron 4 splice donor and acceptor sites could be performed to determine if a splicing defect in that region has any effect on the efficiency *of miR-10b* biogenesis and on the localization of *Hoxd4* transcripts to nuclear bodies. In addition, RNA immunoprecipitation with P1 transcripts coupled with mass spectrometry analysis could help identify possible interacting proteins. We also note that the *Hoxd4* P1 probe may potentially detect any other transcripts that also span this region, such as pri-*miR-10b* and possibly the unspliced *Hoxd4* and *Hoxd3* nascent transcripts, all of which are produced in the nucleus. However, splicing of these nascent transcripts would occur co-transcriptionally, while mature capped and spliced transcripts such as *Hoxd4* P2 transcripts should be efficiently exported to the cytoplasm. To rule out the possibility that the *Hoxd4* P1 probe may have detected transcripts other than P1, we could repeat the RNA FISH with an additional control probe that extends 5’ from the

Drosha cleavage site. We predict that such a probe would yield no or low signal due to the baseline amounts of complementary transcripts.

We have shown that while the expression domain of both *Hoxd4* and *miR-10b* coincides, *Hoxd4* has a more anterior border of expression in the mouse hindbrain, suggesting that post-transcriptional mechanisms that help distinguish the two genes must exist. It will also be important to perform whole mount immunohistochemistry on developing mouse embryos to determine the expression domain of HOXD4 protein with respect to the other *Hoxd4*-associated transcripts. In addition, it will be interesting to look at the translation of HOXD4 protein from various alternatively spliced P2 transcripts.

The Drosha-cleaved P1 transcript may have a regulatory or structural role independent of its HOXD4 protein coding ability. To test for this possibility, we would need to study the effects of depleting these P1 transcripts without affecting the production of either *miR-10b* or HOXD4 protein. One approach would be to mutate the Drosha cleavage site at the *Hoxd4* locus and study the phenotypes of E8.5-10.5 transgenic mouse embryos ($Hoxd4^{D/D}$). Any phenotype observed would be attributed to the loss of both mature *miR-10b* and the P1 transcript. As a control, we could also either mutate the *miR-10b*

seed sequence or simply generate *miR-10b* knockout mice while retaining the Drosha cleavage site. Any defects in *miR-10b*$^{-/-}$ mice would then be due to the loss of *miR-10b* only. If there are no observable phenotypes, transcription profiling using microarrays could be employed to detect any subtle differences in gene expression between *Hoxd4*$^{D/D}$ and *miR-10b*$^{-/-}$ mice.

The Drosophila *miR-10* is predicted to target the neighbouring Hox gene, *Sex combs reduced* (*Scr*), while the zebrafish *miR-10* has been shown to repress *hoxb1a* and *hoxb3a* within the spinal cord of the developing embryo. Therefore, it is conceivable that the murine *miR-10b* could function by regulation of the corresponding mouse Hox genes, *Hoxb1* and *Hoxb3*. To test this possibility, we could knock down *miR-10b* in differentiating P19 cells with morpholinos or siRNAs and determine if there is any significant effect on *Hoxb1* and *Hoxb3* expression levels. In addition, we could also determine the effect of *miR-10b* and *Hoxd4* knockdown in differentiating P19 cells by analyzing changes in neuronal gene expression markers such as nestin.

Moving forward, the zebrafish is an attractive and extremely amenable vertebrate model organism and it would be a suitable model in which we could extend our current *Hoxd4* and *miR-10b* studies. As a start, we would

need to identify the *hoxd4a* and adjacent *miR-10b-2* promoter(s) and characterize the various transcripts to determine if they share the same promoter and whether a potentially bicistronic P2-like transcript exist. Further, we could also attempt to detect and map Drosha-cleaved P1-like transcripts in the zebrafish, as well as determine their spatial distribution in the embryo using *in situ* hybridization. Additional experiments to test the putative functions of these P1-like transcripts would be more easily performed in the zebrafish. For example, we could inject the zebrafish embryos with these uncapped P1-like transcripts and examine both embryos and adults for phenotypic changes.

Aside from investigating a possible co-regulatory relationship between the adjacent *hoxd4a* and *miR-10b-2* genes, we could also examine other pairs of Hox and Hox-associated miRNAs such as *hoxb4a*/*miR-10c* or *hoxa9a*/*miR-196a-2* to determine whether the conservation of their physical proximity on the genome contributes to any functional significance.

# CHAPTER V. CONCLUSION

All life depends on proper gene expression and more complex life forms have evolved many ingenious ways to tightly regulate and control gene expression at the transcriptional and post-transcriptional levels. This thesis has shown that the mouse *Hoxd4* and adjacent *miR-10b* gene share the same promoter and is may be regulated by common *cis*-regulatory elements such as the 3' neural enhancer. This may explain the high level of conservation of the genomic locations of these two genes. The precise mechanism of intronic microRNA biogenesis coupled with co-transcriptional control of its host gene is not very well-studied. Particularly fascinating is the observation that both *Hoxd4* and *miR-10b* share the same promoter and are thus expressed together at the same time and yet, not all cells express both mature *miR-10b* and *Hoxd4* transcripts. This points to the importance of post-transcriptional regulation of both *miR-10b* and *Hoxd4* during early development in the embryo. The origin of *Hoxd4* transcripts at the anterior-most portion of its expression domain up to r6/7 in the developing mouse embryo and the functional relevance of Drosha-cleaved Hox-associated transcripts remains to be elucidated.

