Function and Regulation of Distal-less-Related Homeobox Genes During Visceral Arch Development In Zebrafish

by

Debra L. Ellies

A thesis submitted to the School of Graduate Studies and Research in partial fulfillment for the degree

-of-

Master of Science

Department of Anatomy and Neurobiology

University of Ottawa

Ottawa. Ontario

July 24, 1995

ISBN 0-612-11554-2

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

ABSTRACT

Vertebrate *distal-less* homeobox genes are known to play a role in embryonic pattern formation. Six zebrafish *distal-less* homeobox genes have been isolated; *dlx1, dlx2, dlx3, dlx4, dlx6,* and *dlx7*. Their combinatorial embryonic expression patterns seen in the forebrain, visceral arches, and fins are suggestive of a new "*homeobox code*". In zebrafish, visceral arch *distal-less* expression patterns correlate with future neural crest derived craniofacial cartilage derivatives, insinuating that they may be involved in visceral arch development and differentiation. Retinoic acid, a potent teratogen, is known to cause abnormal craniofacial cartilage development, and abnormal *distal-less* expression in the visceral arches. Various doses of *all-trans* retinoic acid were used to disrupt *distal-less* expression during different visceral arch developmental time points. The abnormal visceral arch *distal-less* expression was further correlated with craniofacial cartilage development.

Zebrafish embryos at various stages, treated with 10^{-6}M RA, displayed a loss of all *distal-less* expression in the mandibular and hyoid arches. The embryos exhibiting a loss of *distal-less* expression, also developed with a loss of mandibular and hyoid arch derived craniofacial cartilage components.

Embryos treated with 10^{-7}M RA displayed a stage dependent loss of *distal-less* expression, mainly in the branchial arches. The visceral arch craniofacial cartilage differentiated abnormally in embryos displaying a loss of *distal-less* expression. When *distal-less* expression was lost in the branchial arches, then only the branchial arch cartilage was lost. However, abnormal *distal-less* expression in any of the visceral arches could also be correlated with abnormal visceral arch cartilage development.

These experiments suggest that *dlx* genes are part of the multi-step process leading to visceral arch cartilage development. Zebrafish genes involved in this process are sensitive to retinoic acid (10^{-7}M) until 24hpf, where RA does not seem to affect the process of craniofacial cartilage development. Accordingly, a loss in visceral arch *dlx* expression during any stage (up to 24hpf) of this process leads to abnormal craniofacial cartilage development.

ACKNOWLEDGMENTS

I would like to thank my supervisor, Dr. Marc Ekker, who guided and often challenged my scientific ability. I would also like to thank him for teaching me how to think objectively, work independently, and more importantly, to consider all possible experimental outcomes. He also taught me another very important ability, to think before I talk, and to be able to back up what I say with scientific proof.

I would also like to thank Dr. Marie-Andre Akimenko for her support throughout my degree, for her patience when she was teaching me new techniques, and also for her guidance on new experimental ideas and approaches.

I would also like to take the time to thank the most important members of the laboratory; Lucille, Lynda, and Wei. Lucille, you were the most patient with me during these two years, thank you for your tremendous help and guidance (you will have to find another sequencing partner). Lynda, thank you for teaching me how to work with zebrafish embryos, and for helping me with those *!@# calculations in the morning (don't forget to use protection). Wei, I would like to thank you for teaching me how to perform whole mount *in situ* hybridizations and cross sections on zebrafish embryos. I would also like to thank Wei for letting me use her *in situ* probes. This thesis would have been extremely difficult without your help, thank you for making it an enjoyable experience.

One person who deserves a lot of thanks is Genny Giroux, who mapped the *dlx1* genomic loci, isolated the *dlx1* cDNA clone, and who did a lot of sequencing for the project.

Another person I must thank for their guidance, is Dr. Robert Langille. His suggestions, support, and knowledge on craniofacial neural crest cartilage formation was greatly appreciated.

Last but not least, I would like to thank Dr. Y. Lefevbre and Dr. R. Hache for their support and guidance during the last two years.

In all, I would like to thank everyone who had the patience to endure; my stubbornness, my stupidities, and long hours in the lab.

INDIVIDUAL CONTRIBUTIONS TO THE WORK PRESENTED IN THIS THESIS

The work described in this thesis is the result of a collaborative effort carried out by the Author and members of the Ekker and Akimenko laboratories at the Loeb research Institute, University of Ottawa, unless otherwise stated. The contributions of the various individuals who took part in this work are acknowledged below.

Chapter 2; The genomic clones for dlx2 and dlx4 were isolated by the Author. They were mapped and subcloned by G. Giroux who also identified the dlx1 gene. The dlx1 cDNA was isolated by W. Lin, who also identified and sequenced the dlx6 gene. The dlx7 gene was isolated and sequenced in part in the laboratory of Dr. R. Brietbart, Harvard U.

Sequence for the dlx1 gene was carried out by the Author and G. Giroux. The Author completed the sequence of dlx7 and performed the sequence and phylogenetic analyses of the dlx gene family.

Expression studies for the new dlx genes were carried out entirely by the Author.

Chapter 3; The genomic clones for dlx3, msxA, msxB, and msxD were isolated and mapped by Marie-Andree Akimenko and Marc Ekker in the laboratory of Monte Westerfield, University of Oregon. Some of the sequencing was also carried out by M.A.-A. and M.E. in Dr. Westerfield's laboratory.

The genomic clone for msxC was isolated by W. Lin. The genomic clone for msxE was isolated by the Author. The msxC genomic clone was mapped and sequenced by the Author.

The ExoIII deletion mutants were performed by the Author and M.A. Akimenko. Most of the sequencing of the putative regulatory regions was carried out by the Author, except for dlx1 and dlx4 which were done by G. Giroux. Sequence analysis of the putative regulatory regions was carried out entirely by the Author.

Chapter 4; The Author performed almost all the experiments described in this chapter and analyzed their results. W. Lin carried out some of the embryo treatments and in situ hybridizations. Lynda Laforest helped with the double labeling. The help of Dr. R. Langille in the analysis of the cartilage phenotypes is acknowledged.

TABLE OF CONTENTS

LIST OF FIGURES

List of Tables

LIST OF ABBREVIATIONS

72 hpf	72 Hours Post Fertilization (Zebrafish)
9 dpc	9 Days Postcoitum (Mouse)
Abd-A	Abdominal A Gene (*Bx-C* Gene)
Abd-B	Abdominal B Gene (*Bx-C* Gene)
AER	Apical Ectodermal Ridge
Ant-C	Antennapedia Homeotic Complex (*Drosophila*)
Ant	Antennae Gene (*Ant-C* Gene)
bp	Base Pair
Bx-C	Bithorax Homeotic Complex (*Drosophila*)
cDNA	Complementary DNA
CRABP	Cytosolic Retinoic Acid Binding Protein
D1-D4	Diencephalic Segment 1 To 4 (Forebrain Segments)
Dfd	Deformed Gene (*Ant-C* Gene)
Dll	Distal-Less Gene (*Drosophila*)
Dlx	Distal-Less Gene (Vertebrate)
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
Dwnt-5	*Drosophila Wnt* Gene Homolog
Emx	Empty Spiracles Gene (*Drosophila*)
Eng	Engrailed Gene
FGF	Fibroblast Growth Factor
GCG	Genetics Computer Group Search, Analysis, and Display Software
Gsc	Goosecoid Gene
Hom-C	Homeotic Cluster (*Drosophila*)

Hox	Clustered Homeobox Genes (Vertebrate)
Kb	Kilo Base Pair
Kr	Kreisler (Mouse Hindbrain Mutant)
Krx20	Krox 20 Vertebrate Homeotic Gene
Lab	Labial Gene (*Ant-C* Gene)
M	Molarity
mRNA	Messenger Ribonucleic Acid
Msx	Muscle Segmented Homeobox Gene (vertebrate)
N-CAM	Neural Cell Adhesion Molecule
NRC	National Research Council of Canada
Otx	Orthodenticle Gene (*Drosophila*)
PCR	Polymerase Chain Reaction
R1-R9	Rhombomere 1 To Rhombomere 9
RA	Retinoic Acid
RAR	Retinoic Acid Receptor
RARE	Retinoic Acid Response Element
RXR	Retinoic X Receptor
RXRE	Retinoic X Response Element
Scr	Sex Comb Reduced Gene (*Ant-C* Gene)
TGF	Transforming Growth Factor
THR	Thyroid Hormone Receptor
Ubx	Ultrabithorax (*Hom-C* Gene)
VDR	Vitamin D_3 Receptor
Wg	Wingless Segment Polarity Gene

1. INTRODUCTION

1.1. Developmental Model

The genetics of invertebrate embryonic pattern formation is traditionally studied using the *Drosophila* developmental model; and the mouse developmental model is used to study vertebrate embryonic pattern formation. However, the uneasiness of the mouse model prompted other vertebrate developmental models to be sought out. We use the zebrafish, *Brachydanio rerio*, as a vertebrate developmental model to study the genetics of pattern formation. Why zebrafish? The development of the zebrafish enables the use of both experimental embryology and molecular genetic methodologies.

More specifically zebrafish have distinct features which make it a preferred vertebrate developmental model (figure 1-1).[1] They are small (3-4cm long) tropical freshwater fish, easy to maintain in large numbers. They reach sexual maturity within 3 months. Females lay anywhere from 300 to 500 eggs per spawning therefore facilitating genetic analyses. Eggs are fertilized externally and the embryos are lucent, facilitating the observation of individual cells during development. The translucent embryo allows one to view the cell movements during gastrulation, the formation of the brain, and the beating heart. Large-scale genetic screens are presently underway to saturate the zebrafish genome for mutations that affect pattern formation during development. [2] Furthermore, gynogenetic zebrafish embryos can be created with ease. [2]

1.2. Developmental Genetics

The development of the *Drosophila* embryo is controlled by a network of regulatory genes. The first group of genes involved in this regulatory network are the coordinate genes which are expressed maternally during oogenesis. They create a gradient pattern of expression along the anterior-posterior/dorso-ventral axis of the developing embryo. The next group of genes are the segmentation genes which convert the gradient patterns into repetitive

ZEBRAFISH, *Brachydanio rerio*

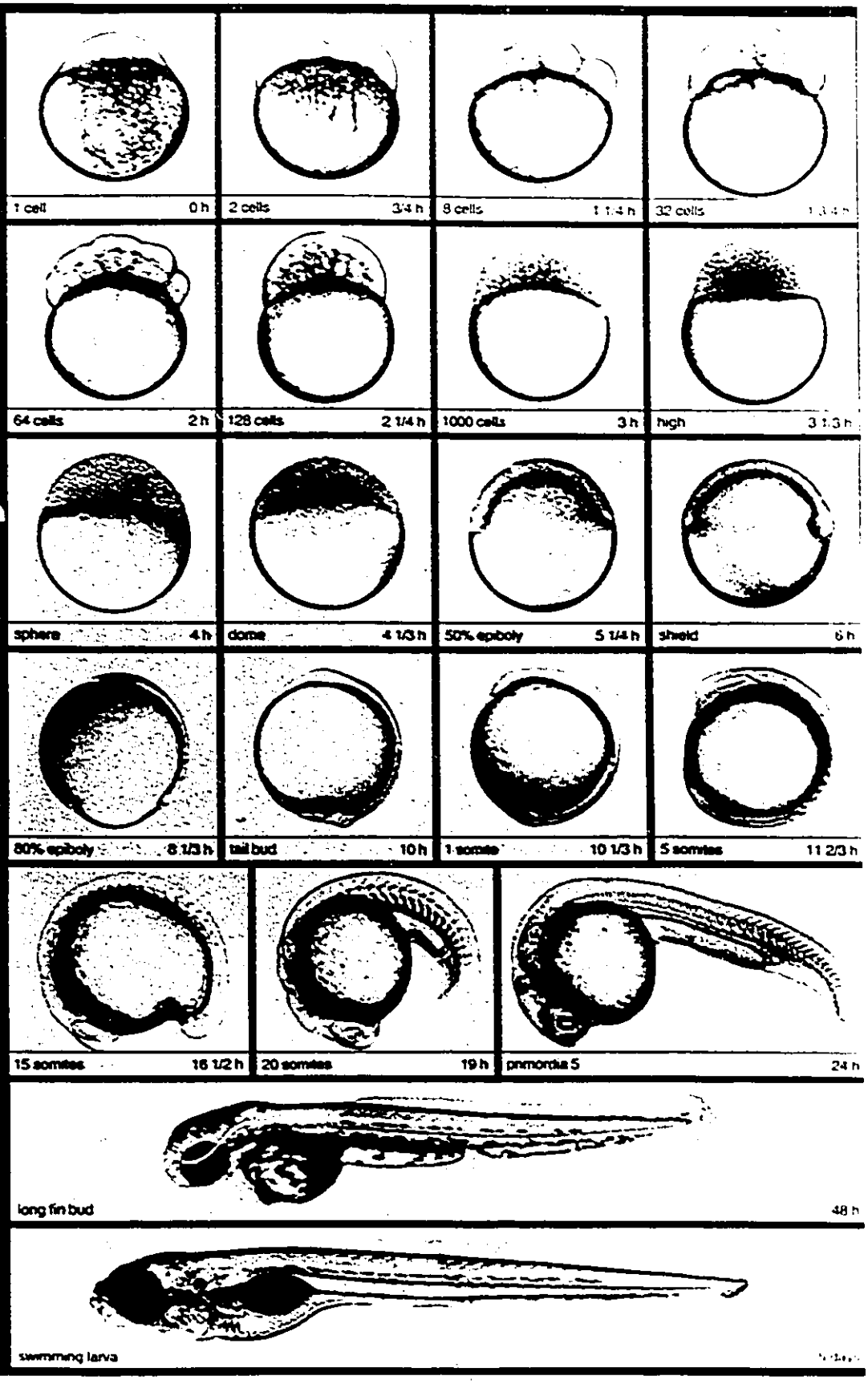


Figure 1-1: Zebrafish development.

The panels on the left show the developmental stages of a zebrafish embryo from the 1 cell stage (0hpf) to a swimming larva (5dpf).

patterns (segments). The segmentation genes are divided into three classes[3]; i) Gap Genes, are expressed in domains that are several segments wide, ii) Pair-Rule Genes, are expressed in alternate segments, and iii) Segment Polarity Genes, which are expressed in parts of every segment. Homeotic genes are the third group of genes in this network. They are needed to convert the repetitive pattern found in similar segments to a sequential pattern providing each segment with its unique identity.

The homeotic genes were first discovered by studying spontaneous mutations found in *Drosophila*, termed homeotic phenotypes. Homeosis refers to the substitution of one body part for another. The first homeotic mutation, discovered by Bridges in 1915, resulted in a partial or complete segmental transformation. [4] He suggested that the homeotic genes were involved in the genetic control of the developing body plan.

1.3. Homeodomain

The discovery of other homeotic genes revealed that they all contain a homologous sequence of 60 amino acids, the homeodomain (homeobox refers to the DNA sequence encoding the homeodomain). The homeodomain is a helix-turn-helix motif known to bind specific DNA sequences[5], thus categorizing homeotic genes as transcriptional regulators. The homeodomain is found to be highly conserved throughout evolution from insects to man. [6]

Using amino acid sequence similarities within the homeodomain and in flanking protein sequence, it is possible to divide homeodomain protein and the genes that encode them into various classes.

1.4. Hom-C/ Hox Genes

In *Drosophila*, homeotic genes conferring segment identity are organized on two genomic homeotic clusters *(Hom-C)*. [4] The two drosophila genomic clusters are found on chromosome 3, and are called the Bithorax Complex (*Bx-C*) and the Antennapedia Complex (*Ant-C*).

Vertebrate homeobox genes (*Hox*) were identified based on sequence similarities to the *Drosophila* homeotic genes, such as the *Ant-C* and *Bx-C* (Figure 1-2). Vertebrate *Hox* genes are organized on four genomic clusters (A,

B, C, D) on chromosomes 6, 11, 15 and 2 in the mouse and on chromosomes 7, 17, 12 and 2 in human. [7] The order and position of the *Hox* genes within a genomic cluster found in mice and in humans is very similar, indicating a high level of conservation during evolution. The vertebrate *Hox-A* cluster found in the mouse contains 11 *Hox* genes, and is the longest *Hox* cluster characterized to date. The genes are spaced by a distance of 5-10kb and are transcribed in the same direction.

The homeobox genes within each of the four vertebrate genomic clusters can be aligned into 13 groups of genes, called paralogous groups, that share a distinctive similarity in the homeobox region. These paralogous groups can be related to specific homeotic genes from the *Drosophila Hom-C* cluster (Fig. 1-2). [8]

1.4.1 Evolution of Clusters

The organization of the vertebrate *Hox* cluster suggests that there was 2 chromosomal region duplication events in the ancestor of mammals. [9] More and more data suggests that the organization of homeobox clusters in the *Xenopus* and zebrafish is similar to the cluster arrangements found in mammals, therefore, establishing the time point for the alleged duplication events early during the evolution of vertebrates. The ancestral linkage group most likely arose from an expansion process involving local gene duplications in the middle and at the extremes of the primordial cluster. [9]

Comparisons between *Drosophila* and vertebrate homeobox genes suggest that the *Drosophila* cluster diverged before the vertebrates from the distant ancestor. For example, the vertebrate and *Drosophila* homeobox genes are all transcribed in the same direction, except for the *Drosophila Dfd gene*. The *Drosophila Hom-C* probably arose from an independent duplication and inversion event. [9]

1.4.2 Hom-C /Hox Code

Current models relating the expression of homeotic genes to the control of segment development are based on the pioneering work of Lewis[7] . The *Hom-C* code, suggested by Lewis, proposes that the *Bx-C* contains a series of

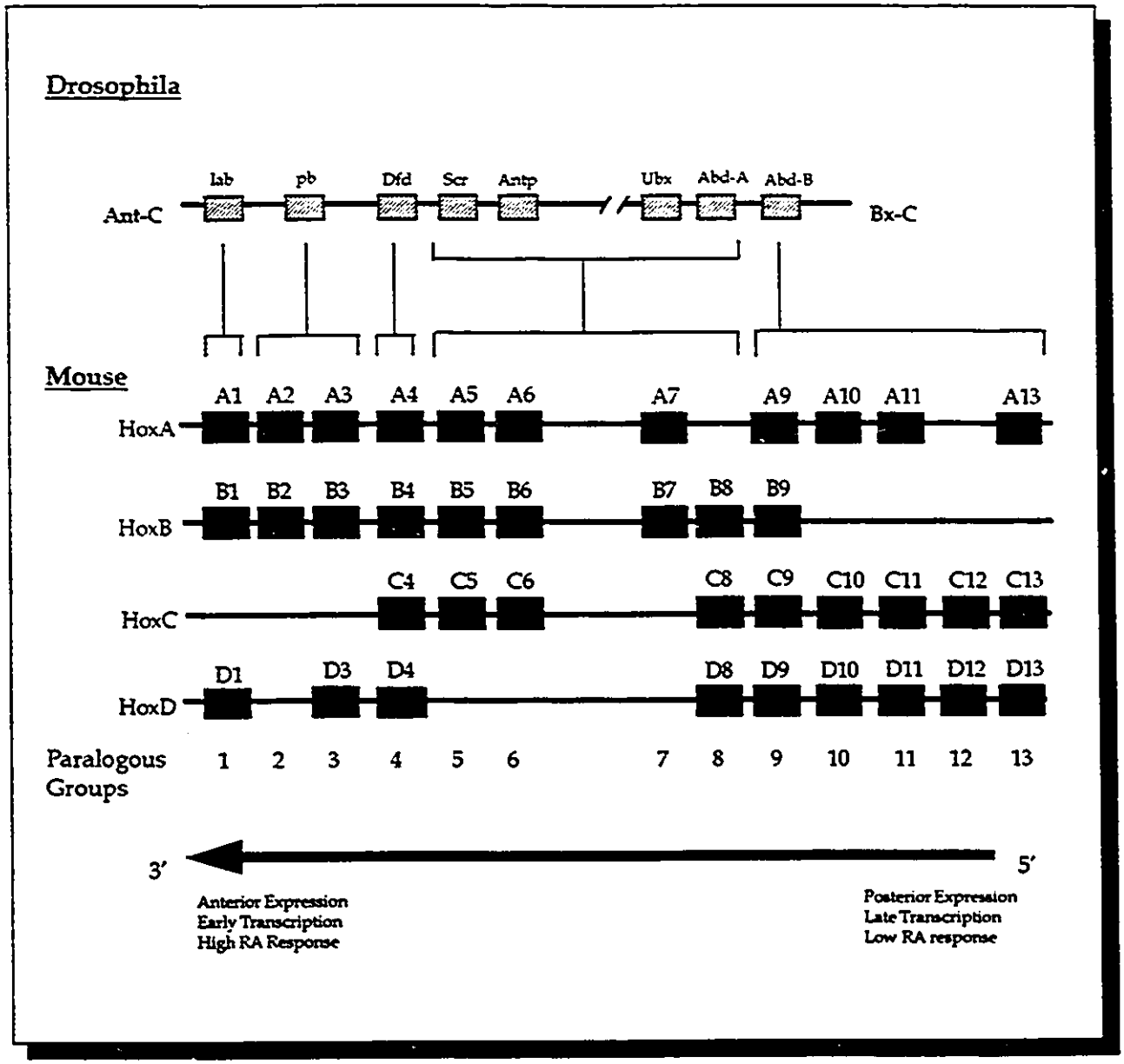


Figure 1- 2. Alignment of the four vertebrate *Hox* clusters with the *Drosophila Hom-C* homeotic complex.

Figure 1-2: *Hom*-C/ *Hox* Clusters.

Schematic showing the relationship between the *Drosophila Hom*-C genes and the four mouse *Hox* clusters. The four mouse clusters are named *HoxA* to *HoxD*. The individual homeotic genes within each mouse cluster can be aligned vertically with the other mouse and *Drosophila* clusters. The *Drosophila Hom*-C can be divided into 2 families; the bithorax complex (*Bx-C*) and the antennapedia complex (*Ant-C*). There are 13 vertical vertebrate groups called paralogous groups. For example, the *Drosophila Dfd* gene has paralogues in all four mouse clusters, *Hox-A4*, *Hox-B4*, *Hox-C4*, and *Hox-D4*. The genes paralogous to the *Drosophila Dfd* are in one paralogous group, group 4. The large arrow beneath the clusters represents the linear relationship of the homeotic genes with their respective expression time & domains and retinoic acid sensitivity. The arrow also indicates the direction of transcription for the vertebrate *Hox* genes. Adapted from R. Krumlauf (TIGS, 1993; 9(4) 106-111).[8]

genes, each of which would be turned on at a different position along the antero-posterior axis of the developing embryo, in partially overlapping domains. Each segment would contain a combination of partially overlapping expression domains of different *Hom-C* genes. This combination would then define segment identity (figure 1-3). [10]

The *Hox* code's functional organization has been conserved from insects to mammals. The *Drosophila* 3' *Hom-C* genes are expressed in the anterior part of the embryo. The same is true for the vertebrate 3' *Hox* genes, however there are a maximum of four vertebrate *Hox* genes for one *Drosophila Hom-C* gene, and the vertebrate *Hox* code's anterior expression domain is limited to the level of the hindbrain.

1.4.2.1 Opposing The Hom-C Code

Lewis' model must be revised because the molecular data present today suggests that this model is incorrect due to the following three points. First, there are not enough genes, and probably not enough gene products, to give a unique combination of active homeotic genes to each segment. Three genes (*Abd-A, Abd-B,* and *Ubx)* make up the *Drosophila Bx-C* which specifies nine segments. Together these genes are supposed to control, according to Lewis, the different identity of at least nine parasegments.

Second, the cells of a segment do not express the same combination of homeotic genes throughout development. Segments do initially appear as domains of uniform homeotic gene expression but, later in development, a pattern of different cell types is generated within each segment. The segment therefore becomes a mosaic of different cell types expressing different homeotic genes.

Finally, certain cells switch from expressing one homeotic gene during very early development to expressing another or none at all, during later stages of development, thus confronting the strict lineage model of homeotic gene expression.

Another model was then proposed by Welcome Bender and colleagues named, *"open for business"*. [11] This model is based on the observation that there are only three gene products (*abd-A, abd-B,* and *Ubx*) from the *Bx*-C, and nine segments in *Drosophila.* These three gene products

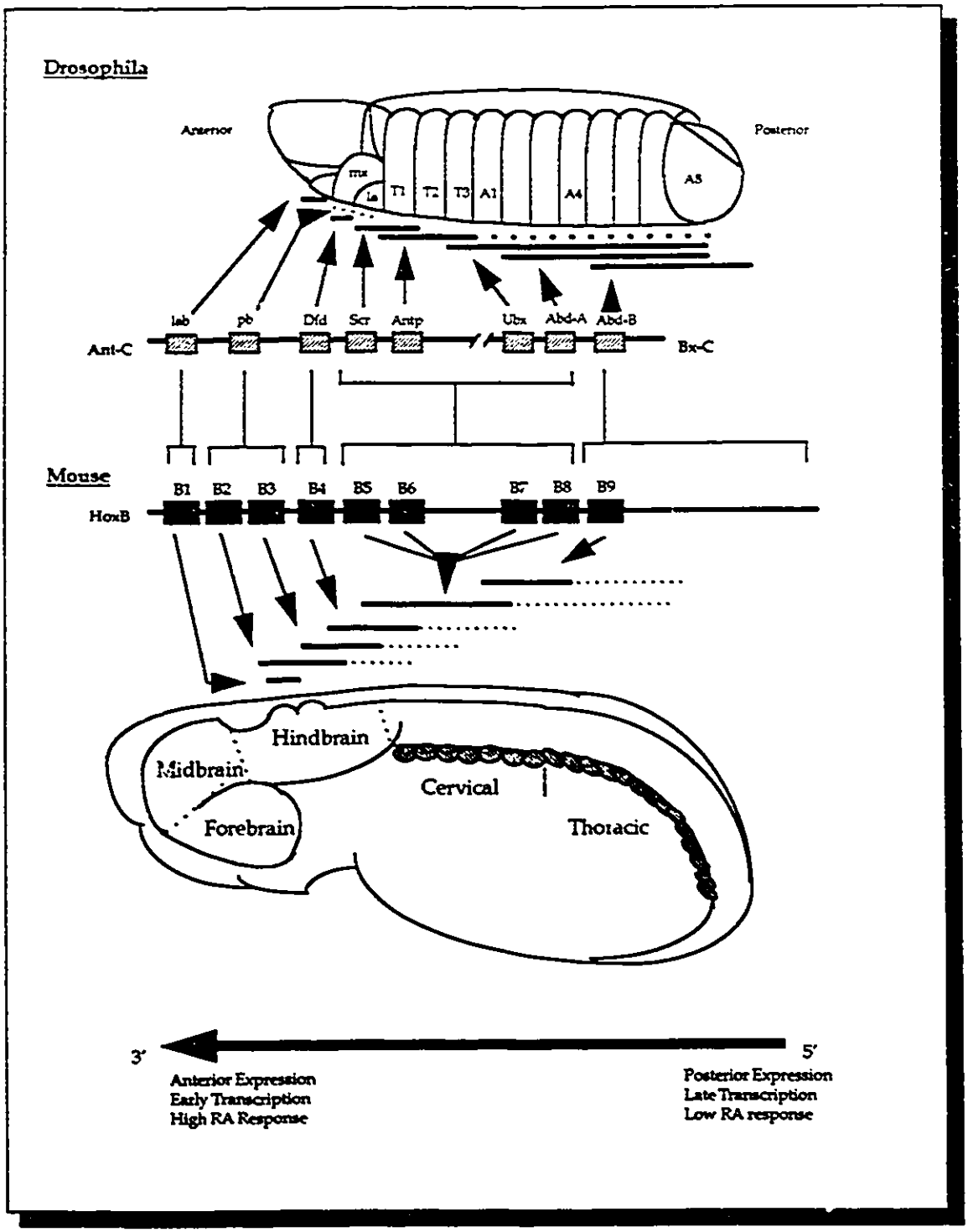


Figure 1-3. Summary of *Drosophila Hom-C* and mouse *Hox* B expression patterns.

Figure 1-3: *Hox* Code.

The upper panel shows a 10h *Drosophila* embryo with its *Hom-C* expression patterns. The lower panel illustrates a 12 dpc mouse embryo and its *Hox-b* expression domains which are paralogous to the *Drosophila Hom-C* genes. The large arrow beneath the illustrations shows the direction of transcription 5' to 3', but more importantly shows the collinear relationship between the *Hox* gene expression and their location along the cluster. In other word the 3' cluster genes have a more anterior expression domain than the 5' cluster genes. Adapted from McGinnis et al. (1992; Cell 68; 283-302). [12]

must then be controlled by multiple, independent regulatory elements. For example, *Ubx* expression would be controlled by multiple regulatory elements, thereby facilitating differential transcription depending on the segment. The differential transcription would make it more possible for three gene products to specify nine segments. The essence of the model is that each segment decides, early in development, which regulatory domains are *'open for business'* according to the segment's position in the embryo.

1.4.3 Homeotic Mutations

Lewis[10] pointed out that there are two types of homeotic mutations with opposite effects; loss-of-function mutants resulting from a deletion or inactivation of a gene, and a dominant gain-of-function mutant where the gene product is unaltered but expressed ectopically.

1.4.3.1 Loss-of-Function Mutants

Most loss-of-function *Drosophila Hom-C* mutants result in homeotic transformations, ie. for the *Scr, Ant, Ubx, Abd-A, Abd-B* homeotic genes. [13] For others (*Lab, Dfd*), however, loss-of-function mutations are accompanied by structural deficiencies, and there are no detectable segmental transformations. [14]

Homeotic transformations due to a loss-of-function are only seen with those genes which have an overlapping expression domain with other *Hom*-C genes. For both *Lab* and *Dfd* genes there are no other *Hom*-C genes with overlapping expression patterns, consequently is no 'backup' gene to compensate for their functional annihilation.

The same applies for vertebrate loss of *Hox* gene function[15], however, it is more complex due to cluster duplications. The duplications have created paralogous genes with nearly identical domains of expression (discussed in more detail in section 1.5.2.5). However, it is interesting that cognate *Hox* genes on different clusters do not show identical expression patterns. For example, *Hox-c8* and *Hox-b8* belong to the same paralogous group, but their anterior expression limit varies. [16]

1.4.3.2 Gain-of-Function Mutants

Ectopic expression of a *Hom*-C gene is capable of transforming other segmental identities towards an identity normally specified by the ectopically expressed gene. In most cases, regions anterior to the normal domain of expression are homeotically transformed. The idea that ectopic expression always causes a posterior transformation is a 'rule' that has been violated in many circumstances, for example, ectopic expression of either *Scr* or *Dfd* proteins causes an anterior transformation. [17]

The best case reporting a homeotic transformation of axial structures in vertebrates comes from ectopic expression of *Hox* genes by transgenic means or by stimulation with exogenous factors, such as retinoic acid. Ectopic expression of the *Hox-a7* gene causes craniofacial abnormalities and alterations to the axial skeleton in transgenic mice. [18] Homeotic transformations of vertebrae have also resulted from exposure to teratological doses of retinoic acid (RA). [19]

1.4.4 Retinoic Acid Regulation of Hox Genes

Retinoic acid was long thought to play a role in the regulation of vertebrate *Hox* genes. The first evidence came from studies of the human teratocarcinoma cell line NT2/D1, where all four vertebrate *Hox* clusters were shown to be RA inducible and that the location of a *Hox* gene within the cluster determined its responsiveness to RA (Figure 1-2). [20] Genes near the 3' end of the cluster are induced extremely rapidly by low concentrations of RA, whereas genes located closer to the 5' end of the cluster required higher doses and longer treatment times of RA in order to achieve equivalent gene induction. Thus, RA *Hox* gene inducibility is collinear with the organization of the genes on the clusters.

The limb bud is a unique developmental system because it has both an anterior-posterior and a proximal-distal developmental axis. The *Hox* genes would be expected to play an integral role in the development of the anterior-posterior limb axis because they are known to be involved in the development of the embryonic anteroposterior body axis. The control of the anterior-posterior axis polarity comes from the limb bud's posterior region called the *zone of polarizing activity* (ZPA). When this region is grafted to the anterior edge of the

limb bud, mirror image digit duplications develop. When RA beads are implanted into the same anterior bud's location, digit duplications occur, [21] suggesting that RA is part of the ZPA's pathway.

RA was shown to induce ectopic expression of certain Hox genes only after 18h. [22] This would imply that either: 1) RA acts indirectly on the limb bud's *Hox* genes because pharmacokinetic studies have shown administered RA to drop to endogenous levels within 8h, [23] and/or 2) RA may act directly with very slow kinetics.

Increasing evidence links RA with direct regulation of the *Hox* genes during embryogenesis in many vertebrates. Retinoic acid response elements (RAREs) have been identified downstream of the murine *Hox-a1* and human *Hox-A1* genes and the murine and chick *Hox-b1* genes. [24] This data is highly suggestive of a retinoic acid inducible *'locus control center'* which would be located in the 3' end of each cluster. A conserved RARE in the promoter regions for the murine and human *Hox-D4* genes has also been reported. [25]

In 1993, NT2/D1 cell culture studies reported a dosage dependent RA induction of *Hox-D4* transcription which originated from two different, synchronously activated major transcriptional start sites. [26] Expression of these transcripts was shown to be developmentally regulated according to tissue- and region- specific patterns in human embryonic development. The higher 'RA sensitivity' proximal start site required RA doses of at least 0.5uM whereas the lower sensitivity required 10nM RA.

1.5. Head Homeobox Genes

The most anteriorly-expressed gene of the vertebrate *Hox* cluster has its anterior expression limit at the border between rhombomeres 2 and 3 of the hindbrain. [27-29] One can then conclude that the clustered *Hox* genes are not involved in the patterning of the anterior brain. However, homeobox genes involved in *Drosophila* head development are thought to play similar roles during vertebrate head development. Vertebrate head homeobox genes were identified using the *Drosophila* head gene homologs, for example *distal-less, orthodenticle,* and *empty spiracles.* [30] The vertebrate homologs to the *Drosophila* head homeotic genes were found to be expressed in a region-specific way in mid- and forebrain.

1.5.1 Forebrain Model

A model for early brain development was proposed by Boncinelli (1994) which involves vertebrate homeobox genes not found organized on the known *Hox* clusters, but are rather dispersed in the genome. [31] Vertebrate homeobox genes with anterior and posterior brain expression domains have been identified; *distal-less* -like *(dlx), orthodenticle*-like (*otx), empty spiracles*-like (*emx*), *sek* (tyrosine kinase receptor), and *krx20*. The expression domains of these anterior homeobox genes have also been shown to concur with neuromeric boundaries which implicates them in segmental determination.

The forebrain is divided into the anterior telencephalon and more caudal diencephalon. Recently it has been reported that the diencephalon is further subdivided into four neuromeres, D1 to D4. [32] These neuromeres seem to conform with all the criteria necessary for them to be termed *'segments'*. A number of homeobox genes have been identified which may play a role in identifying each diencephalic segment (figure 1-4). The analysis of the expression domains of the numerous distal-less (*dlx*) genes in the developing brain suggest that they may furnish instructions for the early subdivision and/or regional differentiation in the forebrain. In fact, *distal-less* genes are expressed fairly late during vertebrate development, which suggests that they may not be involved in the subdivision of the brain. Comparing the expression profiles of *dlx* and other neuromeric genes suggests that *dlx* genes are more likely to respond to previously established neuromeric boundaries, whereas other gene families, including *Otx* and *Emx* genes, could be involved in the early subdivision of the forebrain.

1.5.2 Craniofacial Development

1.5.2.1 Neurulation

Initially in all vertebrates, the ectoderm thickens and a neural plate forms (figure 1-5). Neural folds appear along the lateral sides of the neural plate.

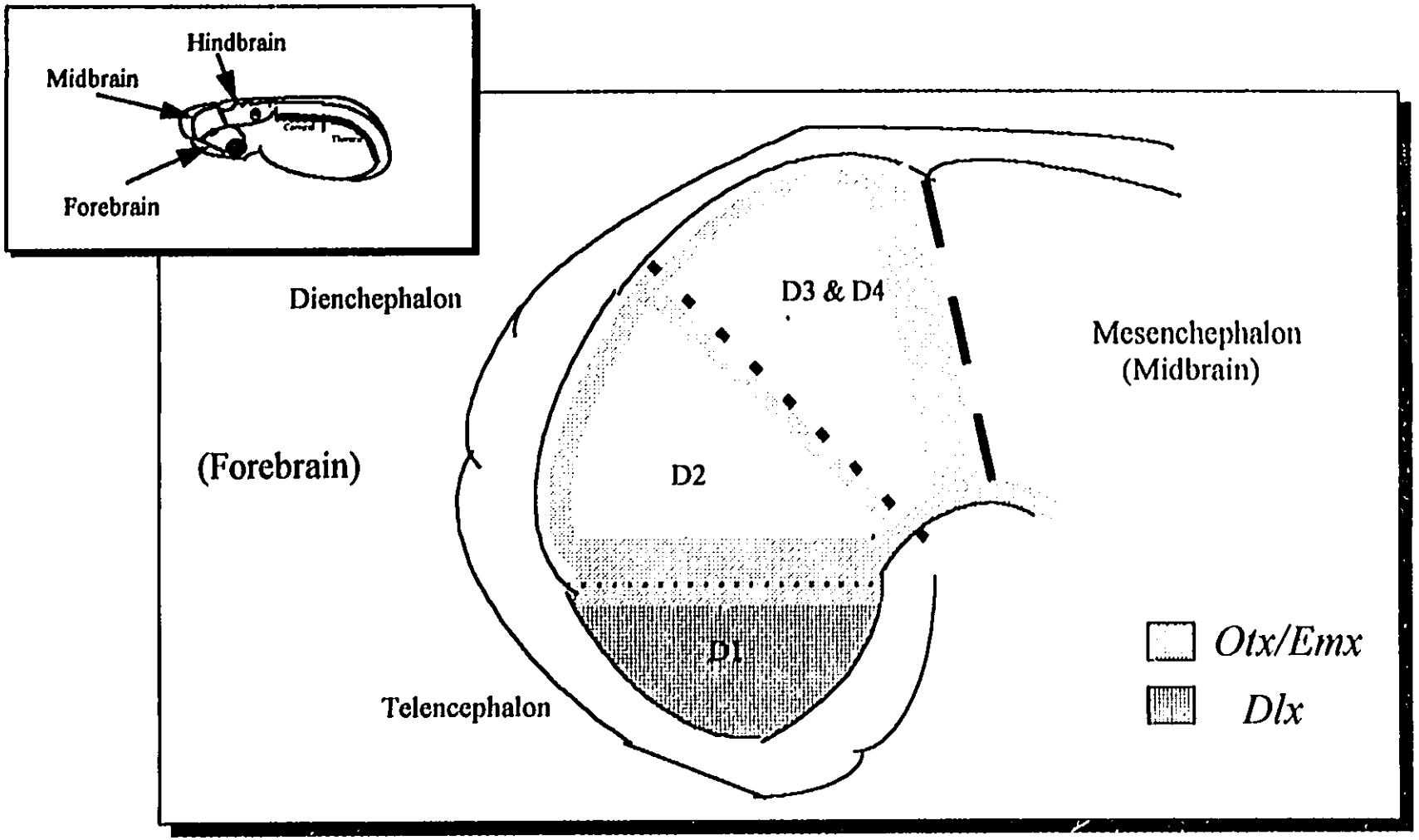


Figure 1-4. Boncinelli's Model on Forebrain Development

Figure 1-4: Boncinelli's Forebrain Model.

Illustration of a 12.5 dpc mouse forebrain. The expression patterns of *otx/emx* and *dlx* are shown, see legend. These expression patterns seem to concur with the forebrain diencephalic segments, D1 - D4. Adapted from Boncinelli, E (Current Opinion in Neurobiology, 1994, 4; 29-36).[31]

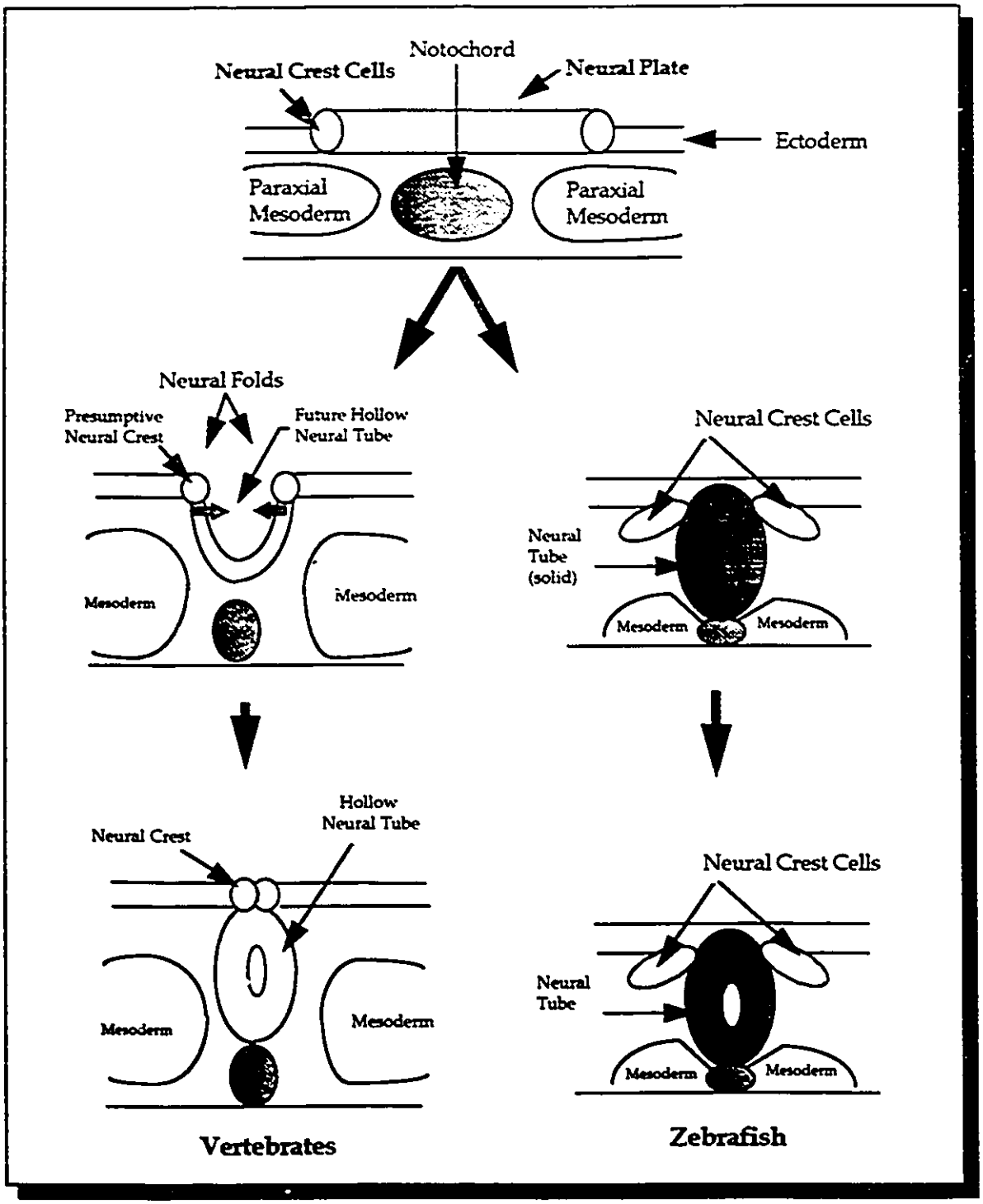


Figure 1-5. Comparative neurulation in zebrafish and other vertebrates

Figure 1-5: Comparative Neurulation.

