CHAPTER II

REVIEW OF THE LITERATURE

The review is divided into eight major parts as follows:

2.1 *MSX1* and its Structure and Functions

2.1.1 Mapping and structure

2.1.2 Conservation of *MSX1* gene

2.1.3 Mapping *MSX1* mutations by domain

2.1.4 Msx1 protein structure

2.1.5 Knockout of *Msx1* $^{-/-}$ in mice

2.1.5.1 Perinatal lethality in *Msx1* $^{-/-}$ homozygotes

2.1.5.2 Abnormal tooth development

2.1.5.3 Cleft palate in *Msx1* deficient mice

2.1.5.4 Maxillary and mandibular defects in *Msx1* $^{-/-}$ mice

2.1.5.5 Abnormal skull and nasal bone development

2.1.5.6 Middle ear abnormalities

2.1.5.7 Other organs

2.1.6 Roles of *MSX1* in development

2.1.6.1 *Msx1* gene encode transcription repressor

2.1.6.2 *Msx1* in the role of antisense transcription factor

2.1.6.3 *Msx1* induces dedifferentiation of myotubes

2.1.6.4 *Msx1* function in repression of myogenic gene expression

2.5 Orofacial clefts

2.5.1 Genes known to cause non-syndromic orofacial cleft in mice and human

2.6 *Msx1*-associated human syndromic malformations

2.6.1 Witkop syndrome

2.6.1.1 Clinical manifestations

2.6.1.2 Etiology

2.6.2 Wolf-Hirschhorn syndrome (WHS)

2.6.2.1 Clinical manifestations

2.6.2.2 Etiology

2.7 *Msx1*-associated human non-syndromic malformations

2.7.1 *MSX1* mutations and pattern of non-syndromic hypodontia

2.7.2 *MSX1* mutations and pattern of non-syndromic orofacial clefts

2.8 Previously reported *MSX1* mutations

2.1 *MSX1* and its structure and functions

MSX1 (Muscle Segment Homeobox, Drosophila, homolog of, 1) is a member of the *Msx* family. The vertebrate *Msx* genes were originally cloned from mice and determined as being homologous to the Drosophila *msh* (muscle segment homeobox) gene (Hill et al., 1989; Robert et al., 1989). *Msx* genes have been isolated from a variety of organisms, including ascidians (Ma et al., 1996), sea urchin, zebrafish (Akimenko et al., 1995), frogs (Su et al., 1991), birds (Takahashi and Le Douarin, 1990), and humans (Padanilam et al., 1992).

The mammalian *Msx* gene family comprises of 3 physically unlinked members, named *Msx1*, *Msx2*, and *Msx3*. From a genomic cosmid library, isolation a cosmid containing the human sequence homologous to the mouse homeobox gene (*Msx1* or *Hox7* in the past) revealed close homology in structure and sequence between human and mouse *Msx1*. *Msx1* is expressed in the heart valves, mandibular and hyoid arches, and limb buds during normal development.

2.1.1 Mapping and structure

MSX1 gene locates at chromosome 4p16.1. This region reveals homology of synteny with part of mouse chromosome 5 where the *Msx1* gene is located (Robert et al., 1989).

MSX1 gene is 4271 bp long and it consists of two exons of 704 bp and 1229 bp (Figure 2.1). This gene has two primary transcripts, both are of 1933 bp long, begin in the start codon at 236 or 254, producing two proteins of 303 and 297 amino acids, respectively. The open reading frame is inserted by a single intron of approximately 2332 bp (Pawlowska et al., 2009).

The first part of gene (upstream of exon 1) consists of 5' untranslated region while the end of gene (downstream of exon 2) contains 3' untranslated region (Jezewski et al., 2003). The sequence upstream of the coding sequence of exon 1 composes of a GC-rich putative promoter region. There is no TATA box, but a CCAAT and numerous GC boxes are revealed. The area encompassing the promoter region, exon 1, and the 5' region of exon 2 have a higher than expected frequency of CpG dinucleotides; numerous sites for rare-cutter restriction enzymes are present, a characteristic of HTF islands (Hewitt et al., 1991).

Moreover, an antisense transcript plays important role by limiting *MSX1* function. Its length and direction are shown in Figure 2.1 by the red arrow. It involves the intron, all of exon 2 and 3' untranslated region (Blin-Wakkach et al., 2001). Blast comparison between human and mouse sequences revealed that antisense transcript terminated within the MIHR (*MSX1* Intronic Human-mouse Homology Region). The antisense transcript consists of 98 amino acid open reading frame (Jezewski et al., 2003).

The human *MSX1* coding sequence has a very high degree of similarity with the mouse *Msx1* cDNA (Jezewski et al., 2003). The two sequences share 94% identity at the DNA level, all substitutions being silent. This high level of sequence identity is not limited to the homeodomain only, but overall the human and mouse *Msx1* gene products also show 80% identity at the amino acid level. Both the 5' and 3' untranslated regions also show significant similarity to the mouse gene, with 79 and 70% sequence identity, respectively.

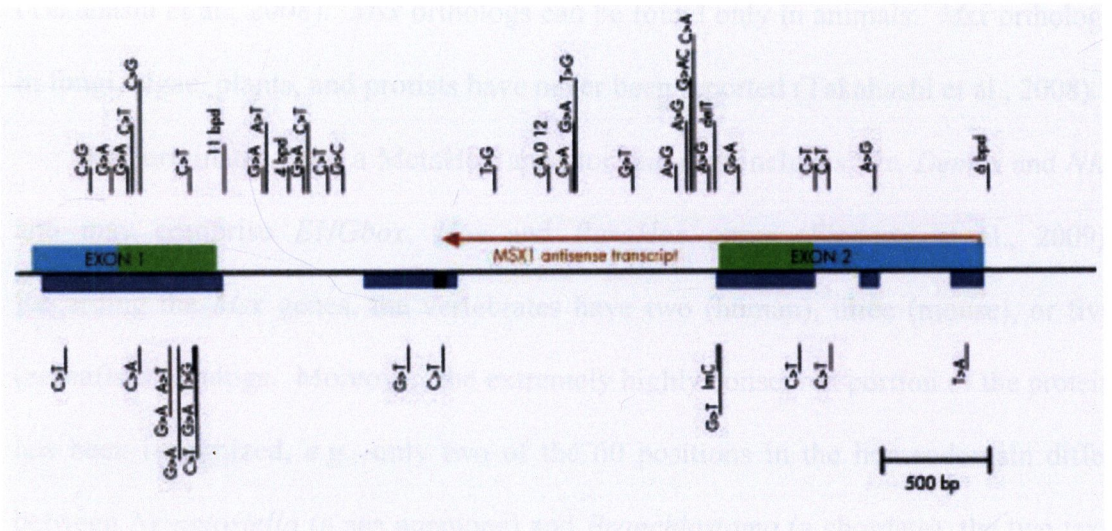


Figure 2.1 Structure of the *MSX1* gene. *MSX1* gene has two coding regions of exons (Green boxes). Light blue indicates 5' and 3' untranslated regions. Dark blue determines the regions of human/mouse sequence homology. Red arrow shows the antisense-*MSX1* transcript length/direction. Black box within the dark blue is the range where antisense transcript terminates. Description of the variants found is shown by vertical lines with the letters of the nucleotide change indicated. Common alleles are shown above and potential etiological mutations are shown below the gene structure (Jezewski et al., 2003).

2.1.2 Conservation of *MSX1* gene

In evolution, the *Msx* gene family was one of the earliest animal-specific homeodomain transcription factors. *Msx* genes have been identified in sea anemones (Finnerty and Martindale, 1997; Ryan et al., 2006), corals (de Jong et al., 2006), hydras (Schummer et al., 1992), jellyfish (Galle et al., 2005), and sponges (Seimiya et al., 1994). The sponges are commonly regarded as the most basal animal phylum, indicating that *Msx* gene is important for the existence of animals. The *Msx* gene family has also been described in eleven different phyla of triploblastic animals

(Takahashi et al., 2008). *Msx* orthologs can be found only in animals. *Msx* orthologs in fungi, algae, plants, and protists have never been reported (Takahashi et al., 2008).

Msx originates from a MetaHox ancestor that also includes *Tlx*, *Demox* and *NK*, and may comprise *EHGbox*, *Hox* and *ParaHox* genes (Finnerty et al., 2009). Regarding the *Msx* genes, the vertebrates have two (human), three (mouse), or five (zebrafish) paralogs. Moreover, the extremely highly conserved portion of the protein has been recognized, *e.g.*, only two of the 60 positions in the homeodomain differ between *Nematostella* (a sea anemone) and *Branchiostoma* (a chordate), the two taxa that diverged over 600 million years ago.

The functional evolution of Msx occurred from duplication, divergence, or loss of domains (Finnerty et al., 2009). Duplicated domains allow expressed proteins to develop new functions without disrupting existing interaction networks.

The conserved Msx Homology (MH) domains have been divided into seven domains (Figure 2.2) (Finnerty et al., 2009). These consist of two Groucho-binding domains, MH1N and MH1C domains, MH2 (a Pbx binding motif), MH3 (the eight amino acids immediately upstream of the homeodomain), MH4 (the homeodomain), MH5 (the twelve amino acids downstream of homeodomain) and MH6 (a PIAS-binding domain located at the carboxy terminus).

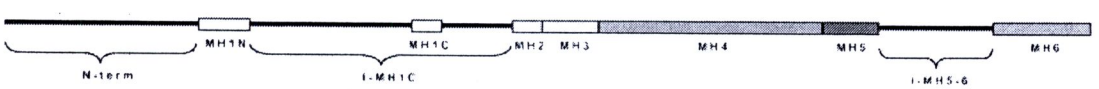


Figure 2.2 Seven Msx Homology (MH) domains, including two Groucho repression domains, MH1N and MH1C domains, MH2, MH3, MH4, MH5, and MH6 (Finnerty et al., 2009).

The MH1N and MH1C domains, or Groucho-binding domains, are the N-terminal region amino acid sequences that have been evolutionarily conserved in a wide range of species (Takahashi et al., 2008). They are the most conserved sequences located at the N-terminal end. The consensus sequence in MH1N can be summarized as FSVEALMAD. MH1C is located between the MH1N and PWM sequences, and its consensus sequence is FSVGGLLKL, which is similar to that of MH1N. In the exon–intron organization of *Msx* gene, or the alignment of the N-terminal region amino acid sequence by adjusting the intron position as the cardinal point, there is a weakly conserved sequence near the intron position. This conserved sequence is summarized as PWM, where tryptophan is strongly conserved (Takahashi et al., 2008).

The same MH1N sequence exists in Xenopus, Ambyostoma, Notophthalamus, Gallus, Mus, Rattus, Bos, and all primate species (Takahashi et al., 2008). Whereas MH1N is conserved in all species, MH1C shows variable distribution among the invertebrate Msx proteins investigated. The MH1C sequence is maintained in the Vertebrata, Cephalochordata, Hemichordata, Echinodermata, Mollusca, and Anthozoa (Cnidaria) but not in the Urochordata or Hydrozoa (Cnidaria) Msx homologs. Furthermore, in vertebrate Msx2 and Msx3, the MH1C is missing. Finnerty and coworkers suggested that the MH1C diverged substantially, implying its novel function (Finnerty et al., 2009). Because of the sequence similarity between MH1N and MH1C, it is possible that Groucho proteins can bind to either MH1N or MH1C domain, exhibiting functional redundancy.

The MH2 domain includes the "hexapeptide", a motif first determined in Hox proteins and that increases DNA binding specificity by binding to Pbx family proteins

as a cofactor (Peltenburg and Murre, 1996). This motif consists of PWM sequence. The MH2 domains in the Msx proteins of basal metazoans and nonvertebrate deuterostomes most often have double tryptophans, whereas the vertebrate Msx proteins have only a single tryptophan that resembles the hexapeptide sequences of anterior Hox genes (Finnerty et al., 2009).

The MH3 domain is the part between the MH2 domain and the homeodomain (Finnerty et al., 2009). This linker consists of the eight highly conserved amino acids immediately upstream of the homeodomain.

The MH4 domain corresponds to the homeodomain, which is involved in DNA binding and protein-protein (homo- or hetero-) dimerization (Finnerty et al., 2009). The homeodomain is strongly conserved (Figure 2.3) (Takahashi et al., 2008). Particularly, the residues that are responsible for the molecular interaction with DNA bases (R2, R5, K46, I47, Q50, N51, and R58) or with DNA phosphoribosyl backbones (K3, T6, F8, Y25, R31, W48, R53, K55, and K57) (Hovde et al., 2001) are absolutely conserved among the sequences in the homeodomain, suggesting that interactions between DNA or other proteins and the residues in the homeodomain are essential for many animals to survive against natural selection (Takahashi et al., 2008).

The MH5 domain, which is contiguous with the carboxy-terminus of the homeodomain, has been shown to be involved in transcriptional repression (Finnerty et al., 2009). Comparison of the downstream sequence of several Demox homeodomain proteins (Richelle-Maurer et al., 2006) and the first portion of the MH5 sequence (EAELEKLKMAAKPMLPP), reveals a strong match between them. It has

been reported that the first thirteen amino acids of the MH5 domain are the most conserved residues (Figure 2.3) (Takahashi et al., 2008).

The MH6 domain is the PIAS protein-binding domain (Finnerty et al., 2009). Although MH6 has not been examined directly in the initial analysis of the non-vertebrate taxa, multi-sequence alignments have demonstrated that many core amino acids within the MH6 domain are also conserved (Finnerty et al., 2009).

Furthermore, a manually-assembled sequence collection, comprising two chordate Msx proteins, two Msx proteins, and two poriferan Msx proteins in addition to full-length NK and Tlx homeobox proteins, including all of the published full-length MetaHox protein sequences from sponges, revealed strong matches to the MH3/4 domains and much weaker matches to the MH1 and MH2 domains, resulting from the lack of core amino acids (Finnerty et al., 2009).

Regarding Msx paralogs (Msx1 and Msx2), the greatest divergence between them is found in the N-terminal segment, and the least divergence is found in the MH1N and MH4 domains (Finnerty et al., 2009). Finnerty and coworkers suggested that the subsequent genome and *Msx* gene duplications in vertebrates may indicate an additional level of functional redundancy, perhaps promoted by paralog co-evolution (of the Groucho and PIAS families) (Finnerty et al., 2009). This may have permitted the divergence of the MH1N and MH1C domains, ultimately resulting in the loss of MH1C in both the Msx2 and Msx3 lineages. Moreover, Msx2 may be redundant for mutations in *Msx1* where these two proteins share a domain (e.g., MH1N), but it may not be redundant for mutations in a domain of Msx1 that is lacking in Msx2.

From these studies of *MSX1* conservation, the conserved Msx Homology (MH) domains have been divided into seven domains. Each of the conserved domains has

different level of conservation and different important functions. For the example, there are more conserved amino acids in the MH3 and MH4 domains than in the MH1 and MH2 domains. Highly conserved amino acids are often required for basic cellular function and protein stability. Therefore, mutations of highly conserved amino acids most likely to interrupt critical functions are found within the most conserved domains.

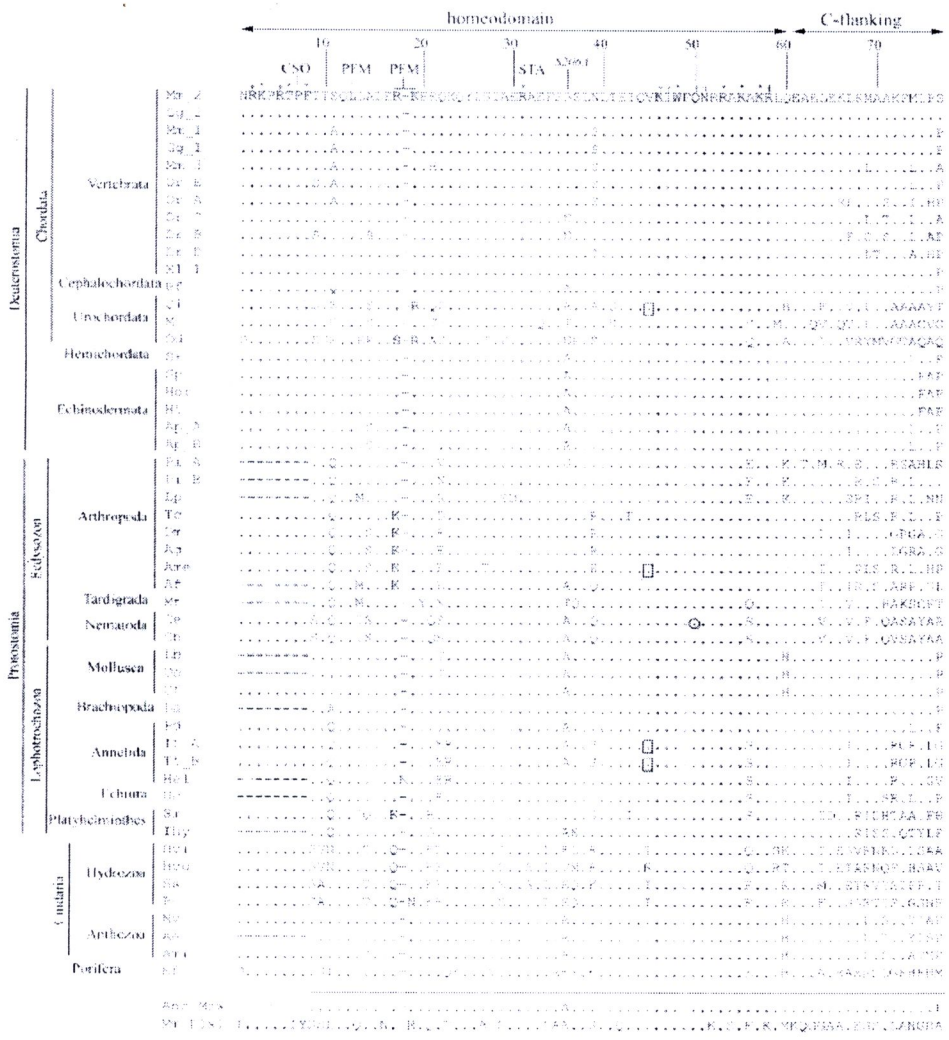


Figure 2.3 The amino acid sequences of the homeodomain and a C-terminal flanking region of the Msx1 protein. "Dots" indicate the same amino acids as those in the top sequence (Mm_2), and "hyphens" show the spaces artificially inserted in the multiple alignment program. ''Plus'' signs indicate the amino acids that are responsible for molecular interaction with DNA bases, and

''asterisks'' indicate the amino acids that are responsible for molecular interaction with DNA phosphoribosyl backbones. ''Shading'' indicates the locations of DNA-associated or disease-causative aa residues in the alignment. ''Open boxes'' and ''open circles'' indicate the presence of phase-0 and phase-2 introns, respectively. The positions of residues responsible for human genetic disorders are shown at the top alignment. Animal species names are given in Table 2.1. S206A indicates an artificial missense mutation introduced into Xl-Msx1. Abbreviations: Anc_Msx: the metazoan ancestral Msx sequence deduced by ANCESCON; CSO: Craniosynostosis; PFM: Parietal foramina; STA: Selective tooth agenesis (Takahashi et al., 2008).