## REFERENCES


Abate-Shen, C. (2002). Deregulated homeobox gene expression in cancer: cause or consequence? Nat Rev Cancer *2*, 777-785.

Abe, M., Hamada, J., Takahashi, O., Takahashi, Y., Tada, M., Miyamoto, M., Morikawa, T., Kondo, S., and Moriuchi, T. (2006). Disordered expression of HOX genes in human non-small cell lung cancer. Oncol Rep *15*, 797-802.

Aboobaker, A.A., Tomancak, P., Patel, N., Rubin, G.M., and Lai, E.C. (2005). Drosophila microRNAs exhibit diverse spatial expression patterns during embryonic development. Proc Natl Acad Sci U S A *102*, 18017-18022.

Agirre, X., Jimenez-Velasco, A., San Jose-Eneriz, E., Garate, L., Bandres, E., Cordeu, L., Aparicio, O., Saez, B., Navarro, G., Vilas-Zornoza, A., *et al.* (2008). Down-regulation of hsa-miR-10a in chronic myeloid leukemia CD34+ cells increases USF2-mediated cell growth. Mol Cancer Res *6*, 1830-1840.

Aravin, A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J., and Tuschl, T. (2003). The small RNA profile during Drosophila melanogaster development. Dev Cell *5*, 337-350.

Baron, A., Featherstone, M.S., Hill, R.E., Hall, A., Galliot, B., and Duboule, D. (1987). *Hox-1.6*: A mouse homeobox-containing gene member of the *Hox-1* complex. EMBO J *6*, 2977-2986.

Bashkirov, V.I., Scherthan, H., Solinger, J.A., Buerstedde, J.M., and Heyer, W.D. (1997). A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates. J Cell Biol *136*, 761-773.

Baskerville, S., and Bartel, D.P. (2005). Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. RNA *11*, 241-247.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes Dev *20*, 1885-1898.

Bell, E., Wingate, R.J., and Lumsden, A. (1999). Homeotic transformation of rhombomere identity after localized Hoxb1 misexpression. Science *284*, 2168-2171.

Berezikov, E., Chung, W.J., Willis, J., Cuppen, E., and Lai, E.C. (2007). Mammalian mirtron genes. Mol Cell *28*, 328-336.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., *et al.* (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799-816.

Bloomston, M., Frankel, W.L., Petrocca, F., Volinia, S., Alder, H., Hagan, J.P., Liu, C.G., Bhatt, D., Taccioli, C., and Croce, C.M. (2007). MicroRNA expression patterns to differentiate pancreatic adenocarcinoma from normal pancreas and chronic pancreatitis. JAMA *297*, 1901-1908.

Bomgardner, D., Hinton, B.T., and Turner, T.T. (2001). Hox transcription factors may play a role in regulating segmental function of the adult epididymis. J Androl *22*, 527-531.

Bridges, C.B., and Morgan, T.H. (1923). The third-chromosome group of mutant characters of Drosophila melanogaster (Washington,, Carnegie Institution of Washington).

Bushati, N., Stark, A., Brennecke, J., and Cohen, S.M. (2008). Temporal reciprocity of miRNAs and their targets during the maternal-to-zygotic transition in Drosophila. Curr Biol *18*, 501-506.

Cai, X., Hagedorn, C.H., and Cullen, B.R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA *10*, 1957-1966.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., *et al.* (2005). The transcriptional landscape of the mammalian genome. Science *309*, 1559-1563.

Carpenter, E.M., Goddard, J.M., Chisaka, O., Manley, N.R., and Capecchi, M.R. (1993). Loss of Hox-A1 (Hox-1.6) function results in the reorganization of the murine hindbrain. Development *118*, 1063-1075.

Chatterjee, S., and Grosshans, H. (2009). Active turnover modulates mature microRNA activity in Caenorhabditis elegans. Nature *461*, 546-549.

Chen, C.Z., Li, L., Lodish, H.F., and Bartel, D.P. (2004). MicroRNAs modulate hematopoietic lineage differentiation. Science *303*, 83-86.

Ciafre, S.A., Galardi, S., Mangiola, A., Ferracin, M., Liu, C.G., Sabatino, G., Negrini, M., Maira, G., Croce, C.M., and Farace, M.G. (2005). Extensive modulation of a set of microRNAs in primary glioblastoma. Biochem Biophys Res Commun *334*, 1351-1358.

Cianetti, L., Di Cristofaro, A., Zappavigna, V., Bottero, L., Boccoli, G., Testa, U., Russo, G., Boncinelli, E., and Peschle, C. (1990). Molecular mechanisms underlying the expression of the human HOX-5.1 gene. Nucleic Acids Res *18*, 4361-4368.

Cordes, S.P. (2001). Molecular genetics of cranial nerve development in mouse. Nat Rev Neurosci *2*, 611-623.

Davis, A.P., and Capecchi, M.R. (1994). Axial homeosis and appendicular skeleton defects in mice with a targeted disruption of hoxd-11. Development *120*, 2187-2198.

Davis, A.P., and Capecchi, M.R. (1996). A mutational analysis of the 5' HoxD genes: dissection of genetic interactions during limb development in the

mouse. Development *122*, 1175-1185.

Davis, A.P., Witte, D.P., Hsieh-Li, H.M., Potter, S.S., and Capecchi, M.R. (1995). Absence of radius and ulna in mice lacking hoxa-11 and hoxd-11. Nature *375*, 791-795.

Debernardi, S., Skoulakis, S., Molloy, G., Chaplin, T., Dixon-McIver, A., and Young, B.D. (2007). MicroRNA miR-181a correlates with morphological sub-class of acute myeloid leukaemia and the expression of its target genes in global genome-wide analysis. Leukemia *21*, 912-916.