Schematic representation comparing zebrafish and vertebrate neurulation. The zebrafish does not develop like other vertebrates. There is no neural fold closure in zebrafish. Rather, a solid neural keel forms instead of a hollow neural tube. Also note the difference between the zebrafish and the vertebrate neural crest cell's position.

The subsequent step during vertebrate neurulation pertains to the closure of the neural folds and migration of the neural crest from the dorsal neuroepithelium. However, in zebrafish, no neural folds appear and instead of a neural tube forming with a hollow lumen, a solid neural keel emerges at first, then a hollow lumen forms within the neural keel. [33] Lateral to the zebrafish's developing anterior neural keel (12 hpf) appears a mass, uniform in structure, called the premigratory neural crest. Many cranial neural crest cells never reach the dorsal midline, like other vertebrates, but remain laterally segregated with a distinct lateral border near the optic vesicle. This mass extends laterally to the otic placode.

1.5.2.2 Brain Segmentation

Segmentation which succeeds neurulation includes the subdivision of the anterior neural keel (zebrafish) into three components, the forebrain, midbrain, and hindbrain. [34] After the initial division of the brain (16 hpf), the hindbrain further subdivides into nine segments (18 hpf), referred to as rhombomeres (r) or hindbrain neuromeres (r1-r9). Rhombomere segmentation has been very well conserved throughout the evolution of all vertebrates. When the full complement of rhombomeres has been generated, the neuroepithelium is subdivided into repeating units. Each rhombomere unit contains boundaries that restrict the cells' movement within each rhombomere, thus categorizing each rhombomere as a unique individual segment.

Embryonic manipulations, essentially in the avian embryo, have demonstrated the importance of the hindbrain and the neural crest derived from the hindbrain neuroepithelium. The crest cells seem to be specified before migration as to the structures they will form. [35] It has been suggested that neural crest cells are patterned according to their rhombomeric origins. During migration the anterior-posterior order in which they arise is maintained therefore the segmental patterning is extended to regions outside the neuroectoderm. [36]

1.5.2.3 Neural Crest Migration

Forebrain and midbrain neural crest cells migrate to the frontonasal mass and contribute to the frontonasal mesenchyme and derivatives, whereas the

hindbrain neural crest cells migrate to the branchial arches in a segmental manner.

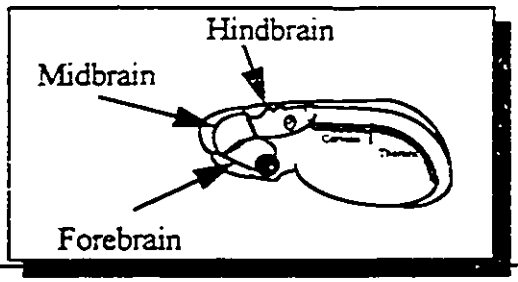
The onset of cranial crest migration has been shown to vary depending on the vertebrate species being analyzed. In the chick, cranial crest cells migrate from the neuroepithelium at or shortly after neural tube closure. [37] In rodents, cranial crest migration begins while the neural folds are still open at the midbrain region. [38] The rostral otic region crest cells (hyoid crest) migrate from an open V-shaped neural tube and the caudal otic region crest cells migrate from a closing or closed neural tube. In mammals, cranial crest cells migrate before, during, and after neural tube closure depending on their location along the anterior-posterior neuroaxis, but not in a simple rostral-caudal fashion. [39] There is no neural tube closure in the zebrafish. The anterior neural crest cells just caudal to the developing eye simply change form (15-16 hpf) and leave the lateral neural keel mass. Neural crest cells migrate from progressively more caudal regions thereafter. [40]

1.5.2.4 Hindbrain Neural Crest Migration

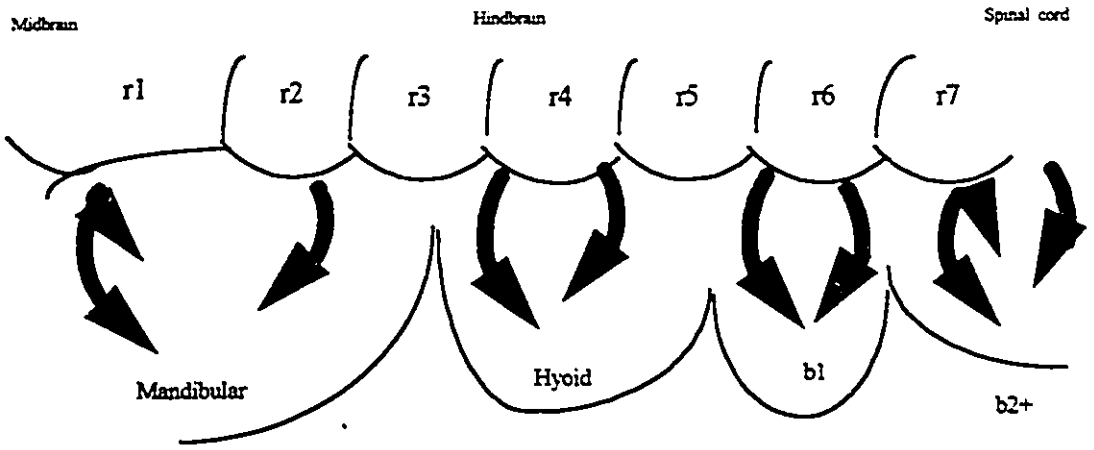
Lineage tracing experiments in chick embryos were used to define the fate of hindbrain crest cells along their anterior-posterior migratory paths (figure 1-6). [41] The avian hindbrain crest cells from r2, r4, and r6 mainly populate branchial arches 1, 2, and 3. It has also been reported in avian models that r3 and r5 rhombomeric locations do not produce neural crest cells that populate the branchial arches. [42] In zebrafish, the term *'visceral arches'* includes all arches from the mandibular and hyoid arches to the last branchial arch 4. Recently, Shilling and Kimmel (1994) used lineage tracers to follow the migration of zebrafish hindbrain neural crest cells. [40] Their results were somewhat contradictory to the avian neural crest cell fate map. The zebrafish neural crest cells from rhombomeric locations' r2/r3 migrate to the mandibular arch, and r4/r5 crest cells populate the hyoid arch.

1.5.2.5 Hindbrain Hox Model

Hunt (1991) proposed the following model for patterning of the vertebrate branchial arch region (figure 1-7). [29] Homeobox gene expression



Hindbrain Neural Crest Cell Migration in the Chick



Hindbrain Neural Crest Migration In The Zebrafish

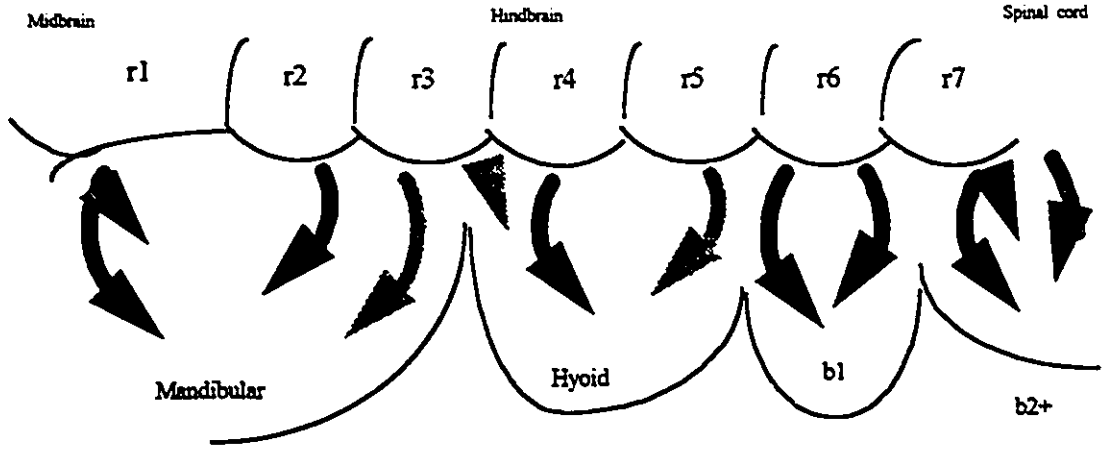


Figure 1-6. Hindbrain neural crest cell migration.

Figure 1-6: Hindbrain Neural Crest Migration.

Schematic illustration comparing the zebrafish and chick hindbrain neural crest cell migration into the visceral arches. In the chick and zebrafish r1/r2 crest cells migrate into the mandibular arch, r4 populate the hyoid arch, and r6/r7 crest cells migrate into the branchial arches. The main difference is in r3 and r5. In the chick no crest cells migrate from these rhombomeres, whereas in the zebrafish, r3 and r5 crest cells populate the mandibular and hyoid arches. In the zebrafish, neural crest cells from rhombomeric locations' r2/r3 migrate to the mandibular arch, and r4/r5 crest cells supply the hyoid arch.

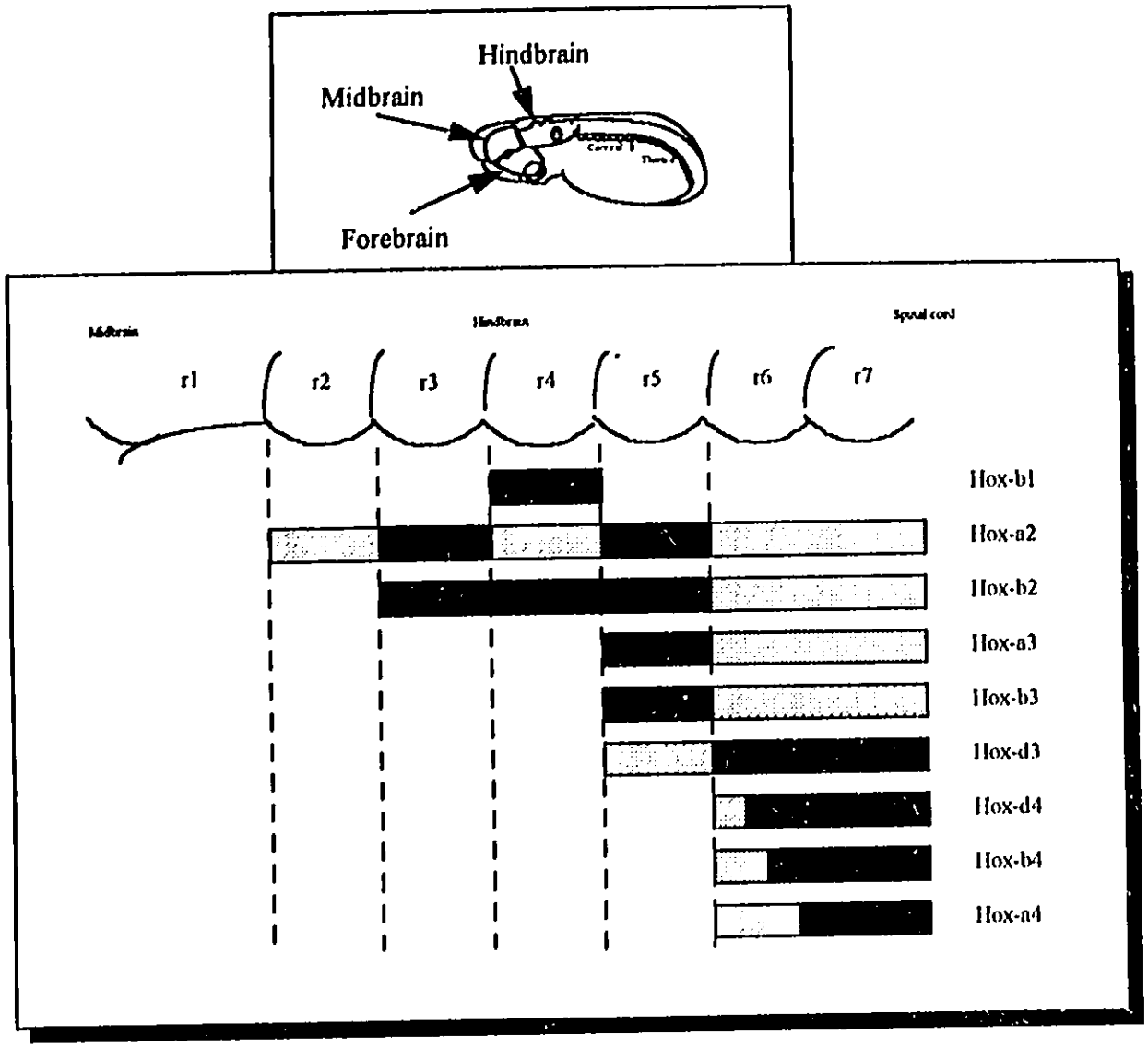


Figure 1-7. Hunt's Branchial Arch *Hox* Code.

Figure 1-7: Hunt's Hindbrain/Branchial Arch *Hox* Code.

Schematic representation of the hindbrain of a 9.5 dpc mouse embryo. Shown below the hindbrain is Hunt's *Hox* code which is rhombomere restricted. The dark shading represents high levels of expression whereas gray shading represents low expression. Note, no *Hox* genes are expressed in rhombomere 1. Adapted from Krumlauf, R. (TIGS 9(4) 106-111, 1993).[8]

(Zen-like, pb-like, lab-like) emerges in the neuroectoderm and presumptive neural crest cells giving the cells a unique segmental identity. A little later during development there is a segmental expression of the *Hoxb* cluster in the hindbrain, thought to assign a differentiation program to the presegmented crest cells. This *Hoxb* code is sustained during neural crest migration, resulting with a *Hoxb* code in the cranial ganglia and branchial arch mesenchyme that reflects the crest's rhombomeric origin. Therefore, each branchial arch contains a unique *Hoxb* code (except arch 1) and this specific *Hoxb* code is present in the hindbrain crest cells prior to migration.

The overlapping expression domains of the *Hox* paralogous groups might suggest redundancy. However, their specific functions may arise after primary specification of the rhombomeres and neural crest cells. One possible role is in specifying the range of different structures after primary specification has been accomplished. Thus, paralogous *Hox* genes would act in parallel and independently to specify regional identity in the different derivatives of the hindbrain neural plate. Krumlauf (1993) suggested that these paralogous *Hox* gene's expression patterns differ in two aspects. [8] First, the transverse plane of expression becomes dorsally restricted to regions which later develop into neural structures. Second, during hindbrain development the different genes within the paralogous groups have elevated expression domains that differ depending on the rhombomeric segments. The paralogous *Hox* expression could be required for inscribing the final differentiation program for the presegmented neural crest cells.

However, the clustered *Hox* genes can not be solely responsible for the specification of the hindbrain neural crest cells. In the mouse, hindbrain crest cells are given a segmental identity between 7-8.5 dpc. The earliest clustered *Hox* genes, the *pb-like* and *lab-like* paralogous groups, are expressed in the hindbrain at 7.75 dpc, which implies that they may be involved in the *segmentation competence signal*. However, the clustered *Hox* genes are not expressed in the hindbrain r1/r2 crest cells. This would suggest that there are other genes required for the establishment of the neural crest cell lineage's.

1.5.2.6 Neural Crest Differentiation

By 72 hpf the zebrafish neural crest cells have finished migrating and begin responding to their local environmental cues and ultimately differentiate into cartilage derivatives appropriate for their respective terminal location in the visceral arches. [43] The corresponding contributions from the neural crest cells to the formation of the skull and associated cartilage derivatives have been deciphered using both chick and quail marker systems. [44] The results concerning the origin of the visceral arch cartilage derivatives in teleost fish embryos[45] are summarized in figure 1-8.

The branchial (visceral in zebrafish) arch region of the vertebrate head is one part of the body whose early regional specification is thought to involve *Hox* genes. The prepatterned neural crest cells that lead to specific branchial arch structures always arise from particular rhombomeres[41], suggesting that the developmental patterning of the hindbrain and branchial arch regions are linked. Furthermore, the association between rhombomeres and branchial arches is conserved in all vertebrate species at similar developmental stages. However, the nature of the regulatory genes that govern pattern formation of the rhombomeres boundaries remains unknown.

1.5.2.7 Craniofacial Chondrogenesis

There are five distinct processes involved in craniofacial chondrogenesis; 1) origin of neural crest cells, 2) migration of the neural crest cells, 3) localization of the cells at their final destination (pattern formation), 4) differentiation, and 5) morphogenesis. [46] The neural crest cells arise from the neuroepithelium as typical epithelial cells but as they migrate through the basal lamina, they transform from epithelial to typical mesenchymal cells, ectomesenchymal cells. As these neural crest cells start to migrate they also begin to loose their polarity and secrete a matrix thereby enabling their migration.

A key property found in migrating neural crest cells is the downregulation of N-CAM molecules which lowers cell adhesion therefore aiding cell migration. A two fold increase in the number of cell adhesion molecules produces a thirty fold increase in rate of cell adhesion, therefore inhibiting neural crest migration. Injection of a single dose of RA during the primitive

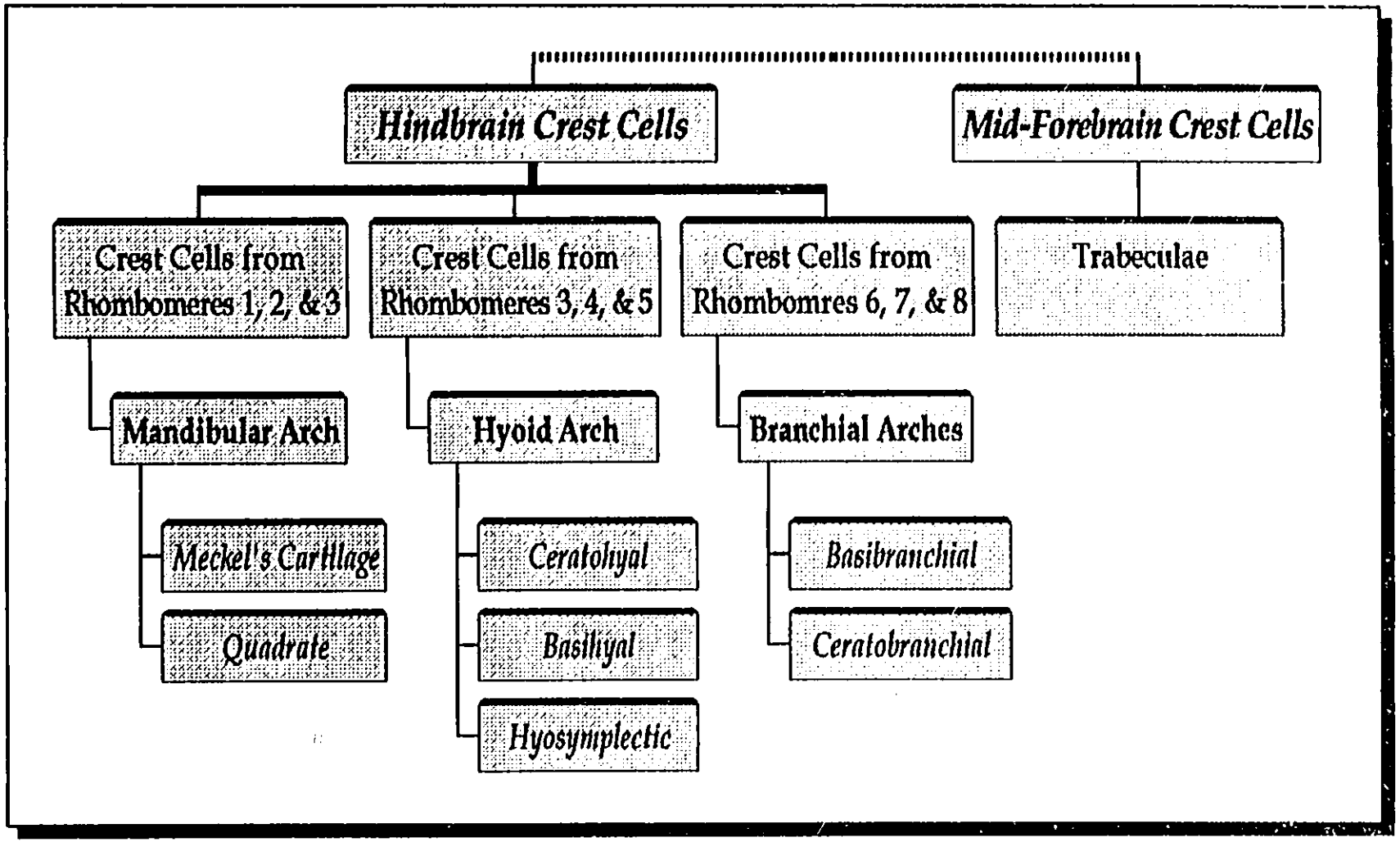


Figure 1-8. Neural Crest Derivatives.

Figure 1-8: Neural Crest Derivatives.

This flow chart illustrates the neural crest cells that give rise to particular craniofacial cartilage. The top of the chart originates with the neuroepithelia which divides in the brain into the fore-midbrain crest and hindbrain crest cells. The forebrain crest cells migrate anteriorly and differentiate, one component of their differentiation is the trabeculae. The hindbrain crest cells migrate ventrally into the visceral arches. By 72 hpf in zebrafish, the hindbrain crest cells have differentiated into their respective cartilage components. For example, the crest cells from r1, r2, and r3 migrate into the mandibular arch and later differentiate into the hyosymplectic, meckel's, and the quadrate cartilage components.

streak stage in mice results in an upregulation of N-CAM and a change in the neural crest cell fate that is irreversible. [47]

Once they have migrated to their final destination, an interaction between the crest cell and overlying epithelium is required for the subsequent initiation of differentiation of a particular determined cell state. How do the neural crest cells know when they have reached their terminal location? It is known that neural crest cells express a cell surface protein called anchorin C-II, which is a receptor for type II collagen. Type II collagen is found to be expressed in the forebrain, midbrain, hindbrain, optic and otic vesicles. It is believed that the crest cells migrate until they reach an area expressing type II collagen, where they get trapped. Some of these areas correspond to facial primordia. These trapped neural crest cell contain region specific patterning information from their original locations (*how many skeletal elements and what shape*), therefore they are mosaics of predetermined cells waiting for further instruction.

The vertebrate face contains a series of small facial primordia. [48] For example, in most vertebrates the upper jaw develops from five facial primordia: the frontonasal mass, two lateral nasal primordia, and two maxillae primordia. Unlike the lower jaw which develops from a pair of mandibular primordia. In the chick it has been shown that each facial primordia gets populated by neural crest cells from various origins. [46] For example, the frontonasal mass is populated by crest from the prosencephalic and anterior mesencephalic regions, whereas the maxillae primordia get populated from the posterior mesencephalic region.

Differentiation is thought to be a process involving both 1) epithelial-mesenchymal interactions, and 2) interactions among the differentiating cells. These interactions cause the neural crest cells (ectomesoderm) to condense and differentiate into cartilage. It has been reported that agents which disrupt collagen synthesis also prevent the epithelium from initiating bone differentiation. [49] Transplant experiments have shown that there exists a differentiation default pathway for the frontonasal mass. [50] The absence of the epithelial signals in the frontonasal mass causes the neural crest cells to differentiate into a lower jaw pattern.

Once the phase of cytodifferentiation has been initiated, morphogenesis begins. Morphogenesis is the development of change in cell shape. The shape of

the developing cartilage is determined by its own population of chondrocyte cell shape. Morphogenesis is accompanied with growth along a proximal distal axis.

The development of the face is very similar to the development of the limb. The same cell types differentiate even though these cells have very different developmental origins. The above data on craniofacial development coupled with the homeotic gene expression information to follow will hopefully answer some question as to how the craniofacial skeleton develops.

1.5.3 Craniofacial Dysmorphogenesis

1.5.3.1 Hox Loss and Gain-of-Function

Much of the documentation on the potential function of *Hox* genes is descriptive, and comes from studies using *in situ* hybridization. The function of a gene is not illustrated by its expression pattern, but is defined by its biochemical properties. For example, *Hox* genes regulate the transcription of other genes, and their function can indirectly be seen in the effects of the cell's physiology and/or development. Loss of a *Hox* gene's expression can only be correlated with a perturbed cellular differentiation, which would only confirm that its expression is important for that cell's normal development.

Certain *Hox* genes which are ectopically expressed in transgenic mice can alter craniofacial morphologies. Ectopic expression of *Hox-a7* or *Hox-b7* in transgenic mice results in death shortly after birth. [18] This ectopic expression of *Hox-a7/b7* also leads to multiple craniofacial abnormalities including cleft palate. [51, 52] This phenotype is similar to the effects seen with RA administration in mice[18], suggesting that ectopic *Hox* expression and RA administration share a common pathogenic pathway. One can hypothesize that the similar craniofacial effects seen with RA administration and ectopic Hox gene expression are the result of ectopic Hox expression induced by teratological doses of RA.

Hox loss-of-function mutants result in homeotic transformations, [13] but more interestingly, the parts associated with the transformations are primarily related to rhombomeric and branchial expression domains of the endogenous gene. Abnormalities mainly affect the neural crest derivatives.

Changing the expression of a single homeotic gene can disturb the overall *Hox* code, and create a new combination that is interpreted to form alternative structures.

1.5.3.1.1 Hox-A1

One example is the gene *Hox-a1*, which is normally expressed in the hindbrain's r4. It was shown in 1993 that a loss of the *Hox-a1* gene in the mouse resulted in altered rhombencephalon development. [53] More specifically, the first three rhombomeres were normal in development, but r4 and r5 were absent. [54] Therefore, r6 forms the fourth rhombomere. Recently (1994), ectopic expression of *Hox-a1* was reported to reorganize select regions of the developing hindbrain by inducing partial transformations of several rhombomeres into r4-like identity. The ectopic *Hox-a1* expressing transgenic mice also show ectopic *Hox-b1* expression in r2. [55] Therefore, the action of *Hox-a1* on *Hox-b1* regulation in the hindbrain suggests that *Hox-a1* directly regulates *Hox-b1* during normal development. The *Hox-d1* locus does not share the *Hox-a1* regulatory elements found within the *Hox-b1* locus, therefore *Hox-d1* expression is not induced by *Hox-a1*.

1.5.3.1.2 Kreisler Mutant

Recently a hindbrain mutant was identified in the mouse called the kreisler (*kr*) mutant. The homozygous mutant mouse displays an unsegmented neural tube in the location of rhombomeres 4 to 7 (figure 1-9). [56] It appears that the fundamental defect is in the loss of r5 and r6. One model proposes that the cells that would normally become specified as r5 and r6 adopt an r4 identity, producing an excess of r4 cells that are disposed of by cell death. To investigate whether the unsegmented hindbrain would be associated with abnormal branchial arch differentiation, Frohman (1993) examined neural crest derived tissues of the second and third branchial arches, which normally arise from r4 and r6. They found that the hyoid bone in *kr/kr* animals exhibited an accessory process on the greater horn (third arch structure) most easily explained by ectopic development of the second arch in an area normally derived from the third arch. [57] This would suggest that neural crest cells from r6

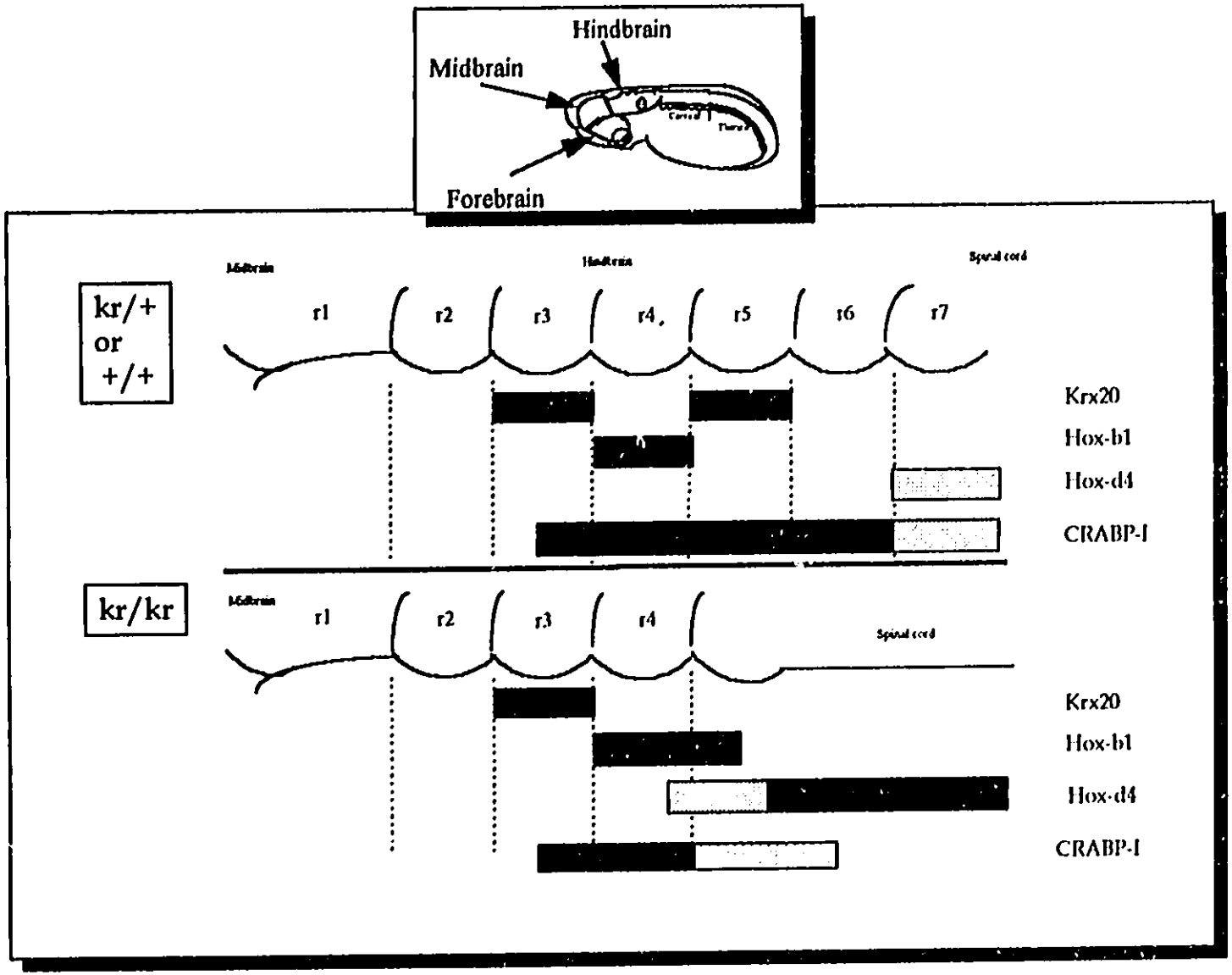


Figure 1-9. The kriesler mouse and the altered hindbrain *Hox* code.

Figure 1-9: Kreisler Mutant.

Illustration comparing a wildtype and *kr/kr* homozygote mouse hindbrain. The top panel shows a normally segmented hindbrain and normal gene expression. The bottom panel depicts the *kr/kr* homozygous mutant mouse hindbrain and gene expression. The abnormal *krx20* gene expression implies that r5 is absent. Whereas *Hox-d4* is expressed just posterior to r4 indicating a loss of r5-r7, which is confirmed by CRABP I expression. Adapted from McKay et al. (Development, 1994, 120: 2199-2211).[56]

are not given the differentiation signal either because the crest cells are not present or they can not interpret the signal.

1.5.3.2 Retinoic Acid and Craniofacial Development

Retinoic acid has been shown to have dramatic effects on craniofacial and neural development *in vivo*. Evidence that RA-induced phenotypes in the mouse embryo are related to changes in *Hox* expression has been provided by *in situ* hybridization analysis. [8]

When RA is applied to chick embryos at early stages of facial primordia development, craniofacial defects occur. [58] The upper beak (derived mainly from the frontonasal mass) is missing whereas the lower beak (derived from the mandibular arch) is normal regardless of the treatment stage. [59, 60] The frontonasal mass fails to enlarge and fuse with the lateral nasal processes and maxillae. This defect is classified as severe clefting of the primary palate. The loss of the upper beak and development of a normal lower beak is reminiscent of the frontonasal default pathway (section 1.5.2.7).

Retinoic acid has been shown to interfere with pattern formation in the limb. Epithelial-mesenchymal interactions governing pattern formation and growth of the craniofacial skeleton are sensitive to RA treatment. [21] Removal of the facial ectoderm results in truncation of the facial primordia derivatives. Recombination experiments have demonstrated that the mesenchyme, not the epithelium, of the facial primordia is the target for RA action. It has been shown that the facial epithelium has similar properties to the apical ectodermal ridge (AER) of limb bud, suggesting that proximal-distal pattern formation in the facial primordia involves a *"progress zone"* mechanism. [22]

Retinoic acid has also been shown to inhibit chondrogenesis in micromass cultures. More specifically, RA affects procollagen synthesis in a dose dependent manner. [61]

1.5.3.2.1 Rhombomere Respecification

The use of genetic markers present during normal brain development has provided information as to what structures are absent or present in RA treated embryos. For example, both *pax2* and *eng* genes are necessary for normal mid-hindbrain development, and their elimination correlates with a loss of the mid-

hindbrain region. A loss in *krx20* expression is associated with a loss in the development of the anterior rhombomeres r1-r3. G*oosecoid* (*gsc*) is another example; its expression is required for normal anterior forebrain development and its elimination is related to a loss of forebrain development.

Marshall (1992) suggested a model for the effects of RA on rhombomere development. [62] At 7.5 dpc in the mouse, RA is unable to override the specification choice of an odd- or even- number rhombomere character, but RA induces a change in their identity. RA treatment at 7 dpc causes a compression of the most rostral rhombomeres (r1-r3) which leads to extensive alterations in the midbrain/forebrain structures. Thus, RA can induce rapid changes in *Hox* expression that eventually leads to transformations in craniofacial regions, suggesting that low doses of RA might act as an endogenous morphogen required for normal craniofacial development.

Wood (1994), showed that RA treatment prior to murine somatogenesis (at 7.75 dpc) resulted in suppression of rhombomeric segmentation, and the genetic identity of the whole preotic hindbrain was converted to a r4-like identity. [63] In contrast, exposure to RA during early somatogenesis (at 8.25 dpc) resulted in nearly normal rhombomeric segmentation. However the rhombomeric *Hox* gene expression indicated that only r2 had changed its genetic identity to that of r4 (figure 1-10). Thus, these results indicate that RA has separable effects; 1) on the genes mediating the processes of rhombomere segmentation, and 2) on the genes that influence the identity of the rhombomere.

1.5.3.2.1.1 RA and Hindbrain Neural Crest Cells

The fate of the epithelium to become neural epithelium and crest cells are established around 7 dpc in the mouse, because when mice are treated with RA prior to 7 dpc no anterior neural structures develop, consistent with the above hypothesis that this treatment prevents the epithelium from becoming neural epithelia.

The signal which imprints the competence of a neuroepithelial cell to become a neural crest cell is determined by an unknown set of regulatory genes. Once the cells have been specified to become neural crest cells, they must receive a signal for their segmentation properties. This segmentation signal

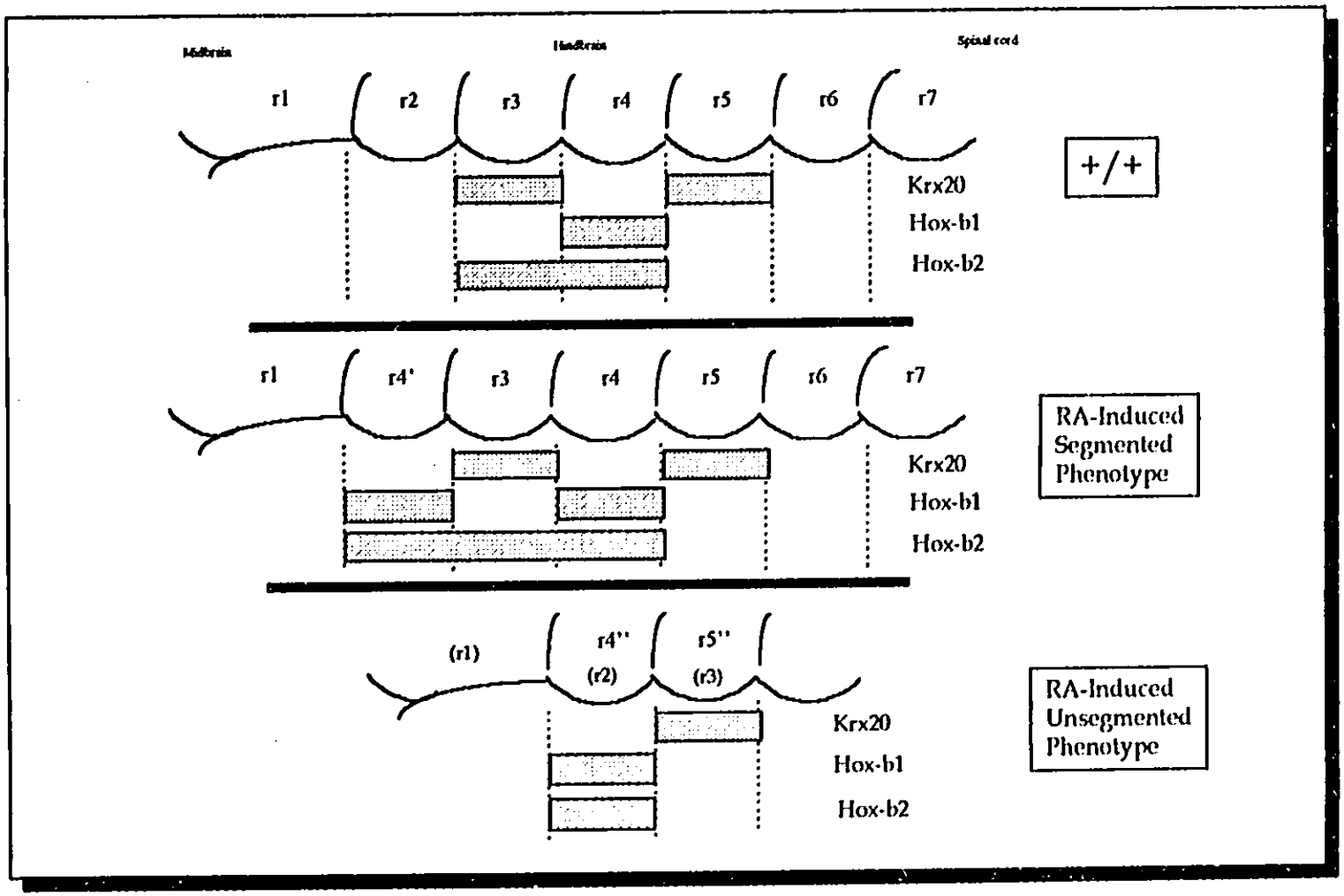


Figure 1-10. Effects of RA on Hindbrain Segmentation and *Hox* gene expression.

Figure 1-10: RA Effects on Hindbrain Development.

Schematic of a 9.5 dpc mouse segmented hindbrain. Top panel, Wildtype hindbrain and gene expression. Middle panel, mouse hindbrain after RA treatment at 8.25 dpc. The hindbrain retains its segmentation but r2 is respecified into r4-like determined by *Hox-b* expression. Bottom panel, mouse hindbrain after RA treatment at 7.75 dpc. Wildtype r3 rhombomere expression includes a unique combination of *Hox-b2* and *krx20*, whereas wildtype r2 has neither expression of *Hox-b* nor *krx20*. Therefore, in RA treated 7.75dpc mice r2 and r3 have a new identity resembling r4 and r5. Adapted from Wood et al. (1994, Development, 120; 2279-2285). [63]

has been shown to occur before crest migration (8 dpc mouse, 13 hpf in zebrafish). If RA is administered before the crest cells receive their segmentation signal, the phenotype would be an unsegmented or compressed hindbrain.

The segmented neural crest cells require information about their differentiation program, which seems to arise before crest cells begin to migrate. RA treatment at 8 dpc results in rhombomere respecification. By 8 dpc the crest cells would have been signaled to become segmented, so RA administration only causes a failure in the final differentiation program of unique segments.

1.5.3.2.2 RA and Zebrafish

Retinoic acid administration has also been shown to have dramatic effects on the developmental patterning of the zebrafish embryos. The known effects include; loss of midbrain-hindbrain border (treated between 5-10hpf) [64], loss of anterior heart structures (treated between 5-13.5 hpf) [65], and a duplication of the retina (treatment during optic primordial development). [66] Recently (1994), a group reported that RA treatment during early gastrula stages produced an abnormal mid-hindbrain development, but more interestingly, r2 is respecified into r4-like. [67] This rhombomeric respecification has also been shown to occur in mice when RA is administered during early somitogenesis. [63]

1.5.3.2.3 RA and Gene Suppression

Retinoic acid is known to cause developmental dysmorphologies by regulating gene expression. The normal developmental patterns of gene expression are altered when RA is deficient or is present in excess amounts. RA mediates gene expression through two families of nuclear steroid receptors, the retinoic acid receptor (RAR) and the retinoid X receptor (RXR). Four isoforms of both the RAR and RXR have been isolated; α, β, γ, and d. The retinoic acid receptor has a high affinity for both *all-trans* RA and *9-cis* RA (metabolite of *all-trans* RA), whereas the RXR receptor has affinity for only *9-cis* RA. When embryos are treated with *all-trans* RA, a proportion of *all-trans* RA is metabolized into *9-cis*

RA. These steroid receptors are found in the nucleus bound to genomic DNA as either homodimers (ie. RAR/RAR) or heterodimers (ie. RXR/RAR). RXR has been shown to also heterodimerize with other members of the steroid receptor super family, namely the thyroid hormone receptor (RXR/THR) and the vitamin D receptor (RXR/VDR). The steroid receptors bind the genomic DNA at sites called response elements, each type of receptor recognizes its unique DNA sequence. The response elements are composed of two half-sites, therefore two receptors bind to one response element. To cause transactivation of the target gene, two steroid hormone receptors must be bound to their response elements and also be complexed with their ligands. Retinoic acid response elements (RAREs) and RXREs have been found in regulatory regions of many endogenous genes, for example the alcohol dehydrogenase 3 gene, laminin B1 gene, and the complement factor H gene. Other genes have also been shown to be RA-sensitive *in-vitro*, ie. phosphoenolpyruvate carboxykinase, oxytocin, apolipoprotein A1, major histocompatability complex class I, and osteocalcin genes. It is important to note that these *in-vitro* studies do not necessarily apply *in-vivo*.

