

# **HUMAN TBX22 EXPRESSION AND PROTEIN-DNA INTERACTIONS**

**Steven Lisgo**

A thesis submitted for the degree of Doctor of Philosophy.

Institute of Human Genetics  
Newcastle University

October 2010

## **DECLARATION**

This is to certify that this thesis titled "**Human TBX22 expression and protein-DNA interactions**" presented for the Degree of Doctor of Philosophy at the University of Newcastle is my own work, except where stated, and has not been presented at any other University.

Steven Lisgo

October 2010

## ACKNOWLEDGEMENTS

I have much to thank Prof. Lindsay for: not only did Susan give me the encouragement and belief to undertake this project; she also gave me my first opportunity in the field of academic science, and for some reason continues to stick with me! Without Susan's guidance and support this Thesis would never have made it this far.

The work, help and camaraderie of the HDBR team and research midwives, is greatly appreciated and whose efforts continue to enable the smooth running of the HDBR collection.

Thanks to all members of the Institute of Human Genetics, past and present, whose friendship I have been lucky enough to enjoy whose help has been invaluable over the last few years, but are too numerous to list by name here.

Thanks go to my parents for all their encouragement and support - I finally did it! The assistance with the editorial input into this thesis from my father is also very much appreciated.

Finally, my sincerest gratitude belongs to my wife Jane, for her companionship, support, encouragement and most of all patience and understanding during the writing of this thesis. The completion of which was assisted by her reminders (and occasional gentle nagging!) to remain focused from other distractions.

MANY THANKS TO YOU ALL.

## ABSTRACT

Cleft palate is one of the most common birth abnormalities. Figures published in 2006 by the American Centres for Disease Control and Prevention, report the incidence of those born in the United States with a cleft palate without the presence of a cleft lip (CPI) to be 6.39 for every 10000 in the three years between 1999 to 2001 and for cleft lip in association with a cleft palate (CLP) to be even greater - 10.48 per 10000 live births. In 2001, Braybrook and colleagues reported that mutations in the *TBX22* gene cause X-linked cleft palate (CPX), a disease characterised by a cleft of the secondary palate and is often seen in association with ankyloglossia (tongue-tie) (Braybrook *et al.* 2001).

A cleft of the secondary palate arises as a consequence of disturbance to correct development during palatogenesis: an anomaly in palatal shelf growth; delayed or failed shelf elevation; defective shelf fusion or a failure of medial edge epithelium cell death. This thesis reveals that the expression of *TBX22* during these key developmental events in human embryos is consistent with the phenotype seen in CPX.

To enable an investigation for *TBX22* target genes, a DNA binding sequence is determined for the *TBX22* protein. This sequence is used to generate a generic *TBX22* DNA binding site, the presence of which is screened for in promoter regions, defined as 2kb upstream of transcription start sites. 132 genes were selected as candidate *TBX22* targets on the basis that they underlie human disorders that include a cleft palate. The screen shows that 28 of these genes have at least one perfect or near perfect match to the generic *TBX22* DNA binding site. Of these, only two both contained a perfect *TBX22* generic DNA binding site and mouse mutants also had cleft palates: *SUMO1* and *MSX1*. Interaction between *SUMO1* and *TBX22* has already been shown (Andreou *et al.* 2007). This study investigated *MSX1* as a downstream target of *TBX22* using a luciferase reporter gene construct *in vitro*. The results showed that in the presence of *TBX22*, the luciferase signal was reduced and support *MSX1* being a downstream target gene of *TBX22*.

These findings further the understanding of the molecular networks regulating craniofacial development. Unravelling these complex interactions is crucial to identifying the mechanisms of oro-facial clefting, important steps towards improved methods of counselling, treatment and prevention of these common birth disorders.

## TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION.....	1
1.1 BRACHYURY, TBX22 AND THE T-BOX GENE FAMILY.....	2
1.1.1 <i>TBX22</i> Chromosomal location and Gene Structure.....	7
1.2 EARLY MORPHOGENESIS OF THE HUMAN FACE.....	10
1.2.1 The pharyngeal arches .....	13
1.2.2 Formation of the facial processes.....	18
1.2.3 Formation of the secondary palate.....	22
1.2.4 Extension of the palate.....	27
1.2.5 Removal of the midline epithelia cells. ....	30
1.2.6 Palate formation in other vertebrates.....	34
1.3 CLEFT PALATE.....	35
1.3.1 Syndromic and non-syndromic CL/P and CPI.....	35
1.3.2 The aetiology of cleft palate.....	36
1.3.3 Cleft palate risk factors.....	37
1.3.4 Detection, treatment and outcomes.....	44
1.4 AIMS AND STRUCTURE OF THE THESIS.....	46
CHAPTER 2. EXPRESSION OF TBX22 DURING EARLY HUMAN DEVELOPMENT.....	47
2.1 INTRODUCTION.....	47
2.2 MATERIALS AND METHODS.....	51
2.2.1 Embryo collection and processing.....	51
2.2.2 Tissue <i>in situ</i> hybridisation.....	53

2.3 RESULTS.....	58
2.4 DISCUSSION.....	71
2.4.1 <i>TBX22</i> expression correlates with the CPX phenotype.....	72
2.4.2 A mechanism for a Cleft Palate.....	72
2.4.3 Expression outside of the palate.....	75
2.4.4 Comparison with mouse and chick <i>Tbx22</i> expression.....	77
 CHAPTER 3. THE TBX22 DNA BINDING SEQUENCE.....	81
3.1 INTRODUCTION.....	81
3.1.1 T-box domains.....	81
3.1.2 Generating the TBX22 protein.....	84
 3.2 MATERIALS AND METHODS.....	85
3.2.1 Producing TBX22 protein with an N-terminal 6xHis Tagged in <i>E. coli</i> .....	85
3.2.2 Rabbit Reticulate Transcription Translation.....	92
3.2.3 Protein verification.....	94
3.2.4 <i>In vitro</i> oligonucleotide selection assay.....	97
3.2.5 Electrophoretic Mobility Shift Assay (EMSA).....	100
 3.3 RESULTS.....	103
3.3.1 Producing a TBX22 protein.....	103
3.3.2 Determining a TBX22 preferential DNA binding sequence.....	108
 3.4 DISCUSSION.....	116
 CHAPTER 4. TBX22 DOWNSTREAM TARGETS.....	122
4.1 INTRODUCTION.....	122

4.2 MATERIALS AND METHODS.....	124
4.2.1 Electrophoretic Mobility Shift Assay of <i>MSX1</i> oligonucleotide.....	124
4.2.2 Non-radioactive <i>in situ</i> hybridisation.....	124
4.3 RESULTS.....	128
4.3.1 Computational search for Potential TBX22 target genes.....	128
4.4 DISCUSSION.....	141
 CHAPTER 5. MSX1: A POTENTIAL TBX22 TARGET GENE.....	148
5.1 INTRODUCTION.....	148
5.2 MATERIALS AND METHODS.....	151
5.2.1 Preparation of the expression and reporter constructs.....	151
5.2.2 Co-transfection of the expression and reporter constructs.....	154
5.2.3 Detection of <i>TBX22</i> and <i>MSX1</i> expression in HeLa and 293Tcells.....	159
5.2.4 Detection of luciferase by Western blot.....	160
5.3 RESULTS.....	161
5.4 DISCUSSION.....	172
 CHAPTER 6. CONCLUSION AND FUTURE DIRECTIONS.....	177
6.1 CONCLUSION.....	117
6.2 FUTURE DIRECTIONS.....	186
 BIBLIOGRAPHY.....	188

APPENDICES.....211

APPENDIX 1: OLIGONUCLEOTIDES AND ANTIBODIES

APPENDIX 2: MALDI-TOF ANALYSIS OF THE TBX22 RECOMBINANT PROTEIN

APPENDIX 3: THE TBX22 DNA CODING AND PROTEIN SEQUENCE

APPENDIX 4: ALIGNMENT OF THE CLONES FROM THE *IN VITRO* DNA BINDING ASSAY AND AN EXAMPLE CHROMATOGRAM OF CLONED DNA FROM *IN VITRO* OLIGONUCLEOTIDE BINDING STUDY

APPENDIX 5: THE GENES SHOWN TO CAUSE A CLEFT PLATE IN HUMAN (TAKEN FROM OMIM 10 FEBRUARY 2010).

APPENDIX 6: LUCIFERASE ASSAY REPORT FROM FLUOROSKAN ASCENT FL LUMINOMETER

APPENDIX 7: LUCIFERASE ASSAY REPORT FROM FLUOROSKAN ASCENT FL LUMINOMETER OF PREVIOUS DATA.

## TABLE OF FIGURES

<b>FIGURE 1: PHYLOGENETIC TREE OF HUMAN T-BOX GENES.....</b>	4
<b>FIGURE 2: THE GENOMIC STRUCTURE OF TBX22.....</b>	9
<b>FIGURE 3: THE DEVELOPMENT OF THE HUMAN FACE.....</b>	13
<b>FIGURE 4: THE MAMMALIAN PHARYNGEAL ARCHES.....</b>	14
<b>FIGURE 5: THE NEURAL CREST FROM EACH RHOMBOMERE MIGRATES TO A SPECIFIC POSITION ALONG THE NEUROAXIS.....</b>	16
<b>FIGURE 6: THE PHARYNGEAL ARCHES AND FRONTONASAL PROCESS.....</b>	19
<b>FIGURE 7: THE MANDIBULAR PROCESSES ARE THE FIRST OF THE FACIAL PROMINENCES TO FUSE.....</b>	20
<b>FIGURE 8: THE FORMATION OF THE NASAL PROCESSES.....</b>	21
<b>FIGURE 9: THE FORMATION OF THE SECONDARY PALATE AND PRE-MAXILLA.....</b>	24
<b>FIGURE 10: MODELS OF THE MOLECULAR AND MORPHOGENETIC ACTIVITY ASSOCIATED WITH THE RUGAE GROWTH ZONE (RGZ) (TAKEN FROM (WELSH AND O'BRIEN 2009)).....</b>	28
<b>FIGURE 11: TBX22 IS NOT DETECTED AT CS14 OR CS15.....</b>	58
<b>FIGURE 12: TBX22 IS EXPRESSED IN THE NASAL AND MAXILLARY PROCESSES OF A CS 16 EMBRYO.....</b>	60
<b>FIGURE 13: EXPRESSION OF TBX22 IN THE CS17 EMBRYO.....</b>	62
<b>FIGURE 14: TBX22 IS EXPRESSED DURING THE FORMATION OF THE PRIMARY AND SECONDARY PALATE.....</b>	64
<b>FIGURE 15: TBX22 EXPRESSION IN THE EXTENDING PALATAL SHELVES AND IN THE FRENULUM OF THE TONGUE.....</b>	66
<b>FIGURE 16: TBX22 IS DOWN REGULATED AS THE PALATAL SHELVES FUSE.....</b>	69
<b>FIGURE 17: A RECOMBINANT PROTEIN EXPRESSION TIME COURSE STUDY.....</b>	103

<b>FIGURE 18: PURIFICATION OF THE 6xHis-TBX22 PROTEIN BY Ni-NTA AFFINITY COLUMN.....</b>	104
<b>FIGURE 19: A PROBABILITY BASED MOWES SCORE OF THE PROTEIN SAMPLES COMPARED TO THE HUMAN TBX22 PROTEIN SEQUENCE.....</b>	105
<b>FIGURE 20: LYSING THE BACTERIAL CULTURES UNDER NON-DENATURING CONDITIONS.....</b>	106
<b>FIGURE 21: DETECTION OF THE 6xHis-TBX22 PROTEIN FROM A RABBIT RETICULOCYTE LYSATE.....</b>	108
<b>FIGURE 22: AN ELECTROMOBILITY SHIFT ASSAY OF THE TBX22 PROTEIN.....</b>	110
<b>FIGURE 23: TBX22 PROTEIN BINDS SPECIFICITY WITH THE TBX22 DNA BINDING CONSENSUS SEQUENCE.....</b>	112
<b>FIGURE 24: T-BOX PROTEINS CONTACT DNA AT SPECIFIC RECOGNITION SITES.....</b>	116
<b>FIGURE 25: THE TBX22 BINDING SITE IDENTIFIED BY THE IN SILICO SCREEN OF THE 2KB OF THE MSX1 PROMOTER IS ABLE TO SHIFT THE TBX22 PROTEIN IN AN EMSA.....</b>	136
<b>FIGURE 26: mRNA EXPRESSION OF MSX1 AND TBX22 IN THE DEVELOPING PALATE.....</b>	137
<b>FIGURE 27: A SCHEMATIC DIAGRAM SHOWING THE POSITION OF THE 983BP FRAGMENT FROM THE MSX1 PROMOTER WITHIN THE PUC57 CLONING VECTOR.....</b>	153
<b>FIGURE 28: A REPRESENTATION OF THE LAYOUT OF THE 96-WELL PLATE USED FOR EACH TRANSFECTION EXPERIMENT.....</b>	157
<b>FIGURE 29: HE LA CELLS EXPRESS ENDOGENOUS TBX22 AND MSX1, 239T CELLS EXPRESS MSX1 ONLY.....</b>	162
<b>FIGURE 30: LUCIFERASE EXPRESSION DRIVEN BY THE MSX1 PROMOTER IN HE LA AND 239T CELLS.....</b>	164

<b>FIGURE 31:</b> AN INTERVAL PLOT SHOWING THE NO SIGNIFICANT DIFFERENCE BETWEEN ABSOLUTE RELATIVE LIGHT UNITS EMITTED FROM THE CONTROL RENILLA LUCIFERASE.....	169
<b>FIGURE 32:</b> A BOX-PLOT OF THE NORMALISED RATIOS OF LUCIFERASE/RENILLA FOR NO TBX22, 26.7NG TBX22 AND 53.4NG TBX22 IN HE LA CELLS.....	170
<b>FIGURE 33:</b> A BOX-PLOT OF THE NORMALISED RATIOS OF LUCIFERASE/RENILLA FOR NO TBX22, 26.7NG TBX22 AND 53.4NG TBX22 IN 293T CELLS.....	171
<b>FIGURE 34:</b> A PROPOSED MODEL FO REPRESSION OF <i>MSX1</i> BY <i>TBX22</i> IN THE FORMATION OF PALATAL SHELVES.....	183

## TABLE OF TABLES

<b>TABLE 1:</b> A SUMMARY OF THE KEY EVENTS DURING THE FORMATION OF THE HUMAN PALATE.....	10
<b>TABLE 2:</b> A SIMPLIFIED SUMMARY OF THE STRUCTURES THAT WILL FORM FROM EACH OF THE PHARYNGEAL ARCHES.....	14
<b>TABLE 3:</b> A SUMMARY OF THE EVIDENCE FOR AND AGAINST THE PRINCIPLE MECHANISMS INVOLVED WITH MES DISAPPEARANCE. (TAKEN FROM DUDAS <i>ET AL.</i> 2007).....	31
<b>TABLE 4:</b> A LIST OF KNOWN LOCI FOR CL/P (ADAPTED FROM CARINCI <i>ET AL.</i> 2007 AND OMIM ( <a href="http://WWW.NCBI.NLM.NIH.GOV/OMIM/">HTTP://WWW.NCBI.NLM.NIH.GOV/OMIM/</a> [28 FEB 2010])).....	39
<b>TABLE 5:</b> GENES IMPLICATED IN HUMAN SYNDROMIC OROFACIAL CLEFTING BASED ON EVIDENCE FROM HUMAN GENETIC STUDIES, MOUSE MODELS, AND EXPRESSION DATA IN OROFACIAL PRIMORDIAL (FROM GRITLI-LINDE 2008; JUGESSUR <i>ET AL.</i> 2009).....	41
<b>TABLE 6:</b> A CROSS-SPECIES COMPARISON OF DEVELOPMENTAL TIMINGS (BUTLER AND JUURLINK 1987; O'RAHILLY AND MULLER 1987; HAMBURGER AND HAMILTON 1992).....	50
<b>TABLE 7:</b> EMBRYO STAGING GUIDE (FROM BULLEN AND WILSON, 1997).....	52
<b>TABLE 8:</b> FOETAL STAGING GUIDE (FROM HERN, 1984).....	53
<b>TABLE 9:</b> SITES OF <i>TBX22</i> EXPRESSION DURING DEVELOPMENT OF THE HUMAN, MOUSE AND CHICK (MOUSE DATA FROM (BUSH <i>ET AL.</i> 2002) AND CHICK DATA FROM (HAENIG <i>ET AL.</i> 2002)).....	78
<b>TABLE 10:</b> A SUMMARY OF DIFFERENT EXPERIMENTALLY DETERMINED T-BOX PROTEIN DNA BINDING SITES.....	83
<b>TABLE 11:</b> THE RECOMMENDED IUPAC ABBREVIATIONS FOR NUCLEIC ACIDS (IUPAC-IUB 1971).....	83

<b>TABLE 12:</b> CELL LYSIS BUFFERS USED TO SOLUBILISE THE 6XHis TAGGED PROTEIN UNDER DENATURING CONDITIONS.....	92
<b>TABLE 13:</b> THE OLIGONUCLEOTIDES USED TO GENERATE THE DOUBLE-STRANDED OLIGOS USED IN THE ELECTROPHORETIC MOBILITY SHIFT ASSAY OF TBX22 PROTEIN.....	100
<b>TABLE 14:</b> DETERMINING THE TBX22 BINDING SITE.....	109
<b>TABLE 15:</b> SEQUENCES OF T-BOX PROTEIN DNA BINDING SITES INCLUDED WHEN DETERMINING GENERIC TBX22 BINDING SITE.....	129
<b>TABLE 16:</b> A LIST OF THE HUMAN CLEFT PALATE GENES CONTAINING A HIT TO THE GENERIC TBX22 BINDING SITE.....	131
<b>TABLE 17:</b> HUMAN CLEFT PALATE GENES SHOWING THREE OR MORE GENERIC TBX22 BINDING SITES WITH A ONE BASE MIS-MATCH.....	133
<b>TABLE 18:</b> THE KNOWN HUMAN CLEFT PALATE CAUSING GENES THAT CONTAIN EITHER AN EXACT MATCH TO THE GENERIC TBX22 BINDING SITE, OR AT LEAST 3 COPIES OF A 1 BASE MIS-MATCH TO IT, IN A 2KB REGION FROM THE START OF TRANSCRIPTION AND ALSO GIVE RISE TO A CLEFT PALATE PHENOTYPE WHEN DISRUPTED IN MOUSE.....	135
<b>TABLE 19:</b> A SUMMARY OF THE PLASMIDS USED IN THE TRANSFECTION EXPERIMENTS.....	157
<b>TABLE 20:</b> RELATIVE LIGHT UNITS FROM LUCIFERASE AND RENILLA ASSAY IN HE LA CELLS.....	166
<b>TABLE 21:</b> RELATIVE LIGHT UNITS FROM LUCIFERASE AND RENILLA ASSAY IN 239T CELLS.....	167
<b>TABLE 22:</b> A SUMMARY OF THE ASSUMPTIONS MADE IN THE INTERPRETATION OF RESULTS FOR THE TRANSFECTION ASSAY AND THE CONTROLS THAT SHOULD BE INCLUDED TO VERIFY THAT THEY HOLD TRUE.....	174

## TABLE OF ABBREVIATIONS

BMP	bone morphogenetic protein
bp	base pairs
BSA	bovine serum albumin
cDNA	complementary DNA
CL/P	cleft lip with or without cleft palate
CMV	<i>cytomegalovirus</i>
CPI	isolated cleft palate
cpm	counts per minute
CPX	X-linked cleft palate
CS	Carnegie stage
Da	Dalton
DAB	3,3'-diaminobenzidine tetrahydrochloride
DEPC	diethylpyrocarbonate
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
dpc	days post conception
dpo	days post-ovulation
DTT	dithiothreitol
EDTA	ethylene diamine tetraacetic acid
EMSA	electrophoretic mobility shift assay
g	gravitational acceleration (9.8m/s <sup>2</sup> )
HDBR	Human Developmental Biology Resource
HH stage	Hamburger-Hamilton stage
His	histidine
HSV	Herpes simplex virus
HRP	horse radish peroxidase
IPTG	isopropyl $\beta$ -D-1-thiogalactopyranoside
kb	kilobases
LB	Luria Bertani medium
M	molar
m	milli
MALDI-TOF	matrix-assisted laser desorption/ionization-time of flight
MEE	medial edge epithelial cells
MEM	minimum Essential medium
MES	medial epithelial seam
mRNA	messenger RNA
Milli-Q	deionised and purified water to a resistance of 18.2 M $\Omega$ cm at 25 °C, using Milli-Q Q-POD unit from Millipore
MSX1	muscle segment homeodomain homeobox 1
n	nano
NCC	neural crest cells
Ni-NTA	nickel-nitrilotriacetic acid
N-terminal	amino-terminal
OD	optical density
OPT	Optical Project Tomography
p	pico
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PNK	polynucleotide kinase

PVDF	polyvinylidene Fluoride
RA	retinoic acid
RNA	ribonucleic acid
rpm	revolutions per minute
SSC	sodium chloride, sodium citrate buffer
T	Brachyury
TBX22	T-box 22
TBE	T-box binding element
TBE buffer	tris base, boric acid and EDTA buffer
TE buffer	tris-EDTA buffer
TISH	tissue <i>in situ</i> hybridisation
TK	thymidine kinase
tRNA	transfer RNA
U	units
UV	ultraviolet

# CHAPTER 1

## INTRODUCTION

The thesis centres upon TBX22 - the underlying genetic cause of X-linked cleft palate (Braybrook *et al.* 2001). More specifically, this thesis will consider the mRNA expression pattern of *TBX22*; the DNA binding specificity of the TBX22 protein and the investigation of a possible downstream target gene of the TBX22 protein.

An orofacial cleft is a frequent birth defect. Recent figures for the population of the USA estimate that the average prevalence of cleft lip with or without a cleft palate to be 10.48 per 10000 live births and those born with a cleft palate without the presence of a cleft lip to be 6.39 for every 10000 (C.D.C. 2006). Several genes have been uncovered where a cleft palate is present together with multiple congenital malformations as part of a syndrome. Whilst studies are uncovering genes associated with non-syndromic oral clefts all the time (Moreno *et al.* 2009; Carter *et al.* 2010; Ingersoll *et al.* 2010), *TBX22* remains one of only a handful of genes that have been shown to underlie non-syndromic cleft (Carinci *et al.* 2007). Although some patients with mutations in the *TBX22* gene display ankyloglossia (tongue-tie) as well as a cleft of the secondary palate, *TBX22* mutations account for a significant proportion of non-syndromic cleft palate cases (Marcano *et al.* 2004; Suphapeetiporn *et al.* 2007).

This thesis attempts to further the understanding of the biological role of *TBX22* through its interaction with other genes. This is a necessary step in order that *TBX22* can be placed in relation to other transcription factors and signalling molecules within the complex mechanisms underpinning palate formation.

The following introduction aims to provide a background to the study and to highlight the context of the work from the start of the project 2001 to our current understanding. Topics introduced here will be expanded upon in specific chapters as indicated in 1.4. This introduction will begin with a brief overview of the *TBX22* gene and the T-box

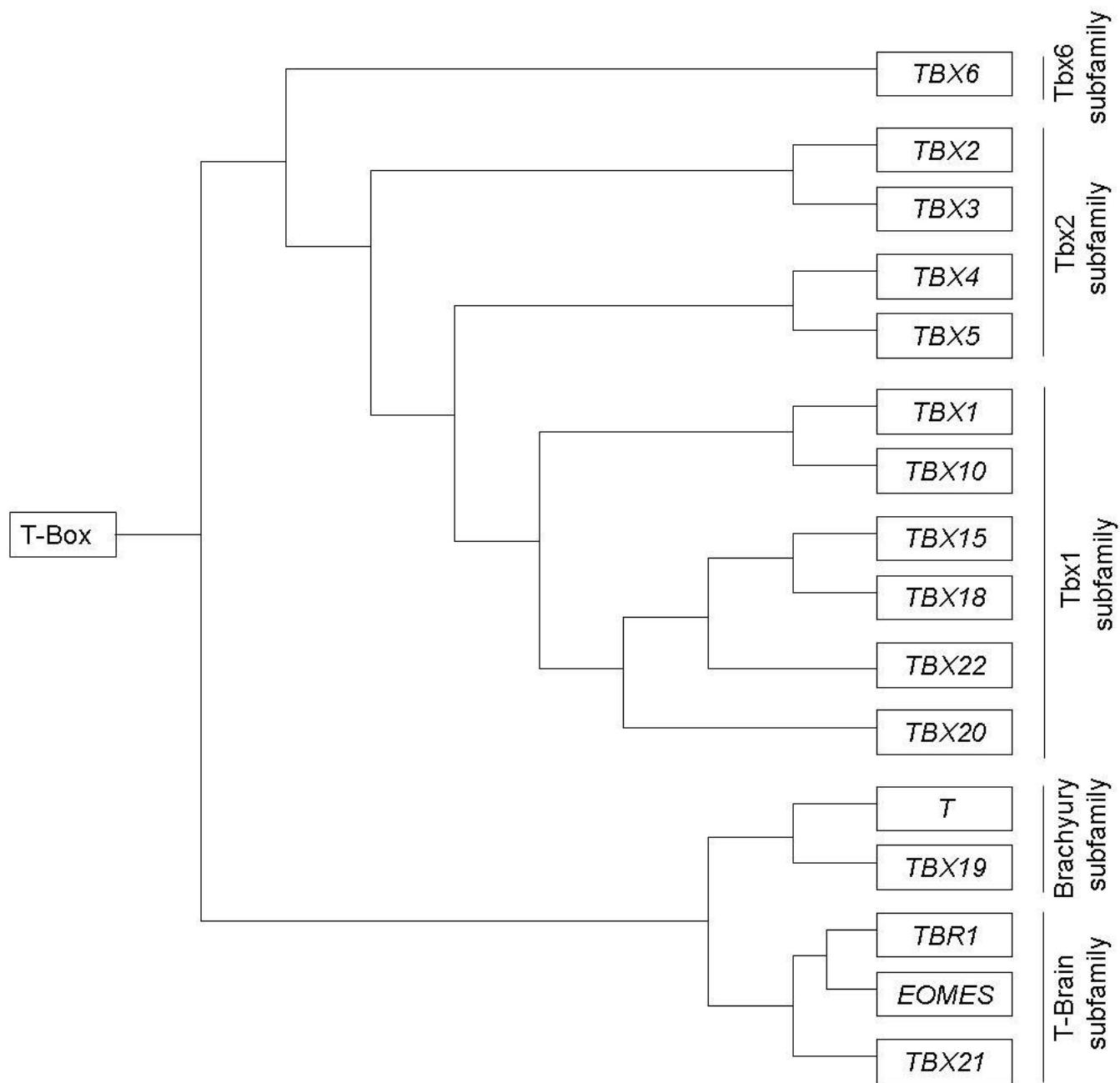
gene family, followed by a review of the development of the head and face with particular emphasis on the formation of the palate; relating how defects in the correct formation of these structures may give rise to a cleft, giving an overview of known cleft palate genes; as well as giving more detail about the X-linked cleft palate phenotype.

## 1.1 BRACHYURY, TBX22 AND THE T-BOX GENE FAMILY

In 1927, a mouse mutant was described as having a short tail phenotype and the mutation responsible named T, for tail, or more elaborately, Brachyury, the Greek for short tail (Dobrovolskia-Zavadskaa 1927). The *T* gene was identified when the Brachyury mutation was isolated more than sixty years later (Herrmann *et al.* 1990). Soon after, the *Drosophila* gene, optomotor-blind (omb), was shown to be related to the *T* gene (Pflugfelder *et al.* 1992), when an homology domain was identified between the central region of the omb protein and the amino-terminal region of *T*. Despite clearly not being orthologues, these two gene products certainly belonged to the same protein family. The homologous region was named the T-box, and since the discovery of sequence similarity between *omb* and *T*, a host of other genes have been identified in many diverse animal species that share this same homology, known collectively as T-box genes. T-box genes are part of the larger protein clan, P53-like, which share similarity in their immunoglobulin-like, beta-sandwich DNA-binding domains (Berardi *et al.* 1999); pfam [<http://pfam.sanger.ac.uk/>]).