Table 2.1 Animal species names (Takahashi et al., 2008).

Phylum Subphylum Class	Species	Abbreviation	Genome	cDNA	Msx-Related Genes (Synonym, Accession Number)
Chordata					
Vertebrata					
Mammalia	*Mus musculus*	Mm	DB	DB	Msx1 (AAH16426); Msx2 (Q03358); Msx3 (P70354)
Aves	*Gallus gallus*	Gg	DB	DB	Msx1 (Hox7, P502223); Msx2 (Hox8, P28362)
Amphibia	*Xenopus laevis*	Xl		DB	Msx1 (AAH81101)
Osteichthyes	*Danio rerio*	Dr	DB	DB	Msh-A (Q03357); Msh-B (Q03356); Msh-C (Q11703); Msh-D (Q11704); MsxE (AAB03273)
Cephalochordata					
Leptocardia	*Branchiostoma floridae*	Bf		DB	Msx (CAA10201)
Urochordata					
Ascidiacea	*Ciona intestinalis*	Ci	DB	DB	CAD56691
	Molgula oculata	Mo		DB	AAA87223
Appendicularia	*Oikopleura dioica*	Od		DB	AAW24005
Hemichordata					
Enteropneusta	*Saccoglossus kowalevskii*	Sk		DB	Msx (ABD97280)
Echinodermata					
Echinoidea	*Strongylocentrotus purpuratus*	Sp	DB	DB	SpMsx (AAB97688)
	Hemicentrotus erythrogramma	Her		DB	Msx (AAY86177)
	Hemicentrotus tuberculata	Ht		DB	Msx (AAY86178)
Asteroidea	*Asterina pectinifera*	Ap	Fosmid	PCR	MsxA (Ast-1 (AB302953a)), MsxB (Ast-3 (AB302954))
Arthropoda					
Chelicerata	*Limulus polyphemus*	Lp		PCR	MsxA (AB302959)
	Pandinus imperator	Pi		PCR	MsxA (AB302960), MsxB (AB302961)
Crustacea	*Artemia franciscana*	Af		PCR	Msx (AB302962)
Insecta	*Tribolium castaneum*	Tc	DB		Muscle segment homeoprotein (AAW21975)
	Drosophila melanogaster	Dm	DB	DB	msh (Q03372)
	Anopheles gambiae	Ag	DB	DB	EAA08817
	Apis mellifera	Ame	DB		Msx (AB362784)
Tardigrada	*Milnesium tardigradum*	Mt		PCR	Msx (AB302966)
Nematoda	*Caenorhabditis elegans*	Ce	DB	DB	vab-15 (Q09604)
	Caenorhabditis briggsae	Cb		DB	CAE58718
Mollusca					
Cephalopoda	*Loligo bleekeri*	Lb		PCR	Msx (AB302963)
	Octopus ocellatus	Oo		PCR	Msx (AB302964)
Bivalvia	*Corbicula fluminea*	Cf	Fosmid	PCR	Msx (Cor-2/Cor-3 (AB302955))
Brachiopoda					
Inarticulata	*Lingula anatina*	La		PCR	Msx (AB302965)
Annelida					
Polychaeta	*Platynereis dumerilii*	Pd		DB	Msx (CAB38810)
Oligochaeta	*Tubifex tubifex*	Tt	Fosmid	PCR	MsxA (Tt-1 (AB302956)), MsxB (Tt-3 (AB302957))
Hirudinida	*Helobdella* sp.	Hel		DB	Msx (AAB37254)
Echiura					
	Urechis unicinctus	Uu		PCR	Msx (AB302967)
Platyhelminthes					
Tubellaria	*Schmidtea mediterranea*	Sm	DB		Msx (AB362785)
	Dugesia sp.	Dsy		PCR	Msx (AB302968)
Cnidaria					
Hydrozoa	*Hydra viridis*	Hvi		DB	CAA45912
	Hydra vulgaris	Hvu		DB	CAB88390
	Scolionema suvaense	Ss	Fosmid	PCR	Msx (Sco-1/Sco-6 (AB302958))
	Podocoryne carnea	Pc		DB	AAX58756
Anthozoa	*Nematostella vectensis*	Nv	PCR		Msx (AB362783)
	Acropora nobilis	An		PCR	Msx (AB302969)
	Acropora millepora	Ami		DB	msh3 (ABK41269)
Porifera					
Demospongiae	*Ephydatia fluviatilis*	Ef		DB	prox3 (AAA20151)

NOTE. DB, sequence collected from NCBI databases; fosmid, isolated as fosmid genomic clones in this study; PCR, cloned by PCR amplification of the cDNA

[a] Sequence newly determined in this study.

[b] Sequences edited from the publicized draft sequences (http://www.ncbi.nlm.nih.gov/BLAST/Tax4fam.shtml).

2.1.3 Mapping *MSX1* mutations by domain

MSX1 mutations associated with tooth agenesis with or without orofacial clefting disorders are mapped to the conserved domains in a non-random pattern. The mutations causing orofacial clefting (OFC) and the mutations causing tooth agenesis or ectodermal dysplasias (ED) are mapped to the conserved domain in a non-overlapping pattern (Figure 2.4) (Finnerty et al., 2009). OFC mutations have been reported in and around the MH1C, MH3 and MH6 domains, whereas ED mutations are found within or upstream of MH1N and within MH4 domains (Finnerty et al., 2009).

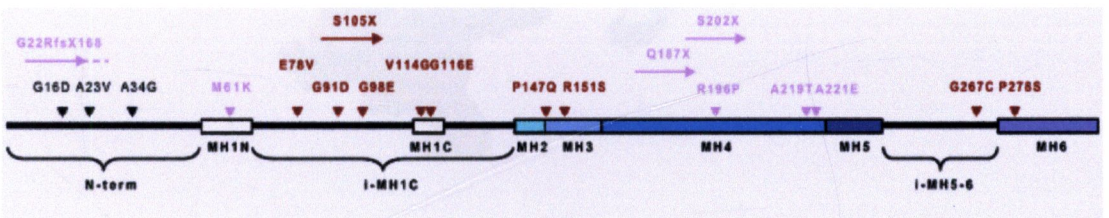


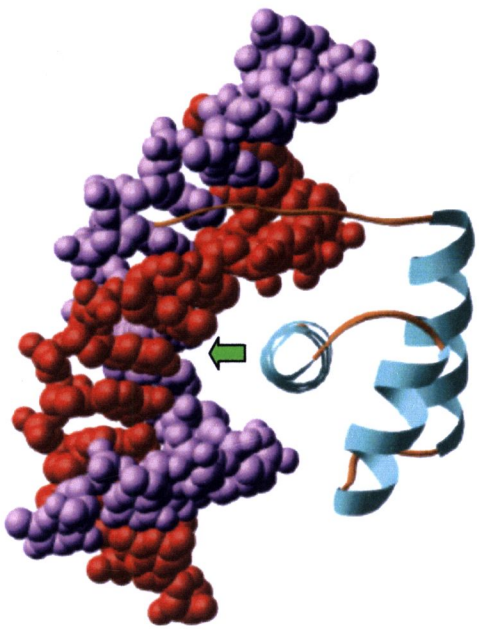
Figure 2.4 The mutations map to the conserved domains. OFC mutations are shown in dark red, ED mutations are shown in light pink (Finnerty et al., 2009).

2.1.4 Msx1 protein structure

The *Msx1* gene encodes protein that consists of 297 amino acids. The Msx1 protein (Figure 2.5) functions in several biological processes, such as negative regulation of transcription from RNA polymerase II promoter, skeletal development, muscle development, organ morphogenesis, embryonic limb morphogenesis, forebrain development, and midbrain development (Catron et al., 1996). In these processes, the Msx1 protein binds to specific DNA sequences, a TAAT core motif which is typical of homeoproteins (Catron et al., 1993), and thereby repressing transcription. Moreover, several proteins play an essential role in selective protein-

protein interactions that can facilitate or antagonize the function of Msx1 protein. The thorough understanding of Msx1 structure will be necessary for the better understanding of its function.

A.



B.

```
          1  2  3  4  5  6  7  8  9  10 11 12 13 14 15 16
5'-T  G  T  C  A  C  T  A  A  T  T  G  A  A  G  G
      C  A  G  T  G  A  T  T  A  A  C  T  T  C  C  T-5'
     32 31 30 29 28 27 26  25 24 23 22 21 20 19 18 17
```

Figure 2.5 (A) The Msx1 homeodomain-DNA complex. The Msx1 protein is shown in blue, while the DNA is shown in red and purple. The green arrow indicates the core binding site for Msx1 protein (TAAT site). (B) The sequence of DNA, the bases in bold (**TAAT**) are the core binding site for Msx1 protein (Hovde et al., 2001).

Msx1 protein consists of three domains: an N-terminal domain (amino terminal end), a homeodomain and a C-terminal domain (carboxy terminal end), respectively as shown in Figure 2.6. Each domain has different functions and interacts with different proteins to regulate the biological process.

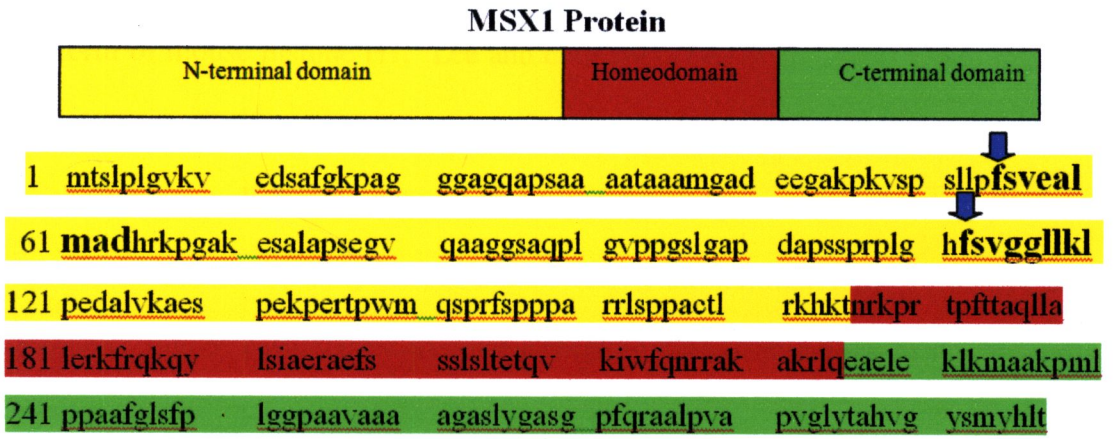


Figure 2.6 MSX1 protein domains and amino acid sequences in each domain. "Arrows" indicate the two highly conserved amino-acid regions (FSVEALMAD and FSVGGLLKL) in the N-terminal domain.

In Msx1, Msx2, and Msx3, in many animals, there are two highly conserved amino-acid regions (FSVEALMAD and FSVGGLLKL) in the N-terminal domain (Ekker et al., 1997). These motifs have also been suggested to be similar to the engrailed homology (EH-1 and EH-2) repression domains (Smith and Jaynes, 1996). These conserved Msx1 regions, which are quite hydrophobic, may also have repression activity and interact with Groucho (Grg). Grg is a basic-helix-loop-helix protein, which interacts with several motifs in other transcription repression proteins (Rave-Harel et al., 2005), suggesting that transcription repression by these domains is mediated by Grg protein. Grg1 can physically interact with Msx1 and can regulate an Msx target gene.

Moreover, Histone 1b protein can also bind to the specific motif in the N-terminal domain (Figure 2.7). One study revealed that Msx1 proteins lacking residues 105 to 139 were also defective for interaction with Histone 1b (Lee et al., 2004).

There is a physical interaction between Msx1 and Histone 1b which is a specific isoform of mouse histone H1. Lee and collegues (Lee et al., 2004) found that Msx1 and Histone 1b bind to a key regulatory element of *MyoD*, a central regulator of skeletal muscle differentiation, where they induce repressed chromatin. Moreover, Msx1 and Histone 1b cooperate to inhibit muscle differentiation in cell culture and in *Xenopus* animal caps. Jezewski and coworkers found mutations in the N-terminal domain and suggested that Msx1 protein lacking the N-terminal domain was not able to upregulate cyclin D1, nor to inhibit differentiation. They also suggested that early differentiation in the progress zone of the facial processes could reduce outgrowth (Jezewski et al., 2003).

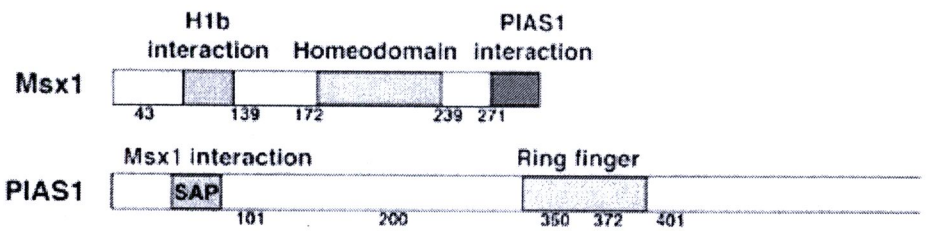


Figure 2.7 Histone H1b interacts with Msx1 protein at the N-terminal region, whereas PIAS1 interacts at the C-terminal region (Lee et al., 2006).

The conserved sequence of amino acids just N-terminal to the homeodomain, previously called the extended homeodomain (EHD). This EHD region consists of a PBX (pre-B-cell leukemia homeobox) binding sequence (TPWMQ), several potential phosphorylation sites, a potential nuclear localization signal, and conserved residues that mediate homo- and hetero-dimerization of Msx1 with other transcription factors (Suzuki et al., 2004). PBX is the protein that raises the DNA binding specificity for Msx1 protein (Morgan et al., 2000).

The next important part is the homeodomain. The homeodomain containing 60 amino acids numbered 166–225 (Figure 2.8) is subdivided into the N-terminal arm (N-term) and Helices I, II and III, which conduce to protein stability, DNA binding specificity, transcriptional repression and protein interactions (Zhang et al., 1996; Zhang et al., 1997). In the Msx1 homeodomain, there are several conserved amino acids that establish a consensus sequence, which facilitates tertiary structure and mediates DNA binding activity. Consensus amino acids in Helices I–III form the hydrophobic core and keep the amphipathic nature of the alpha helices, whereas other consensus amino acids in Helix III and in the N-terminal arm mediate DNA binding activity.

Stability of the homeodomain is provided by three salt bridges, which function as "electrostatic cross-links" between each of the three helices (Hovde et al., 2001). The salt bridge, between E30 and K23, links Helix I with Helix II; the salt bridge between E42 and R31 links Helix II with Helix III; the salt bridge between E17 and R52 links Helix III with Helix I. Similar, though not identical, salt bridges are observed in most of the other homeodomain structures, indicating the importance of these electrostatic interactions for domain stability. In the third helix, the highly conserved residue 50 deserves special mention because it is one of the critical amino acids involved in discriminative DNA recognition for distinct classes of homeodomains.

A comparison of the structure of Msx1 homeodomain-DNA complex and that of other homeodomain-DNA complexes, revealed that the most significant differences are an extension, by several amino acids of homeodomain proteins, of the interaction of the N-terminal arm with the minor groove, and the degree of DNA bending seen in this complex (Hovde et al., 2001). The structure of the Msx-1 homeodomain is

similar to those of other homeodomains. This structural homology is maintained due to the dispensation of several highly conserved hydrophobic amino acids that constitute the small hydrophobic core of the domain. Mutations of these amino acids severely compromise both the DNA-binding affinity as well as the transcriptional repression activity of the Msx1 homeodomain, indicating the importance of these amino acids for structural integrity (Zhang et al., 1996). The comparison found that the N-terminal arm is the most significant structural difference between homeodomains (Hovde et al., 2001).

In addition to its interaction with DNA, the Msx1 homeodomain has physical interaction with proteins that are important for the biological process of Msx1. One study of interactions and transcriptional repression by the general transcription factor TFIIF found that the specific amino acids, K3, R5, and F8, are required for these mechanisms (Zhang et al., 1996).

It has been reported that Msx1 interacts with distal-less homeobox (Dlx) and TATA box binding protein (TBP) (Catron et al., 1995; Zhang et al., 1996). Both of these interactions have been identified to be mediated by the homeodomain, specifically amino acids in the N-terminal part of homeodomain. A study of the overlapping expression patterns of Msx1 and Dlx genes and their involvement in epithelial–mesenchymal signaling cascades of murine odontogenesis suggested that Msx and Dlx proteins form hetero-dimeric complexes *in vivo,* which provide a mechanism for transcriptional regulation by functional antagonism (Zhang et al., 1997).

Both the Helix I part and part of homeodomain adjacent to the N-terminal domain play important roles in transcriptional repression (Hovde et al., 2001). On the

other hand, mutants in Helix III are still capable of transcriptional repression, indicating that the DNA binding part is not required for transcription activity.

A similar interaction surface is used in the interactions between Msx1 and other homeodomain proteins such as Pax3 (Bendall et al., 1999), Pax9 (Ogawa et al., 2006), Lhx1 (Bendall et al., 1998); and Dhx2 (Bendall et al., 1998; Bendall et al., 1999). Moreover, Msx1 is also able to bind with Msx1 as homodimeric complexes or with Msx2 as heterodimeric complexes (Zhang et al., 1997).

The C-terminal region of Msx1 is required for interaction with PIAS1. The N-terminal region (amino acids 1–200) of PIAS1, including the SAP domain, is necessary for interaction with Msx1 (Figure 2.7) (Lee et al., 2006). The interaction between Msx1 and PIAS1 is highly specific, because PIAS1 does not interact with Msx2. It is noteworthy that this is the first biochemical activity that discriminates Msx1 and Msx2, which are otherwise virtually indistinguishable (Catron et al., 1996).