In normal conditions RA is not supplied to the embryo directly, rather it is synthesized from retinol. The developmental requirement for RA is dependent on the embryonic tissue, some tissues require high levels of RA whereas others require less. The level of RA in embryonic tissues is therefore regulated in a tissue dependent manner. This regulation comes from the cytoplasmic retinoid binding proteins (CRABP). To date only three CRAB protein isoforms from murine and Xenopus have been isolated, one that is specific for retinol and the other two specific for retinoic acid. Binding studies showed that CRABP I has a 3-5 fold higher affinity for RA than CRABP II. [68] During embryonic development all three CRAB proteins are expressed in a clearly distinct spatial and temporal fashion (figure 1-11). In the *axolotl*, CRABP is first developmentally expressed after neural tube closure (somite 1-3 stage) in the hindbrain region, its expression spreads caudally as the neural tube develops. CRABP expression is also found in the 8.5 dpc mouse embryo in regions that correspond to the frontonasal mesenchyme, visceral arches, and rhombomeres, especially r4-r6. [69] It has also been reported that RA induces the expression of both CRABP I and CRABP II, which act specifically on RA. [70]

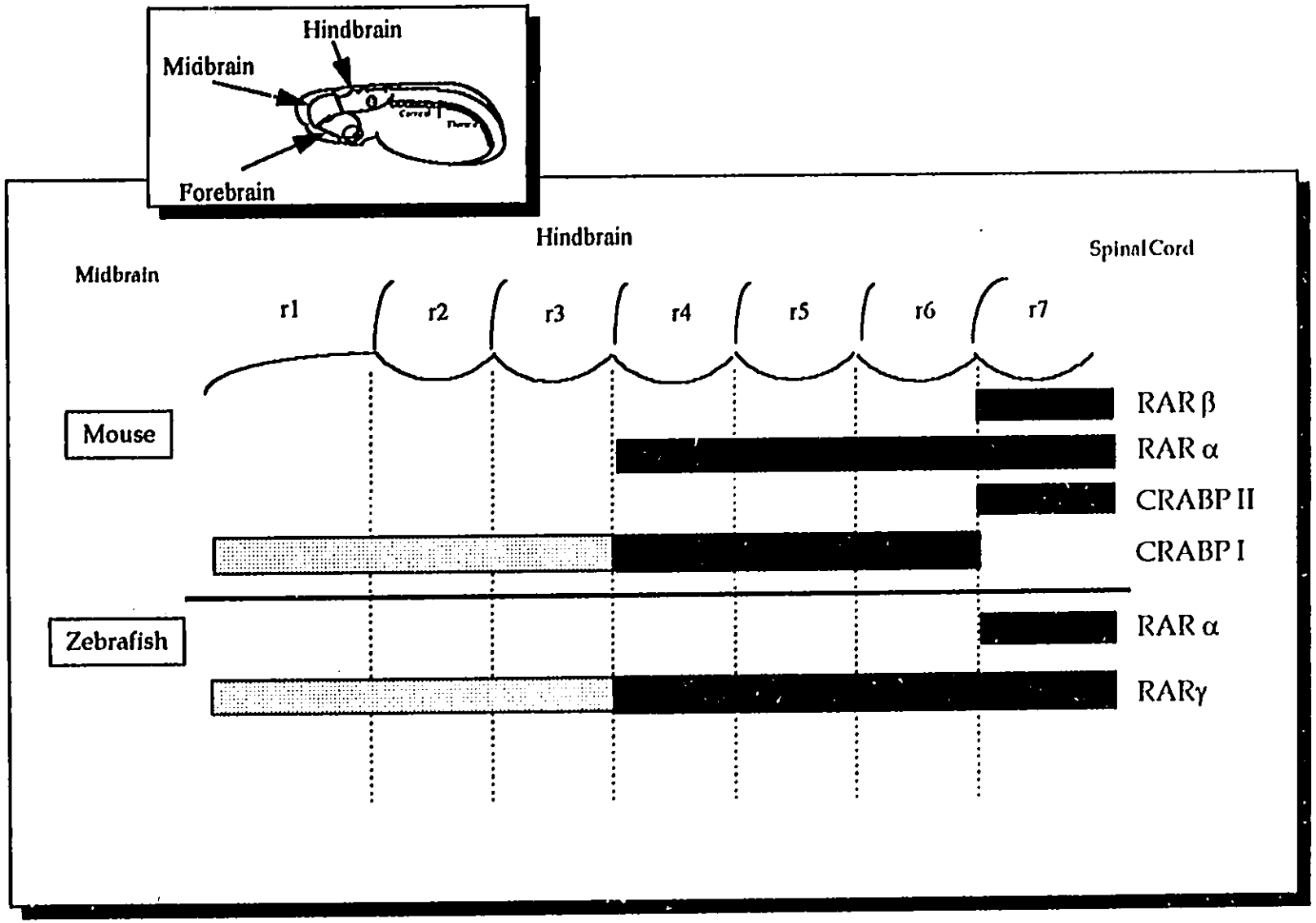


Figure 1-11. Diagramatic summary of the retinoic acid receptors and CRABP expression in zebrafish and mouse embryos.

Figure 1-11: Developmental Expression of RARs, RXRs, and CRABPs in the Hindbrain.

Schematic representation of a 9 dpc mouse hindbrain. Top panel, wildtype expression of RARs and CRABPs. Note low expression of CRABP and no RAR expression in r1-r3. Bottom panel, expression of RARs in the zebrafish hindbrain at 24 hpf. RARb has not yet been analyzed and CRABP have not yet been isolated in the zebrafish. Dark shading represents high gene expression whereas, gray shading is low gene expression. Adapted from Morris-Kay, G. (Bioessays, 1993, 15(1) 9-15).[72]

It has been postulated that CRABP I and II act positively on RA by capturing RA and therefore delivering it to the nucleus. [71] However, CRABP action is not required for the RAR/RA activated transcription, since RA can act in tissues that have no CRABP expression. Contrarily, it seems more likely that CRABP acts negatively on RA by sequestering and degrading cytosolic RA, therefore reducing the amount of RA reaching the nucleus. [73]

It has been recently reported (1993) that ectopic CRABP I expression in mice resulted in defective differentiation of the lens fiber, ultimately causing cataract formation. [74] The effect of CRABP on lens fiber differentiation could be explained by a reduction in the level of free RA in the lens fiber, either because CRABP is sequestering free RA, or CRABP is catabolizing the free RA. The F-crystalline gene has previously been reported to contain a RAR response element in its promoter region. [75] One would expect to see a downregulation of this gene due to a decrease in cellular RA with ectopic CRABP expression, however none was detected. One possible explanation is that this gene is not involved in the above affected process.

When there are no CRAB proteins in a cell, RA enters the cell nucleus where it binds to its receptor. It has been shown that RARβ needs 5 - 10 times less RA than RARα, and RARα requires 5 - 10 times less RA than RARγ in order to activate target gene expression. One target gene for RA is the RARβ gene, which is upregulated in the presence of RA, and causes fetal dysmorphogenesis in most vertebrate species. Fetal dysmorphogenesis has been reported to be caused by the RARβ mRNA level remaining elevated for a length of time, 6-9h in the mouse. [76]

A model proposed by Maden et al, 1992, suggests that CRABP I is not involved in transducing the teratological effects caused by RA. [77] This model is based on the evidence that the mandibular arch (no CRABP expression) is the primary site for RA action, however this not always the case. As seen in the previous section, the primary site for RA action is also in the mesenchyme of the frontonasal mass, and the mandibular arch is unaffected. Low levels of CRABP are found in the mandibular arch, unlike the other arches which have high CRABP expression. One could then hypothesize that the high levels of CRABP in r2-r8 may protect the hyoid and branchial arches from RA treatment. However,

low CRABP levels in r1 may explain why only the mandibular arch is affected at certain RA treated developmental stages.

In *Xenopus* embryos, RA has been shown to be a *'competence modifier'* of mesodermal cell fate. RA modifies the action of mesodermal inducing molecules like activin, TGFb, or FGF by repressing the mesoderm-inducer's target genes like *goosecoid*. For example, TGF b's are part of the growth factor family and are known to be involved in cell growth, differentiation, and morphogenesis. Exposure of early neurula stage embryos to RA showed a down-regulation of TGF- b1 and TGF- b2. [78] TGF- b1 was specifically reduced in the neuroepithelium and cranial neural crest cells for up to 48hrs after RA administration, whereas TGF- b2 expression was reduced for up to 60hrs after RA treatment. Pharmacokinetic studies in mice have reported that orally administered RA reaches a peak in 1h and falls to near normal levels within 8h. [23] This implies that either RA is acting indirectly from a sustained elevated level of RAR expression or that RA acts directly on TGF which would in turn autoregulate itself.

Many studies have been focused on the genetics of muscle differentiation and on the effects of RA on muscle differentiation. The conversion of mesodermal undetermined cells into myoblasts requires the action of the *MyoD* family of transcriptional regulators, four basic helix-loop-helix proteins; *MyoD, Myf5, myogenin,* and *Mrf4/Mrf6/herculin*. *Myf5* is the first of these genes to be expressed in the muscle precursor cells. *Myf5* has been shown to be downregulated by RA treatment *in-vitro*[79], suggesting that RA is required to maintain mesodermal precursor cells in an undetermined state.

Myf5 downregulation was shown to be specifically due to the action of *9-cis* RA on RXR, rather than *all-trans* RA on RAR. [79] The action of RAR and RXR on target gene activation may therefore be very different. For example, it is possible that high cellular RAR levels cause RAR homodimerization and activation of gene transcription, whereas high RXR levels cause a preferential heterodimerization with either RAR, VDR, or THR, which may cause target gene suppression.

The *COUP* receptors are vertebrate homologs of the *Drosophila seven-up* gene which is required for determining *Drosophila* photoreceptor cell fate. *COUP* orphan (unknown ligand) receptors are a class of transcriptional repressors that restrict RA, vitamin D_3, and thyroid hormone response pathways. [80] It is believed

that they bind to the RXR which then heterodimerizes with RAR, VDR, and THR. This complex of COUP/heterodimerized receptors would thereby compete for the binding of the receptor's response elements. The competition would thereby inhibit DNA binding of both heterodimeric steroid receptors. A new orphan receptor was recently identified, RVR, and shown to be a negative regulator of gene transcription. [81] RVR is highly similar to the viral form of THR, V-erbA. THR has also been shown to be a transcriptional repressor in the absence of ligand, whereas V-erbA is a potent constitutive repressor. The competition of various repressors and activators for the same DNA binding site may explain why administration of RA causes both activation and repression of different genes. Therefore, the transcriptional control switch may be regulated by the ratio of repressor-activator expression in any given cell.

Some homeobox genes have also been shown to be downregulated by RA. It is still uncertain whether or not these interactions with RA are direct or indirect. Local RA treatment in the limb bud generates mirror image digit duplications and has been shown to rapidly suppress both ectodermal and mesenchymal expressions of the *msx*1 homeobox gene. [82] *Hox-a10, Hox-a11,* and *Hox-d10* have been reported to be down regulated by RA *in-vitro*. [83] One example of how RA can repress *Hox* gene expression is shown in a study by Studer (1994). [84] The molecular mechanism behind the restricted r4 *Hox-b1* expression was analyzed. It was found that the regulatory region of the *Hox-b1* gene contains both enhancers and repressors, where the enhancer generates expression extending from r3 to r5, and a repressor which limits this domain to r4. The repressor contains a conserved RA response element. This suggests that RA is directly involved in sharpening segment restricted expression on *Hox-b1* during rhombomere boundary formation.

1.6. The Distal-Less Family Of Homeobox Genes.

The first member of the *distal-less* family of homeobox genes was identified in *Drosophila (Dll)*. [85] Since then nineteen vertebrate *distal-less* (*dlx*) genes have been isolated. Five *distal-less* genes (*dlx)* belong to the mouse, [86-90] one from the rat (*rdlx*), [91] two from the newt (*NvHBox*), [92] six from *Xenopus* (*Xdll*), [93-95] and six from zebrafish (*dlx*). Three zebrafish *dlx* genes *(dlx2, dlx3,*

dlx4) were previously identified by Akimenko and colleagues. [96] The other three zebrafish *dlx* genes *(dlx1, dlx6 and dlx7)* are identified in this manuscript.

Clearly vertebrate species contain multiple *distal-less* genes, in contrast to the single *distal-less* gene found in *Drosophila*. These genes are classified as *distal-less* genes because their homeoboxes share a high degree of similarity with that of the *Drosophila distal-less* gene. Phylogenetic analysis has shown that the *distal-less* gene family is made up of at least six orthologous groups (figure 1-12).

1.6.1 Genomic Organization

The best characterized group of *Dlx* genes are those of the mouse. It has been reported that the mouse *Dlx1&Dlx2* and *Dlx5&Dlx6 genes* are linked in an inverted fashion, so they are transcribed towards each other. [86] The *Dlx1/Dlx2* locus is found on chromosome 2, whereas, the *Dlx5/Dlx6* locus is located on chromosome 7. [86] Many human congenital disorders resulting in a combination of facial and limb defects have been mapped to the chromosomal location 7q21, which is very close to the *Dlx5/Dlx6* 7q22 locus. Interestingly, the *Dlx1/Dlx2* locus at 2q32 is located very close to the *HoxD* locus at 2q31. [86] The genomic organization of *dlx* genes is very different from that of the *Hox* genes which are found clustered in tandem arrays.

The genomic organization of the *distal-less* related genes is very well conserved from insects to mammals. Each gene contains three exons, and the homeodomain spans exons two and three.

1.6.2 Distal-less Evolution

The six orthologous *dlx* groups can further be divided into; *Dlx1/Dlx6, Dlx2/Dlx5, Dlx7* and *Dlx3*. These groups are situated as three diverging branches in the phylogenetic tree (figure 2-4). It is then possible that these six vertebrate genes arose from a common ancestral *distal-less* gene. The ancestral *Dll*-like gene would have undergone a duplication, thus creating two vertebrate *Dlx* genes (possibly *Dlx1* and *Dlx6*). These vertebrate Dlx genes would have then undergone an inversion / duplication event thus

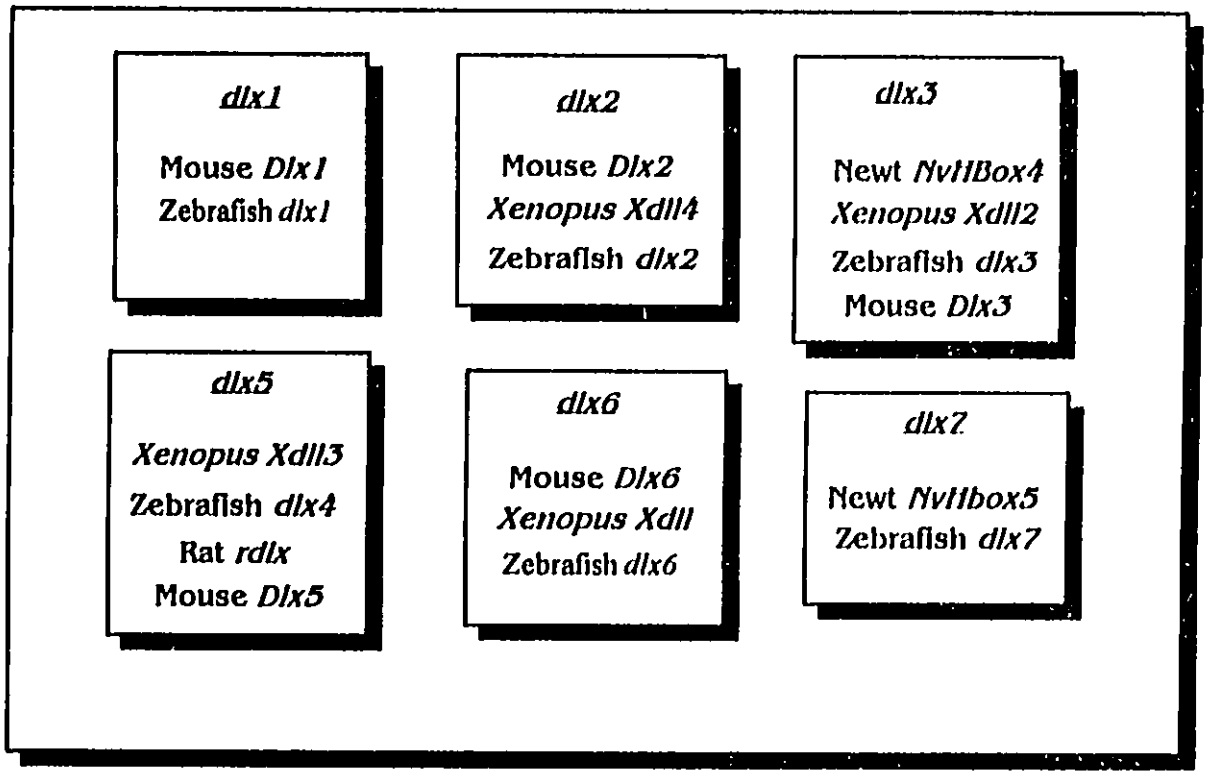


Figure 1-12. Vertebrate *distal-less* Orthologous Groups

Figure 1-12: *Dlx* orthologous Groups

The following figure depicts the six orthologous *dlx* groups. The sixth Dlx orthologous group will be discussed in Chapter 2, with the isolation of the three new zebrafish *dlx* genes. The mouse *Dlx5* and *Dlx6* full length amino acid sequences have not yet been published.

creating two inverted *Dlx* genes which are genomically linked, like the mouse *Dlx1* and *Dlx2* locus. These two events would have created the *Dlx* linkage already documented in the mouse (*Dlx1/Dlx2, Dlx5/Dlx6*). *Dlx3* and *Dlx7* would have appeared from a duplication of either *Dlx5* or *Dlx2*.

1.6.3 Distal-less Gene Expression

1.6.3.1 Drosophila

Distal-less (Dll) is first expressed in the head of the *Drosophila* blastoderm stage in two incomplete dorsal stripes. One stripe corresponds to the presumptive antennal limb primordia, whereas the other stripe is for the presumptive maxillary and labial primordia. During early gastrulation *Dll* is expressed anterior to the antennal primordia corresponding to the presumptive labrum. Germ band extension brings about distention of the maxillary/labial *Dll* stripe resulting in two distinct *Dll* patches corresponding to each primordia. During late germ band extension in the drosophila, *Dll* begins to be expressed in maxillary cirri, in optic formation (brain), and at the posterior base of the spiracles. In conclusion, *Dll* is involved in the development of appendages, and its expression is most abundant in appendages deriving from the gnathal (head) segments. [85]

Dll is required for limb development in the adult. This gene acts as a developmental switch between becoming a body wall or becoming a limb. If *Dll* expression is eliminated the fly develops a body wall instead of a limb. Interestingly, the limbs of the gnathal segments are associated with the larval limb primordia. Later during development *Dll* is also expressed in the head skeleton. In *Dll Drosophila* mutants, the larval head skeleton is collapsed or twisted.

Many groups have been involved in deciphering the mode of *distal-less* action in *Drosophila* because of its important role in limb development. This includes isolating the upstream and downstream target genes from *distal-less*. The first clue as to what upstream gene was involved in regulating *distal-less* came from a study conducted by Cohen, 1990. [85] In this study it was shown that the segment polarity gene *wingless (wg)* plays a key role in defining the positions at which limb primordia will develop along the anterior-posterior axis. In 1992,

the isolation of *Dwnt-5* was reported, a gene which shares high homeobox similarity to the vertebrate *Wnt* family of genes. [97] This gene is expressed in the gnathal and thoracic limb primordia, and it was shown that the expression of *Dwnt-5* is dependent on the expression of *Dll*.

In embryo's mutant for the *Bx-C* genes, *Dll* is derepressed in the abdomen correlating with the formation of abdominal limbs. To determine which *Bx-C* gene was responsible for the repression of *Dll* in the abdomen, Vachon and collaborators (1992) dissected the 5' regulatory region of the *Dll* gene. [98] The *Bx-C* gene *Ubx* was reported to bind to a repressor element found 5' to the *Dll* gene in-vitro.

In 1993, another putative *Dll* upstream gene was reported, namely *deformed (Dfd)*. [99] *Distal-less* was shown to be required for the development of a subset of maxillary epidermal identities which are specified by the *deformed* gene. This regulatory effect was shown to be mediated in a tissue specific manner (maxillary) by an enhancer region located 3' to the *Dll* gene.

1.6.3.2 Vertebrates

Embryonic expression patterns for *Dlx1, Dlx2, Dlx3, Dlx5,* and *Dlx6* have been studied in the mouse. Four of the five mouse *Dlx* genes (*Dlx1, Dlx2, Dlx5* and Dlx6) show a very similar expression pattern in the forebrain (ventral thalamus). *Dlx1* is expressed in the branchial arches, tooth buds, during ocular development, in neural crest derivatives, and in the limb bud apical ectodermal ridge. *Dlx2* is expressed 2 days before and in the same tissues as *Dlx1*, but with a lower intensity. *Dlx5* and *Dlx6* also show a similar pattern of expression, however *Dlx6* is expressed with a lower intensity. Both *Dlx5* and *Dlx6* genes are expressed in the branchial arches, otic vesicle, and the frontonasal ectoderm around the olfactory placodes. Later during development *Dlx5* and *Dlx6* are expressed in all tissues undergoing chondrogenesis. Recently, a group documented that the mouse *Dlx3* expression patterns are different from the other mouse *Dlx* genes. [87] For example, *Dlx3* is not expressed in the developing central nervous system, rather it is expressed in a highly restricted pattern in the branchial arches.

Distal-less expression in other vertebrate species closely resembles that found in the mouse. These genes are found in the developing forebrain, with the

exception of the *Dlx3* orthologous group members. The *Xenopus Xdll3,* Dlx5 ortholog, appears to be expressed very early during development in the anterior transverse ridge of the open neural plate. Another exception, is the *Xenopus Xdll* gene, tentitively a Dlx6 ortholog, which is found to be highly expressed in the oocyte, unfertilized eggs, in the very early zygote, and during late neurula stage.

The *distal-less* gene's upstream and downstream target genes have not yet been deciphered in vertebrates. Full length homeotic proteins from *Drosophila* have been successfully expressed and purified, unlike their vertebrate homologs. However, the main reason for the advancement in *Drosophila,* is the number of mutant fly strains available compared to that found in vertebrates.

The high level of functional and structural conservation between invertebrate and vertebrate homeobox genes may allow one to extrapolate information already documented for invertebrates to vertebrates. The vertebrate homologs to the *Drosophila* homeobox genes, that were discussed in section 1.6.3.1, are shown below- the *Drosophila* gene is followed by its vertebrate homolog: *deformed* (*Dfd*) is homologous to the fourth vertebrate paralogous group, ie. *Hox A4-D4; Wingless* (*wg*) is *Wnt-1; Dwnt-5* is *Wnt-5; Ubx* is homologous to the seventh vertebrate paralogous group. Experiments in Drosophila have shown that the regulatory elements necessary for vertebrate *Hox* gene expression are conserved with *Hom-C* genes in *Drosophila.* [100] They demonstrated that *Hom-C* genes were able to regulate *Hox* gene's that were exogenously present in Drosophila.

One can speculate that during vertebrate development, *Wnt-1* binds to its receptor which would be found on cells from the anterior developing brain, it's signal would then pass through the cell and activate *distal-less* expression. *Distal-less* would in turn activate *Wnt-5* expression. The vertebrate homolog to the *Drosophila Dfd* gene would not be involved in the development of the anterior brain, because in vertebrates the fourth paralogous *Hox* group nas an anterior expression limit at r7 in the hindbrain. In *Drosophila, Dll* is required for limb development. One could then speculate that vertebrates may also require *distal-less* expression for limb development. The areas that do not develop limbs in the vertebrate, would develop similarly to the *Drosophila's* abdomen,

where *Ubx* is required to repress *dll* expression. In vertebrates, maybe either or both, *Hox-a7* or *Hox-b7*, would be responsible for the repression of *dlx*.

1.6.3.2.1 Zebrafish Distal-less Expression

Three members of the *distal-less* homeobox gene family have been isolated in the zebrafish. [96] These genes fit into the orthologous groups *Dlx2*, *Dlx3*, and *Dlx5*, and are named *dlx2*, *dlx3*, and *dlx4*. The earliest of the three genes to be expressed is *dlx3*, which is expressed in two bilateral anterior-posterior stripes during gastrulation. The onset of *dlx2* and *dlx4* expression is later at about 13 hpf, where the transcripts are expressed in the ventral forebrain rudiment. Unlike most *distal-less* genes, *dlx3* is not expressed in the ventral forebrain. However, this lack of expression is characteristic of the orthologous *Dlx3* group. [96] Shortly after the onset of *dlx2* expression, *dlx2* transcripts are detected in the migrating neural crest cells of the hindbrain. Pharyngeal arch cells express all the *dlx* genes in a combinatorial fashion with different temporal and spatial expression pattern. A striking feature of zebrafish *dlx* expression patterns is that diverse regions of the embryo express particular patterns of the different members of the given family (figure 1-15).

Akimenko and collaborators (1994) suggested that *dlx* genes are part of a unique *"homeobox code"* responsible for specifying cell fate within the forebrain, peripheral regions of the head, and in the fins. [96]

1.7. Area of Investigation/ Aims of This Research

Our long term goal is to understand the role of transcriptional regulators in the generation of embryonic patterning, and the mechanisms by which homeobox genes with similar homeodomains exert different roles during development. The following projects are centered on pursuing a greater understanding on the combinatorial function of the *distal-less* family of homeobox genes during embryonic patterning in the zebrafish. One section deals with the isolation of three new *dlx* genes *(dlx1, dlx6, and dlx7)* from the zebrafish, the next section investigates the putative *dlx* and *msx* gene regulatory

	Dlx EXPRESSION		
Ectodermal Srtipes in Gastrula	-	dlx3	-
Ventral Forebrain	dlx2	-	dlx4
Olfactory Placode	-	dlx3	dlx4
Migrating Neural Crest	dlx2	-	-
Pharyngeal Arches	dlx2	dlx3	dlx4
Otic Vesicle	-	dlx3	dlx4
Sensory Hair Cells	dlx2	-	dlx4
Pectoral Fin Buds	dlx2	dlx3	dlx4
Median Fin Fold	dlx2	dlx3	dlx4

Figure 1-13: Combinatorial embryonic *Dlx* expression in zebrafish.

Figure 1-13: Zebrafish *Distal-less* Gene Expression.

The following figure shows the embryonic expression patterns for the zebrafish *dlx* genes. [96] Their expression patterns will be discussed in more detail, in chapter 2.

regions, and the last section focuses on the function of *distal-less* genes during visceral arch development in the zebrafish.

The characterization of the zebrafish *dlx1* and *dlx7* genes is dealt with in chapter 2, which includes the isolation and embryonic expression patterns for both *dlx* genes. In order to understand how *distal-less* genes are combinatorially involved in the development of the forebrain, visceral arches, and limb buds, one must first identify all the possible genes from the zebrafish *dlx* family. Their embryonic expression patterns are important because they show the spatial and temporal location of the gene's transcripts which can be correlated with the expression of other *dlx* family members.

The isolation of *dlx* and *msx* genomic DNA clones and analysis of their 5' regulatory regions is dealt with in chapter 3. Regulatory regions are important to understand the mechanism by which these genes are controlled at the molecular level.

The forebrain model proposed by Boncinelli is based on mouse *distal-less* genes *(Dlx1, Dlx2, Dlx5,* and *Dlx6),* however, *distal-less* genes may be involved in the development of other parts of the embryo, for example, the visceral arches and limb buds. In zebrafish, *dlx3* is expressed in neuroectodermal cells prior to neural crest cell formation, and *dlx2* is found in neural crest cells migrating into the arches. Once crest cells have migrated to the arches, all other zebrafish *dlx* genes are turned-on in the arches, in a specific temporal and spatial combination. The mouse *dlx* genes have also been detected in mesenchymal chondrocyte condensations derived from the branchial arches, whereas, in the zebrafish no *dlx* genes have yet been detected in chondrocyte condensations. [96] More *dlx* genes from the zebrafish must be isolated and characterized in order to determine whether *dlx* genes play a combinatorial role in visceral arch development and differentiation.

The expression of *distal-less* genes during visceral arch development and differentiation is the focus of the fourth chapter. *Distal-less* expression was altered using a teratogen, RA, and correlated with craniofacial dysmorphologies. RA was chosen because we had previously found that it was able to downregulate *dlx* expression and also have dramatic effects on craniofacial development. The question that was addressed in this section was, whether or not, a loss of *dlx* expression in the visceral arches results in craniofacial dysmorphogenesis?

2. CHARACTERIZATION OF THREE NEW ZEBRAFISH DLX GENES

Chapter 2

2.1. Introduction

Three *distal-less genes, dlx2, dlx3* and *dlx4,* have previously been isolated from the zebrafish. [96] *Dlx3* is the first *distal-less* gene to be expressed in the zebrafish. It is expressed in two bilateral stripes, corresponding to ectodermal cells, along the anteroposterior axis of the gastrulating (6hpf) zebrafish embryo. Later during development (12hpf- 5 somites) these bilateral stripes converge towards the dorsal midline, but never reach the dorsal midline. Later during development these bilateral stripes become more dense in the frontonasal region & presumptive otic placode, and less dense in the median fin fold & in a thin lateral stripe running rostrally from the presumptive otic placode. This thin stripe correlates with the position of the presumptive neural crest cells. At the same developmental stage, 12hpf, *dlx2* is expressed in the presumptive ventral forebrain and in the neural crest cells migrating into the visceral arches. *Dlx4* is expressed beginning at 13hpf in the presumptive ventral forebrain and median fin fold. By 24hpf (28 somites) *dlx2* and *dlx4* are expressed in the ventral forebrain, *dlx3* and *dlx4* in the olfactory and otic placodes, and *dlx2, dlx3 & dlx4* are expressed in overlapping patterns in the median fin fold and visceral arches.

Orthologous groups for vertebrate *dlx* genes are divided upon overall protein sequence similarities. As seen in figure 1-12, five *distal-less* genes have been identified in the mouse, and only three *distal-less* genes had been previously isolated from zebrafish. [96] The three zebrafish *distal-less* genes were found to be members of the orthologous groups *Dlx2, Dlx3,* and *Dlx5.* Phylogenetic analysis showed that neither the *Dlx1* orthologous group branch, nor the *Dlx6* orthologous group contained zebrafish *distal-less* genes (figure 1-13), suggesting that there were *distal-less* genes missing in the zebrafish. *Distal-less* expression patterns are fairly conserved between species for a given orthologous group. The mouse *Dlx5* and *Dlx6* genes are known to be expressed in all chondrocyte condensations, [86] and the zebrafish *dlx2, dlx3,* and *dlx4* genes are not, [96] suggesting that the zebrafish *distal-less* family contains more than three genes.

It was previously suggested that *distal-less* genes act in a combinatorial manner, [96] because they appear to be expressed in overlapping regions during embryonic development. For example, the zebrafish's mandibular arch expresses a temporal combination of *dlx2, dlx3,* and *dlx4*. This combinatorial pattern of *distal-less* expression suggests that 1) *distal-less* genes may autoregulate each other, and/or 2) *distal-less* genes may control the fate of either the same cells or different co-existing cells. The isolation of all existing *distal-less* genes from zebrafish is important to understand the mechanism by which the combinatorial "*distal-less code*" functions during embryonic patterning.

2.2. Methods And Materials

2.2.1 Animals.

Wildtype zebrafish embryos were bought from a pet store and were maintained under conditions conforming to standard procedures. [101] Embryos were staged at 28.5°C according to hours (hpf) and days (dpf) post fertilization. [101]

2.2.2 Isolation Of Zebrafish Dlx Genes.

Dlx1 was isolated from a lambda Dash II genomic library (provided by M. Petkovich) using the zebrafish *dlx3* homeobox fragment as a probe. The *dlx3* homeobox fragment was labelled using the random priming protocol. The DNA fragment is denatured in the presence of random hexamers (pdN6) in solution for 3 minutes at 100°C. Radioactive dCTP (α-P^{32}), dNTP's (excluding dCTP), and Klenow Polymerase is added to the annealed primer-template mixture. The labelling reaction proceeds for 2 to 4 hours at room temperature. The reaction is stopped using 0.5M EDTA (pH 8). The incorporated radioactive dCTP is separated from the unincorporated dCTP by running the whole labelling reaction through a spin column. A spin column is made by placing glass wool in a 1ml syringe (to prevent the sephadex from leaking out), and then packed with sephadex G-25 to top up the syringe by centrifugation at a G-force (G) of 373. The column is then rinsed with TE (pH 8) two to three times. A solution of blue dextran (10mg/ml) to the labelling reaction, and the mixture is centrifuged

through the column for 2 minutes at 373G. The eluate should contain your incorporated (radioactive) dCTP probe.

The *dlx1* cDNA was isolated by PCR from a Lambda ZapII cDNA library (provided by D. Grunwald) prepared from 6-16hpf/20-28hpf zebrafish embryos. The PCR primers used to isolate the *dlx1* cDNA corresponded to sequences immediately upstream of the putative "ATG" start codon and immediately downstream from the predicted termination codon, as determined by the sequence analysis of the genomic clones. *Dlx7* was isolated by R. Breitbart and kindly given to us. The zebrafish *dlx6* gene was isolated by probing a zebrafish *dlx4* genomic clone with the zebrafish *dlx3* homeobox fragment. For more details on the molecular cloning of the dlx genes, please refer to the results section entitled "Molecular cloning of the zebrafish dlx genes".

Restriction fragments from *dlx1, dlx6* and *dlx7* were subcloned into the Bluescript KSII plasmid vector. DNA sequencing was performed according to the dideoxy-termination method using Sequenase (U.S. Biochemicals, Inc.). Briefly, the double stranded DNA template is denatured using an alkaline solution (0.5M NaOH). The denaturation is preformed at 37°C for 30 minutes, afterwards the reaction is neutralized with 3M sodium acetate (0.1X volume). The denatured DNA template is precipitated with an ethanol/glycogen mixture at -20°C. Sequencing is preformed using the Sequenase, T7 DNA Polymerase Kit. The denatured DNA template is annealed to a chosen primer (M13, T7, T3, Rev, ect.) at 65°C for 2 minutes. This reaction is then allowed to cool to room temperature for about 30 minutes, then it is chilled on ice. The annealed template-primer mixture is then ready for the labelling reaction. The sequenase enzyme (3 units), dNTP mixture (excluding the radioactive nucleotide), and the radioactive nucleotide (dATP-α-P^{32}) is added to the annealed template-primer mixture. This reaction is carried out at room temperature for 5 minutes. The labelling reaction is terminated by mixing the labelling reaction with the appropriate dideoxytermination mixture (either ddGTP, ddCTP, ddATP, or ddTTP). The termination reaction is carried out at 37°C for 5 minutes. Sequences were analyzed using the GCG package (Genetic Computer Group's Wisconsin Sequence Analysis Package) and the evolutionary tree was constructed using the Phylip package (Phylogeny Interface Package, provided by J. Felsenstein) according to sequence similarities from the GCG Gap program.

2.2.3 Whole mount in situ hybridization.

Zebrafish embryos were fixed in 4% paraformaldehyde (PFA) for 12h at 4°C. *In situ* hybridization's were performed on whole mount embryos using antisense riboprobes as described below. [102, 103] The following templates were used for synthesizing the antisense riboprobes: *dlx1*, a cDNA fragment of ~900bp that includes the coding region; *dlx7*, a 1Kb EcoRI fragment containing exons 1, 2, and part of exon 3.

An antisense RNA digoxigenin (DIG)-11-UTP (Boehringer Mannheim) probe was synthesized from a linearized cDNA template. Briefly, the linearized DNA template is incubated for 2 hours at 37°C with NTP's, DIG-11-UTP, and either T3 or T7 RNA polymerase. The DNA template is then removed from the mixture by adding DNase (RNase free). The DIG labelled fragment is then precipitated with ethanol/LiCl overnight at minus 20°C. The precipitated probe is then hydrolyzed with sodium bicarbonate/sodium carbonate at 60°C. The time of hydrolysis depends on the length of the labelled fragment.

$$\text{Time of hydrolysis} = \frac{\text{start length - desired length}}{0.11 \text{ X (start length X final length)}}$$

All steps below are preformed using RNase free solutions and at room temperature unless otherwise indicated. The fixed embryos are rehydrated in serial baths of methanol/RNase Free water. Once rehydrated, use 1/6 of the embryos for preabsorption. Place embryos in a 1.5ml eppendorf with the anti-DIG antibody conjugated to alkaline phosphatase diluted 1/1000 in PBST, 2% calf serum, 2mg/ml Bovine serum albumin. The preabsoption step is carried out at room temperature for 1 hour then overnight at 4°C.

The *in situ* embryos are permeabilized with proteinase K (10ug/ml in PBS-tween (PBST)) for either 2 minutes (embryos younger than 27 hpf) or for 5 minutes (embryos older than 27 hpf). The embryos are then postfixed for 20 minutes in 4% PFA, and rinsed two times in PBST. To minimize background, an acetylation step is preformed for 10 minutes. The acetylation mixture contains triethanolamine, acetic anhydride, and Rnase-free (RF) water. The embryos are rinsed in PBST for 10 minutes two times.

Embryos are then prehybridized for 1 hour at 65°C in 50% formamide, 5X SSC, 0.1% Tween-20, 50ug/ml heparin, 10ug/ml yeast tRNA, water and citric acid to pH 6. The hybridization step is carried out at 65°C, overnight, in the same solution supplemented with 100ng (0.5ng/ul) of labelled RNA probe. The hybridized embryos are then washed at 65°C for 10 minutes in sequential baths containing hybridization solution and increasing percentages of 2XSSC, until the embryos are in a solution of 2XSSC. Then they are washed two times for 30 minutes each in 0.2X SSC at 60°C. The embryos are then brought to room temperature and put through sequential baths containing 0.2XSSC and increasing percentages of PBST, until they are immersed in PBST.

The hybridized embryos are preincubated for 1 to 4 hours in PBST, 2% calf serum, and 2mg/ml bovine serum albumin. After the preincubation, the embryos are incubated for 3 to 4 hours with 1/5000 dilution of the preabsorbed antibody. This reaction allows the anti-DIG antibody to bind to the DIG labelled probe. The embryos are washed in PBST, and again preincubated in a prestain buffer {100mM TrisHCl pH 9.5; 50mM MgCl2; 100mM NaCl; 0.1% Tween-20; 1mM lavamisol} for 5 minutes. The embryos are then stained using the substrates 5-Bromo-4-chloro-3-indolyl-phosphate (BCIP) and 4-Nitrophenyl phosphate (NBT). A chromogenic precipitate forms when the stain reacts with the alkaline phosphatase conjugated to the anti-DIG antibody, leaving a blue-purple color precipitate. A 2 hr post fixation in 4% paraformadehyde followed by a wash with PBS is carried out. The embryos are then stored at 4°C in PBS/ 5mM sodium azide.

2.3. Results

2.3.1 Molecular Cloning Of The Zebrafish Dlx Genes.

It was previously reported that the mouse *dlx1* and *dlx2* genes are linked in an inverted fashion, ie. transcribed towards each other, with a intergenic region of about 8Kb. [86] A lambda Dash II genomic zebrafish library was screened at high stringency with a probe containing the entire *dlx2* cDNA. One clone, *dlx2:7*, was isolated and found to contain only exons 2 and 3 of *dlx2* and 12Kb flanking 3' to the *dlx2* gene. A southern analysis using a *dlx3* homeobox probe revealed that the *dlx2:7* clone contained two highly homologous *distal-less*

homeobox sequences. The hybridized bands were then subcloned into the Bluescript KS II plasmid vector. Another southern analysis, using a probe containing the entire mouse *dlx1* gene, on the *dlx2* genomic subclones showed that one of these homeobox containing bands was highly homologous to the mouse *dlx1* gene. Further sequencing identified this band as the zebrafish *dlx1* gene. Two oligonucleotide primers were synthesized that correspond to sequences outside the start codon and stop codon. PCR was performed on DNA prepared from a lambda Zap II cDNA library of total 20-28hpf zebrafish embryo mRNA (provided by D. Grunwald). A 900bp fragment corresponding to part of the *dlx1* cDNA was isolated.

Dlx7 was isolated by R. Breitbart (whom generously sent it to us) from a 30hpf zebrafish lambda gt-11 cDNA library using the *Nkx2.5* cardiac homeobox gene as a probe. [104] Sequencing of the 1.3Kb cDNA clone identified it as a member of the *distal-less* gene family. [104]

The mouse *Dlx5* and *Dlx6* genes are also linked in an inverted fashion, like *Dlx1* and *Dlx2*. The ortholog to the mouse *Dlx5* is the zebrafish *dlx4*, personal communication from J. Rubenstein. A zebrafish *dlx4* genomic clone was previously isolated from a lambda DashII library using a probe which contained the entire length of the zebrafish *dlx4* cDNA. This genomic clone was identified as having another homeodomain which was shown to be similar to that of the *dlx* gene family. Sequence analysis confirmed this second homedomain to be contained within the zebrafish *dlx6* gene.

The amplified cDNA for *Dlx1* has length of 834bp, whereas that of the *dlx7* cDNA is 1130bp. The predicted protein lengths would be 252AA for *dlx1*, and 258AA for *dlx7* (Figures 2-1, and 2-2).

2.3.2 Phylogenetic Analysis.

The protein encoded by the zebrafish *dlx1* cDNA shows the highest percentage of similarity at the amino acid level with the mouse *dlx1* gene (84%). The *Xdll* (*Dlx1* ortholog) has 75% similarity, *NvHBox5* (*Dlx6* ortholog) has 67% similarity, and the rest of the *distal-less* family members have less than 60% similarity. The *dlx1* homeodomain also shows highest percent similarity with the mouse *dlx1* (98%), 60 of the 61 amino acids are identical. *Dlx7* on the other hand shows the highest percentage of similarity to *NvHBox5* (75%). *Xdll* shows

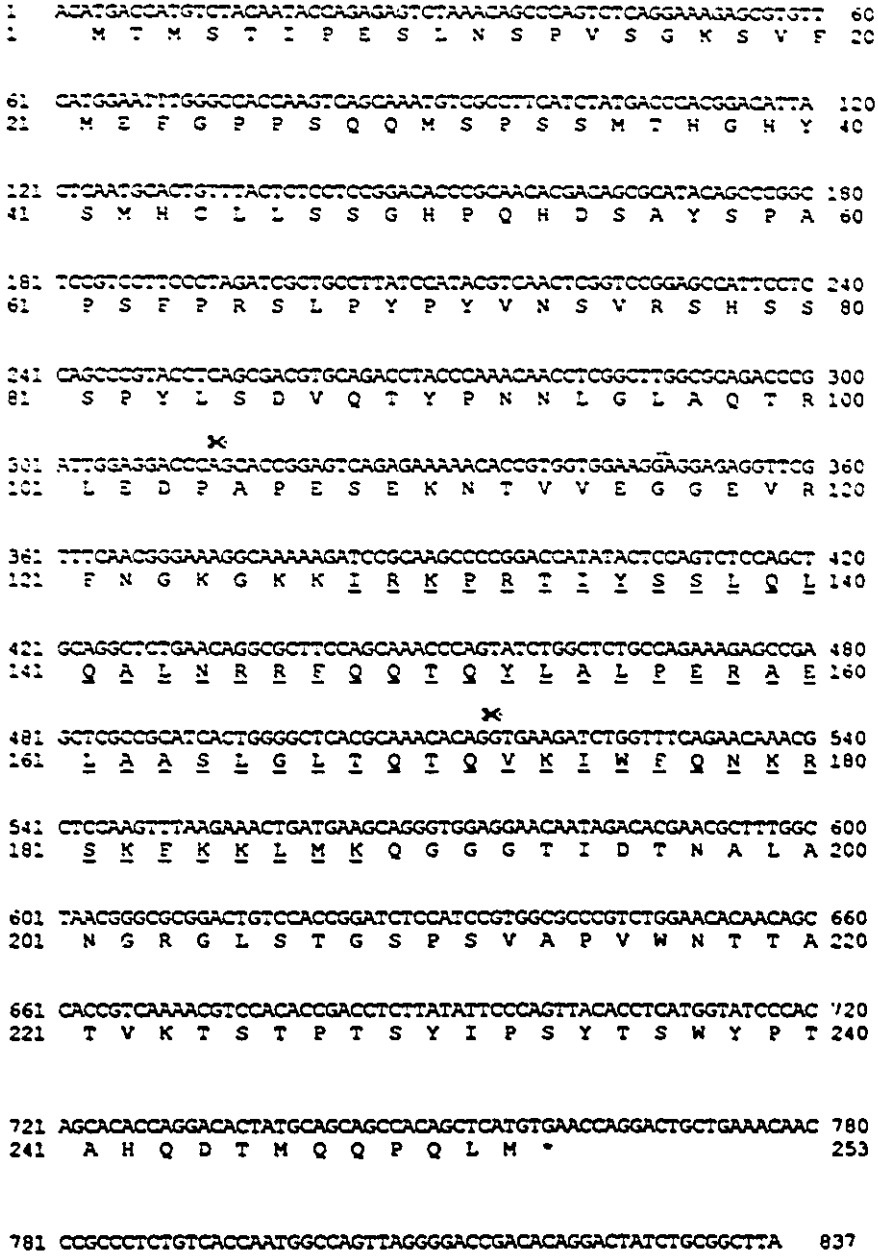

1 ACATGACCATGTCTACAATACCAGAGAGTCTAAACAGCCCAGTCTCAGGAAAGAGCGTGTT 60
1 M T M S T I P E S L N S P V S G K S V F 20
61 CATGGAAATTTGGGCCACCAAGTCAGCAAATGTCGCCTTCATCTATGACCCACGGACATTA 120
21 M E F G P P S Q Q M S P S S M T H G H Y 40
121 CTCAATGCACTGTTTACTCTCCTCCGGACACCCGCAACACGACAGCGCATACAGCCCGGC 180
41 S M H C L L S S G H P Q H D S A Y S P A 60
181 TCCGTCCTTCCCTAGATCGCTGCCTTATCCATACGTCAACTCGGTCCGGAGCCATTCCTC 240
61 P S F P R S L P Y P Y V N S V R S H S S 80
241 CAGCCCGTACCTCAGCGACGTGCAGACCTACCCAAACAACCTCGGCTTGGCGCAGACCCG 300
81 S P Y L S D V Q T Y P N N L G L A Q T R 100
301 ATTGGAGGACCCAGCACCGGAGTCAGAGAAAAACACCGTGGTGGAAGGAGGAGAGGTTCG 360
101 L E D P A P E S E K N T V V E G G E V R 120
361 TTTCAACGGGAAAGGCAAAAAGATCCGCAAGCCCCGGACCATATACTCCAGTCTCCAGCT 420
121 F N G K G K K I R K P R T I Y S S L Q L 140
421 GCAGGCTCTGAACAGGCGCTTCCAGCAAACCCAGTATCTGGCTCTGCCAGAAAGAGCCGA 480
141 Q A L N R R F Q Q T Q Y L A L P E R A E 160
481 GCTCGCCGCATCACTGGGGCTCACGCAAACACAGGTGAAGATCTGGTTTCAGAACAAACG 540
161 L A A S L G L T Q T Q V K I W F Q N K R 180
541 CTCCAAGTTTAAGAAACTGATGAAGCAGGGTGGAGGAACAATAGACACGAACGCTTTGGC 600
181 S K F K K L M K Q G G G T I D T N A L A 200
601 TAACGGGCGCGGACTGTCCACCGGATCTCCATCCGTGGCGCCCGTCTGGAACACAACAGC 660
201 N G R G L S T G S P S V A P V W N T T A 220
661 CACCGTCAAAACGTCCACACCGACCTCTTATATTCCCAGTTACACCTCATGGTATCCCAC 720
221 T V K T S T P T S Y I P S Y T S W Y P T 240
721 AGCACACCAGGACACTATGCAGCAGCCACAGCTCATGTGAACCAGGACTGCTGAAACAAC 780
241 A H Q D T M Q Q P Q L M * 253
781 CCGCCCTCTGTCACCAATGGCCAGTTAGGGGACCGACACAGGACTATCTGCGGCTTA 837

Figure 2-1: Partial cDNA sequence of the zebrafish *Dlx1* gene and its predicted protein sequence. The homeodomain is underlined and the star represents a stop codon. The splice sites between the exons are indicated by scissors.