The T-box domain is a conserved region encompassing ~200 amino acids, and has been shown to display binding activity to specific DNA sequences (Kispert and Herrmann 1993). All T-box genes show ability to bind variations of a similar DNA sequence (Tada and Smith 2001); known as the T-box binding element. It is this DNA-binding capability that enables T-box genes to function as transcriptional regulators. The importance of these transcription factors during development is highlighted by the effect of mutations in T-box genes in different organisms and the resultant phenotypes observed (Chapman and Papaioannou 1998; Bruneau *et al.* 2001; Jerome and Papaioannou 2001; Szabo *et al.* 2002; Davenport *et al.* 2003; Naiche and Papaioannou 2003; Bussen *et al.* 2004; Harrelson *et al.* 2004; Singh *et al.* 2005a;

Singh *et al.* 2005b). The developmental functions of T-box genes are varied: having a role in very early embryogenesis – specifying the primary germ layers as in the case of Brachyury and VegT (reviewed in Showell *et al.* 2004); as well as later in development, for example in conveying identity and patterning the limbs (reviewed in Simon 1999).



**FIGURE 1: Phylogenetic tree of human T-box genes.** The evolution of the human T-box genes is shown in relation to the common ancestral genes from which they arose during gene duplication events (adapted from Ruvinsky *et al.* 2000; Larroux *et al.* 2008).

By comparative analysis of the genomes of lower organisms, particularly that of the cephalochordate amphioxus which is believed to be the closest living invertebrate relative to vertebrates (Holland *et al.* 2004), to those of higher organisms the evolution of the T-box genes from a single ancestral gene has been determined (see Fig. 1). The amphioxus genome only contains a basic set of chordate genes involved in development and many of these have retained their identity through evolution through to higher organisms. However, whereas the amphioxus genome only has one of these genes, as a consequence of whole genome duplication, vertebrates will often have two, three, or four paralogues of the same genes (Holland *et al.* 2008).

The duplication of pre-existing genes, whether through whole genome, or whole or partial chromosome duplication, had a major role in evolution and has lead to the emergence of the various human T-box genes. Following a gene duplication event, one copy may inherit a null mutation, and that copy is likely to be subsequently lost, providing the function is maintained by the other copy. However, if one or both copies of the gene gain mutations such that a new function is acquired, then this gene will be retained through evolution through positive selection. A third possibility is that mutations result in both gene copies having reduced expression and/or functional alterations such that both copies of the gene must be retained in order for the function of the ancestral gene to be preserved (Minguillon and Logan 2003; Innan and Kondrashov 2010).

The phylogenetic study of the T-box genes has enabled all of the human T-box genes to be grouped together into one of five subfamilies: *Brachyury* (*T*), *T-Brain*, *TBX1*, *TBX2* and *TBX6* (see Fig. 1) based upon the ancestral gene from which they arose. *TBX22* is a member of the *Tbx1* subfamily and is most closely related to *TBX15* and *TBX18* (Ruvinsky *et al.* 2000; Larroux *et al.* 2008). Genes of the same subfamily often display over-lapping expression patterns. Several members of the *Tbx1* subfamily - *Tbx1*, *Tbx10*, *Tbx15*, *Tbx18* and *Tbx22* have been shown to be expressed during craniofacial development (Chapman *et al.* 1996; Agulnik *et al.* 1998; Kraus *et al.* 2001a; Bush *et al.* 2002; Bush *et al.* 2003; Herr *et al.* 2003), however this is not apparent for *Tbx20* (Kraus *et al.* 2001b).

The possibility of over-lapping but distinct functions is also highlighted by the human disorders that mutations within related T-box genes cause. Mutations in *TBX5* have been shown to cause Holt-Oram syndrome (Basson *et al.* 1997; Li *et al.* 1997). Characteristically this syndrome involves atrial septal defects and anomalies of the thumbs (MIM #142900) (Holt and Oram 1960). Mutations in *TBX3*, similarly to *TBX5* a member of the Tbx2 subfamily, lead to Ulnar-Mammary syndrome (MIM #181450) (Bamshad *et al.* 1997). The phenotype of this disorder is usually characterised by complete absence or malformation of the fingers, rather than the thumbs as seen in Holt-Oram syndrome, often accompanied with delayed growth and onset of puberty, obesity, hypogenitalism and hypoplasia of nipples (Schinzel *et al.* 1987).

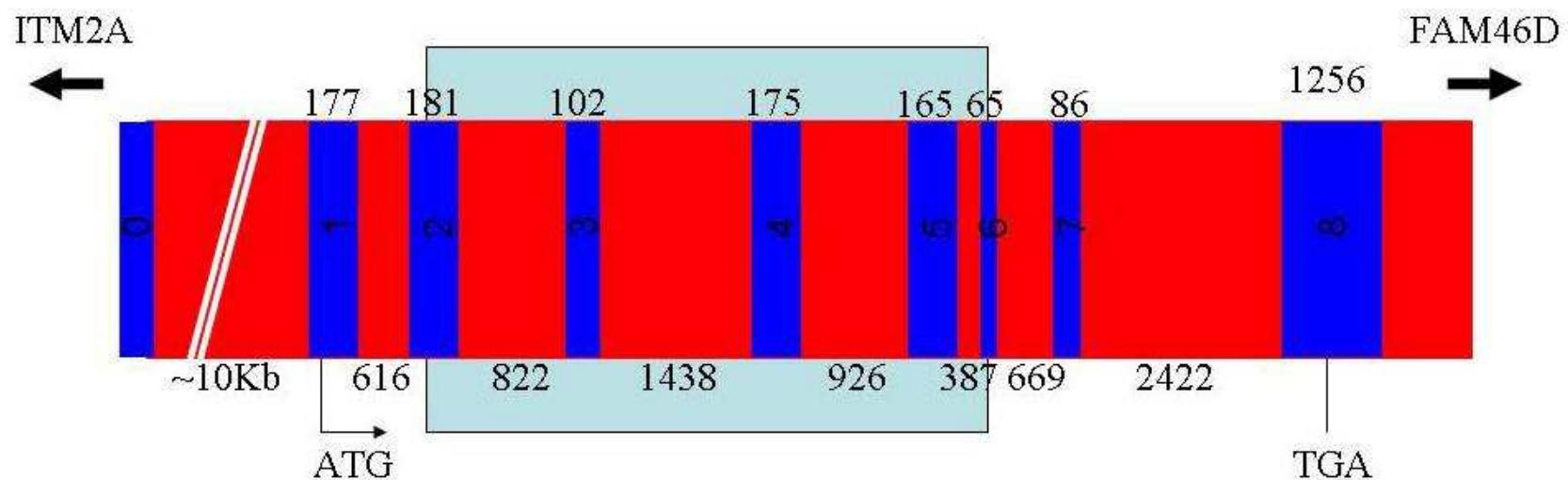
The importance of T-box genes during normal development is clear given the disease phenotypes described above and further human diseases caused by mutations in other T-box genes have been described. Mutations in *TBX1* have been found in patients with conotruncal anomaly face syndrome (MIM #217095) (Yagi *et al.* 2003) and (Paylor *et al.* 2006; Zweier *et al.* 2007) and have also been linked to DiGeorge syndrome (MIM # 188400) (Yagi *et al.* 2003; Stoller and Epstein 2005). Mutations in *TBX4* were shown to cause small patella syndrome (MIM #147891) (Bongers *et al.* 2004); mutations in *TBX15* give rise to Cousin syndrome (MIM #260660) (Lausch *et al.* 2008); mutations to *TBX19* cause adrenocorticotropic hormone deficiency (MIM #201400) (Lamolet *et al.* 2001) and disruption to the regulatory region surrounding *EOMES* co-segregated with microcephaly with polymicrogyria and corpus callosum agenesis in one family (Baala *et al.* 2007). A single nucleotide polymorphism in the promoter region of *TBX21* has been shown to lead to an increased risk to aspirin-induced asthma (Akahoshi *et al.* 2005) and carrying an allelic variant of the *T* locus, TIVS7-2, has been associated with increased risk of neural tube defects and spina bifida (Morrison *et al.* 1996; Jensen *et al.* 2004)

### **1.1.1 TBX22 Chromosomal location and Gene Structure**

Whilst trying to identify candidate genes involved in X-linked mental retardation, through searching the human genome databases and using gene prediction tools,

Laugier-Anfossi and Villard isolated a new T-box gene, namely, *TBX22* (Laugier-Anfossi and Villard 2000). This gene was originally thought to have a truncated T-box domain, which was missing the first 20 amino acids and in a zoo blot was found only in the porcine genome and absent in all others, including the mouse, leading the authors to suggest that *TBX22* may have appeared relatively recently in the evolution of higher mammals. Additional upstream exons in *TBX22* were later identified, increasing the *TBX22* protein by 120 amino acids, showing that *TBX22* did indeed contain a full length T-box domain (Braybrook *et al.* 2001) and far from being a gene found exclusively in higher mammals, phylogenetic analysis has shown that since its diversion from *TBX15/18*, *TBX22* is in fact conserved throughout metazoan evolution (Papaioannou 2001).

The *TBX22* gene consists of eight coding exons and one non-coding exon and is located between the genes *ITMA* and *FAM46D* on the chromosome Xq21.1. Three mRNA transcripts exist; the first contains exons 1-7, whilst a second contains the additional exon 0 spliced to the 5' end. However, both transcripts encode the same protein, using the same translation start site located in exon 1 (Andreou *et al.* 2007). A third smaller transcript includes the 5' exon 0, but is translated from an AUG site further downstream resulting in a shorter isoform (Andreou *et al.* 2007). The T-box domain is located between exons 2 and 6 (Fig. 2).

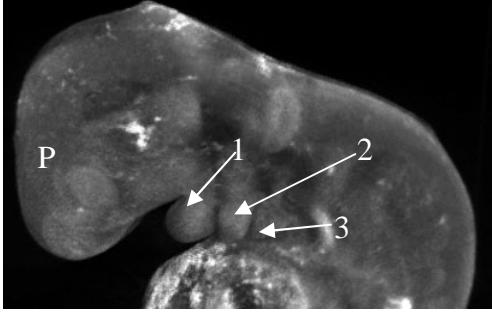
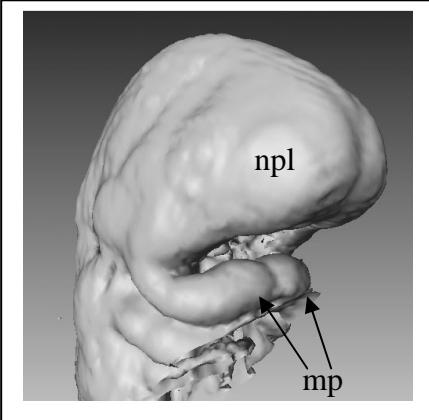


6

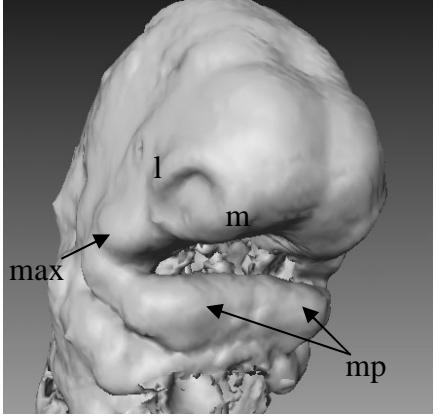
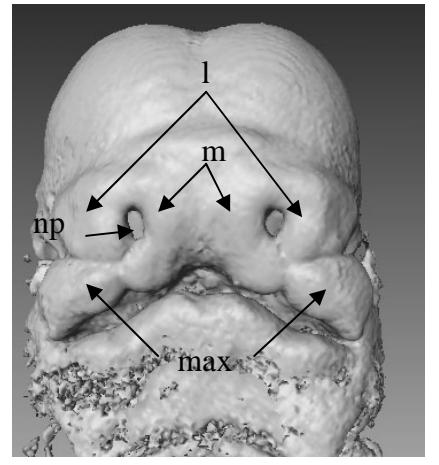
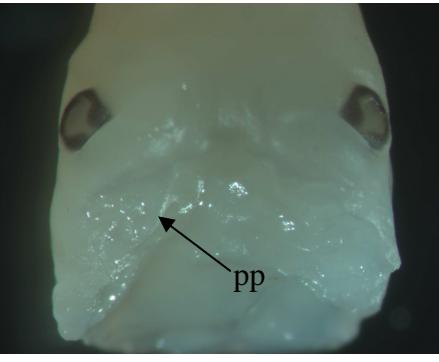
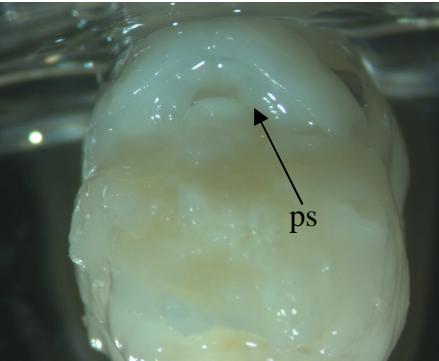
**Figure 2: The genomic structure of *TBX22*.** A schematic diagram of the genomic structure of *TBX22*, based on Braybrook *et al.* 2001; Andreou *et al.* 2007. *TBX22* is comprised of eight coding exons (1-8) and one non-coding exon (0). The numbers below the red bar indicate the numbers of bases of intronic sequence and the numbers above the blue bars indicate the size of each intron, the light blue region indicates the T-box domain. The start and stop codons are also shown.

## 1.2 EARLY MORPHOGENESIS OF THE HUMAN FACE

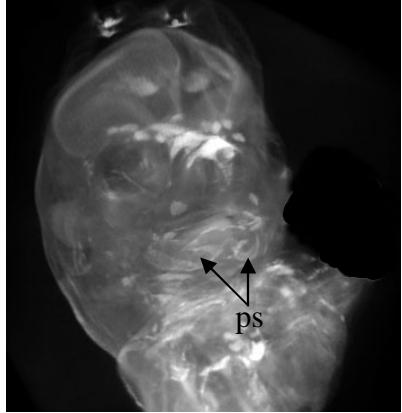
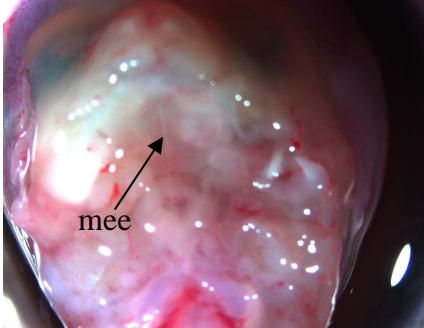
A disturbance to any of the mechanisms involved in the correct fusion of the palatal shelves, from the very foundations of head formation right through to the final fusion of the palatal shelves themselves may result in a cleft of the palate (Satokata and Maas 1994). As mutations in *TBX22* result in a cleft palate, then it is highly likely that *TBX22* will be involved with one or more of these mechanisms during development. It is for this reason that the following background on the formation of the human face and palate is given; Table 1 provides a summary of the main events and timings and Fig. 3 summarises the precursors of the future face. The source and staging systems used for embryos and foetuses are given in section 2.2.1.

Carnegie Stage	Embryo or OPT Image	Summary of events
CS12	 <p>OPT image of a Carnegie Stage 12 embryo. The image shows the prosencephalon (p) at the top, with three pharyngeal arches labeled 1, 2, and 3. Arrows point from the labels to the corresponding arches.</p>	<p>Between ~CS9 and CS12, neural crest from specific neuromeres migrate to a predetermined position; either a specific pharyngeal arch (arches 1, 2 and 3 indicated in image), or the mesoderm surrounding the prosencephalon (p).</p>
CS13	 <p>3D rendering of a Carnegie Stage 13 embryo. The image shows the nasal placodes (npl) in the frontonasal process and the mandibular processes (mp) forming from the first pharyngeal arch.</p>	<p>Nasal placodes (npl) form in the frontonasal process. Paired swellings (mandibular processes) form from the first pharyngeal arch (mp).</p>

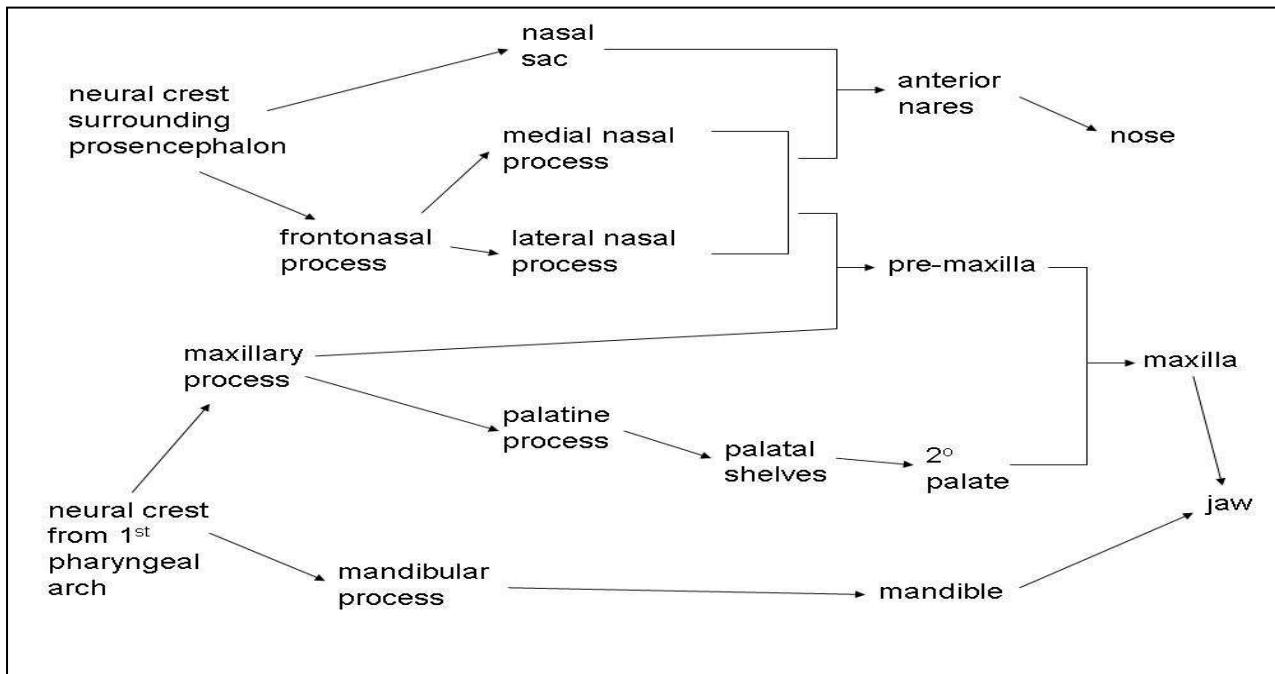
continued overleaf ...

vCS14		<p>The mandibular processes have fused in the midline (mp). Appearance of the facial processes - medial (m) and lateral (l) nasal processes and the maxillary process (max) can be seen.</p>
CS16		<p>By CS15-CS16 fusion of the nasal and maxillary processes – forms the premaxilla. Fusion of the medial and lateral nasal processes, combined with deepening of the nasal pit (np) forms the anterior nares.</p>
CS17		<p>From CS16 to CS17, the palatine processes (pp) emerge from the maxillary processes.</p>
CS20		<p>Between CS19 and CS21, there is further proliferation of the palatine processes forming palatal shelves (ps) which extend vertically either side of the tongue.</p>

continued overleaf ...

CS22		From around CS20 to CS22, the palatal shelves (ps) move to take a horizontal position above the tongue and continue to grow towards each other.
9 wpc		The palatal shelves contact and fuse with each other. The remains of the medial edge epithelia (mee) can be seen in the midline of the secondary palate.

**Table 1: A summary of the key events during the formation of the human palate.** Images representing CS12, CS13, CS14, CS16 and CS22 are optical projection tomography reconstructions (Sharpe *et al.* 2002) generated from intact embryos. Volume renders of the OPT reconstructions are shown at CS12 and CS22 while surface renders of the reconstructions are shown at CS13, CS14 and CS16. The CS17, CS20 and the 9 week post conception (wpc) foetus images are of the head with the mandible removed to aid viewing using a Zeiss stereo microscope. All specimens are from the Newcastle HDBR collection (see 2.2.1).

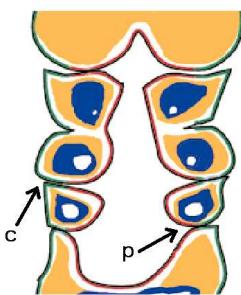


**Figure 3: The development of the human face.** A schematic representation showing the structures formed during the development of the human face

### 1.2.1 The Pharyngeal Arches

The first external evidence of the pharyngeal arches, are the paired swellings of the mandibular (arch 1) and hyoid (arch 2) in the CS10 embryo. The arches form in cranio-caudal sequence and the five pharyngeal arches that develop in human correspond to numbers 1, 2, 3, 4 and 6 of the primitive complement in the evolutionary line leading to land vertebrates. Arch 5 never develops in humans, or else forms as a short lived rudiment and regresses (Larsen 1997).

Initially, the arches resemble gills, except that, the gill slits never become perforated. Instead, the internal pharyngeal pouches (p, in Fig. 4) are separated from the external pharyngeal clefts (c, in Fig. 4) by three cell layers - an outer covering of ectoderm, an inner layer of endoderm and a central core of mesenchymal tissue derived from both neural crest and mesoderm.



**Figure 4: The mammalian pharyngeal arches.** A frontal diagram representing the mammalian pharyngeal arches (adapted from Graham and Smith 2001). The components of the arches indicated by colour: Neural crest derived mesenchyme, yellow; ectoderm, green; mesoderm, blue; endoderm, red. The external pharyngeal clefts (c) and the internal pharyngeal pouches (p) are highlighted.

Each of these different cell populations will eventually become distinct tissues: the ectoderm giving rise to the epidermis and sensory neurons associated with that arch (Couly and Le Douarin 1990); the endoderm will become the oral epithelia lining (Haworth *et al.* 2004); the mesenchyme derived from the neural crest and mesoderm will form primarily the smooth and skeletal muscles of the face and connective and skeletal tissues of the skull and jaw (Bhattacherjee *et al.* 2007).

It is important that the arches are patterned correctly, as distinct structures develop from different pharyngeal arches. The eventual structures that arise from each arch are summarised in Table 2.

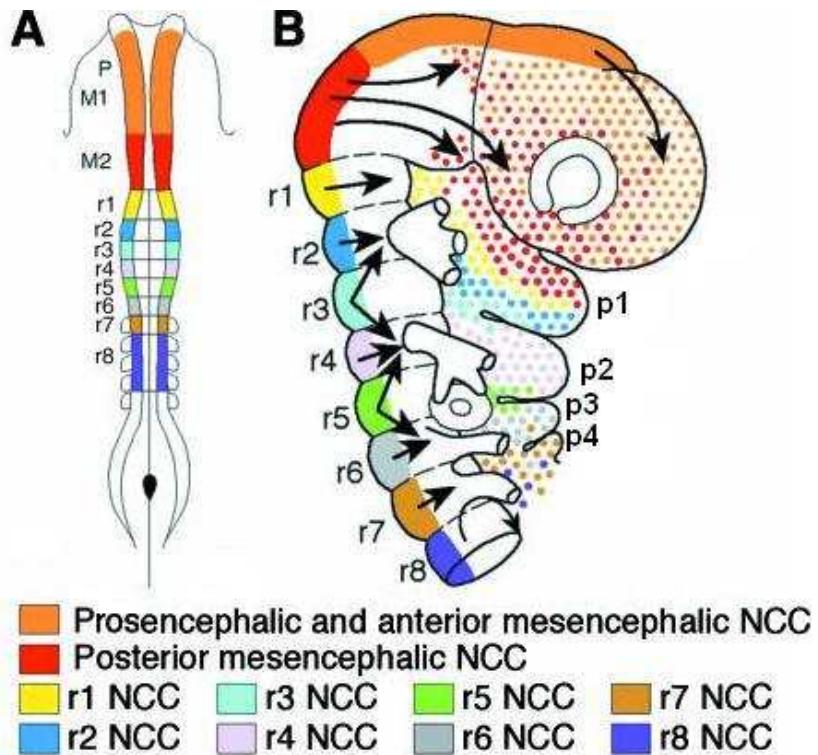
Pharyngeal arch	Structures
1st (mandibular)	Mandible and muscles of the lower jaw
2nd (hyoid)	Hyoid, stapes bone of the ear, the throat and neck
3rd	Elevator muscles of pharynx and tongue
4th	Constrictor muscles of the pharynx and vocal chords
6th	Muscles of the larynx

**Table 2: A simplified summary of the structures that will form from each of the pharyngeal arches.** The actual structures that are derived from each arch are more extensive than is represented here (Drake *et al.* 2009).

The embryonic neural tube (in the brain) is segmented anterior-posterior into prosencephalon, mesencephalon, and rhombencephalon. These are then further subdivided into discrete segments; known as neuromeres. Each neuromere has a specific identity in relation to its position along the neuroaxis (Fig. 5). The prosencephalon is divided into six prosomeres (p1-p6) (Rubenstein *et al.* 1994; Puelles and Rubenstein 2003), and the mesencephalon into two mesomeres (m1 and m2) which are separated from the rhombomeres by the isthmic neuromere (Muller and O'Rahilly 1997). Rhombomeres 1-7 (rh1-rh7) can usually be distinguished while rh8 incorporates rhombomeres 8-11 (see Fig. 5). This early segmentation of the embryo provides the blueprint for patterning the developing head and neck, largely dependent upon neural crest cells (NCC) that migrate from neuromeres and intermingle with other cell types in specific pharyngeal arches. NCC arising from the neural folds of the developing embryo are multipotent and migrate to form many varied structures (Bronner-Fraser 1995). As a general rule: neural crest migrating to cell populations of mesoderm generally induce membranous bone; in the absence of mesoderm, neural crest in contact with an epithelial derivative will induce only cartilage (Carstens 2002).

The majority of experimental work investigating the role of neural crest in pharyngeal arch patterning has been performed in the chick due to the relative ease of transplantation of embryonic tissue. It is inferred from these experiments that the processes will be similar in human, although it must be remembered that in the following discussion the data comes from the chick, and must only be used as a model for development in human.

The tissues that will arise from the each pharyngeal arch into which the neural crest cells migrate, is somewhat predetermined by the NCC themselves, prior to their migration from their neuromeres. This was demonstrated in chick and quail embryos by first exercising the NCC from pharyngeal arches 2 and 3 and replacing with transplanted NCC that would normally migrate to the first pharyngeal arch, and then following the subsequent migration of these cells into the pharyngeal arches (Noden 1983). This resulted in the animals developing tissues normally associated with the first pharyngeal arch, including a beak-like structure, being superimposed onto structures that would normally develop from pharyngeal arches 2 and 3.



**Figure 5: The neural crest from each rhombomere migrates to a specific position along the neuroaxis.** The neural crest cells (NCC) migrate from a specific neuromere (prosomere (P), mesomere (M1 and M2) or rhombomere (rh1-8)). **A** shows the position of the neuromeres along the neuroaxis and **B** indicates where the NCC from each neuromere migrates to. p1-p4 represent pharyngeal arches 1-4 (adapted from Creuzet *et al.* 2005).

Each of the neuromeres can be defined by the gene expression that they exhibit (for review see (Trainor and Krumlauf 2001; Creuzet *et al.* 2005; Iimura and Pourquie 2007; Wellik 2007) and references therein), and each neuromere shares a common gene expression pattern with the structures to which the neural crest from that neuromere migrates. For example, the nasal and orbital mesoderm shares a common gene coding with prosomeres p6 and p5 respectively, as this is where the neural crest in these structures originated from. It is because the same NCC populations are present in both of these structures that clinical conditions affecting naso-orbital structures are often accompanied by structural abnormalities of the prosencephalon (Carstens 2002).

Using quail-chick chimeras to study the long-term fate of neural crest subpopulations of individual neuromeres (Kontges and Lumsden 1996) showed that each pharyngeal arch mesenchyme is derived from a particular neuromere of the cranial neural crest (see Fig. 5). The most distal regions of the first arch, including those regions that will form the maxilla and palatine bones, are exclusively derived from the midbrain neural crest. The rhombomere 1 and 2 neural crest forms the more proximal regions of the 1<sup>st</sup> arch, with rhombomere 2 derivatives surrounding the middle ear cavity. The second arch was shown to be derived mainly from rhombomere 4 NCC, with a few cells from rhombomeres 3 and 5 that formed small cell islands. The NCC from rhombomeres 6 and 7 were shown to mix freely with each other in the third and fourth pharyngeal arches. This separation of neural crest into discrete migratory streams (i.e. neural crest from rhombomere 4 migrating to arch 2, and neural crest from rhombomeres 6 and 7 to arches 3 and 4) is aided by the majority of neural crest cells from rhombomeres 3 and 5 undergoing apoptosis (Lumsden *et al.* 1991). The apoptosis seen in rhombomeres 3 and 5 is directed by the up regulation of *Msx1* and *Msx2* (Graham *et al.* 1993).