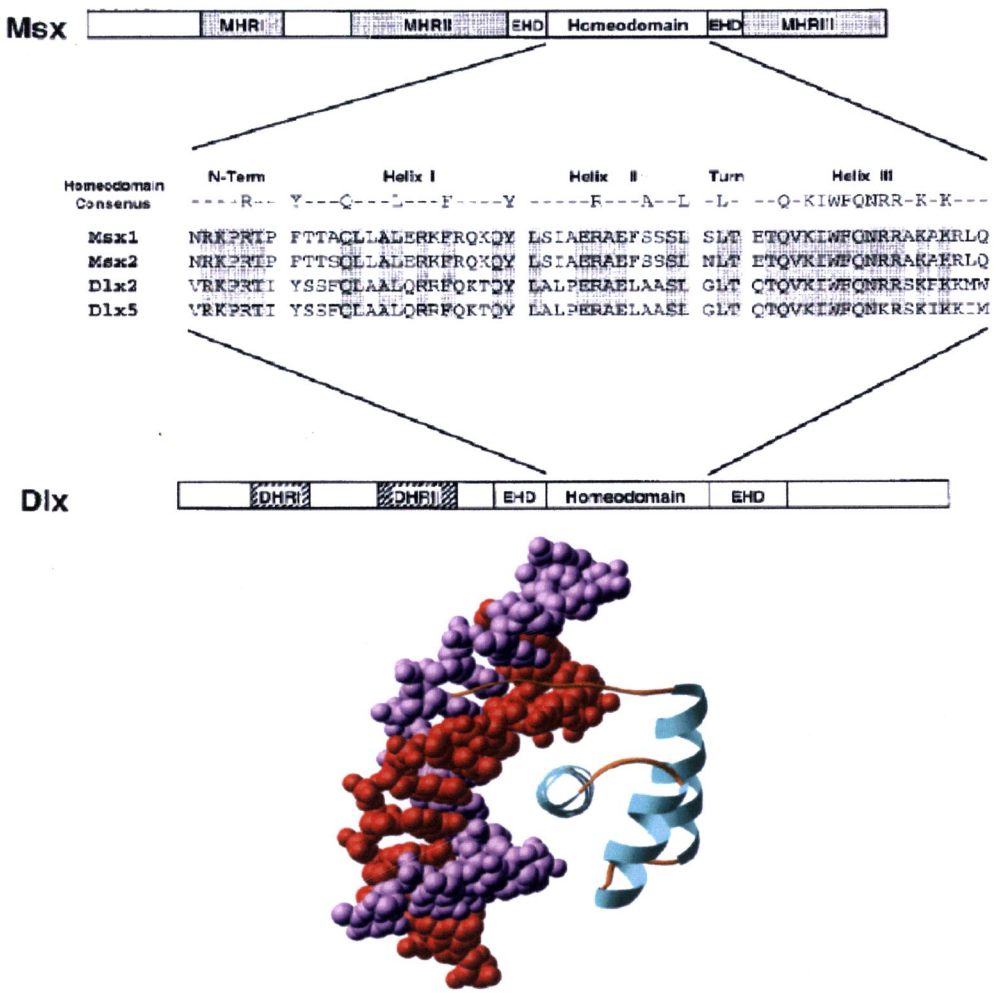


Figure 2.8 Comparison of Msx and Dlx proteins. The upper figure shows amino acid sequences in the homeodomain and indicates the position of the N-terminal part (N-term) and Helices I, II, and III (Zhang et al., 1997). The lower figure shows Msx1 protein model, demonstrating three blue helices (Hovde et al., 2001).

2.1.5 Knockout of *Msx1* $^{-/-}$ in mice

In an *Msx1* knockout study, Satokata and colleagues have revealed that homozygous *Msx1* $^{-/-}$ mice died immediately after birth and showed severe craniofacial abnormalities. Their phenotypes consisted of a complete cleft of the secondary palate, a failure of tooth and alveolar bone development in the mandible

and maxilla, and abnormalities of the skull, malleus, nasal bones and conchae. Analysis of *Msx1* transcripts has revealed that there are no noticeable *Msx1* transcripts in homozygous mice, with a moderate amount of transcripts in heterozygotes (Satokata and Maas, 1994).

2.1.5.1 Perinatal lethality in *Msx1*$^{-/-}$ homozygotes

All homozygous mice died within 24 hours of birth. Before death, they exhibited gasping respirations and cyanosis. Moreover their stomachs were lacking of milk and enlarged with air (Satokata and Maas, 1994).

2.1.5.2 Abnormal tooth development

In *Msx1* deficient mice, the incisor tooth bud could not be seen and first and second molars have failed to develop past the bud stage. Tooth development also exhibited a 24 hour delay in progression when compared to the wild type controls, which were at the early bell stage (Figure 2.9). The retardation of odontogenesis at the bud stage strongly suggests that *Msx1* is necessary for mesenchyme to respond to the induction by epithelial Bmp4. *Msx1* expression in mesenchymal cells is evidently important not only for differentiation of the mesenchyme into the dental papilla and dental follicle, but also reciprocally for the subsequent development of the epithelial tooth bud to the cap and bell stages (Satokata and Maas, 1994).

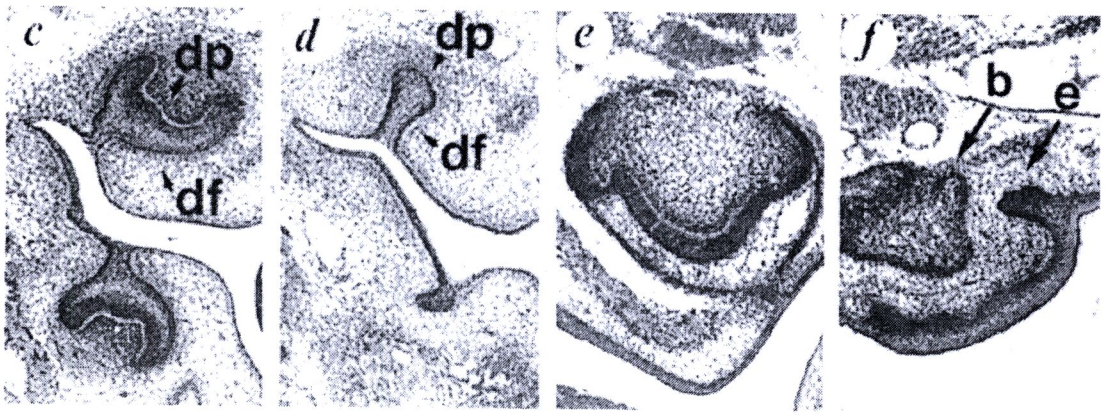


Figure 2.9 Odontogenesis in mice. c,d, Wild type (c) and mutant (d) molar development. The wild type molars are at the early bell stage, but the mutant molar tooth buds are delayed at the bud stage, despite a small amount of mesenchymal condensation can be observed. dp, dental papilla; df, dental follicle. e,f, Wild type (e) and mutant (f) newborn upper first molar development. In mutant, the ectodermal invagination labeled "e" stands for the residual tooth bud which has failed to develop beyond this stage; "b" represents ectopic bone which forms a rudimentary palatal ridge in the mutant. Magnification, 100x (Satokata and Maas, 1994).

2.1.5.3 Cleft palate in *Msx1* deficient mice

Homozygotes exhibited a complete cleft of the secondary palate (Figure 2.10). Maxillary and palatine shelves were absent (Figure 2.11). At E14.75, the palatal shelves were normally elevated, but were not fused with each other or with the nasal septum (Figure 2.12) (Satokata and Maas, 1994).

Expression of *Msx1* was observed in the early palatal shelves at E12, but only in the anterior region. This expression could be related fusion of the palate because fusion starts anteriorly at the same time as *Msx1* expression, and then proceeds posteriorly. Regarding tooth abnormalities, Satokata and Maas noted that the dental follicle mesenchyme is close to the developing palatal shelves (Satokata and Maas, 1994). Therefore, it is possible that the absence of *Msx1* expression in the dental

follicle could be associated with the impaired palatal progression exhibited in *Msx1* mutant mice and these would suggest the relationship between the occurrence of cleft palate and tooth anomalies (Satokata and Maas, 1994).

One important difference between cleft palate in humans and in the *Msx1* mutant mouse is that in the mouse, the cleft palate is a completely penetrant recessive mutation. In contrast, the mode of inheritance in humans is autosomal dominant. Partial loss of function could be the cause of an incompletely penetrant cleft palate phenotype in the mouse. In addition, different genetic backgrounds may cause different phenotypes (Satokata and Maas, 1994).

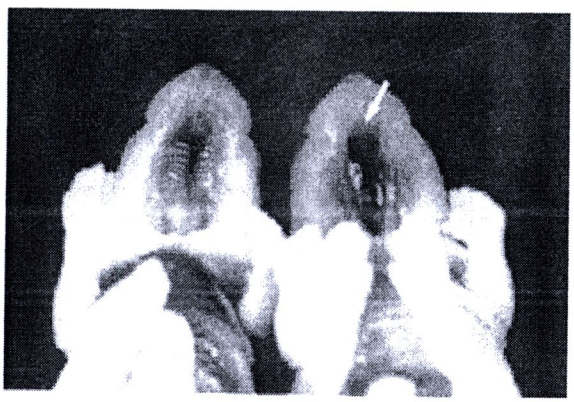


Figure 2.10 Palate of wild type (left) and mutant (right). The arrow indicates the anterior nasal septum. The primary palate is intact. The cleft is a complete cleft of the secondary palate, but the limbs of the homozygote appear normal (Satokata and Maas, 1994).

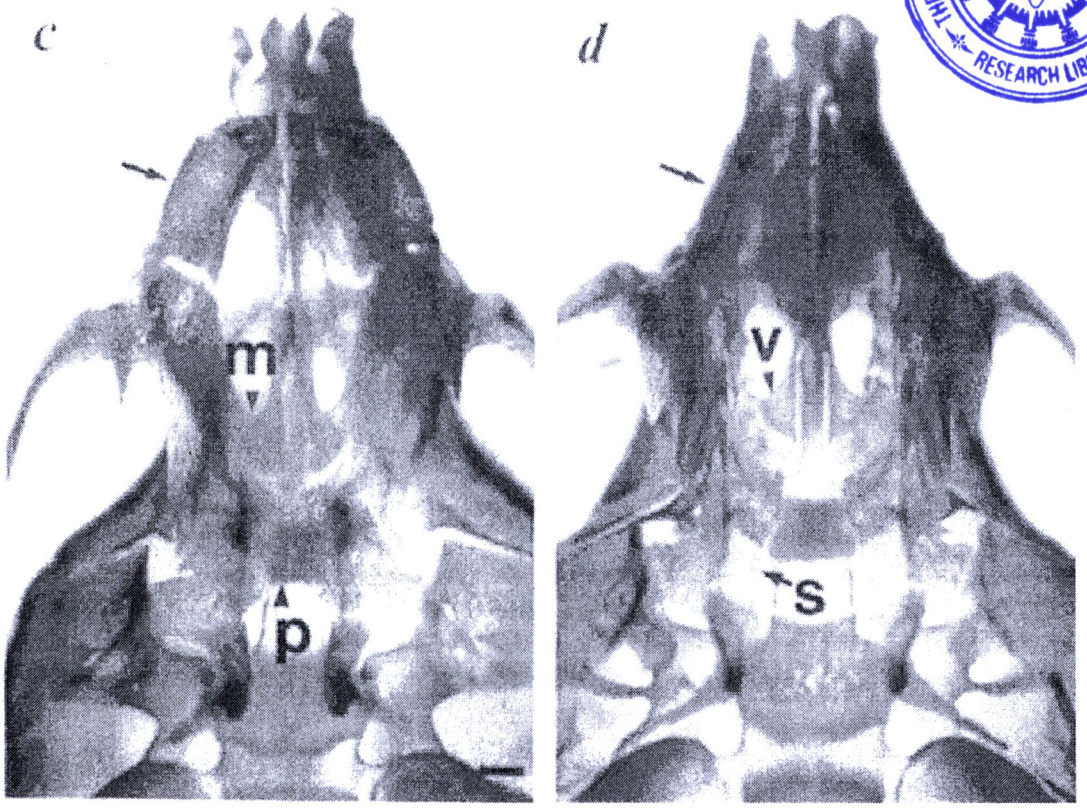


Figure 2.11 Skeletal stain of wild type (c) and mutant (d), exhibiting the maxillary (m) and palatine (p) shelves. In the mutant, both shelves are absent, permitting direct view of the sphenoid bone (s) and vomer (v). Appearance of the premaxilla in the wild type looks rounded when compare to the pointed appearance in mutant (Satokata and Maas, 1994).

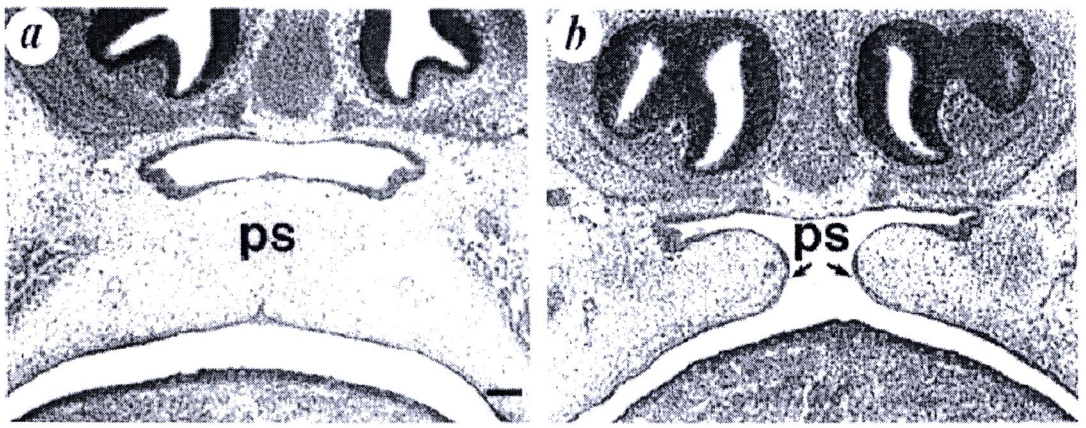


Figure 2.12 Palate development in wild type (a) and mutant (b) E14.75 specimens revealing failure of palatal fusion in mutant. The palatal shelves (ps) are elevated in the mutant, but 24h retarded in their development in this stage. Magnification, 100x (Satokata and Maas, 1994).

2.1.5.4 Maxillary and mandibular defects in *Msx1*$^{-/-}$ mice

The main deficient portion was alveolar bone related to the tooth bearing areas for the incisors and molars (Figure 2.13, 2.14). In maxilla, the deficiency of bone was seen in the premaxillary area. The upper incisors and alveolar bone were also absent (Satokata and Maas, 1994).

Examination of the mandible revealed alveolar process deficiency, foreshortening of the tooth-bearing mandible, and absence of incisor tooth buds. The alveolar bone osteoblasts differentiate from the dental follicle mesenchyme. It seems that the alveolar deficiency is related to a failure of the dental follicle mesenchyme to differentiate into alveolar bone osteoblasts (Satokata and Maas, 1994).

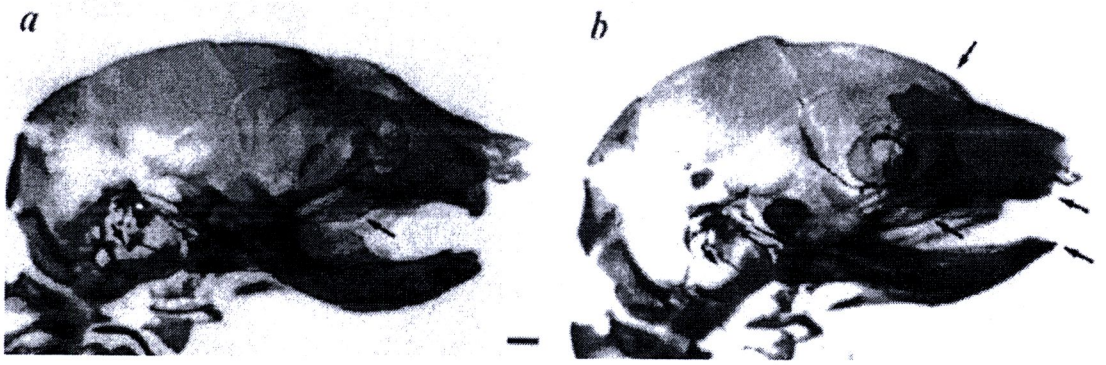


Figure 2.13 Skeletal stain of the heads of wild type (a) and mutant (b). Arrow show the more rounded appearance in mutant and hypoplasia of the maxilla and mandible , due to an absence of upper and lower incisors and alveolar bone. Magnification, 12x (Satokata and Maas, 1994).

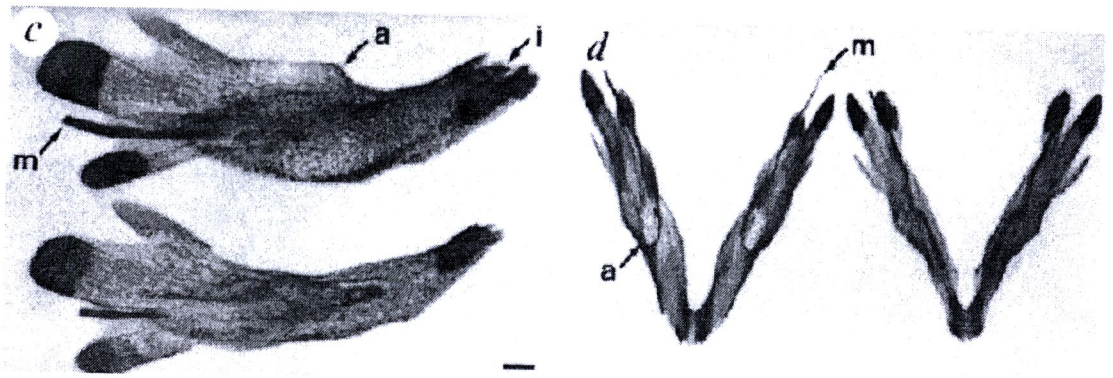


Figure 2.14 Lateral view (c) in wild type (upper) and mutant (lower). Overall mandibular length is reduced and there are abnormalities in shape of the lower border in mutant. Upper view (d) in wild type (left) and mutant (right). Note the deficiency in the height of alveolar ridge in mutant, due to an absence of alveolar bone in molar region. There are defects in anterior segment of mandible and symphysis menti in mutant. Abbreviations: a: alveolar ridge; i: lower incisor; m: Meckel's cartilage. Magnification, 20x (Satokata and Maas, 1994).

2.1.5.5 Abnormal skull and nasal bone development

Deficiency of the medial portions of the frontal bones caused a widening of the anterior fontanelle, and widening of the sutural (wormian) bone in the metopic suture leading to subtle differences in the form of the middle and anterior facial regions in the affected mice (Figure 2.15). Furthermore, the nasal bone looked more rectangular than in wild type. These were related to the fact that in the developing skull, *Msx1* is expressed in the cellular precursors (Satokata and Maas, 1994).