Dolle, P., Dierich, A., LeMeur, M., Schimmang, T., Schuhbaur, B., Chambon, P., and Duboule, D. (1993). Disruption of the Hoxd-13 gene induces localized heterochrony leading to mice with neotenic limbs. Cell *75*, 431-441.

Dolle, P., Izpisua-Belmonte, J.C., Falkenstein, H., Renucci, A., and Duboule, D. (1989). Coordinate expression of the murine Hox-5 complex homoeobox-containing genes during limb pattern formation. Nature *342*, 767-772.

Du, H., and Taylor, H.S. (2004). Molecular regulation of mullerian development by Hox genes. Ann N Y Acad Sci *1034*, 152-165.

Duboule, D., and Dolle, P. (1989). The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. EMBO J *8*, 1497-1505.

Duboule, D., and Morata, G. (1994). Colinearity and functional hierarchy among genes of the homeotic complexes. Trends Genet *10*, 358-364.

Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. (2007a). P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. Mol Cell Biol *27*, 3970-3981.

Eulalio, A., Huntzinger, E., and Izaurralde, E. (2008). Getting to the root of miRNA-mediated gene silencing. Cell *132*, 9-14.

Eulalio, A., Rehwinkel, J., Stricker, M., Huntzinger, E., Yang, S.F., Doerks, T., Dorner, S., Bork, P., Boutros, M., and Izaurralde, E. (2007b). Target-specific requirements for enhancers of decapping in miRNA-mediated gene silencing. Genes Dev *21*, 2558-2570.

Favier, B., Le Meur, M., Chambon, P., and Dolle, P. (1995). Axial skeleton homeosis and forelimb malformations in Hoxd-11 mutant mice. Proc Natl Acad Sci U S A *92*, 310-314.

Featherstone, M.S., Baron, A., Gaunt, S.J., Mattei, M.G., and Duboule, D. (1988). Hox-5.1 defines a homeobox-containing gene locus on mouse chromosome 2. Proc Natl Acad Sci U S A *85*, 4760-4764.

Filipowicz, W., Bhattacharyya, S.N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet *9*, 102-114.

Folberg, A., Kovacs, E.N., and Featherstone, M.S. (1997). Characterization and retinoic acid responsiveness of the murine Hoxd4 transcription unit. J Biol Chem *272*, 29151-29157.

Franks, T.M., and Lykke-Andersen, J. (2008). The control of mRNA decapping and P-body formation. Mol Cell *32*, 605-615.

Friedman, R.C., Farh, K.K., Burge, C.B., and Bartel, D.P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. Genome Res *19*, 92-105.

Fritzsch, B. (1998). Of mice and genes: evolution of vertebrate brain development. Brain Behav Evol *52*, 207-217.

Fromental-Ramain, C., Warot, X., Messadecq, N., LeMeur, M., Dolle, P., and Chambon, P. (1996). Hoxa-13 and Hoxd-13 play a crucial role in the patterning of the limb autopod. Development *122*, 2997-3011.

Garzon, R., Garofalo, M., Martelli, M.P., Briesewitz, R., Wang, L., Fernandez-Cymering, C., Volinia, S., Liu, C.G., Schnittger, S., Haferlach, T., *et al.* (2008). Distinctive microRNA signature of acute myeloid leukemia

bearing cytoplasmic mutated nucleophosmin. Proc Natl Acad Sci U S A *105*, 3945-3950.

Garzon, R., Pichiorri, F., Palumbo, T., Iuliano, R., Cimmino, A., Aqeilan, R., Volinia, S., Bhatt, D., Alder, H., Marcucci, G., *et al.* (2006). MicroRNA fingerprints during human megakaryocytopoiesis. Proc Natl Acad Sci U S A *103*, 5078-5083.

Gaunt, S.J., Krumlauf, R., and Duboule, D. (1989). Mouse homeo-genes within a subfamily, Hox-1.4, -2.6 and -5.1, display similar anteroposterior domains of expression in the embryo, but show stage- and tissue-dependent differences in their regulation. Development *107*, 131-141.

Gehring, W.J., Qian, Y.Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A.F., Resendez-Perez, D., Affolter, M., Otting, G., and Wuthrich, K. (1994). Homeodomain-DNA recognition. Cell *78*, 211-223.

Giampaolo, A., Acampora, D., Zappavigna, V., Pannese, M., D'Esposito, M., Care, A., Faiella, A., Stornaiuolo, A., Russo, G., Simeone, A., *et al.* (1989). Differential expression of human HOX-2 genes along the anterior-posterior axis in embryonic central nervous system. Differentiation *40*, 191-197.

Giraldez, A.J., Cinalli, R.M., Glasner, M.E., Enright, A.J., Thomson, J.M., Baskerville, S., Hammond, S.M., Bartel, D.P., and Schier, A.F. (2005). MicroRNAs regulate brain morphogenesis in zebrafish. Science *308*, 833-838.

Gould, A., Itasaki, N., and Krumlauf, R. (1998). Initiation of rhombomeric Hoxb4 expression requires induction by somites and a retinoid pathway. Neuron *21*, 39-51.

Graham, A., Papalopulu, N., and Krumlauf, R. (1989). The murine and Drosophila homeobox gene complexes have common features of organization and expression. Cell *57*, 367-378.

Greer, J.M., Puetz, J., Thomas, K.R., and Capecchi, M.R. (2000). Maintenance of functional equivalence during paralogous Hox gene evolution. Nature

*403*, 661-665.