The majority of the pharyngeal arch mesenchyme is derived from neural crest originating in the rhombencephalon (Noden 1983). The *HOX* gene family show discrete expression domains within the rhombomeres, displaying a definite boundary of expression between the rhombomeres (Nieto *et al.* 1992). This expression is replicated in the pharyngeal arch that the NCC from these rhombomeres migrate to, such that each arch has a specific combination of *HOX* gene expression (Trainor and Krumlauf 2001; Wellik 2007). This shows that the mesenchyme within the arches is derived from neural crest from specific neuromeres, and implies that specific sets of *HOX* genes impart the molecular specification to allow correct positioning and identity to the arches, often referred to as the *Hox* Code (Hunt *et al.* 1991). It has been shown that mutations in these *HOX* genes result in phenotypes with disrupted arch formation (Condie and Capecchi 1993; Gendron-Maguire *et al.* 1993; Rijli *et al.* 1993).

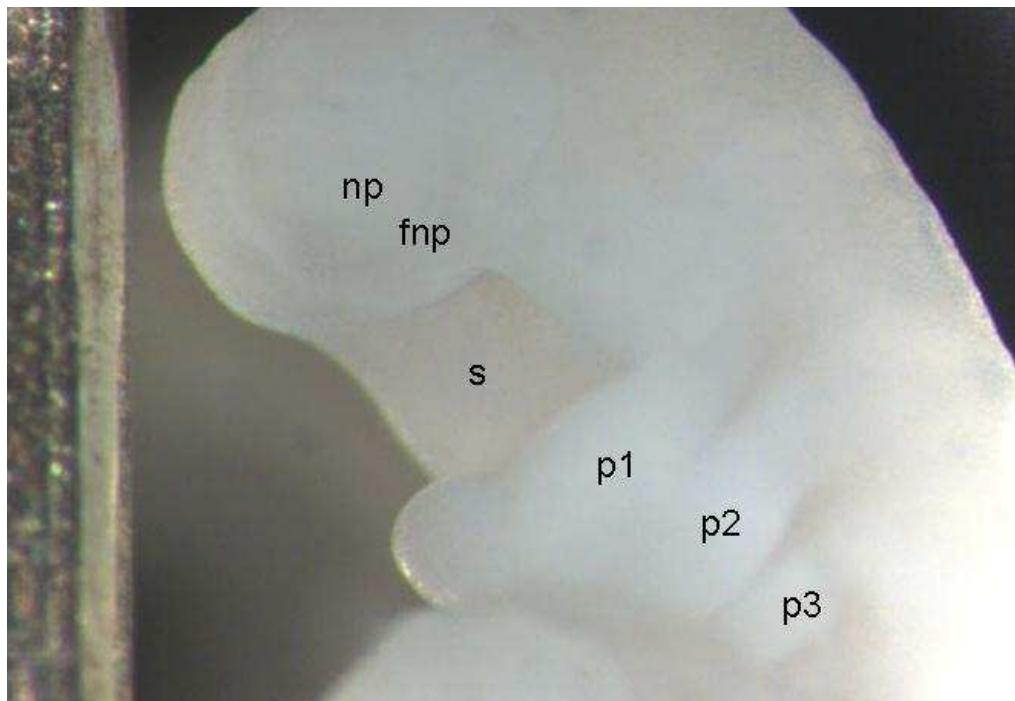
The experiments of Noden and colleagues, described above, suggested that neural crest would be pre-programmed and that the NCC would themselves confer the

identity of the arch to the surrounding cells. However, by transposing neural crest from one rhombomere to a pharyngeal arch that it was not intended to migrate to, and by the subsequent expression of the correct HOX for that arch - despite the presence of the foreign neural crest - it was demonstrated that NCC display a certain amount of plasticity (Trainor and Krumlauf 2000; Schilling *et al.* 2001). It was suggested therefore, that patterning of the pharyngeal arches was determined by signalling between the neural crest and other cells in the surrounding environment. Indeed, it has been shown that pharyngeal arches are capable of forming in the absence of neural crest (Veitch *et al.* 1999), whilst maintaining regionalised identity. However, in order that the complex structures that arise from the pharyngeal arches that lower animals do not possess are formed, it is likely that interactions are necessary between signalling molecules arising from both the neural crest of the neuromeres, and the ectoderm, endoderm and mesoderm of the pharyngeal arches to which they migrate (Graham and Smith 2001; Trainor and Krumlauf 2001).

### **1.2.2 Formation of the Facial Processes**

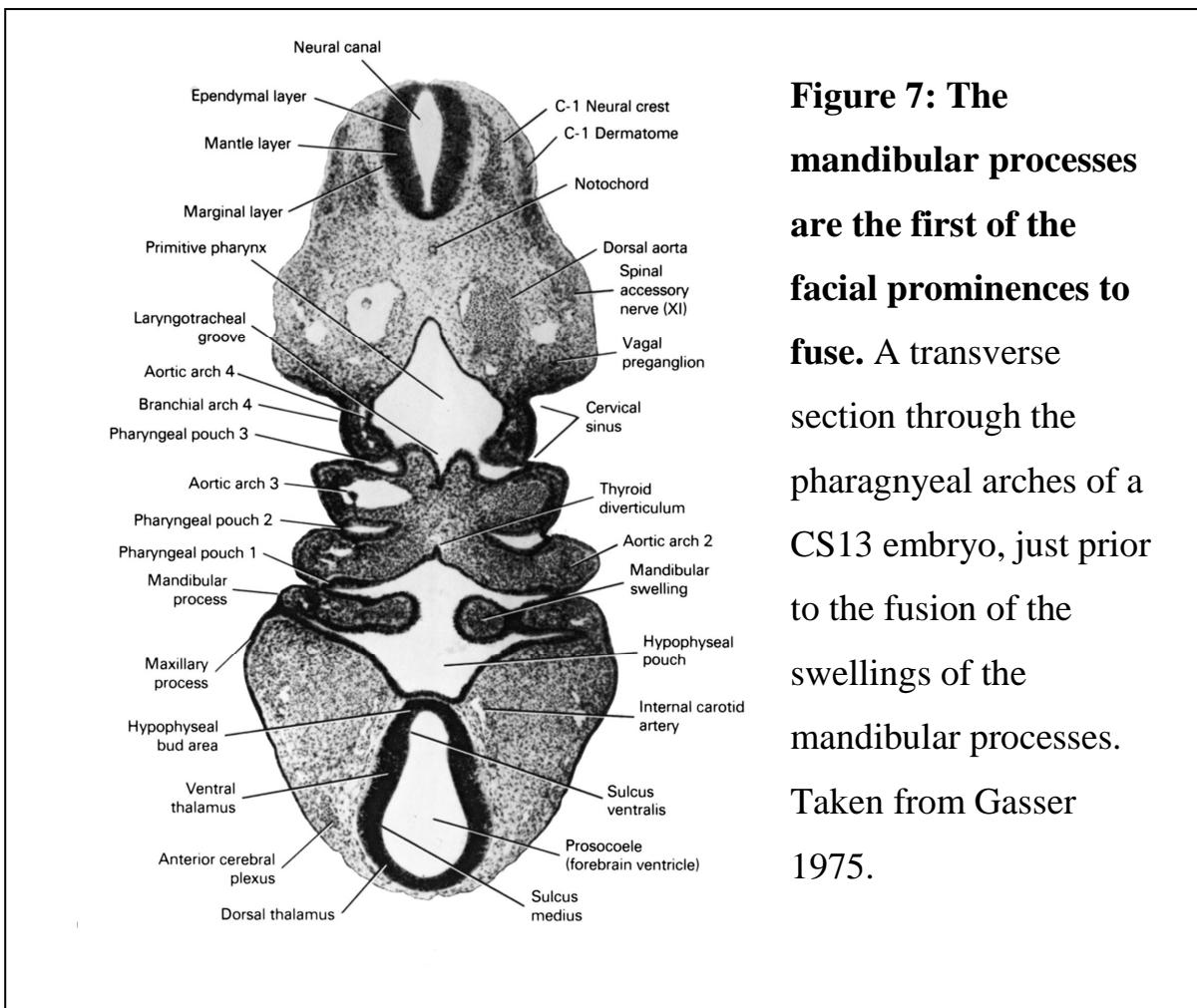
The tissue in front of prosomeres 5 and 6 is invaded by migrating NCC to establish a new layer of mesenchymal-ectomesenchymal tissue known as the frontonasal process (fnp in Fig. 6). The oronasal cavity (or stomodeum) is bound rostrally by the frontonasal process and caudally by the developing gut and cardiac plate (Fig. 6). The stomodeum is sealed from the outside amniotic cavity by the buccopharyngeal (or oropharyngeal) membrane, which will later breakdown around CS11 (Yoon *et al.* 2000) which allows amniotic fluid to come into contact with the internal structures of the developing embryo (McLachlan 1994). Two regions of thickened surface ectoderm appear in the frontonasal process of the CS13 embryo, the nasal discs, or placodes (O'Rahilly and Muller 1987). This is accompanied by the formation of a groove in the mandibular pharyngeal arch which gives rise to a pair of maxillary and mandibular processes (Yoon *et al.* 2000).

These facial processes (or prominences): maxillary, mandibular and the frontonasal processes; were first described by the 19<sup>th</sup> Century Swiss anatomist Wilhelm His and will form the basic morphology of the human face. Carstens (2002) noted that each process is not a singular anatomical unit in itself, but is composed of many cell populations and therefore proposed the field theory of facial midline development. Field theory defines structures by the neuromeres from which they are coded: all structures derived from prechordal mesoderm are known as A fields and are coded by prosomeres p5 and p6; structures that are associated with paraxial mesoderm receive their coding from the rhombomeres and are known as the B fields. Therefore one must be mindful when describing fusion of the facial processes, that each process is not in itself a singular structure, rather it is composed of several cell types all interacting with each other and their surroundings, and that the facial processes that we see are the result of these inter-cellular interactions.



**Figure 6: The Pharyngeal arches and frontonasal process.** A stereomicroscopy image of the head region of a CS12 embryo from the Newcastle HDBR collection (see 2.2.1) indicating the frontonasal process (fnp), the recessed nasal placodes (np) and stomodeum (s). The first (1), second (2) and third (3) pharyngeal arches are also evident. The first pharyngeal arch is also called the mandibular arch.

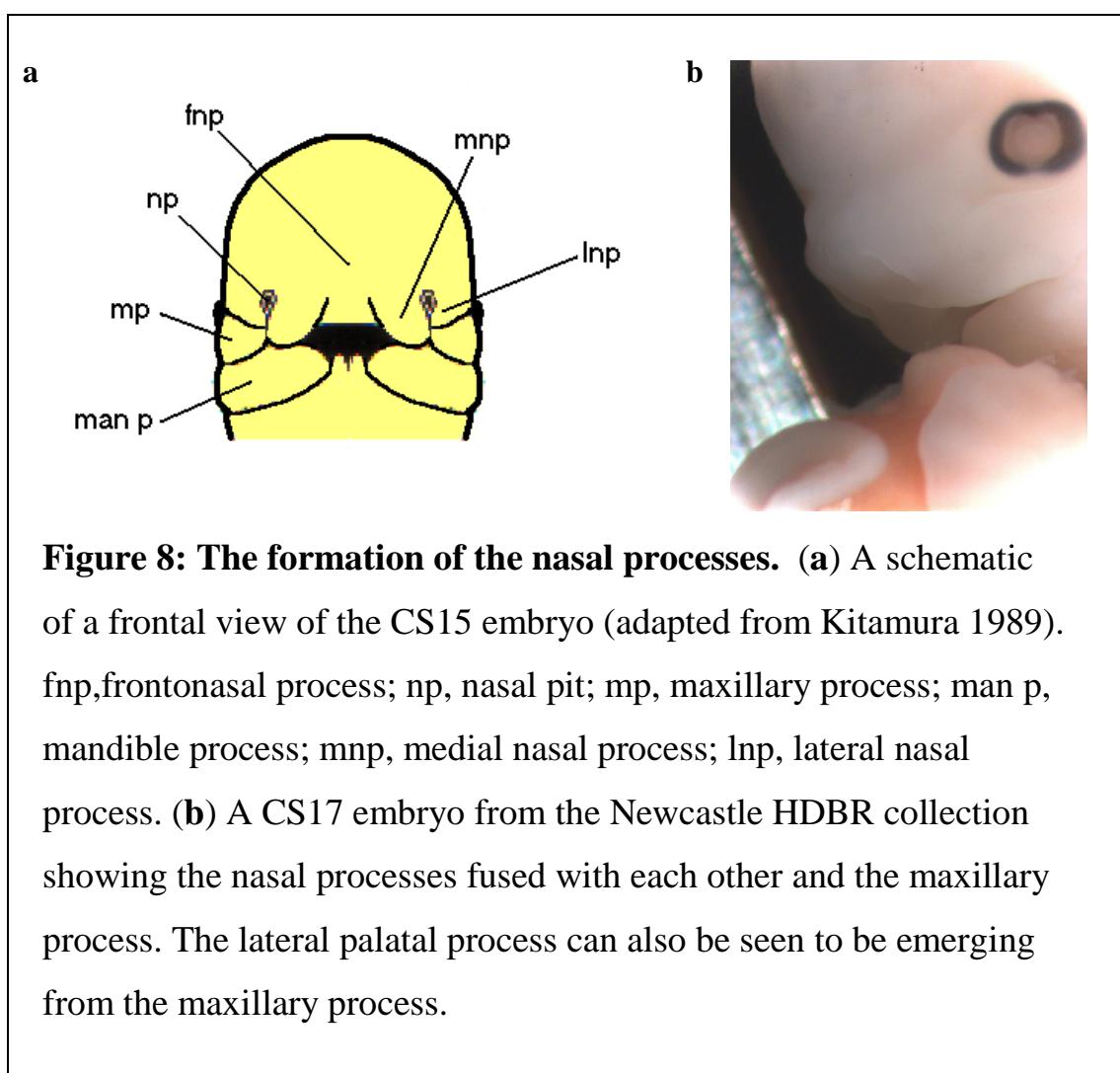
The first of the facial prominences to fuse are the mandibular processes (Fig. 7). Extending from the lateral swellings on the mandibular pharyngeal arch, the two mandibular processes expand in a horizontal direction eventually merging with each other in the midline of the embryo. This process is almost complete in the CS14 embryo “such that only a shallow depression remains central to a continuous structure on the rostral surface of the stomodeum” (Yoon *et al.* 2000). It is from this fused structure that the lower mandible, lip and other features associated with the lower jaw will develop. Proliferation in the mesenchyme of the frontonasal processes surrounding the nasal placodes (np in Fig. 6) produces a pair of swellings forming a horseshoe ridge around the placodes. These swellings are known as the medial nasal



process and lateral nasal process (mnp and lnp in Fig. 8a). The maxillary process is separated from the lateral nasal process by the nasolacrimal groove. The emergence of these nasal processes allows the nasal placodes to deepen forming a nasal pit.

The formation of the upper lip and maxilla are formed by the continued growth and fusion of the nasal and maxillary processes. The medial and lateral nasal processes continue to swell which cause the nasal pits to deepen further and to fuse, forming a single ectodermal nasal sac. A nasal fin is formed by proliferation of ectoderm in the floor and posterior wall of the nasal sac (Moore and Persaud 2003). Vacuoles develop within the nasal fin, which forms the oronasal membrane. The oronasal membrane is then ruptured, thus forming a continuous chamber between the primary nasal cavity and primitive oral cavity (Kitamura 1989).

As the nasal processes are enlarging, so too are the maxillary processes, which grow towards each other in a horizontal direction (mp in Fig. 8a), forcing the medial nasal processes towards each other and into the medial plane of the face.



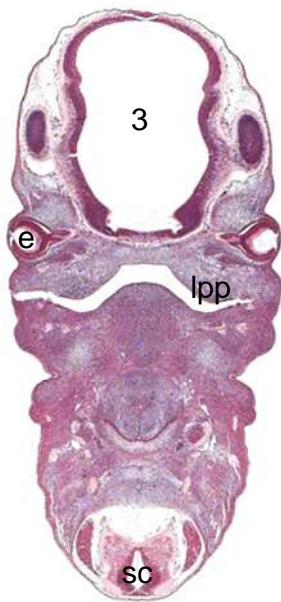
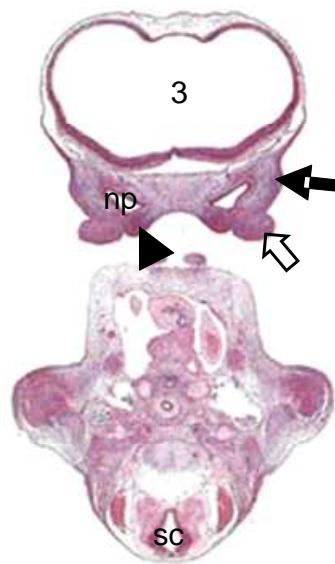
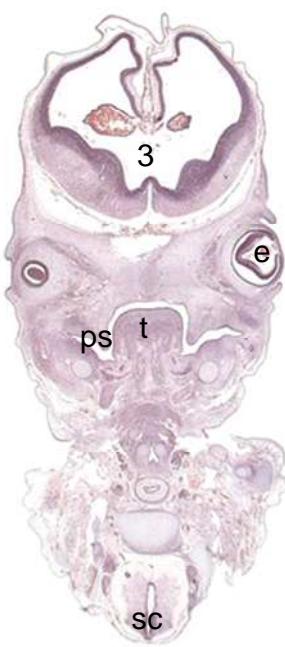
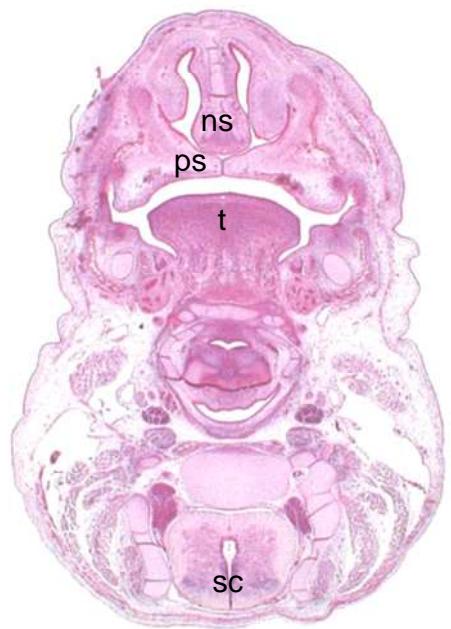
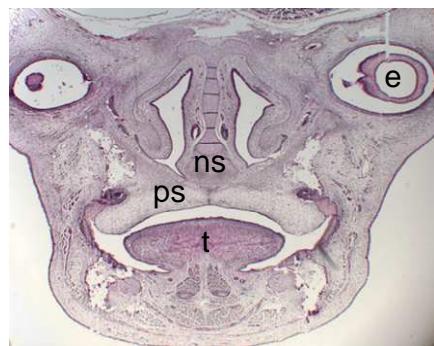
Following further enlargement and fusion of the facial processes, the morphology of the face begins to take shape (Fig. 9b). The medial and lateral nasal processes merge - closing the opening to the nasal sac, now known as the anterior nares. The medial nasal processes merge with each other and with the maxillary processes, making the groove between them indistinct in the CS19 embryo (Yoon et al. 2000). These combined fusions form a continuous structure above the mandible; namely the pre-maxilla, which separates the anterior nares from the stomodeum. The pre-maxilla will give rise to the primary palate and the incisors. The region in which the medial nasal processes and maxillary processes fuse is known as the intermaxillary segment and from this the indent in the middle of the upper lip, the philtrum, is formed. The maxillary processes also fuse with the lateral nasal processes along the nasolacrimal groove forming a continuation between the lateral side of the face and the nose (Fig. 8a). During the fusion of the facial processes, the epithelial cell covering of the adjoining processes must disintegrate to allow a structure composed of continuous mesenchyme cells to be formed (see 1.2.5).

### **1.2.3 Formation of the Secondary Palate**

At around 40 days of development from each of the maxillary processes emerge a further pair of swellings, the lateral palatal processes (Fig. 9a). These swellings proliferate, forming palatal shelves and grow vertically extending down into the oral cavity (Fig. 9c). The shelves then elevate above the tongue to a horizontal position and continue to extend until they come to abut each other and fuse (Luke 1976). The process of shelf elevation occurs within a few hours and there are several theories on the mechanism behind it, as reviewed by (Ferguson 1981; Shuler 1995), which can be categorised into two broad groups.

In the first, the palatal shelves are thought to elevate through an involuntary process mediated by some extrinsic factor. This extrinsic factor is thought to act through forces exerted on the shelves as a consequence of pressure differences caused by the growth of the lower jaw, movement of the tongue and the head rising from the chest, all of which are seen in the developing embryo (Diewert 1983). The second group postulates that the shelves have some kind of intrinsic action, which enables them to

elevate. This intrinsic force is likely to be chiefly generated by the progressive accumulation and hydration of glycosaminoglycans, mainly hyaluronic acid (or hyaluronate). Hyaluronic acid is a highly electrostatically charged molecule capable of binding up to ten times its own weight in water. It has been shown that hyaluronic acid (HA) makes up around 60% of the extra cellular matrix of the palatal *shelves in vitro* and *in vivo* (Pratt *et al.* 1973). The accumulation of hyaluronic acid is regionalised to the anterior region of the palate, where it causes a swelling of extracellular matrix and a decrease in mesenchyme cell density brought about by changes in osmotic pressure (Brinkley and Bookstein 1986). This regionalised swelling, due to the induced osmotic force, is constrained and directed by the epithelial covering of the palatal shelves (Brinkley 1984; Brinkley *et al.* 1992), and partly by the anchoring of fibronectin and collagen III molecules throughout the extracellular matrix of the vertical palatal shelf (Ferguson 1988). This has the consequential effect of causing a swinging flip-up mechanism in the anterior third of the shelves (Brinkley and Morris-Wiman 1987).

**a****b****c****d****e**

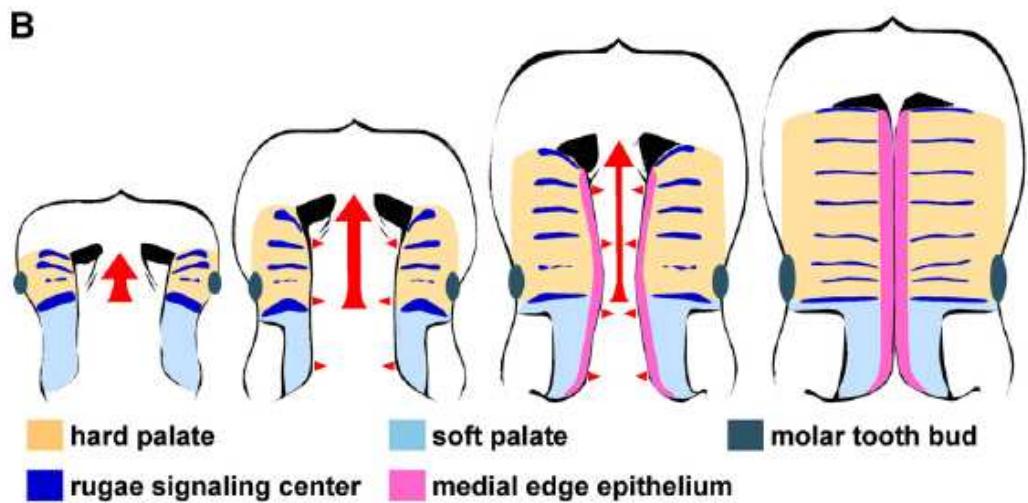
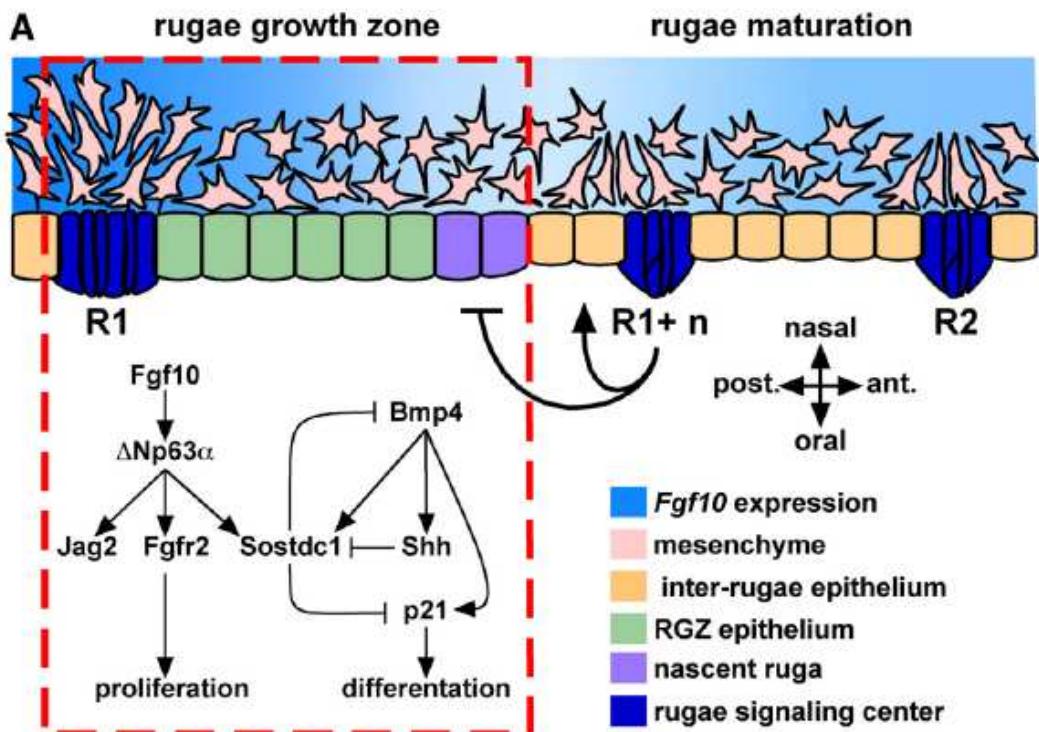
**Figure 9: The formation of the secondary palate and pre-maxilla.** (a) Transverse sections through a CS17 embryo at the level of the eye (e) and (b) through the nasal pit (np); sections through the developing secondary palate in a CS21 (c); CS23 (d) and 9 week foetus (e). The fusion of the lateral nasal process (filled arrow), medial nasal process (arrow outline) and maxillary process (arrow head) form the pre-maxilla (b). The secondary palate is formed by the elongation of the lateral palatal processes (lpp), (b), which form vertical palatal shelves (ps), (c), either side of the tongue (t) which then grow horizontally toward each other (d) and eventually fuse with each other and the nasal septum (e). 3, 3<sup>rd</sup> ventricle; sc, spinal cord.

Transforming Growth Factor Beta, TGF $\beta$ , (D'Angelo and Greene 1991); Transforming Growth Factor Alpha, TGF $\alpha$  and Epidermal Growth Factor, EGF, (Turley *et al.* 1985) have been shown to stimulate the synthesis of HA in the palatal shelves. (Yoshikawa *et al.* 1987) showed that a Vitamin A derivative; retinoic acid (RA), was capable of reducing the production of glycosaminoglycans in cultured fetal mouse palatal shelf cells. Later, (Degitz *et al.* 1998) showed in whole palatal shelf cultures a significant delay of the presence of HA in palatal shelves grown in the presence of RA compared to that of the controls. These RA treated palatal shelf cultures showed delayed palatal shelf elevation – although the mesenchyme appeared to have undergone hydration. The expansion of the mesenchyme within the shelves altered the curvature of the palate pushing downward towards the oral cavity. These shelves were substantially smaller than the controls and never fused. The authors postulated that: precise timing of palate elevation is required for correct fusion. Interestingly, RA has been shown to be crucial to the proper formation of the facial prominences (Song *et al.* 2004). By blocking RA synthesis in the nasal placodes, *Fgf8* failed to be correctly regulated in this region, resulting in a significant increase in programmed cell death in the lateral nasal processes.

After elevation of the palatal shelves, growth continues in a horizontal direction, until contact is established between the two palatal shelves and with the nasal septum: a structure which has arisen as a vertical protrusion into the oral cavity by growth of the merged medial nasal processes (Fig. 9e). The initial point of contact of the palatal shelves is the middle third region of the palate, with fusion spreading in both directions from here (Ferguson 1988). Fusion is completed along the entire length of the palate and with the primary palate in the anterior region of the mouth.