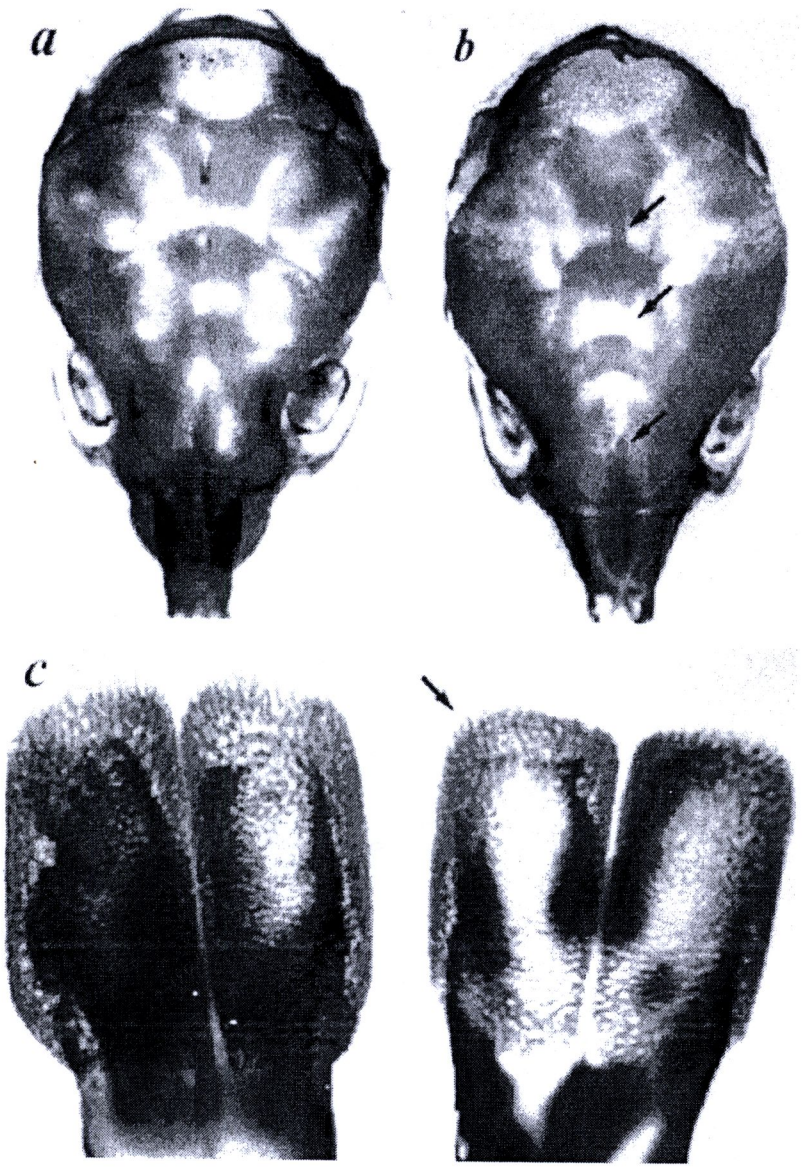


Figure 2.15 Upper view of skull in wild type (a) and mutant (b). The abnormalities are: Overlapping parietal bones (upper arrow), deficiency of the medial portion of frontal bones, resulting in large open area in skull (middle arrow), and enlargement of the frontal sutural bone (lower arrow), which is small and unclearly present in wild type. c, the nasal bones are more rectangular and smaller in mutant (arrow) (Satokata and Maas, 1994).

2.1.5.6 Middle ear abnormalities

Regarding the phenotype of middle ear (Figure 2.16), the incus and stapes were normal, but the short process (prosessus brevis) of the malleus failed to develop in all homozygous mice. The malleus was also decreased in height. These manifestations were related with the first pharyngeal arch derivation of the malleus and indicated the role of *Msx1* in the bone of the middle ear (Satokata and Maas, 1994).

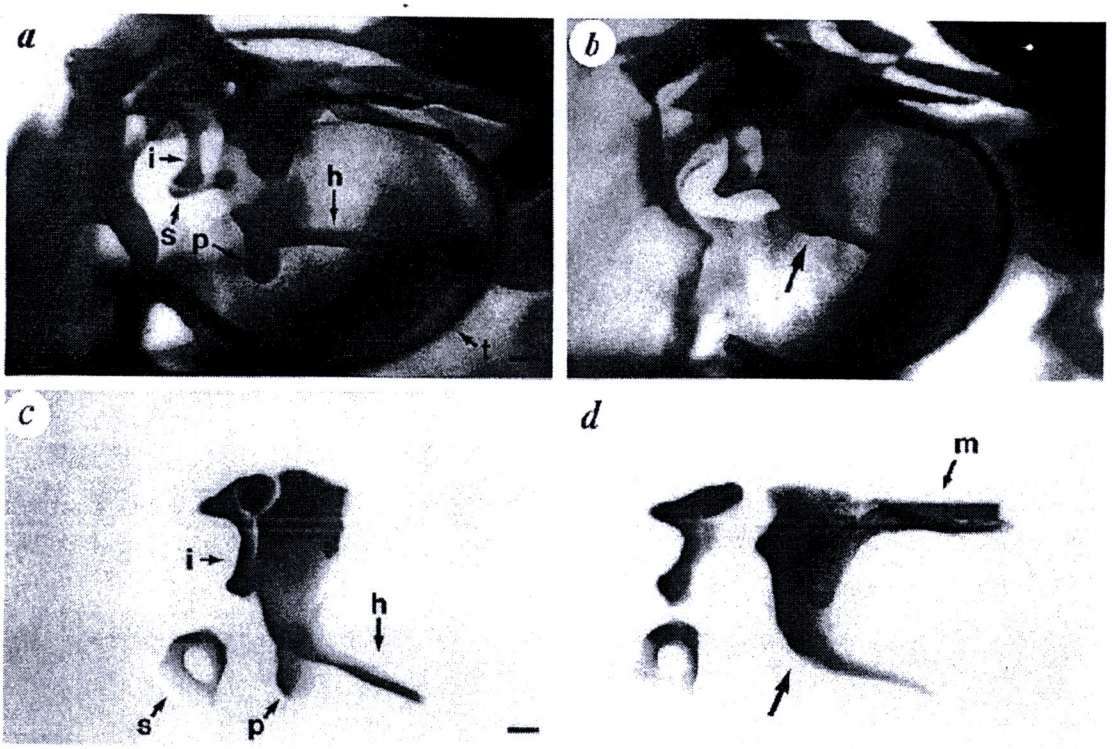


Figure 2.16 Skeletal stain analyses of wild type (a, c) and mutant (b, d). The short process or processus brevis of the malleus (p) is absent in mutant (arrow). Abbreviation: p: malleus; i: incus; s: stapes; h: mallear handle; t: tympanic ring; m: Meckel's cartilage. Magnification, 50x (Satokata and Maas, 1994).

2.1.5.7 Other organs

Msx1 also express in several organs such as the limbs, ciliary body of eye, genitourinary tract and brain, but the *Msx1* mutant mice had no detectable phenotypes in these organs. This investigation included the choroid plexus, meninges, pituitary and regions derived from the dorsal neural tube. The major anomalies were seen in the organs that are derived from the first pharyngeal arch (Satokata and Maas, 1994).

2.1.6 Roles of *MSX1* in development

2.1.6.1 *Msx1* gene encode transcription repressor

The Msx proteins are necessary regulators of craniofacial, limb, and nervous system development. They are regulatory proteins that function as transcriptional repressors *in vitro* and *in vivo* (Catron et al., 1993; Catron et al., 1995; Catron et al., 1996).

Transcriptional repression is a mechanism that involves protein-protein interactions. It leads to target gene selection and transcription regulation by the involvement of the residues within their homeodomain (Zhang et al., 1996). The Msx homeodomain functions directly with the TATA binding protein (TBP) and the core content of the general transcription complex in order to perform transcription repression. Significantly, the ability to interact with the basal transcriptional complex and affect transcription is not dependent on their DNA-binding function (Catron et al., 1995).

In order to regulate transcription, Msx proteins also interact with other homeodomain proteins. Heterodimers formed between *Msx1* and other homeodomain proteins such as *Dlx2*, *Dlx5*, *Lhx2* and *Pax3* cause reciprocal functional antagonism *in*

vitro (Zhang et al., 1997; Bendall et al., 1998; Bendall et al., 1999). It is believed that there may be a regulatory mechanism in tissues in which expression of *Msx1* overlaps those of other proteins.

It has been reported that both Msx1 and Msx2 (another member of the Msx family) had a similar DNA binding site preference and the ability to repress transcription (Catron et al., 1996). They exhibited different biochemical properties by virtue of unique N-terminal domains; Msx2 had a greater affinity for DNA, whereas Msx1 was a more potent repressor.

An investigation of the three-dimensional structure of the Msx1 homeodomain and the DNA complex demonstrated two major distinctions from that of the other homeodomains (Hovde et al., 2001). At first, Msx1 has two proline residues, resulting in great stability and order in the N-terminal arm of the homeodomain, the arm which tracks the minor groove of the DNA. Secondly, the DNA, combined with the *Msx1* homeodomain, exhibited a 28° bend, whereas the normal 21° was observed when combined with other homeodomain proteins.

2.1.6.2 *Msx1* in the role of antisense transcription factor

In situ analysis revealed the presence of an endogenous Msx1 antisense RNA in mice, rats, and humans. This RNA is expressed only in differentiated dental and bone cells with an opposite relationship with Msx1 protein. The balance between the levels of the Msx1 sense RNA and Msx1 antisense RNA is important to the expression of Msx1 protein (Blin-Wakkach et al., 2001).

The effect of this balance is also seen in the Msx-Dlx homeoprotein pathway. Overexpression of *Msx1*, *Msx2*, and *Dlx5* affects the skeletal differentiation. The

expression of Msx1 antisense RNAs is abolished by Dlx5. Msx1 is shown to downregulate a master gene of skeletal cell differentiation, *Cbfa1*. These data have suggested that the ratio between Msx1 sense and antisense RNAs is an important factor in the control of skeletal terminal differentiation.

Location of the initiation site for Msx1 antisense RNA transcription by primer extension in both mouse and human is an identical region, including a consensus TATA box, suggesting evolutionary conservation of the antisense RNA-mediated regulation of *Msx1* gene expression (Blin-Wakkach et al., 2001).

2.1.6.3 *Msx1* induces dedifferentiation of myotubes

It has been reported that terminally differentiated mammalian myotubes can dedifferentiate when stimulated with the appropriate signals and that *MSX1* can contribute to the dedifferentiation process. *Msx1* is expressed in the early regeneration blastema and its expression in the developing mouse limb demarcates the boundary between the undifferentiated (*Msx1*-expressing) and differentiating (no *Msx1* expression) cells (Hill et al., 1989). Ectopic expression of *Msx1* in C2C12 myotubes decreased the nuclear muscle proteins Myod, myogenin, Mrf4, and p21 to insensible levels in 20 to 50% of the myotubes. Approximately 9% of the myotubes cleaved to produce smaller multinucleated myotubes or proliferating, mononucleated cells. Clonal populations of the myotube-derived mononucleated cells could be induced to redifferentiate into cells expressing osteogenic, chondrogenic, adipogenic and myogenic markers (Odelberg et al., 2000).

2.1.6.4 *Msx1* function in repression of myogenic gene expression

By identification a physical interaction between Msx1 and H1b, found that Msx1 and H1b bind to a key regulatory component of Myod, a central regulator of skeletal muscle differentiation, where they induce repressed chromatin. Furthermore, Msx1 and H1b coordinated to inhibit muscle differentiation in cell culture and in Xenopus animal caps. These revealed a theretofore unknown function for linker histones in gene-specific transcriptional regulation (Lee et al., 2004).

2.1.6.5 Msx1 cooperates with Pias1

Interaction of Msx1 with Pias1 is required for Msx1 to function as an inhibitor of myoblast differentiation through repression of myogenic regulatory genes, such as *Myod*. Msx1 sumoylation is not demanded for its inhibitory function or its interaction with Pias1. To localize and retain of Msx1 at the nuclear periphery in mouse myoblast cells, Pias1 is required (Lee et al., 2006).

2.1.6.6 Msx1 cooperates with Lmx1a

Msx1 and Lmx1a are critical intrinsic dopamine neuron determinants. They are determinants of midbrain dopamine neurons in mouse and chicken embryos. Lmx1a is important and sufficient to trigger dopamine cell differentiation, and initial activity of Lmx1a induces expression of *Msx1*, which complements Lmx1a by inducing expression of *Ngn2* and neuronal differentiation. In embryonic stem cells, expression of Lmx1a results in robust generation of dopamine neurons with midbrain identity (Andersson et al., 2006).

2.1.7 Functional redundancy

Msx1 and *Msx2* function redundantly in craniofacial development. *Msx1* and *Msx2* have overlapping expression at several sites of tissue-tissue interaction including the craniofacial structures (Mackenzie et al., 1992; Jowett et al., 1993; Catron et al., 1996). In the developing craniofacial structures, *Msx1* and *Msx2* are expressed in the suture mesenchyme, dura mater, the distal part of the facial primordia, the associated sense organs, and teeth (Mackenzie et al., 1991; Mackenzie et al., 1992). *Msx1* expression reaches into the postnatal stages of skull morphogenesis, whereas rapid decrease of *Msx2* expression is seen after birth (Alappat et al., 2003).

Several studies have suggested that *Msx1* and *Msx2* are functionally redundant. Examination at the biochemical level showed that *Msx1* and *Msx2* have common transcriptional properties and DNA-binding sites (Catron et al., 1996). They both function as transcriptional repressors and recognize the same DNA consensus sites. Moreover, *Msx1* and *Msx2* only differ in one amino acid in their homeodomains. Spatial and temporal expressions of *Msx1* and *Msx2* in mice have revealed a certain degree of co-localization. About co-localization pattern, many examinations found synergistic defects in skull, tooth, ear, limb, hair follicle, and mammary gland development in *Msx1/Msx2* double mutant mice (Bei and Maas, 1998; Satokata et al., 2000; Zhang et al., 2003). So, whereas the *Msx2* single mutant mice show either incomplete or delayed ossification of the skull resulting in calvarial patency, the double mutant mice are deficient in calvarial ossification.

Msx1 and *Msx2* also reveal functional redundancy during the early stages of organogenesis in tooth, hair follicle and mammary gland, coinciding with their overlapping expression pattern. Mice heterozygous for *Msx1* but lacking functional *Msx2* do not manifest early tooth anomalies (Bei and Maas, 1998), suggesting that the participation of *Msx2* in the early stages of tooth development appears nonessential. Odontogenesis develops normally through the lamina, bud, and cap stages in the *Msx2*-deficient mice. A requirement for *Msx2* is presented in the late stages of organogenesis following the downregulation of *Msx1* expression. Defective tooth development in the *Msx2*-deficient mice happens at E16.5 when the stellate reticulum and stratum intermedium fail to develop normally resulting in the degeneration of the ameloblast and ultimately the enamel organ. *Msx2* was postulated to participate in the regulation of the spatiotemporal expression of the amelogenin gene during odontogenesis (Zhou et al., 2000). In the *Msx1* and *Msx2* deficient mice (double knockout mice), the phenotype is greatly amplified. Molar tooth development arrests at the dental lamina stage in *Msx1*/*Msx2* compound mutants unlike *Msx1*-/- mice that reveal a bud-stage arrest (Satokata and Maas, 1994; Bei and Maas, 1998). Thus, *Msx1* and transiently expressed *Msx2* function redundantly at the initiation of odontogenesis.

Likewise, additive phenotype is seen in the developing middle ear of the compound *Msx* mutants. Mice double null for *Msx1* and *Msx2* genes are lacking the manubrium as well as the processus brevis, whereas *Msx1*-/- mice lack the processus brevis only. There is a failure in hair follicle induction in the *Msx1*/*Msx2* double mutant mice, unlike *Msx2*-/- mice where the pelage hairs form but are lost prematurely owing to abnormalities in hair maintenance (Satokata et al., 2000; Ma et

al., 2003). In the same way, the mammary gland epithelium does not invaginate in *Msx1* and *Msx2* double mutant mice while a deficiency of *Msx2* alone results in a sprout stage arrest of the mammary glands. Therefore, a loss of *Msx1* and *Msx2* in the tooth, hair and mammary gland results in a more severe phenotype than those resulting from the loss of either one. Functional redundancy between *Msx1* and *Msx2* is also evident from limb deformities that are exhibited by double knockout mice (Chen and Zhao, 1998). According to the functional redundancy, the single mutant mice have normal limbs. However, *Msx1* and *Msx2* mutant mice develop limbs with lack of antero-posterior polarity, absence of radius, marked inhibition of apoptosis in the interdigital regions, and polydactyly. These studies have suggested that functional redundancy exists between the *Msx* genes.

2.1.8 Expression of *Msx1* genes

In vertebrates, the expression of *Msx1* is seen in many organs which are associated with epithelial-mesenchymal interactions during morphogenesis such as craniofacial structures, tooth buds, and nail beds (Mackenzie et al., 1991). Expression is also seen in the limbs, atrioventricular cushion, Rathke's pouch, genital tubercle, and developing mandible and teeth (Hill et al., 1989; Robert et al., 1989). These organs are derived from the neural folds and adjacent crest cells (Robert et al., 1989)

There has been the study about the spatial and temporal expression of *Msx1* during craniofacial development in the mouse from embryonic day 9.5 (E9.5) to E15.5 by *in situ* hybridization. Expression patterns of *Msx1* in the developing craniofacial parts reveal a possible involvement in cellular differentiation leading to the formation of a number of craniofacial structures, including the anterior pituitary,

external ear, choroid plexus, brain meninges, and skull, nasal pits and forming Jacobson's organs (Mackenzie et al., 1991).

In the developing craniofacial structures, *Msx1* is expressed in the sutural mesenchyme, dura mater, the distal part of the facial primordia, the associated sense organs, and teeth (Mackenzie et al., 1991; Mackenzie et al., 1992). Expression of *Msx1* is overlapped with *Msx2* expression. *Msx1* expression continues into the postnatal stages of skull morphogenesis, whereas rapid decrease of *Msx2* expression is evident after birth (Alappat et al., 2003).

2.1.9 Pathways of *Msx1*

Msx1 is involved in several growth factor signaling pathways and serves in the arrangement of inductive pathways essential for organogenesis. Therefore repeated function of *Msx1* in the *Bmp*, *Fgf*, Endothelin and *Shh* signaling pathways can be found.

Bmp2, Bmp4, Fgf2, Fgf4, Fgf8, and Fgf9 are the growth factors from the oral and/or the dental epithelium that are capable of stimulating *Msx1* expression in the underlying mesenchyme of the maxilla and mandible. The mesenchymal expression of multiple growth and transcription factors such as Bmp4, Fgf3, Dlx2, syndecan-1, and Ptc are downstream of *Msx1* (Chen et al., 1996; Bei and Maas, 1998; Zhang et al., 1999).

2.1.9.1 *Bmp4* and *Fgf8* signaling pathways

A genetic pathway for early odontogenesis is shown to describe association between *Msx1*, *Bmp4*, and *Fgf8* (Figure 2.17). At E11.5 epithelial *Bmp4* and *Fgf8*

require *Msx1* to stimulate members of their own gene families in dental mesenchyme. Both epithelial *Bmp4* and *Fgf8* can induce *Msx1*. Then *Msx1* is able to act as the signaling mediator for mesenchymal downstream genes. Moreover, epithelial *Bmp4* cannot induce mesenchymal *Fgfs*, and epithelial *Fgfs* cannot induce mesenchymal *Bmp4* expression, suggesting that *Bmp4* and *Fgf8* function through separate *Msx1*-dependent pathways to induce expression of these downstream genes. At the bud stage of development, *Fgf3* is placed downstream of *Msx1* because its expression is decreased in the *Msx1* deficient dental mesenchyme. *Dlx2* is placed downstream of mesenchymal *Bmp4* because *Bmp4* induces *Dlx2* expression even in the *Msx1* mutant mice, and *Dlx2* expression is reduced in the *Msx1* deficient dental mesenchyme (Bei and Maas, 1998).