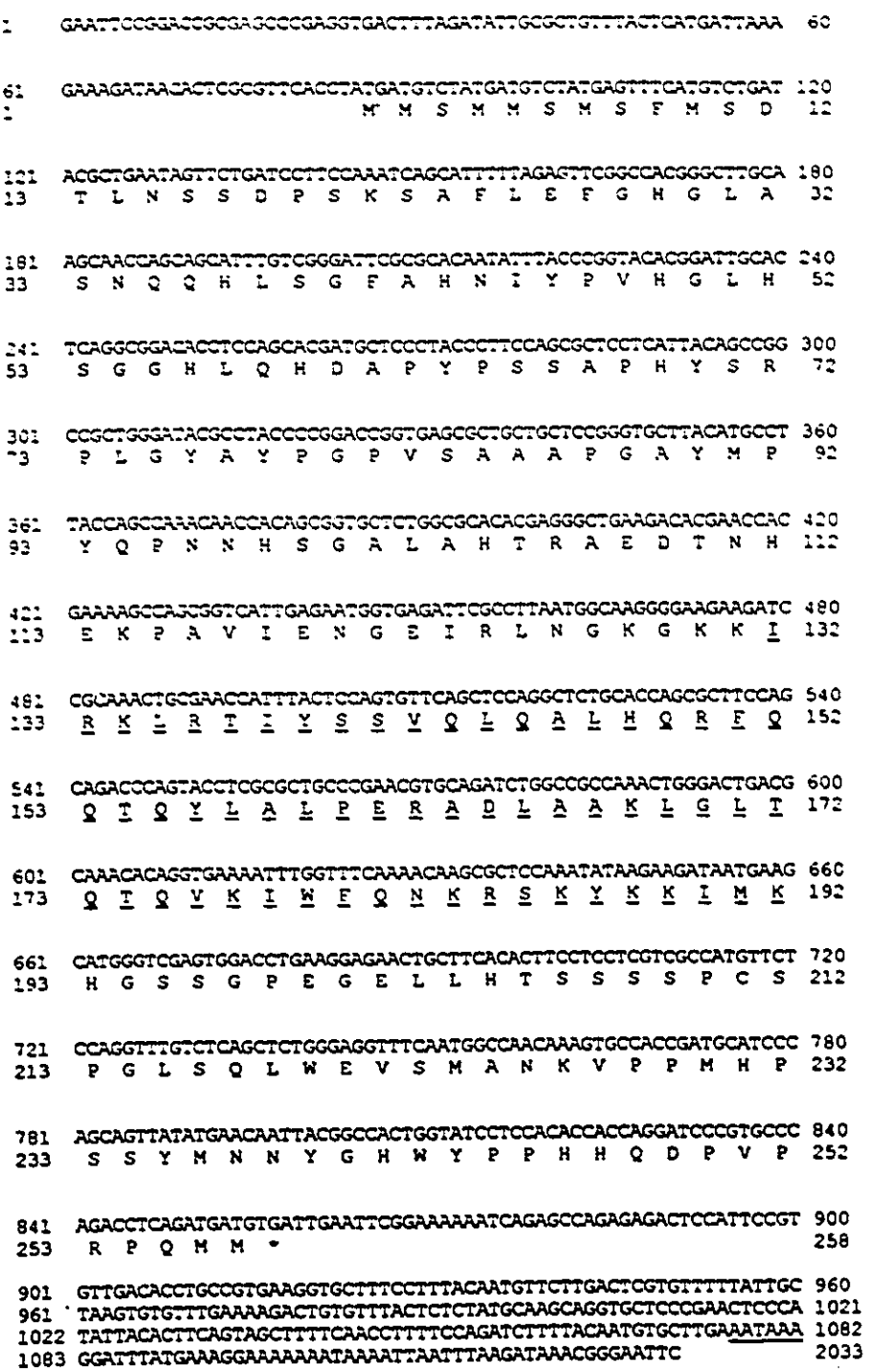
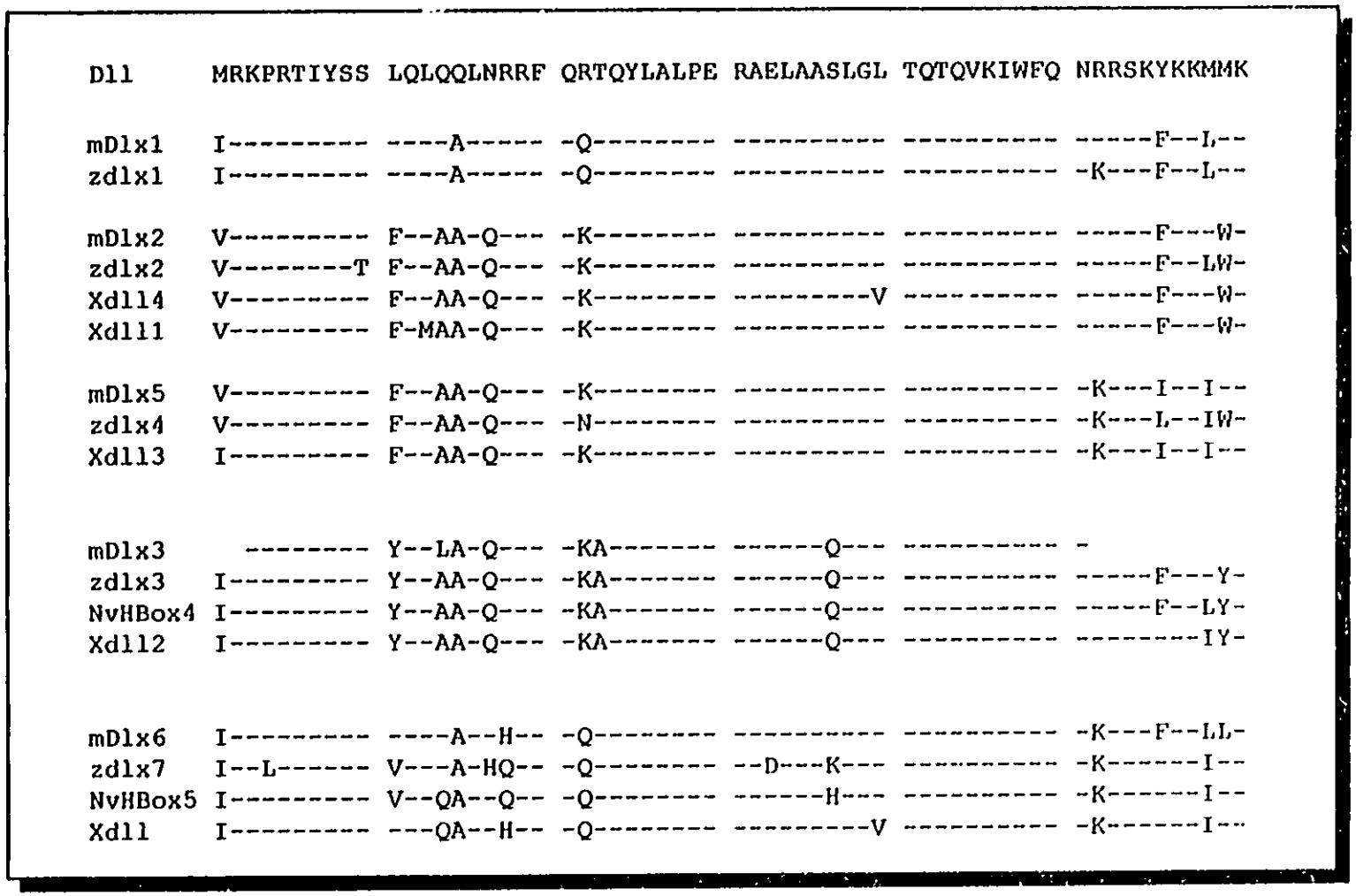


Figure 2-2: cDNA sequence for the zebrafish *Dlx7* gene and its predicted protein sequence. A potential polyadenylation signal is underlined. The homeodomain's amino acids are underlined, and the star corresponds to the stop codon.



```
Dll      MRKPRTIYSS LQLQQLNRRF QRTQYLALPE RAELAASLGL TQTQVKIWFQ NRRSKYKKMMK

mDlx1    I--------- ----A----- -Q-------- ---------- ---------- -----F--L--
zdlx1    I--------- ----A----- -Q-------- ---------- ---------- -K---F--L--

mDlx2    V--------- F--AA-Q--- -K-------- ---------- ---------- -----F---W-
zdlx2    V--------T F--AA-Q--- -K-------- ---------- ---------- -----F--LW-
Xdll4    V--------- F--AA-Q--- -K--------- ---------V ---------- -----F---W-
Xdlll    V--------- F-MAA-Q--- -K-------- ---------- ---------- -----F---W-

mDlx5    V--------- F--AA-Q--- -K-------- ---------- ---------- -K---I--I--
zdlx4    V--------- F--AA-Q--- -N-------- ---------- ---------- -K---L--IW-
Xdll3    I--------- F--AA-Q--- -K-------- ---------- ---------- -K---I--I--


mDlx3      -------- Y--LA-Q--- -KA------- ------Q--- ---------- -
zdlx3    I--------- Y--AA-Q--- -KA------- ------Q--- ---------- -----F---Y-
NvHBox4  I--------- Y--AA-Q--- -KA------- ------Q--- ---------- -----F--LY-
Xdll2    I--------- Y--AA-Q--- -KA------- ------Q--- ---------- --------IY-


mDlx6    I--------- ----A--H-- -Q-------- ---------- ---------- -K---F--LL-
zdlx7    I--L------ V---A-HQ-- -Q-------- --D---K--- ---------- -K------I--
NvHBox5  I--------- V--QA--Q-- -Q-------- ------H--- ---------- -K------I--
Xdll     I--------- ---QA--H-- -Q-------- ---------V ---------- -K------I--
```

Figure 2-3: *Distal-less* homeodomain pileup.

Figure 2-3: Multiple *Distal-Less* Homeodomain Alignment.
The *distal-less* homeodomains, from each specie, are aligned to the *Drosophila distal-less* homeodomain. The homeodomains are divided into the six *Dlx* orthologous groups. *Dll* refers to the *Drosophila distal-less* gene, *mDlx* refers to the mouse *Dlx*, *zdlx* refers to the zebrafish *dlx*, *Xdll* refers to *Xenopus dlx*, and *NvHbox5/NvHBox4* refers to the newt *dlx*. Conserved amino acids are shown by a dash.

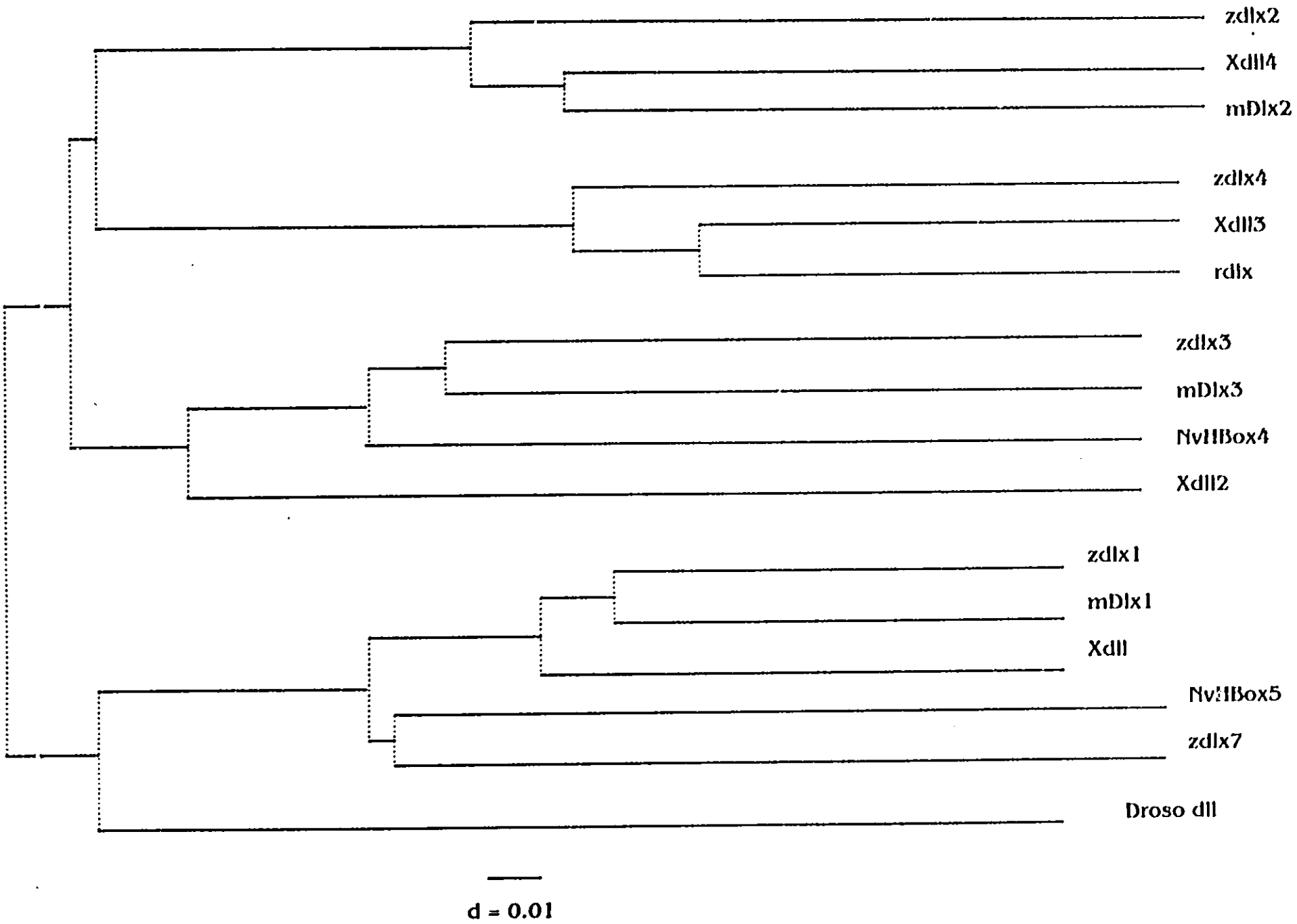

zdlx2
Xdll4
mDlx2
zdlx4
Xdll3
rdlx
zdlx3
mDlx3
NvHBox4
Xdll2
zdlx1
mDlx1
Xdll
NvHBox5
zdlx7
Droso dll
d = 0.01

Figure 2-4: Phylogenetic Tree Showing The New *Distal-Less* Gene Family Members.

The evolutionary distances are indicated by the scale beneath the tree. The tree was constructed with complete protein sequences using the Neighbour-Joining method. Complete amino acid sequences are not yet availible for the mouse *Dlx5* and *Dlx6*.

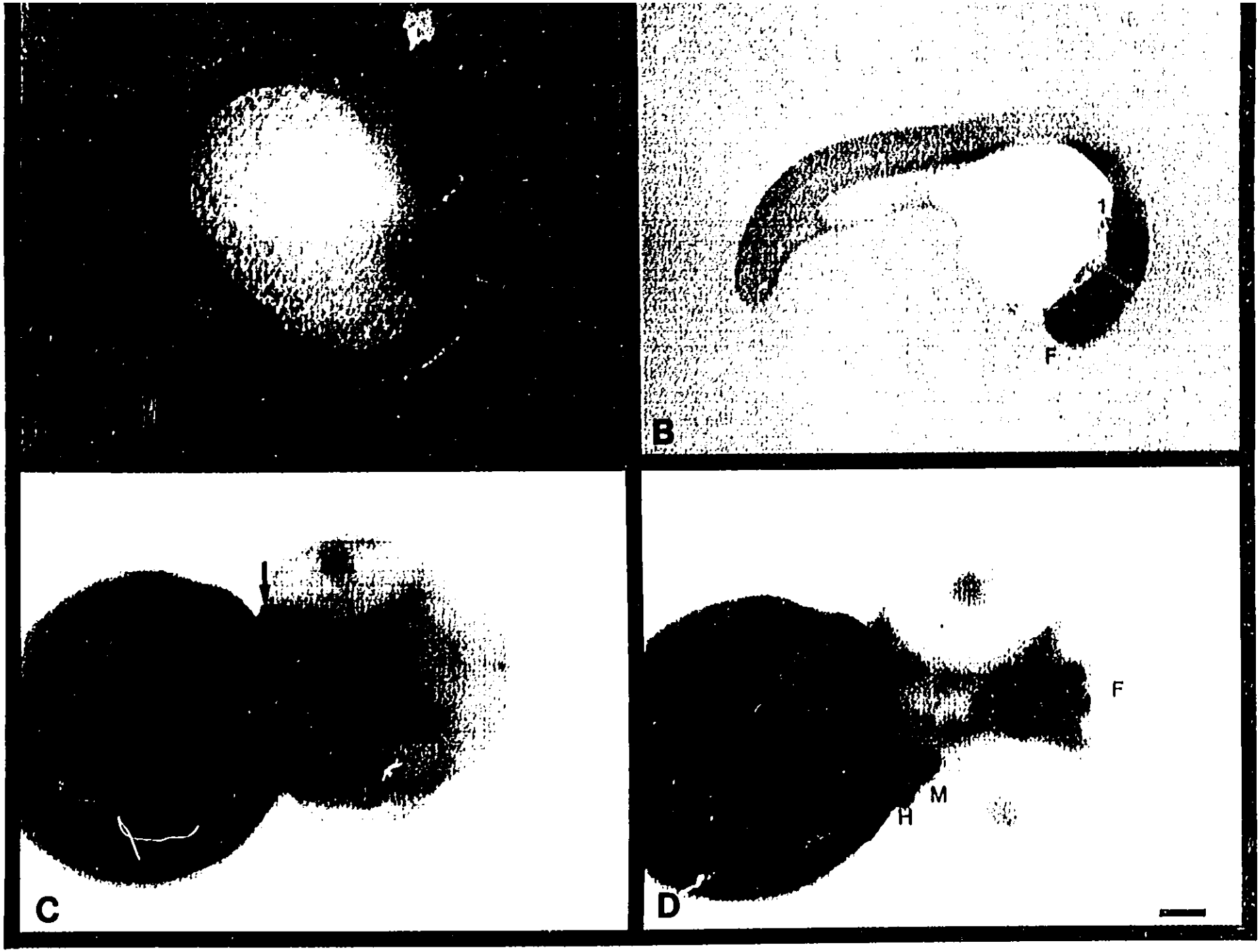
70%, and the mouse *dlx1* shows 68% similarity to the zebrafish *dlx7*. The rest of the *distal-less* family shows less than 60% similarity to the zebrafish *dlx7*. Within the homeodomain, 56 of the 61 amino acids are identical between the *NvHBox5* and zebrafish *dlx7* (92%). See figure 2-3 for the *distal-less* family homeodomain alignment.

A phylogenetic tree was constructed from the predicted *distal-less* protein sequences using the Neighbor Joining method (figure 2-4). The percent similarity between each sequence was calculated with the Gap program, the distances were then calculated using the formula Distance = 100 - % similarity. These distances were entered into a distance matrix in the Phylip phylogenetic program.

2.3.3 Embryonic Expression Patterns Of The Zebrafish Dlx1 And Dlx7 Genes.

Dlx1 expression is not detectable prior to 17hpf. The transcripts appear at 17hpf in the ventral forebrain unlike *dlx2* and *dlx4* which are expressed in the presumptive ventral forebrain at 13hpf. [96] At 22hpf *dlx1* transcripts are detected at a lower intensity in the lateral component of branchial arch 1, and *dlx1* is also expressed with a higher intensity in the ventral forebrain. At 2dpf *dlx1* is expressed in the ventral forebrain, and in the distal portion of the pectoral fin buds (apical ectodermal ridge). There also seems to be expression in the areas that later condensate to form the trabeculae (cartilage derived from mid-forebrain crest cells), and visceral arch cartilage derivatives. At 3dpf, *dlx1* expression in the ventral forebrain is still very intense (figure 2-5). 5dpf and 6dpf embryos do not show any detectable *dlx1* expression.

Dlx7 expression during epiboly (8hpf) is very faint, but resembles the expression pattern of the *dlx3* gene. [96] *Dlx7* hybridization signals become stronger from 12hpf to 16hpf in the frontonasal area and presumptive otic placode area. Expression in the posterior edge of the median fin fold is also apparent. *Dlx7* expression at 19hpf is the same as for 16hpf except that the frontonasal expression becomes restricted to the presumptive olfactory placodes and expression moves anteriorly along the edges of the fin fold. At 19hpf, the visceral arch primordia express *dlx7* in the medial and lateral components of the mandibular and hyoid arches, and in branchial arch 2. A striking aspect of *dlx7* expression, is that it is found in the same cells that are expressing *dlx3*. By 48hpf *dlx7* expression is undetectable (figure 2-5).


B
1
F
C
D
F
M
H

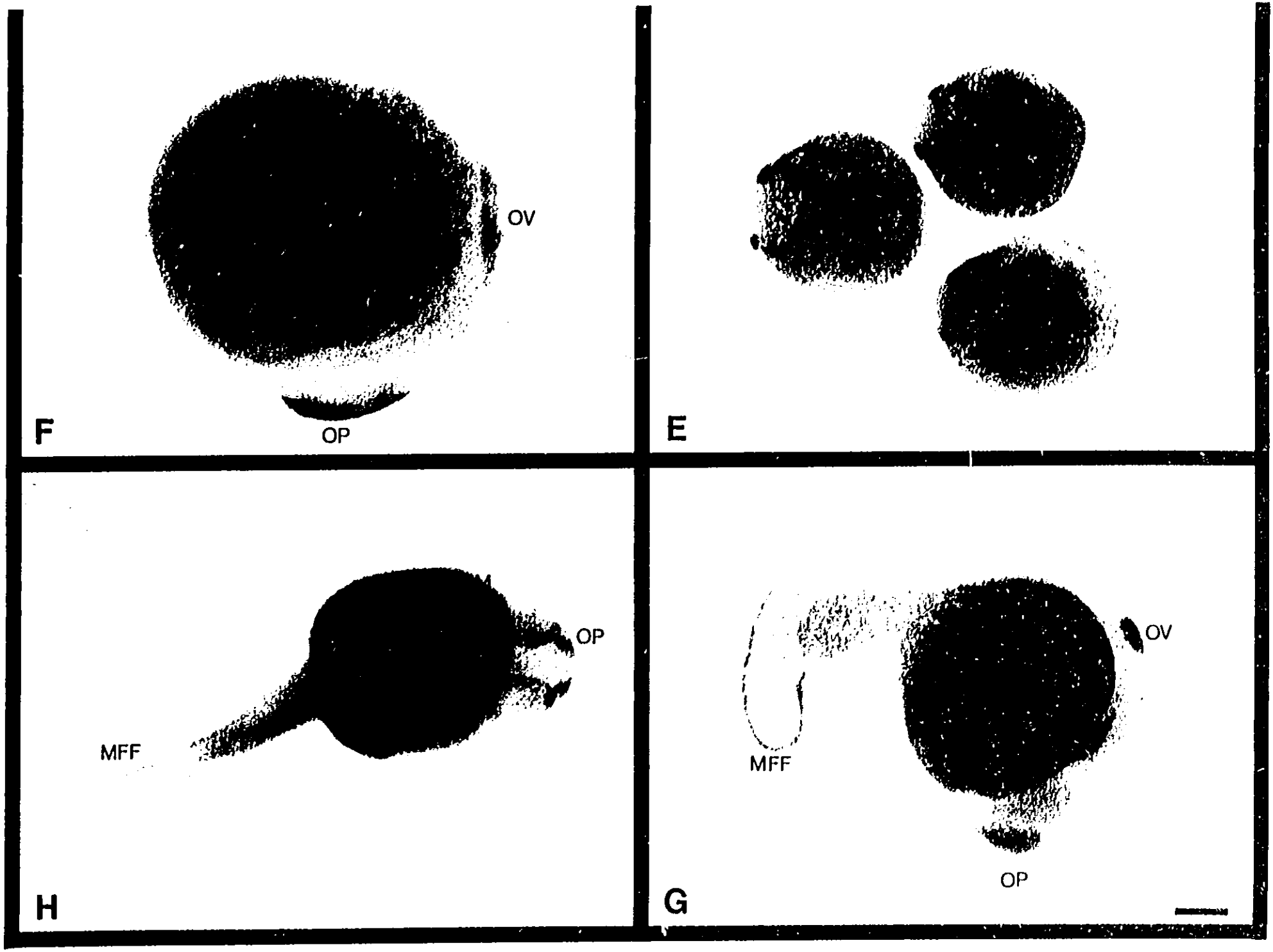

F
OV
OP
E
H
MFF
OP
G
MFF
OV
OP

Figure 2-5: *Dlx1* And *Dlx7* Expression Patterns.

The left page shows *dlx1* expression during zebrafish embryo development. A) 17hpf, *dlx1* expression in the presumptive ventral forebrain, "F". B) 22hpf, *dlx1* expression in the ventral forebrain and branchial arch 1, "1". C) 48hpf ventral view, *dlx1* expression in the ventral forebrain, visceral cleft (arrow). D) 73hpf (3dpf) ventral view, *dlx1* expression in the ventral forebrain and visceral clefts, "M" (mandibular) & "H" (hyoid). The right page shows *dlx7* expression patterns in zebrafish embryos. E) *dlx7* expression during epiboly (arrow heads), 8hpf. F) *dlx7* expression is confined to the frontonasal mass ("OP") and presumptive otic placode ("OV") at 12hpf. G) at 19hpf, *dlx7* is expressed in the median fin fold ("MFF"), frontonasal mass ("OP"), otic placode ("OV"), and visceral arch primordia ("VA"). H) *dlx7* expression at 24hpf is found in the olfactory placodes ("OP"), otic placodes ("OV"), and median fin fold ("MFF"). Forebrain, F; branchial arch 1, 1; visceral cleft, VC; frontonasal mass, OP; median fin fold, MFF; and otic placode, OV. Scale bar: 96um for A; 89um for B; 173um for C; 141 um for D; 80um for E; 119um for F; 118um for G; and 83um for H.

2.4. Discussion

2.4.1 Zebrafish Dlx Expression

The first zebrafish *distal-less* genes to be expressed are *dlx3* and *dlx7*, in ectodermal cells. The same cells seem to express both *dlx7* and *dlx3*. Their expression begins in two bilateral stripes running along the anterior-posterior axis during gastrulation (6hpf). The difference between *dlx3* and *dlx7* can be seen at 48hpf, where *dlx3* expression continues, and *dlx7* expression is undetectable.

The Dlx3 orthologous group is the only *Dlx* group to date which is not expressed in the forebrain of developing vertebrate embryos. The zebrafish *dlx7* is also not expressed in the developing forebrain, it is therefore very likely that members of the *dlx7* and the *dlx3* orthologous groups are related. These two genes would not be related by sequence but rather by their function. However, it is unknown whether or not the two genes are linked.

The onset of *dlx1* expression is found at 17hpf, where it is expressed in the presumptive forebrain. *Dlx2* on the other hand, has an onset of expression at 12hpf. A time lag between the onset of *dlx1* and *dlx2* is also found in the mouse. The time lag in the mouse is 2dpc, whereas in the zebrafish it is 5hpf. Both of these genes are expressed in the same cells until 48hpf, except that *dlx1* is not expressed in migrating crest cells. At 48hpf, *dlx1* and *dlx2* are expressed in neural crest derived visceral arch components. *Dlx2* expression is undetectable at 3dpc, whereas *dlx1* continues to be expressed past 5dpc.

2.4.2 Six Orthologous Dlx Groups

Members of the *Dlx1* and *Dlx2* orthologous groups are expressed in similar areas, but the onset of *Dlx1* expression lags behind that of its *Dlx2* ortholog. In the mouse, the onset of *Dlx1* expression lags behind the onset of *Dlx2* expression by 2dpc, and in the zebrafish by 5hpf. These genes are expressed in the forebrain, visceral arches, and apical ectodermal ridge (AER) of the limb bud. The zebrafish and mouse orthologs, *Dlx1 and Dlx2*, are expressed in areas which resembles chondrocyte condensations which will derive from the visceral arch neural crest cells.

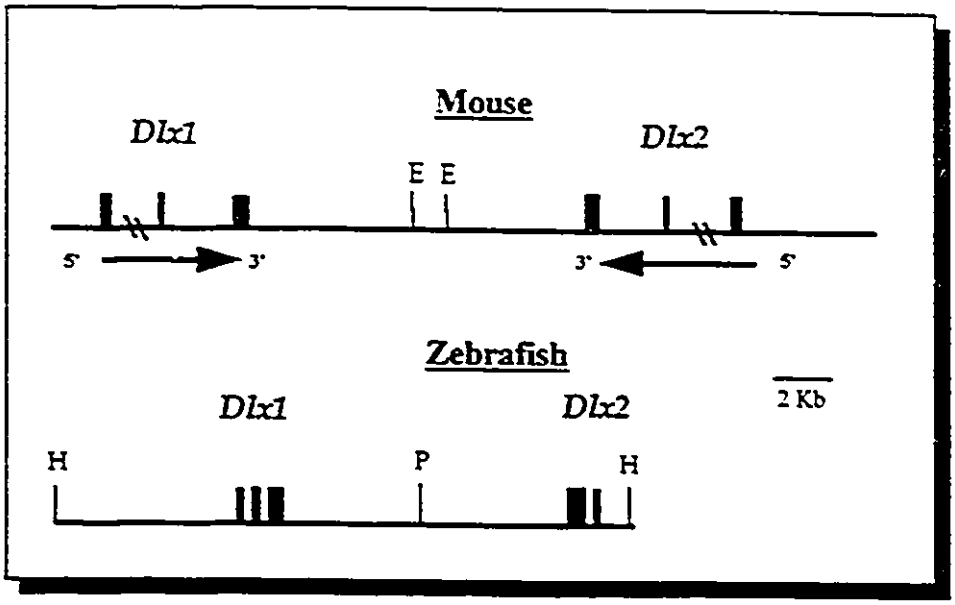
A third *Dlx* orthologous group, *Dlx3*, has representatives from four species. This group of genes is unique among vertebrate *dlx* genes because they are not expressed in the forebrain. Rather, they are expressed in the visceral arches, otic vesicle, olfactory placode, in the limb bud's AER and median fin fold. The mouse *dlx3* was also recently shown to be expressed in chondrocyte condensations derived from the branchial arches. [87] Similarily, as will be shown in the next chapter, the zebrafish *dlx3* gene is also expressed in areas which coincide with neural crest derived chondrocyte condensations.

The fourth *Dlx* orthologous group (Dlx5) is made up of four genes, and their transcripts are found in the forebrain, branchial arches, olfactory placodes, otic vesicles, and the limb bud's AER. The rat *rdlx* and mouse *Dlx5* transcripts are also found in developing cartilage, perichondria of mature cartilage, and mesenchymal cells of developing membranous bones. [86, 91] The zebrafish dlx4 ortholog is also found in areas which coincide with visceral arch neural crest derived cartilage condensations. The mouse *Dlx5* gene is expressed in the same areas as the mouse *Dlx6* gene, however its expression seems to be stronger than that of its *Dlx6* ortholog.

The sixth *Dlx* orthologous group, *Dlx6*, is made up of three genes; mouse *dlx6*, *Xenopus Xdll*, and zebrafish *dlx6*. The group's expression patterns have not all been characterized. The mouse *Dlx6* gene is expressed in the otic, olfactory placodes, forebrain, visceral arches, and all developing cartilage. The zebrafish and *Xenopus* genes have not yet been fully analyzed by whole mount *in situ* hybridization. Northern analysis has shown that the *Xenopus* gene is predominately expressed in oocytes.

The zebrafish *dlx7* gene is found in another orthologous group, Dlx7. The zebrafish *dlx7* protein sequence has a high degree of similarity with the newt *NvHBox5* protein sequence. The zebrafish *dlx7* gene is not expressed in the forebrain and developing cartilage, rather, its expression pattern (up to 48hpf) is identical to that of the zebrafish *dlx3* gene. Northern analysis has shown that the newt *NvHBox5* gene is expressed predominantly in the brain and limb bud's AER.

The expression patterns for each *Dlx* orthologous group (except for the *Dlx6* & *Dlx7* orthologous groups) are very well conserved between species. It is possible that the conserved expression patterns are due to the evolutionary



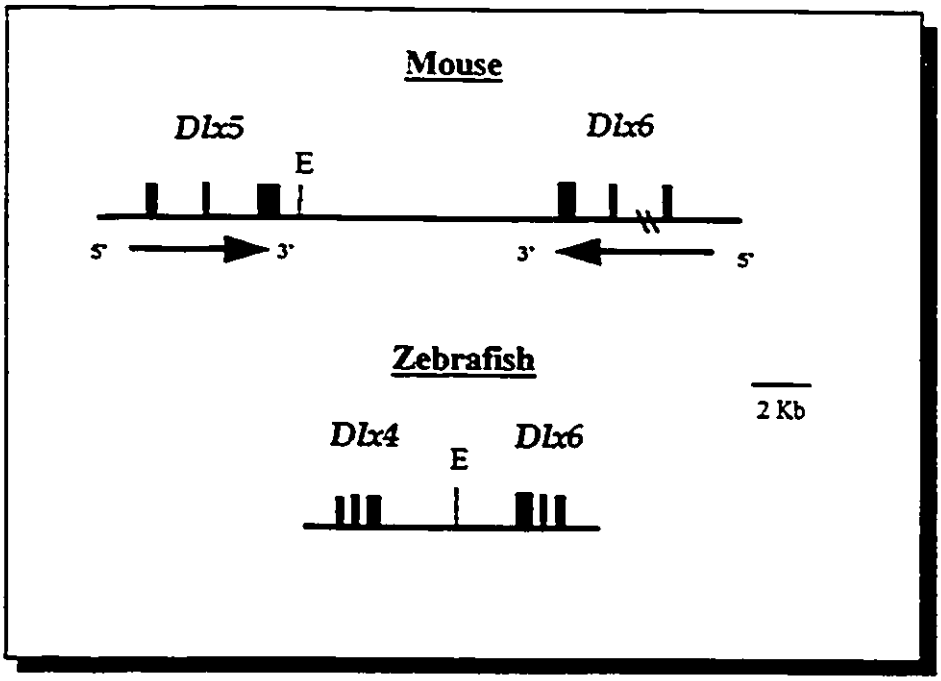


Figure 2-6: *Dlx*-1/ *Dlx*-2 and *Dlx5/Dlx6* Loci.

Figure 2-6: Genomic Organization Of The Mouse And Zebrafish *Dlx1Dlx2* & *Dlx5/Dlx6* Loci.

The mouse *Dlx1/Dlx2 & Dlx5/Dlx6* loci are shown above that of the zebrafish. The exons are depicted by solid black boxes, endonuclease sites (letters indicate which endonuclease) are shown by perpendicular lines. E, *EcoRI;* H, *HindIII;* and P, *PstI.* The arrows under the mouse *Dlx* genes shows the direction of transcription, which is conserved with the zebrafish.

sequence conservation within one orthologous group. This sequence conservation would also include the preservation of regulatory elements necessary to maintain their tissue-specific expression.

2.4.3 Dlx1/Dlx2 and Dlx5/Dlx6 Genomic Organization

There are six *distal-less* genes to date that have been isolated in the zebrafish, unlike the mouse which only has five *dlx* genes isolated. The zebrafish *dlx6* gene has not been fully isolated and characterized thus far. However, the evolutionary conservation of gene arrangment has been maintained between zebrafish and mice.

The mouse *Dlx1/Dlx2* and *Dlx5/Dlx6* genes were previously shown to be linked in an inverted fashion; transcribed towards each other. [86] This genomic organization was found to be conserved in the zebrafish (figure 2-6). The zebrafish *dlx1* is linked to *dlx2*, and *dlx4* (Dlx5 ortholog) is linked to *dlx6*. Another interesting feature of the *Dlx1* and *Dlx2* genes, is that they are expressed in similar locations with the onset of *Dlx1* lagging behind that of *Dlx2*, by 2dpc in mouse and 5hpf in the zebrafish. One would then expect that the regulatory elements necessary for their tissue-specific expression patterns to also be conserved and located between the two genes, ie. found 3' to both genes. The mouse and zebrafish *Dlx1/Dlx2* loci are very different in size, however their intergenic region is fairly conserved in size, ~8Kb. It is therefore very likely that the conserved *Dlx1/Dlx2* intergenic region contains some conserved regulatory elements necessary for their tissue specificity.

2.4.4 Phylogenetic Analysis of Vertebrate Distal-less Genes

The exact mode by which the *distal-less* genes evolved is still uncertain. However, one part of the evolutionary process that is adhered to is that vertebrate *distal-less* genes probably arose from an ancestor containing only one *distal-less* gene.

This ancestral *distal-less* gene could have evolved either by simple duplication or by duplication and inversion events. The former would create two *distal-less* genes of the same orientation, whereas the latter would create two *distal-less* genes, of which one would be in an inverted orientation (like the *Dlx1* and *Dlx2* locus).

The most probable path of evolution for the vertebrate *dlx* genes would be the inversion-duplication event. The common ancesteral-*Dll* gene would have duplicated and inverted to create a locus similar to the mouse *Dlx1/Dlx2* because these two genes are linked in an inverted fashion.

The human and mouse *Dlx1* and *Dlx2* genes are found in close proximity to the *HoxD* cluster on chromosome 2, and the human *Dlx5* and *Dlx6* genes are located on the same chromosome as the human *HoxA* cluster. However, in *Drosophila*, the *Antp/Ubx* cluster is found on chromosome 3 and *Dll* is found on chromosome 2. It is possible that during the course of evolution, the vertebrate ancestral *Dll*-like gene translocated to the same chromosome which contains the ancestral *Hom-C* -like cluster, or vise versa. Therefore, a vertebrate *Dlx* common ancester would have a single chromosome with the *Dlx* genes linked to the *Hox* cluster. This single ancesteral vertebrate chromosome would have been the common ancester to the known vertebrate *Hox/Dlx* chromosomal arrangment. This ancesteral vertebrate chromosome could have further evolved by duplications. Using this hypothesis, the *Hox* cluster's duplications would have duplicated the linked *Dlx1/Dlx2* loci to create a linked *Dlx5/Dlx6*. This however, is uncertain, because the *HoxA* cluster is not in close proximity to *Dlx5/Dlx6* loci like that of the *HoxD* cluster to the *Dlx1/Dlx2* loci. However, It is possible that both the *HoxD/Dlx1&Dlx2* and the *HoxA/Dlx5&Dlx6* loci evolved from a common ancester, and then an insertion event took place between the *HoxA* cluster and the *Dlx5&Dlx6* locus. This insertion event would place the *Dlx5&Dlx6* locus further from the *HoxA* cluster. It is also possible that the the common ancester was more similar to the *HoxA/Dlx5&Dlx6* genomic arrangment, and the *HoxD/Dlx1&Dlx2* loci underwent a deletion event. Another possibility is that these two events are independent.