Griffiths-Jones, S. (2004). The microRNA Registry. Nucleic Acids Res *32*, D109-111.

Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A., and Enright, A.J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res *34*, D140-144.

Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.C., Hung, T., Argani, P., Rinn, J.L., *et al.* (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature *464*, 1071-1076.

Hadrys, T., Prince, V., Hunter, M., Baker, R., and Rinkwitz, S. (2004). Comparative genomic analysis of vertebrate Hox3 and Hox4 genes. J Exp Zool B Mol Dev Evol *302*, 147-164.

Hagan, J.P., Piskounova, E., and Gregory, R.I. (2009). Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nat Struct Mol Biol *16*, 1021-1025.

Han, J., Lee, Y., Yeom, K.-H., Nam, J.-W., Heo, I., Rhee, J.-K., Sohn, S.Y., Cho, Y., Zhang, B.-T., and Kim, V.N. (2006). Molecular Basis for the Recognition of Primary microRNAs by the Drosha-DGCR8 Complex. Cell *125*, 887-901.

Hart, K., and Bienz, M. (1996). A test for cell autonomy, based on di-cistronic messenger translation. Development *122*, 747-751.

Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V.N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Mol Cell *32*, 276-284.

Horan, G.S., Kovacs, E.N., Behringer, R.R., and Featherstone, M.S. (1995a). Mutations in paralogous Hox genes result in overlapping homeotic transformations of the axial skeleton: evidence for unique and redundant function. Dev Biol *169*, 359-372.

Horan, G.S., Ramirez-Solis, R., Featherstone, M.S., Wolgemuth, D.J., Bradley, A., and Behringer, R.R. (1995b). Compound mutants for the paralogous hoxa-4, hoxb-4, and hoxd-4 genes show more complete homeotic transformations and a dose-dependent increase in the number of vertebrae transformed. Genes Dev *9*, 1667-1677.

Horan, G.S., Wu, K., Wolgemuth, D.J., and Behringer, R.R. (1994). Homeotic transformation of cervical vertebrae in Hoxa-4 mutant mice. Proc Natl Acad Sci U S A *91*, 12644-12648.

Houbaviy, H.B., Murray, M.F., and Sharp, P.A. (2003). Embryonic stem cell-specific MicroRNAs. Dev Cell *5*, 351-358.

Huang, H., Xie, C., Sun, X., Ritchie, R.P., Zhang, J., and Chen, Y.E. (2010). miR-10a contributes to retinoid acid-induced smooth muscle cell differentiation. J Biol Chem *285*, 9383-9389.

Humphreys, D.T., Westman, B.J., Martin, D.I., and Preiss, T. (2005). MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. Proc Natl Acad Sci U S A *102*, 16961-16966.

Hunt, P., and Krumlauf, R. (1992). Hox codes and positional specification in vertebrate embryonic axes. Annu Rev Cell Biol *8*, 227-256.

Hunt, P., Whiting, J., Muchamore, I., Marshall, H., and Krumlauf, R. (1991a). Homeobox genes and models for patterning the hindbrain and branchial arches. Dev Suppl *1*, 187-196.

Hunt, P., Wilkinson, D., and Krumlauf, R. (1991b). Patterning the vertebrate head: murine Hox 2 genes mark distinct subpopulations of premigratory and migrating cranial neural crest. Development *112*, 43-50.

Isono, K., Mizutani-Koseki, Y., Komori, T., Schmidt-Zachmann, M.S., and Koseki, H. (2005). Mammalian polycomb-mediated repression of Hox genes requires the essential spliceosomal protein Sf3b1. Genes Dev *19*, 536-541.

Jenny, A., Hachet, O., Zavorszky, P., Cyrklaff, A., Weston, M.D., Johnston, D.S., Erdelyi, M., and Ephrussi, A. (2006). A translation-independent role of oskar RNA in early Drosophila oogenesis. Development *133*, 2827-2833.

Jung, C., Kim, R.S., Lee, S.J., Wang, C., and Jeng, M.H. (2004). HOXB13 homeodomain protein suppresses the growth of prostate cancer cells by the negative regulation of T-cell factor 4. Cancer Res *64*, 3046-3051.

Jungbluth, S., Bell, E., and Lumsden, A. (1999). Specification of distinct motor neuron identities by the singular activities of individual Hox genes. Development *126*, 2751-2758.

Kai, Z.S., and Pasquinelli, A.E. (2010). MicroRNA assassins: factors that regulate the disappearance of miRNAs. Nat Struct Mol Biol *17*, 5-10.

Kappen, C., Schughart, K., and Ruddle, F.H. (1989). Two steps in the evolution of Antennapedia-class vertebrate homeobox genes. Proc Natl Acad Sci U S A *86*, 5459-5463.

Katoh, T., Sakaguchi, Y., Miyauchi, K., Suzuki, T., Kashiwabara, S., and Baba, T. (2009). Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. Genes Dev *23*, 433-438.

Kawahara, Y., Zinshteyn, B., Chendrimada, T.P., Shiekhattar, R., and Nishikura, K. (2007a). RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex. EMBO Rep *8*, 763-769.

Kawahara, Y., Zinshteyn, B., Sethupathy, P., Iizasa, H., Hatzigeorgiou, A.G., and Nishikura, K. (2007b). Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. Science *315*, 1137-1140.

Kawazoe, Y., Sekimoto, T., Araki, M., Takagi, K., Araki, K., and Yamamura, K. (2002). Region-specific gastrointestinal Hox code during murine embryonal gut development. Dev Growth Differ *44*, 77-84.