There have been several studies about genetic relationship between *Msx1* and *Bmp4*. They exhibited that *Bmp4* expression was decreased in the *Msx1* deficient mesenchyme but is not affected in *Msx1* deficient epithelium (Chen et al., 1996; Bei and Maas, 1998). These suggest that *Msx1* is required for the expression of *Bmp4* in the dental mesenchyme and that *Bmp4* functions downstream of *Msx1* in the dental mesenchyme. *Msx1* does not influence the expression of epithelial *Bmp4* and therefore acts upstream of *Msx1*. Investigation has revealed that epithelial *Bmp4* can induce the *Msx1* and *Bmp4* expression in the dental mesenchyme. Despite the fact that epithelial *Bmp4* can induce its own expression and that of *Msx1* in the dental mesenchyme, *Bmp4* is not able to substitute for all the inductive functions of the dental epithelium.

Fgf8 acts as an inductive signal in several developmental organs (Vogel et al., 1995; Lee et al., 1997; Richman et al., 1997). *Fgf*-dependent signaling pathway

involve in odontogenesis. *Fgfs* may be an additional content of the signaling pathway mediating odontogenic epithelial-mesenchymal interactions. It is possible that *Fgf8* function is analogous to the function proposed for mesenchymal *Bmp4* (Chen et al., 1996), mesenchymal *Fgfs* bind to epithelial receptors to induce further epithelial development. *Fgf8* is able to induce a translucent zone in dental mesenchyme and has been proposed to function antagonistically with *Bmp4* to determine the regions of tooth initiation (Cam et al., 1992; Neubuser et al., 1997; Kettunen and Thesleff, 1998).

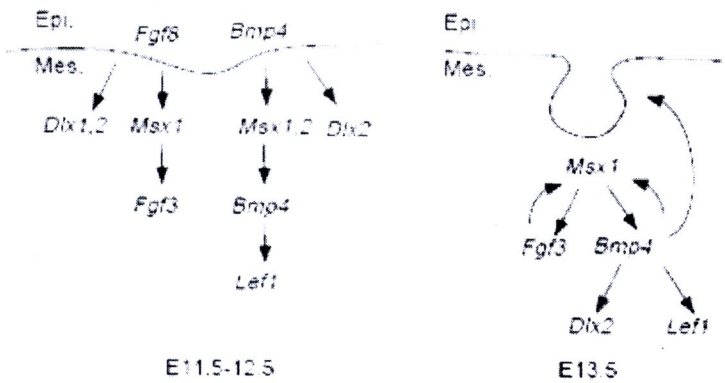


Figure 2.17 BMP4 and FGF8 signaling pathways (Bei and Maas, 1998).

2.1.9.2 Endothelin signaling pathway

Endothelins and their receptors were initially known for their roles in the regulation of blood pressure. Endothelins are members of the family of conserved 21-amino acid peptides that includes three members encoded by different genes (*Edn1, Edn2, Edn3*). It has been described that homozygous mutant mice for Endothelin receptor type A (*Ednra*), Endothelin-1 (*Edn1*) or Endothelin Converting Enzyme-1

(*ECE-1*) died immediately after birth due to severe anomalies in the formation of neural crest derivatives (Clouthier et al., 1998; Yanagisawa et al., 1998). It has been demonstrated in the mouse embryo that Edn1/Ednra is important for the patterning process in the pharyngeal arch by regulating the expression of Dlx6/Hand2 transcription factors (Thomas et al., 1998; Charite et al., 2001; Yanagisawa et al., 2003) through Gαq/Gα11 proteins (Ivey et al., 2003). In the mouse embryo, there is the expression of *Ednra* in migrating cranial neural crest cells and ectomesenchymal cells in the pharyngeal arches (Clouthier et al., 1998; Yanagisawa et al., 1998). *Ednra* function lies downstream the transcription factor *Msx1* in the genetic cascade of neural crest induction (Figure 2.18).

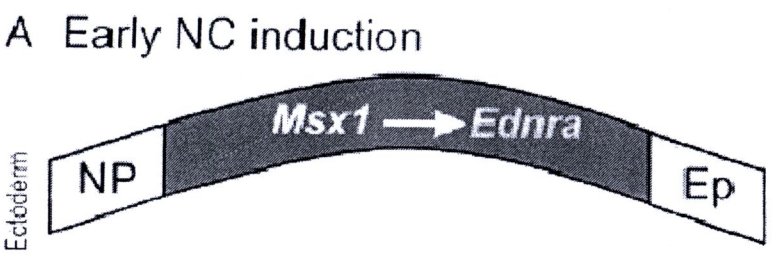


Figure 2.18 *Msx1* is expressed in the neural fold and works upstream of *Ednra*. This pathway is required for neural crest maintenance and for cell survival (Bonano et al., 2008).

Moreover, *Msx1* is also expressed downstream of *dHAND*, a mesenchymal transcription factor in the Endothelin-1 signaling pathway (Figure 2.19). *Msx1*, *Msx2* and the bHLH transcription factor *dHAND* are expressed in a similar pattern in the mesenchyme of the branchial arches, which is the leading edge of the growing arch. *Msx1* may promote cellular proliferation. In *dHAND* mutant mice, *Msx1* expression is absent and the branchial arches become hypoplastic. Together, dHAND and Msx1

play a role in the earliest stages of development of the distal ectomesenchyme of the branchial arches (Thomas et al., 1998).

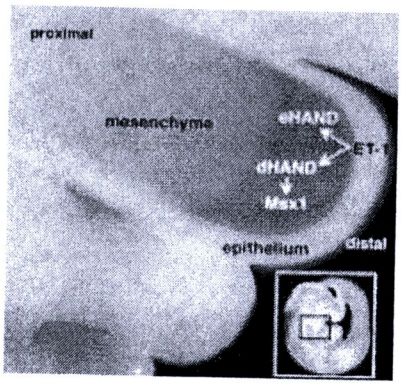


Figure 2.19 *Endothelin-HAND-Msx1* pathway regulates branchial arch growth. Endothelin-1 (ET-1) is secreted from the branchial arch epithelium into the mesenchyme and induces *dHAND* and *eHAND* expression in the distal mesenchyme. *dHAND* in turn controls expression of *Msx1* also in the distal branchial arch. An appropriate balance of Msx1, Msx2 and Dlx2 in the distal arch may be critical for normal development of branchial arch proliferation, differentiation and cell death (Thomas et al., 1998).

2.1.9.3 Sonic hedgehog (*Shh*) signaling pathway

Shh is expressed in the dental epithelium of the early developing molar tooth bud, and the transcripts of *Shh* downstream target genes *Ptc* and *Gli1* are expressed in dental epithelium as well as underlying mesenchymal tissue (Figure 2.20). At this time, the homeobox gene *Msx1* is also expressed in the dental mesenchyme of the molar tooth germ. The expression of *Ptc*, but not *Gli1*, was downregulated in the dental mesenchyme of *Msx1* deficient mice. In wild-type E11.0 molar tooth mesenchyme Shh-soaked beads stimulated the expression of *Ptc* and *Gli1*. However, in *Msx1* deficient dental mesenchyme Shh-soaked beads had the ability to induce *Gli1* but the ability to induce *Ptc* expression was not found, determining a requirement for

Msx1 in the induction of *Ptc* by Shh. Furthermore, another signaling molecule, *Bmp4*, could induce *Ptc* expression in wild-type dental mesenchyme, but induced a distinct expression pattern of *Ptc* in the *Msx1* mutant molar mesenchyme. These suggest that *Msx1* is a component of the *Shh* signaling pathway that induces *Ptc* expression (Figure 2.20). In addition, the pattern of *Ptc* expression in the tooth-developing regions is regulated and coordinated by at least two inductive signaling pathways (Zhang et al., 1999).

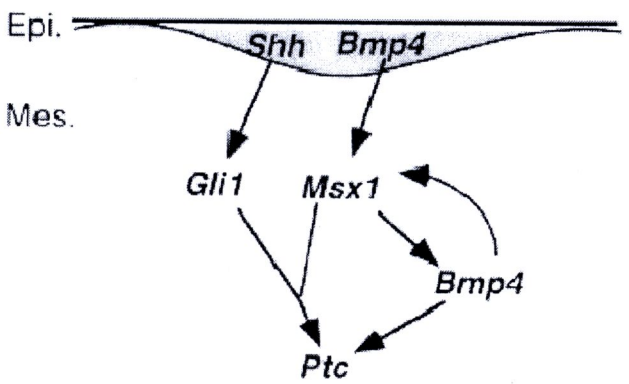


Figure 2.20 Genetic pathway regulating *Ptc* expression in the early odontogenesis. Epithelial *Bmp4* induces the mesenchymal *Msx1* which is required for the mesenchymal *Bmp4* expression. Epithelial *Shh* may induce *Ptc* expression via the induction of *Gli1*. *Msx1* participates in this induction of *Ptc* by *Shh* probably by interacting with *Gli1*. Mesenchymal *Bmp4* can also regulate *Ptc* expression. *Ptc* expression is thus regulated by at least two separated pathways in the early mouse tooth germ (Zhang et al., 1999).

2.1.9.4 Retinoid regulation of *Msx* genes

The regulation of *Msx* gene expression is completed by several mechanisms involving retinoids. Retinoid regulation of *Msx* genes is recognized following the

identification of a retinoic acid-responsive enhancer part in the 5′ flanking region of human *MSX1*. Endogenous retinoids regulate the spatial expression of *Msx1* by restricting its expression to the posterior regions of quail embryos at the gastrulation and neurulation stages, suggesting that retinoids are important regulators of normal *Msx1* expression in avians. In contrast, retinoic acid seems to inhibit *Msx1* expression in mouse embryonic palate mesenchymal cells (Nugent and Greene, 1998). Moreover, the 5' upstream part of mouse *Msx1* have several enhancer elements including three potential NFκB-binding sites and the Msx1 consensus binding site (Takahashi et al., 1997; Shetty et al., 1999). These studies have indicated several regulations of *Msx1* expression.

2.1.9.5 The association of *Msx1* and *Dlx5*

Specifying transcriptional activities of homeoproteins require protein-protein interactions. For example, Msx and Dlx homeoprotein families form homo and heterodimeric complexes. Msx and Dlx dimerized protein is mediated through their homeodomains and that the residues required for this interaction correspond to those necessary for DNA binding. Association of Msx and Dlx proteins does not promote cooperative DNA binding, unlike most other known examples of homeoprotein interactions; instead, dimerization and DNA binding are mutually exclusive activities. Msx and Dlx proteins interact independently and noncooperatively with homeodomain DNA binding sites and that dimerization is specifically inhibited by the presence of such DNA sites. Furthermore, the transcriptional properties of Msx and Dlx proteins display reciprocal inhibition. Msx proteins function as transcriptional repressors and Dlx proteins function as activators, whereas in combination, Msx and

Dlx proteins counteract each other's transcriptional activities. In mouse embryogenesis during limb bud and craniofacial development, the expression patterns of *Msx* and *Dlx* genes (*Msx1*, *Msx2*, *Dlx2*, and *Dlx5*) overlap, consistent with the potential for their protein products to interact. These suggest that mechanism for regulating the transcriptional actions of Msx and Dlx homeoproteins results from functional antagonism through heterodimer formation (Zhang et al., 1997).

2.1.10 Sumoylation and Msx1

Sumoylation [The reversible conjugation of Small Ubiquitin-like Modifier (SUMO) to protein substrates] is a post-translational modification associated with various cellular processes such as transcriptional regulation, nuclear-cytosolic transport, gene expression, protein stability, apoptosis, response to stress, and progression through the cell cycle (Hay, 2005). SUMO is a small protein (100 amino-acids in length and 11 kDa in mass) that can be covalently attached to and detached from specific target proteins to modify their functions. There are three isoforms in humans: SUMO-1, SUMO-2 and SUMO-3. SUMO1 binds target proteins as a monomer, whereas SUMO2 and SUMO3 harbor sumoylation sites themselves, making it possible to bind target proteins as polymers (Johnson and Gupta, 2001).

Most target proteins contain the tetrapeptide consensus motif Ψ-K-x-D/E, where Ψ is a large hydrophobic amino acid, most commonly isoleucine or valine, K (lysine) is the site of SUMO- conjugation, x is any amino acid, D or E is an acidic residue (Pauws and Stanier, 2007). SUMO attachment to its target uses four enzymatic reactions (Figure 2.21). First, the pre-protein is cleaved by a SUMO protease. Next, SUMO is inserted in an ATP-dependent step to an E1-activating

enzyme. After that, activated SUMO forms a thioester, intermediate with the E2-conjugating enzyme Ubc9. Finally, SUMO is covalently bound to its target, mediated by an E3 ligase, which can be one of several protein inhibitors of STAT (PIAS) proteins. This attachment is reversible through a family of sentrin/SUMO-specific proteases (SENPs).

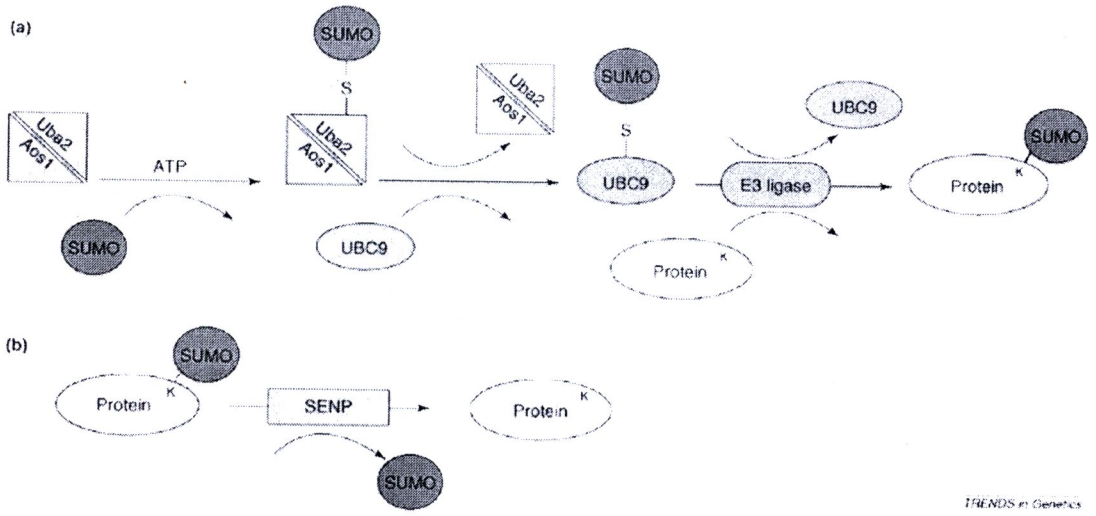


Figure 2.21 Sumoylation of a target protein. (a) SUMO is first attached to a complex of Uba2 and Aos1 proteins (the E1-activating enzyme) by a thioester bond (S) using ATP as a catalyst. Second, SUMO is transferred to UBC9 (the E2-conjugating enzyme). Finally, transfer of SUMO to the appropriate lysine (K) on the target protein is accelerated by an E3 ligase. (b) Proteins can be desumoylated through the activity of sumo-specific proteases (SENPs) (Pauws and Stanier, 2007).

Sumoylation also plays a key role in craniofacial development (Pauws and Stanier, 2007). Several genes, such as *Pax9, Eya1, Satb2, p63, Tbx22, Trps1, Sox2, Sox9, Sox10, Smad4,* and *Msx1* have been identified that, when mutated, cause either nonsyndromic CL/P or syndromic forms with other features such as hypodontia, ectodermal dysplasia or ankyloglossia. The resulting proteins of these several cases have been shown to be targets of SUMO modification (Pauws and Stanier, 2007).

Msx1 is one of the target proteins of SUMO (Gupta and Bei, 2006). Msx1 sumoylation establishes a regulatory mechanism modulating Msx1 function during organogenesis. Msx1 is sumoylated *in vivo* at two lysine residues (Lys9 and Lys127) (Gupta and Bei, 2006) as shown in Figure 2.22. Both target lysines are located in the Msx1 N-terminal domain, which is associated with transcriptional repression. It is possible that SUMO conjugation may play an important role in regulating the ability of Msx1 to interact with distinct transcriptional co-factors to regulate transcriptional output. Some studies suggest that SUMO1 haploinsufficiency leads to cleft lip and palate (Gupta and Bei, 2006). Sumoylation of the other cooperating factors of Msx1 is also important in controlling their transcriptional synergy in a promoter and cell-specific manner. Lee and colleagues have suggested that the major effect of sumoylation is not on its repression activity but on its subnuclear localisation (Lee et al., 2006). In addition, these lysine residues are conserved across several species, suggesting that this modification is likely to be conserved throughout mammalian evolution.

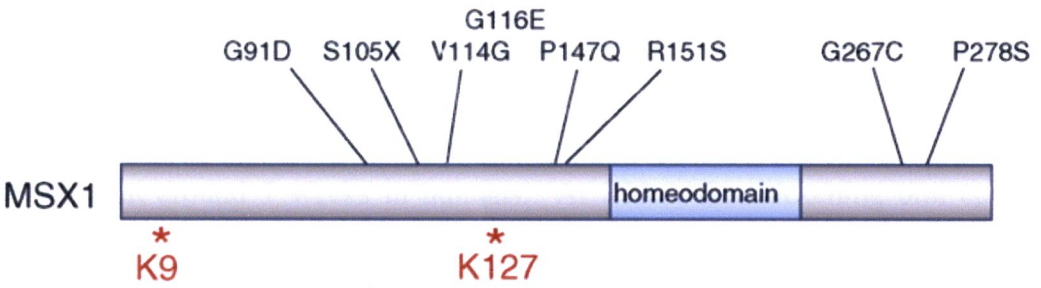


Figure 2.22 Lysine residues (Lys9 and Lys127) that are covalently attached to SUMO are marked by an asterisk (in red). Recognized homeodomain is shown in blue. Missense and nonsense mutations found in patients with CL/P are also shown (Pauws and Stanier, 2007).

In the part of pathways of *Msx1*, reveal that *Msx1* is common to several growth factor signaling pathways and provides in the arrangement of inductive pathways essential to organogenesis. MSX1 is important for several organs, including teeth. Understanding mechanism of tooth development will emphasize the important of MSX1.

2.2 Tooth development

2.2.1 Stages of tooth development

Dental morphology consists of interactions between epithelial and mesenchymal cells (Jernvall and Thesleff, 2000) control cell activities like proliferation, condensation, adhesion, migration, differentiation and secretion that result in the formation of a tooth organ. Organogenesis primarily consists of three elementary processes: 1) initiation; a group of cells translates positional information provided by other cells to begin organ formation at both the right place and time, 2) morphogenesis; the cells establish an organ element and 3) differentiation; cells form organ-specific structures (Peters and Balling, 1999).