The most probable evolutionary path followed is that the ancesteral-like *Dll* gene underwent an inversion-duplication event, thus creating two vertebrate *Dlx* genes which are inverted. The inverted vertebrate *Dlx* genes would have duplicated to produce the known *Dlx* loci, *Dlx1/Dlx2* & *Dlx5/Dlx6*. The unlinked *Dlx3* gene would have evolved from a duplication of a single *Dlx* gene. The single *Dlx* gene that would have undergone a duplication event, would have to be either *Dlx2* or *Dlx5*, because the *Dlx2/Dlx5* evolutionary branch distance is closer (d=0.35) to that of *Dlx3* than the *Dlx1/Dlx6* branch (d=0.44). (See Figure 2-7)

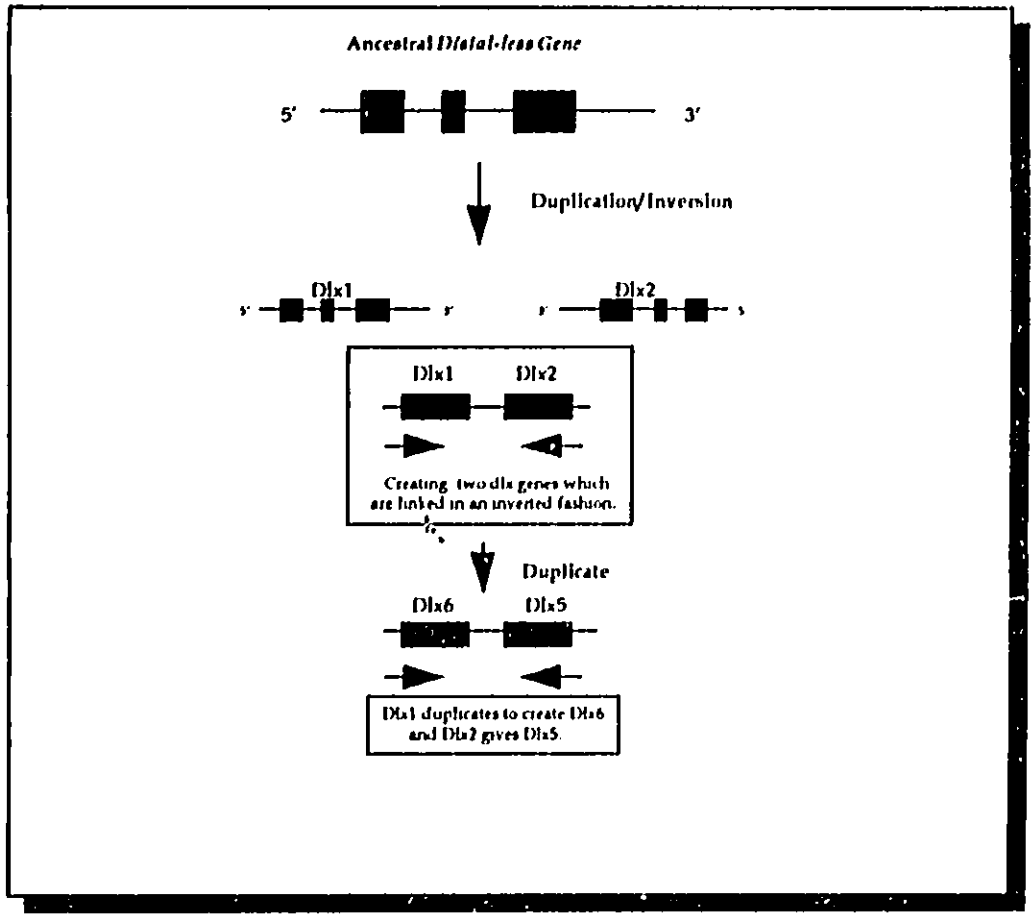


Figure 2-7. Evolutionary Paths for the *Dlx* gene family.

Figure 2-7: Possible Evolutionary Paths Taken By The *Distal-less* Gene family.

3. *Putative 5′ Regulatory Regions of Dlx and Msx Genes*

Chapter 3

3.1. Introduction

Genetic interactions are important to understand the mechanisms involved in embryonic pattern formation. The cascade of genetic interactions involving homeotic selector genes has been extensively studied in the *Drosophila*. However, little is known about the genetic interactions that will lead to homeobox gene expression in vertebrates. The regulation of *msx* and *dlx* gene expression is still poorly understood, although, the signaling molecule *BMP* has been reported to participate in the control of *msx* gene expression.[82]

Retinoic acid is known to mediate gene expression through two families of nuclear steroid receptors, the retinoic acid receptor (RAR) and the retinoid X receptor (RXR). [106-109] These steroid receptors are found in the nucleus bound to genomic DNA as either homodimers (ie. RAR/RAR) or heterodimers (ie. RXR/RAR). RXR has been shown to also heterodimerize with other members of the steroid receptor super family, namely the thyroid hormone receptor (RXR/THR) and the vitamin D_3 receptor (RXR/VDR). The steroid receptors bind the genomic DNA at sites called response elements, each type of receptor recognizes its unique DNA sequence. The response elements are composed of two half-sites (each consisting of 6bp), therefore two receptors bind to one response element (6bp-space-6bp, the space between the half sites is unique for each combination of receptors). To cause transactivation (induction) of the target gene, two steroid hormone receptors must be bound to their response elements and also be complexed with their ligands.

RA has been shown to repress genes like *goosecoid, Hoxb1,* and *Myf5.* [79] *Myf5* downregulation was shown to be specifically due to the action of *9-cis* RA on RXR, rather than *all-trans* RA on RAR. [79] The action of RAR and RXR on target gene activation may therefore be very different. For example, it is possible that high cellular RAR levels cause preferential RAR homodimerization and activation of gene transcription, whereas high RXR levels cause a preferential heterodimerization with either RAR, VDR, or THR, which may cause target gene suppression.

It is well known that RA induces Hox gene expression, [19, 25, 62] however, recently it has also been shown to downregulate certain genes. [79] However, the nature of this negative regulation is still unknown. It is uncertain whether or not these interactions with RA are direct or indirect. Recently, *Hoxb1* gene

nature of this negative regulation is still unknown. It is uncertain whether or not these interactions with RA are direct or indirect. Recently, *Hoxb1* gene expression within the rhombomeres was shown to be restricted by RA acting on a RARE repressor element found in the mouse and chick *Hoxb1* flanking regions. [84]

The direct action of RA on *distal-less* genes is likely because RA has been shown to directly induce overexpression of *Hox* genes in vertebrates. Furthermore, *Drosophila* homeobox genes were reported to downregulate *distal-less* genes. Therefore, vertebrate *Hox* genes may be responsible for directly downregulating *distal-less* genes.

The mechanism of RA action on vertebrate homeobox genes like *distal-less*, may therefore involve ectopic expression and/or overexpression of *Hox* genes, which in turn would directly downregulate (*dlx* or *msx*) homeobox gene expression.

To address the question of whether retinoic acid acts directly or indirectly on homeobox genes like *dlx* and *msx*, their 5' flanking sequences were isolated and characterized.

3.2. Methods & Materials

Genomic clones for each member of the *dlx* and *msx* homeobox gene family were isolated from either a zebrafish lambda Dash II library (provided by M. Petkovich) or a zebrafish EMBL-4 lambda phage library (provided by A. Fjose). The EMBL-4 library was initially screened with a mouse *Hox-7* homeobox probe.

The *Hox-7* homeobox fragment was labelled using the random priming protocol (refer to section 2.2.2 for more details). The zebrafish *msxA*, *msxB*, *msxD*, and *dlx3* genes were isolated from the EMBL-4 library. Three genes (*msxC*, *dlx2*, and *dlx4*) were isolated from the lambda DashII library using the corresponding cDNAs as probes (probes were labeled as described in section 2.2.2). Exon 1 and 5' flanking sequence from each genomic clone were subcloned into a plasmid vector, Bluescript KSII, and confirmed by sequencing (dideoxy-termination method, see section 2.2.2).

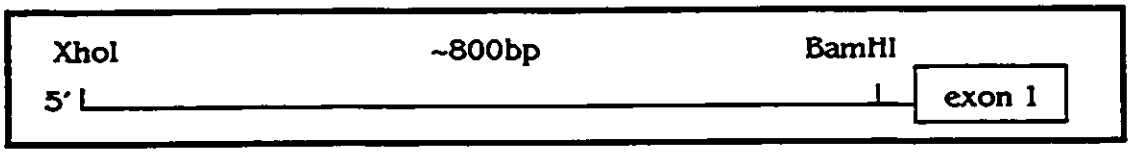
Further sequencing of the large genomic subclones was carried out following production of a series of deletions with ExoIII. [110] The pKSII subclones were linearized by double digests within the first exon and mutliple cloning site.

multiple cloning site and leave a 3' protrusion, which is necessary to protect the plasmid from an ExoIII bidirectional digest. The second enzyme must digest the DNA more 5' from the initial digest, within the first exon, and leave a 5' protrusion. Exonuclease III (ExoIII) only digests 5' protrusions. Once the template DNA is fully linearized and 3' protected, the ExoIII reaction is carried out. Timed aliquotes (every 30 seconds) were removed to ice. The terminated ExoIII reactions were then subjected to an S1 nuclease digest (for 30 minutes at 30°C), removing the 3' DNA overhangs where ExoIII digested. Both of the enzymes (ExoIII, and S1 nuclease) were inactivted by incubation at 70°C for 10 minutes. The DNA was then ligated to itself, and transfected into E.coli. The resulting E.coli colonies were grown up in culture and the plasmid DNA used as the template for sequencing, was isolated.

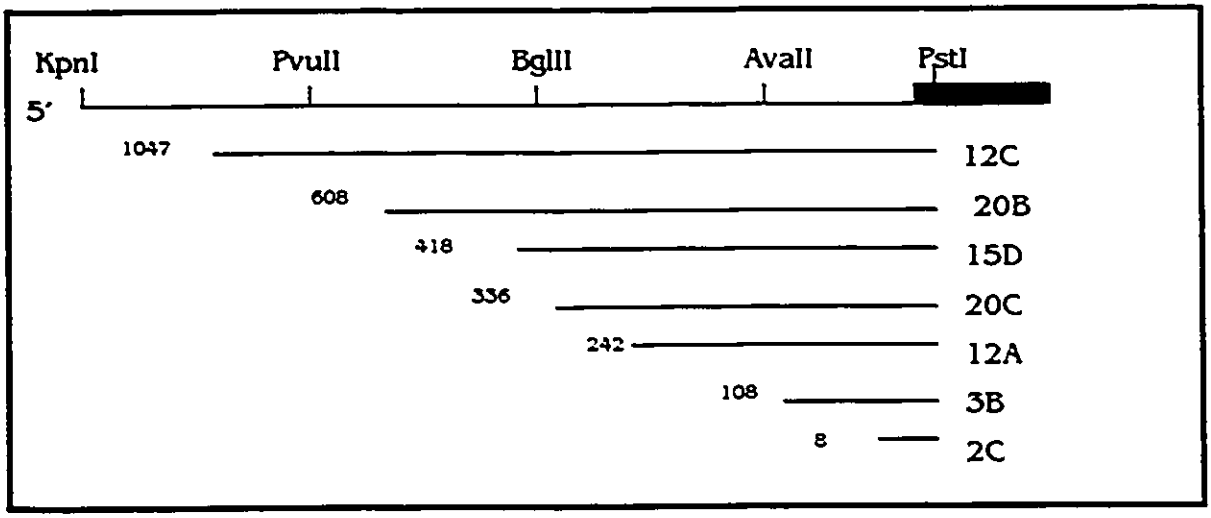
3.3. Results

The genomic 5' flanking sequences from *msxA, msxB, msxD, dlx1, dlx3,* and *dlx4* are shown in figures 1 to 7. *MsxC* and *msxE* genomic clones were isolated, however, their genomic 5' flanking sequences were not sequenced. Anther exception is *dlx2*, the two genomic clones only contain sequences from the first intron and downstream. The genomic clone for dlx7 has not yet been isolated.

Dlx1. An 800bp genomic XhoI-BamHI fragment was isolated and sequenced, where the BamHI site is found 90bp upstream from the initiation codon (figure 3-1). An estrogen receptor's response element half site, and many potential TATA boxes were found. Five chick retinoic acid repressor elements were also found, however, their half site sequences had a 4/6bp and 3/6bp match. Another cis-element was found, a mouse retinoic acid repressor element at position -79bp with a spacer of 7bp between the half sites, one of which had 3/6bp match and the other had a 5/6bp sequence match. See below for a schematic of the genomic fragment.

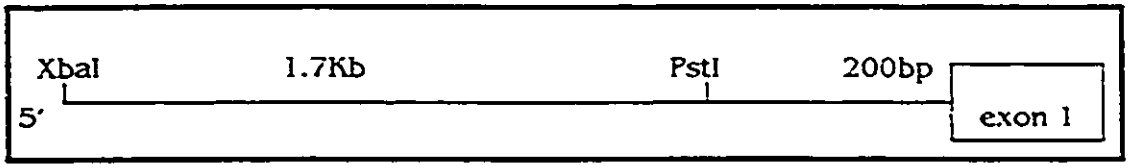


Dlx3. A 1.5Kb genomic KpnI-PstI fragment was isolated, where the PstI site is in exon 1 (figure 3-2). The entire 1.5Kb genomic fragment was sequenced. Many cis-acting elements were located in the 5' flanking sequence. Two "CAAT" sites, and four potential "TATA" boxes were found. One very interesting finding is a cis-element which is highly homologous to the *Drosophila Ubx* repressor element, BX1. One estrogen half site, matching 5/5bp, and three progesterone/glucocorticoid MMTV half sites, matching 6/6bp, were found. One chick retinoic acid repressor element contained one perfect 6/6 bp half site sequence match, and the second half site half site had 3/6bp match, the half sites are spaced by 8bp. Eight other mouse retinoic acid repressor element were found, however, one half site matches 4/6bp and the other 3/6bp. The schematic below indicates the subclones used to assemble the complete 1.5Kb sequence.

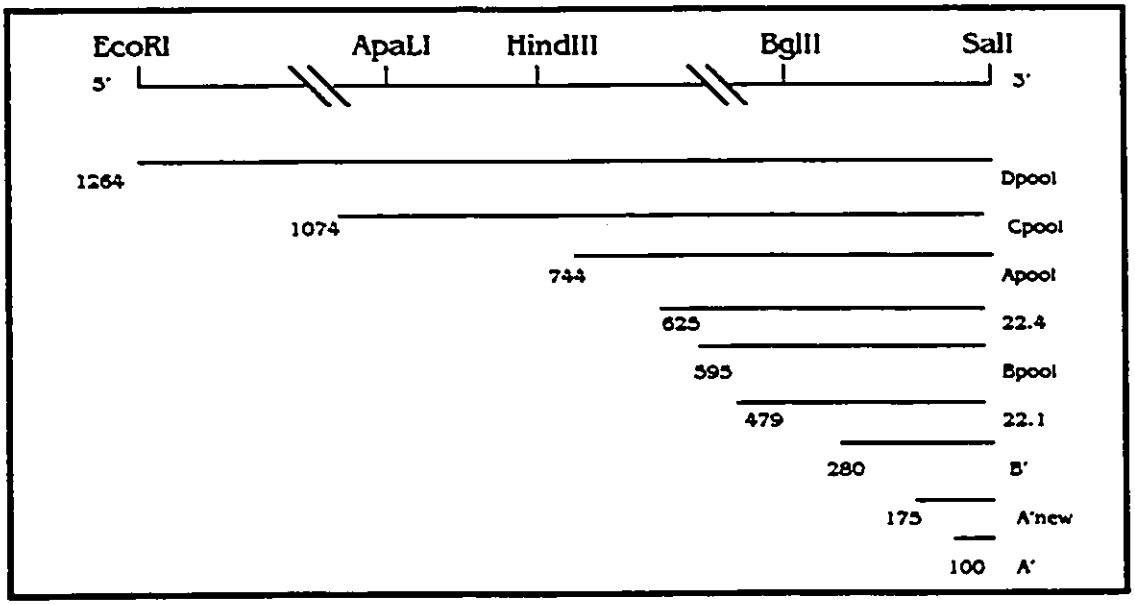


Dlx4. A 1.7Kb genomic PstI-XbaI fragment was isolated (see schematic below), where the PstI site is located 200bp 5' from the PstI site in the first exon. The sequence shown in figure 3-3 includes 1.1Kb upstream from the PstI site in the 5' flanking region. One chick retinoic acid repressor element spaced by 7bp with 5/6bp matching in the first half site, and 3/6bp matching in the

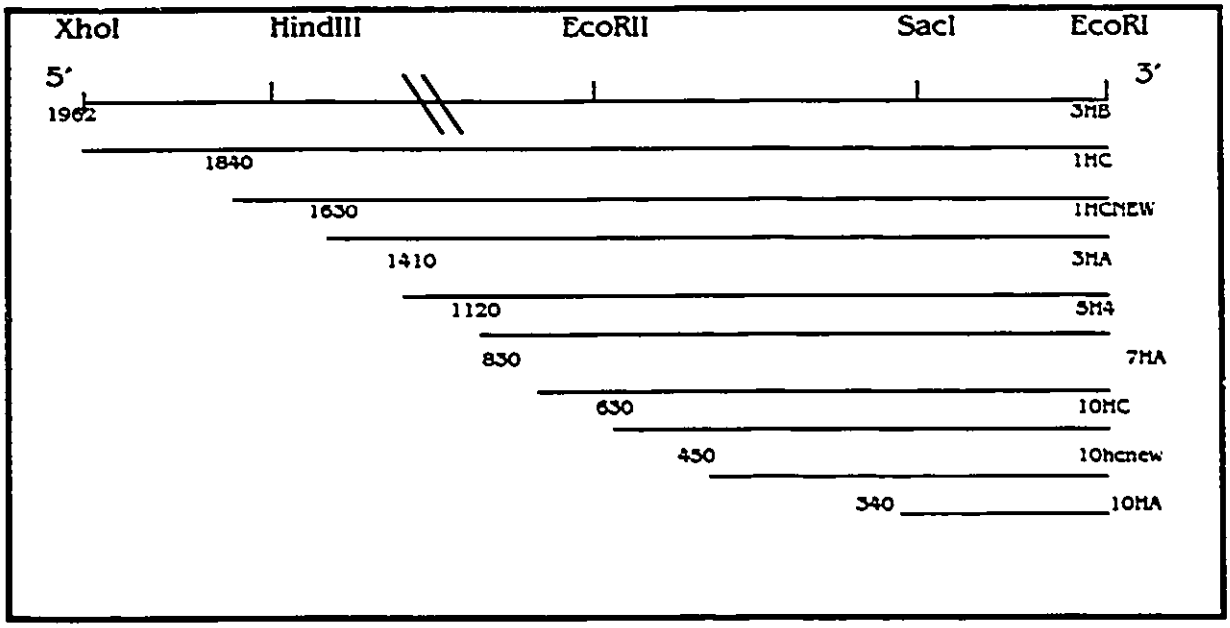
second half site was found. Seven mouse retinoic acid repressor elements were found with 4/6bp and 3/6bp half site sequence matches. The schematic below shows the 5′ genomic fragment that was isolated.



MsxA. The genomic fragment (EcoRI-SalI) contains 2.0Kb of 5′ flanking sequence (figure 3-4), where the SalI site is 5′ to exon 1. There are a few hundred base pairs of sequence missing between deletion clones Cpool-Dpool and 22.1-Bpool. A CAAT box and a TATA box were found. Other cis-acting regulatory elements were found; an OctA1 site, two chick retinoic acid repressor elements spaced by 8bp (half sites matching 5/6bp and 3/6bp), three mouse retinoic acid repressor elements spaced by 7bp (half sites matching 5/6bp and 3/6bp, or 4/6bp and 4/6bp). The complete 1.2Kb sequence was generated from deletion clones, as indicated below in the schematic.



MsxB. The genomic fragment (XhoI-EcoRI) that was sequenced contains 1.9Kb, where the EcoRI site is 5' to exon 1 (figure 3-5). There is one hole at position -1120bp, with very few base pairs missing (insert size was 1.9Kb). Some interesting features of the *msxB* 5' flanking region, 1) there is a repeat of 34 "AC" bases at position -730bp, and 2) there is a stretch of "A"'s found at position -950bp, and at position -1560bp. Many cis-elements were located in the 5' flanking sequence; two estrogen response element half sites, one glucocorticoid/ progesterone MMTV response element, and three chick retinoic acid repressor elements which have different half-site sequence matches. Three mouse repressor retinoic acid elements were also found. The schematic below shows the arrangement of subclones used to assemble the 5' flanking sequence.



MsxC. Two genomic fragments containing part of exon 1 and sequences 5' to the first exon were isolated and verified by sequencing (clone C8 (EcoRI-SacI), 2.8Kb 5'; and C5 (XbaI-EcoRI), 1.2Kb 5'). Another EcoRI-SacII genomic fragment (clone 1.2) containing 1.2Kb was isolated and identified as the intron joining exons 1 and 2 of *msxC* (figure 3-6).

MsxD. The genomic fragment was sequenced from the XbaI site within the first exon going 5'. The overall sequence reported contains 1kb of 5' flanking region (figure 3-7). Many cis acing elements were found in the 5' flanking region; a "CAAT" box and TATA box, two estrogen half sites, one OctA1 site, one OctB1 site, two chick retinoic acid repressor elements, one of which had 5/6bp and 4/6bp half site sequence matches with n=8bp, the other had 5/6 and 3/6bp matches and n=7bp. Two mouse retinoic acid repressor elements were also found, one of which had 4/6 and 5/6bp half site sequence matches with n=8bp, and the other had 4/6 and 4/6bp matches with n=8bp.

3.4. *Discussion*

The putative cis-acting regulatory elements found in the 5' flanking regions of the *dlx* and *msx* genes are not necessarily relevant. These elements were found by searching a databank consisting of published cis-acting elements. Direct evidence for the binding of steroids to these putative response elements will be provided by further experiments, of which, have not been conducted. However, it is possible to speculate that steriod receptors will bind to these 5' flanking sequences.

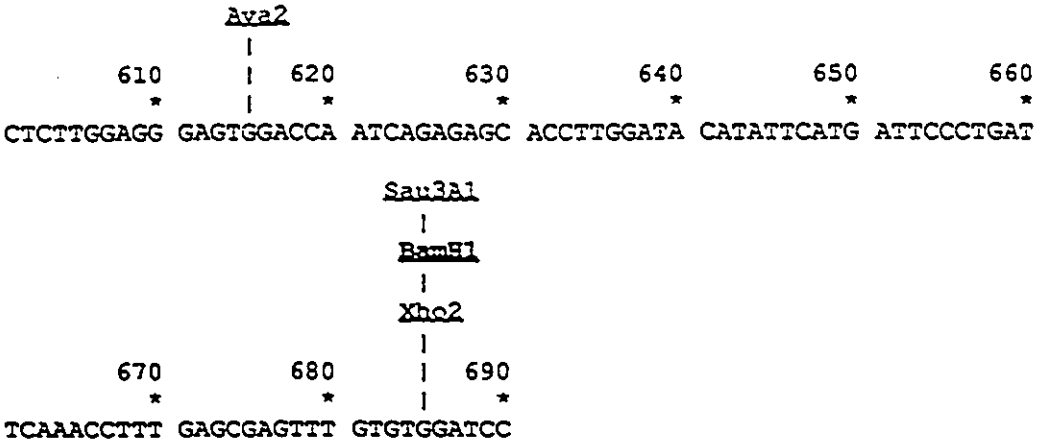
The downregulation of the *distal-less* homeobox family, by the *Hom-C* genes, has been shown to be due to a repressor element found in the 5' flanking sequence of the *distal-less* gene. The repressor element, Bx1, was shown to have a high affinity for the *Hom-C Ubx* gene. Therefore, direct downregulation of *distal-less* expression is accomplished by *Hom-C* genes in *Drosophila*. One very interesting finding is the BX1 element in the *dlx3* 5' sequence. Therefore, it is possible that the vertebrate homologs of *Ubx* recognize the same cis-element, because of their sequence conservation.

The other interesting cis-acting element found in the *msx* and *dlx* 5' regions, was a sequence similar to the mouse and chick retinoic acid repressor element naturally found in the *Hoxb1* regulatory region. This *Hoxb1* element is necessary to restrict the expression of *Hoxb1* to its specific rhombomere.

It is likely that embryos treated with retinoic acid will display a dose dependent induction of *Hox* genes, including the seventh *Hox* paralogous group. This induction of *Hox* genes may results in a direct downregulation of *dlx3* expression, and a restiction of certain *Hox* gene expression within the rhombomeres. These RARE's responsible for *Hox* gene downregulation may also be directly responsible for certain *dlx* and/or *msx* gene downregulation. The downregulation of *dlx3* expression may in turn affect other *dlx* genes because they are thought to cross regulate each other. Therefore, RA may indirectly downregulate certain *dlx* gene expression through the *Hox* genes, and directly downregulate other *dlx* gene expression through the repressor RARE cis-elements.

Experiments aimed at elucidating the mechanism of retinoic acid action on *dlx* and *msx* homeobox genes will shed further light into the exact nature of this interaction.

81	mRARErepres				
	(part #1)	79	3/6	+	AatGCt
	(part #2)	86	5/6	+	tGTTCA
98	mRARErepres				
	(part #2)	93	3/6	-	cGTgaA
	(part #1)	100	4/6	-	AcGGCt
103	mRAPErepres				
	(part #2)	98	3/6	-	gGcTCc
	(part #1)	105	4/6	-	tGGtCA
126	mRARErepres				
	(part #2)	122	3/6	-	tGcTCc
	(part #1)	128	4/6	-	AtaGCA
183	mRARErepres				
	(part #2)	177	4/6	-	AGTaaA
	(part #1)	185	3/6	-	AaaaCA
183	mRARErepres				
	(part #2)	177	4/6	-	AGTaaA
	(part #1)	185	3/6	-	AaaaCA



Pos.	Subseq		Loc.	#Res. Matched		Subseq Found
30	cRARE					
		(part #2)	26	4/6	-	AGaTgA
		(part #1)	32	3/6	-	aaGCaA
36	cRARE					
		(part #1)	34	4/6	+	GaGaCA
		(part #2)	42	3/6	+	caTTtA
96	cRARE					
		(part #1)	94	4/6	+	GaGCCg
		(part #2)	100	3/6	+	tGacCA
102	cRARE					
		(part #1)	100	4/6	+	tGaCCA
		(part #2)	108	3/6	+	AGaaaA
103	cRARE					
		(part #2)	98	3/6	-	gGcTCc
		(part #1)	105	4/6	-	tGGtCA
104	ER_half-si					
		(part #1)	104	5/5	-	GGTCA
30	mRARErepres					
		(part #2)	26	4/6	-	AGaTgA
		(part #1)	32	3/6	-	AaGcaA
31	mRARErepres					
		(part #2)	26	4/6	-	AGaTgA
		(part #1)	33	3/6	-	caaGCA
54	mRARErepres					
		(part #2)	50	3/6	-	gGaTaA
		(part #1)	56	4/6	-	AGtaCA

Xho1

```
          10         20         30         40         50         60
           *          *          *          *          *          *
AAATGAATCA TCAAAGTAAT TCATCTTTGC TTGGAGACAC ACATTTATCC TGTACTTAAT

          70         80         90        100        110        120
           *          *          *          *          *          *
CATATTTGTT AGGTGTTAAA TGCTTTGTTC ACGGAGCCGT GACCAGCAGA AAATTAGGAG

         130        140        150        160        170        180
           *          *          *          *          *          *
CATGCTATCT TCGTTCACTA TTACGAAAGG TTTGCTTTTT ATATAACAGC CTTTACTTTT

         190        200        210        220        230        240
           *          *          *          *          *          *
GTTTTTTATT CTGATTTGTT TATTTTAAGA GATGTATTTA TAATTAATAT AATAGTAATA

         250        260        270        280        290        300
           *          *          *          *          *          *
GTAATAATAA TAATATAATT TACTTATAAT TACACATATA TGCATAAGAA TTACCATTTC
```

TaqI EcoRS

```
         310        320        330        340        350        360
           *          *          *          *          *          *
CCTTGGTTAT ATAATTTAAT TTTACTGTAC TTTCGATAAT TTGTAATTGA TATCAGGCCT

         370        380        390        400        410        420
           *          *          *          *          *          *
TACACTGATG GCACTGATAT AAAAACGAAT GTAATTATAA AAAAATTTAT ATATTTACAC

         430        440        450        460        470        480
           *          *          *          *          *          *
TGTTGAGACA AATTTGGTGG TTTCAGCGAA ATGTGTTTGA GCCTCCCAAG TACGGCCTAT
```

Ava2

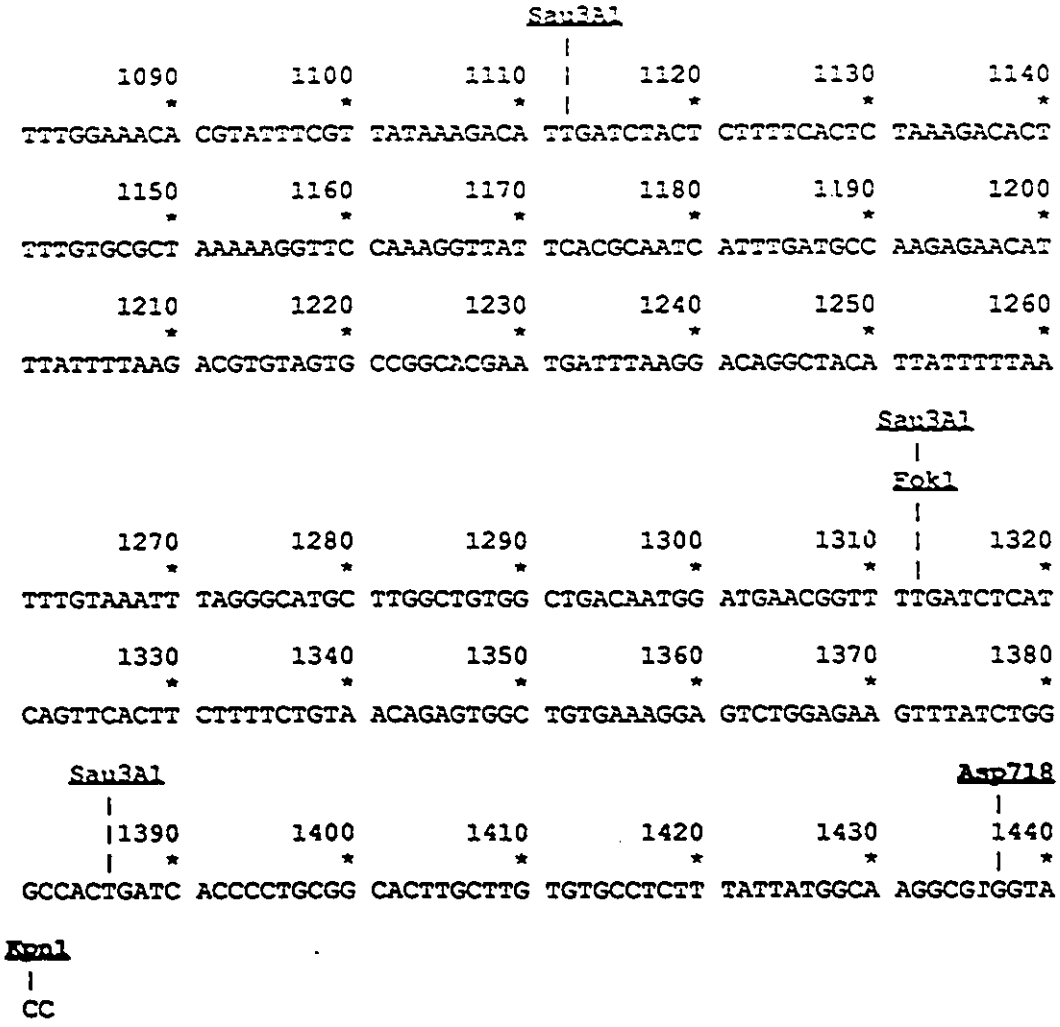
```
         490        500        510        520        530        540
           *          *          *          *          *          *
GCGTCTCGGT CCATTCCCTA AATTTCTACA TTAGCATTGA CACAATAGAG TCACAGGTTG
```

Ava2

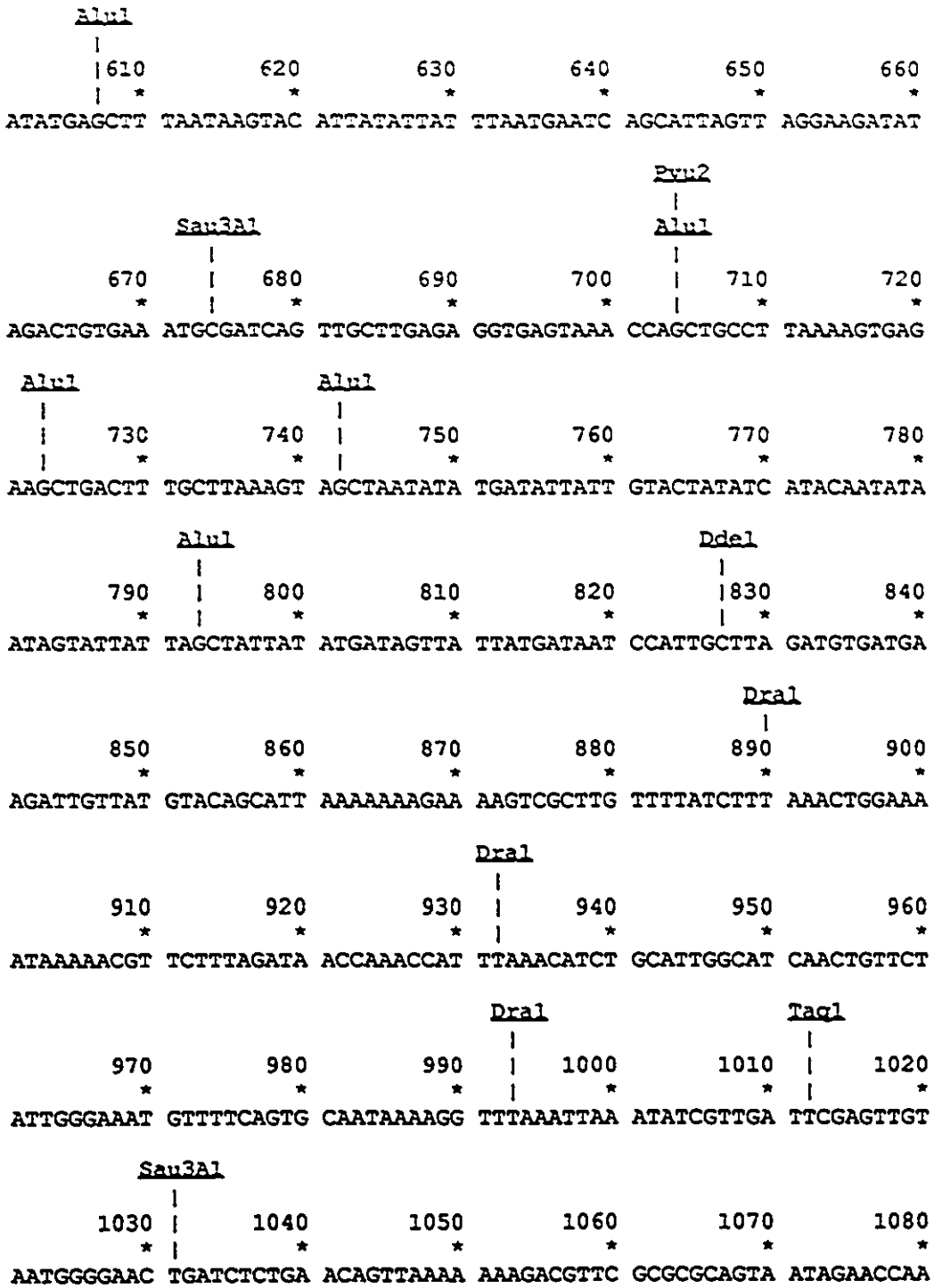
```
         550        560        570        580        590        600
           *          *          *          *          *          *
GTCTTATTTC TTGAGTTTAT CAGTCCGTCG GACCCGACCC TCGGATTACC AATCGCATCG
```

Figure 3-1. *Dlx1* genomic 5′ sequence.
A BamHI-XhoI fragment, 800bp, was sequenced from the XhoI site.

146	cRARE				
	(part #2)	141	3/6	-	cGTTgt
	(part #1)	148	4/6	-	GGaCgA
199	cRARE				
	(part #2)	195	3/6	-	AaccCA
	(part #1)	201	4/6	-	GtcCCA
199	cRARE				
	(part #1)	197	4/6	+	GGGaCg
	(part #2)	203	3/6	+	AtgcCA
205	cRARE				
	(part #1)	203	4/6	+	atGCCA
	(part #2)	211	3/6	+	AacaCA
238	cRARE				
	(part #2)	232	6/6	-	AGTTCA
	(part #1)	240	3/6	-	GaGaaA
410	ER_half-si				
	(part #1)	410	5/5	-	GGTCA
534	GR/PR-MMTV				
	(part #1)	534	6/6	+	TGTTCT
955	GR/PR-MMTV				
	(part #1)	955	6/6	+	TGTTCT
1199	GR/PR-MMTV				
	(part #1)	1199	6/6	-	TGTTCT
84	mRARErepres				
	(part #1)	82	4/6	+	cGGGtA
	(part #2)	88	3/6	+	tcTTCc
116	mRARErepres				
	(part #2)	112	4/6	-	AGTggA
	(part #1)	118	3/6	-	AGGatg
117	mRARErepres				
	(part #2)	112	4/6	-	AGTggA
	(part #1)	119	3/6	-	AaGGat
129	mRARErepres				
	(part #1)	127	4/6	+	AaGGaA
	(part #2)	133	3/6	+	AaaaCA
149	mRARErepres				
	(part #2)	145	3/6	-	cGaTCg
	(part #1)	151	4/6	-	AGtGgA
169	mRARErepres				
	(part #1)	167	3/6	+	AccGaA
	(part #2)	175	4/6	+	AagTCA
203	mRARErepres				
	(part #1)	203	4/6	+	AtGcCA
	(part #2)	211	3/6	+	AacaCA
207	mRARErepres				
	(part #2)	202	3/6	-	cGTcCc
	(part #1)	209	4/6	-	ctGGCA
	Ubx-BX1repres				
	(part #1)	1260	7/7	-	AAATTAA



Pos.	Subseq		Loc.	#Res. Matched		Subseq Found
58	cRARE					
		(part #2)	52	4/6	-	AGTTgc
		(part #1)	60	3/6	-	GctCtA
72	cRARE					
		(part #2)	67	4/6	-	ctTTCA
		(part #1)	74	3/6	-	GacCgA
73	cRARE					
		(part #2)	67	4/6	-	ctTTCA
		(part #1)	75	3/6	-	cGaCCg
73	cRARE					
		(part #1)	71	4/6	+	GGtCgA
		(part #2)	77	3/6	+	AaTaCc
107	cRARE					
		(part #2)	101	3/6	-	AtaTgA
		(part #1)	109	4/6	-	GGaCCg


AluI
610 620 630 640 650 660
ATATGAGCTT TAATAAGTAC ATTATATTAT TTAATGAATC AGCATTAGTT AGGAAGATAT
PvuII
Sau3AI
AluI
670 680 690 700 710 720
AGACTGTGAA ATGCGATCAG TTGCTTGAGA GGTGAGTAAA CCAGCTGCCT TAAAAGTGAG
AluI
AluI
730 740 750 760 770 780
AAGCTGACTT TGCTTAAAGT AGCTAATATA TGATATTATT GTACTATATC ATACAATATA
AluI
DdeI
790 800 810 820 830 840
ATAGTATTAT TAGCTATTAT ATGATAGTTA TTATGATAAT CCATTGCTTA GATGTGATGA
DraI
850 860 870 880 890 900
AGATTGTTAT GTACAGCATT AAAAAAAGAA AAGTCGCTTG TTTTATCTTT AAACTGGAAA
DraI
910 920 930 940 950 960
ATAAAAACGT TCTTTAGATA ACCAAACCAT TTAAACATCT GCATTGGCAT CAACTGTTCT
DraI
TaqI
970 980 990 1000 1010 1020
ATTGGGAAAT GTTTTCAGTG CAATAAAAGG TTTAAATTAA ATATCGTTGA TTCGAGTTGT
Sau3AI
1030 1040 1050 1060 1070 1080
AATGGGGAAC TGATCTCTGA ACAGTTAAAA AAAGACGTTC GCGCGCAGTA ATAGAACCAA

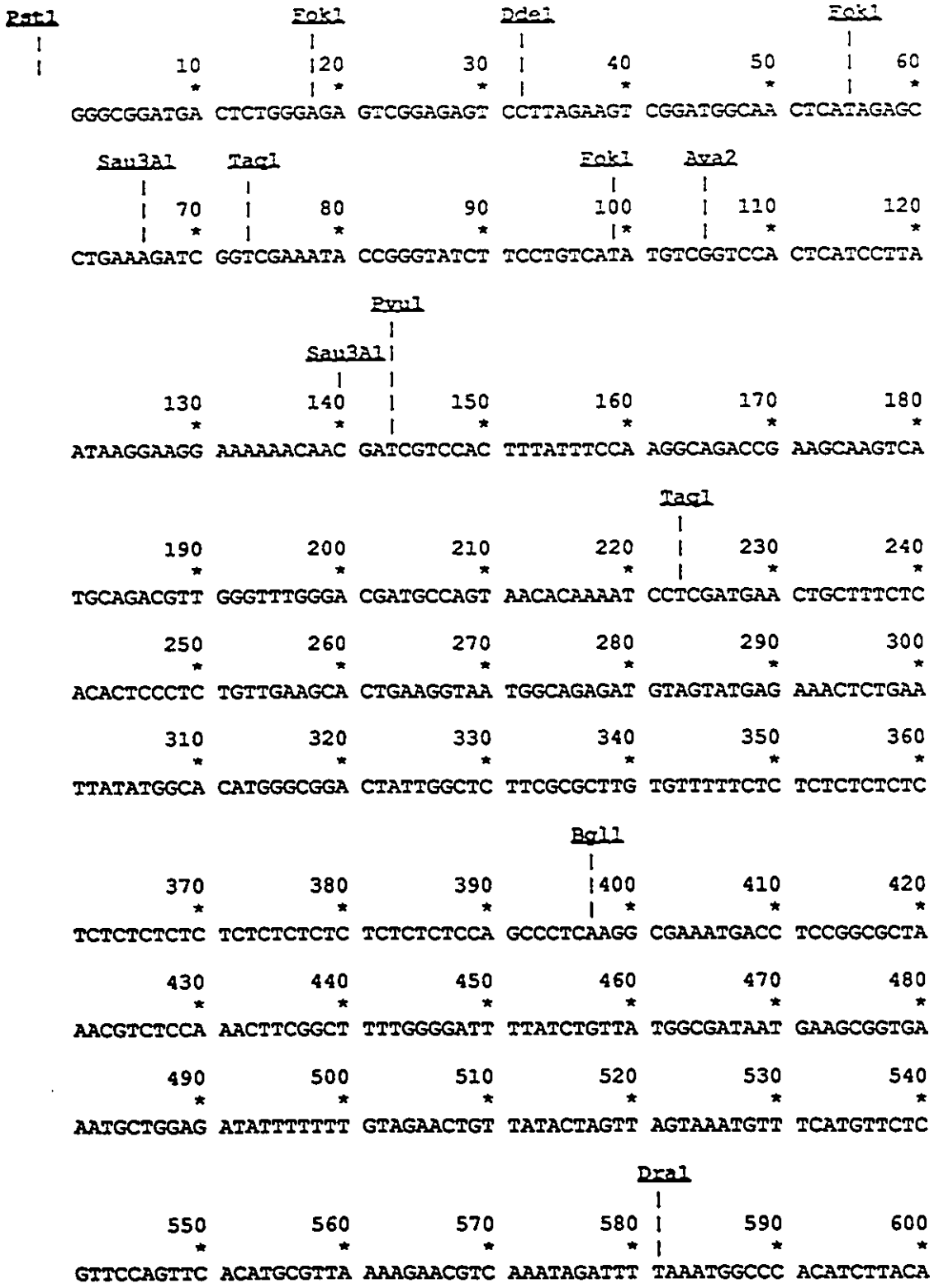

PstI
FokI
DdeI
FokI
10
20
30
40
50
60
GGGCGGATGA CTCTGGGAGA GTCGGAGAGT CCTTAGAAGT CGGATGGCAA CTCATAGAGC
Sau3AI
TaqI
FokI
Ava2
70
80
90
100
110
120
CTGAAAGATC GGTCGAAATA CCGGGTATCT TCCTGTCATA TGTCGGTCCA CTCATCCTTA
PvuI
Sau3AI
130
140
150
160
170
180
ATAAGGAAGG AAAAAACAAC GATCGTCCAC TTTATTTCCA AGGCAGACCG AAGCAAGTCA
TaqI
190
200
210
220
230
240
TGCAGACGTT GGGTTTGGGA CGATGCCAGT AACACAAAAT CCTCGATGAA CTGCTTTCTC
250
260
270
280
290
300
ACACTCCCTC TGTTGAAGCA CTGAAGGTAA TGGCAGAGAT GTAGTATGAG AAACTCTGAA
310
320
330
340
350
360
TTATATGGCA CATGGGCGGA CTATTGGCTC TTCGCGCTTG TGTTTTTCTC TCTCTCTCTC
BglI
370
380
390
400
410
420
TCTCTCTCTC TCTCTCTCTC TCTCTCTCCA GCCCTCAAGG CGAAATGACC TCCGGCGCTA
430
440
450
460
470
480
AACGTCTCCA AACTTCGGCT TTTGGGGATT TTATCTGTTA TGGCGATAAT GAAGCGGTGA
490
500
510
520
530
540
AATGCTGGAG ATATTTTTTT GTAGAACTGT TATACTAGTT AGTAAATGTT TCATGTTCTC
DraI
550
560
570
580
590
600
GTTCCAGTTC ACATGCGTTA AAAGAACGTC AAATAGATTT TAAATGGCCC ACATCTTACA

Figure 3-2. *Dlx3* genomic 5′ flanking sequence.
A 1.7Kb KpnI-PstI fragment was isolated, and sequence from the PstI site within the first exon.