Kessel, M., and Gruss, P. (1991). Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. Cell

*67*, 89-104.

Khraiwesh, B., Arif, M.A., Seumel, G.I., Ossowski, S., Weigel, D., Reski, R., and Frank, W. (2010). Transcriptional control of gene expression by microRNAs. Cell *140*, 111-122.

Khvorova, A., Reynolds, A., and Jayasena, S.D. (2003). Functional siRNAs and miRNAs Exhibit Strand Bias. Cell *115*, 209-216.

Kim, D.H., Saetrom, P., Snove, O., Jr., and Rossi, J.J. (2008). MicroRNA-directed transcriptional gene silencing in mammalian cells. Proc Natl Acad Sci U S A *105*, 16230-16235.

Kim, V.N., Heo, I., Joo, C., Kim, Y.K., Ha, M., Yoon, M.J., Cho, J., Yeom, K.H., and Han, J. (2009). TUT4 in Concert with Lin28 Suppresses MicroRNA Biogenesis through Pre-MicroRNA Uridylation. Cell *138*, 696-708.

Kim, Y.K., and Kim, V.N. (2007). Processing of intronic microRNAs. EMBO J *26*, 775-783.

Kiriakidou, M., Tan, G.S., Lamprinaki, S., De Planell-Saguer, M., Nelson, P.T., and Mourelatos, Z. (2007). An mRNA m7G cap binding-like motif within human Ago2 represses translation. Cell *129*, 1141-1151.

Kloosterman, W.P., Wienholds, E., de Bruijn, E., Kauppinen, S., and Plasterk, R.H. (2006). In situ detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes. Nat Methods *3*, 27-29.

Kosman, D., Mizutani, C.M., Lemons, D., Cox, W.G., McGinnis, W., and Bier, E. (2004). Multiplex detection of RNA expression in Drosophila embryos. Science *305*, 846.

Kostic, D., and Capecchi, M.R. (1994). Targeted disruptions of the murine Hoxa-4 and Hoxa-6 genes result in homeotic transformations of components of the vertebral column. Mech Dev *46*, 231-247.

Krumlauf, R. (1994). Hox genes in vertebrate development. Cell *78*, 191-201.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001).

Identification of novel genes coding for small expressed RNAs. Science *294*, 853-858.

Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W., and Tuschl, T. (2002). Identification of tissue-specific microRNAs from mouse. Curr Biol *12*, 735-739.

LaRosa, G.J., and Gudas, L.J. (1988). Early Retinoic Acid-Induced F9 Teratocarcinoma Stem Cell Gene ERA-I: Alternate Splicing Creates Transcripts for a Homeobox-Containing Protein and One Lacking the Homeobox. MOLECULAR AND CELLULAR BIOLOGY *8*, 3906-3917.

Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science *294*, 858-862.

Le Mouellic, H., Lallemand, Y., and Brulet, P. (1992). Homeosis in the mouse induced by a null mutation in the Hox-3.1 gene. Cell *69*, 251-264.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.

Lee, Y., Jeon, K., Lee, J.T., Kim, S., and Kim, V.N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. EMBO J *21*, 4663-4670.

Lempradl, A., and Ringrose, L. (2008). How does noncoding transcription regulate Hox genes? Bioessays *30*, 110-121.

Leung, A.K., Calabrese, J.M., and Sharp, P.A. (2006). Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules. Proc Natl Acad Sci U S A *103*, 18125-18130.

Lumsden, A., and Keynes, R. (1989). Segmental patterns of neuronal development in the chick hindbrain. Nature *337*, 424-428.

Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. Science *303*, 95-98.

Ma, L., Reinhardt, F., Pan, E., Soutschek, J., Bhat, B., Marcusson, E.G., Teruya-Feldstein, J., Bell, G.W., and Weinberg, R.A. (2010). Therapeutic silencing of miR-10b inhibits metastasis in a mouse mammary tumor model. Nat Biotechnol *28*, 341-347.

Ma, L., Teruya-Feldstein, J., and Weinberg, R.A. (2007). Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. Nature *449*, 682-688.

Mainguy, G., Koster, J., Woltering, J., Jansen, H., and Durston, A. (2007). Extensive polycistronism and antisense transcription in the mammalian Hox clusters. PLoS One *2*, e356.

Mansfield, J.H., Harfe, B.D., Nissen, R., Obenauer, J., Srineel, J., Chaudhuri, A., Farzan-Kashani, R., Zuker, M., Pasquinelli, A.E., Ruvkun, G., *et al.* (2004). MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. Nat Genet *36*, 1079-1083.

Maroney, P.A., Yu, Y., and Nilsen, T.W. (2006). MicroRNAs, mRNAs, and translation. Cold Spring Harb Symp Quant Biol *71*, 531-535.

Mathonnet, G., Fabian, M.R., Svitkin, Y.V., Parsyan, A., Huck, L., Murata, T., Biffo, S., Merrick, W.C., Darzynkiewicz, E., Pillai, R.S., *et al.* (2007). MicroRNA inhibition of translation initiation in vitro by targeting the cap-binding complex eIF4F. Science *317*, 1764-1767.

McGinnis, W., Garber, R.L., Wirz, J., Kuroiwa, A., and Gehring, W.J. (1984a). A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. Cell *37*, 403-408.

McGinnis, W., Levine, M.S., Hafen, E., Kuroiwa, A., and Gehring, W.J. (1984b). A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. Nature *308*, 428-433.