#### 1.2.4 Extension of the palate

Restricted expression of some genes to either the anterior – *Msx1*, *Bmp2*, *Bmp4*, *Fgf10*, *Shox2* (Zhang et al. 2002; Alappat et al. 2005) or posterior – *Meox2*, *Baxx1* (Peters et al. 1998; Jin and Ding 2006) palate leads to the possibility of an anterior/posterior axis within the palatal shelves, distinguishing the anterior hard palate from the posterior soft palate (Hilliard et al. 2005). The anterior/posterior gene expression boundary was later shown to be linked to the formation of palatal rugae (Hilliard et al. 2005). Palatal rugae are regions of epithelia thickening which form transversal ridges on the hard palate in mammals (Peterkova et al. 1987; Pantalacci et al. 2008). The number of rugae varies according to species and also within species – 3 or 4 in human, 7-10 in mouse depending on the strain and 18 in the horse (Pantalacci et al. 2008; Welsh and O'Brien 2009) and function to aid in mastication (Peterkova et al. 1987). The rugae of mouse and hamster have been shown to develop in a sequential manner within the anterior region of the palatal shelves. The first formed of which (R1) is positioned at the junction of the future hard and soft palate (Welsh and O'Brien 2009). Additional rugae are then formed in the region anterior to first formed rugae, with the next ruga being positioned between the first and the most recently formed rugae. The addition of the forming rugae is accompanied with expansion of the region between R1 and the most recently formed ruga, such that all the previously developed and the most recently formed rugae are extended anteriorly away from the anterior/posterior junction thus extending the future hard palate (Welsh and O'Brien 2009). A proposed molecular mechanism for the sequential addition of rugae within the oral epithelia are formed and how this relates to the accompanied extension of the anterior palate has been proposed (Welsh and O'Brien 2009) and is outlined in Fig. 10.



**Figure 10: Models of the molecular and morphogenetic activity associated with the rugae growth zone (RGZ) (taken from Welsh and O'Brien 2009).**

**(A)** A suggested model of molecular interactions integrating FGF10 and BMP4 signalling during rugae formation within the RGZ has been proposed (Welsh and O'Brien 2009). Both *Fgf10* and *Bmp4* are required for epithelial expression of *Shh*. *Fgf10* is expressed in a gradient extending from R1 to the site of nascent rugae formation. It is proposed that FGF10 signalling, mediated through  $\Delta Np63\alpha$  and its targets *Jag2* and *Fgfr2b*, maintains proliferation of epithelial progenitors at the posterior end of the RGZ. Epithelial expression of the *Bmp4* antagonist *Sostdc1* in the anterior palate also requires *Fgf10*. *Sostdc1* acts to restrict induction of *Shh* in the RGZ. Induction of *Shh* and *p21* expression at the site of the next forming rugae results in epithelial differentiation while continued proliferation of inter-rugae epithelium moves the R1+n rugae away from the RGZ. Signals from the R1+n rugae (bar and arrow) are also proposed to influence the fate of RGZ epithelium.

**(B)** The anterior growth of the anterior palate (ant) proceeds from the first formed rugae (red arrow) and is coincident with the establishment of segmental signalling domains (rugae). Fusion of the bilateral shelves requires medially directed growth (red arrowheads) and patterning of the medial edge epithelium (MEE, pink). The lateral boundary of the MEE coincides with the medial edge of the rugae, suggesting that signals from the rugae also participate in the intrinsic program that patterns the MEE. Thus, rugae and the RGZ provide a reference frame for visualizing the organization of signalling domains with respect to the anterior-posterior and medial-lateral patterning and growth of the palatal shelves.

### 1.2.5 Removal of the midline epithelia cells

The epithelia covering the palatal shelves are divided into oral, nasal and medial edge epithelia (MEE), depending on their regional position, and each cell type has a different fate: the nasal epithelia differentiate into pseudo-stratified epithelia; the oral epithelia into squamous epithelia; the epithelial cells covering the shelves in the region of fusion – namely, the medial edge epithelia – are removed (Chai and Maxson 2006).

Upon fusion, the opposing palatal shelves adhere to each other through a coat of glycoprotein molecules secreted by the MEE (Greene and Kochhar 1974) and by the formation of desmosomes (DeAngelis and Nalbandian 1968; Mogass *et al.* 2000), thus forming a midline epithelial seam (MES). It has been shown that the cell adherence properties of the MEE is specific and will not fuse with other epithelia under normal conditions (Ferguson *et al.* 1984).

The palatal shelves, similar to every other soft tissue in the oronasal region, are covered by two layers of epithelial cells: an inner layer of basal cells on the basement membrane; and an outer layer, consisting of periderm cells (Dudas *et al.* 2007). The periderm cells either peel out from the epithelial surface layer prior to contact of the palatal shelves, or else they become trapped by the two opposing shelves after fusion and undergo apoptosis (Fitchett and Hay 1989). The basal epithelia form the MES following fusion.

For complete unison of the shelves to occur and a continuous secondary palate of confluent mesenchyme cells to be formed, the MES must be removed. The process through which the MES is removed is still the subject of some debate and even the same technique has revealed contradictory results. The evidence for and against the principle mechanisms proposed for MEE removal is shown in Table 3.

Method Used	Observations and Interpretations	Suggestion or Evidence	References
CCFSE – intravital permanent epithelial staining (whole palates and living palatal slices.	No midline-crossing migration of green cells into the wild type tissue, suggesting that the MES cells die and/or migrate away from midline.	No epithelia-mesenchyme transformation.	(Cuervo and Covarrubias 2004)
	After fusion, labelled cells identified as fibroblasts in the mesenchyme	Epithelia-mesenchyme transformation.	(Griffith and Hay 1992; Sun <i>et al.</i> 1998; Kang and Svoboda 2002)
Tracing with adenovirus expressing lacZ.	No midline-crossing migration of labelled cells into the mesenchyme	No Epithelia-mesenchyme transformation.	(Cuervo and Covarrubias 2004)
Tracing with retrovis expressing lacZr.	Clusters of labelled cells into the mesenchyme	Epithelia-mesenchyme transformation.	(Martinez-Alvarez <i>et al.</i> 2000)
Lineage tracing with transgenic reporters.	No epithelial cells migrate into the mesenchyme. No labelled cells of the epithelial origin are present in the palatal midline after fusion.	No Epithelia-mesenchyme transformation.	(Vaziri Sani <i>et al.</i> 2005; Dudas <i>et al.</i> 2006)
GFP transgenic shelves fused with wild-type shelves.	No midline-crossing migration of green cells into the wild type tissue, suggesting that the MES cells die and/or migrate away from midline. Intraepithelial interactions (mixing) of green and wild type cells in the MES at the time of peak occurrence of apoptosis. Limited to shelf-to-shelf migration of oral and nasal epithelial cells recorded.	No epithelia-mesenchyme transformation, Epithelial motility.	(Cuervo and Covarrubias 2004)
Cell tracing after labelling with lipophilic fluorescent dyes DiO and/or Dil.	MEE cells migrate vertically from the midline towards the oral and nasal surfaces, but not into the mesenchyme or across the midline.	No epithelia-mesenchyme transformation migration.	(Carette and Ferguson 1992; Tudela <i>et al.</i> 2002; Cuervo and Covarrubias 2004)
	MEE cells migrate vertically from the midline towards the oral and nasal surfaces, but not into the mesenchyme or across the midline.	Epithelia-mesenchyme transformation.	(Shuler <i>et al.</i> 1991; Shuler <i>et al.</i> 1992)
Blocking cell migration by cytochalasin D.	Oral and nasal epithelial triangles not found, and cell death did not occur in any of MES cells	Periderm migration.	(Cuervo and Covarrubias 2004)

continued overleaf ...

Molecular detection of apoptosis,	Multiple apoptotic cells detected by TUNEL.	Cell death.	(Mori <i>et al.</i> 1994; Taniguchi <i>et al.</i> 1995; Martinez-Alvarez <i>et al.</i> 2000; Cuervo <i>et al.</i> 2002; Holtgrave and Stoltenburg-Didinger 2002; Martinez-Alvarez <i>et al.</i> 2004)
	No apoptotic cells detected by TUNEL.	Epithelia-mesenchyme transformation.	(Nawshad <i>et al.</i> 2004)
Blocking of cell death ( $\pm$ combined with epithelial cell labelling),	MES persists, no fusion occurs, the basal lamina stays intact, and no migration of labelled cells detected in the mesenchyme. This suggests either that the occurrence of EMT is low and insufficient to allow for palatal fusion, or that EMT does not occur prior to the death of the MES cells at all. Also suggests that basement membrane degradation is secondary to cell death (“cataptosis”) and not vice versa (“anoikis”).	Cell death.	(Cuervo <i>et al.</i> 2002; Cuervo and Covarrubias 2004)
	Complete fusion occurred	No cell death	(Honarpour <i>et al.</i> 2000; Takahara <i>et al.</i> 2004)
	Palatal fusion failed (after <i>Apaf1</i> gene inactivation in mice)	Cell death	(Cecconi <i>et al.</i> 1998)

continued overleaf ...

Blocking of degradation of the basement membrane (BM) by MMP inhibitor,	All cells disappear from the midline.	Cell death, no anoikis.	(Cuervo and Covarrubias 2004)
	Apoptosis occurs naturally, but the mesenchyme of apposed shelves remain separated by intact basement membranes without cells, suggesting basement membrane degradation is secondary to cell death (“cataptosis”) and not vice versa (“anoikis”).	Cell death, no anoikis.	(Cuervo and Covarrubias 2004)
	Cells did not disappear from the midline.	Cell death, anoikis.	(Blavier <i>et al.</i> 2001)
Removal of the periderm by trypsin,	No triangles formed, and resulting palate is thinner.	Periderm migration.	(Cuervo and Covarrubias 2004)
Electron and/or light microscopy observations,	No cell death or EMT seen, only epithelium migrating orally and nasally.	Migration.	(Carette and Ferguson 1992; Bittencourt and Bolognese 2000)
	No cell death seen, only epithelium changing to mesenchyme.	Epithelia-mesenchyme transformation.	(Fitchett and Hay 1989; Nawshad <i>et al.</i> 2004)
	Signs of cell death seen: dense bodies, autophagic vacuoles, lysomes.	Cell death.	(Pourtois 1966; DeAngelis and Nalbandian 1968; Farbman 1968; Hayward 1969; Smiley 1970; Holtgrave and Stoltenburg-Didinger 2002)
	Bulging cells observed, protruding from the MEE towards the opposite shelf.	Epithelial cell motility.(Martinez-Alvarez <i>et al.</i> 2000; Martinez-Alvarez <i>et al.</i> 2000)	(Martinez-Alvarez <i>et al.</i> 2000; Martinez-Alvarez <i>et al.</i> 2000)
	All three: EMT, death, and migration occur (statements based purely on static images). MEE cells can disappear from shelves before contact when amniotic fluid is not present in culture.	Epithelia-mesenchyme transformation, apoptosis, migration.	(Tsai and Verrusio 1977; Schupbach and Schroeder 1983; Takigawa and Shiota 2004)
Examination of cell differentiation markers.	MEE cells express cytokeratin and vimentin, and profusion midline mesenchyme just vimentin	Epithelia-mesenchyme transformation.	(Fitchett and Hay 1989; Shuler <i>et al.</i> 1991; Shuler <i>et al.</i> 1992)
Assessing metabolic health of midline cells.	Midline cells are alive and healthy.	No cell death	(Gartner <i>et al.</i> 1978; Gartner <i>et al.</i> 1978)

**Table 3: A summary of the evidence for and against the principle mechanisms involved with MES disappearance. (Taken from Dudas *et al.* 2007).**

### 1.2.6 Palate formation in other vertebrates

Much of the understanding of human palate development has been derived from work using animal models in which a wide range of experimental and genetic tools are available (e.g. Jheon and Schneider 2009) and experimental manipulation is possible that for obvious ethical reasons would not be possible in human subjects. Although the information obtained from experiments using animal models is often very revealing (reviewed in Gritli-Linde 2008), it should be noted that the formation of the secondary palate varies in birds, amphibians and reptiles to that of mammals, reviewed in Ferguson (1988).

The palatal shelves of birds arise from the maxillary processes, but instead of extending vertically they develop in a horizontal direction. The palatal shelves contact each other but the MEE never adhere. The MEE cells also do not produce glycoproteins or desmosomes, instead, these cells keratinise - leaving birds with a naturally occurring cleft palate (Koch and Smiley 1981).

In amphibians and some reptiles, the roof of the mouth consists largely of an extended primary palate; although other reptiles have a secondary palate similar to that found in birds, and the palate of the crocodilians is fused, similar to the mammalian secondary palate. The formation of a secondary palate seems to be an evolutionary feature exclusive to higher vertebrates and is absent in species lower on the evolutionary tree, including fish.

Although the development of the palate proceeds in mostly the same fashion in all mammals, there are differences between species. The floor of the nasal chamber in human, for example, recedes following the disruption of the oronasal membrane (see 1.2.2), to form a singular oro-nasal cavity; in non-primate mammals, including mouse, the floor of the nasal region consists of a “continuous cartilaginous flange” (Carstens 2002).

## 1.3 CLEFT PALATE

Stedman's medical dictionary describes a cleft palate as "a congenital fissure in the median line of the palate, often associated with cleft lip" (Stedman 2005). Any orofacial cleft, either of the palate or lip, is formed as a consequence of two embryonic structures that would normally fuse during development, not being able to properly do so. Improper fusion of any of the lateral nasal, medial nasal or maxillary processes (Fig. 8) may result in a child being born with a cleft lip or with a cleft of the primary palate, or in some instances both. As there are two pairs of these processes at either side of the developing face, such clefts may be present as either unilateral or bilateral. A cleft of the secondary palate is formed when there has been incomplete fusion of the palatal shelves.

### 1.3.1. Syndromic and Non-syndromic CL/P and CPI

A cleft lip, with or without a cleft palate, is conventionally notated as CL/P and isolated cleft palate CPI. A cleft of either the palate or the lip is referred to as "non-syndromic" in incidences where no other anomalies are present; or conversely, if an orofacial cleft is present together with other clinical features then the cleft is classed as "syndromic".

Several hundred syndromes are known to have a cleft lip or palate along with other abnormalities (OMIM; <http://www.ncbi.nlm.nih.gov/omim/>) and in many of these cases, a specific gene has been identified, for example: *IRF6* - Van der Woude and popliteal pterygium syndromes (Kondo *et al.* 2002); *MID1* – Opitz syndrome (Quaderi *et al.* 1997); *ESCO2* - Roberts syndrome (Vega *et al.* 2005) and many more have been shown to be caused by chromosomal abnormalities, such as DiGeorge syndrome which is caused by a deletion in chromosome 22 (Kelley *et al.* 1982). It is now thought that some of the genes responsible for syndromic clefts may also underlie non-syndromic clefts, either through incomplete penetrance or through the action of additional genetic or environmental factors (Stanier and Moore 2004). In addition to the syndromes mentioned above, *IRF6* has also been attributed to non-syndromic CL/P (Zuccheri *et al.* 2004; Blanton *et al.* 2005; Scapoli *et al.* 2005). Van

der Woude is the most common form of syndromic CL/P and is often seen together with a pit of the lower lip. Although people affected by Van der Woude syndrome may also have brain abnormalities, they are often undetectable without an MRI scan and in cases where lower lip pits are absent, Van der Woude is impossible to differentiate from non-syndromic CL/P.

### **1.3.2 The Aetiology of Cleft palate**

A cleft palate may arise as a consequence of incorrect development at any stage of palatogenesis: an anomaly in palatal shelf growth; delayed or failed shelf elevation; defective shelf fusion or a failure of medial edge epithelium cell death (reviewed in (Kerrigan *et al.* 2000; Murray and Schutte 2004; Rice 2005; Meng *et al.* 2009). These events are orchestrated by restricted gene expression and protein signalling, requiring precisely timed cell interactions and modifications. Disruptions to these processes lead to the failure of the correct fusion of one or more of the facial prominences, ultimately leading to facial anomalies, which in many cases display as a facial cleft.

A cleft of the secondary palate at birth can be caused by an anomaly at any time during the patterning of the pharyngeal arches, as well as incorrect growth or fusion of the palatal shelves, or incomplete disintegration of the MEE. Abnormal timing of head movements and tongue position in the mouth during embryogenesis have also been proposed as potentially being involved in forming a cleft palate (Ferguson 1981). The mammalian tongue occupies the space of the oral cavity to such an extent that the vertical palatal shelves cannot extend horizontally until the tongue has physically moved to enable sufficient space for the shelves to extend horizontally. It has therefore been suggested that incorrect movement of the tongue during development may in turn lead to insufficient space in the oral cavity for the horizontal extension of the palatal shelves, thus resulting in a cleft.

The actual aetiology of why normal development of the head and palate are disrupted and a cleft is formed appears to have both genetic and environmental contributing factors. The multifactorial threshold model of inheritance (Roberts 1961; Roberts 1964) used to model the inheritance of pyloric stenosis, has since been applied to the

inheritance of a cleft palate (Fraser 1976; Fraser 1996; Fraser 1998; Lin *et al.* 1999). The multifactorial threshold model of inheritance (MFT) postulates that the various genetic and environmental factors are interacting to provide a “continuous distribution of liability”. This liability is separated by a threshold value. In the case of cleft palate, Fraser explains that the liability can be thought of as the embryonic stages over which the palatal shelves are fusing. The threshold is the time point at which the shelves can no longer fuse with each other. Any factor that contributes to delaying the movement or fusion of the palatal shelves will increase the frequency of liability. The closer that an embryo is to the threshold, the more likely any delay in palate formation will result in the threshold stage being missed and therefore the susceptibility to a cleft will be increased. Although this model has been challenged recently by a Norwegian study (Sivertsen *et al.* 2008), the data produced from a much larger study (Grosen *et al.* 2009), concluded that their data supported a multifactorial threshold model of inheritance.

### **1.3.3 Cleft Palate Risk Factors**

#### **A. ENVIRONMENTAL FACTORS**

There have been many studies into the various environmental factors that contribute to the liability of a cleft palate. That is to say; non-genetic causes influencing the disruption, or delaying, the correct formation of the head, face and palate. Such environmental factors have been shown to include: parental age and maternal weight (Bille *et al.* 2005; Cedergren and Kallen 2005); geographic location (Poletta *et al.* 2007), which the authors ascribe to variance in altitude and socioeconomic status; seasonal variance (Krost and Schubert 2006), attributed to differences in exposure to ultra-violet light, fluctuations in diet and infectious disease cycles; exposure to various teratogens, including cigarette smoke, alcohol, multivitamins and anti-epilepsy drugs (Wyszynski and Beaty 1996; Shaw and Lammer 1999; Heilbronner 2005).

Various genetic risk factors have also been discovered (1.3.3.B), but it must be remembered that inheritance of cleft palate is multifactorial and as such, the effect of

environmental factors and genetic susceptibility to a cleft palate must be considered together. Despite this, gene-environment studies remain relatively few and difficult to interpret given the nature of these studies where so many unknown variables can influence findings. Indeed, whilst one study has shown the incidence of cleft palate to be higher in children born with a mutation in the cleft palate candidate gene *TGFα* if they had a maternal parent who smoked (Shaw *et al.* 1996), other studies have shown no such association between these two factors (Beatty *et al.* 1997; Christensen *et al.* 1999; Zeiger *et al.* 2005).

## B. GENETIC FACTORS

It was postulated relatively early that CL/P may have an hereditary component (Fogh-Anderson 1942). Reports have shown that gender is a determining factor; with cleft lip and palate being more common in males than females and cleft palate only, being more likely to be seen in females than males (Fraser 1970). In the most recent, and largest study to date, studied data from a Danish cohort of over 54000 relatives of individuals with an oral cleft (Grosen *et al.* 2009), the risk of oral cleft recurrence between first, second and third degree relatives was established. This study found that recurrence of a non-syndromic isolated cleft lip or palate between first degree relatives was 2.7%, with a relative risk – the ratio of the probability of a first degree relative of an individual with non-syndromic cleft lip being born with the same condition to that of the probability an individual being born with the condition in the background population – of 13. This effectively means that a first degree relative of an individual with isolated cleft palate would be 13 times more likely to be born with the same malformation. Relative risk for first degree relatives were established for non-syndromic cleft lip and palate (risk – 3.5%; relative risk 17) and non-syndromic isolated cleft palate (risk 3.1%, relative risk 15) (Grosen *et al.* 2009). Relative risks were also calculated for second and third degree relatives. Such information is invaluable in the counselling of families with these conditions (Gagnon *et al.* 2009), however with more complete understanding of the genetics involved in the inheritance of an oro-facial cleft, more accurate calculations of the risk of recurrence can be made e.g. depending on the specific genetic risk factor involved.

To date, there have been 12 distinct loci identified for non-syndromic CL/P designated orofacial cleft (OFC) 1-12 (Table 4) and in some of these cases a specific gene has been identified (Stanier and Moore 2004; Carinci *et al.* 2007). Of the six genes that have been confirmed as being causative agents in CL/P, four of those are also involved in syndromic CL/P diseases.

Non-syndromic loci	MIM number	Gene	Chromosomal Location
OFC1	119530		6p23-p24
OFC2	602966	<i>TGFα</i> ?	2p13
OFC3	600757	<i>BLC3</i> ? <i>/PVR?/PVRL2</i>	19q13
OFC4	608371	<i>SCD5</i> ?	4q21-q31
OFC5	608874	<i>MSX1</i> #	4p16.1
OFC6	608864	<i>IRF6</i> #	1q32-q41
OFC7	225060 *	<i>PVRL1</i> #	11q23.3
OFC8	129400 *	<i>TP63</i> #	3q27
OFC9	610361		13q33.1-q34
OFC10	601912	<i>SUMO1</i>	2q33
OFC11	600625	<i>BMP4</i>	14q22-q23
OFC12	612858		8q24.31

**Table 4: A list of known loci for CL/P** (Adapted from Carinci *et al.* 2007 and OMIM (<http://www.ncbi.nlm.nih.gov/omim/>) [28 Feb 2010]). A ? denotes that linkage studies have suggested a link between CL/P and that gene in the region, but has not been confirmed to date. # denotes that this gene has also been implicated in a syndromic CL/P condition, \* denotes that the MIM number actually refers to a syndrome in which an orofacial cleft is a feature of the phenotype, but the entry also includes reference to the highlighted non-syndromic OFC loci.

Of particular interest is the OFC10 locus, *small ubiquitin-like modifier 1 - SUMO1*. The SUMO1 protein is known to target specific proteins through a process known as sumoylation. This is a reversible action, capable of modifying the protein in question altering protein stability, facilitating transcriptional regulation and nuclear transport and apoptosis(Su and Li 2002; Meulmeester and Melchior 2008). There are 4 known *SUMO* genes in human, *SUMO1*, *SUMO2*, *SUMO3* (Su and Li 2002) and *SUMO4* (Bohren *et al.* 2004). A balanced translocation in the *SUMO1* gene was identified in individual with non-syndromic cleft lip and palate (Alkuraya *et al.* 2006). Further evidence of the importance of SUMO1 in palatogenesis has been provided by gene expression studies in mice, which demonstrated that *SUMO1* is expressed within the primary and secondary palate and in the developing lip (Alkuraya *et al.* 2006). Moreover, a number of the heterozygous *Sumo1*<sup>Gt/+</sup> mice displayed a cleft palate or oblique facial cleft and in *Sumo1*<sup>Gt/+</sup>; *Eyal*<sup>+-</sup> compound heterozygotes the occurrence of cleft palate (36%) was significantly increased (Alkuraya *et al.* 2006). However, a separate *Sumo1* gene trap mutant generated phenotypically normal mice, suggesting that SUMO proteins may compensate for one another (Evdokimov *et al.* 2008). SUMO1 has been shown to interact with other proteins implicated in palate formation, including SATB2, SMAD, MSX1, SOX9, EYA1, p53, p63 and TBX22 (Rodriguez *et al.* 1999; Dobreva *et al.* 2003; FitzPatrick *et al.* 2003; Lin *et al.* 2003; Huang *et al.* 2004; Taylor and Labonne 2005; Alkuraya *et al.* 2006; Gupta and Bei 2006; Andreou *et al.* 2007) and reviewed in (Pauws and Stanier 2007; Gritli-Linde 2008). This extensive number of proteins points towards a common genetic pathway regulating palatogenesis. Interestingly, sumoylation has been shown to be affected by various environmental factors including: heat shock, oxidative and osmotic stress and viral infection (Meulmeester and Melchior 2008), thus providing a link between the known genetic and environmental influences involved in the CL/P.

Although the known causative loci for non-syndromic genes is still relatively limited (Table 4), there are other genes that have been implicated in human non-syndromic oro-facial clefts on the basis of human gene linkage association studies, gene expression and phenotypic analysis gained from animal models and mutants (reviewed in Gritli-Linde 2008; Jugessur *et al.* 2009) and shown in Table 5.

Gene	References
<i>CHD7</i>	(Felix <i>et al.</i> 2006)
<i>CRISPLD2</i>	(Chiquet <i>et al.</i> 2007)
<i>ESR1</i>	(Osoegawa <i>et al.</i> 2008)
<i>ESCO2</i>	(Vega <i>et al.</i> 2005)
<i>FGF3</i>	(Riley <i>et al.</i> 2007a)
<i>FGF8</i>	(Riley <i>et al.</i> 2007a)
<i>FGF10</i>	(Riley <i>et al.</i> 2007a)
<i>FGF18</i>	(Riley <i>et al.</i> 2007a)
<i>FGFR1</i>	(Riley <i>et al.</i> 2007a)
<i>FGFR2</i>	(Riley <i>et al.</i> 2007a; Riley <i>et al.</i> 2007b; Osoegawa <i>et al.</i> 2008)
<i>FGFR3</i>	(Riley <i>et al.</i> 2007a)
<i>FOXE1</i>	(Vieira <i>et al.</i> 2005)
<i>GABRB3</i>	(Scapoli <i>et al.</i> 2002; Inoue <i>et al.</i> 2008)
<i>GAD1</i>	(Kanno <i>et al.</i> 2004)
<i>GLI2</i>	(Vieira <i>et al.</i> 2005)
<i>IRF6</i>	(Kondo <i>et al.</i> 2002)
<i>JAG2</i>	(Vieira <i>et al.</i> 2005)
<i>LHX8</i>	(Vieira <i>et al.</i> 2005)
<i>MSX1</i>	(van den Boogaard <i>et al.</i> 2000)
<i>MSX2</i>	(Vieira <i>et al.</i> 2005)
<i>MYH9</i>	(Martinelli <i>et al.</i> 2007)
<i>PTCH</i>	(Mansilla <i>et al.</i> 2006; Carter <i>et al.</i> 2010)
<i>PAX9</i>	(Ichikawa <i>et al.</i> 2006)
<i>PVR</i>	(Warrington <i>et al.</i> 2006)

Gene	References
<i>PDGFC</i>	(Ding <i>et al.</i> 2004)
<i>PVRL1</i>	(Suzuki <i>et al.</i> 2000)
<i>PVRL2</i>	(Warrington <i>et al.</i> 2006)
<i>RARA</i>	(Chenevix-Trench <i>et al.</i> 1993)
<i>RUNX2</i>	(Sull <i>et al.</i> 2008a)
<i>RYK</i>	Watanabe <i>et al.</i> , 2006
<i>SATB2</i>	(Brewer <i>et al.</i> 1999; FitzPatrick <i>et al.</i> 2003; Vieira <i>et al.</i> 2005)
<i>SKI</i>	(Vieira <i>et al.</i> 2005)
<i>SPRY2</i>	(Vieira <i>et al.</i> 2005)
<i>TBX1</i>	(Yagi <i>et al.</i> 2003; Paylor <i>et al.</i> 2006; Zweier <i>et al.</i> 2007)
<i>TBX10</i>	(Vieira <i>et al.</i> 2005)
<i>TBX22</i>	(Braybrook <i>et al.</i> 2002; Marcano <i>et al.</i> 2004)
<i>TCOF1</i>	(Sull <i>et al.</i> 2008b)
<i>TFAP2A</i>	(Milunsky <i>et al.</i> 2008)
<i>TGFA</i>	(Carter <i>et al.</i> 2010)
<i>TGFB1</i>	(Nawshad <i>et al.</i> 2004; Stoll <i>et al.</i> 2004)
<i>TGFB3</i>	(Lidral <i>et al.</i> 1998)
<i>TP63 (p63)</i>	(Celli <i>et al.</i> 1999; McGrath <i>et al.</i> 2001)
<i>WNT3A</i>	(Chiquet <i>et al.</i> 2008)
<i>WNT5A</i>	(Chiquet <i>et al.</i> 2008)
<i>WNT11</i>	(Chiquet <i>et al.</i> 2008)
<i>WNT9B</i>	(Juriloff <i>et al.</i> 2006)

**Table 5: Genes implicated in human syndromic orofacial clefting based on evidence from human genetic studies, mouse models, and expression data in orofacial primordial (from Gritli-Linde 2008; Jugessur *et al.* 2009).**

Although there are fewer loci genes identified for CPI, the first reported study suggesting that the genetic heredity of cleft palate may follow an X-linked pattern was described in 1966 (Weinstein and Cohen 1966); although cleft palate was not the only clinical feature present in this family. An X-linked inheritance for CPI was described in an British Colombia Indian family (Lowry 1970) and then a single causative gene for non-syndromic cleft palate was eventually mapped to a locus on the X chromosome in an Icelandic population (Moore *et al.* 1987; Bjornsson *et al.* 1989). In affected individuals in the Icelandic pedigrees presenting with problems of the secondary palate, the severity varied from a complete cleft of the secondary palate, through a fused but submucous cleft with an associated bifid uvula, to a high vaulted palate. In some cases there was no obvious palate cleft at all, but patients showed signs of ankyloglossia. Indeed, ankyloglossia was often seen in many cases in conjunction with the observed secondary palate malformations and it was therefore speculated that the cleft palate may in fact be a secondary consequence caused by inappropriate cell death in the tongue during palatogenesis (Gorski *et al.* 1992; Stanier *et al.* 1993). As the inheritance of this trait followed an X-linked pattern, the disease was given the notation CPX – X-linked cleft palate (MIM# 303400). Linkage analysis mapped the CPX locus to marker DXYS1X (Moore *et al.* 1991) on the long arm of the X chromosome in the region Xq21.1. This CPX locus was then further refined (Gorski *et al.* 1992; Stanier *et al.* 1993; Gorski *et al.* 1994) and finally *TBX22* was identified as the causative gene (Braybrook *et al.* 2001). Usually affected males are seen with cleft palate and ankyloglossia (CPA), although they do present CPI 17%, and rarely with ankyloglossia alone 4%. Female carriers vary from being fully affected, to displaying an entirely normal phenotype - CPA 11%, CPI 6%, ankyloglossia alone 43% and unaffected 40% (figures from Stanier and Moore 2004). Loss-of-function mutations have been found in coding regions of *TBX22* in CPX patients (Braybrook *et al.* 2002; Marcano *et al.* 2004; Suphapeetiporn *et al.* 2007). A haplotype containing risk single nucleotide polymorphisms within the promoter region of the *TBX22* gene and consequently reducing transcription of *TBX22*, has also been associated with cleft palate and ankyloglossia (Pauws *et al.* 2009b). The severity of the palate cleft seen in these CPX patients varies greatly, with the most severe cases presenting with a complete cleft of both the hard and soft palate, whilst others display a submucous cleft palate (SMCP), a condition where a layer of mucosa covers the roof of the mouth, but the underlying muscles do not join correctly leading to

velopharyngeal insufficiency. SMCP may also present with uvula bifida – a splitting of the uvula, or a v-shaped notch in the hard palate, a translucent line in the midline of the soft palate and a short palate (Wales *et al.* 2009). Cases where a SMCP has not given rise to velopharyngeal insufficiency are referred to as occult SMCP (Pauws *et al.* 2009a). CPX patients may also present with ankyloglossia together with any of these features, or indeed it may be the only clinical feature present.