217	cRARE				
	(part #1)	215	3/6	+	GtGCgt
	(part #2)	221	4/6	+	AGTTtt
217	cRARE				
	(part #1)	215	3/6	+	GtGCgt
	(part #2)	221	4/6	+	AGTTtt
39	mRARErepres				
	(part #2)	33	4/6	-	AGTatA
	(part #1)	41	3/6	-	tGGtgA
120	mRARErepres				
	(part #2)	114	3/6	-	AcTaaA
	(part #1)	122	4/6	-	AGGatA
120	mRARErepres				
	(part #2)	114	3/6	-	AcTaaA
	(part #1)	122	4/6	-	AGGatA
185	mRARErepres				
	(part #2)	180	4/6	-	AGTctA
	(part #1)	187	3/6	-	AaaGaA
209	mRARErepres				
	(part #2)	204	4/6	-	AcaTCA
	(part #1)	211	3/6	-	AttGtA
216	mRARErepres				
	(part #1)	214	3/6	+	tGtGCg
	(part #2)	221	4/6	+	AGTTtt
216	mRARErepres				
	(part #1)	214	3/6	+	tGtGCg
	(part #2)	221	4/6	+	AGTTtt

PstI

```
        10          20          30          40          50          60
         *           *           *           *           *           *
GGACTGCTGT TTACATAAAT ATCCCGATAT ACTTATCACC AGCAGGAATG TCTGATTCCT

        70          80          90         100         110         120
         *           *           *           *           *           *
TAAATGTGTT TGTAACACAA TTCATATGTA ATTACTTATA TTTCTGTGTT TAGTTATATC

 DdeI        DdeI
  |           |
  |   130     |   140       150         160         170         180
  |    *      |    *         *           *           *           *
CTCTTAGACG GCCTCAGGAT TCAGTATTTT ATAGTTTTCT GCTTTAATCC GCATTAGACT

                                                    DraI
                                                     |
       190         200         210         220     | 230         240
         *           *           *           *     |  *          *
CTTCTTTTGT GTTTGTGTTG ATGTTTACAA TAATGTGCGT AGTTTTAAAA CCACATATTT

       250         260         270         280         290         300
         *           *           *           *           *           *
TGTTTAATTC AGACTAGAAT AATCTATCAC CAAAACTGTA TTAAATAATC TCTATTTGTT

       310         320                     XbaI
         *           *
TAGATTGTAA TCATTGCAAT TTGCA
```

Pos.	Subseq		Loc.	#Res. Matched		Subseq Found
3	cRARE					
		(part #1)	1	3/6	+	GGaCtg
		(part #2)	8	4/6	+	tGTTtA
37	cRARE					
		(part #2)	33	4/6	-	AGTatA
		(part #1)	39	3/6	-	GtGatA
39	cRARE					
		(part #2)	33	4/6	-	AGTatA
		(part #1)	41	3/6	-	tGGtgA
74	cRARE					
		(part #1)	72	3/6	+	GtaaCA
		(part #2)	79	5/6	+	AaTTCA
132	cRARE					
		(part #1)	130	3/6	+	GGcCtc
		(part #2)	138	4/6	+	gaTTCA
133	cRARE					
		(part #1)	131	3/6	+	GcctCA
		(part #2)	138	4/6	+	gaTTCA
140	cRARE					
		(part #1)	138	3/6	+	GattCA
		(part #2)	146	4/6	+	AtTTtA
217	cRARE					
		(part #1)	215	3/6	+	GtGCgt
		(part #2)	221	4/6	+	AGTTtt

Figure 3-3. *Dlx4* genomic 5′ flanking sequence.

A 1.7Kb PstI-XbaI fragment was isolated, but only 1kb was sequenced from the PstI site.

118	cRARE				
	(part #1)	116	3/6	+	tctCCA
	(part #2)	123	4/6	+	AGTTtg
130	cRARE				
	(part #2)	125	3/6	-	AcTTtg
	(part #1)	132	4/6	-	GGctCA
157	cRARE				
	(part #1)	155	3/6	+	GcGtCt
	(part #2)	162	4/6	+	AaaTCA
44	mRARErepres				
	(part #2)	38	4/6	-	gGTTCt
	(part #1)	46	4/6	-	AGGttA
55	mRARErepres				
	(part #2)	51	3/6	-	ttTaCA
	(part #1)	57	4/6	-	AGcGCg
59	mRARErepres				
	(part #2)	53	4/6	-	cGTTtA
	(part #1)	61	3/6	-	AGcGag
61	mRARErepres				
	(part #2)	57	3/6	-	AGcgCg
	(part #1)	63	4/6	-	AGaGCg
77	mRARErepres				
	(part #2)	72	5/6	-	AGTTaA
	(part #1)	79	3/6	-	AGctaA
77	mRARErepres				
	(part #2)	72	5/6	-	AGTTaA
	(part #1)	79	3/6	-	AGctaA
101	mRARErepres				
	(part #2)	97	4/6	-	AGTTat
	(part #1)	103	3/6	-	ttGtCA
704	OCTA1.1				
	(part #1)	704	8/8	+	ATTTGCAT
704	OCTA1.3				
	(part #1)	704	8/8	+	ATTTGCAT

```
      1090        1100        1110        1120        1130        1140
        *           *           *           *           *           *
AACTAGAAAA TTAGAGGAAG GGGGAAAAAG ACACTTAAGA CAAGGCAAGG CAGATTATTA

      1150        1160        1170        1180        1190        1200
        *           *           *           *           *           *
TACAGTTAGA GCCCAAAGGA ATGATTGTTA CTGAAATGCT ACAGTATTGG AAAACGCTTC

      1210        1220        1230        1240        1250        1260
        *           *           *           *           *           *
AGATATTGTG TG-------AA GACAAGGCAA GGCAGATTAT TATACAGTTA GAGCCCAAAG

      1270        1280        1290        1300        1310        1320
        *           *           *           *           *           *
GAATGATTGT TACTGAAATG CTACAGTATT GGAAAACGCT TCAGATATTG TGTGATTTTG
```

```
                                                          TaqI
                                                           |
      DraI                                          EcoR1  |  EcoR5
       |                                             |     |    |
      1330        1340        1350        1360      1370  |   |1380
       |*           *           *           *        | *  |   |  *
ATTTAATTTA AAACCTAGTT AAGTCAAGAC AATTGTTAAA AAAAAAGGAA TTCGATATCA
```

```
    TaqI       TaqI
     |          |
AluI |       SalI    AvaI
 |   |          ||     |
 |   |1390      || 1400 |     1410
 |   |  *       ||    * |       *
GCTTATCGAT ACCGTCGACC CTCGGGGGCC CCCT
```

Pos.	Subseq	Loc.	#Res. Matched		Subseq Found
43	cRARE				
	(part #2)	38	4/6	-	gGTTCt
	(part #1)	45	3/6	-	GGttaA
44	cRARE				
	(part #2)	38	4/6	-	gGTTCt
	(part #1)	46	3/6	-	aGGttA
78	cRARE				
	(part #2)	72	5/6	-	AGTTaA
	(part #1)	80	3/6	-	taGCtA
78	cRARE				
	(part #2)	72	5/6	-	AGTTaA
	(part #1)	80	3/6	-	taGCtA
101	cRARE				
	(part #2)	97	4/6	-	AGTTat
	(part #1)	103	3/6	-	ttGtCA
118	cRARE				
	(part #1)	116	3/6	+	tctCCA
	(part #2)	123	4/6	+	AGTTtg

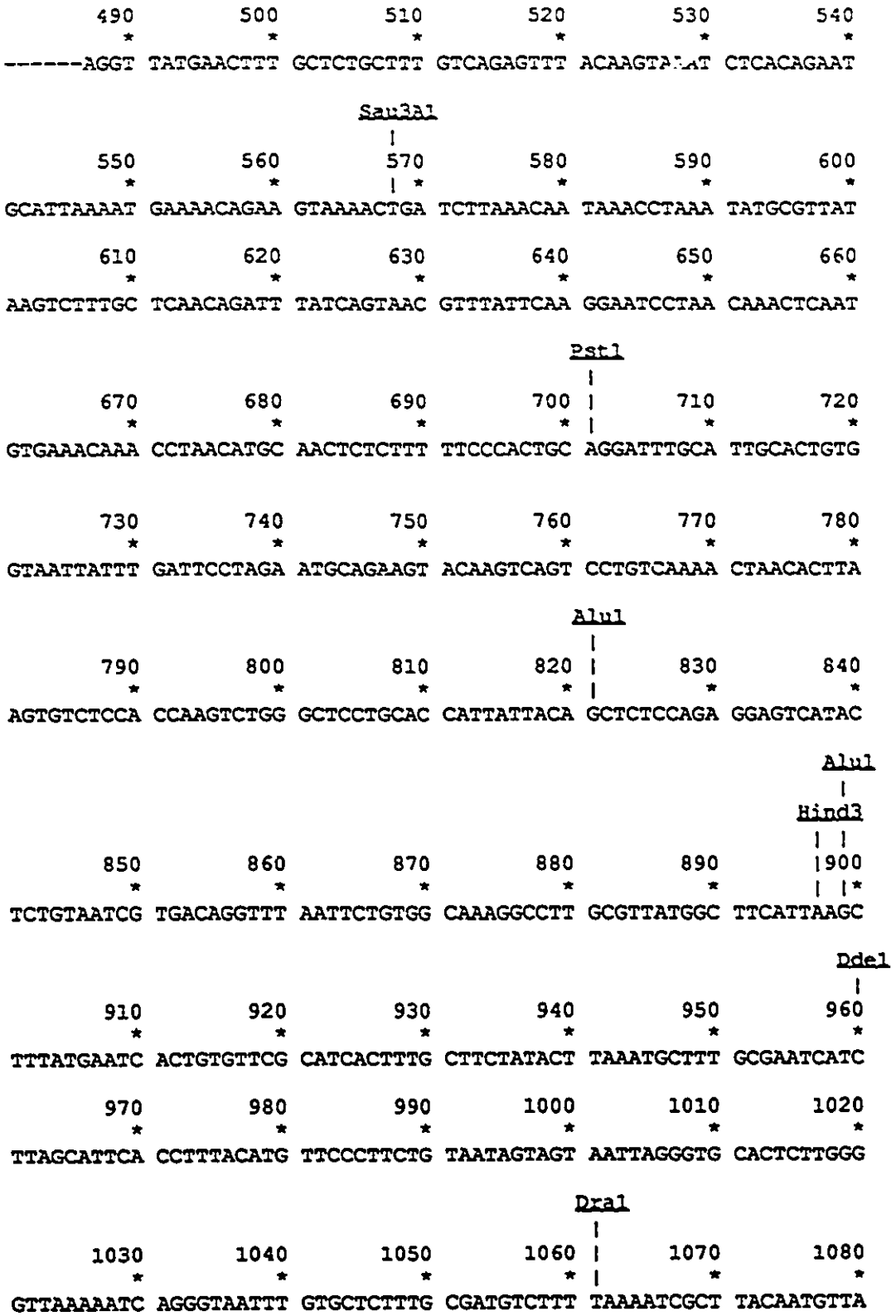
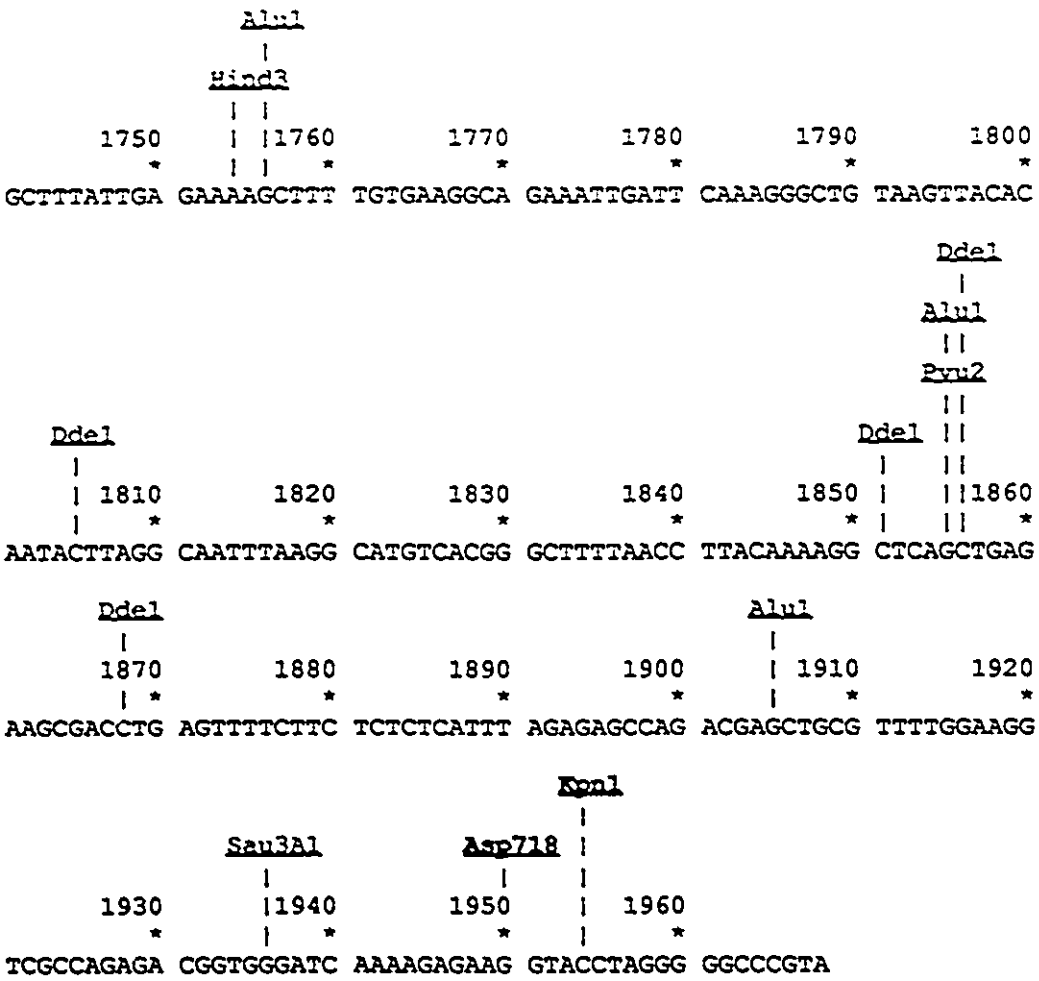

490 500 510 520 530 540
------AGGT TATGAACTTT GCTCTGCTTT GTCAGAGTTT ACAAGTA[illegible]T CTCACAGAAT
Sau3A1
550 560 570 580 590 600
GCATTAAAAT GAAAACAGAA GTAAAACTGA TCTTAAACAA TAAACCTAAA TATGCGTTAT
610 620 630 640 650 660
AAGTCTTTGC TCAACAGATT TATCAGTAAC GTTTATTCAA GGAATCCTAA CAAACTCAAT
PstI
670 680 690 700 710 720
GTGAAACAAA CCTAACATGC AACTCTCTTT TTCCCACTGC AGGATTTGCA TTGCACTGTG
730 740 750 760 770 780
GTAATTATTT GATTCCTAGA ATGCAGAAGT ACAAGTCAGT CCTGTCAAAA CTAACACTTA
AluI
790 800 810 820 830 840
AGTGTCTCCA CCAAGTCTGG GCTCCTGCAC CATTATTACA GCTCTCCAGA GGAGTCATAC
AluI
Hind3
850 860 870 880 890 900
TCTGTAATCG TGACAGGTTT AATTCTGTGG CAAAGGCCTT GCGTTATGGC TTCATTAAGC
DdeI
910 920 930 940 950 960
TTTATGAATC ACTGTGTTCG CATCACTTTG CTTCTATACT TAAATGCTTT GCGAATCATC
970 980 990 1000 1010 1020
TTAGCATTCA CCTTTACATG TTCCCTTCTG TAATAGTAGT AATTAGGGTG CACTCTTGGG
DraI
1030 1040 1050 1060 1070 1080
GTTAAAAATC AGGGTAATTT GTGCTCTTTG CGATGTCTTT TAAAATCGCT TACAATGTTA

Figure 3-4. *MsxA* genomic 5' flanking sequence.
A 2Kb EcoRI-SalI genomic fragment was isolated, and 1.5Kb were sequenced. Two regions of the 2Kb sequence are missing, at position 474bp, and 1206bp.

166	mRARErepres				
	(part #1)	164	3/6	+	gtGGaA
	(part #2)	170	4/6	+	AGgTCc
166	mRARErepres				
	(part #1)	164	3/6	+	gtGGaA
	(part #2)	170	4/6	+	AGgTCc
176	mRARErepres				
	(part #2)	170	3/6	-	ttTcCA
	(part #1)	178	4/6	-	gGGGgA
176	mRARErepres				
	(part #2)	170	3/6	-	ttTcCA
	(part #1)	178	4/6	-	gGGGgA
177	mRARErepres				
	(part #2)	171	3/6	-	ctTTCc
	(part #1)	179	4/6	-	AGGGgg
182	mRARErepres				
	(part #1)	180	3/6	+	gaGGCg
	(part #2)	186	4/6	+	AGcTCg
192	mRARErepres				
	(part #1)	190	3/6	+	cGGaCc
	(part #2)	196	4/6	+	AaTcCA
194	mRARErepres				
	(part #2)	189	4/6	-	AGcTCg
	(part #1)	196	3/6	-	tGGtCc
218	mRARErepres				
	(part #1)	216	4/6	+	AaGtCA
	(part #2)	224	4/6	+	tGgTCA
225	mRARErepres				
	(part #2)	219	4/6	-	AcTTCt
	(part #1)	227	3/6	-	AccaCA
226	mRARErepres				
	(part #1)	224	4/6	+	tGGtCA
	(part #2)	230	5/6	+	AGTTgA
236	mRARErepres				
	(part #2)	232	4/6	-	AcTTgA
	(part #1)	238	3/6	-	gtGtCA
237	mRARErepres				
	(part #1)	235	3/6	+	AcacCA
	(part #2)	243	4/6	+	AcaTCA
251	mRARErepres				
	(part #1)	249	5/6	+	AGGGgA
	(part #2)	255	3/6	+	taTTgA

218	cRARE				
	(part #1)	216	3/6	+	aaGtCA
	(part #2)	224	4/6	+	tGgTCA
226	cRARE				
	(part #1)	224	4/6	+	tGGtCA
	(part #2)	230	5/6	+	AGTTgA
236	cRARE				
	(part #2)	232	4/6	-	AcTTgA
	(part #1)	238	4/6	-	GtGtCA
237	cRARE				
	(part #1)	235	3/6	+	acaCCA
	(part #2)	243	4/6	+	AcaTCA
39	ER_half-si				
	(part #1)	39	5/5	+	GGTCA
225	ER_half-si				
	(part #1)	225	5/5	+	GGTCA
487	GR/PR-MMTV				
	(part #1)	487	6/6	+	TGTTCT
4	mRARErepres				
	(part #1)	2	3/6	+	AattCA
	(part #2)	9	4/6	+	cGTTaA
21	mRARErepres				
	(part #2)	16	4/6	-	gGTTaA
	(part #1)	23	3/6	-	gGcGgA
21	mRARErepres				
	(part #2)	16	4/6	-	gGTTaA
	(part #1)	23	3/6	-	gGcGgA
40	mRARErepres				
	(part #1)	38	4/6	+	cGGtCA
	(part #2)	45	3/6	+	gtcTCA
73	mRARErepres				
	(part #2)	67	3/6	-	AGTgtg
	(part #1)	75	4/6	-	AGGGac
107	mRARErepres				
	(part #2)	102	4/6	-	AGTTtg
	(part #1)	109	3/6	-	gGaGaA
108	mRARErepres				
	(part #2)	102	4/6	-	AGTTtg
	(part #1)	110	3/6	-	tGGagA
124	mRARErepres				
	(part #2)	120	3/6	-	cGcTgA
	(part #1)	126	4/6	-	cGcGCA
126	mRARErepres				
	(part #2)	120	3/6	-	cGcTgA
	(part #1)	128	4/6	-	AGcGCg
140	mRARErepres				
	(part #2)	134	3/6	-	tGgTCg
	(part #1)	142	4/6	-	gGGaCA
165	mRARErepres				
	(part #1)	163	3/6	+	gGtGgA
	(part #2)	170	4/6	+	AGgTCc
165	mRARErepres				
	(part #1)	163	3/6	+	gGtGgA
	(part #2)	170	4/6	+	AGgTCc

62	cRARE					
		(part #2)	56	4/6	-	AGTTgt
		(part #1)	64	3/6	-	GtGttA
72	cRARE					
		(part #2)	67	3/6	-	AGTgtg
		(part #1)	74	4/6	-	GGGaCg
106	cRARE					
		(part #2)	102	4/6	-	AGTTtg
		(part #1)	108	3/6	-	GaGaaA
107	cRARE					
		(part #2)	102	4/6	-	AGTTtg
		(part #1)	109	3/6	-	GGagaA
108	cRARE					
		(part #2)	102	4/6	-	AGTTtg
		(part #1)	110	3/6	-	tGGagA
131	cRARE					
		(part #1)	129	4/6	+	cGaCCA
		(part #2)	137	3/6	+	tGTcCc
131	cRARE					
		(part #2)	126	3/6	-	cGcgCA
		(part #1)	133	4/6	-	GGtCgA
140	cRARE					
		(part #2)	134	3/6	-	tGgTCg
165	cRARE					
		(part #1)	163	3/6	+	GGtggA
		(part #2)	170	4/6	+	AGgTCc
165	cRARE					
		(part #1)	163	3/6	+	GGtggA
		(part #2)	170	4/6	+	AGgTCc
166	cRARE					
		(part #1)	164	3/6	+	GtGgaA
		(part #2)	170	4/6	+	AGgTCc
166	cRARE					
		(part #1)	164	3/6	+	GtGgaA
		(part #2)	170	4/6	+	AGgTCc
174	cRARE					
		(part #2)	170	3/6	-	ttTcCA
		(part #1)	176	4/6	-	GGGaCc
176	cRARE					
		(part #2)	170	3/6	-	ttTcCA
		(part #1)	178	4/6	-	GGGggA
176	cRARE					
		(part #2)	170	3/6	-	ttTcCA
		(part #1)	178	4/6	-	GGGggA
182	cRARE					
		(part #1)	180	3/6	+	GaGgCg
		(part #2)	186	4/6	+	AGcTCg
192	cRARE					
		(part #1)	190	3/6	+	cGGaCc
		(part #2)	196	4/6	+	AaTcCA
193	cRARE					
		(part #2)	189	4/6	-	AGcTCg
		(part #1)	195	4/6	-	GGtCCg
194	cRARE					
		(part #2)	189	4/6	-	AGcTCg
		(part #1)	196	3/6	-	tGGtCc



Pos.	Subseq	Loc.	#Res. Matched		Subseq Found
21	cRARE				
	(part #2)	16	4/6	-	gGTTaA
	(part #1)	23	3/6	-	GGcggA
21	cRARE				
	(part #1)	19	4/6	+	ccGCCA
	(part #2)	27	3/6	+	AGcTgc
21	cRARE				
	(part #2)	16	4/6	-	gGTTaA
	(part #1)	23	3/6	-	GGcggA
22	cRARE				
	(part #2)	16	4/6	-	gGTTaA
	(part #1)	24	3/6	-	tGGCgg
22	cRARE				
	(part #2)	16	4/6	-	gGTTaA
	(part #1)	24	3/6	-	tGGCgg
40	cRARE				
	(part #1)	38	4/6	+	cGGtCA
	(part #2)	45	3/6	+	gtcTCA

DraI

1150 1160 1170 1180 1190 1200

TATATATCTT AATTTAAACA TGTTGGTGAC TAAAATATTA TTTAATAAAT ATATATGTTT

DraI

1210 1220 1230 1240 1250 1260

AATGTGTTTT ATTTAAATAC ACTAAAATAT ATTGCCTATA TTCACTGCAT GAGAAATGGA

DdeI DdeI

1270 1280 1290 1300 1310 1320

TAAAATATCA AAATGGGGTC TACTCAGTTA TGCTGAGCAC TGTATATGGC CTTGCAGCGC

1330 1340 1350 1360 1370 1380

TTATGTAGGC TAACTCATTT GTTCATTCAT TTTCTTCGGC TTCTATAAAT TAACTCCTAA

AluI DdeI

1390 1400 1410 1420 1430 1440

AGCTATAACC GAATTGCCTT ATTTGTGAAA TATTAGACTC ATGTAGCCTA AGAAACTGTC

AluI

1450 1460 1470 1480 1490 1500

AGTAAATAAT AAAGTCATTA CAGTAGCTAA TAGCGAATAA TACACCTGTT TATGGTGACT

AluI

1510 1520 1530 1540 1550 1560

GCTCTCAACT AGCTTTGAGG TAAGTCCAGC GACCAGAAAA GTGTAATTAT AGGGCTCGTT

1570 1580 1590 1600 1610 1620

AAAAATACAA AAAAAAAAAA AAAAAAAAGA TATTTAATGT AGGGAAAAAA AAACAAACAA

1630 1640 1650 1660 1670 1680

CTCCTGCAAA CAACAAACCT TATACTAATT TTCATAAAAT CAATCTTTTT TTTCCTGCTA

1690 1700 1710 1720 1730 1740

AAATAAAAAG CCTGTCTTTT GAAGAAGCAT TTCTAAAGTA TATGGAAACT TTTCATGGTT

```
       430         440         450         460         470         480
        *           *           *           *           *           *
AGGCACAATG GAGAGGAATT ATCTCCAGAC CAAAACTTTG GTCTATACGA TACAGATACG

       490         500         510         520         530         540
        *           *           *           *           *           *
GCAGCCTGTT CTATCTAAAA TATTGTATTT TTAATCATTA TTATTATAAT TATTATTATT

       550         560         570         580         590         600
        *           *           *           *           *           *
AAATTAAAGA TGATTATTAT CTTTATAAAA AAGTATTCTC ATGTAAAATA ATTAGACTAA

       610         620         630         640         650         660
        *           *           *           *           *           *
TTGCTACTAT ATATATATAT AAAGGGATAT TTCACTTTGA AAACCAGAAA AAATGACTTT

                                                          TaqI
                                                           |
       670         680         690         700         710         720
        *           *           *           *           *          |*
TCTTTGCCAG GAGCCTGATA AGGAAACTTT TAGATGGCAT TTAGAGGTTT TTGCATCGAC

       730         740         750         760         770         780
        *           *           *           *           *           *
CCAAACTCTC ATACACACAC ACACACACAC ACACACACAC ACACACACAC ACACACACAC

       790         800         810         820         830         840
        *           *           *           *           *           *
ACACACACAC ACACACACAC GTTCCATGTT CATATACATT AGATTAGTCA ATACTGCAAT

                  AluI
                   |
       850         860         870         880         890         900
        *          |*           *           *           *           *
ACTGAAGCAA AATCTGGAGC TTATCTAACA AACTAACTTA TGGTAAACTT ATGTCCAAAA

       910         920         930         940         950         960
        *           *           *           *           *           *
AGTAGTACAC CCTAAATTTA TGTTTTACAA AAAATTACCA ACACACAAAT TGAAAAATAA

       970         980         990        1000        1010        1020
        *           *           *           *           *           *
GAAAAAAGAA AAAAAAGAAA AAAAAAAAAC AAAAAAAGAA ACAATTTTGT TAAAAACTTT

                                                    DraI
                                                     |
      1030        1040        1050        1060       |1070        1080
        *           *           *           *        |  *           *
TATTGTTTTT TGCAATATTT TGCTTGAATT TAATTATATT ATCTTTAAAT TTCTAAACAT

                            AluI
                             |
      1090        1100       | 1110        1120        1130        1140
        *           *        |  *           *           *           *
GTTTGGTGAC TAAAATATTA TTCAGCTTTT GTTCCCTTTA GT------TA TTGCTGATTA
```

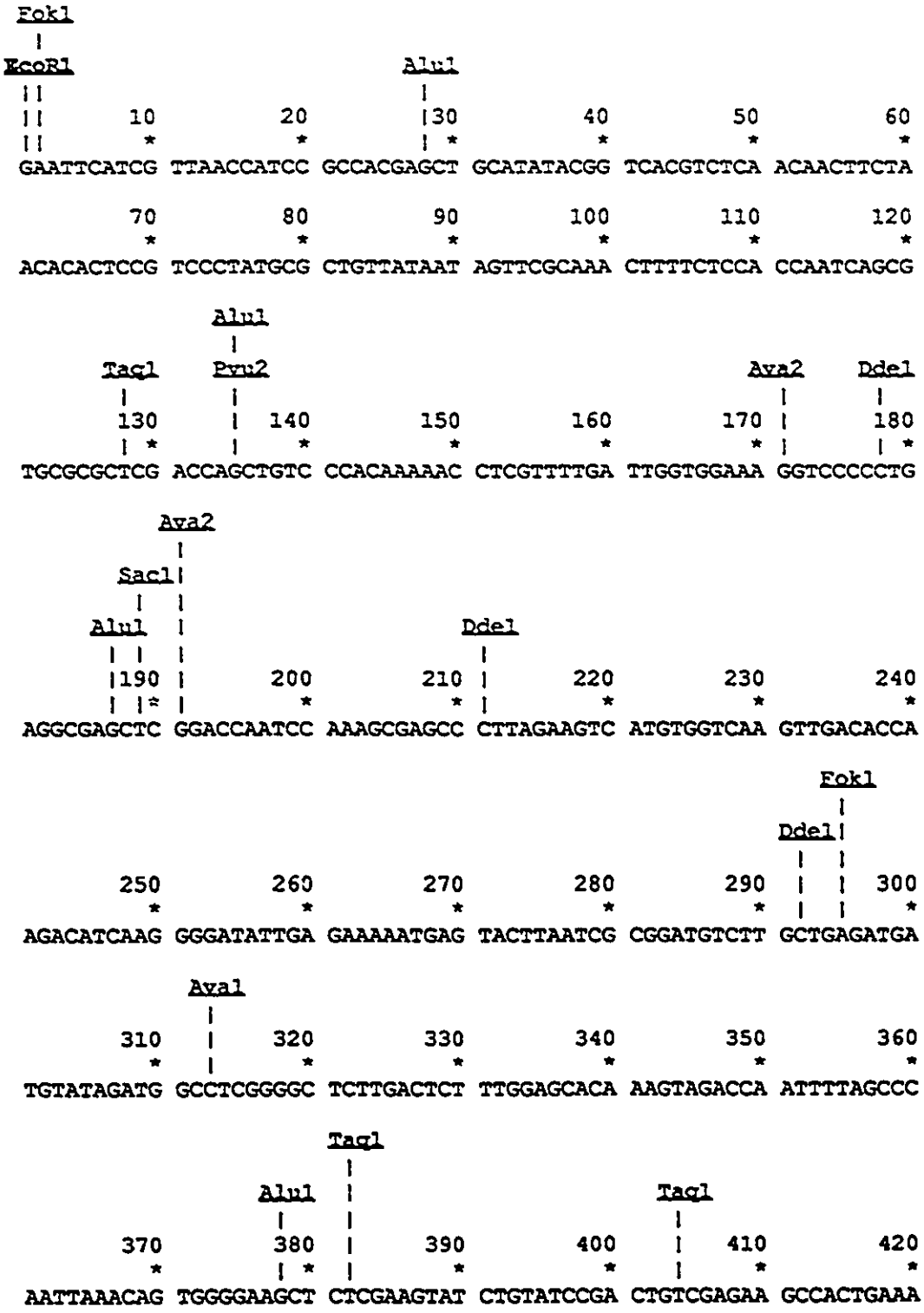

FokI
EcoR1
AluI
10 20 30 40 50 60
GAATTCATCG TTAACCATCC GCCACGAGCT GCATATACGG TCACGTCTCA ACAACTTCTA
70 80 90 100 110 120
ACACACTCCG TCCCTATGCG CTGTTATAAT AGTTCGCAAA CTTTTCTCCA CCAATCAGCG
AluI
TaqI
Pvu2
Ava2
DdeI
130 140 150 160 170 180
TGCGCGCTCG ACCAGCTGTC CCACAAAAAC CTCGTTTTGA TTGGTGGAAA GGTCCCCCTG
Ava2
SacI
AluI
DdeI
190 200 210 220 230 240
AGGCGAGCTC GGACCAATCC AAAGCGAGCC CTTAGAAGTC ATGTGGTCAA GTTGACACCA
FokI
DdeI
250 260 270 280 290 300
AGACATCAAG GGGATATTGA GAAAAATGAG TACTTAATCG CGGATGTCTT GCTGAGATGA
AvaI
310 320 330 340 350 360
TGTATAGATG GCCTCGGGGC TCTTGACTCT TTGGAGCACA AAGTAGACCA ATTTTAGCCC
TaqI
AluI
TaqI
370 380 390 400 410 420
AATTAAACAG TGGGGAAGCT CTCGAAGTAT CTGTATCCGA CTGTCGAGAA GCCACTGAAA

Figure 3-5. *MsxB* genomic 5' sequence.

A 1.9Kb XhoI-EcoRI fragment was isolated and sequenced, only one region of the sequence can not be assembled with the other part, position 1120bp.

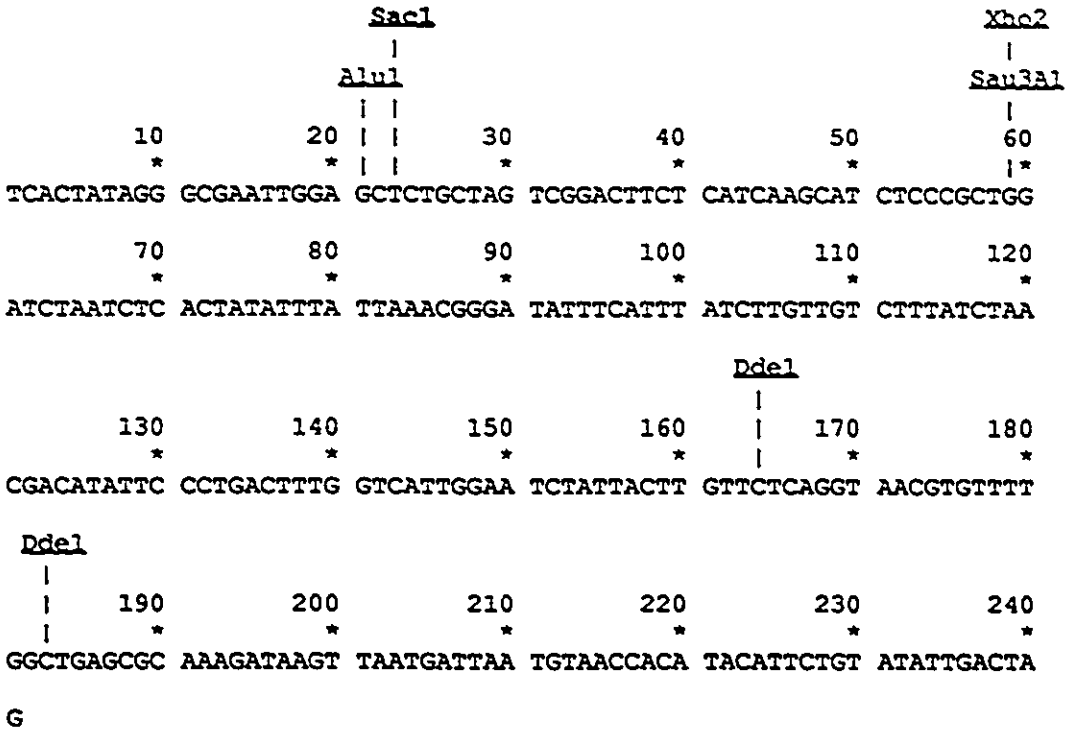

Sac1
Xho2
Alu1
Sau3A1
10 20 30 40 50 60
TCACTATAGG GCGAATTGGA GCTCTGCTAG TCGGACTTCT CATCAAGCAT CTCCCGCTGG
70 80 90 100 110 120
ATCTAATCTC ACTATATTTA TTAAACGGGA TATTTCATTT ATCTTGTTGT CTTTATCTAA
Dde1
130 140 150 160 170 180
CGACATATTC CCTGACTTTG GTCATTGGAA TCTATTACTT GTTCTCAGGT AACGTGTTTT
Dde1
190 200 210 220 230 240
GGCTGAGCGC AAAGATAAGT TAATGATTAA TGTAACCACA TACATTCTGT ATATTGACTA
G

Figure 3-6. *MsxC* sequence corresponding to the intron joining exon 1 and 2. The fragment is an EcoRI-SacI sequence spanning 1.2Kb.

155	mRARErepres				
	(part #2)	151	3/6	-	AGaggA
	(part #1)	157	4/6	-	AaGGCg
218	mRARErepres				
	(part #2)	212	5/6	-	cGTTCA
	(part #1)	220	4/6	-	AaGcCA
244	mRARErepres				
	(part #2)	239	4/6	-	AGTcgA
	(part #1)	246	3/6	-	AtGGag
249	mRARErepres				
	(part #2)	244	3/6	-	gGagCA
	(part #1)	251	4/6	-	gGGaCA
353	oct-B1A-SV				
	(part #1)	353	8/8	-	CTTTGCAT
346	Octa1.4				
	(part #1)	346	8/8	+	ATGCAAAG

Pos.	Subseq	Loc.	#Res. Matched		Subseq Found
15	cRARE				
	(part #2)	9	3/6	-	AacTCt
	(part #1)	17	4/6	-	GGGttA
60	cRARE				
	(part #1)	58	4/6	+	GGGtgA
	(part #2)	64	3/6	+	gGTTtg
60	cRARE				
	(part #1)	58	4/6	+	GGGtgA
	(part #2)	64	3/6	+	gGTTtg
71	cRARE				
	(part #2)	66	4/6	-	AccTCA
	(part #1)	73	3/6	-	GcttCA
184	cRARE				
	(part #1)	182	4/6	+	tGGCtA
	(part #2)	189	3/6	+	AaaTaA
185	cRARE				
	(part #2)	179	3/6	-	AaacCA
	(part #1)	187	4/6	-	taGCCA
218	cRARE				
	(part #2)	212	5/6	-	cGTTCA
	(part #1)	220	4/6	-	aaGCCA
242	cRARE				
	(part #1)	240	4/6	+	GctCCA
	(part #2)	246	3/6	+	tGTcCc
249	cRARE				
	(part #2)	244	3/6	-	gGagCA
	(part #1)	251	5/6	-	GGGaCA
595	ER_half-si				
	(part #1)	595	5/5	-	GGTCA
725	ER_half-si				
	(part #1)	725	5/5	-	GGTCA
8	mRARErepres				
	(part #1)	6	4/6	+	AGttCA
	(part #2)	13	3/6	+	AaccCA
30	mRARErepres				
	(part #2)	24	4/6	-	AcaTCA
	(part #1)	32	3/6	-	AtGaCt
73	mRARErepres				
	(part #1)	71	3/6	+	AGcatA
	(part #2)	78	4/6	+	AGTgCt
102	mRARErepres				
	(part #1)	100	4/6	+	AccGCA
	(part #2)	108	3/6	+	AaTcCc
117	mRARErepres				
	(part #2)	111	3/6	-	gaTTCt
	(part #1)	119	4/6	-	AGatCA
127	mRARErepres				
	(part #1)	125	4/6	+	tGcGCA
	(part #2)	131	3/6	+	AacTgA
147	mRARErepres				
	(part #2)	143	4/6	-	tGTcCA
	(part #1)	149	3/6	-	AGGatg
149	mRARErepres				
	(part #2)	143	4/6	-	tGTcCA
	(part #1)	151	4/6	-	AGaGgA

```
       550         560         570         580         590         600
        *           *           *           *           *           *
AAACAGAGGT GTACAATGTA GATAAATGAA TATAAATCGC TTCTTTACAG TGACCTAGAA

             DraI          DdeI
              |             |
       610    |  620        630         640         650         660
        *     |   *         | *          *           *           *
GTTAAAACTT ATTTAAAGTT TATTTAACTT AGTTTGTAGT TCTGTAATGG GTAATATAAT

       670         680         690         700         710         720
        *           *           *           *           *           *
TATAACAAAC GCCAACAGAT ATTAGTCTTC AGTTGAGGGT CTGTTTAATC AATCGGAACA

       730         740         750         760         770         780
        *           *           *           *           *           *
TGACCGTTTG TTTCTTTGCA ACTTTATGTG ACATGTTGTG GGGTCTTATT GTTAGCCACC

       790         800         810         820         830         840
        *           *           *           *           *           *
CTTATGAATA TTATTCATGG TTTTATTAGT TGTAGTAAAA GTGCAGTAAC CATGTTTTGT

                                                                  DdeI
                                                                   |
       850         860         870         880         890        900
        *           *           *           *           *         | *
GTGTGTTTAC CTTTTGTATA ACCCCAGTTT TGCCACTAAT GCTGGTAACA TTATGGACTG

           AluI
            |
       910  |      920         930         940         950         960
        *   |       *           *           *           *           *
AGGTGAGGTT AGCTTTATAT TCACACATTC AGACTTTCTC CTTAATAATG TATTTTCGGG

       970         980         990        1000        1010        1020
        *           *           *           *           *           *
AAAAGTATTC TTCGCAAATT CTAATTAGCA AAATCATATT AGCAGCCAAT GACTATCAAA

      1030        1040        1050
        *           *           *
GTACACATAT TCACAGTGAT AATAGTGCTA ATGT
```

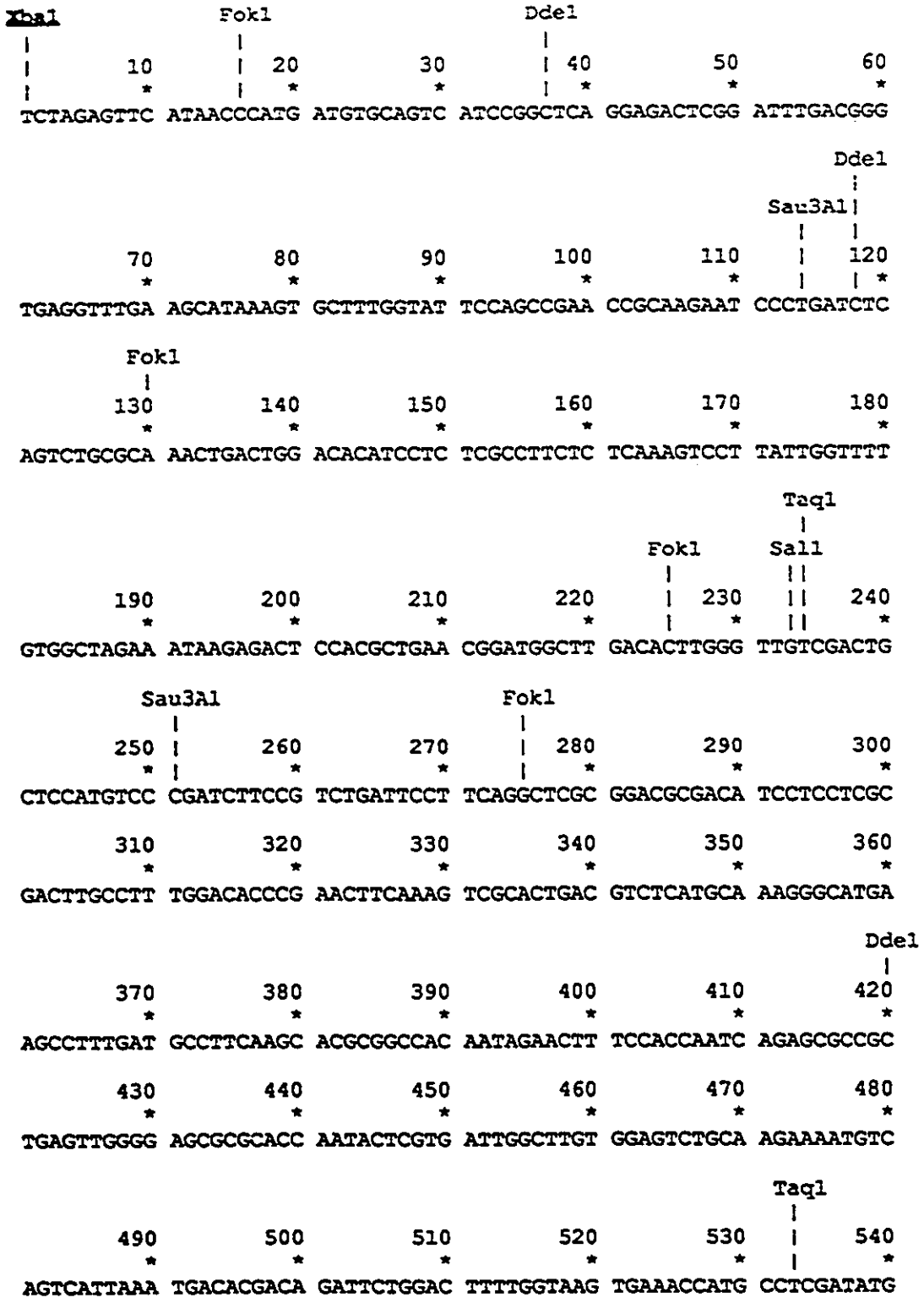

XbaI
FokI
DdeI
10 20 30 40 50 60
TCTAGAGTTC ATAACCCATG ATGTGCAGTC ATCCGGCTCA GGAGACTCGG ATTTGACGGG
DdeI
Sau3AI
70 80 90 100 110 120
TGAGGTTTGA AGCATAAAGT GCTTTGGTAT TCCAGCCGAA CCGCAAGAAT CCCTGATCTC
FokI
130 140 150 160 170 180
AGTCTGCGCA AACTGACTGG ACACATCCTC TCGCCTTCTC TCAAAGTCCT TATTGGTTTT
TaqI
FokI
SalI
190 200 210 220 230 240
GTGGCTAGAA ATAAGAGACT CCACGCTGAA CGGATGGCTT GACACTTGGG TTGTCGACTG
Sau3AI
FokI
250 260 270 280 290 300
CTCCATGTCC CGATCTTCCG TCTGATTCCT TCAGGCTCGC GGACGCGACA TCCTCCTCGC
310 320 330 340 350 360
GACTTGCCTT TGGACACCCG AACTTCAAAG TCGCACTGAC GTCTCATGCA AAGGGCATGA
DdeI
370 380 390 400 410 420
AGCCTTTGAT GCCTTCAAGC ACGCGGCCAC AATAGAACTT TCCACCAATC AGAGCGCCGC
430 440 450 460 470 480
TGAGTTGGGG AGCGCGCACC AATACTCGTG ATTGGCTTGT GGAGTCTGCA AGAAAATGTC
TaqI
490 500 510 520 530 540
AGTCATTAAA TGACACGACA GATTCTGGAC TTTTGGTAAG TGAAACCATG CCTCGATATG

Figure 3-7. *MsxD* 5′ flanking region.
1Kb of 5′ sequence starting at the XbaI site within the first exon.