Human tooth development (Figure 2.23) begins as a thickening of the oral epithelium that appears at around the 11.5th day of embryonic development (E11.5) in mice and at around 7 weeks in humans. At E13.5, the thickening oral epithelium starts to invaginate into the underlying neural crest-derived mesenchyme to form the tooth bud. Then this mesenchyme condenses around the invaginating epithelium. At E14.5, the condensing mesenchyme is wrapped around by the epithelium to form a cap. At E18.5, a 'bell'-stage tooth germ is formed. This process is regulated by a signaling center that develops at the enamel knot (the tip of the tooth bud at late bud

stage). The cytodifferentiation occurs during the bell stages, the epithelial cells adjacent to the dental mesenchyme are differentiated to be enamel-forming ameloblasts, and the adjacent dental mesenchymal cells are differentiated to be dentin-forming odontoblasts (Tucker and Sharpe, 2004).

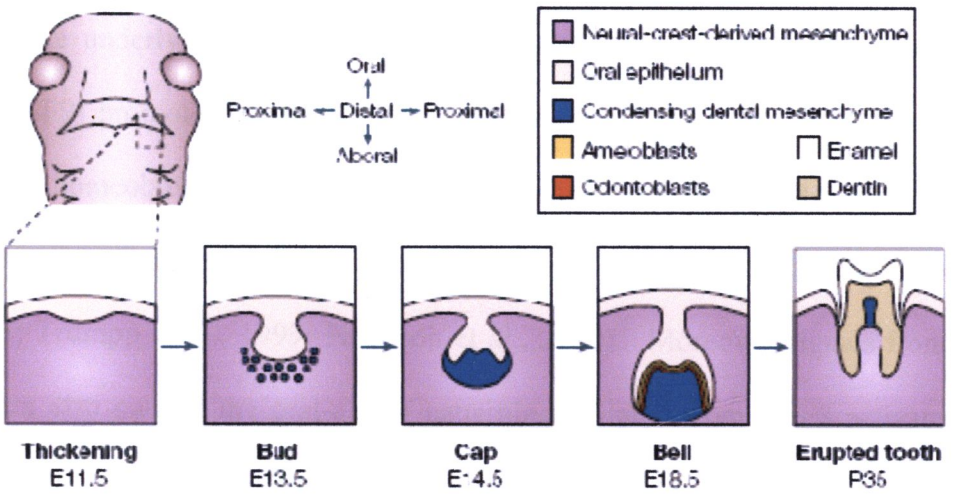


Figure 2.23 Stages of tooth development (Tucker and Sharpe, 2004).

2.2.2 Molecular basis in odontogenesis

Recent advances in our knowledge about molecular aspects of odontogenesis (Cobourne, 1999; Peters and Balling, 1999; Tucker and Sharpe, 1999; Jernvall and Thesleff, 2000) reveal that the development of teeth is under strict genetic regulation, which affects the positions, numbers, and shapes of different types of teeth (Arte et al., 2001). To date, more than 200 genes are known to have a function in odontogenesis. Proteins coded by these genes involve in many stages of dental development and function as transcription factors, signaling molecules and its receptors and extracellular matrix molecules (http://bite-it.helsinki.fi/).

The basic pattern of the dentition, type and size of the teeth are determined to develop by the oral epithelium can be separated into two domains; molar area (proximal part) has the expression of *Fgf8* and *Fgf9* (fibroblast growth factors 8 and 9) and incisor area (distal part) has the expression of *BMP4* (bone morphogenetic protein 4). These signaling molecules then regulate the expression of homeobox genes in the underlying neural crest-derived mesenchyme. FGF8 (and to a lesser extent, FGF9) promotes the expression of *Barx1* (BarH-like homeobox 1) and *Dlx2* (distal-less homeobox 2), while BMP4 promotes the expression of *Msx1* and *Msx2*, and at the same time inhibits the expression of *Barx1* (Bei and Maas, 1998; Tucker et al., 1998; Trumpp et al., 1999; Ferguson et al., 2000). Moreover signals from oral epithelium also control the oral-aboral patterning. Signal from FGF8 activates the expression of *Lhx6/7* (LIM homeobox genes) in the oral region. Then FGF8 positively regulate the expression of homeobox gene *Gsc* (goosecoid) in the aboral mesenchyme of the developing jaw. So the jaw is separated into a tooth-forming Lhx-positive domain and a nontooth-forming Gsc-positive domain (Figures 2.24, 2.25) (Tucker et al., 1999; Thesleff, 2003).

During the initial time of tooth development, the epithelium invaginates into the mesenchyme. In the epithelium (E11.5), at the specific sites of the developing incisor and molar tooth germs are determined by *Shh* (sonic hedgehog) expression and Shh also involves in the tooth bud development by regulating the proliferation of dental epithelial cells (Hardcastle et al., 1998). The interaction between Shh and members of the Wnt family produce the boundary between the oral epithelium and the dental epithelium (Sarkar et al., 2000). *Wnt7b is* expressed in the non-dental oral epithelium (mutual pattern of expression to *Shh*). In the mesenchyme, at the four small sites of

the future invaginations are marked by the paired box gene *Pax9* expression. *Pax9* is positively regulated by Fgf8 and negatively regulated by Bmp4. At bud stage *Pax9* is required for the mesenchymal expression of Bmp4 and Msx1, demonstrating that its function is essential to produce the inductive capacity of mesenchyme (Peters et al., 1998). Moreover Msx1 and Bmp4 also act together in a positive-feedback loop (Tucker and Sharpe, 2004).

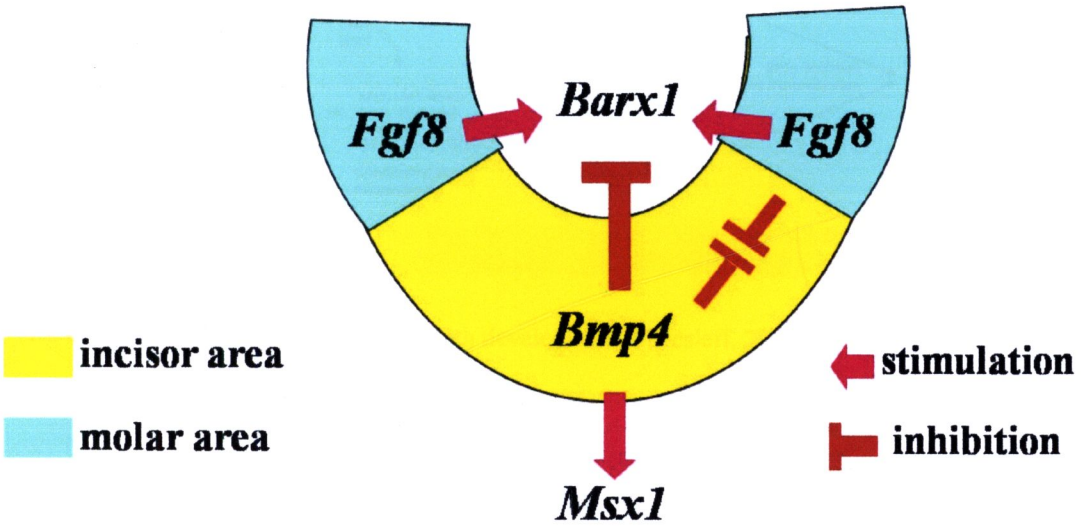


Figure 2.24 Different gene expression in early tooth development. In the epithelium of molar area *Fgf8* upregulates *Barx1* in mesenchyme. In the epithelium of incisor area, *Bmp4* activates *Msx1* and inhibits *Barx1* in mesenchyme. *Bmp4* is antagonistic against *Fgf8*.

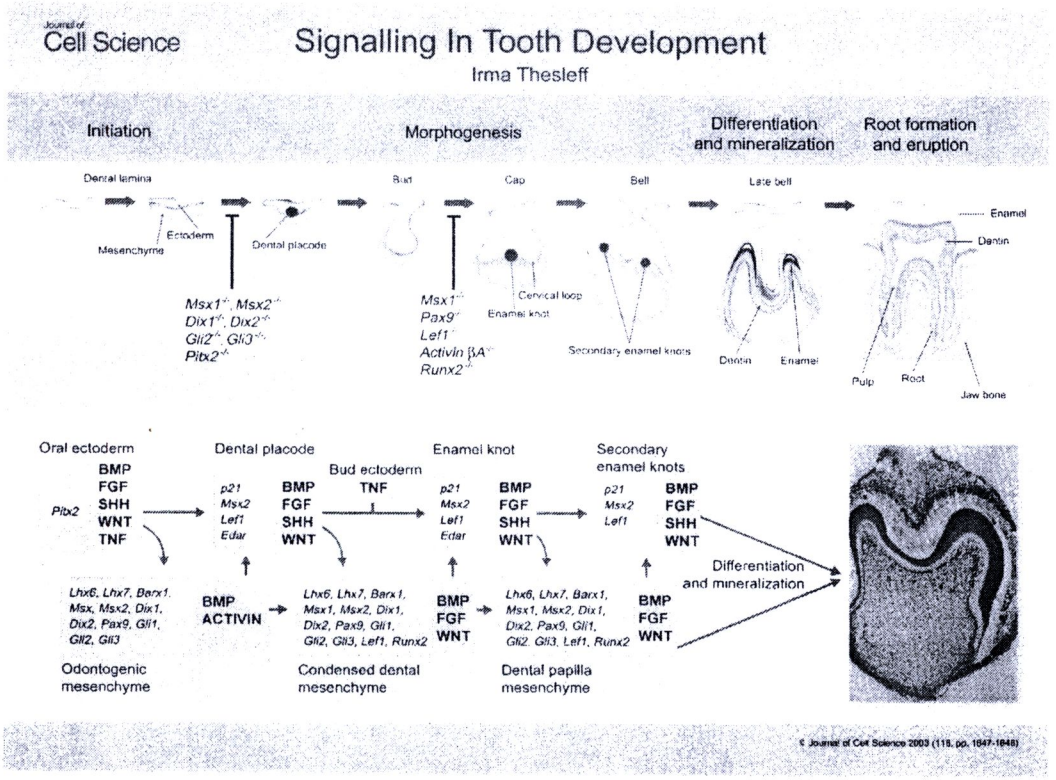


Figure 2.25 Molecular signaling in tooth development. (Thesleff, 2003)

Msx1 gene has an early role in forming the tooth bud in incisor region and also has a later role in the mesenchyme that condenses around each forming tooth bud (Figure 2.26) (Chen et al., 1996). Mutations of *Msx1* result in a block at the bud stage of tooth development and cause tooth agenesis (Satokata and Maas, 1994).

This phenotype has been exhibited to partly result from loss of *Bmp4* in the mesenchyme around the tooth bud. *Msx1* and *Bmp4* act together in a positive-feedback loop. In *Msx1* knockout mice, the early expression of *Bmp4* in the oral epithelium can be found but subsequent expression in the mesenchyme is disappeared. The addition of *Bmp4* to the mesenchyme can rescue molar tooth development in *Msx1* knockout mice at least to the cap stage (Bei et al., 2000; Zhang et al., 2000; Zhao et al., 2000). Whereas the incisor development is consistent with the earlier

epithelial role of *Bmp4*, therefore addition of *Bmp4* to the mesenchyme could not rescue incisor development (Zhao et al., 2000).

2.2.3 *MSX1* and tooth development

Msx genes have a general function in several organs during morphogenesis by responding to and regulating the reciprocal expression of inductive signaling molecules in epithelial and mesenchymal tissues. It is interesting to describe the *Msx* gene expression patterns in terms of a functional requirement for *Msx* in inductive signaling, because this is an alternating and repeated incident in organogenesis. The obvious relation of *Msx* gene expression with that of other growth factors suggests that *Msx* genes may participate in the regulation of a large and diverse set of growth factors in organ forming including odontogenesis.

At the initial step to form the tooth, Bone Morphogenetic Protein 4 (Bmp4) is expressed in the dental epithelium. Bmp4 is the component of the inductive signal that transfers tooth inductive potential from dental epithelium to mesenchyme. *Bmp4* expression is first detected in the molar tooth at E11.5 in the dental lamina epithelium, but then is shifted to the dental mesenchyme at E12.5, coincident with the shift in dental developmental potential between tissue layers (Tureckova et al., 1995). Bmp4 in epithelium can induce the expression of *Msx1* in dental mesenchyme.

Msx1 is absolutely expressed in the dental mesenchyme and absent from the dental epithelium throughout the bud, cap and bell stages of odontogenesis (Mackenzie et al., 1992). The expression of *Msx1* is first seen in the dental mesenchyme at the lamina stage, and increases in the condensing dental mesenchyme at the bud stage. At the morphogenetic cap stages both the dental papilla and follicle

express *Msx1* maximally. The expression begins to decrease before the differentiation of the odontoblasts and ameloblasts. In the late stages of odontogenesis, *Msx1* expression is absolutely absent from the root sheath epithelium and is rather weak in the dental pulp. By assumption, it appears that *Msx1* does not support root morphogenesis in the developing tooth (Yamashiro et al., 2003).

Msx1 is required in early tooth development for the transfer of *Bmp4* expression from dental epithelium to dental mesenchyme. This is important for progression of molar tooth development beyond the bud stage. There have been studies about the genetic pathways of *Msx1* and found that *Msx1* expression is the upstream of mesenchymal *Bmp4* and *Lef1* expression (Figure 2.26). Their expression is specifically decreased in *Msx1* deficient tooth mesenchyme. In addition, at the first stage, the pathway exhibits *Msx1* downstream of epithelial *Bmp4*, because *Bmp4* is expressed in dental epithelium earlier than expression of *Msx1* in the beneath mesenchyme (Tureckova et al., 1995) and BMP4 can stimulate *Msx1* expression in explanted dental mesenchyme (Chen et al., 1996). Furthermore, in contrast to mesenchymal *Bmp4* expression, epithelial *Bmp4* expression is preserved in *Msx1* mutant tooth germs (Chen et al., 1996). After epithelial *Bmp4* induces *Msx1* in mesenchyme, *Msx1* subsequently induces mesenchymal *Bmp4* expression. Mesenchymal *Bmp4* facilitates the re-induction of *Msx1* expression throughout the dental mesenchyme by a positive feedback loop. Mesenchymal *Bmp4* continues to induce mesenchymal *Lef1* expression for completing odontogenesis. *Msx1* is required for Bmp4 mediated induction of mesenchymal *Bmp4* expression, but not for Bmp4 mediated induction of mesenchymal *Lef1* expression. *Msx1* might function as an amplifier to allow the stronger and more rapid spread of the Bmp4 inductive signal

throughout the dental mesenchyme than would find in the case of simple diffusion alone (Chen et al., 1996).

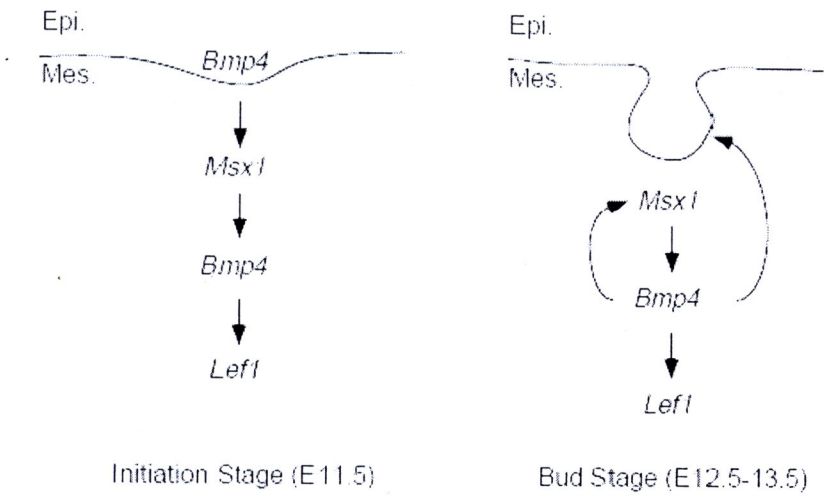


Figure 2.26 *Msx1*, *Bmp4* and *Lef1* into a genetic pathway in early tooth formation. At the initiation stage, epithelial *Bmp4* induce mesenchymal *Msx1* and *Msx1* subsequently induce mesenchymal *Bmp4* expression. After that mesenchymal *Bmp4* continue to induce mesenchymal *Lef1* expression for completing odontogenesis. At the bud stage, Mesenchymal *Bmp4* facilitates the re-induction of *Msx1* expression by a positive feedback loop (Chen et al., 1996).

It has been demonstrated that FGFs are potent inducers of *Msx* expression in dental mesenchyme and that the induction of syndecan-1 by FGFs is mediated at least in part by *Msx* genes (Chen et al., 1996). The role of *Msx1* in the induction of syndecan-1 expression is associated with the condensation of dental mesenchyme. The reduction of syndecan-1 expression may imply the failure of mesenchymal condensation in *Msx1* mutant tooth germs (Chen et al., 1996).

The tooth phenotype of *Msx1* mutant mice can be partially rescued by exogenous BMP4. Although *Bmp4* is more potent than *Msx1*, *Msx1* is still necessary for expression of its downstream targets. These facts suggest that specific combinations

of inductive signals and transcription factors are required for the progressive specification of odontogenesis (Chen et al., 1996).

2.3 Hypodontia (Tooth agenesis)

Agenesis of at least one tooth is the most common developmental anomaly of humans (Vastardis, 2000). The congenitally missing teeth results from interferences during the early stages of tooth development: initiation and proliferation. The prevalence of permanent tooth agenesis varies between 1.6% and 9.6%, depending on the studied population and it reaches 20% if third molars are accounted (Vastardis, 2000).

The prevalence of hypodontia in primary dentition is lower, varying between 0.5% and 0.9% (Vastardis, 2000). Otherwise, mice have only one set of dentition equivalent to primary dentition in humans, so comparing between the mice and humans is rather difficult (Tucker and Sharpe, 2004). Hypodontia in the primary dentition is more common in the upper arch and is often associated with the lateral incisors. Males and females are equally affected. If a primary tooth is missing, the permanent tooth is most likely to be absent. Hypodontia in the permanent dentition can be found with similar frequency in the maxilla and mandible. The most common missing teeth are third molars, lateral maxillary incisors or second mandibular premolars, respectively (Vastardis, 2000). Prevalence of hypodontia in children who have a cleft lip, cleft palate or both is high (Vieira et al., 2003; Slayton et al., 2003).

Hypodontia can be isolated, as the only phenotypic variation in a person, or associated to other variations as part of a syndrome. Isolated or non-syndromic tooth agenesis can be familial or sporadic; and may be inherited in autosomal dominant or

recessive, or X-linked modes (Vastardis, 2000). Expressivity is quite variable, with a wide range of missing teeth, regardless of the fact that a typical phenotype may be determined for each known syndrome. Grahnen has suggested that tooth agenesis is typically transmitted as an autosomal dominant trait with incomplete penetrance and variable expressivity (Grahnen, 1956). Some studies have considered the variations in shape and size, like peg-shaped teeth, as a variation in the expression of the mutated genes (Arte et al., 2001; Lidral and Reising, 2002). The activity of modifying genes or epigenetic factors is regulated by the interaction with other molecules that can be tissue-specific and have allelic variants (Jumlongras et al., 2001). The interaction could result in different phenotype. Many of the genes involve in tooth forming also have important functions in the development of other organs. This is evident as hypodontia as a part of at least 45 syndromes (Sharpe, 1995).