There is also statistical evidence for a second CPI locus at 2q32 (Brewer *et al.* 1999). Two genes in this region have been suggested as possible candidates - *TGFα* and *SATB2*. Several studies have tried to link *TGFα* and *SATB2* to CPI with varying degrees of success (FitzPatrick *et al.* 2003; Vieira 2006; Leoyklang *et al.* 2007). It is likely that sampling biases lack of statistical power, and genuine population diversity have all contributed to the various studies leading to differing conclusions as to whether or not these genes are implicated in CPI.

As previously mentioned all of the syndromic cases of CL/P must be considered as potential candidates for non-syndromic CL/P too (Stanier and Moore 2004). Due to the multifactorial nature of CL/P, different genetic or environmental modifiers could affect the clinical presentation of a mutation in such a gene. As well as the known syndromic and non-syndromic CL/P causing genes, targeted gene disruption in animal models provide us with yet more possible CL/P candidate genes, as do gene expression profiles that reveal genes which are expressed in the facial prominences during development.

#### **1.3.4 Detection, Treatment and Outcomes**

In countries where routine antenatal ultrasound checks are common place, an oral cleft is often detected prior to birth. Even where antenatal ultrasound checks are not routine, a cleft lip being an external feature will be recognised immediately after birth. However a cleft palate, particularly a submucous cleft palate, in isolation is sometimes harder to detect and is sometimes not diagnosed until some weeks after birth. However, with the arrival of 3-dimensional ultrasound techniques, antenatal diagnosis is improving (Maarse *et al.* 2010). Suckling problems and faltering weight gain in children born with a CL/P are common (Beaumont 2008) and advice is often given by a specialised health practitioner on techniques to help mothers breastfeed and specialist aids can be offered to assist with this (Cole *et al.* 2009). In some cases, particularly with children born with a cleft of the secondary palate, children are unable to produce sufficient negative pressure in the oral cavity and therefore cannot move the bolus backward to the pharynx; rendering suckling impossible. In such cases the child will be fitted with a palatal lift obturator to aid nutritional intake prior to surgery (Karayazgan *et al.* 2009). In extreme cases feeding via a nasogastric tube is necessary (Oliver and Jones 1997).

After a period of around three to six months surgical repair of the cleft can be undertaken (De Mey *et al.* 2009). Follow up surgery is also sometimes required and a bone graft may be required in later childhood. A significantly higher proportion of children born with a facial cleft present with otitis media with effusion, or glue ear, than children who do not have a facial cleft (Flynn *et al.* 2009). This often leads to long term hearing difficulties, even following successful palate repair (Sharma and Nanda 2009) and may require separate surgical intervention and/or require the use of a hearing aid (Zambonato *et al.* 2009).

Whilst children born with a cleft lip usually develop speech normally, those born with a cleft palate are more likely to experience difficulties with speech development. Such speech development problems have also shown to affect their intelligibility, social competence and emotional development (Rullo *et al.* 2009), although the intervention of a speech therapist can reduce these effects.

Many cleft palate diseases are also associated with tooth abnormalities (van den Boogaard *et al.* 2000) and children often have to undergo orthodontic treatment to correct for teeth malformations. A higher proportion of tooth decay has also been observed in children born with a cleft lip or palate than those without (Al-Dajani 2009).

The emotional and physiological impairment to a child born with an oro-facial cleft is not to be underestimated. Their physical appearance and social embarrassment from not being understood contribute to reducing, to varying degrees, their quality of life into adulthood even following a complete treatment plan including surgery, speech therapy and audiology management (Mani *et al.* 2010). The reality exists that whilst children born in the Western World receive intervention from multidisciplinary teams offering both primary and secondary care (Austin *et al.* 2010), many of those children born with a cleft in underdeveloped countries will be denied access to any of this treatment due to the prohibitive costs of treatment in these countries. Left untreated these unfortunate children are often destined for a life of rejection due to social ignorance of their deformity and many never go to school. Without education and lacking social skills these children grow up never being unable to find employment (Mendoza 2009).

## 1.4 AIMS AND STRUCTURE OF THE THESIS

The overarching goal of the thesis is to increase our understanding of the role that *TBX22* plays during normal development and in CPX.

There are four specific aims of this thesis:

1. To investigate the spatial and temporal expression of *TBX22* during early human development, with a particular focus on the developing face and palate (Chapter 2). Potential functions of *TBX22* can be postulated from the timing of and tissue type where *TBX22* is expressed.
2. To synthesise a *TBX22* protein, in its native conformation, so that it can be used in functional studies to determine a preferential DNA binding site for *TBX22* (Chapter 3). Finding genes that contain this sequence is one approach to identifying potential downstream targets of *TBX22*.
3. To define potential *TBX22* downstream targets and screen their promoters *in silico* for the presence of possible *TBX22* binding sites (Chapter 4). Identifying potential downstream targets will suggest pathways in which *TBX22* may be involved and hence possible mechanisms for its action(s).
4. To investigate the possibility that one of the genes identified by the *in silico* search is a *TBX22* downstream target *in vivo* (Chapter 5). Characterising the interaction of *TBX22* and one of the potential downstream targets will validate the results gained *in silico* and provide stronger evidence for one possible role for *TBX22*.

## CHAPTER 2

### EXPRESSION OF TBX22 DURING EARLY HUMAN DEVELOPMENT

#### 2.1 INTRODUCTION

Since mutations in the *TBX22* gene had been shown to cause CPX ((Braybrook *et al.* 2001); reviewed in Chapter 1) it was likely that *TBX22* played a significant role in the normal development of the palate. Elucidating its expression pattern would be an important first step towards understanding exactly when and where *TBX22* could be exerting its effect(s). It might distinguish, for example, between the possibilities that the cleft palate seen in CPX arose as a secondary consequence of the incorrect separation of the tongue from the floor of the mouth during palatogenesis (leading to a decrease in the amount of space available in the oral cavity to allow correct fusion of the shelves) or alternatively, if the cleft might be due to a failure of the growth or fusion of the palatal shelves themselves (see 1.3.2).

*TBX22* expression was examined at the RNA level by tissue *in situ* hybridisation. Studying the RNA, rather than the protein, has the advantage that primers or probes can be easily synthesised to be specific for a particular region of the mRNA. Although when working with RNA, the additional technical issues of eliminating RNase contamination become important. There are various methods to study the RNA expression profile of a gene of interest. RT-PCR amplifies cDNA which has been generated from RNA within the sample tissue using the reverse transcriptase enzyme with specific primers ensuring detection of the gene of interest. However, this approach only gives a “yes or no” answer as to whether an mRNA is present in the experimental sample or not. Any indication of the tissue specificity of expression is dependent on the degree to which the tissue sample can be sub-divided by dissection but there is no indication of cellular distribution of the RNA within the dissected tissue. As detection is performed by amplifications of the RNA within tissues, some low level transcripts or even cross-contamination from nearby tissues can produce

unexpected results. Early RT-PCR studies of the full length *TBX22* transcript indicated that *TBX22* was expressed in all of the fetal tissues examined (Braybrook *et al.* 2001), perhaps indicating that *TBX22* did not have a tissue specific expression. Real time quantitative PCR (Chiang *et al.* 1996; Gibson *et al.* 1996; Heid *et al.* 1996), whilst still being limited to the same sample size restraints as traditional RT-PCR has, as the name suggests, the additional benefit of measuring the abundance of an RNA transcript within a particular sample allowing the level of expression in one tissue to be compared to that of another.

Insights into whether there are different RNA transcripts within a tissue may be gained by Northern blot (Alwine *et al.* 1977). This entails blotting a labelled probe to a membrane blot of an RNA gel of a tissue sample. Information about the spatial distribution is again limited by the range of tissues tested and the specificity of the tissue sample dissections. This was highlighted when the *TBX22* gene was first cloned. As tissue from the palate was not included in the test tissues, no mRNA transcript was detected in any of the human adult or fetal tissues tested in a northern blot hybridisation (Laugier-Anfossi and Villard 2000). However, this technique does give insights into the size and number of RNA transcripts (Law *et al.* 1998).

Originally using RNA or DNA probes to detect specific DNA molecules within a tissue sample, tissue *in situ* hybridisation, first described in 1969 (Gall and Pardue 1969; Pardue and Gall 1969) enabled detection directly within a histological sample. There have been many refinements to the technique since 1969, including the use of RNA probes to detect RNA molecules within tissue sections and as whole-mount preparations (Hargrave *et al.* 2006). By directly hybridising probes to tissue preparations of various ages, the temporal and spatial expression patterns can be visualised in the host tissue. Hence, this technique was chosen to characterise the temporospatial expression pattern of *TBX22* during facial development.

It was important that this study be undertaken directly in human. It has been shown that although a gene's coding sequence may be highly conserved during evolution, this does not necessarily indicate that its expression pattern will also be conserved (Fougerousse *et al.* 2000). It has also been noted that there are differences between gene expression patterns in human and other vertebrates. For example, *Tbx5*

expression has been shown to be heart chamber specific in chick and mouse (Bruneau *et al.* 1999) although this specificity has not been reported in humans (Li *et al.* 1997). Despite potential differences in expression patterns between, for example, mouse and human, the expression profiles of many genes have been studied using other species as a model for human development, partly due to the relative ease with which such tissue is available to researchers and to the variety and extent of experimental studies that can be performed. To facilitate cross-species comparisons, studies have defined equivalent developmental stages in different species (summarised in Table 6 and see <http://embryology.med.unsw.edu.au/>) although these are estimates based on the development of a number of organ systems and specific organs and tissues may develop at differential rates across species (Clancy *et al.* 2001; Clancy *et al.* 2007) including the formation of the face and secondary palate (see 1.2).

Much valuable information has been obtained by studying mouse gene expression patterns (Bush *et al.* 2002; Herr *et al.* 2003; Kim *et al.* 2009b; Pauws *et al.* 2009a). However, in some aspects mouse is not an ideal model of human oral development. For example, although displaying a fused palate, tooth development differs and they only have one set of teeth. The formation of rugae, which have been proposed as a signalling centre determining anterior and posterior palate (Pantalacci *et al.* 2008; Welsh and O'Brien 2009), occurs during palatal shelf growth in mice but are only reported in humans after fusion has occurred (Patil *et al.* 2008).

Animal models, especially mouse, have provided many useful tools for the understanding of human disease, for example by knocking out, reducing or altering the expression of specific genes. However, the resultant phenotypes do not always match the human disorder they were designed to mimic. Indeed, the *Tbx22* mouse null mutant displayed differences to the human CPX phenotype it models (Pauws *et al.* 2009a). A major phenotypic feature of this mutant was choanal atresia, which has not been described in the CPX phenotype. Also, ankyloglossia was found in all of the *Tbx22* mutants examined. Ankyloglossia is variably penetrant and often not seen at all in CPX. Therefore comparing the expression pattern of human to that of mouse will help to assess how much reliance can be placed upon a mouse model of CPX.

*TBX22* expression was investigated during the time when the face is forming, from ~33 days – 9 weeks (see 1.2). Particular attention was paid to the formation of the palatal shelves; as developmental delays in the formation of the shelves are thought to determine whether a cleft of the secondary palate, as seen in CPX patients, will occur (see 1.2.3).

Human		Mouse		Chick	
Carnegie Stage	Age (dpo)	Theiler Stage	Age (dpc)	Hamburger-Hamilton (HH) Stages	Age (dpc)
9	20	12	8	7-8	1
10	22	13	8.5	9-10	1.5
11	24	14	9	11-13	2
12	28	15	9.5	14-17	2.25
13	30	16	10	18-21	2.5
14	33	17	10.5	21-22	3
15	36	18	11	23-24	3.25
16	40	19	11.5	24-25	3.75
17	42	20	12	26	4.75
18	44	21	13	27-28	5.5
19	48	22	14	29	6.25
20	52	23	15	30	7.25
21	54	24	16	31-32	7.75
22	55	25	17	33-38	8.5
23	58	26	18	39-44	10

**Table 6: A cross-species comparison of developmental timings** (Butler and Juurlink 1987; O'Rahilly and Muller 1987; Hamburger and Hamilton 1992)

## 2.2 MATERIALS AND METHODS

Tissue *in situ* hybridisation (TISH) using  $^{35}\text{S}$ -UTP labelled riboprobes was employed to determine the temporal and spatial mRNA expression pattern of *TBX22* directly in human embryonic tissue. These expression studies were undertaken using material made available from the MRC-Wellcome Trust Human Developmental Biology Resource ([www.HDBR.org](http://www.HDBR.org)).

### 2.2.1 Embryo collection and Processing

Human embryonic and fetal tissues were collected by the MRC-Wellcome Trust Human Developmental Biology Resource (HDBR) in Newcastle (Lindsay and Copp 2005), with ethical approval by the Newcastle and North Tyneside Local Research Ethics Committee and appropriate maternal consent. Following either surgical or medically induced termination of pregnancy (Bullen *et al.* 1998), embryos were staged using the Carnegie stage classification; using the definitions in Table 7, which is taken from the staging guidelines in (Bullen and Wilson 1997). These staging guidelines are a revision of the original definitions of (O'Rahilly and Muller 1987) and are based on external characteristics of single embryos, rather than a comparison of several embryos at the same Carnegie stage. This staging system describes specimens prior to fixation, a process that often leads to artificial changes in morphology and is especially useful as no dissection is necessary and minimal manipulation of the embryo is needed in order to stage it, which is obviously desirable given the delicate nature of the unfixed embryo.

The fetal human samples (those older than CS23) that were collected were staged using the guidelines set out in Table 8. The fetal staging guide uses foot and knee to heel lengths to estimate developmental age and is adapted from (Hern 1984).

Where possible, all embryos are subjected to karyotype analysis using a small amount of placental tissue or, in cases where there is insufficient tissue, a skin sample from the embryo itself. Only embryos that have a normal chromosomal complement and arrangement are used in this study and two embryos are used at each stage to verify the findings.

The embryos were fixed in 4% PFA/PBS for 24 hours and then transferred to 70% ethanol. After successive changes through a graded series of 80%, 90% and 100% ethanol (AnalaR, BDH) the embryos were then embedded in paraffin wax (Fibrowax, BDH). Serial sections were taken using a standard microtome at 5 $\mu$ m intervals and mounted onto Marienfeld HistoBond slides. To aid with orientation through the embryo, every ninth section was stained with haematoxylin and eosin, following the standard procedures for paraffin sections (Wilson and Gamble 2002).

Carnegie Stage	External Characteristics – “diagnostic”	Secondary features	Crown – Rump length (mm)
10	Mainly open neural groove; up to 12 somites		1.5 – 3
11	Anterior neuropore closing, posterior open		2.5 – 4.5
12	No limb buds, or minimal upper bud	Posterior neuropore may be open	3 – 5
13	4 limb buds, tiny lowers		5 – 7
14	Tapering upper limb, no hand plate	Open lens pit	7 – 9
15	Hand plate now	Lens pit closed	11 - 14
16	Foot plate and retinal pigment	Pharyngeal arch 3 receding. Auricular hillocks	11 - 14
17	Digital rays of hand; full round foot plate	Full auricular hillocks now	11 - 14
18	Notched hand +/- elbow	Eyelid folds in later specimens	13 - 17
19	Prominent toe rays	Limbs all extended nearly directly forward	17 - 20
20	Stubby fingers; toe notches; vascular plexus of head	Elbows bent. Hands still well apart	21 - 23
21	Fingers longer, getting closer. Toes not just notches, may touch	“Tactile pads” on ends of fingers	22 - 24
22	Fingers touch/overlap. Vascular plexus $\frac{3}{4}$ up head	Ear features more prominent	25 - 27
23	Vascular plexus almost at vertex. Limbs generally more mature.	Eyelids may be starting to fuse at margins	28 – 30

**Table 7: Embryo staging guide** (from Bullen and Wilson, 1997)

Fetal Age (completed menstrual weeks)	Developmental Age (weeks post fertilisation)	Foot length (mm)	Knee – heel length (mm)
10	8 (CS23)	5 - 6	8
11	9 (F1)	7	11
12	10 (F2)	8 - 9	13
13	11 (F3)	10 – 12	17
14	12 (F4)	13 – 16	24
15	13 (F5)	17 – 19	31
16	14 F6)	20 – 22	36
17	15 (F7)	23 – 24	40
18	16 (F8)	25 – 27	43

**Table 8: Fetal staging guide** (from Hern, 1984)

### 2.2.2 Tissue in Situ Hybridisation

TISH was performed using radio labelled RNA probes and visualised by coating the slides with a photo-sensitive emulsion. The method was devised by Moorman and his colleagues (Moorman *et al.* 1993) and is described below. For this procedure all glassware was pre baked at 180°C for 4 hours to denature any potential contaminating RNase enzymes.

#### A. DNA TEMPLATE PREPARATION

A 336-bp fragment (nucleotides 640-976 of Genbank accession no. NM\_016954) - which includes regions of exon 6 and 7 outside of the T-box domain to increase specificity - was cloned into a pGEM-T vector between the T3 and T7 RNA

polymerase promoter sites. This plasmid was a kind gift from Dr. Laurent Villard and was used as a template to synthesise RNA probes. 1µg of plasmid DNA were digested to completion with either SalI restriction enzyme (Promega) to generate the sense probe, or ApaI (Promega) to generate the antisense probe. The sequences of both the sense and antisense probes were determined (3130XL, Applied Biosystems) and then verified using the BLAST alignment program (Altschul *et al.*, 1990), housed at the National Centre for Biological Information (NCBI (<http://www.ncbi.nlm.nih.gov/>)) to ensure that they were specific to *TBX22* mRNA and that there was minimal similarity to other sequences. As expected 100% similarity was seen between the probe template sequence and all three *TBX22* mRNA transcripts, with 94% similarity between the probe and the *TBX22* genomic sequence. The highest similarity of the probe to a sequence other than *TBX22* was to a 61 nucleotide region of *TBX1* (NM\_005992 797-857). As *TBX22* shares the highest sequence homology to *TBX15* and *TBX18* (see 1.1), the probe sequence was blasted directly against the mRNA sequences of these genes (NM\_152380 and NM\_001080508). This resulted in a 74% match covering 250 nucleotides for the *TBX15* mRNA and 68% similarity over a 231 nucleotide region for *TBX18*. These results meant that it was very unlikely that either sense or antisense probe would cross-hybridise to other RNA targets.

The DNA was purified by 1:1 phenol/chloroform extraction, precipitated with 96% ethanol and 0.3M NaAc, washed with 70% ethanol and resuspended in 10µl TE buffer (10mM Tris, 1mM EDTA, pH 7.6). The linearised plasmid DNA was verified by visualisation on 1% agarose check gel, and the size of the linearised DNA was confirmed by comparison to a 1kb ladder (Promega). DNA concentration was estimated by comparing the intensity of the electrophoresed DNA band to a known amount of  $\lambda$  phage DNA pre-digested to completion with a HindIII restriction enzyme (Promega).

#### **B. RNA TRANSCRIPTION AND INCORPORATION OF RADIOACTIVE $^{35}$ S-UTP**

The linearised template DNA was used to synthesise sense and antisense RNA probes via transcription with either T3 or T7 RNA polymerase respectively. 500ng of linearised template DNA was added to a transcription reaction containing a final

concentration of 500 $\mu$ M ATP (Roche), 500 $\mu$ M CTP(Roche), 500 $\mu$ M GTP(Roche), 500 $\mu$ M (50  $\mu$ Ci)  $^{35}$ S-UTP (Amersham Biosciences), 1X RNA transcription buffer (Roche), 1U/ $\mu$ l RNAsin (RNAase inhibitor, Sigma), 50 $\mu$ M DTT, 1U specific RNA polymerase (Roche) and double distilled DEPC treated H<sub>2</sub>O in a 10 $\mu$ l volume. The reaction was incubated at 37°C for 2 hours, and then the DNA template was digested by the addition of 1U DNase (Roche) and incubated at 37°C for a further 15 minutes. Unincorporated NTPs and the template DNA fragments were removed by adjusting the volume to 100 $\mu$ l, adding the equivalent volume of 1:1 ratio phenol (Sigma)/chloroform (BDH), mixing and following centrifugation at 12000g for 5 minutes in a microcentrifuge, applying the aqueous phase to a NICK column (Amersham). The probe was eluted from the column by the addition of TE buffer and further purified by ethanol precipitation (see 2.2.2.A) with the addition of 1 $\mu$ l Yeast tRNA (Sigma) as a carrier, and finally resuspended in 10 $\mu$ l TE/10mM DDT.

#### C. QUANTIFICATION

The amount of RNA probe used in each experiment was adjusted to a standard working concentration, based on the incorporation of  $^{35}$ S-UTP. This was achieved by measuring 0.5 $\mu$ l aliquots of the probe in 1ml of scintillation fluid and counting the isotopic activity using a liquid scintillation counter (Camberra-Parkard, Tricarb 2100TR) and adjusting the volume of TE buffer until the probe reached a concentration of 10<sup>6</sup> cpm/ $\mu$ l.

#### D. PROBE HYBRIDISATION

DEPC treated Milli-Q water was added to the required amount of probe estimated to give 5x10<sup>4</sup> cpm/ $\mu$ l and the RNA probe denatured by heating for 5 minutes at 96°C. The denatured probe was added to a hybridisation mix of final concentration; 50% deionised formamide, 10% dextran sulphate, 2X SSC, 2X Denhardt's solution (Sigma) , 0.1% Triton X-100 (Sigma), 10mM DTT (Sigma) and 2000ng/ml salmon sperm DNA (Sigma).

Paraffin was removed from the tissue sections by immersing the slides in 3 changes of xylene for 5 minutes and rinsing for 3 minutes in 1:1 xylene/ethanol and twice in absolute ethanol. The slides were dried for 1 hour before being soaked in 2N HCl for 20 minutes, rinsed in water and washed in 2X SSC at 70°C for 10 minutes. The tissue was subjected to proteolytic digestion in a 1% pepsin solution at 37°C for varying times dependent upon the age of the tissue: CS16, 3 minutes; CS17, 3.5 minutes; CS19, 4 minutes; CS20, 4.5 minutes; CS23, 6 minutes and 9 week fetal, 6.5 minutes. This was necessary to permeate the cells to increase the accessibility of the target mRNA. The pepsin was neutralised by rinsing in a 0.2% glycine/PBS solution. The sections were then fixed in 4% PFA/PBS for 20 minutes, to maintain the localisation of the nucleic acids, before being air-dried. 6µl of the probe/hybridisation mix was placed onto each section and hybridised in a humidified chamber at 52°C overnight.

#### E. POST HYBRIDISATION WASHES

Following hybridisation, the sections were rinsed with 5X SSC and washed twice with 50% formamide/1X SSC 10mM DTT and once with 1X SSC 10mM DTT for 15 minutes at 52°C. Slides were next washed in RNase buffer at 37°C then transferred to fresh 1X RNase buffer containing 1U RNase for 30 minutes at 37°C. Finally the slides were washed with 1X SSC, then 0.1X SSC for 15 minutes each at 52°C. The slides were then dried in a filtered airstream for 1 hour.

#### F. VISUALISATION

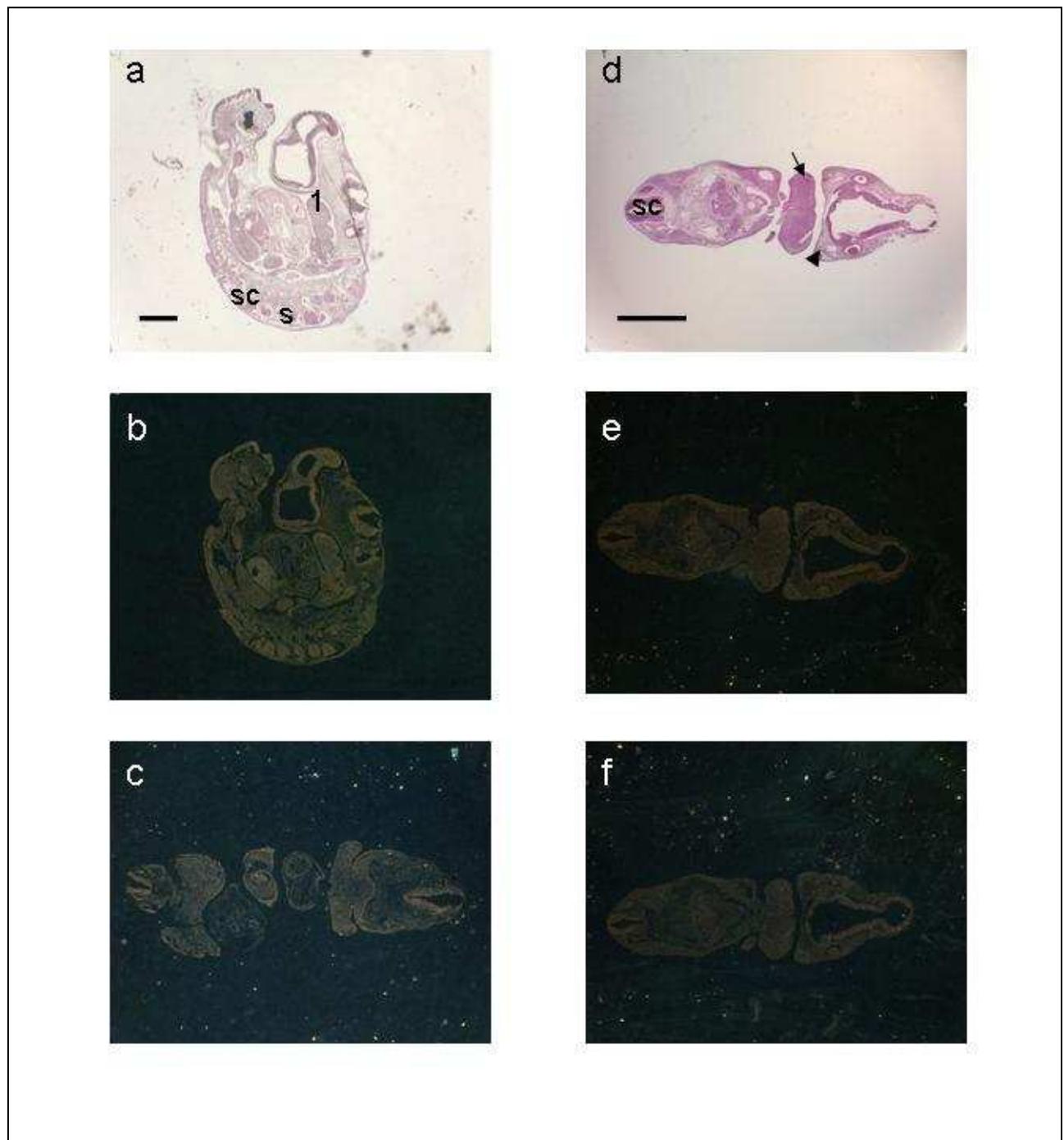
The dry slides were coated in Ilford K5 photographic emulsion under dark room conditions. The coated slides were then exposed to the emulsion for ten days. The emulsion was developed using Kodak D-19 developer and fixed in 30%  $\text{Na}_2\text{S}_2\text{O}_3 \cdot 5\text{H}_2\text{O}$ . The sections were then counter stained with 0.1 % nuclear fast red kernechtrot (R. A. Lamb) with 5% aluminium sulphate, dehydrated in ethanol, cleared in xylene and mounted in mananol (Klinipath) and covered with a glass cover slip.