4. CRANIOFACIAL DYSMORPHOGENESIS CORRELATES WITH A LOSS OF DLX EXPRESSION IN RETINOIC ACID TREATED ZEBRAFISH.

Chapter 4.

4.1. Introduction

Development of the head skeleton depends in part upon interactions between neural crest cells that populate the visceral arches and the overlying epithelial cells of the arch primordia. Neural crest cells originate from the hindbrain rhombomeres and from the mesencephalon. Once they have reached their terminal location within the visceral arch primordia, an interaction between crest cells and the overlying epithelium is required for the ectomesenchyme that originates from neural crest cells to condense and differentiate into cartilage.[49] This phase of cytodifferentiation is then followed by morphogenesis.

The morphology of the cartilage elements formed in a given arch depends on the identity of the neural crest cells that migrate into that arch.[46] Considerable attention has been paid to the genetic mechanisms that regulate this process. Cells from specific rhombomeres will populate defined arches and positional information carried by the neural crest cells seems to be transferred to the arches and will influence the morphology of the skeletal elements. Premigratory neural crest cells of the hindbrain express a rhombomere-specific combination of *Hox* genes which is thought to give them their unique segmental identity along the antero-posterior axis of the hindbrain. [27-29] This unique combination of neural crest cell subpopulations, carry their prepatterned *Hox* information into each visceral arch.

Selective inactivation of *Hox* genes leads to homeotic transformations by which skeletal elements that originate from one arch will take the likeness of the elements that normally originate from another arch. Furthermore, ectopic *Hox* gene expression can lead to craniofacial abnormalities. [32, 33] Some of these craniofacial phenotypes are similar to effects seen in mice following administration of retinoic acid (RA). [18] Such craniofacial defects involve a failure of the frontonasal mass to enlarge and fuse with the nasal processes and maxillae.[59, 60] This defect is classified as severe clefting of the primary palate, or cleft palate.

Retinoic acid is known to directly regulate *Hox* gene expression by binding to regulatory elements found within the *Hox* clusters. The retinoic acid

response elements (RAREs) found 3' to the *Hox* clusters are known to upregulate *Hox* gene expression. However, RAREs have also been found in the mouse and chick *Hoxb1* flanking sequences, and these RAREs were shown to downregulate the expression of *Hoxb1* within the hindbrain.[84]

Although a lot is known about the role of *Hox* genes in conferring positional identity for the different rhombomeres and arches, little is known about other genetic factors that determine which subset of neural crest cells will contribute to the visceral arches and which factors influence the differentiation program of these neural crest cells. We have studied the role of another family of homeobox genes, the *distal-less* (*dlx*) genes, on craniofacial development. It was previously shown that migrating crest cells that will apparently populate the visceral arches express the *dlx2* gene, both in zebrafish and in mice.[94,88] Following crest cell migration, the lateral mesenchymal cells of the visceral arches express a complex combination of several *dlx* genes. In zebrafish, this combination is temporally regulated. [96] At later stages of development, expression of *dlx* genes has been shown in arch mesenchymal cells at the time of condensation and cartilage differentiation. [86, 87, 91] In the present study, we report that treatment of zebrafish embryos with exogenous retinoic acid results in a loss of *dlx* expression in several regions of the embryo but that *dlx* expression in the migrating crest cells and in the visceral arches is particularly sensitive to RA treatment. The loss of *dlx* expression in migrating crest cells and arch primordia correlates closely with the loss or abnormal development of craniofacial cartilage components suggesting an essential role for *dlx* genes in development of the head cartilage.

4.2. **Materials And Methods**

4.2.1 Animals

Adult zebrafish embryos were obtained from a local supplier and were maintained according to standard procedures. [101] Embryos were staged at 28.5°C according to hours (hr) and days (d) post fertilization.

4.2.2 Retinoic acid treatment

A 1 mM stock solution of all-trans retinoic acid (Sigma) was prepared in DMSO. For treatment of zebrafish embryos, RA was adjusted to final concentrations ranging from 0.1 µM to 1 µM in 0.5% DMSO. Embryos were transferred to the retinoic acid solution for two hr after which they were washed twice with tank water. The schedules of RA administration, which extended from late gastrulation to arch segmentation, are shown in Figure 3-1. Control embryos were taken from the same spawning as those treated with RA and left to develop for 2 hr in 0.5% DMSO. RA treated embryos and controls were fixed at 27 hr for whole mount in situ hybridization's, or at 5 d for cartilage staining. All groups contained a minimum of 5 embryos.

4.2.3 Whole-mount in situ hybridizations

Embryos were fixed in a phosphate buffer saline solution containing 4% paraformaldehyde. In situ hybridizations were performed on whole mount embryos using antisense riboprobes as previously described. [102] The templates used for synthesizing the antisense riboprobes for *dlx*2, *dlx*3, and *dlx*4 have been previously described. [96] The template for the *dlx*1 probe consisted of a 900 bp cDNA fragment which will be described in details elsewhere (Ekker et al., unpublished observations).

Double labelling was carried out on whole mount embryos using the same protocol as described above with the following modifications. A fluorescein-12-UTP (#1427857, Boehringer Mannheim) probe was synthesized like the digoxygenin (DIG) probe, however the DIG precipitation step was omitted. Instead, the fluorescein probe was filtered through a Nuctrap TM Push Column (Stratagene). Hybridization was carried out synchronously using both probes (digoxygenin and fluorescein). Once reactions with the anti-DIG antibody and alkaline phosphatase detection were complete, but prior to post-fixation, the embryos were immersed for 10 min in glycine-HCl (pH 2.2)/0.1% Tween-20 to inactivate the anti-digoxygenin antibody. The embryos were then washed four times for 5 min with 1X phosphate buffered saline/0.1% Tween-20 (PBST) and incubated with the antifluorescein antibody conjugated to alkaline

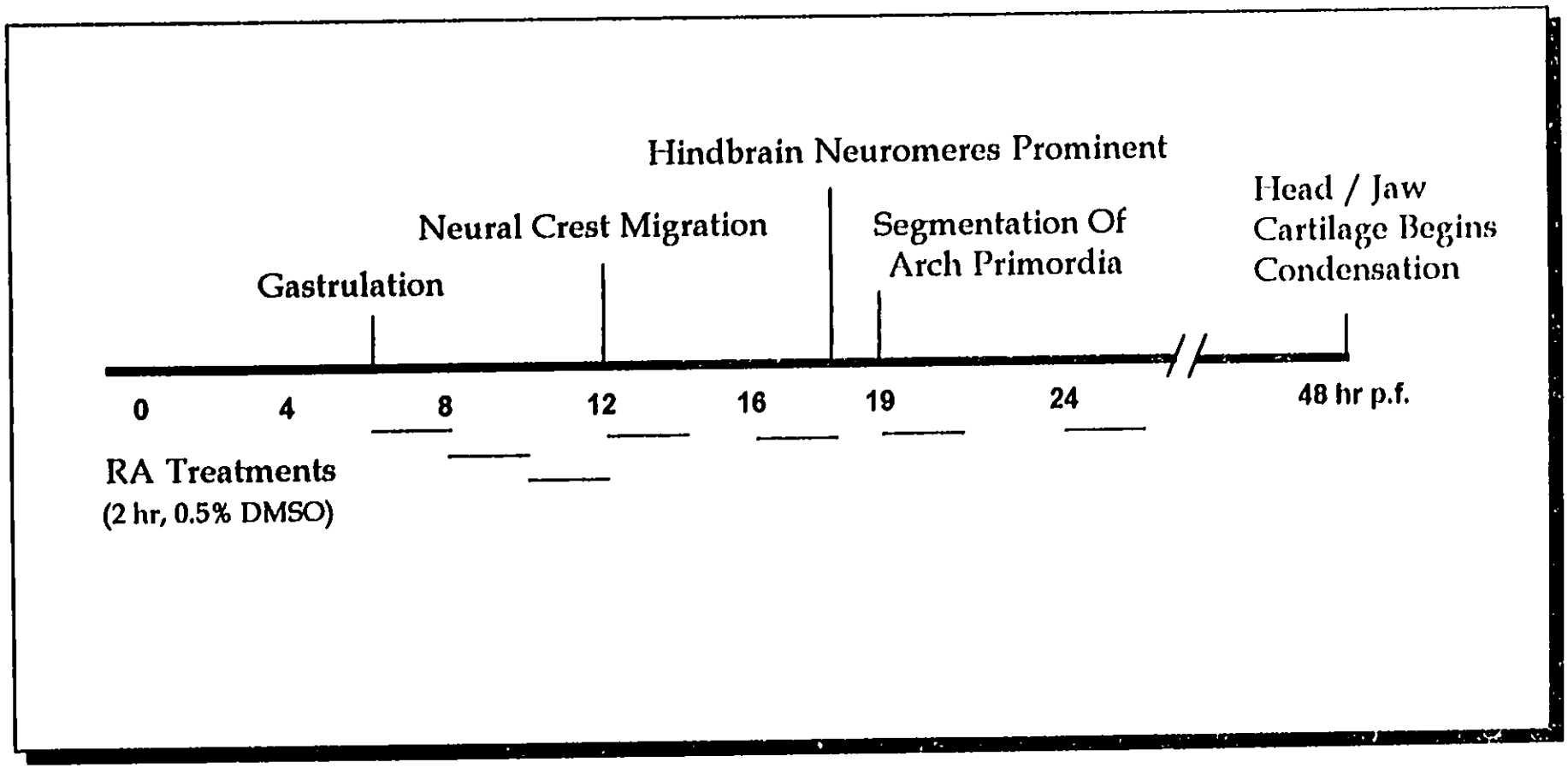


Figure 1. Retinoic acid administration schedule.

Figure 4-1: Retinoic Acid Treatment Schedule.

The horizontal line is representative of zebrafish embryonic hours post fertilization from 0 to 48 hr. Above the line are developmental events that happen during the corresponding developmental stages. Below the developmental stages are shorter lines which correspond to the RA treatment periods. The lines beneath the RA treatment schedule depict important visceral arch developmental stages.

phosphatase(#1426338, Boehringer). The embryos were washed 6 times for 15 min in PBST, and twice for 5 min in 0.1M Tris-HCl (pH 8.2)/0.1% Tween-20. Detection of the fluorescin probe was done using two tablets of Fast Red (#1496549, Boehringer) dissolved in 4ml of 0.1M Tris-HCl (pH 8.2)/0.1% Tween-20. A 2 hr post fixation in 4% paraformadehyde followed by a wash with PBS were carried out. The embryos were then stored at 4°C in PBS/ 5mM sodium azide.

4.2.4 Cartilage Staining

Embryos were grown in a 0.2mM solution of PTC (Phenylthiocarbamide, Sigma) in tank water, to inhibit pigment formation and were stained with alcian blue. [66] Briefly, 5 d larvae were fixed in 4% paraformaldhyde and washed as above. They were stained for 1-12 hr in a filtered 0.1% alcian blue solution in 30% acetic acid, 70% ethanol, rehydrated in a water-ethanol series, and cleared with a 5% KOH solution overnight at 4°C. Stained larvae were kept in H_2O at 4°C.

4.3. RESULTS

4.3.1 *Dlx* Expression During Visceral Arch Cartilage Differentiation.

A specific temporal and spatial combinatorial expression of zebrafish *dlx* genes during development of the visceral arches up to 55-60 hr has been reported. [96] To better correlate patterns of *dlx* gene expression with the development of cartilage elements, we examined *dlx* expression in visceral arches at the time of arch differentiation. Craniofacial cartilage begins to condense at 2 d in zebrafish embryos. At this time, expression patterns of *dlx* genes can be seen in the mandibular, hyoid, and branchial arches (Fig. 4-2 A, B). From 2.5 d, the visceral arch primordia express *dlx* genes in areas that correlate with cranial crest-derived chondrocyte condensations. Thus, during the third day, cells in the position of the developing Meckel's cartilage express *dlx*2 (Fig. 4-2C) and *dlx*3 (Fig. 4-2D). Cells forming the quadrate cartilage express *dlx*1 (not shown), *dlx*2 (Fig. 4-2C), and *dlx*3 (Fig. 4-2D); cells forming the ceratohyal cartilage express*dlx*1 (not shown), *dlx*3 (Fig. 4-2D), and *dlx*4 (not shown); and

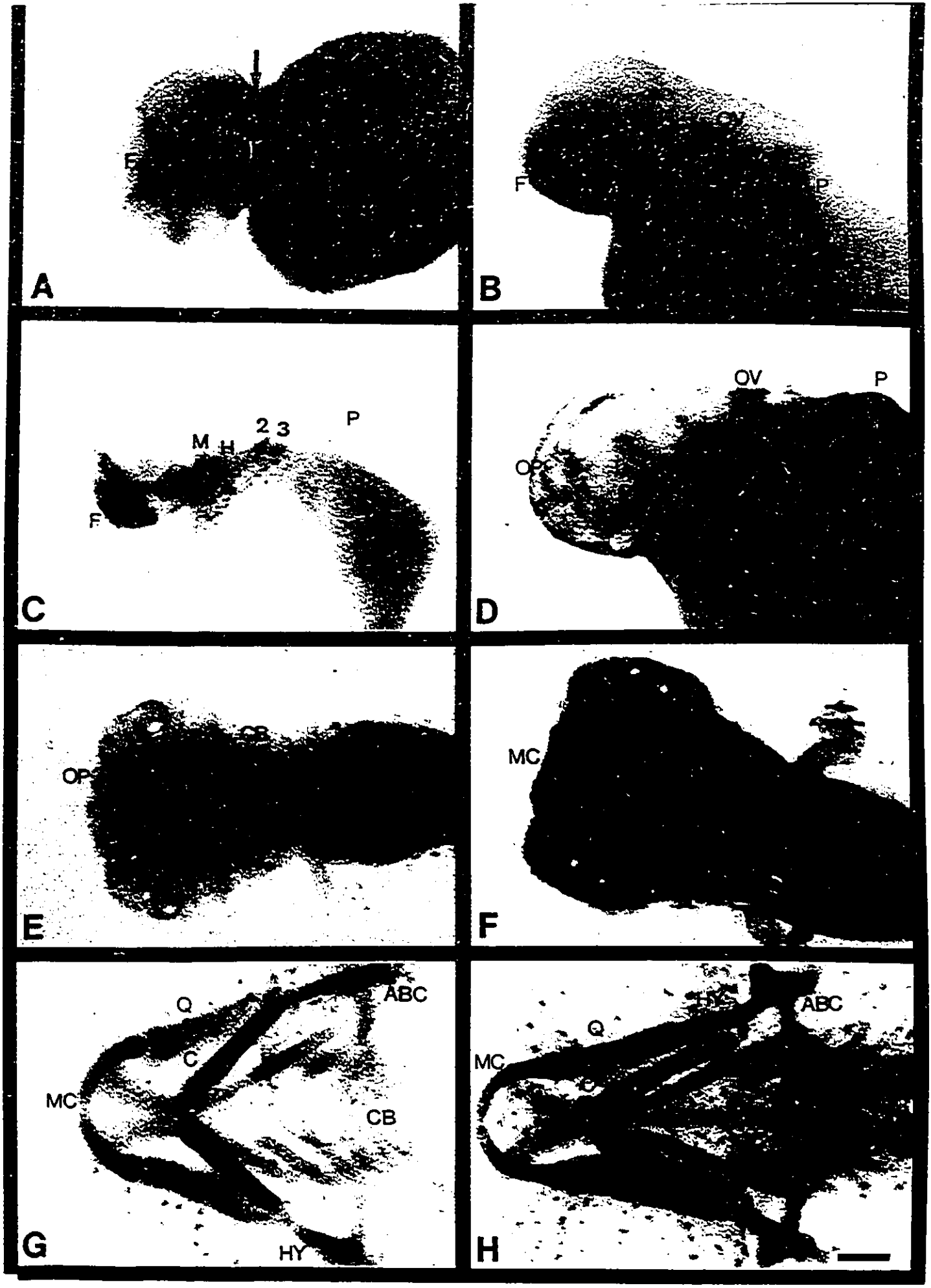

A
F
B
F
P
C
M
H
2
3
P
F
D
OV
P
OP
E
CB
OP
F
MC
G
Q
ABC
C
MC
CB
HY
H
HY
ABC
Q
MC

Figure 4-2: *Dlx* Expression Correlates With Neural Crest Derived Chondrocyte Condensations. A) ventral view of a 2 d embryo showing *dlx*4 gene expression in the visceral cleft (arrow), B) a sagittal view of the same embryo showing *dlx*4 expression in the mandibular, hyoid and in branchial arches 1, 2, and 3. C) saggital view of *dlx*2 expression in the visceral cleft (M, H) and branchial arches 2 and 3 of a 2.5 d embryo. D) *Dlx*3 expression at 3 d. E)*dlx*4 expression in a 4 d larvae. F) *Dlx*3 expression in a 5 d larvae. (G- H) Craniofacial cartilage elements derived from the visceral arches at 4 d (G), and 7 d (H). F, forebrain; M, mandibular arch; H, hyoid arch; 1-3, branchial arches 1 to 3; OV, otic vesicle; P, pectoral fin bud; OP, olfactory placode; PT, presumptive pterygoid; MC, Meckel's cartilage; Q, quadrate; HY, hyosymplectic; CB, ceratobranchials; O, operculum; C, ceratohyal; ABC, anterior basicranial commissure. Scale bar: 107 µm for A-C & E; 83 µm for D and F; 51µm for G; and 61µm for H.

cells forming the hyosymplectic cartilage express *dlx*1 (not shown) and *dlx*3 (Fig. 4-2D). *Dlx*3 is also expressed in the presumptive pterygoid cartilage(Fig. 4-2D).

Morphogenesis of the various cartilage elements continues during the fourth day and the trabeculae, basihyal, quadrate, hyosymplectic, and ceratohyal components can be seen elongating. Meckel's cartilage now articulates with the anterior tip of the quadrate, and the hyosymplectic articulates with the anterior basicranial commisure (Fig. 4-2G). The basibranchial and ceratobranchial cartilage components become partially visible between 4 and 5d. During this time, *dlx* expression is progressively down-regulated and transcripts are only detectable in areas that correspond to those cartilage elements that condense last. Thus, *dlx*2 expression is no longer detectable after 3 d and *dlx*4 expression is then restricted to the ceratobranchials (Fig. 4-2E). Only *dlx*3 expression is detectable at 5 d. *Dlx*3 transcripts are detected in the chondrifying Meckel's cartilage, ceratobranchials, operculum, as well as the circumoral area (Fig. 4-2F). The craniofacial cartilage components are well developed at 7 d (Fig. 4-2H). *Dlx* gene expression in visceral arch primordia is particularly sensitive to exogenous retinoic acid.

In addition to the primordia of the visceral arches, cells in the ventral forebrain, olfactory placodes, otic placode and vesicle, pectoral fin buds, and median fin fold express multiple *dlx* genes.[94] However, embryos that were treated with a dose of 1 µM RA at 6 hr show a complete loss of all *dlx* gene expression in the head region when measured at 24 hr (Fig. 4-3 A, B) while weak expression of *dlx*3 persists at the caudal end of the developing median fin fold (Fig. 4-3B). Embryos treated with the same RA dose at either 8 hr or at 12 hr show similar results (not shown). When the dose of RA is reduced to 0.1 µM and administered at either 6 hr, 8hr, or 12 hr, expression of *dlx* genes in structures such as the ventral forebrain, the otic vesicle, and the olfactory placodes is progressively reestablished in a dose-dependent manner, although the morphology of the structures containing *dlx*-expressing cells is sometimes abnormal (data not shown). However, RA treatments which have no effect on *dlx* gene expression in areas such as the olfactory placodes and otic vesicle, show an abolishment of *dlx* expression in the primordia of the visceral arches (Fig. 4-3C, D). Furthermore, in embryos treated at 10 hr hr with 0.1 µM RA, the otic vesicle is located in a position closer towards the optic vesicle, consistent

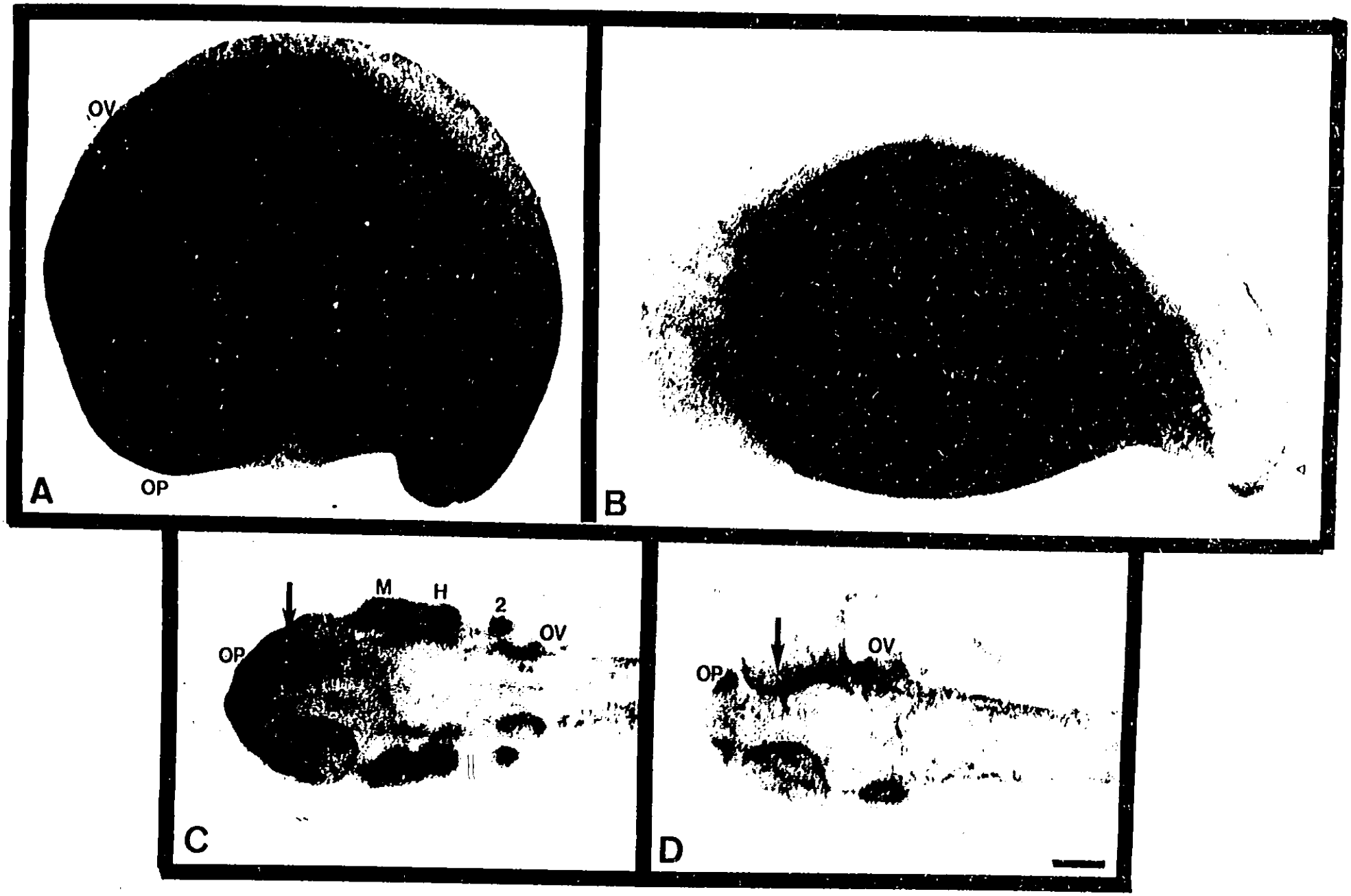

OV
OP
A
B
M
H
2
OV
OP
C
OV
OP
D

Figure 4-3: Effects of early RA treatments on *dlx* 3 expression.
Panels (A) and (C) are control embryos (DMSO-treated) that show the normal patterns of *dlx*3 expression at 16 hr (A) and 27 hr (C). (B), embryo treated with 1 µM RA at 6 hr. *Dlx*3 expression was determined at 27 hr and is undetectable in the olfactory placode area, otic vescicle and visceral arches. Some *dlx*3 transcripts are detected at the caudal end of the median fin fold (arrowhead). Development of RA-treated embryos is severely retarded. Compare the embryo in (B) with those in (A) and (C). (D), embryo treated treated with 0.1 µM RA at 10 hr. *Dlx*3 expression in the viseceral arches is lost whereas it is reduced in the olfactory placodes and normal in the otic vesicle. Note that the distance between the otic vesicle and the optic cup (arrow) is shorter than in control embryos (C). Other symbols are as described in Fig. 4-2 . Scale Bar: 79 µm for A, 74 µm for B, 24 µm for C and D.

with the loss of the midbrain-hindbrain border region that was previously described. [64, 65] The same RA treatment also results in selective loss of *dlx2* and *dlx4* gene expression in the visceral arch primordia (not shown), but neither in the ventral forebrain (*dlx2* and *dlx4*) nor in the olfactory and otic placodes (*dlx4*). Therefore expression of *dlx* genes in visceral arch primordia seems to be more sensitive to RA treatment.

4.3.2 RA treatment affects *dlx2* expression in migrating neural crest cells

Zebrafish hindbrain neural crest cells express *dlx2* during their migration into the visceral arches from 12 hr to 19 hr. [96] To determine whether the loss of *dlx* gene expression in the visceral arch primordia can be related to altered *dlx2* expression in migrating neural crest cells, we treated embryos with varying doses of RA before and during neural crest cell migration. To better localize the position of *dlx2*-expressing cells we double labeled embryos with *dlx2* and with either *Krox20*, a gene expressed in cells of rhombomeres 3 and 5 or with *eng2*, which is expressed in cells at the midbrain-hindbrain border. [111, 112] In control embryos, neural crest cells expressing *dlx2* migrate into the periphery along pathways adjacent to rhombomeres 1, 2, 4 and 6 (Fig. 4-4A, C, E). [96] Furthermore, double-labelling with *dlx2* and *eng2* shows that cells that will populate the mandibular arch primordium originate from the hindbrain but that their pathway deviates rostrally (Fig. 4-4C).

In embryos treated at 12 hr with 1 µM RA, *dlx2* expression, measured at 14 h, is lost (Fig. 4-4F). Similar results are obtained when the embryos are treated at either 8 hr or at 10 hr (not shown). Krox 20 expression in rhombomeres 3 and 5 is also abolished by this treatment. At a dose of 0.1 µM RA, administered at either 8 hr (Fig. 4-4D), 10 hr (not shown) or 12 hr (Fig. 4-4B), *dlx2* expression is no longer detected in migrating crest cells when measured at 14 hr. *Krox20* expression is still detected in embryos that received this dose of RA, although earlier treatments (8 hr and 10 hr) result in one or two diffuse stripes of *Krox20*-expressing cells (Fig. 4-4D). *Krox20* expression appears to be normal in embryos treated at 12 hr (Fig. 4-4B). When 0.1µM or 1 µM RA was administered at either 14 hr or at 16 hr, *dlx2* expression is restricted to thin stripes of cells, one anterior and one posterior to the otic vesicle (data not shown).

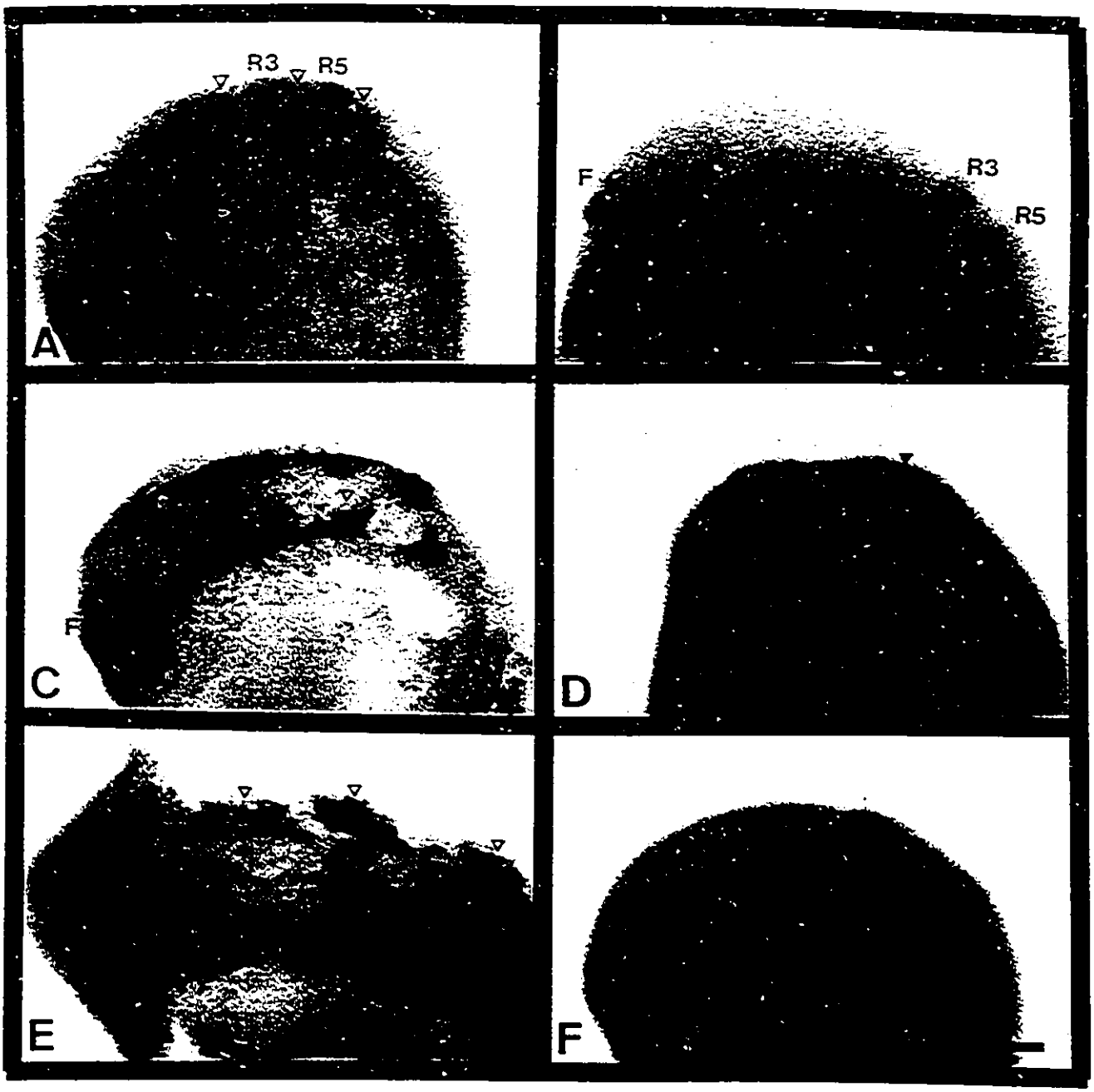

R3
R5
A
F
R3
R5
C
D
E
F

Figure 4-4: RA effects on *dlx2* expression in migrating neural crest cells
Panels (A), (C), and (E) are control embryos. A) 14hr embryo showing *dlx2* and *Krox20*, (C) *dlx2* and *eng2* at 16hr, and (E) dorsal view of a 20hr embryo showing *dlx2* and *Krox20*. Panels (B), (D), and (F) show *dlx2* and *Krox20* expression at 14hr in embryos that were treated with RA. (B) treated at 12hr with 0.1 µM RA, (D) treated at 8hr with 0.1 µM RA, and (F) treated at 12hr with 1 µM RA. A 1 uM RA treatment leads to a loss of both *dlx2* and *Krox20* expression, whereas a 0.1 uM treatment leads to abnormal *Krox20* and a loss of *dlx2* expression. F, forebrain; R3, rhombomere 3; R5, rhombomere 5; hollow arrowhead, *dlx2* expression in migrating neural crest cells; solid arrow head, *eng2* expression in panel (C), *Krox20* expression in panels (D) and (E). Scale bar: 100 µm for A, C, and F; 68 µm for B; 75 µm for D; and 89 µm for E.

4.3.3 RA affects *dlx* expression and cartilage development in the visceral arches.

4.3.3.1 Mandibular Arch

Zebrafish embryos treated with 1 µM RA at 10 hr, 12 hr, 16 hr, 19 hr, and 24 hr, all display a loss of *dlx*2, *dlx*3, and *dlx*4 expression in the mandibular arch (Table I, Fig 3-5D). At a dose of 0.1 µM, we observe a loss of *dlx*3 expression in the medial region of the arch regardless of the schedule of RA administration (Table I), whereas *dlx*2 expression was found to be weaker than controls but retained across the whole arch (Table I, Fig 3-5B, C). Intermediate doses ranging from 0.2 µ M to 0.8 µM result in a loss of expression for all *dlx* genes in the mandibular arch (data not shown).

The two cartilage components derived from the mandibular arch, the quadrate and Meckel's cartilage, are also lost in embryos treated with 1 µM RA (Table I, Fig. 4-5H). At a dosage of 0.1 µM RA, 5 d larvae show a loss of cartilage components which is dependent upon the schedule of RA administration. Thus, embryos treated at 12 hr, 16 hr and 19hr show either a partial loss or altered morphology of the mandibular cartilage components (Table I, Fig. 4-5F, G), whereas embryos treated at 24 hr display no loss of cartilage (Table I). Intermediate doses (0.2 µM to 0.8 µM) result in the loss of both the quadrate and Meckel's cartilage (data not shown).

4.3.3.2 Hyoid Arch

Embryos treated with 1 µM RA display a loss of expression of *dlx*2, *dlx*3, and *dlx*4 in the hyoid arch regardless of the schedule of administration (Table II, Fig. 4-5 D). Embryos that were treated with 0.1 µM RA show an overall weaker *dlx* gene expression even though the expression patterns sometimes appear to be normal. When we observe abnormal *dlx* expression patterns, they are often restricted to either the medial or lateral region of the arch. Thus, expression of *dlx*2 and *dlx*3 is restricted to the lateral part of the arch in embryos treated either at 12 hr (Fig. 4-5B) or at 24 hr for *dlx*2 , and in embryos treated at either 16 hr or at 19 hr for *dlx*3. Interestingly, in embryos treated at 24 hr, *dlx*3

Table I. Mandibular Arch *Dlx* Expression And Cartilage Development Following Retinoic Acid Treatment

		Retinoic Acid Administration Dose & Stage								
	Wildtype *Dlx* Expression	10^{-7}M 12h	10^{-7}M 16h	10^{-7}M 19h	10^{-7}M 24h	10^{-6}M 10h	10^{-6}M 12h	10^{-6}M 16h	10^{-6}M 19h	10^{-6}M 24h
Mandibular *Dlx* Expression	*dlx2*	+	+	+	+	-	-	-	-	-
	dlx3	+ L	+ L	+ L	+ L	-	n.a.	-	-	-
	dlx4	n.a.	-	+	n.a.	-	n.a.	-	-	-
Number of embryos displaying phenotype		6/9	21/25	20/31	9/11	5/5	5/5	27/27	13/13	8/8
Quadrate		Curved	Asy.	Asy.	Normal	Absent	Absent	Absent	Absent	Absent
Meckel's Cartilage		Partial	Partial	Asy.	Normal	Absent	Absent	Absent	Absent	Absent

Table I. *Dlx* expression was analyzed at 27 h. At this time, only *dlx2, dlx3* and *dlx4* are expressed in the mandibular arch. *Dlx* expression in the mandibular arch is; +, normal; +L, present in the lateral half of the arch; -, absent; or n.a., not assessed. Gene expression was determined on a minimum of five embryos per gene and was always identical for all embryos in a given group. The craniofacial cartilage was analyzed at 5 d. The mandibular cartilage phenotype is characterized from "normal" to "absent". "Partial" cartilage component refers to the component being partially present, whereas "curved" refers to the component being present but abnormally curved bilaterally (ventral), and a "Asy." component is bilaterally asymmetrical. The preponderant phenotype for each treatment is shown and the number of embryos displaying this phenotype is indicated. Embryos that did not show the preponderant phenotype often resembled those that were treated with the same dose of retinoic acid in immediately earlier or immediately later stages.

Table II. Hyoid Arch *Dlx* Expression And Cartilage Development Following Retinoic Acid Treatment

		Retinoic Acid Administration Dose & Stage								
	Wildtype *Dlx* Expression	10^{-7}M 12h	10^{-7}M 16h	10^{-7}M 19h	10^{-7}M 24h	10^{-6}M 10h	10^{-6}M 12h	10^{-6}M 16h	10^{-6}M 19h	10^{-6}M 24h
Hyoid *Dlx* Expression	*dlx2*	+ L	+	+	+ L	-	-	-	-	-
	dlx3	-	+ L	+ L	+ M	-	n.a.	-	-	-
	dlx4	n.a.	-	+	n.a.	-	n.a.	-	-	-
Embryos displaying phenotype										
Ceratohyal		Absent	Partial	Asy.	Normal	Absent	Absent	Absent	Absent	Absent
Hyosymplectic		Absent	Partial	Abnormal	Normal	Absent	Absent	Absent	Absent	Absent
Basihyal		Absent	Absent	Absent	Normal	Absent	Absent	Absent	Absent	Absent

Table II. *Dlx* expression in the hyoid arch was analyzed at 27 h. At this time, only *dlx2*, *dlx3*, and *dlx4* genes are expressed. A "+M" designates the *dlx* expression pattern was found in the medial half of the arch. "Abnormal" cartilage phenotype refers to the cartilage element having an abnormal positional symmetry. The craniofacial cartilage was analyzed at 5 d. The other symbols and phenotypes are defined as in Table I. A minimum of five embryos were analyzed for each attribute shown in the table.

expression in the hyoid arch is shifted medially (Table II). *Dlx* gene expression within the hyoid arch was abnormal in embryos that received intermediate doses (data not shown).

At RA doses higher than 0.1 μM, the ceratohyal, basihyal, and hyosymplectic cartilage components are lost irrespective of the time at which RA is administered (Table II, Fig. 4-5H). A dose of 0.1 μ M RA, administered at 10 hr or 12 hr, also results in the loss of both cartilage elements (Table II, Fig. 4-5F) whereas their morphology is abnormal in embryos treated at either 16 hr or at 19 hr with the same dose (Fig. 4-5G). Treatment at 24 hr has no effect on either the ceratohyal or the hyosymplectic cartilage elements. The basihyal is lost in embryos that received intermediate doses (0.2 μM to 0.8 μM) at 16 hr, and in all embryos treated with 0.1 μM RA (Fig. 4-5 F, G) except for those embryos treated at 24 hr.

4.3.3.3 Branchial Arches

Development of the branchial arch cartilage seems to be less sensitive to RA treatment but displays a pleiotropic response. In embryos treated with 1 μM RA at 10 hr, expression of all *dlx* genes is abolished in the branchial arch primordia when measured at 27 hr (Table III). However, cells in the branchial arch primordia of embryos that received the same dose of RA at intermediate times (12 hr, 16 hr, 19 hr) express a subset of *dlx* genes, with *dlx*1 and *dlx*4 being more sensititve. Exposure to RA at 24 hr again abolishes expression of all *dlx* genes. Embryos treated with 0.1 μM RA at embryonic stages varying from 10 hr to 24 hr show a complex loss of *dlx* expression in a subset of arches (Table III). Expression in branchial arches 1 and 2 seems to be sensitive to RA treatments in manners that are dependent upon the gene and the schedule of administration. Expression of *dlx*2 in branchial arches 2, 3 and 4 is more sensitive to early RA treatments (Table III, Fig 3-5B, C).