2.3.1 Genes known to cause non-syndromic hypodontia in mice and human

Molecular genetics is used for identify the etiology of hypodontia. Many studies reported candidate genes that involve in tooth development. Hypodontia is possibly caused by independent mutant genes, working alone or in association with other genes, leading to a specific phenotypic pattern (Vieira, 2003).

To date, there have been studies about the association between mutations in *MSX1* and *PAX9* and an isolated form of familial or sporadic hypodontia. Protein products of these genes function as transcription factors that are responsible for the interactions between dental tissues and are necessary for the formation of the odontogenic potential of the mesenchyme (Tucker and Sharpe, 2004). In

homozygous *Msx1* or *Pax9* mutant embryonic mice, dental development is arrested at the bud stage, suggesting that the expression of these genes is critical for the development of the dentition (Das et al., 2002; Vieira, 2003; Tucker and Sharpe, 2004).

MSX1 contains a highly conserved homeobox sequence encoding a 60 amino acid-long DNA-binding homeodomain. Protein product of *MSX1* is a transcription factor expressed in numerous embryonic structures, including the dental mesenchyme (Mostowska et al., 2003). Mice homozygous for *Msx1* deletion reveal secondary cleft palate, deficiency of alveolar process of the maxilla and mandible, failure of incisor development, and an arrest of molar development at the bud stage. Abnormalities of nasal, frontal and parietal bones as well as of the malleus in the middle ear are also found in the mice. *Msx1* has an important role in epithelial-mesenchymal interactions during craniofacial and tooth development (Satokata and Maas, 1994).

PAX9 is a member of a gene family encoding transcription factors, 128-amino acid long DNA-binding paired domain, function during embryogenesis. In mouse embryos, the expression of *Pax9* is in mesenchyme prior to any morphological signs of odontogenesis, and maintain thereafter in the developing tooth mesenchyme. Mice homozygous for *Pax9* deletion die shortly after birth because of respiration problems. Mice also lack all teeth and show a wide range of developmental disorders, including secondary cleft palate as well as defects in craniofacial bones and cartilages. In *Pax9*-deficient mouse embryos, tooth forming is arrested at the bud stage (Peters et al., 1998; Mostowska et al., 2003).

2.4 Facial and palate development

2.4.1 Embryology of orofacial development

Development of the face and jaws is the product of growth and fusion of processes (mandibular processes, maxillary processes and frontonasal process) and associated cell migration, proliferation, differentiation and apoptosis. Palatal tissue consists of a mesenchymal core derived mainly from the cranial neural crest and of an ectodermally derived epithelial outer covering (Carlson, 2004).

During the 4th gestational week, a frontonasal prominence is seen in the midline to form the upper part of the stomodeum. Following the formation of the nasal placodes, locate besides of the frontonasal prominence, the medial and lateral nasal processes develop within the frontonasal prominence (Figure 2.27).

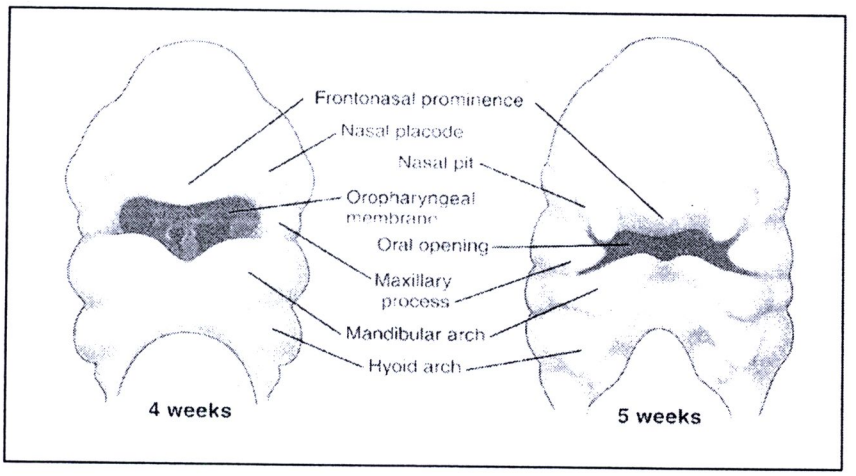


Figure 2.27 Frontal view of facial development during the 4th week. Frontonasal prominence forms the upper part of stomodeum. Nasal placode is seen before medial and lateral nasal process development (adapted from Carlson, 2004).

During the 5th and 6th weeks, the medial and lateral nasal processes are developed from two nasal placodes that elevate into a horseshoe shape (Carlson, 2004). The medial nasal processes enlarge more rapidly than the lateral nasal

processes, and fuse with each other in the midline. The fused medial nasal process is developed to be the central part of upper lip, anterior (primary) palate in front of the incisive foramen, maxillary incisors and alveolar process. The lateral nasal processes become the alae of the nose. The nasal pits are formed by invaginations of the center of the medial and lateral nasal processes and consequently progress to be the nostrils (Figure 2.28). Palatal shelves are developed from the inner parts of the maxillary process (Figure 2.29). Two mandibular processes enlarge and move to the midline, and fuse with each other to form the mandibular arch and lower lip (Figure 2.30).

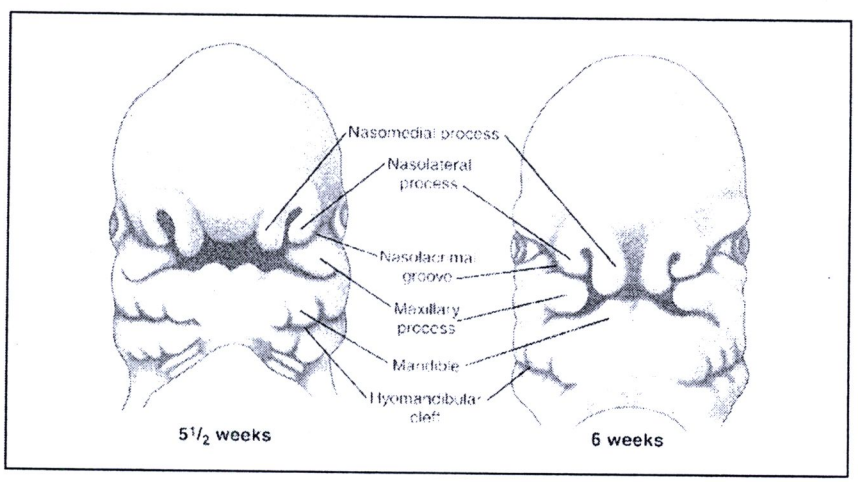


Figure 2.28 Frontal view of facial development during the 5th and 6th week. The medial and lateral nasal processes are developed (adapted from Carlson, 2004).

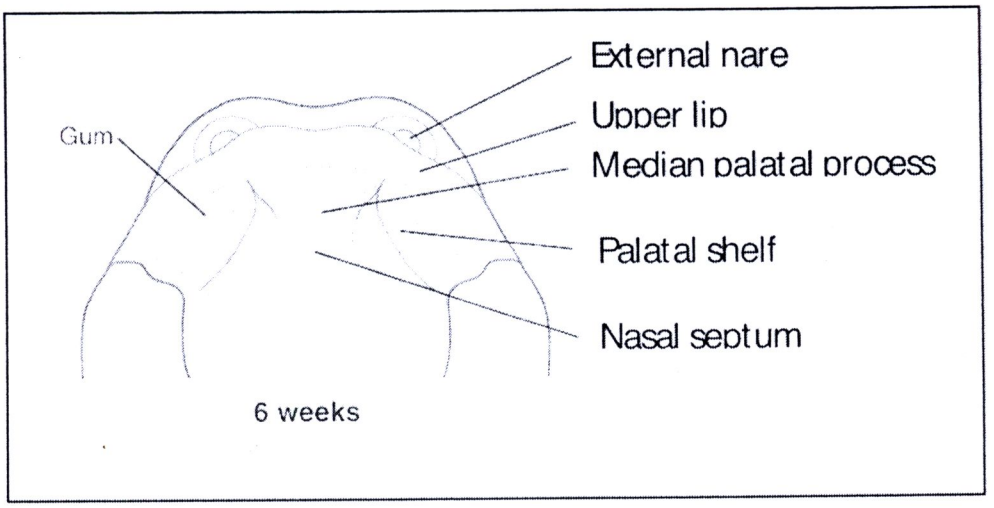


Figure 2.29 Hard palate formation during the 6th gestational week. Two lateral palatal shelves of the maxillary processes begin to grow (adapted from Carlson, 2004).

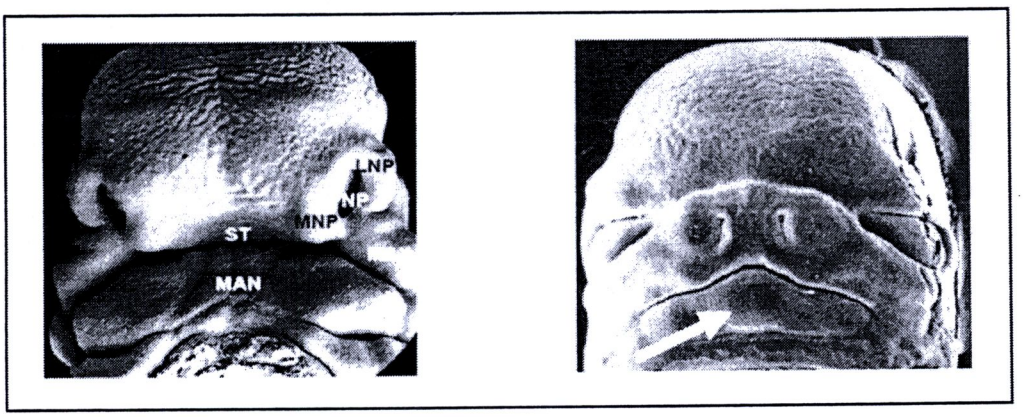


Figure 2.30 Mandibular formation. A, Mandibular processes are separated from each other in the 4th week. B, Mandibular processes are fused with each other (arrow) in the 6th week (adapted from Carlson, 2004).

Tongue begins to develop at the initiation of the 7th gestational week. Initially, palatal shelves grow vertically and downward along the sides of the tongue (Carlson,

2004). After that in the 8th week, the tongue moves downward, following the growth of the mandible. Subsequently the palatal shelves are elevated into the horizontal position, grow and fuse to each other during the 9th week (Figure 2.31-2.33)

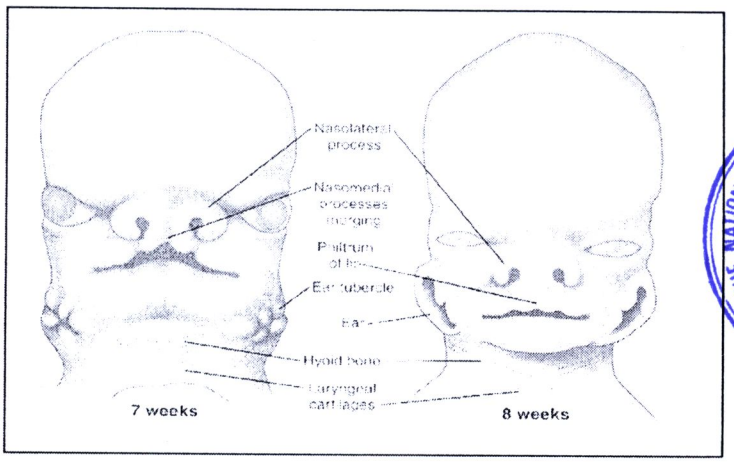


Figure 2.31 Craniofacial development during the 7th and 8th week. The medial nasal, lateral nasal, maxillary and mandibular processes fuse together (adapted from Carlson, 2004).

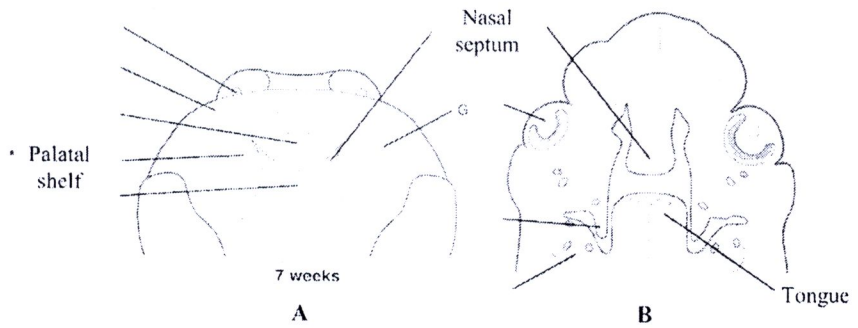


Figure 2.32 Hard palate formation during the 7th gestational week. Palatal shelves of the maxillary processes grow vertically and downward along the sides of the tongue. A, Lower view. B, Frontal view (adapted from Carlson, 2004).

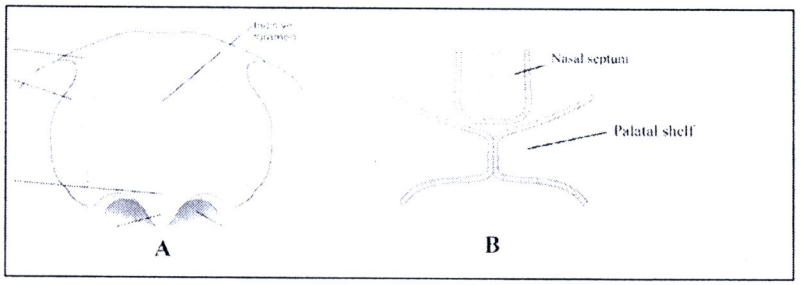


Figure 2.33 Hard palate formation during the 8th week. Palatal shelves are elevated into horizontal position. A, Lower view. B, Frontal view (adapted from Carlson, 2004).

Palatal shelves adhere horizontally with the other side and with the primary palate at the end of 8th week (Carlson, 2004). Fusion of the palatal shelves begins from anterior to posterior and complete at the 11th gestational week. The contact part form an epithelial seam, is later replaced by mesenchyme, to form the definitive palate. The nasal septum grows downward and fuses to the palate superiorly in the midline at the same time with the fusion of the secondary palate. So two nasal cavities are completely separated (Figure 2.34).

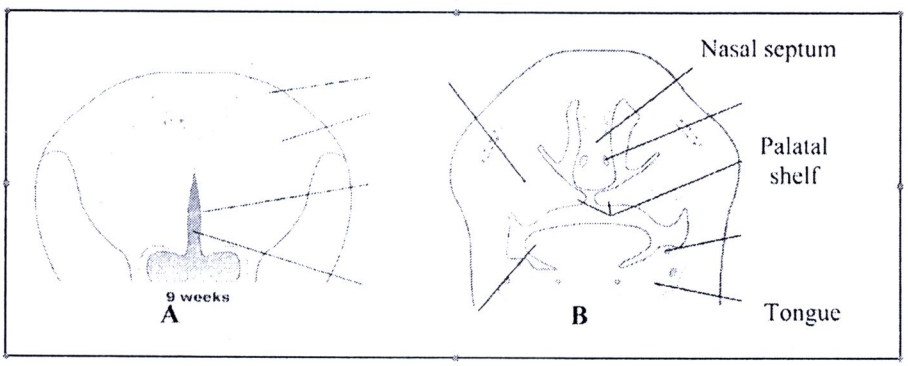


Figure 2.34 Hard palate formation at the end of the 8th gestational week. A, Lower view of secondary palate. B, Frontal view of junction of two palatal shelves and nasal septum (adapted from Carlson, 2004).

2.4.2 *MSX1* and palate development

Msx1 expression has been demonstrated in the developing palate. Identification of a diffuse expression of *Msx1* in the palatal mesenchyme indicated the first evidence that *Msx1* may have a direct function in palate development (Mackenzie et al., 1991; Tureckova et al., 1995). There has been the report that *Msx1* expression in the palatal mesenchyme is located at the anterior portion of the developing palatal shelves (Zhang et al., 2002).

Regarding primary palate development during the 5^{th} and 6^{th} weeks, the medial nasal processes enlarge more rapidly than do the lateral nasal processes, and fuse with each other in the midline. The fused medial nasal process develops to become the central part of the upper lip and the primary palate in front of the incisive foramen (Carlson, 2004). Development of the normal lip requires the posterior parts of the medial and lateral nasal processes to fuse with each other and with the medial portion of the maxillary process (Trasler, 1968). These stages require a normal mesenchymal proliferation. Andersen and Matthiessen (Andersen and Matthiessen, 1967) proposed that cleft lip will occur if mesenchymal proliferation is retarded in the medial nasal, lateral nasal and maxillary processes. *Msx1* expression is important for a normal mesenchymal proliferation (Hu et al., 2001). It has been reported that *Msx1* is expressed early in the mesenchyme of medial nasal, lateral nasal, maxillary and mandibular processes (Mackenzie et al., 1991).

Associated signaling pathways consist of *Bmp4*, *Msx1*, *Shh* and *Bmp2*, are important for normal growth and fusion of palatal shelves in the mouse (Zhang et al., 2002). *Msx1* expression is localized in the anterior palatal mesenchyme. A feedback loop containing *Msx1* and *Bmp4* stimulates expression of *Shh* in the medial edge

epithelium of the palatal region. Epithelial *Shh* induces *Bmp2* in the underlying anterior mesenchyme, and both *Shh* and *Bmp2* control tissue proliferation of the anterior palate (Zhang et al., 2002). *Bmp4* and *Bmp2* are expressed in the epithelium and mesenchyme of the developing palate at embryonic day 12.5 (E12.5). At E13.5, expression of *Bmp2* remains the same. *Bmp4* is identified only in the anterior mesenchyme beneath the medial edge epithelium. *Msx1* is obviously expressed in mesenchyme at E12.5 and E13.5. In homozygous *Msx1* deficient mice, disturbances of mesenchyme proliferation may result in cleft palate. *Shh* is expressed in the anterior and posterior palate, localized in the medial edge epithelium of the anterior palate at E12.5 and E13.5, and in the oral epithelium of the posterior palate. *Shh* has been shown to play a role in palatal shelf mesenchymal proliferation in vitro (Rice et al., 2004).

2.5 Orofacial clefts

Orofacial clefts are the common congenital craniofacial abnormalities. These can be found in syndromic and non-syndromic forms of cleft lip with or without cleft palate (CL/P) and cleft palate only (CP). Genetic and environmental factors are involved in etiology. Known teratogens consist of diphenylhydantoin (dilantin), thalidomide, dioxin, ethyl alcohol, trimethadione, retinoids, aminopterin and methotrexate, and hyperthermia (Gorlin et al., 2001). The associated factors include geographical origin, racial and ethnic backgrounds, and socio-economic status. Syndromic forms of cleft palate found approximately 50–55% of cases (Koillinen et al., 2005). Frequency of non-syndromic forms of cleft palate is 1: 1000 live human births, whereas cleft lip with or without cleft palate occurs 1: 700 live human births in

Caucasian (Koillinen et al., 2005). It has a prevalence of approximately 1 in 600 among Thai newborns (Tongkobpetch et al., 2006).