#### G. IMAGING

The glass mounted slides were examined using a Zeiss Axioplan microscope using a dark field viewing stage and illumination (Leica). Whole embryos were viewed using a Zeiss stereo-microscope MC80. On both microscopes digital images were recorded using an Axicocam camera and Axiovision (version 3.2) software package (Imaging Associates).

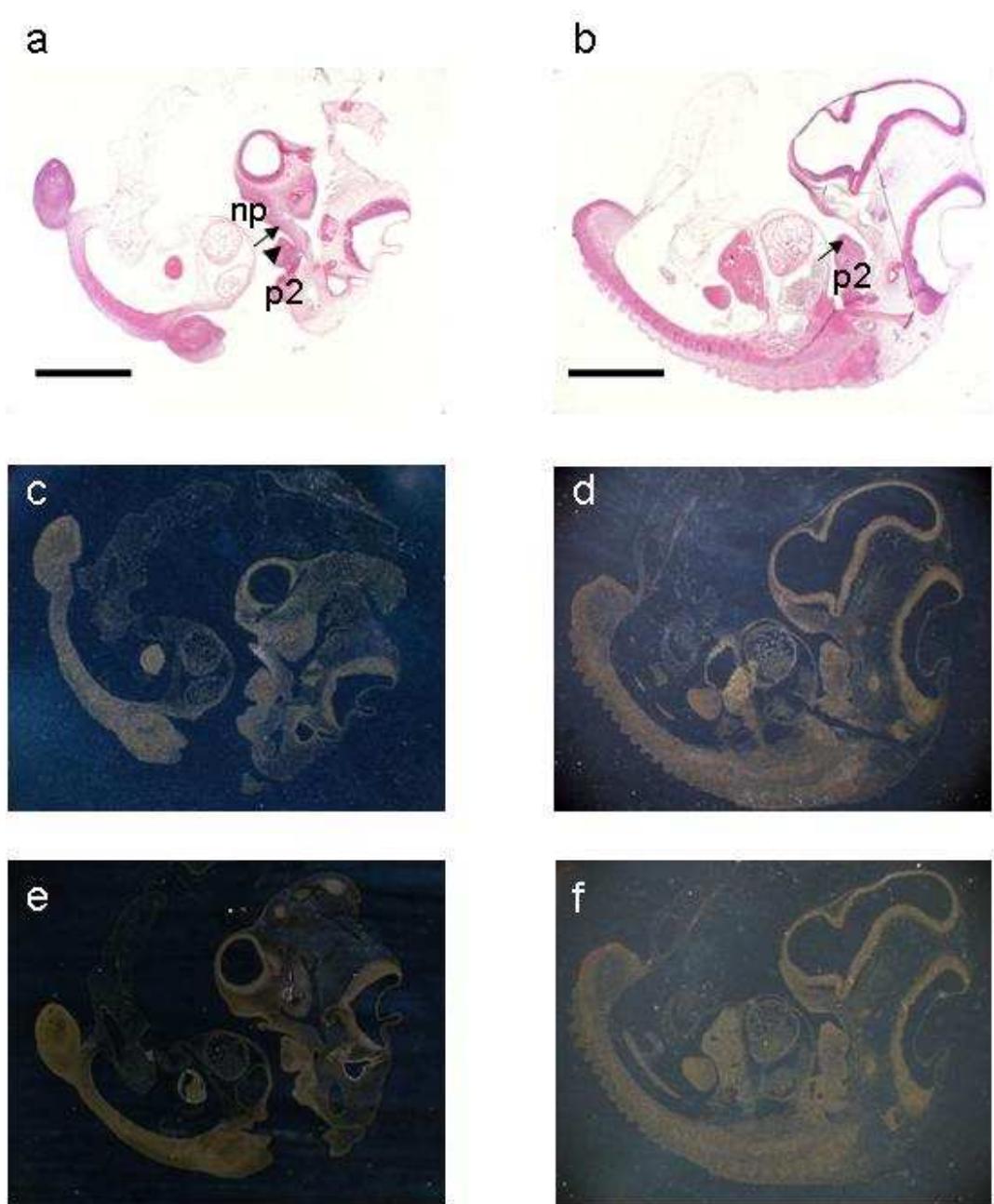
## 2.3 RESULTS

The sense probe did not show any signal in any of the *in situ* hybridisation experiments (Fig. 11f and Fig. 12e and f and data not shown). During the early formation of the jaw and palate, at CS14 and CS15, *TBX22* was not detected, or was expressed at levels below those which could be detected by *in situ* hybridisation (Fig. 11).



**Figure 11: *TBX22* is not detected at CS14 or CS15.** A CS14 embryo, sagittal section (**a** and **b**) and transverse section (**c**); a CS15 embryo sectioned transversely (**d – f**). **a** and **d** are bright field images of haematoxylin and eosin stained sections, **b**, **c** and **e** are dark-field images of antisense *TBX22* probe **f** is a dark-field images of sense probe. Arrow, mandibular process; arrow head, maxillary process; 1, first pharyngeal arch; s, somites; sc, spinal cord. Bar = 100 $\mu$ m.

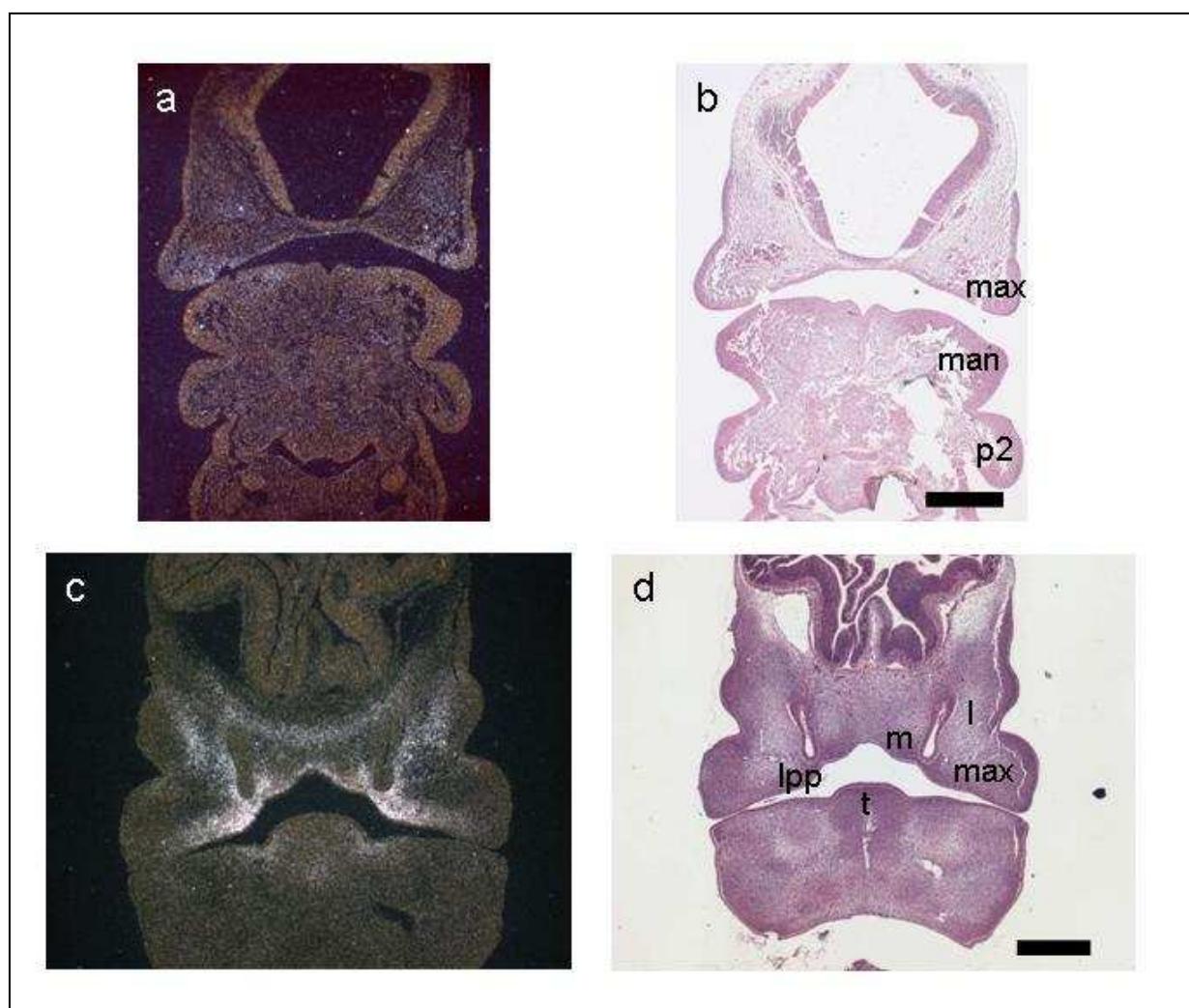
It was not until CS16 that the first definitive expression of *TBX22* during the formation of the face was detected (Fig. 12). The expression in the CS16 embryo was detected most strongly in the maxillary process (the arrow in Fig. 12) and appeared stronger in lateral (Fig. 12c) than medial sections (Fig. 12d), although weaker expression was seen surrounding the nasal pit in both medial and lateral sections (Fig. 12c and d).



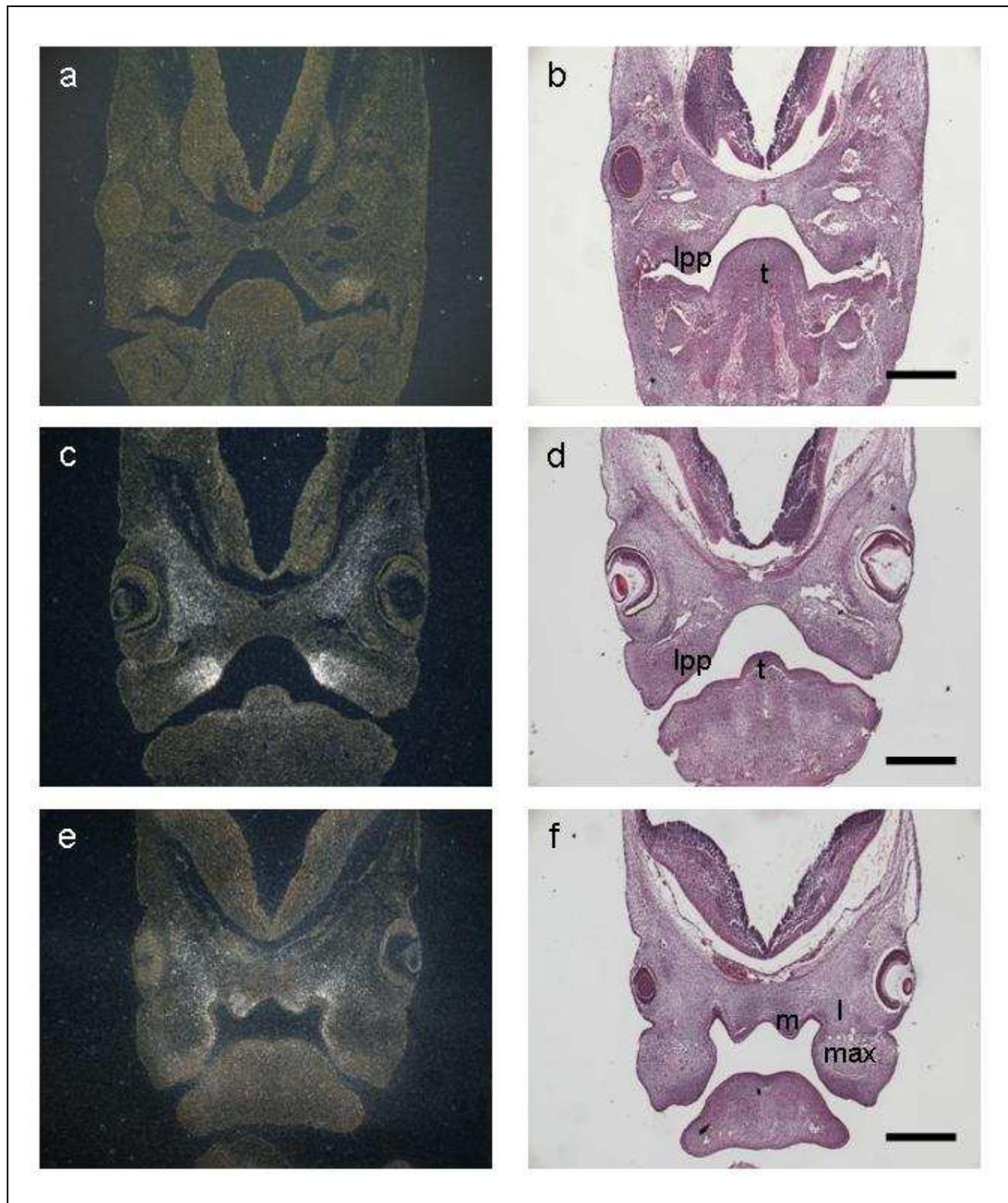
**Figure 12: *TBX22* is expressed in the nasal and maxillary processes of a CS16 embryo.** Transverse sections of a CS16 embryo, lateral (a, c and e) and midline sections (b, d and f). a and b are bright field images of haematoxylin and eosin stained sections, c and d are dark-field images of antisense *TBX22* probe and e and f are dark-field images of sense probe. Bar = 100 $\mu$ m. np, nasal pit; arrow, maxillary process; arrow head, mandibular process; p2, second pharyngeal arch.

The site of this expression in the CS16 embryo corresponded to the maxillary process and also the area surrounding the nasal pit. There was no obvious expression detected in the mandibular process in any of the embryonic stages studied (figs. 11 – 16 and data not shown).

At CS17, strong, restricted expression was observed in the mesenchyme of the medial and lateral nasal processes and the lateral palatal processes (Fig. 13). Signal was also detected in the frenulum of the tongue and the mesenchyme of the future skull beneath the forebrain and surrounding the eye. This expression was observed in the medial and lateral nasal processes at the front of the stomodeum in the CS19 embryo (Fig. 14e), as they swelled to form the nasal sacs and the primary palate. (see 1.2.2).

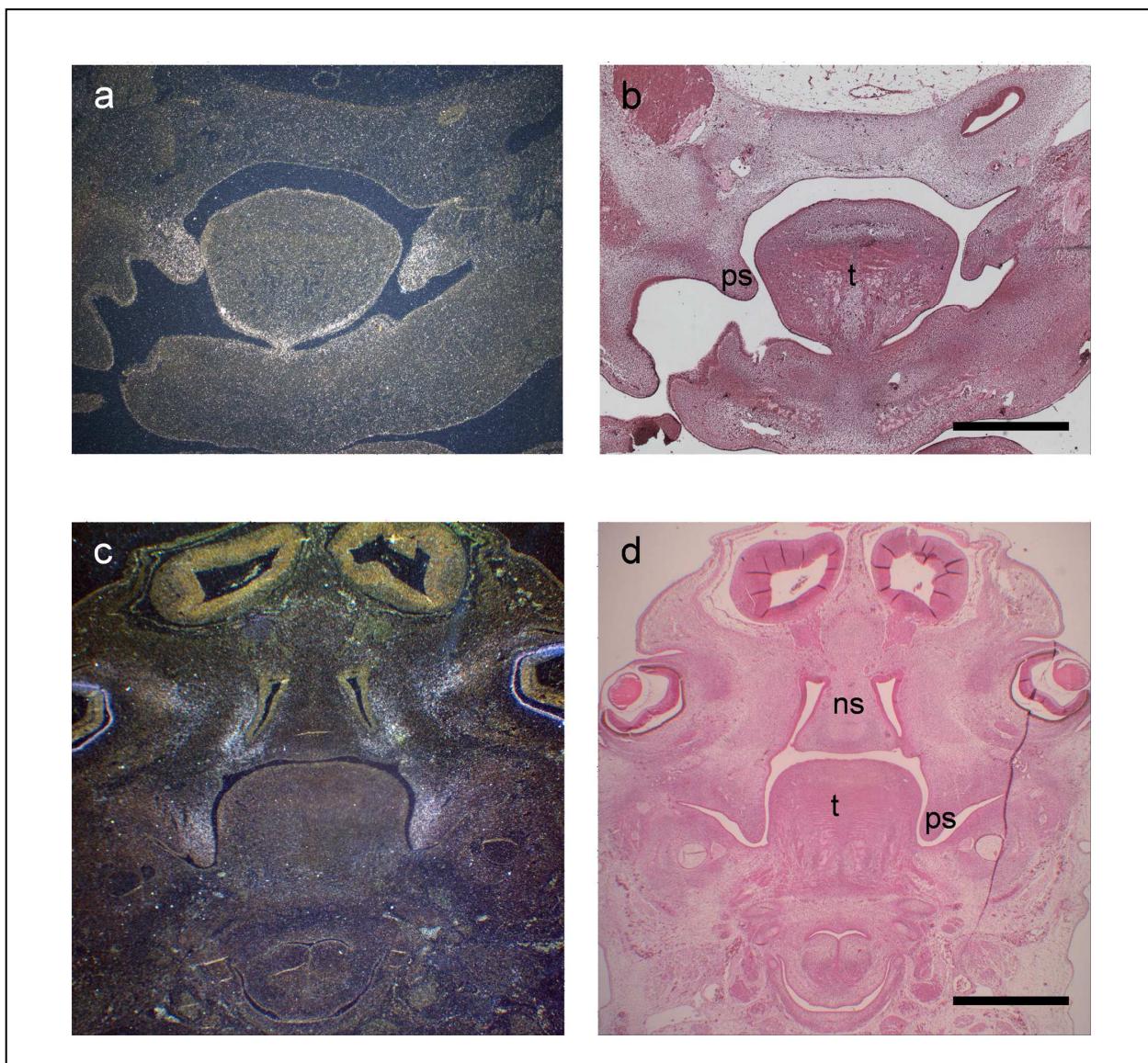


**Figure 13: Expression of *TBX22* in the CS17 embryo.** *TBX22* is expressed in the maxillary processes and medial and lateral nasal processes. Transverse sections of a CS16 (a) and (b) and CS17 (c) and (d) through the stomodeum. (a) and (c) are dark-field images of antisense *TBX22* probe and (b) and (d) are bright field images of haematoxylin and eosin stained sections. Bar = 400 $\mu$ m. max, maxillary process; man, mandibular process; p2, second pharyngeal arch; lpp, lateral palatal process; m, medial nasal process; l, lateral nasal process; t, tongue. No expression was seen with sense probes (data not shown).



**Figure 14: *TBX22* is expressed during the formation of the primary and secondary palate.** Transverse sections of a CS19 embryo, at the back (**a and b**), middle (**c and d**) and front of the stomodeum (**e and f**). (**a**), (**c**) and (**e**) are dark-field images of antisense *TBX22* probe and (**b**), (**d**) and (**f**) are bright field images of haematoxylin and eosin stained sections. Bar = 400 $\mu$ m. lpp, lateral palatal process; m, medial nasal process; l, lateral nasal process; t, tongue. No expression was seen with sense probes (data not shown).

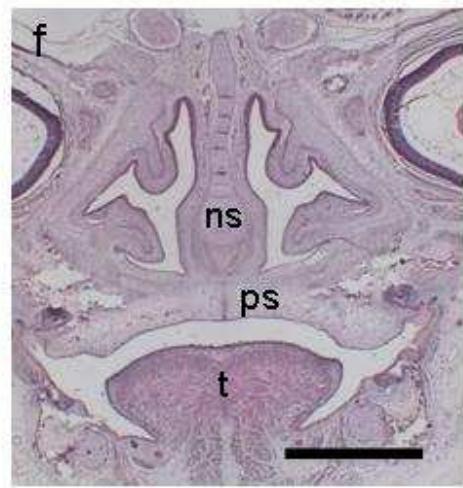
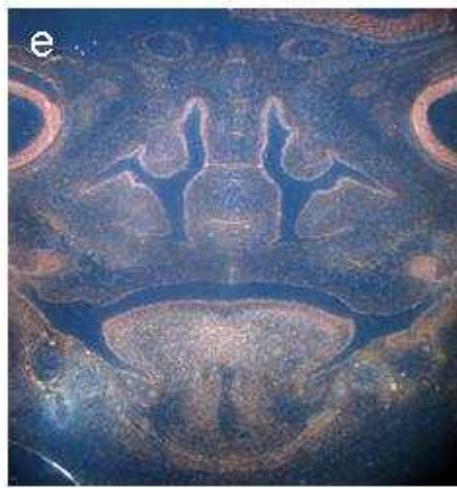
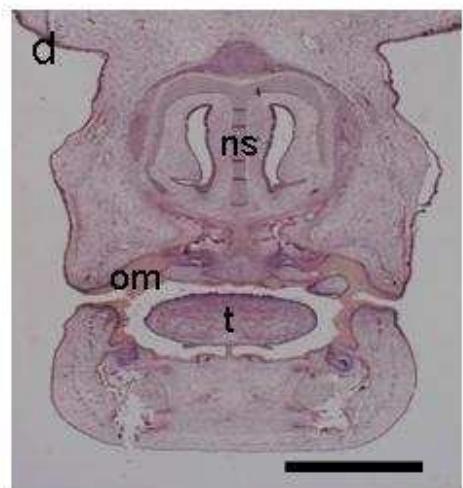
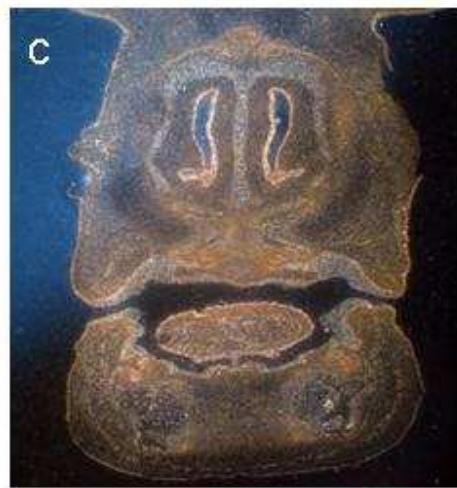
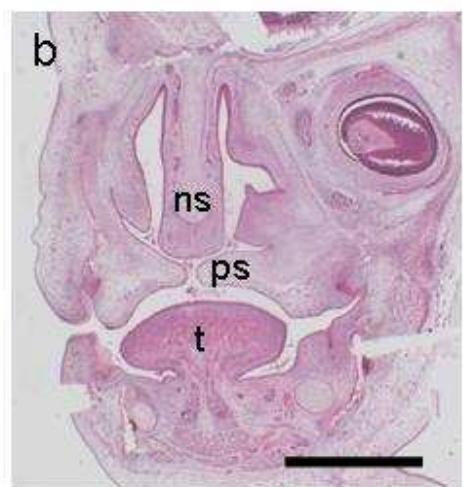
Very intense expression was detected in the swellings of the lateral palatal processes of the CS19 embryos (Fig. 14c). The regions of the maxillary processes where the palatine processes were not expanding showed reduced expression (figs. 14a and e); although localised expression was detected in the mesenchyme surrounding the nasal pit (Fig. 14e) and the region where the maxillary and mandibular processes will fuse along the nasolacrimal groove (Fig. 14a, see 1.2.2 for more detail). The signal in the base of the tongue, first seen at CS17 (Fig. 13c), continued to be detected at CS19 (figs. 1c and e).



**Figure 15: *TBX22* expression in the extending palatal shelves and in the frenulum of the tongue.** Transverse sections of a CS20 (a) and (b) and a CS21 (c) and (d) embryo. (a) and (c) are dark-field images of antisense *TBX22* probe and (b) and (d) are bright field images of haematoxylin and eosin stained sections. Bar = 1000 $\mu$ m. ns, nasal septum; ps, palatal shelf; t, tongue. No expression was seen with sense probes (data not shown).

By CS20 (figs. 15a and b) and CS21 (figs. 15c and d), intense expression was most evident in the regions closest to the oral cavity of the vertically extending palatal shelves and the floor of the tongue, which has by now mostly separated from the mandible.

At CS23, TBX22 expression was down regulated in the palatal shelves (figs. 16a); although, there was a clear band of expression in the mesenchyme of the nasal septum adjacent to the epithelial cells where fusion with the palatal shelves will occur (Fig. 16e). After 9 weeks of development the shelves have fused and *TBX22* was completely undetected in the palatal shelves (Fig. 16c and e). At this time the restricted expression seen at the base of the tongue had become more widespread through the whole of the tongue and the expression in the nasal septum was now seen throughout the nasal cartilage. The forming odontogenic mesenchyme was now expressing *TBX22* as well as the regions where cartilage will form in the septum and wings of the nose (om in Fig. 16d).



**Figure 16: *TBX22* is down regulated as the palatal shelves fuse.**

Transverse sections of a CS23 embryo (a) and (b), and a 9 week foetus (c – f). (a), (c) and (e) are dark-field images of antisense *TBX22* probe and (b), (d) and (f) are bright field images of haematoxylin and eosin stained sections. Bar = 1000 $\mu$ m. ns, nasal septum; ps, palatal shelf; t, tongue. No expression was seen with sense probes (data not shown). (c) and (d) are sections through the anterior region of the stomodeum in the region of the pre-maxilla and (e) and (f) are more posterior sections through the secondary plate.

## 2.4 DISCUSSION

As TBX22 functions as a transcription factor (see 1.1), the ideal methodology to reveal the tissues where TBX22 was exerting its effect would be by investigating the protein expression pattern, using an antibody specific to the TBX22 protein. However, at the time of performing these studies no suitable commercial antibody was available to do this. Such antibodies are now available, although they have not been validated for use in immunohistochemistry experiments. Therefore, there were two options available: either produce an antibody to use for this purpose, or use an alternative method. A major obstacle in the production of antibody is to verify its specificity, especially in relation to other related proteins, which in this case would be other T-box family members. To allay doubts as to cross-hybridisation with other T-box proteins and to ensure that an accurate temporal and spatial expression pattern was determined, rather than investigate the TBX22 protein expression, the alternative approach of determining the *TBX22* mRNA was undertaken. A probe was designed such that only 74% similarity existed between the nearest human T-box homologs *TBX15/18* and then only over a 250 nucleotide region of a 336 nucleotide probe. Although radioactive tissue *in situ* hybridisation was undertaken to study the expression of *TBX22*, non-radioactive methods have now been developed (Darby *et al.* 2009) which retain the sensitivity of radioactive methods but have the added advantage of cellular resolution. This would now be the method of choice and was employed in later tissue *in situ* hybridisation experiments (Chapter 4).