The two cartilage components which have developed sufficiently at 5 d, the time at which embryos were fixed for cartilage staining, are the basibranchials and the ceratobranchials. At a dose of 1 μM RA , the branchial arch cartilage derivatives are completely absent (Table III, Fig. 4-5H). As the RA dosage decreased to 0.1 μM RA, the cartilage components are retained in a dose dependent manner (Table III). With a dose of 0.1 μM RA, we observe a loss of

Table III. Branchial Arch *Dlx* Expression And Cartilage Development Following Retinoic Acid Treatment

		Retinoic Acid Administration Dose & Stage								
	Wildtype *Dlx* Expression	10^{-7}M 12h	10^{-7}M 16h	10^{-7}M 19h	10^{-7}M 24h	10^{-6}M 10h	10^{-6}M 12h	10^{-6}M 16h	10^{-6}M 19h	10^{-6}M 24h
Branchial Arch 1 *Dlx* Expression	*dlx1* *dlx2* *dlx4*	n.a. + L n.a.	- + +	n.a. + +	- - n.a.	n.a. - -	n.a. + n.a.	- + -	n.a. + -	- - -
Branchial Arch 2 Dlx Expression	*dlx2* *dlx3* *dlx4*	- - n.a.	+ + L +	+ + L +	+ - n.a.	- - -	+ n.a. n.a.	+ + +	+ + +	- - -
Branchial Arch 3 *Dlx* Expression	*dlx2*	-	-	-	+	-	+	+	+	-
Branchial Arch 4 *Dlx* Expression	*dlx2*	-	-	-	+	-	+	+	+	-
Embryos displaying phenotype										
Basibranchial & Ceratobranchials		Absent	Abnormal	Abnormal	Normal	Absent	Absent	Absent	Absent	Absent

Table III. *Dlx* expression was determined at 27 h. The symbols are defined as in Table I. At 27 h, *dlx3* expression is only found in branchial arch 2; *dlx1* expression is only found in branchial arch1; *dlx4* expression in branchial arches 1 and 2, and *dlx2* expression is found in all of the branchial arches. The craniofacial cartilage was analyzed at 5 d. The phenotypes are defined as in Table I and II. A minimum of five embryos were analyzed for each attribute shown in the table.

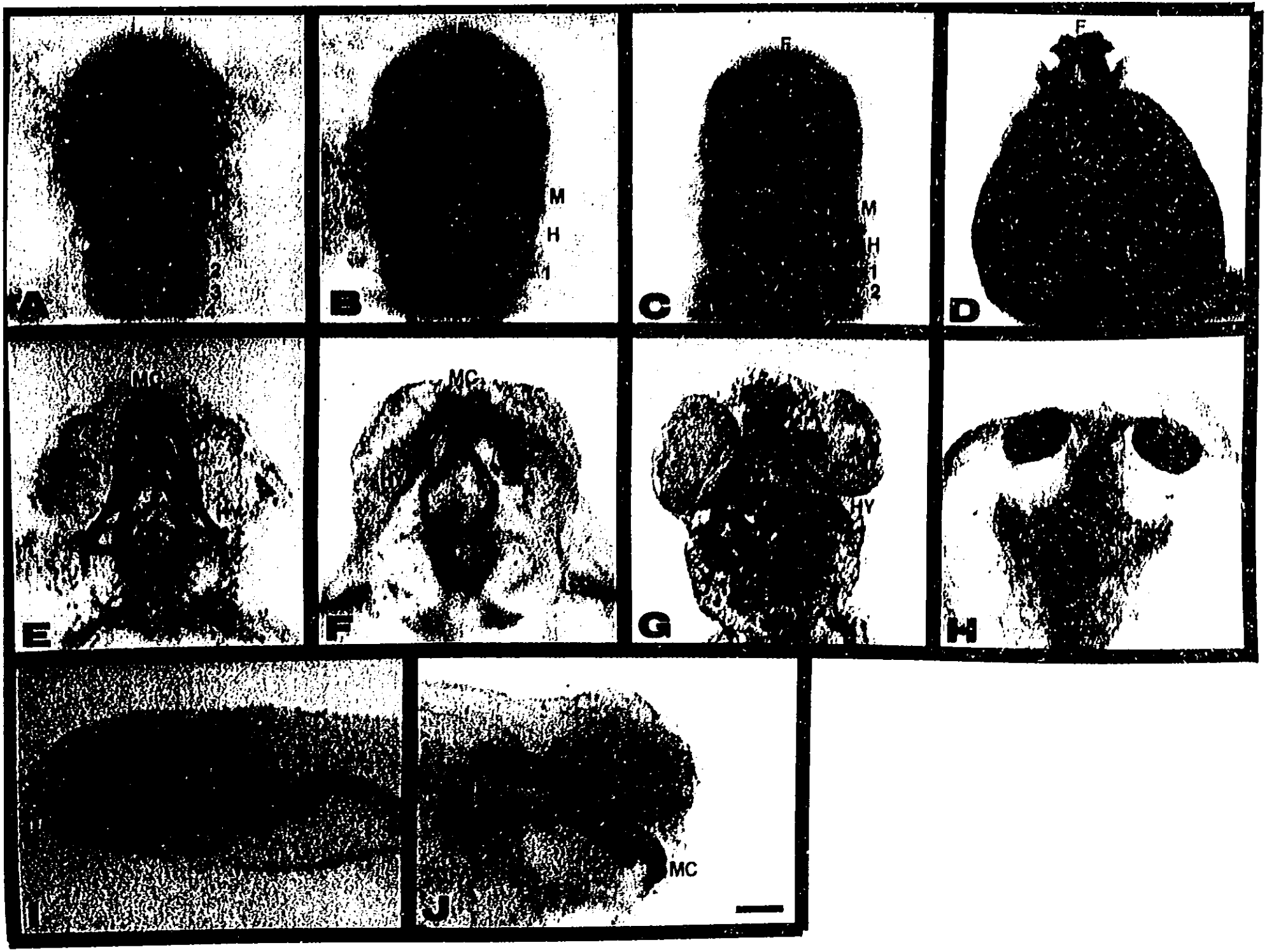

A
B
M
H
1
C
M
H
1
2
D
E
MC
F
MC
G
HY
H
I
J
MC

Figure 5. J

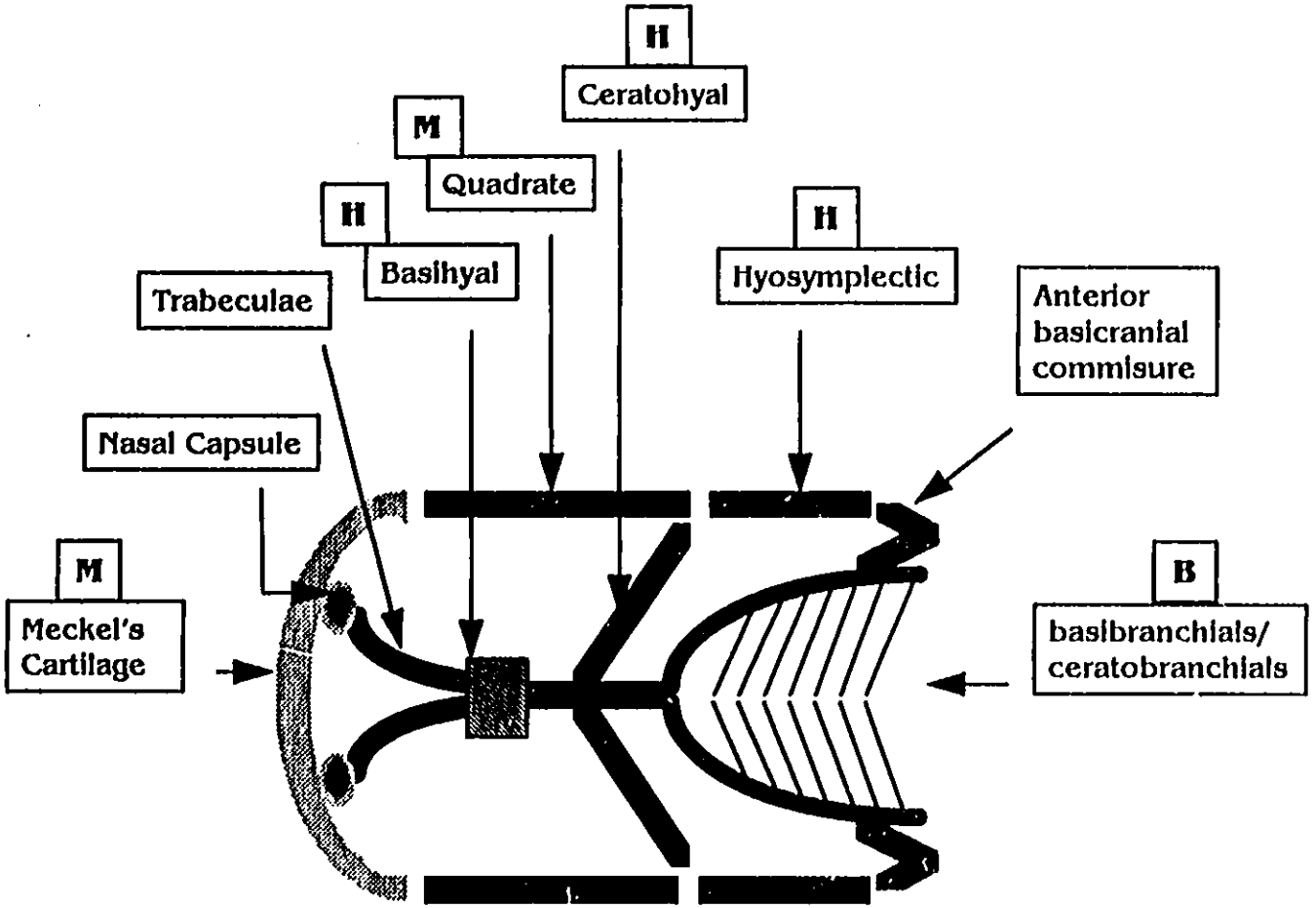

H
Ceratohyal
M
Quadrate
H
Basihyal
Trabeculae
H
Hyosymplectic
Anterior basicranial commisure
Nasal Capsule
M
Meckel's Cartilage
B
basibranchials/ ceratobranchials

Figure 4-5: RA affects visceral arch *dlx* expression and cartilage development . Panels (A-D) show *dlx2* expression in embryos at 27 hr, panels (E-J) show craniofacial cartilage from 5 d embryos. Panels (A) and (E) are control embryos. Panels (B) and (F) are embryos treated at 12 hr with 0.1 µM RA. Panels (C) and (G) embryos treated at 19 hr with 0.1 µM RA. A 0.1 uM RA treatment leads to abnormal mandibular & hyoid *dlx* expression and cartilage differentiation, but most of all a treatment dependent loss of branchial arch *dlx* expression and cartilage differentiation. Panels (D) and (H) are embryos treated at 16 hr with 1 µM RA. A 1 uM RA treatment results in a loss of mandibular & hyoid *dlx* expression and cartilage differentiation. (I) sagittal view of a wildtype embryo showing craniofacial cartilage elements. (J) sagittal view of an embryo treated at 14 hr with 0.1 µM RA showing abnormal craniofacial cartilage structure. (K) ventral view schematic representation of all neural crest derived craniofacial cartilage components. M, mandibular arch; H, hyoid arch; B, branchial arches. Other symbols as in Fig. 4-2. Scale bar: 147 µm for A, 109 µm for B, 90 µm for C, 119 µm for D, 159 µm for E, 145 µm for F, 175 µm for G & J, 133 µm for H, and 114 µm for I.

cartilage derivatives in embryos that were treated at 12 hr (Fig. 4-5F), whereas those treated at later stages with the same dose display normal morphology but show an apparent posterior shift in their position at the time of differentiation (Fig. 4-5G).

4.4. DISCUSSION

Treatment of zebrafish embryos with retinoic acid at various stages of their development affects both the expression of *dlx* genes in the primordia of the visceral arches and the morphology of the cartilage that originates from these arches. We observe a strong correlation between the effects of a particular RA treatment on *dlx* gene expression and cartilage morphogenesis. When zebrafish embryos are treated with 1 µM RA before neural crest cell migration, expression of *dlx2* is no longer observed in the migrating crest cells which will populate the visceral arches and subsequently, expression of all *dlx* genes is lost in the primordia of all visceral arches. These treatments also result in loss of cartilage components or severe morphological abnormalities. When the same dosage is administered during or immediately after neural crest cell migration into the arches, expression of *dlx* genes is again abolished in the mandibular and hyoid arches with a subsequent loss of cartilage components. However, expression of *dlx* genes is less severely affected in the branchial (gill) arches and this loss is correlated with more moderate effects on cartilage formation. Lower doses of RA have correspondingly decreased effects, both on the expression of *dlx* genes and on cartilage formation. Nevertheless, the effects of RA on gene expression and on cartilage formation remain coupled over the dose range and with the different schedules of administration.

It is not clear how treatment with RA impairs expression of *dlx* genes. We cannot at this time exclude the possibility that loss of gene expression is simply the result of absence of the cells that normally express *dlx* genes. However, we consider this unlikely because treatments with the higher doses of RA result in the loss of expression of *dlx* genes in several structures in addition to the visceral arches. These include the presumptive olfactory placodes, the otic vesicle, and the ventral forebrain. Interestingly, the only region of the embryo where *dlx* expression patterns does not seem to be affected is the median fin

fold. This suggests that RA's effects on *dlx* expression are restricted to specific tissues or to populations of cells such as in the neural tube, head neural crest and placode derivatives.

An intriguing finding is that expression of *dlx* genes and cartilage formation are affected by RA treatment that occured before the onset of neural crest cell migration and before the onset of *dlx2* expression, the first *dlx* gene to be expressed in mmigrating crest cells and arch primordia. One possible explanation is that RA treatments result in severe perturbations in neural tube patterning leading to the absence or respecification of neural crest cells. RA treatments could also affect the migration of hindbrain crest cells. The absence, respecification or impaired migration of crest cells would serverely perturb visceral arch development and result in the concomitant abnormal expression of all *dlx* genes. However, we have shown that RA treatments at developmental stages that coincide with neural crest cell migration with the end of migration also lead to a loss of both gene expression and of cartilage formation. Although it does not eliminate the impact of RA treatments on neural crest specification, this last observation further supports a direct effect for RA on the expression of *dlx* genes and a requirement for normal expression of *dlx* genes for proper cartilage formation.

An alternative explanation for the effects of RA administered at stages that precede the onset of *dlx2* expression is that, although embryos treated at 6, 8, or 10 hr are washed extensively at the end of the RA treatment, some RA is retained, pehaps in the yolk sac, diffuses gradually, and would still be able to block induction of *dlx2* expression in migrating crest cells.

The lack of *dlx2* expression in migrating crest cells could in turn prevent induction of the other *dlx* genes in the arch primordia. It was previously suggested that expression of genes like *dlx3* and *dlx4* in the visceral arch primordia is activated following migration, in the same crest cells that express *dlx2* and/or in mesodermally derived cells as a result of interactions with these crest cells.[94] Such cross-regulation of *dlx* gene expression would be consistent with the observation that *dlx* genes are, in general, equally affected in the visceral arches following treatment with RA.

In addition to the possible roles of *dlx* genes during crest cell migration and patterning in the early primordia as described above, the patterns of *dlx* expression at later stages coincide with cartilage formation. In the present study

we have shown a strong correlation between the expression patterns of the zebrafish *dlx*1, *dlx*2, *dlx*3, and *dlx*4 genes and subsequent visceral arch cartilage differentiation. Moreover, when exogenous RA abolished *dlx* expression in specific facial regions, there is a direct one to one loss or abnormal modification of the visceral arch cartilage elements in those areas alone. Tellingly, two regions of the head skeleton which derived from regions that never displayed *dlx* expression, the posterior neurocranial base (the basilar plate), derive from the prechordal plate mesoderm and the trabaculae, which are derived from the more anterior mesencephalic crest, are largely unaffected by exogenous RA.

Studies on other vertebrates also point to a role for *dlx* genes in cartilage differentiation in the face. Thus, the mouse ortholog of the zebrafish *dlx*2 is expressed in migrating neural crest cells that will populate the visceral arches.[90, 96] *Dlx*2 and other *dlx* genes are expressed in the primordia of visceral arches. [89, 95, 96] Mammalian *dlx* genes are expressed during chondrocyte condensation. [86, 87, 91] Mouse *Dlx*5 and *Dlx*6 genes are known to be expressed in all tissues undergoing cartilage condensation. [86] A more extensive cartilage analysis was performed by Zhao et al (1994) with the rat r*dlx* gene.[91] They reported the expression of r*dlx* in the primordium of Meckel's cartilage. The r*dlx* transcripts were more abundant in cells on the periphery of the cartilage than in the center. It was suggested that r*dlx* plays a major role in either the morphogenesis and pattern formation of the skeleton or in determining the fate and phenotype of cells that are constituents of the skeleton. Recently, the mouse *Dlx*3 gene was also shown to be expressed in a group of condensed cells of the mandibular arch, in the proximity of Meckel's cartilage.[87] Orthologous zebrafish *distal-less* genes (*dlx*3, *dlx*4) are also found to be expressed in regions which correspond to condensing craniofacial cartilage. Thus, *dlx* genes may be an integral part of the pathway which restricts skeletogenic populations. The abolition of the expression of these genes by factors such as exogenous RA, as reported here, would then impede skeletogenesis as we have described.

5. DISCUSSION

Chapter 5

The *distal-less* family of homeobox genes is subdivided into six orthologous groups; *Dlx1, Dlx2, Dlx3, Dlx4, Dlx6,* and *Dlx7.* Each orthologous group has its own distinct expression pattern which is conserved between species.

The conserved expression patterns seen within each orthologous group is also seen between certain orthologous groups. For example, the *Dlx1* and *Dlx2* orthologous groups also share similar expression patterns, however, the onset of *Dlx1* expression lags behind that of *Dlx2.* This difference in onset of expression is conserved between zebrafish and mouse. The mouse *Dix1* and *Dlx2* genes are also found expressed in developing cartilages, the same is true for the zebrafish *dlx1* and *dlx2* genes. These two genes have also been reported to be genomically linked in an inverted fashion in the mouse, this manuscript shows that the zebrafish homologs are also linked in the same fashion. This remarkable conservation in expression patterns between the zebrafish and mouse must be due to a conservation of regulatory elements. It is thus possible that the conserved regulatory elements necessary for their expression patterns are located within the *Dlx1* and *Dlx2* intergenic region.

The third orthologous group, *Dlx3,* is not expressed in the developing ventral forebrain. The mouse *Dlx3* and zebrafish *dlx3* genes also show a conserved pattern of expression in the condensations of the craniofacial cartilage. Another interesting feature is that both the zebrafish *dlx3* and *dlx7* transcripts are found in the same cells. Therefore, it is highly possible that *dlx3* and *dlx7* are linked as are the orthologous groups Dlx1/Dlx2 & Dlx5/Dlx6. The zebrafish *dlx7* gene only has one gene which is closely related to it, the newt *NvHBox5* gene.

The mouse *Dlx5* and *Dlx6* orthologous groups are linked like the *Dlx1* and *Dlx2* genes, they also share a similar pattern of expression in the mouse. In the zebrafish, *dlx4* is the mouse Dlx5 ortholog, and it is linked as in the mouse to *dlx6.* The expression patterns for the zebrafish *dlx6* gene have not yet been determined, therefore the similarities in expression patterns can not be compared between the zebrafish and mouse.

According to Zhao,[89] *distal-less* genes play a major role in cartilage development of the rat. In zebrafish, *distal-less* genes are expressed during visceral arch development and differentiation in a specific temporal and spatial combination. Their visceral arch expression patterns correlate with future

neural crest derived craniofacial cartilage derivatives, insinuating that they may be involved in visceral arch development and differentiation.

Retinoic acid, a potent teratogen, is known to cause abnormal craniofacial cartilage development. In the fourth chapter, retinoic acid was shown to cause abnormal *dlx* expression in the visceral arches. Embryos at various stages, treated with 10^{-6}M RA, displayed a loss of all *dlx* expression in the mandibular and hyoid arches. The embryos exhibiting a loss of *dlx* expression, also developed with a loss of mandibular and hyoid arch derived craniofacial cartilage components.

Embryos treated with 10^{-7}M RA displayed a stage dependent loss of *distal-less* expression, mainly in the branchial arches. When *dlx* expression was lost in the branchial arches, then only the branchial arch cartilage was lost. However, abnormal *dlx* expression in any of the visceral arches could also be correlated with abnormal visceral arch cartilage development. Embryos treated at 24hpf with 10^{-7}M RA exhibited near normal *dlx* expression in the visceral arches, which could be correlated with normal visceral arch cartilage development.

These experiments support Zhao's hypothesis, suggesting that *dlx* genes are part of the multi-step process leading to visceral arch cartilage development. Zebrafish genes involved in this process are sensitive to retinoic acid (10^{-7}M) until 24hpf, where RA does not seem to affect the process of craniofacial cartilage development. Accordingly, a loss in visceral arch *dlx* expression during any stage (up to 24hpf) of this process leads to abnormal craniofacial cartilage development.

References

1. Streisinger, G., Walker, C., Dower, N., Knauber, N., and Singer, F., *Production of clones of homozygous diploid zebrafish (brachydanio rerio).* Nature, 1981. **291**: p. 293.
2. Mullins, M., Hammerschmidt , M., Haffter, P., and Nusslein-Volhart, C. *Mutagenesis and screen for embryonic mutants in the zebrafish, branchydanio rerio.* in *International Society Of Dev. Bio. 12th International Congress.* 1993. Vienna, Austria.
3. Gilbert, S., *Developmental biology,* ed. Sunderland. 1988, Massachusetts: Sinauer Ass. Press.
4. Gehring, W., *The discovery of the homeobox.*, in *Guidebook to the Homeobox Genes,* D. Duboule, Editor. 1994, Oxford University Press: Toronto. p. 3.
5. Kissinger, C., Liu, B., Martin-Blanco , E., Kornberg, T., and Pabo, C., *Crystal structure of an engrailed homeodomain DNA complex at 2.8 A resolution: a framework for understanding homeodomain-DNA interactions.* Cell, 1990. **63**: p. 579.
6. Scott, M., Tamkun, J., and Hartzell, W., *The structure and function of the homeodomain.* Biochemica Et Biophysica Acta, 1989. **989**: p. 25.
7. Scott, M., *Vertebrate homeobox gene nomenclature.* Cell, 1992. **71**: p. 551.
8. Krumlauf, R., *Hox genes and the pattern formation in the branchial region of the vertebrate head.* TIG, 1993. **9**(4): p. 106.
9. Kappen, C. and F, R., *Evolution of a regulatory gene family: Hom/Hox genes.* Current Opinion In Genetics And Development, 1993. **3**: p. 931.
10. Lewis, E., *A gene complex controlling segmentation in Drosophila.* Nature, 1978. **276**: p. 565.
11. Pfiefer, M., Karch, F., and Bender , W., *The bithorax complex controlling segmental identity.* Genes and Development, 1987. **1**: p. 891.
12. Mcginnis, W. and Krumlauf, R., *Homeobox genes and axial patterning.* Cell, 1992. **68**: p. 283.
13. Hafen, E., Levine, M., and Gehring, W., *Regulation of antennapedia transcript distribution by the bithorax complex in Drosophila.* Nature, 1984. **307**: p. 287.
14. Bridges, C. and Morgan, T., *The third chromosome group of mutant*

characters of the Drosophila melanogaster. Carnegie Inst. Washington Pub., 1923. **327**: p. 93.

15. Kostic, D. and Capecchi, M., *Targeted disruption of the murine Hoxa-4 and Hoxa-6 genes result in homeotic transformations of components of the vertebral column.* Mech. Dev., 1994. **46**: p. 231.
16. Izpisua-Belmonte, J., Dolle, P., Renucci, A., Zappavigna, V., Falkenstein, H., and Duboule, D., *Primary structure and embryonic expression pattern of the mouse Hox 4.3 homeobox gene.* Development, 1990. **110**: p. 733.
17. Kuziora, M. and Mcginnis, W., *Autoregulation of the Drosophila homeotic selector gene.* Cell, 1988. **55**: p. 477.
18. Kessel, M., Balling, R., and Gruss, P., *Variations of cervical vertebrae after expression of a Hox-1.1 transgene in mice.* Cell, 1990. **61**: p. 301.
19. Kessel, M. and Gruss, P., *Homeotic transformations of murine prevertebrae and concomitant alterations of Hox codes induced by retinoic acid.* Cell, 1991. **67**: p. 89.
20. Simone, A., Acompora, D., Arcioni, L., Andrews, P., Boncinelli, E., and Mavillio, F., *Sequential activation of Hox2 homeobox genes by retinoic acid in human embryonal carcinoma cells.* Nature, 1990. **346**: p. 763.
21. Tickle, C., Alberts, B., Wolpert, L., and Lee , J., *Local application of retinoic acid to the limb bond mimics the action of the polarizing region.* Nature, 1982. **296**: p. 564.
22. Tabin, C., *Retinoids, homeoboxes, and growth factors: toward molecular models for limb development.* Cell, 1991. **66**: p. 199.
23. Creech-Kraft, J., *Pharmacokinetics, placental transfer, and teratogenicity of 13-Cis retinoic acid, its isomer and metabolites.* In Retinoids In Normal Development And Teratogenesis, 1992. **OUP Oxford Press**: p. 267.
24. Langston, A. and Gudas, L., *Retinoic acid and homeobox gene regulation.* Current Opinion In Genetics And Development, 1994. **4**: p. 550.
25. Popperl, H. and Featherstone, M., *Identification of a retinoic acid response element upstream from murine Hox 4.2 gene.* Mol. Bio. Cell., 1993. **13**: p. 257.
26. Moroni, M., Vigano, M., and Mavilio, F., *Regulation of the Hoxd4 gene by*

retinoids. Mechanisms of Development, 1993. **44**: p. 139.

27. Hunt, P., Whiting, J., Muchamore, I., Marshall, H., and Krumlauf, R., *Homeobox genes and models for patterning the hindbrain and branchial arches*. Development, 1991. **1 Supplement**: p. 187.
28. Hunt, P. and Krumlauf , R., *Deciphering the Hox code: clues for patterning the branchial regions of the head*. Cell, 1991. **66**: p. 1075.
29. Hunt, P., Whiting, J., Nonchev, S., Sham, M., Marshall, H., Graham, A., Cook, M., Allemann, R., Rigby, P., Gulisano, M., Faiella, A., Boncinelli, E., and Krumlauf, R., *The branchial Hox code and its implications for gene regulation, patterning of the nervous system and head evolution*. Development Supplement, 1991. **2**: p. 63.
30. Finkelstein, R. and Perrimon , N., *The molecular genetics of head development in Drosophila*. Development, 1991. **112**: p. 899.
31. Boncinelli, E., *Early CNS development: distal-less related genes and forebrain development*. Current Opinion In Neurobiology, 1994. **4**: p. 29.
32. Figdor, M. and Stern, C., *Segmental organization of the embryonic diencephalon*. Nature, 1993. **363**: p. 630.
33. Schmitz, B., Papan, C., and Campos-Ortega, J., *Neurulation in the anterior trunk region of the zebrafish*. Roux's Arch. Dev. Bio., 1993. **202**: p. 250.
34. Papan, C. and Campos-Ortega, J., *On the formation of the neurla keel and neural tube in the zebrafish*. Roux's Arch. Dev. Bio., 1994. **203**: p. 178.
35. Le Douarin, N., Dupin, E., and Ziller, C., *Genetic and epigenetic control in neural crest development*. Current Opinion In Genetics And Development, 1994. **4**: p. 685.
36. Couly, G. and Le Douarin, N., *Head morphogenesis in embryonic avian chimeras: evidence for a segmental pattern in the ectoderm corresponding to neuromeres*. Development, 1990. **108**: p. 543.
37. Johnston, M., *A radiographic study of the migration and fate of cranial neural crest cells in the chick embryo*. Anat. Rec., 1966. **156**: p. 143.
38. Nicols, D., *Formation and distribution of neural crest mesenchyme to the first pharyngeal arch region of the mouse embryo*. Americain Journal of Anatomy, 1986. **176**: p. 221.

39. Tan, S. and Morris-Kay , G., *The development and distribution of the cranial neural crest in the rat embryo.* Cell Tissue Res, 1985. **240**: p. 403.

40. Shilling, T. and Kimmel, C., *Segment and cell type lineage restrictions during pharyngeal arch development in the zebrafish embryo.* Development, 1994. **120**: p. 483.

41. Lumsden, A., Sprawson , N., and Graham, A., *Segmental origin and migration of neural crest cells in the hindbrain region of the chick embryo.* Development, 1991. **113**: p. 1281.

42. Lumsden, A., *Segmental patterns of neuronal development in the chick hindbrain.* Nature, 1989. **337**: p. 424.

43. Stemple, D. and Anderson, D., *Lineage diversification of the neural crest: in vitro investigations.* Developmental Biology, 1993. **159**: p. 12.

44. Le Douarin, N., Ziller , C., and Couly , G., *Patterning of neural crest derivatives in the avian embryo: in vivo and in vitro studies.* Dev. Bio., 1993. **159**: p. 24.

45. Langille, R. and Hall, B., *Development of the head skeleton of the Japanese Medeka, Oryzias latipes (teleosei).* Journal of Morphology, 1987. **193**: p. 135.

46. Hall, B., *Cellular interactions during cartilage and bone development.* J. Craniofac. Genet. Dev. Bio., 1991. **11**: p. 238.

47. Shankar, K., Chuong, C., Jaskoll, T., and Melnick, M., *Effect of in ovo retinoic acid exposure on forebrain neural crest: in vitro analysis reveals up-regulation of N-CAM and loss of mesenchymal phenotype.* Developmental Dynamics, 1994. **200**: p. 89.

48. Wedden, S., Ralphs, J., and Tickle, C., *Pattern formation in the facial primordia.* Development, 1988. **103 Supplement**: p. 31.

49. Hall, B., *Genetic and epigenetic control of connective tissues in the craniofacial structures.* Birth Defects, Original Article Series, 1984. **20**(3): p. 1.

50. Noden, D., *Cell movements and control of patterned tissue assembly during craniofacial development.* J. Craniofac. Genet. Dev. Bio., 1991. **11**: p. 192.

51. Mclain, K., Schreiner, C., Yager, K., Stock , J., and Potter , S., *Ectopic*

expression of Hox-2.3 induces craniofacial and skeletal malformations in transgenic mice. Mechanisms of Development, 1992. **39**: p. 3.

52. Balling, R., Mutter, G., Gruss, P., and Kessel, M., *Craniofacial abnormalities induced by ectopic expression of the homeobox gene Hox-1.1 in transgenic mice*. Cell, 1989. **58**: p. 337.

53. Mark, M., Lufkin, T., Vonesch, J., Ruberte, E., Olivo, J., Dolle, P., Gorry, P., Lumsden, A., and Chambon, P., *Two rhombomeres are altered in Hoxa-1 mutant mice*. Development, 1993. **119**: p. 319.

54. Dolle, P., Dierich, A., Le Meur, M., Schimmang, T., Schuhbaur, B., Chambon, P., and Duboule, D., *Disruption of Hoxd-13 gene induces localized heterochrony leading to mice with neotenic limbs*. Cell, 1993. **75**: p. 431.

55. Zhang, M., Kim, H., Marshall, H., Gendron-Maguire, M., Lucas, D., Baron, A., Gudas, L., Gridely, T., Krumlauf, R., and Grippo, J., *Ectopic Hoxa-1 induces rhombomere transformation in mouse hindbrain*. Development, 1994. **120**: p. 2431.

56. Mckay, I., Muchamore, I., Krumlauf, R., Maden, M., Lumsden, A., and Lewis, J., *The Kreisler mouse: a hindbrain segmentation mutant that lacks two rhombomeres*. Development, 1994. **120**: p. 2199.

57. Frohman, M., Martin, G., Cordes, S., Halamek , L., and Barsh, G., *Altered rhombomere-specific gene expression and hyoid bone differentiation in the mouse segmentation mutant, kreisler (kr)*. Development, 1993. **117**: p. 925.

58. Wedden, S., *Effects of retinoids on chick face development*. J. Craniofac. Genet. Dev. Bio., 1991. **11**: p. 326.

59. Wedden, S., Lewin-Smith, M., and Tickle, C., *The effects of retinoids on cartilage differentiation in micromass cultures of chick facial primordia and the relationship to a specific facial defect*. Developmental Biology, 1987. **122**: p. 78.

60. Wedden, S., *Epithelial-mesenchymal interactions in the development of chick facial primordia and the target of retinoid action*. Development, 1987. **99**: p. 341.

61. Oikarinen, A., Oikarinen, H., and Uitto, J., *Demonstration of cellular retinoic acid binding protein in chick embryo tendon cells and effects of*

retinoids on collagen synthesis in tendon and sterna. Biochem. Pharm., 1986. **35**(19): p. 3393.

62. Marshall, H., Nonchev, S., Sham, H., Muchamore, I., Lumsden, A., and Krumlauf, R., *Retinoic acid alters hindbrain Hox code and induces transformation of rhombomeres 2/3 into a 4/5 identity.* Nature, 1992. **360**: p. 737.
63. Wood, H., Pall, G., and Morriss-Kay, G., *Exposure to retinoic acid before or after the onset of somitogenesis reveals separate effects on rhombomeric segmentation and 3'Hoxb gene expression domains.* Development, 1994. **120**: p. 2279.
64. Holder, N. and Hill , J., *Retinoic acid modifies development of the midbrain-hindbrain border and affects cranial ganglion formation in zebrafish embryos.* Development, 1991. **113**: p. 1159.
65. Stanier, D. and Fishman , M., *Patterning the zebrafish heart tube: acquisition of anteroposterior polarity.* Developmental Biology, 1992. **153**: p. 91.
66. Hyatt, G., Schmitt , E., Marsh-Armstrong, N., and Dowling, J., *Retinoic acid induced duplication of the zebrafish retina.* PNAS, 1992. **89**: p. 8293.
67. Hill, J., Clarke , J., Vargesson, N., Jowett, T., and Holder, N., *Exogenous retinoic acid causes specific alterations in the development of the midbrain and hindbrain of the zebrafish embryo including positional repecification of the Mauthner Neuron.* Mechanisms of Development, 1994. **(In Press)**.
68. Ho, I., Mercola, M., and Gudas, L., *Xenopus laevis cellular retinoic acid binding protein: temporal and spatial expression pattern during early embryogenesis.* Mechanisms of Development, 1994. **47**: p. 53.
69. Denker, L., Gustafson, A., Annerwall, E., Busch, C., and Ericksson, U., *Retinoid binding proteins in craniofacial development.* J. Craniofac. Genet. Dev. Bio., 1991. **11**: p. 303.
70. Leid, M., Kastner, P., and Chambon, P., *Mutiplicity generates diversity in the retinoic acid signaling pathways.* TIBS, 1992. **17**: p. 427.
71. Chytil, F. and Ong, D., in *In: The Retinoids.* 1984, Academic Press: Orlando, Florida. p. 90.
72. Morris-Kay, G., *Retinoic acid and craniofacial development: molecules*

and morphogenesis. Bioessays, 1993. **15**(1): p. 9.

73. Maden, M., Ong, D., Summerbell, D., and Chytil, F., *Spatial distribution of Cellular Protein Binding to Retinoic Acid in the chick limb bud.* Nature, 1988. **335**: p. 735.

74. Perez-Castro, A., Tran, V., and Nguyen-Huu, M., *Defective lens fiber differentiation and pancreatic tumorigenesis caused by ectopic expression of the cellular retinoic acid binding protein I.* Development, 1993. **119**: p. 363.

75. Tini, M., Otulakowski, G., Breitman, M., Tsui, L., and Giguere, V., *An everted repeat mediates retinoic acid induction of the Gamma F Crystallin gene: evidence of a direct role for retinoids in lens development.* Genes and Development, 1993. **7**: p. 295.

76. Soprano, D., Gyda, M., Jiang, H., Harnish, D., Ugen, K., Satre, M., Chen, L., Sprano, K., and Kochhar, D., *A sustained elevation in retinoic acid receptor beta 2 mRNA and protein occurs during retinoic acid induced fetal dysmorphogenesis.* Mechanisms of Development, 1994. **45**: p. 243.

77. Maden, M., Horton, C., Graham, A., Leonard, L., Pizzey, J., Siegenthaler, G., Lumsden, A., and Erikson, A., *Domains of Cellular Retinoic Acid Binding Protein expression in the hindbrain and neural crest of the mouse embryo.* Mechanisms of Development, 1992. **37**: p. 13.

78. Mahmood, R., Flanders, K., and Morris-Kay, G., *Interaction between retinoids and TGF betas in mouse morphogenesis.* Development, 1992. **115**: p. 67.

79. Carnac, G., Albagli-Curiel, O., Levin, A., and Bonnieu, A., *9-cis retinoic acid regulates the expression of the muscle determination gene Myf5.* Endocrinology, 1993. **133**(5): p. 2171.

80. Tran, P., Zhang, X., Salbert, G., Hermann, T., Lehmann, J., and Pfahl , M., *COUP orphan receptors are negative regulators of retinoic acid response pathways.* Mol. Cell. Bio., 1992. **12**(10): p. 4666.

81. Retnakaran, R., Flock, G., and Giguere, V., *Identification of RVR, a novel orphan nuclear receptor that acts as a negative transcriptional regulator.* Mol. Endo., 1994. **8**(9): p. 1234.

82. Yokouchi, Y., Ohsugi, K., Sasaki, H., and Kuroiwa, A., *Chicken homeobox gene Msx-1: structure, expression in limb buds and effect of retinoic acid.*

Development, 1991. **113**: p. 431.

83. Kessel, M., *Respecification of vertebral identities by retinoic acid.* Development, 1992. **115**: p. 487.
84. Studer, M., Popperl, H., Marshall, H., Kuroiwa, A., and Krumlauf, R., *Role of a conserved retinoic acid response element in rhombomere restriction of Hoxb-1.* Science, 1994. **265**: p. 1728.
85. Cohen, S., *Specification of limb development in the Drosophila embryo by positional cues from segmentation genes.* Nature, 1990. **343**: p. 173.
86. Simeone, A., Acampora , D., Pannese, M., d'Esposito, M., Stornaiuolo, A., Gulisano, M., Mallamaci, A., Kastury, K., Druck, T., Huebner, K., and Boncinelli, E., *Cloning and characterization of two members of the vertebrate Dlx gene family.* PNAS, 1994. **91**: p. 2250.
87. Robinson, G. and Mahon, K., *Differential and overlapping expression domains of Dlx2 and Dlx3 suggests distinct roles for distal-less homeobox genes in craniofacial development.* Mechanisms of Development, 1994. **48**: p. 199.
88. Robinson, G., Wray, S., and Mahon, K., *Spatially restricted expression of a member of a new family of murine distal-less homeobox genes in the developing forebrain.* New Biologists, 1991. **3**(12): p. 1.
89. Price, M., Lemaistre, M., Pischetola, M., Di Lauro, R., and Duboule, D., *A mouse gene related to distal-less shows a restricted expression in the developing forebrain.* Nature, 1991. **351**: p. 748.
90. Porteus, M., Bulfone, A., Ciaranello , R., and Rubenstein, J., *Isolation and characterization of a novel cDNA clone encoding a homeodomain that is developmentally regulated in the ventral forebrain.* Neuron, 1991. **7**: p. 221.
91. Zhao, G., Zhao, S., Zhou , X., Eberspaecher, H., Solursh, M., and De Crombrugghe, B., *Rdlx, a novel distal-less like homeoprotein is expressed in developing cartilages and discrete neuronal tissues.* Developmental Biology, 1994. **164**: p. 37.
92. Beauchemin, M. and Savard, P., *Two distal-less related homeobox containing genes expressed in regeneration blastemas of the newt.* Developmental Biology, 1992. **154**: p. 55.
93. Asano, M., Emori, Y., Saigo, K., and Shiokawa, K., *Isolation and*

characterization of a Xenopus cDNA which encodes a homeodomain highly homologous to the Drosophila distal-less. Journal of Biological Chemistry, 1992. **267**(8): p. 5044.

94. Dirkson, M., Mathers, P., and Jamrich, M., *Expression of a Xenopus distal-less homeobox gene involved in forebrain and craniofacial development.* Mechanisms of Development, 1993. **41**: p. 121.
95. Papalopulu, N. and Kintner, C., *Xenopus distal-less related homeobox genes are expressed in the developing forebrain and are induced by planar signals.* Development, 1993. **117**: p. 961.
96. Akimenko, M., Ekker, M., Wegner, J., Lin, W., and Westerfield, M., *Combinatorial expression of three zebrafish genes related to distal-less: part of a homeobox gene code for the head.* Journal of Neuroscience, 1994. **14**(6): p. 3475.
97. Eisenberg, L., Ingham, P., and Brown, A., *Cloning and characterization of a novel Drosophila Wnt gene, Dwnt-5, a putative downstream target of the homeobox gene distal-less.* Developmental Biology, 1992. **154**: p. 73.
98. Vachon, G., Cohen, B., Pfiefle, C., Mcguffin , M., Botas, J., and Cohen, S., *Homeotic genes of the bithorax complex repress limb development in the abdomen of the Drosophila embryo through the target gene distal-less.* Cell, 1992. **71**: p. 437.
99. O'Hara, E., Cohen, B., Cohen, S., and Mcginnis, W., *Distal-less is a downstream gene of deformed required for ventral maxillary identity.* Development, 1993. **117**: p. 847.
100. Mcginnis, N., Kuziora, M., and Mcginnis, W., *Human Hox4.2 and Drosophila Deformed encode similar regulatory specificities in Drosophila embryos and larvae.* Cell, 1990. **63**: p. 969.
101. Westerfield, M., *The Zebrafish Book.*, ed. 2. In Westerield M. 1993, Eugene, OR: University of Oregon Press.
102. Puschel, A., Gruss , P., and Westerfield, M., *Sequence and expression pattern of pax-6 are highly conserved between zebrafish and mice.* Development, 1992. **114**: p. 643.
103. Ekker, M., Akimenko, M., Bremiller , R., and Westerfield, M., *Regional expression of three homeobox transcripts in the inner ear of zebrafish embryos.* Neuron, 1992. **9**: p. 27.

104. Breitbart, R., *Personnel communications.* 1994.

105. Held, L., Heup, M., Sappington, J., and Peters, S., *Interactions of decapentaplegic, wingless, and distal-less in the Drosophila leg.* Roux's Arch. Dev. Bio., 1994. **203**: p. 310.

106. Mendlesohn, C., Larkin, S., Mark, M., LeMeur, M., Clifford, J., Zelent, A., and Chambon, P., *RAR beta isoforms: distinct tanscriptional control by retinoic acid and specific spatial patterns of promoter activity during mouse embryonic development.* Mechanisms of Development, 1994. **45**: p. 227.

107. Pratt, M., Langston, A., Gudas, L., and McBurney, M., *Retinoic acid fails to induce expression of Hox genes in differentiation defective murine embryonal carcinoma cells carrying a mutant gene for RAR Alpha.* Differentiation, 1993. **53**: p. 105.

108. Tate, B., Levin, A., and Grippo, J., *The discovery of 9-cis retinoic acid; a hormone that binds the RXR.* TEM, 1994. **5**(5): p. 189.

109. Zimmer, A. and Zimmer, A., *Induction of a RAR beta2-lacZ transgene by retinoic acid reflects neuromiric organization of the CNS.* Development, 1992. **116**: p. 977.

110. Henikoff, S., *Unidirectional digestion with Exonuclease III in DNA sequence analysis.* Methods In Enzymology, 1987. **155**: p. 156.

111. Oxtoby, E. and Jowett, T., *Cloning of the zebrafish Krox-20 (Krx-20) and its expression during hindbrain development.* Nucl. Acids Res., 1993. **21**(5): p. 1087.

112. Ekker, M., Wegner, J., Akimenko, M.A., and Westerfield, M., *Coordinate expression of three zebrafish engrailed genes.* Development, 1992. **116**: p. 1001.