Miranda, K.C., Huynh, T., Tay, Y., Ang, Y.S., Tam, W.L., Thomson, A.M., Lim, B., and Rigoutsos, I. (2006). A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell *126*, 1203-1217.

Morgan, R. (2006). Hox genes: a continuation of embryonic patterning? Trends Genet *22*, 67-69.

Moriarty, C.H., Pursell, B., and Mercurio, A.M. (2010). miR-10b targets Tiam1: implications for Rac activation and carcinoma migration. J Biol Chem *285*, 20541-20546.

Morlando, M., Ballarino, M., Gromak, N., Pagano, F., Bozzoni, I., and Proudfoot, N.J. (2008). Primary microRNA transcripts are processed co-transcriptionally. Nat Struct Mol Biol *15*, 902-909.

Morris, K.V., Chan, S.W., Jacobsen, S.E., and Looney, D.J. (2004). Small interfering RNA-induced transcriptional gene silencing in human cells. Science *305*, 1289-1292.

Morrison, A., Ariza-McNaughton, L., Gould, A., Featherstone, M., and Krumlauf, R. (1997). HOXD4 and regulation of the group 4 paralog genes. Development *124*, 3135-3146.

Morrison, A., Moroni, M.C., Ariza-McNaughton, L., Krumlauf, R., and Mavilio, F. (1996). In vitro and transgenic analysis of a human HOXD4 retinoid-responsive enhancer. Development *122*, 1895-1907.

Mortensen, R.D., Serra, M., Steitz, J.A., and Vasudevan, S. (2011). Posttranscriptional activation of gene expression in Xenopus laevis oocytes by microRNA-protein complexes (microRNPs). Proc Natl Acad Sci U S A *108*, 8281-8286.

Myers, R.M., Stamatoyannopoulos, J., Snyder, M., Dunham, I., Hardison, R.C., Bernstein, B.E., Gingeras, T.R., Kent, W.J., Birney, E., Wold, B., *et al.* (2011). A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol *9*, e1001046.

Nelson, C.E., Morgan, B.A., Burke, A.C., Laufer, E., DiMambro, E., Murtaugh, L.C., Gonzales, E., Tessarollo, L., Parada, L.F., and Tabin, C. (1996). Analysis of Hox gene expression in the chick limb bud. Development *122*, 1449-1466.

Neville, S.E., Baigent, S.M., Bicknell, A.B., Lowry, P.J., and Gladwell, R.T. (2002). Hox gene expression in adult tissues with particular reference to the adrenal gland. Endocr Res *28*, 669-673.

Nolte, C., Amores, A., Nagy Kovacs, E., Postlethwait, J., and Featherstone, M. (2003). The role of a retinoic acid response element in establishing the anterior neural expression border of Hoxd4 transgenes. Mech Dev *120*, 325-335.

Nolte, C., Rastegar, M., Amores, A., Bouchard, M., Grote, D., Maas, R., Kovacs, E.N., Postlethwait, J., Rambaldi, I., Rowan, S., *et al.* (2006). Stereospecificity and PAX6 function direct Hoxd4 neural enhancer activity along the antero-posterior axis. Dev Biol *299*, 582-593.

Nottrott, S., Simard, M.J., and Richter, J.D. (2006). Human let-7a miRNA blocks protein production on actively translating polyribosomes. Nat Struct Mol Biol *13*, 1108-1114.

Nusslein-Volhard, C., and Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. Nature *287*, 795-801.

Oh, S.K., Scott, M.P., and Sarnow, P. (1992). Homeotic gene Antennapedia mRNA contains 5'-noncoding sequences that confer translational initiation by internal ribosome binding. Genes Dev *6*, 1643-1653.

Okamura, K., Hagen, J.W., Duan, H., Tyler, D.M., and Lai, E.C. (2007). The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. Cell *130*, 89-100.

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., *et al.* (2002). Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature *420*, 563-573.

Orom, U.A., Nielsen, F.C., and Lund, A.H. (2008). MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. Mol Cell *30*, 460-471.

Ovcharenko, D., Stolzel, F., Poitz, D., Fierro, F., Schaich, M., Neubauer, A., Kelnar, K., Davison, T., Muller-Tidow, C., Thiede, C., *et al*. (2011). miR-10a Overexpression is Associated with NPM1 Mutations and Mdm4 Downregulation in Intermediate-Risk Acute Myeloid Leukemia. Exp Hematol.

Pawlicki, J.M., and Steitz, J.A. (2008). Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. J Cell Biol *182*, 61-76.

Pawlicki, J.M., and Steitz, J.A. (2009). Subnuclear compartmentalization of transiently expressed polyadenylated pri-microRNAs Processing at transcription sites or accumulation in SC35 foci. Cell Cycle *8*, 345-356.

Petersen, C.P., Bordeleau, M.E., Pelletier, J., and Sharp, P.A. (2006). Short RNAs repress translation after initiation in mammalian cells. Mol Cell *21*, 533-542.

Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science *309*, 1573-1576.

Ramirez-Solis, R., Zheng, H., Whiting, J., Krumlauf, R., and Bradley, A. (1993). Hoxb-4 (Hox-2.6) mutant mice show homeotic transformation of a cervical vertebra and defects in the closure of the sternal rudiments. Cell *73*, 279-294.

Rastegar, M., Kobrossy, L., Kovacs, E.N., Rambaldi, I., and Featherstone, M. (2004). Sequential histone modifications at Hoxd4 regulatory regions distinguish anterior from posterior embryonic compartments. Mol Cell Biol *24*, 8090-8103.