Unlike the results of the RT-PCR screen in which all tissues examined showed expression of *TBX22* (Braybrook *et al.* 2001), tissue *in situ* hybridisation revealed a very specific expression profile in the developing human. Expression is observed from an early period in craniofacial development, prior to the formation of the palatal processes and then in the mesenchyme of the extending palatal shelves until fusion. Significantly, strong expression is also detected in the floor of the tongue. This expression pattern correlates well with the observed phenotype in the CPX disorder: a cleft of the secondary palate and ankyloglossia. However, there are also regions of *TBX22* expression in tissues other than those associated with CPX and this will be discussed in section 2.4.3.

#### **2.4.1 *TBX22* Expression correlates with the CPX phenotype**

*TBX22* expression was first detected at CS16, in the maxillary processes as they extend from the first pharyngeal arch (Fig. 12c). From this stage until fusion, *TBX22* is expressed in the mesenchyme of all developing oronasal processes: the medial and lateral nasal processes and the palatal shelves (figs. 12 – 16). This is most evident in the palatal shelves where the expression is seen at the tip of the expanding shelves as they extend into the oral cavity.

Just prior to fusion of the shelves at CS23, the expression in the palatal shelves is down regulated and the expression is restricted to a thin band in the nasal septum where it will fuse with the shelves. Post fusion of the shelves, after approximately 9 weeks of development, the expression in the palate shelves is completely undetectable as it is in the region where they have fused with the nasal septum. Throughout the formation of the secondary palate, strong expression is detected at the base of the tongue in the region where cell apoptosis is occurring to enable the tongue to separate from the floor of the mouth. This is consistent with the ankyloglossia phenotype seen in CPX patients, whereby the *frenulum linguae* does not recede sufficiently and instead continues to extend the full length of the tongue (Bjornsson *et al.* 1989).

#### **2.4.2 A Mechanism for a Cleft Palate**

As expression is observed in both the palatal shelves and at the base of the tongue, it is likely that the cleft palate and the ankyloglossia features observed in CPX are independent of each other and that the cleft palate seen in CPX patients is not due to incorrect movement or position of the tongue during palatogenesis, but involves errors in the growth or fusion of the palatal shelves themselves. *TBX22* within the palatal shelves is only present during the extension of the palatal shelves and after fusion an mRNA transcript is no longer detected. In the 9 week foetus after the shelves have fused and the medial edge epithelia (MEE) is transforming to mesenchyme (Fig. 16e and f), *TBX22* expression is no longer evident in the shelves, suggesting that *TBX22* has no direct function in this process. The expression in the palatal shelves is at its most prominent in the mesenchymal tissue during the period when they are extending

(figs. 12-15) and the expression appears to be down regulated as this process nears completion (Fig. 16). Thus, from the expression profile of *TBX22* in the palatal shelves during palatogenesis, it would appear that the cleft palate in the CPX is a consequence of insufficient growth of the palatal shelves and not as a result of incorrect fusion of the shelves. However, studies of the *TBX22* mutant mouse (Pauws *et al.* 2009a) revealed that the major role played by *TBX22* during palatogenesis is in orchestrating osteoblast formation, although studies in the chick do suggest that *TBX22* does also have a role in regulating proliferation (Higashihori *et al.* 2010).

#### A. A POSSIBLE FUNCTION OF *TBX22*

As expression of *TBX22* is not detected until the formation of the maxillary processes (Fig. 12, earlier stages not shown), it would appear unlikely that *TBX22* has a role in directing neural crest cells to the first pharyngeal arch. It would seem that *TBX22* exerts its influence on palate formation at a later stage of development, perhaps in the mesenchyme cells in the region in which it is expressed. Due to the intense expression during the extension of the palatal shelves (figs. 12-15) and subsequent down regulation just prior to fusion of the shelves (Fig. 16), it does suggest, that *TBX22* has an influence on a gene or genes which are involved in either establishing or maintaining cell proliferation during the initiation and/or extension of the palatal shelves until such time that the two shelves have extended far enough that fusion between them is possible. While it might be postulated that *TBX22* is regulating a pathway of cell proliferation during the establishment and extension of the medial and lateral nasal processes and the palatal shelves, it would seem that the reverse might be true where expression is seen in the region of the floor of the tongue. At this stage of embryonic development, the cells in this region are undergoing apoptosis. This may suggest that: *TBX22* is in fact a regulator of programmed cell death; however, if this was true in the palatal shelves, one would envisage that a mutation in *TBX22* leading to a loss of protein function (Braybrook *et al.*, 2001) would give rise to an over enlarged palate rather than the underdeveloped palate observed in the CPX phenotype. One explanation is that *TBX22* has more than one target; it may act to regulate different pathways in different tissues. Indeed, it has been shown that T-box genes have several different downstream targets. The *Xenopus* Brachyury protein (Xbra), for example, has been shown to have at least 3 different direct target genes: *Bix1* (Tada *et*

*al.* 1998); *Bix4* (Casey *et al.* 1999) and *eFGF* (Casey *et al.* 1998). Another screen identified a possible 4 additional targets (Saka *et al.* 2000). As the TBX22 protein is known to contain both repression and activation domains (Andreou *et al.* 2007), the regulation of more than one genetic pathway is a plausible theory.

Another possibility is that TBX22 may function as a repressor or an activator depending on which co-factors are interacting with it. Another transcription factor - the Mgα protein – which contains a T-box domain, has been shown to repress the expression of reporter genes containing a Brachyury DNA binding site and also act as a transcription activator when tested with reporter genes carrying a Brachyury and a Myc-like basic helix-loop-helix leucine zipper motif (Hurlin *et al.* 1999). Interaction of T-box proteins with different cofactors has been shown to change the specificity of binding to target genes and hence the same proteins are capable of regulating different pathways during cardiac formation (Boogerd *et al.* 2009).

### 2.4.3 Expression outside of the palate

The typical phenotype of a CPX patient has neither a cleft of the primary palate (the most anterior region of the future hard palate, which includes the first four incisors and extends back to the incisive foramen (see section 1.2), nor a cleft lip. However, *TBX22* is expressed during the expansion of the medial and lateral nasal process, which when fused with each other and the maxillary process forms the primary palate and the upper lip (see 1.2). A cleft of the primary palate or of the lip is caused by failure of mesodermal delivery and proliferation from the maxillary and nasal processes (Talmant 2006). As *TBX22* is expressed in the maxillary and nasal processes (Fig. 13) one might expect that a disruption of the *TBX22* protein would have an effect on these structures also. However, cleft lip is not evident in classical CPX and the mouse knockout also does not have a cleft lip (Pauws *et al.* 2009a).

However, it has been reported that an individual with cleft lip and cleft palate was found in a family which displayed CPX and carried the same mutation in the *TBX22* gene as other family members who only had a cleft palate (Braybrook *et al.* 2001). Also, it has been shown that susceptibility to non-syndromic cleft lip may be linked to the X chromosome. There is also evidence that when *TBX22* is virally over-expressed in the frontonasal process in chick, a cleft lip often results (Higashihori *et al.*). It has been suggested that genetic or environmental disruption in the upstream regulation of *TBX22*, such that a gain of function of *TBX22* was brought about, may cause a cleft lip phenotype in human (Higashihori *et al.* 2010).

Alterations in *TBX22* in the maxillary process of CPX patients may not be critical to the correct formation of the lip as other T-box proteins may be able to functionally compensate for *TBX22* during the development of the upper lip. RNA *in situ* hybridisation experiments in mouse have shown that T-box genes other than *Tbx22* are expressed in the maxillary process: *Tbx1*, *Tbx2*, *Tbx18* and *Tbx19* (Kraus *et al.* 2001a; Bush *et al.* 2003; Gray *et al.* 2004) and perhaps this may explain why, although *TBX22* is expressed in the maxillary process, a cleft lip is not an observed clinical feature of CPX. It has been shown *in vitro* at least that the binding sites of all T-box proteins share a similar recognition sequence ((Tada and Smith 2001) and

references therein) and the possibility of trans-activation of the same downstream target genes cannot be disregarded.

There are other sites of expression, as well as those discussed above, primarily in the region surrounding the eye and at the base of the brain from CS17 – 19 (figs. 13 and 14). Following fusion of the palatal shelves the major sites of expression are the nasal cartilage and the forming orthodontic mesenchyme. However, none of the individuals with CPX have been reported with malformations of the eye, skull or tooth development (Lowry 1970; Moore *et al.* 1987; Bjornsson *et al.* 1989) and compensation by other T-box genes may be the reason defects of these structures have not been reported in CPX patients.

As has already been stated, tongue-tie (ankyloglossia) is not always observed in CPX patients (see 1.3.3.B). Again a compensation effect by another T-box gene could explain this observation. It has been suggested, in chick at least, that the cells expressing *Tbx3* in the tongue (Huang *et al.* 1999) are the same cells that express *Tbx22* (Haenig *et al.* 2002). If the expression patterns are similar in human, then one could envisage that TBX3 was able to functionally compensate for the lack of TBX22 in those CPX patients where tongue tie was not evident.

#### 2.4.4 Comparison with Mouse and Chick *Tbx22* expression

Expression studies have shown that the expression profile of mouse (Braybrook *et al.* 2002; Bush *et al.* 2002; Herr *et al.* 2003) and chick (Haenig *et al.* 2002) *Tbx22* closely resembles that of human *TBX22* and in all species the expression is always restricted to mesenchymal tissue. However, there are subtle differences in the expression of *Tbx22* in each of the three species. Table 9 shows a comparison amongst these three species and the sites of *Tbx22* expression during development, reference to Table 6 may be needed to compare equivalent human, mouse and chick stages.

In both chick and mouse *Tbx22* expression was detected much earlier in development with expression seen in the emerging somites in mouse at E9.25 before being then down-regulated in newly formed somites and then upregulated once more in more mature somites at E10.5. The expression in the somites was also detected in the studies in chick from HH stage 7 through to later stages becoming restricted to the myotome around E3. The *Tbx22* expression seen in the mature somites in the E10.5 mouse and E3 chick was not evident in the equivalent human stage, the CS14 embryo (Fig. 11b).

Even earlier expression of *Tbx22* has been reported during chick development, with signal detected in the mesoderm surrounding the primitive streak as early as HH stage 4 (c.f. mouse ~ 7 dpc, Thieller stage 10; and human ~ 18 dpo, Carnegie stage 8), however comparative stage mouse studies are not reported. Due to the constraints working with human material, very early tissue is incredibly hard to obtain and whether *TBX22* is expressed in these structures at this time is so far unknown.

Another notable difference in the expression pattern between chick and human is in the developing limbs. *Tbx22* expression in both the hind and forelimbs is seen in the E4 chick, which is certainly absent in the hindlimb at an equivalent human stage (Fig. 12) although expression in the forelimb was not examined.

Tissue	Human	Mouse	Chick
Mesoderm surrounding the primitive streak	?	?	From HH stage 4
Nascent somites	?	Evident at E9.25	Evident at HH stage 7
Newly formed somites	?	X No signal in E10.5 newly formed somites	Evident at HH stage 13
Mature somites	X No signal in CS14 embryo (Fig. 11)	Evident at E10.5	Restricted to myotome E3
Mandibular Process	At the tongue boundary only	At the tongue boundary only	Evident at E5
Maxillary Process/Palatal Shelves	Evident from CS16 (Fig. 12)	Evident from E10.5	Evident from HH stage 18
Medial and Lateral nasal Processes	Evident from CS17 (Fig. 13)	Evident from E10.5	Evident from HH stage 18
Mesenchyme surrounding the eye	Evident from CS17 (Fig. 13)	Evident from E13.5	Evident from HH stage 18
Hindlimb	X No signal in CS15 or CS16 embryo (figs. 11 and 12)	?	Evident from E4
Forelimb	?	?	Evident from E4

**Table 9: Sites of *Tbx22* expression during development of the human, mouse and chick.** Mouse data from Bush *et al.* 2002 and chick data from Haenig *et al.* 2002. ? signifies that expression in that tissue hasn't been reported in that species and an X indicates expression is not detected in that tissue in that species.

In all three species during the formation of the palate, a remarkably similar *TBX22* expression pattern is observed. This gives credence to the possibility that the role of *TBX22* in the formation of the secondary palate has been conserved through evolution. This conservation, in turn, suggests that *TBX22* is critical for correct palate development and that *TBX22* has a function that cannot be compensated for by other proteins. This may also form part of the explanation as to why the observed clinical feature of CPX is solely a cleft of the secondary palate, despite expression seen in other tissues.

It appears that the evolutionary pressure to conserve the presence of the protein in regions outside the palate (where *TBX22* expression differs across species) has not been as great - suggesting that it may not be as critical in the formation of other tissues. The absence of *TBX22* expression in the limbs and somites which is seen in chick and mouse may be due to other T-box proteins replacing the function of *TBX22* in these tissues in human.

## CHAPTER 3

### THE TBX22 DNA BINDING SEQUENCE

#### 3.1 INTRODUCTION

Understanding the interaction between the various genetic factors involved in the formation of the palate is fundamental to establishing how normal palatogenesis occurs, which in itself is a necessary prerequisite if we are to fully appreciate how disruptions to these processes may result in a cleft. Elucidating the roles of transcription factors and the genes that they are regulating within these genetic pathways is one of the steps necessary to achieve this goal. Following the discovery that *TBX22* is expressed throughout the facial processes during palate formation (Braybrook *et al.* 2002) and that mutations in the *TBX22* gene lead to CPX (Braybrook *et al.* 2001; Chaabouni *et al.* 2005; Andreou *et al.* 2007), establishing which genes *TBX22* may regulate during palatogenesis would not only provide insights into the genetic factors involved in normal palatogenesis, but also further our understanding of the aetiology of CPX. Furthermore, it may suggest other candidate genes for cleft palate disorders. Although *TBX22* has been shown to be directly regulated by *MN1* (Liu *et al.* 2008), to date the only gene that *TBX22* has been shown specifically to *trans* regulate is *TBX22* itself (Andreou *et al.* 2007), although there is evidence suggesting that it may regulate *MSX2* and *DLX5* (Higashihori *et al.* 2009).

#### 3.1.1 T-box domains

In a similar manner to other T-box family members, *TBX22* is likely to bind to the DNA sequences of the genes it is regulating via the T-box domain. The specific DNA sequence or T-box element (TBE) that *TBX22* binds will determine which genes it will regulate. Finding the genes that *TBX22* regulates would be greatly aided by

knowing the TBE to which it binds, as it would be expected that this TBE, or similar sequence to it, would be found within the regulatory regions of such genes.

Preferential DNA binding sequences have been revealed for several T-box proteins (Table 10) and all have been shown to bind to a core DNA sequence similar to that found for the original Brachyury protein (Kispert *et al.* 1995). However, slight variations in the preferred DNA binding sequences, especially in the flanking nucleotides, have been observed for several T-box proteins and it is postulated that these slight variations in DNA binding sequence is one mechanism by which functional specificity of different T-box proteins is conferred (Conlon *et al.* 2001).

The original Brachyury protein was shown to preferentially bind to the 20 nucleotide DNA sequence TTTCACACCTAGGTGTGAAA (Kispert *et al.* 1995). This is a palindromic sequence comprised of two 10 bp half-sites which are the reverse complement of each other. Since the identification of this DNA binding site several other T-box proteins have been shown to bind to DNA sequences similar to either this full length 20 bp sequence (T-site) or one half of it (T<sup>1/2</sup>-site). Table 10 summarises the experimentally determined variety of T-box DNA binding sites. In some cases the standard IUPAC notation (IUPAC-IUB 1971) has been used to represent ambiguous nucleotides and these definitions are shown in Table 11.

Whilst the T-box domain is a highly conserved region throughout all of the T-box protein family, the way in which these different proteins bind to DNA has been shown to be varied and T-box proteins have been reported as binding to DNA as either a monomer or a dimer (see 3.4). If the protein binds to the DNA as a dimer, then a full length T-site recognition sequence is needed; each monomer binding to one half of the sequence as is the case with Brachyury (Papapetrou *et al.* 1997). However, if the protein binds as a monomer, then only a half site sequence (T<sup>1/2</sup>-site) is required for the protein to bind to DNA, as has been revealed for TBX2 and Tbx3 proteins (Sinha *et al.* 2000; Carlson *et al.* 2001).

Species	Protein	DNA Binding Sequence												Reference								
Human	Brachyury	T	T	T	C	A	C	A	C	C	T	A	G	G	T	G	T	G	A	A	(Papapetrou <i>et al.</i> 1997)	
Human	TBX1	T	T	T	C	A	C	A	C	C	T	A	G	G	T	G	T	G	A	A	(Sinha <i>et al.</i> 2000)	
Human	TBX2											A	G	G	T	G	T	G	A	A	(Sinha <i>et al.</i> 2000; Lingbeek <i>et al.</i> 2002)	
Human	TBX3	T	C	A	C	A	C	A	C	C	T	T	G	G	T	G	C	C	A	A	(Lingbeek <i>et al.</i> 2002)	
Human	TBX5											R	G	G	T	G	T	B	R	N	N	(Ghosh <i>et al.</i> 2001)
Mouse	Brachyury	T	T	T	C	A	C	A	C	C	T	A	G	G	T	G	T	G	A	A	(Kispert <i>et al.</i> 1995)	
Mouse	Tbx6											A	G	G	T	G	T	B	R	N	N	(White and Chapman 2005)
Xenopus	VegT	C	T	T	G	A	C	A	C	C	T										(Casey <i>et al.</i> 1999)	
Xenopus	VegT											A	G	G	T	G	T	G	A	A	G	(Hyde and Old 2000)
Xenopus	Xbra	T	T	T	C	A	C	A	C	C	T											(Casey <i>et al.</i> 1998)
Xenopus	Xbra	C	T	T	G	A	C	A	C	C	T											(Casey <i>et al.</i> 1999)
Ciona	Ci-Bra	G	G	C	A	C	C	T	C	C	T											(Di Gregorio and Levine 1999)

**Table 10: A summary of different experimentally determined T-box protein DNA binding sites.** The DNA binding sequences of the different T-box proteins shown in this table include both the preferred *in vitro* DNA binding sites and known *in vivo* target sequences. Standard IUPAC abbreviations have been used to represent ambiguous bases (see Table 11).

UPAC nucleotide code	A	C	G	T	R	Y	S	W	K	M	B	D	H	V	N
Base	Adenine	Cytosine	Guanine	Thymine	A or G	C or T	G or C	A or T	G or T	A or C	C or G or T	A or G or T	A or C or T	A or C or G	Any Base

**Table 11: The recommended IUPAC abbreviations for nucleic acids (IUPAC-IUB 1971).**

### 3.1.2 Generating the TBX22 protein

In order to study the DNA binding properties of TBX22, a suitable TBX22 protein must be synthesised. However, not only must TBX22 protein be generated, it must be prepared in a suitable manner such that it can be used in protein function studies; that is, a biologically functional TBX22 protein must be produced. This essentially means that protein conformation must be similar to that seen *in vitro*. To achieve this aim, the protein must be recovered under native buffer conditions which do not alter the structure of the protein or else the protein must be recovered and then correctly refolded back to its native state.

Two methodologies were investigated to produce a native TBX22 protein: a bacterial cell approach, whereby the TBX22 protein is expressed by the bacteriophage T7, which can be induced in the BL21 star (DE3) host *E.coli* strain (Invitrogen) in the presence of IPTG; and a coupled cell-free transcription/translation approach whereby plasmid DNA is first transcribed by the addition of a bacteriophage RNA polymerase, and the resultant RNA is then translated into protein using the translation machinery from a Rabbit Reticulate Lysate System (TNT T7 Coupled Reticulocyte Lysate System; Promega). The bacterial cell system can deliver a considerably greater yield, but the recombinant protein is often difficult to recover from the cell lysate in a native form, whereas the Rabbit Reticulate Lysate method overcomes this problem as cells are pre-lysed, but the amount of protein generated in this way is usually much less (Hurst *et al.* 1996).

## 3.2 MATERIALS AND METHODS

### 3.2.1 Producing TBX22 Protein with an N-terminal 6xHis tag in *E. coli*

*TBX22* with an N-terminal 6 amino acid Histidine tag (6xHis) was expressed in *Escherichia coli* by cloning the *TBX22* coding sequence (see 3.2.1.B below) into a pET100/D-TOP<sub>O</sub> vector (Invitrogen). The pET100/D-TOP<sub>O</sub> vector contains the DNA coding sequence for a 6xHis tag which is placed in frame, upstream of the cloned insert. Thus when translated an RNA transcript containing the coding sequence for a 6xHis tagged *TBX22* protein is generated.

The 6xHis tag was necessary to aid subsequent recovery and purification of the plasmid by Ni-NTA. Another reason to generate a 6xHis tagged protein was that although commercial antibodies to *TBX22* are available, the majority have only been tested for use in Western blot applications. As antibodies to a 6xHis tag have been used successfully in several applications, including use in gel-shift assays of T-box genes (Ghosh *et al.* 2001) verification of the presence of a His tagged protein using an anti-6xHis tag antibody should be possible.

The protein was synthesised in bacteria because it was possible to generate a large yield of recombinant protein. The pET100/D-TOP<sub>O</sub> vector utilises a T7 bacteriophage promoter site to drive the expression of the cloned DNA insert in the presence of a T7 RNA polymerase. BL21 Star (DE3) cells (Invitrogen) were used to express the recombinant 6xHis tagged *TBX22* protein. BL21 Star (DE3) cells contain the DE3 bacteriophage lambda ( $\lambda$ DE3) lysogen. The  $\lambda$ DE3 lysogen contains a T7 RNA polymerase gene under the control of the *lacUV5* promoter. Addition of isopropyl  $\beta$ -D-thiogalactoside (IPTG) can induce expression of the RNA polymerase by the *lacUV5* promoter, and therefore drive the expression of the 6xHis-*TBX22* protein from the T7 promoter.

As the 6xHis-*TBX22* protein is foreign to the host cell, it is potentially toxic and could reduce growth of the bacterial culture in which it is expressed. It is preferable to only express the 6xHis-*TBX22* protein following significant growth of the bacterial

culture, thereby limiting these potential effects. However, there will always be some basal level expression of the RNA polymerase from the *lacUV5* promoter. To help prevent the expression of 6xHis-TBX22 protein from this basal expression of T7 RNA polymerase, the pET100/D-TOPO has a *lac* operator sequence placed immediately down stream of the T7 RNA polymerase promoter site, prior to the *6xHis-TBX22* coding sequence. The *lac* operator sequence is a binding site for the lac repressor which when bound to the DNA is capable of preventing the T7 RNA polymerase from binding to DNA and transcribing the *6xHis-TBX22* sequence downstream. Within the  $\lambda$ DE3 lysogen is the *lac I* gene which encodes for the lac repressor. The lac repressor is removed from the *lac* operator sequence in the presence of IPTG, allowing the T7 RNA polymerase to bind to the DNA and transcribe the *6xHis-TBX22* sequence.

#### A. PRIMER DESIGN AND SYNTHESIS

Primer pairs were designed using the software program “Web primer” (<http://seq.yeastgenome.org/cgi-bin/web-primer>) and were commercially synthesised by MWG Biotech. The oligonucleotides were received desiccated and were reconstituted using Milli-Q water to stock concentration of 100pmol/ $\mu$ l, based on the concentration supplied by the manufacturer. The melting temperature for each oligonucleotide was also specified by the supplier. Appendix 1 summarises all the primer pairs used in these studies.

#### B. POLYMERASE CHAIN REACTION

The polymerase chain reaction (Mullis *et al.* 1986) was used to amplify the *TBX22* coding sequence from a plasmid containing the *TBX22* cDNA (clone BC014194; I.M.A.G.E. Consortium) using a Dyad MJ Research Thermo Cycle instrument. A reaction mix contained a final concentration of 100ng plasmid DNA, 10mM dNTPs (Fermentas), 1X PFU polymerase buffer (20mM Tris-HCl (pH 8.8 at 25°C), 10mM KCl, 10mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 2mM MgSO<sub>4</sub>, 0.1% Triton X-100, 0.1mg/ml nuclease-free

BSA; Promega), 1.25U Pfu DNA polymerase (Promega) and 0.5 $\mu$ M forward and reverse primers made to a final volume of 50 $\mu$ l using Milli-Q water.

The Thermo cycler was programmed to perform an initial denaturation step of 5 minutes at 95°C, followed by 25 cycles of 95°C denaturation for 1 minute, 1 minute at the specific annealing temperature, followed by an extension time of 3 minutes. The annealing temperature was determined specifically for each primer pair (Appendix 1 gives details of the primer pairs). To begin with an initial annealing temperature 5°C lower than the lowest melting temperature of primer set was used for the reaction. If non-specific PCR products were produced, the annealing temperature was increased stepwise by 1-2°C. If however, the desired product was not obtained, then the annealing temperature was reduced stepwise by 1-2°C until a single, discrete product could be visualised by UV illumination following electrophoresis on a 1% agarose gel.

An amplicon consisting of the *TBX22* coding sequence - Genbank Accession number NM\_016954, nucleotides 28-1560 (<http://www.ncbi.nlm.nih.gov>) was synthesised using a forward primer 5'-CACCATGGCTCTGAGCTCTGGG-`3 and reverse primer 5'-CTAAAGGTAATGGTTAATTGCTGG-`3 using I.M.A.G.E. clone BC014194 (Geneservice) as a DNA template.

### c. CLONING THE PCR PRODUCT

The *TBX22* coding sequence was cloned in frame into a pET100/D-TOP0 (Invitrogen) plasmid expression vector. The initial CACC at the 5' end of the forward primer immediately prior to the ATG initiation codon of the *TBX22* coding sequence enables the resultant PCR product to anneal to the GTGG overhang of the cloning vector increasing the likelihood of correct orientation of the insert into the vector. The pET100/D-TOP0 vector utilises the DNA binding capabilities of Topoisomerase I from the *Vaccinia* virus, enabling the insert to be cloned into the vector without the need of a ligation reaction. Following PCR, the insert was cloned into the pET100/D-TOP0 according to the manufacturer's protocol, using a 2:1 molar ration of PCR product: TOPO cloning vector, calculated using the formula:-

$$\frac{X \text{ng PCR product} = (Y \text{ bp PCR product}) (Z \text{ ng vector})}{(Z \text{ bp vector})}$$

Where X is the amount of PCR product needed for a 1:1 insert:vector molar ratio. For a molar ration of 2:1, X is multiplied by a factor of 2.

#### D. TRANSFORMATION OF THE PLASMID DNA

The expression plasmid construct was transformed into TOP10 *E. coli* chemically competent cells (Invitrogen). A vial of TOP10 cells was thawed on ice. 3 $\mu$ l of the TOPO cloning reaction was added to the vial and was incubated on ice for 30 minutes. The cells were heat shocked for 30 seconds at 42°C and were then placed in ice water. 250 $\mu$ l of sterilised S.O.C media (2% Tryptone, 0.5% Yeast Extract, 0.05% NaCl, 20mM glucose, 2.5mM KCl, 10mM MgCl, pH 7.5) was added to the vial and incubated horizontally at 37 °C in a shaking incubator with a rotation speed of 200 rpm for 1 hour. 100 $\mu$ l and 200 $\mu$ l volumes of the transformation were plated on to Petri dishes containing Luria-Bertani (LB) media (1% tryptone; 1% NaCl; 0.5% Yeast Extract pH 7.0) with 1.5% agar and with 50 $\mu$ g/ml kanamycin (Sigma) as a selective antibiotic and incubated overnight at 37°C.

#### E. DNA PLASMID PREPARATION

Single bacterial colonies were picked from the selective plates and inoculated in a 5ml culture containing LB and 50 $\mu$ g/ml kanamycin (Sigma) and incubated at 37°C for 16 hours in a shacking incubator at 200 rpm. The bacterial cells were harvested by centrifugation in a Beckman Gene Genius gel doc system from Syngene at 6000g for 10 minutes at 4 °C. Following centrifugation, the plasmid DNA from bacterial pellet was recovered using a Miniprep Kit (Qiagen) in a microcentrifuge, in accordance with the manufacturer's instructions.

#### F. DNA SEQUENCING

DNA from each of the clones was sequenced by the Institute of Human Genetics sequencing service, performed using a MegaBACE 1000 (Amersham/GE Healthcare) machine using T7 promoter and T7 Reverse primers and TBX22 sequencing primers (see Appendix 1). The resulting sequence file was converted to FASTA format using the software programme Chromas (<http://www.technelysium.com.au/chromas.html>) from Technelysium Pty Ltd and aligned against the expected sequence using BLAST (Altschul *et al.* 1990). Clones containing the correct plasmid sequence were used for subsequent protein expression.