Orofacial clefts are the results of failure of the facial processes fusion. The most common form is the failure of the maxillary process and the frontonasal process resulting in cleft lip and a failure of the two palatal shelves of the maxillary processes causing a cleft palate. Every stage of the palatal morphogenesis has important role, disruption of any stage could result in clefts. Deficient or delayed growth of the palatal shelves, impaired shelf elevation, failure of the bilateral shelves to contact, adhere or fuse medially, post-fusion rupture and failure of the mesenchyme to differentiate completely are the common etiology of a cleft palate.

2.5.1 Genes known to cause non-syndromic orofacial cleft in mice and human

Epithelial-mesenchymal interactions are associated with palate development. At initial stage, Mesenchymal *Bmp* and *Fgf* signals induce the epithelial cells to proliferate. Epithelial Shh signals then induce *Ptc* in the underlying mesenchyme. Candidate genes for non-syndromic cleft lip with or without cleft palate in humans including *MSX1, IRF6*, *TGFA*, *TGFβ1*, *TGFβ3*, *RARA* and *BCL3*, have the most supporting evidence. Several candidate genes have been studied for a role in non-syndromic cleft lip with or without cleft palate (Table 2.2) (Lidral and Moreno, 2005). In addition, candidate genes such as *MSX1, PTCH1, IRF6, SUMO1, TGFβ3, PVRL1* and *TBX22* are known to be associated with cleft palate in humans (Satokata and Maas, 1994; Vieira, 2003).

Table 2.2 Candidate genes for non-syndromic cleft lip with or without cleft palate (Modified from Lidral and Moreno, 2005).

Candidate region	Candidate gene
1p36	*MTHFR, SKI, PAX7*
1q32	*IRF6*
2p13	*TGFA*
4p16	*MSX1*
6p23-25	*TFAP2A, OFCC1*
14q24	*TGFB3, BMP4, PAX9*
17q12-21	*RARA, Clf1*
19q13	*BCL3, CLPTM1, PVRL2, TGFB1*

Several genetic networks coordinate during palate development, signals contact between the palatal epithelium and the underlying mesenchyme. These networks consist of multiple signaling molecules and growth factors, such as members of the transforming growth factor β (*Tgfβ*) superfamily [including bone morphogenetic proteins (*Bmp*s) and *Tgfβ*s], Sonic hedgehog (*Shh*), fibroblast growth factors (*Fgf*s) and their receptors, effectors and targets (Gritli-Linde, 2007). In organ patterning, growth and differentiation are required several transcription factors. *Msx1* (Zhang et al., 2002), short stature homeobox *Shox2* (Yu et al., 2005) and odd-skipped relate2 (*Osr2*) genes (Lan et al., 2004) expression have been exhibited in developing palatal shelves.

2.6 *Msx1*-associated human syndromic malformations

2.6.1 Witkop syndrome

Witkop syndrome, also known as "tooth and nail syndrome" or "nail dysgenesis and hypodontia" was first described by Witkop in 1965 as a rare autosomal dominant abnormality (Witkop, 1965). The incidence has been estimated to be 1–2:10,000. It is one of the ectodermal dysplasia syndromes (EDs), a heterogeneous group of disorders characterized by abnormalities in at least two ectodermally derived organs, such as sweat glands, nails, hair, and teeth. Witkop syndrome can be distinguished from other types of EDs by the fact that the disorders associated only teeth and nails.

2.6.1.1 Clinical manifestations

The congenitally missing teeth and nail dysplasia are the main characteristic features in individuals with Witkop syndrome (Hodges and Harley, 1999). The phenotype includes a variable number and variable types of congenitally missing permanent and/or primary teeth, which mostly combines with lip eversion due to loss of occlusion in the vertical dimension. Nails are generally slow-growing, thin, brittle, and spoon-shaped (koilonychia). In some cases, obvious longitudinal ridges and pitting are the only main features. Fingernails are usually less severely affected than Toenails. Nails spontaneously separate from the nail beds or are absent at birth in rare cases. The nail defects are relieved with age and may not be seen during adulthood. The expressivity of tooth and nail abnormalities is highly variable (Witkop, 1965). Almost all Witkop syndrome patients revealed normal hair, sweat glands, and ability to tolerate heat, despite a few reported cases have sparse or fine hair in addition to tooth and nail disorders (Chitty et al., 1996). Variability of the pattern of congenitally

missing teeth and varying degrees of nail abnormalities were revealed in previous reports of families affected with Witkop syndrome and sporadic cases.

Jumlongras and colleagues examined 20 family members, nine were affected. The pedigree showed at least two instances of male-to-male transmission. Affected individuals had 11 to 28 congenitally missing permanent teeth and dysplastic toenails and/or fingernails (Figures 2.35, 2.36). The severity of the phenotype in the family was quite variable. The predominant tooth types affected were premolars, first molars, and third molars. The present permanent teeth appeared smaller in mesiodistal width and had shorter root lengths than normal teeth. Maxilla and mandible appeared to be smaller than normal (Jumlongras et al., 2001).

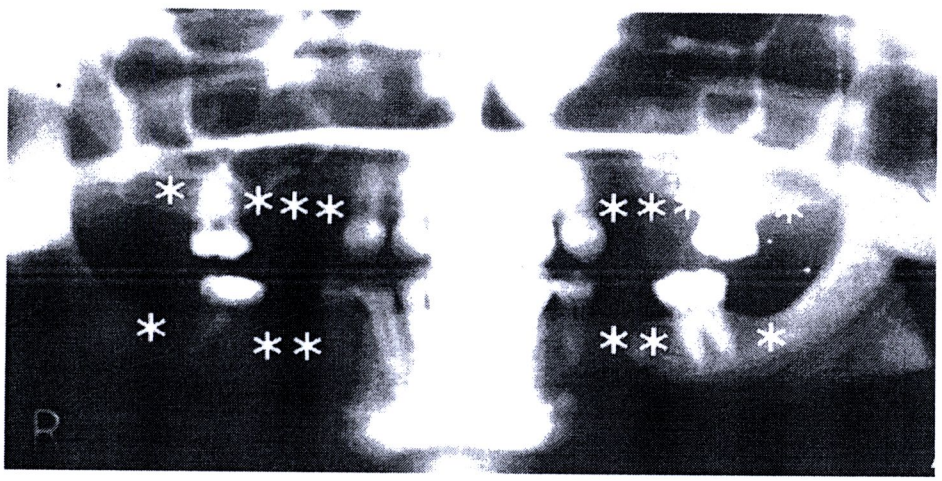


Figure 2.35 Witkop syndrome phenotype, Panoramic radiograph of 33 year-old patient. Asterisks (*) mark congenitally missing permanent teeth. This patient has a total of 14 congenitally missing teeth (Jumlongras et al., 2001).

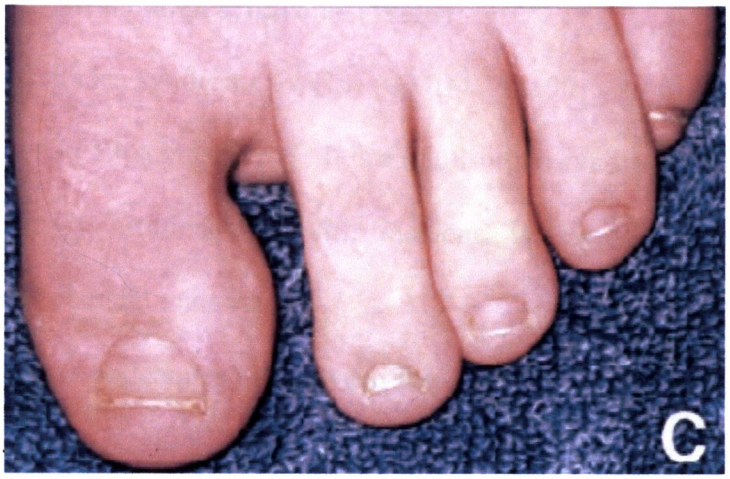


Figure 2.36 Toenail dysplasia in Witkop syndrome patient, All nail plates are concave and hypoplastic, the fifth toenail is almost absent (Jumlongras et al., 2001).

2.6.1.2 Etiology

MSX1 is the responsible gene for Witkop syndrome, since mutations in *MSX1* were revealed to be associated with tooth agenesis in two unrelated families (Vastardis et al., 1996; van den Boogaard et al., 2000). The expression of *Msx1* was also shown in developing mouse tooth buds and nail beds (Mackenzie et al., 1991). In 2001, Jumlongras and collegues did a linkage analysis in a three-generation family affected with Witkop syndrome. They mapped the gene near the *MSX1* locus and mutation analysis of *MSX1* revealed a nonsense mutation (S202X). Additional analysis of *Msx1*-knockout mice revealed that not only was tooth development disrupted in these mice, but nail development was affected as well. The resemblance between the human family affected with Witkop syndrome and the *Msx1*-knockout mice, suggested that a nonsense mutation in *MSX1* is responsible for Witkop syndrome (Jumlongras et al., 2001).

2.6.2 Wolf-Hirschhorn syndrome (WHS)

Wolf-Hirschhorn syndrome (WHS) (OMIM 194190) was first described in 1965 by Wolf *et al.*, and Hirschhorn *et al.*, It is a multi-organ syndrome caused by deletions of the short arm of chromosome 4 (4p). Incidence is estimated to be about 1:50,000 births with a 2:1 female: male ratio.

Nieminen and colleaques investigated the dental manifestations and the presence of the *MSX1* gene in eight Finnish patients with defects of 4p and seven of the WHS patients. Five of the WHS patients exhibited agenesis of several teeth, suggesting that hypodontia may be a common feature of WHS (Nieminen et al., 2003). By FISH analysis, the five patients with hypodontia lacked 1 copy of *MSX1*, while the other 3 had both copies. Only one in the latter group had cleft palate. These suggested that haploinsufficiency for *MSX1* cause selective tooth agenesis but it is not sufficient to cause oral clefts (Nieminen et al., 2003).

2.6.2.1 Clinical manifestations

The characteristic signs (Figure 2.37) include a typical craniofacial abnormalities with prominent glabella, broad and/or beaked nose, hypertelorism, short philtrum, downturned corners of the mouth, cleft lip/palate, micrognathia, maxillary hypoplasia, dysplastic ears and preauricular tags and pits, oral clefts, dental anomalies, microcephaly (90%), mental retardation (75%), seizures (50-85%), short stature (25-66%), low birth weight (77%), muscular hypotonia (90%), congenital heart defects (31-45%), colobomata of iris (30%), genital anomalies (30%), deafness (23%), and renal anomalies (23%).

Figure 2.37 Wolf-Hirschhorn syndrome consists of typical craniofacial abnormalities with prominent glabella, broad and/or beaked nose, hypertelorism, short philtrum, downturned corners of the mouth, cleft lip/palate, micrognathia and maxillary hypoplasia (South et al., 2008).

Mental retardation can be found from mild to severe. Acquirement of walking without support is between 30 and 60 months, many patients are not able to walk. For language skill, first words are spoken between 18 and 48 months. Most of the patients do not develop active speech, only a few can make complex sentences.

The dental abnormalities reported include delayed development and fusion of incisors. Hypodontia involved complete agenesis of second premolars and third molars, whereas there was variability in the number and identity of other missing

teeth (Nieminen et al., 2003). The craniofacial manifestation seems like that of a Greek warrior helmet. The characteristic facial phenotype usually leads the clinical geneticists to use cytogenetic investigations (conventional chromosomal analysis and fluorescence in situ hybridisation/ FISH) to make diagnosis of this syndrome.

2.6.2.2 Etiology

Wolf-Hirschhorn syndrome (WHS) has been reported to be caused by deletions in 4p16.3 (also involve location of *MSX1* gene).

2.7 *Msx1*-associated human non-syndromic malformations

2.7.1 *MSX1* mutations and pattern of non-syndromic hypodontia

There is a similar pattern of tooth agenesis in the several studies about *MSX1* mutations (Vastardis et al., 1996; van den Boogaard et al., 2000; Jumlongras et al., 2001). The lower second premolars are the most commonly affected, followed by upper second premolars, upper first premolars, and upper lateral incisors respectively. Third molar agenesis, in the case of complete absence is suggested that it might be a causal relationship with *MSX1* mutation (Vastardis et al., 1996). There appears to be a pattern of anterior progression of agenesis for each tooth type because the most distal tooth of each tooth type is most often affected, and, as the severity worsens. Another feature of *MSX1* mutations is the large number of teeth missing in the affected individuals, with an average of 16.4/person (Jumlongras et al., 2001), 12.2/person (Lidral and Reising, 2002), 11.0/person (Vastardis et al., 1996) and 8.4/person (van den Boogaard et al., 2000). *MSX1* hypodontia is different from *PAX9* hypodontia, in that first and second molar agenesis were noted in the families with

PAX9 mutations. Some individuals with *PAX9* mutations also lacked maxillary and/or mandibular second premolars as well as mandibular central incisors (Stockton et al., 2000).

Affected individuals have been reported to have normal primary dentition. It is hypothesized that there is a redundancy of homeoproteins signals and/or other genetic mechanisms which are involved in primary tooth formation and patterning (Vastardis et al., 1996).

2.7.2 *MSX1* mutations and pattern of non-syndromic orofacial clefts

Several association studies of the gene with CL/P and cleft palate only (CPO) supported the role of *MSX1* in non-syndromic cleft in different populations. The possibility of an effect of *MSX1* on both types of clefts is supported by study of a Dutch family with an *MSX1* mutation that appears to cause a variety of cleft-related defects (van den Boogaard et al., 2000). Cleft palate has been separated etiologically and embryologically from clefts involving the lip or the lip and the anterior hard palate. A Dutch family in the study of van den Boogaard and coworkers, as well as reported and observations of clinical cases suggested that isolated cleft palate can occasionally occur in the context of cleft lip and palate. Similarly, the linkage disequilibrium data have supported a role for both cleft palate and cleft lip and palate together. One study suggested that the important distinction was not whether the lip was involved but whether the palate was included (Beaty et al., 2001). It may be important to investigate both lip and palate phenotypes as part of a comprehensive search for etiological mutations.

2.8 Previously reported *MSX1* mutations

MSX1 is highly expressed in the mesenchyme of developing tooth bud and palatal tissue. It has been reported that *MSX1* mutations cause non-syndromic congenitally missing teeth and non-syndromic cleft lip with or without cleft palate in several studies. Mutations were found in multiple locations as report in publications (Table 2.3 and Figure 2.38).

Table 2.3 Previous mutations of human *MSX1*

Exon	Mutation	Phenotype	Reference	Mode of inheritance	Note
Exon 1	1-BP DUP, 62G	Tooth agenesis	Kim (2006)	AD	
Exon 1	MET61LYS	Tooth agenesis	Lidral and Reising (2002)	AD	
Exon 1	GLY91ASP	Orofacial cleft	Jezewski (2003)	AD	
Exon 1	SER105TER	Tooth agenesis , Orofacial cleft	van den Boogaard (2000)	AD	
Exon 1	VAL114GLY	Orofacial cleft	Jezewski (2003)	AD	
Exon 1	PRO147GLN	Orofacial cleft	Suzuki (2004)	AD	Found in 8 of 100 Thai controls, Tongkobpetch (2006)
Exon 2	ARG151SER	Orofacial cleft	Jezewski (2003)	AD	
Exon 2	GLN187TER	Tooth agenesis	De Muynck (2004)	AD	
Exon 2	ALA194VAL	Tooth agenesis	Mostowska (2006)	AD	
Exon 2	ARG196PRO	Tooth agenesis	Vastardis (1996)	AD	
Exon 2	SER202TER	Witkop syndrome	Jumlongras (2001)	AD	
Exon 2	ALA219THR	Tooth agenesis	Chishti (2006)	AR	
Exon 2	ALA221GLU	Tooth agenesis	Xuan (2008)	AD	
Exon 2	GLY267CYS	Orofacial cleft	Tongkobpetch (2006)	AD	
Exon 2	PRO278SER	Orofacial cleft	Tongkobpetch (2006)	AD	

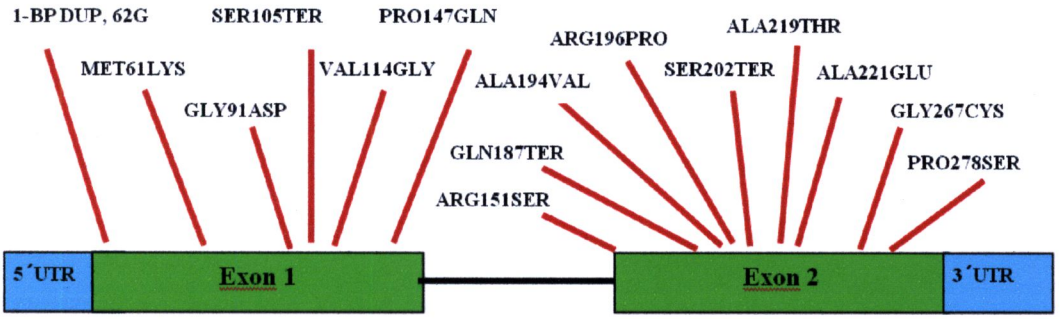


Figure 2.38 Schematic view of human *MSX1* gene. Red lines indicate amino acid changes and location of known mutations.

In addition to allelic variants reported in OMIM (Online Mendelian Inheritance in Man), there have been few reports about *MSX1* mutation in non-coding region, with unclear description, in recessive form and with limb anomalies.

Mostowska and collegues have demonstrated a novel mutation of *MSX1*, which might be associated with the lack of 14 permanent teeth in the proband. Despite this c.581C>T transition was identified also in 2 normal individuals from the proband's family, it was located in a highly conserved homeobox sequence of *MSX1*. They proposed that this transition might be the first described mutation of *MSX1* that might be associated for hypodontia and exhibiting incomplete penetrance. It may also support the opinion that hypodontia might be an oligogenic trait caused by mutations of different genes at the same time (Mostowska et al., 2006).

There has been the study in two consanguineous Pakistani families with an autosomal recessive form of hypodontia. Sequence analysis of *MSX1* showed a novel recessive missense mutation resulting in substitution of alanine to threonine amino acid (p.A219T), localized in the *MSX1* homeodomain, which is important for DNA

binding and protein-protein interaction. This mutation is the first recessive mutation identified in *MSX1* (Chishti et al., 2006).

Furthermore, there was the analysis in three unrelated patients with sporadic, non-syndromic hypodontia. They exhibited a homozygotic deletion of 11 nucleotides in the intron, near 5´splicing site and proposed that the deletion may decrease the expression level of MSX1 protein (Pawlowska et al., 2009).

Besides hypodontia and orofacial cleft, there has been the study of an association between isolated limb anomalies and rare alleles at the *MSX1* locus. They investigated 34 infants with limb anomalies, the frequencies of rare *MSX1* alleles were significantly higher than in 482 infants with other isolated birth defects. Infants containing the rare allele had a 4.81-fold higher risk of limb anomalies when the mother reported smoking during pregnancy, compared to infants who were homozygous for the common allele and whose mother did not smoke (van den Boogaard et al., 2008).