Reinhart, B.J., and Bartel, D.P. (2002). Small RNAs correspond to centromere heterochromatic repeats. Science *297*, 1831.

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature *403*, 901-906.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., *et al.* (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by Noncoding RNAs. Cell *129*, 1311-1323.

Rinn, J.L., Wang, J.K., Allen, N., Brugmann, S.A., Mikels, A.J., Liu, H., Ridky, T.W., Stadler, H.S., Nusse, R., Helms, J.A., *et al.* (2008). A dermal HOX transcriptional program regulates site-specific epidermal fate. Genes Dev *22*, 303-307.

Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., and Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. Genome Res *14*, 1902-1910.

Ruby, J.G., Jan, C.H., and Bartel, D.P. (2007a). Intronic microRNA precursors that bypass Drosha processing. Nature *448*, 83-86.

Ruby, J.G., Stark, A., Johnston, W.K., Kellis, M., Bartel, D.P., and Lai, E.C. (2007b). Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs. Genome Res *17*, 1850-1864.

Rudnicki, M.A., and McBurney, M.W. (1987). Cell culture methods and induction of differentiation of embryonal carcinoma cell lines. In Teratocarcinomas and embryonic stem cells: a practical approach, E.J. Robertson, ed. (Oxford., IRL Press), pp. 19-49.

Sasayama, T., Nishihara, M., Kondoh, T., Hosoda, K., and Kohmura, E. (2009). MicroRNA-10b is overexpressed in malignant glioma and associated with tumor invasive factors, uPAR and RhoC. Int J Cancer *125*, 1407-1413.

Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D.

(2003). Asymmetry in the Assembly of the RNAi Enzyme Complex. Cell *115*, 199-208.

Scott, M.P., and Weiner, A.J. (1984). Structural Relationships among Genes That Control Development - Sequence Homology between the Antennapedia, Ultrabithorax, and Fushi Tarazu Loci of Drosophila. P Natl Acad Sci-Biol *81*, 4115-4119.

Seggerson, K., Tang, L., and Moss, E.G. (2002). Two genetic circuits repress the Caenorhabditis elegans heterochronic gene lin-28 after translation initiation. Dev Biol *243*, 215-225.

Sempere, L.F., Sokol, N.S., Dubrovsky, E.B., Berger, E.M., and Ambros, V. (2003). Temporal regulation of microRNA expression in Drosophila melanogaster mediated by hormonal signals and broad-Complex gene activity. Dev Biol *259*, 9-18.

Shah, N., and Sukumar, S. (2010). The Hox genes and their roles in oncogenesis. Nat Rev Cancer *10*, 361-371.

Shen, X., Fang, J., Lv, X., Pei, Z., Wang, Y., Jiang, S., and Ding, K. (2011). Heparin Impairs Angiogenesis through Inhibition of MicroRNA-10b. J Biol Chem *286*, 26616-26627.

Simeone, A., Pannese, M., Acampora, D., D'Esposito, M., and Boncinelli, E. (1988). At least three human homeoboxes on chromosome 12 belong to the same transcription unit. Nucleic Acids Res *16*, 5379-5390.

Small, K.M., and Potter, S.S. (1993). Homeotic transformations and limb defects in Hox A11 mutant mice. Genes Dev *7*, 2318-2328.

Stark, A., Kheradpour, P., Parts, L., Brennecke, J., Hodges, E., Hannon, G.J., and Kellis, M. (2007). Systematic discovery and characterization of fly microRNAs using 12 Drosophila genomes. Genome Res *17*, 1865-1879.

Steitz, J.A., and Vasudevan, S. (2009). miRNPs: versatile regulators of gene expression in vertebrate cells. Biochem Soc T *37*, 931-935.

Takahashi, Y., Hamada, J., Murakawa, K., Takada, M., Tada, M., Nogami, I., Hayashi, N., Nakamori, S., Monden, M., Miyamoto, M., *et al.* (2004). Expression profiles of 39 HOX genes in normal human adult organs and anaplastic thyroid cancer cell lines by quantitative real-time RT-PCR system. Exp Cell Res *293*, 144-153.

Tan, Y., Zhang, B., Wu, T., Skogerbo, G., Zhu, X., Guo, X., He, S., and Chen, R. (2009). Transcriptional inhibiton of Hoxd4 expression by miRNA-10a in human breast cancer cells. BMC Mol Biol *10*, 12.

Taylor, H.S., Vanden Heuvel, G.B., and Igarashi, P. (1997). A conserved Hox axis in the mouse and human female reproductive system: late establishment and persistent adult expression of the Hoxa cluster genes. Biol Reprod *57*, 1338-1345.

Tian, Y., Luo, A., Cai, Y., Su, Q., Ding, F., Chen, H., and Liu, Z. (2010). MicroRNA-10b promotes migration and invasion through KLF4 in human esophageal cancer cell lines. J Biol Chem *285*, 7986-7994.

Tribioli, C., and Lufkin, T. (1999). The murine Bapx1 homeobox gene plays a critical role in embryonic development of the axial skeleton and spleen. Development *126*, 5699-5711.

Tschopp, P., Tarchini, B., Spitz, F., Zakany, J., and Duboule, D. (2009). Uncoupling time and space in the collinear regulation of Hox genes. PLoS Genet *5*, e1000398.

van Dijk, E., Cougot, N., Meyer, S., Babajko, S., Wahle, E., and Seraphin, B. (2002). Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. EMBO J *21*, 6915-6924.

van Scherpenzeel Thim, V., Remacle, S., Picard, J., Cornu, G., Gofflot, F., Rezsohazy, R., and Verellen-Dumoulin, C. (2005). Mutation analysis of the HOX paralogous 4-13 genes in children with acute lymphoid malignancies: identification of a novel germline mutation of HOXD4 leading to a partial loss-of-function. Hum Mutat *25*, 384-395.

Vasudevan, S., Tong, Y., and Steitz, J.A. (2007). Switching from repression to

activation: microRNAs can up-regulate translation. Science *318*, 1931-1934.

Veerla, S., Lindgren, D., Kvist, A., Frigyesi, A., Staaf, J., Persson, H., Liedberg, F., Chebil, G., Gudjonsson, S., Borg, A., *et al.* (2009). MiRNA expression in urothelial carcinomas: important roles of miR-10a, miR-222, miR-125b, miR-7 and miR-452 for tumor stage and metastasis, and frequent homozygous losses of miR-31. Int J Cancer *124*, 2236-2242.

Wadler, C.S., and Vanderpool, C.K. (2007). A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. Proc Natl Acad Sci U S A *104*, 20454-20459.

Wakiyama, M., Takimoto, K., Ohara, O., and Yokoyama, S. (2007). Let-7 microRNA-mediated mRNA deadenylation and translational repression in a mammalian cell-free system. Genes Dev *21*, 1857-1862.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A., *et al.* (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120-124.

Wang, W., Chan, E.K., Baron, S., Van de Water, T., and Lufkin, T. (2001). Hmx2 homeobox gene control of murine vestibular morphogenesis. Development *128*, 5017-5029.

Wei, Y., Harris, T., and Childs, G. (2002). Global gene expression patterns during neural differentiation of P19 embryonic carcinoma cells. Differentiation *70*, 204-219.

Weiss, F.U., Marques, I.J., Woltering, J.M., Vlecken, D.H., Aghdassi, A., Partecke, L.I., Heidecke, C.D., Lerch, M.M., and Bagowski, C.P. (2009). Retinoic acid receptor antagonists inhibit miR-10a expression and block metastatic behavior of pancreatic cancer. Gastroenterology *137*, 2136-2145 e2131-2137.

Wienholds, E., Kloosterman, W.P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., Horvitz, H.R., Kauppinen, S., and Plasterk, R.H.

(2005). MicroRNA expression in zebrafish embryonic development. Science *309*, 310-311.

Wilkinson, D.G., Bhatt, S., Cook, M., Boncinelli, E., and Krumlauf, R. (1989). Segmental expression of Hox-2 homoeobox-containing genes in the developing mouse hindbrain. Nature *341*, 405-409.

Woltering, J.M., and Durston, A.J. (2006). The zebrafish hoxDb cluster has been reduced to a single microRNA. Nat Genet *38*, 601-602.

Woltering, J.M., and Durston, A.J. (2008). MiR-10 represses HoxB1a and HoxB3a in zebrafish. PLoS One *3*, e1396.

Wutz, A., Rasmussen, T.P., and Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nat Genet *30*, 167-174.

Yahagi, N., Kosaki, R., Ito, T., Mitsuhashi, T., Shimada, H., Tomita, M., Takahashi, T., and Kosaki, K. (2004). Position-specific expression of Hox genes along the gastrointestinal tract. Congenit Anom (Kyoto) *44*, 18-26.

Yamamoto, M., Takai, D., and Yamamoto, F. (2003). Comprehensive expression profiling of highly homologous 39 hox genes in 26 different human adult tissues by the modified systematic multiplex RT-pCR method reveals tissue-specific expression pattern that suggests an important role of chromosomal structure in the regulation of hox gene expression in adult tissues. Gene Expr *11*, 199-210.

Yang, W., Chendrimada, T.P., Wang, Q., Higuchi, M., Seeburg, P.H., Shiekhattar, R., and Nishikura, K. (2006). Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. Nat Struct Mol Biol *13*, 13-21.

Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev *17*, 3011-3016.

Zakany, J., and Duboule, D. (1996). Synpolydactyly in mice with a targeted deficiency in the HoxD complex. Nature *384*, 69-71.

Zeng, Y., Yi, R., and Cullen, B.R. (2003). MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. Proc Natl Acad Sci U S A *100*, 9779-9784.

Zhai, Y., Kuick, R., Nan, B., Ota, I., Weiss, S.J., Trimble, C.L., Fearon, E.R., and Cho, K.R. (2007). Gene expression analysis of preinvasive and invasive cervical squamous cell carcinomas identifies HOXC10 as a key mediator of invasion. Cancer Res *67*, 10163-10172.

Zhang, F., Nagy Kovacs, E., and Featherstone, M.S. (2000). Murine hoxd4 expression in the CNS requires multiple elements including a retinoic acid response element. Mech Dev *96*, 79-89.

Zhang, F., Popperl, H., Morrison, A., Kovacs, E.N., Prideaux, V., Schwarz, L., Krumlauf, R., Rossant, J., and Featherstone, M.S. (1997). Elements both 5' and 3' to the murine Hoxd4 gene establish anterior borders of expression in mesoderm and neurectoderm. Mech Dev *67*, 49-58.

## RELATED PUBLICATION

1. **Phua SLC**, Sivakamasundari V, Shao Y, Cai X, Zhang L-F, et al. (2011) **Nuclear Accumulation of an Uncapped RNA Produced by Drosha Cleavage of a Transcript Encoding *miR-10b* and *HOXD4*.** PLoS ONE 6(10): e25689. doi:10.1371/journal.pone.0025689