#### G. INDUCING EXPRESSION OF RECOMBINANT PROTEIN IN *E. COLI*

5 $\mu$ g of the purified plasmid DNA was used to transform one vial of BL21 Star (DE3) cells, following the transformation protocol in 3.2.1.C, except that the transformation reaction was not plated onto a Petri dish. Instead, the entire reaction volume was used to inoculate 10ml of LB containing 50 $\mu$ g/ml ampicillin and was incubated in a shaking incubator, at 37°C, 250 rpm, overnight until the OD<sub>600</sub> was in the range of 1-2.

For initial pilot studies, 10ml LB containing 50 $\mu$ g/ml ampicillin was inoculated with 500 $\mu$ l of this overnight culture. In purification experiments the volume was increased to 50ml of LB. The bacterial culture was then incubated at 37°C with shaking at 250 rpm until the cells were in mid-log phase (OD<sub>600</sub> = 0.6). At this point the pilot study cultures were split into two 5ml volumes, one of which was induced with IPTG, the other was uninduced to serve as a negative control. Following addition of 1mM IPTG to induce expression of the 6xHis-TBX22 plasmid; the cultures were replaced into the shaking incubator under the same conditions as previously. The cultures were grown for the optimum time determined using the time course pilot study (see 3.2.1.H). At this point the cells were harvested by centrifugation at 4000g for 15 minutes at 4°C and stored at -20°C.

## H. TIME COURSE PROTEIN EXPRESSION PILOT STUDY

In order to establish the optimum growth time for the bacteria following induction, 500 $\mu$ l samples were taken from the induced and uninduced growing cultures immediately after the addition of IPTG and then after 2, 4 and 6 hours. The cells were harvested by centrifugation at 13000g in a microcentrifuge and stored at -20°C.

The cells were then thawed on ice and resuspended in Lysis buffer 1 (50mM potassium phosphate, 400mM NaCl, 100mM KCl, 10% glycerol, 0.5% Triton X-100, 10mM imidazole, 8M urea). The cells were then frozen on dry ice and thawed at 42°C and then re-frozen and thawed again before being centrifuged at 13000g in a microcentrifuge at 4 °C for 1 minute. The supernatant was collected and the resultant pellet was resuspended in 500 $\mu$ l of 1X SDS. 6.5 $\mu$ l of the supernatant was added to sample running buffer and 10 $\mu$ l of the resuspended pellet (without the addition of sample running buffer) was separated on a polyacrylamide gel (see 3.2.3.A).

## I. NI-NTA PURIFICATION

A Nickel-Nitrolotriacetic acid (Ni-NTA) affinity column (Qiagen) was chosen for the recovery of the 6xHis tagged protein. This purification system is centred upon the high affinity for histidine residues to nickel ions which are bound to a NTA resin. However, to be employed successfully, a run of histidine residues must be incorporated into the protein. This histidine tag is commonly placed at either the N- or C-terminal of the recombinant protein. The decision was taken to place a 6x histidine tag at the N-terminal of the TBX22 protein using the pET100/D-TOP0 from Invtrrogen as it is possible to cleave the N-terminal His tag from the expressed protein. It is possible that the addition of a His tag to a protein may affect the natural biological function of the protein, so the possibility to remove it was an attractive feature.

Several cell lysate, wash and elution buffers were employed in order to purify the 6xHis tagged TBX22 protein under native conditions; initially however, the cells

harvested from a 10ml culture were lysed under denaturing conditions. The cell lysate was resuspended in 1ml urea buffer (8M urea, 0.1M  $\text{NaH}_2\text{PO}_4$ , 0.01M Tris-Cl, pH 8.0) loaded onto the Ni-NTA spin column; washed in urea buffer, pH 6.3; and eluted in urea buffer, pH 4.5; following the manufacturer's directions. By reducing the pH of the sample buffer, the affinity of the polyhistidine tag to the Ni-NTA column is reduced, making purification of the 6xHis tag protein possible.

Samples of the original un-purified cell lysate and the flow-through following each wash and elution step were collected and analysed on a Coomassie stained PAGE gel (see 3.2.3.A). The resultant single band in the final elution step was excised from the PAGE gel and used for verification by MALDI-TOF analysis (3.2.3.C).

The samples were also used in a second PAGE analysis which was subsequently used for a Western transfer (see 3.2.3.B) to a PVDF membrane and the N-Terminal 6xHis tag was detected using an Anti-HisG-HRP Antibody (Invitrogen).

In order to purify the 6xHis tagged TBX22 protein under native conditions, various different cell lysis buffers were evaluated for their ability to lyse the bacterial cells and solubilise the 6xHis tagged TBX22 protein. Following lysis, a sample of the cell lysate was analysed for the presence of the 6xHis tagged fusion protein.

The cell culture volume for these experiments was increased to 50ml and following centrifugation to harvest the bacterial cells, 1ml of the various lysis buffers was used to resuspend the cells. To aid cell lysis under these milder conditions, 1mg/ml lysozyme (Sigma) was added to the lysis buffer and the cells were sonicated on ice using a Sonics Vibra Cell machine for six sets of 10 seconds, with 5 second pauses in between each sonication. The various lysis buffers used in this study are detailed in Table 12.

Lysis Buffer	Composition	Reference
Sodium dihydrogen Phosphate *	50 mM Na <sub>2</sub> HPO <sub>4</sub> , 300 mM NaCl (pH 8.0)	Ni-NTA Spin Handbook (Qiagen)
Tris-Cl *	50 mM Tris-Cl (pH 7.6)	(Silva <i>et al.</i> 2003)
PBS *	140 mM NaCl, 2.7 mM KCl, 10 mM Na <sub>2</sub> HPO <sub>4</sub> , 1.8 mM KH <sub>2</sub> PO <sub>4</sub>	HisPur Ni-NTA Resin Instructions (Thermo Scientific)
IP-50 Lysis buffer	Cell Lysis buffer from Immuno Precipitation Kit	IP Kit Handbook (Sigma-Aldrich)

**Table 12: Cell lysis buffers used to solubilise the 6xHis tagged protein under native conditions.** The conditions used for lysing the cells under native conditions. \* These lysis buffers were used separately with the addition of 5mM β-mercaptoethanol (Merck) and 0.5% Tween (Sigma).

#### J. ON COLUMN RE-FOLDING

The bacterial cells were lysed in cell lysis buffer (6M urea, 0.1M NaH<sub>2</sub>PO<sub>4</sub>, 0.01M Tris-Cl, 5mM β-mercaptoethanol) and the cell lysate applied to a Ni-NTA spin column. The bound 6xHis tag protein was then washed in the lysis buffer with a gradual lowering of the urea concentration in the wash buffer, in a 0.5M step-wise gradient until the urea was completely removed from the wash solution. This was performed at 4°C using a previously published protocol (Oganesyan *et al.* 2005).

#### 3.2.2 Rabbit Reticulocyte Transcription translation

The TNT Coupled Reticulocyte Lysate System from Promega enables the transcription and translation of plasmid DNA. The plasmid must contain a prokaryotic phage RNA polymerase promoter upstream of a protein coding sequence, this RNA polymerase is utilised by an RNA polymerase which is added to a rabbit reticulocyte lysate, to transcribe the DNA template. The rabbit reticulocyte lysate contains all of the necessary components for the translation of this RNA to protein.

1 $\mu$ g of purified 6xHis tagged TBX22 pET100/D-TOPO vector DNA was used in a transcription translation reaction mix containing: 25 $\mu$ l TNT Rabbit reticulocyte lysate; 2 $\mu$ l TNT Reaction buffer; 1 $\mu$ l T7 RNA polymerase; 0.5 $\mu$ l 1mM amino acid mix (minus leucine); 0.5 $\mu$ l 1mM amino acid mix (minus methionine); brought to a final volume of 50 $\mu$ l with nuclease free water and incubated at 30°C for 90 minutes.

For use in the EMSA binding studies (see 3.2.5) the *TBX22* coding sequence was cloned into pTNT vector (Promega). This was achieved by amplifying the cDNA sequence, including the 6xHis tag from the pET100/D-TOPO TBX22 expression vector (3.2.1) using primers incorporating *SalI* and *XhoI* restriction enzyme sites respectively (see Appendix 1). Following separate double digestion reactions according to the manufacturers' instructions using *SalI* and *XhoI* restriction enzymes (Promega) of the PCR fragment and pTNT vector, the plasmid was ligated (see 3.2.1.C). The plasmid sequence was verified using T7 forward and reverse primers and the TBX22 sequencing primers outlined in Appendix 1. The pTNT vector, like the that of the pET100/D-TOPO vector contains a T7 RNA polymerase binding site ahead of the multiple coding region, but with the addition of the 5' untranslated region from the 5'  $\beta$ -globin gene positioned upstream of the cloned insert and synthetic poly(A)<sub>30</sub> tail downstream of it. These two elements have both been shown to increase the translation activity (Annweiler *et al.* 1991; Wakiyama *et al.* 1997). The transcription/translation reaction using this plasmid was performed in exactly the same way as outlined above.

### 3.2.3 Protein Verification

#### A. POLYACRYLAMIDE GEL ELECTROPHORESIS (PAGE)

An XCell *Surelock* electrophoresis mini-gel tank (Invitrogen) was used for the electrophoresis and Western transfer of protein samples. All of the gels used contained 12% bis-acrylamide and were precast in a 10X 1mm well format (NuPAGE Mini-Gel; Invitrogen) and were performed under denaturing conditions. An appropriate concentration of protein sample to a maximum of 6.5 $\mu$ l was mixed with sample running buffer (2.5 $\mu$ l 4X NuPAGE LDS Sample buffer (Invitrogen) and 1 $\mu$ l 10X NuPAGE Reducing Agent (Invitrogen) and brought to 10 $\mu$ l with Milli-Q water. The samples were then heated for 10 minutes at 70°C and loaded onto a 12% bis-tris 10X 1mm well NuPAGE Mini-Gel (Invitrogen) together with 6 $\mu$ l of Page Ruler Unstained Protein Ladder 10-200 KDa (Fermentas) which had been heated to 40 °C for 5 minutes. The gels ran with 200ml 1X NuPAGE MES buffer (Invitrogen) containing 500 $\mu$ l NuPAGE Antioxidant (Invitrogen) in the upper chamber and 200ml 1X NuPAGE MES Buffer (Invitrogen) in the lower chamber for 35 minutes at a constant 200 Volts.

To visualise the gel, the gel was removed from its apparatus and covered with Coomassie stain (0.5 % Coomassie Blue R-250 in 40% methanol, 10% acetic acid in Milli-Q water) for 30 minutes with gentle shaking. The gel was then destained in several changes of 40% methanol, 10% acetic acid with gentle agitation until the background was reduced to an acceptable level. The gel was then transferred to 5% acetic acid and imaged using Gene Genius gel doc system from Syngene with a UV filter.

#### B. WESTERN TRANSFER AND IMMUNO-BLOTTING

The PAGE gel was transferred to a blotting membrane using the XCell II Blot Module (Invitrogen). Following the electrophoresis, the PAGE gel was removed from its housing cassette and placed adjacent to a polyvinylidene difluoride (PVDF, Invitrogen) membrane which had been pre-soaked for 30 seconds in methanol, rinsed in deionised H<sub>2</sub>O and then soaked in NuPAGE transfer buffer (Invitrogen) for 10 minutes. Filter paper was rinsed in NuPAGE transfer buffer and placed either side of

the gel and PVDF membrane. Two foam blotting pads were placed in 1X NuPAGE transfer buffer and squeezed whilst submerged to remove all of the trapped air. The blotting pads were then placed on both sides of the filter paper to securely fix the gel and PDVF membrane during protein transfer. The whole assembly was then inserted into the XCell II Bot Module such that the gel was positioned nearest the cathode and the PDVF membrane closest to the anode. The assembled blot module was then placed into the XCell SureLock mini-cell and secured with the fixing wedge. The blot module was filled with transfer buffer (1X NuPAGE transfer buffer with 10% methanol and 1% NuPAGE antioxidant) and the outer tank chamber was filled with H<sub>2</sub>O, to act as a coolant during transfer. The protein was then transferred to the membrane by applying a constant 30V across the blot assembly for 1 hour.

Following transfer, the assembly was dismantled and the PVDF membrane was transferred to a small tray and covered with 10ml PBST (PBS containing 0.05% Tween-20 (Sigma-Aldrich)) and 0.5% Bovine Serum Albumin (BSA, Sigma-Aldrich) as a blocking solution. The membrane was washed in this solution for 1 hour with gentle agitation, then the solution was removed and the membrane was washed in two 20ml washes of PBST for 5 minutes, again using gentle agitation. The Anti-6xHis antibody (Invitrogen) was diluted 1:5000 in PBST with 0.5% Bovine Serum Albumin (Sigma-Aldrich) and 10ml of the diluted antibody was added to the tray containing the membrane. The antibody was incubated for 1.5 hours at room temperature with gentle agitation on a rocker. The antibody solution was removed from the tray and the membrane was again washed in two 20ml washes of PBST for 5 minutes with gentle agitation. The membrane was then transferred to a new tray containing a DAB tablet (Amersham) diluted in 10ml of H<sub>2</sub>O with 10% H<sub>2</sub>O<sub>2</sub>. The membrane was incubated in this solution for 10 minutes and was washed in three rinses of H<sub>2</sub>O for 5 minutes each. The membrane was then air-dried and imaged using a Polaroid digital camera.

### C. MADLI-TOF

Following the identification of a single band on a PAGE gel after Coomassie staining (see 3.2.3.A), the band was excised from the gel and used for protein verification in a

Matrix-assisted laser desorption/ionization-Time of flight (MALDI-TOF) measurement. The measurement, which was performed by the PINNACLE proteomics service at Newcastle University using a Voyager DE-STR mass spectrometer (Applied Biosystems), enabled a statistically significant match to the human TBX22 protein to be made.

### 3.2.4 *In Vitro* Oligonucleotide Selection Assay

An *in vitro* oligonucleotide selection assay was employed (Pollock and Treisman 1990) to identify the preferential TBX22 DNA binding site, as this method had previously been successful in identifying the DNA binding consensus sites for other T-box family members (Kispert and Herrmann 1993; Ghosh *et al.* 2001). The assay involves the selection of a particular DNA sequence from a pool of random oligonucleotides, based upon the binding of a T-box protein to that DNA sequence.

#### A. SYNTHESIS OF DOUBLE STRANDED RANDOM OLIGONUCLEOTIDES

A forward primer (Forward\_Rand) GCTGCAGTTGCACTGAATTGCCT and reverse primer (Reverse\_Rand) CAGGTCAAGTCAGCGGATCCTGTCG were synthesised by MWG- Biotech and used in a PCR reaction (see 3.2.1.B) to amplify a oligonucleotide, also synthesised by MWG-Biotech consisting of 5'- GCTGCAGTTGCACTGAATTGCCTC(NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN)NN)GACAGGATCCGCTGAAC TGACCTG-3'. At each of the 26 "N" bases all four possible nucleotides were represented, thus a 26-mer random oligonucleotide was synthesised with known 5' and 3' end sequences.

10ng of the random oligonucleotide was used as a template in a 50 $\mu$ l PCR reaction using 0.2 $\mu$ M of the Forward\_Rand and Reverse\_Rand primers. Following PCR the DNA was phenol/chloroform extracted and ethanol precipitated (see 2.2.2.A) to remove the protein and salt from the solution. The DNA was then incubated with 10U S1 Nuclease (Promega) in 1X reaction buffer (0.5M sodium acetate pH 4.5, 2.8M NaCl, 45mM ZnSO) at 37°C for 30 minutes to degrade the single-stranded DNA molecules. The nicked, single-stranded DNA was then removed from the solution by applying the DNA solution to a QIAquick Nucleotide Removal Kit (Qiagen) following the manufacturer's instructions.

## B. PROTEIN/OLIGONUCLEOTIDE BINDING REACTIONS

A protein/DNA binding reaction containing 10 $\mu$ l of the transcription/translation reaction (see 3.2.2) was incubated with 200pg of the random oligonucleotide in the

presence of 1X Binding buffer (5X Binding buffer, Promega: 20% glycerol, 5mM MgCl<sub>2</sub>, 2.5mM EDTA, 2.5mM DTT, 250mM NaCl, 50mM Tris-Cl (pH 7.5), 0.25mg/ml poly(dI-dC):poly(dI-dC) in a 20μl reaction volume and incubated at 25°C for 30 minutes.

#### c. CO-IMMUNOPRECIPITATION

A Protein G Immuno-Precipitation kit (Sigma) was used to recover the bound TBX22/DNA complexes. The entire protein/oligonucleotide binding reaction was incubated in an IP spin column (Sigma), preprepared as per the manufacturer's instructions, with the addition of 1μg Anti-6xHisG antibody (Invitrogen) and 10μl protease inhibitor cocktail (Sigma). This was made up to a final volume of 600μl with 1X IP buffer (Sigma). The spin column was incubated overnight at 4°C with inversion.

The following day the spin-column was washed in 700μl of 1X IP buffer and spun in a bench microcentrifuge at 12,000g for 30 seconds at 4°C. This wash and centrifugation step was repeated through 5 further rounds, with a final wash of 0.1% IP buffer and centrifugation at 12,000g for 30 seconds at 4°C. The bound protein/DNA was eluted from the spin columns by the addition of 40μl of 1X modified Laemmeli sample buffer (2% SDS, 15% β-mercaptoethanol, 10% glycerol, 60mM Tris-Cl pH 7.0) and heating the column to 95°C for 5 minutes, before centrifugation at 12,000g for 30 seconds.

#### d. DNA BINDING SITE ISOLATION

The DNA was recovered from the solution by phenol/chloroform extraction, with the addition of 10μg glycogen (VWR) as a carrier, and subsequent ethanol precipitation (see 2.2.2.A) then resuspended in 10μl H<sub>2</sub>O. The DNA was then re-amplified by PCR using 150ng of the forward and reverse primers for 13 cycles (see 3.2.1 B) using 1U Taq polymerase (Promega). This amplified DNA was then used as the substrate for further rounds of protein/DNA binding selection assays. Following PCR amplification from the fifth round of the binding selection assay, the DNA was cloned into a pCRII-

TOPO vector from Invitrogen (see 3.2.1.C) and transformed into Invitrogen's TOP 10 *E. coli* competent cells (see 3.2.1.D). The entire transformation reaction was plated onto LB plates containing 50µg/ml ampicillin which had been previously spread with 40µl of 40mg/ml X-gal (Fermentas) in dimethylformamide. Transformants were picked and cultured, the DNA midi prepped and sequenced by the IHG sequencing service using the M13 forward and reverse primers (see 3.2.1.E and F).

### 3.2.5 Electrophoretic Mobility Shift Assay (EMSA)

#### A. 5' END LABELLING

Oligonucleotides were synthesised by MWG-Biotech and reconstituted in Milli-Q H<sub>2</sub>O according to the suppliers' directions. Four double-stranded oligonucleotides were generated: W=T\_oligo – which represented the TBX22 DNA binding consensus sequence and contains a thymine base at the 8<sup>th</sup> position of the TBX22 consensus sequence (the “W” at this position within the TBX22 consensus sequence implies that this position can be represented equally by either a thymine or an adenine base). The 515-20\_oligo – which was clone 515-20 from the *in vitro* binding assay (see Table 14), is identical to the TBX22 consensus binding sequence at 8 of the 10 positions: the exceptions are a guanine base at the 7<sup>th</sup> position (instead of a cytosine) and a cytosine residue at the 10<sup>th</sup> position (in place of an adenine). The 8<sup>th</sup> position is an adenine base (this position can be either an adenine or thymine in the TBX22 DNA binding consensus sequence). The Mut\_oligo corresponds to the TBX22 DNA binding consensus sequence, except that the guanine residues at the 2<sup>nd</sup> and 3<sup>rd</sup> position in the consensus sequence were replaced by adenine residues and contains an adenine base at the “W” position of the TBX22 consensus sequence. Finally, Cont\_oligo is a positive control oligo which had previously been shown to bind TBX22 (Andreou *et al.* 2007). The oligos used are shown in Table 13.

ds Oligo	Oligo 1	Oligo 2
Cont_oligo	CTAGCAAGGTGTGAAATTGTCACCTCAA	GTTCCACACTTAACAGTGGAGTTCGA
W=T_oligo	CGAGAGGTGTCTTACGAG	CTCGTAAGACACCTCTCG
515-20_oligo	CGAGAGGTGT <b>GAT</b> CCGAG	CTCGGATCACACCTCTCG
Mut_oligo	CGAGA <b>AA</b> TGTCATACGAG	CTCGTATGACATTCTCG
TBX22 DNA binding consensus sequence	AGGTGTCWTA	

**Table 13: The oligonucleotides used to generate the double-stranded oligos used in the electrophoretic mobility shift assay of TBX22 protein.** The position of the TBX22 DNA binding sequence within Oligo 1 of the W=T\_oligo, 515-20\_oligo and Mut\_oligo is shown in bold and variations from this sequence are highlighted in red.

To generate the double-stranded oligos W=T, 515-20, Mut and Cont, the paired oligos (oligos 1 and 2 from Table 13) were heated to 10 °C above their melting temperature for 30 minutes and allowed to anneal by cooling slowly to room temperature. They were then treated with 1U of S1 Nuclease (Promega) for 30 minutes at 37 °C and purified through a QIAquick nucleotide removal kit (Qiagen) following the manufacturer's instructions.

2pmol of the double-stranded oligonucleotides were 5' end-labelled using 10U of T4 polynucleotide kinase (PNK; Promega) in a 10μl reaction containing 1X PNK buffer: 50mM Tris-Cl pH 7.5, 10mM MgCl<sub>2</sub>, 5mM DTT, 0.1mM spermidine (Promega) and 1μl of [ $\gamma$ -<sup>32</sup>P]ATP, 3000Ci/mmol at 10mCi/ml (Amersham Biosciences). The reaction was incubated at 37 °C for 10 minutes after which time the reaction was stopped with the addition of 1μl 5mM EDTA and the PNK enzyme heat-inactivated by incubation at 70 °C for 10 minutes. The volume was adjusted to 100μl with the addition of TE buffer.

#### B. ELECTROPHORETIC MOBILITY SHIFT ASSAY (EMSA)

TBX22 protein/DNA binding site reactions were setup by combining the TBX22 pTNT rabbit reticulocyte lysate with 5' end-labelled oligonucleotide (3.2.5.A). A binding reaction was setup by adding 2μl of the rabbit reticulate lysate with 2μl of 5X binding buffer. In initial experiments the binding buffer was that supplied from the Promega Gel Shift Kit, however in later experiments this was substituted for a binding buffer as described in (Andreou *et al.* 2007). For the positive competitor reaction, 2pmol of unlabeled consensus TBX22 oligonucleotide (W=T\_oligo) was also added to this reaction. To serve as negative controls, a further reaction containing 2μl lysate that contained the pTNT vector which lacked the *TBX22* coding sequence insert and another in which the lysate was absent were also set up. For use in competition experiments 2pmol of the W=T\_oligo was also included in the binding reaction. In all cases, the reaction volume was adjusted to 9μl with nuclease free water and incubated in a water bath at 25 °C for 20 mins.

An anti-TBX22 antibody (sc-17862X; Santa-Cruz) was used for super shift experiments. This antibody was specifically chosen as the manufacturer's advise using

this antibody for gel super shift experiments. Here, 1 $\mu$ l of the TBX22 antibody was added to the EMSA binding reaction containing 2 $\mu$ l of the TBX22 pTNT rabbit reticulocyte lysate from the transcription/ translation and 2 $\mu$ l of the 5X binding buffer. This volume was adjusted to 9 $\mu$ l with H<sub>2</sub>O and incubated at 25°C for 20 mins. Following this, 1 $\mu$ l of the labelled oligonucleotide was added and the reaction was incubated for an additional 30 minutes at 25°C.

### C. VISUALISING THE EMSA

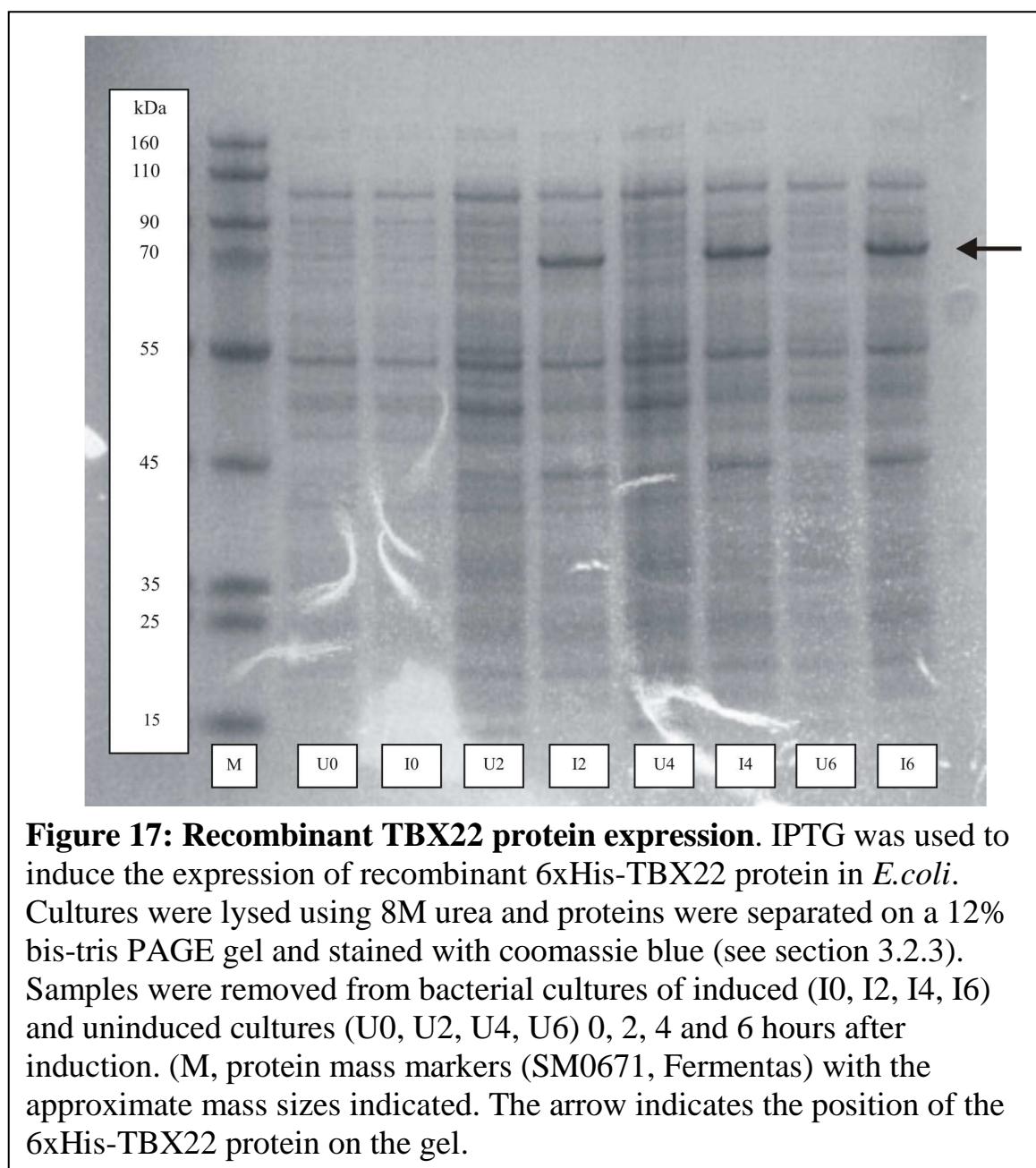
5 $\mu$ l of the reaction from 3.2.5.B was added to 3 $\mu$ l Milli-Q H<sub>2</sub>O and 3 $\mu$ l of Novex Hi-Density TBE sample buffer (Invitrogen) and then loaded on a pre-cast 6% Novex DNA retardation gel (Invitrogen). The samples were electrophoresed at 300V in 0.5X TBE buffer (5X TBE = 5.4% Tris base, 2.75% boric acid, 0.29% EDTA – free acid, pH 8.3; Invitrogen) using the XCell *Surelock* system described in 3.2.3.A, until the marker dye had migrated  $\frac{3}{4}$  of the way down the gel.

The gel was removed from the gel plates and transferred to a sheet of 3MM filter paper (Whatman) overlaid with a thin plastic wrapping and dried using a Flowgen gel drier. Once dry the gel was exposed to Biomax MR-1 X-ray film (Kodac) overnight. The X-ray film was then developed using Xograph's Compact X4 machine according to the manufacturer's instructions using Devalex M and Fixaplus reagents from Champion. The developed X-ray film was imaged using an HP Scanjet 4850 Photo Scanner (Hewlett Packard).

### 3.3 RESULTS

#### 3.3.1 Producing a TBX22 protein

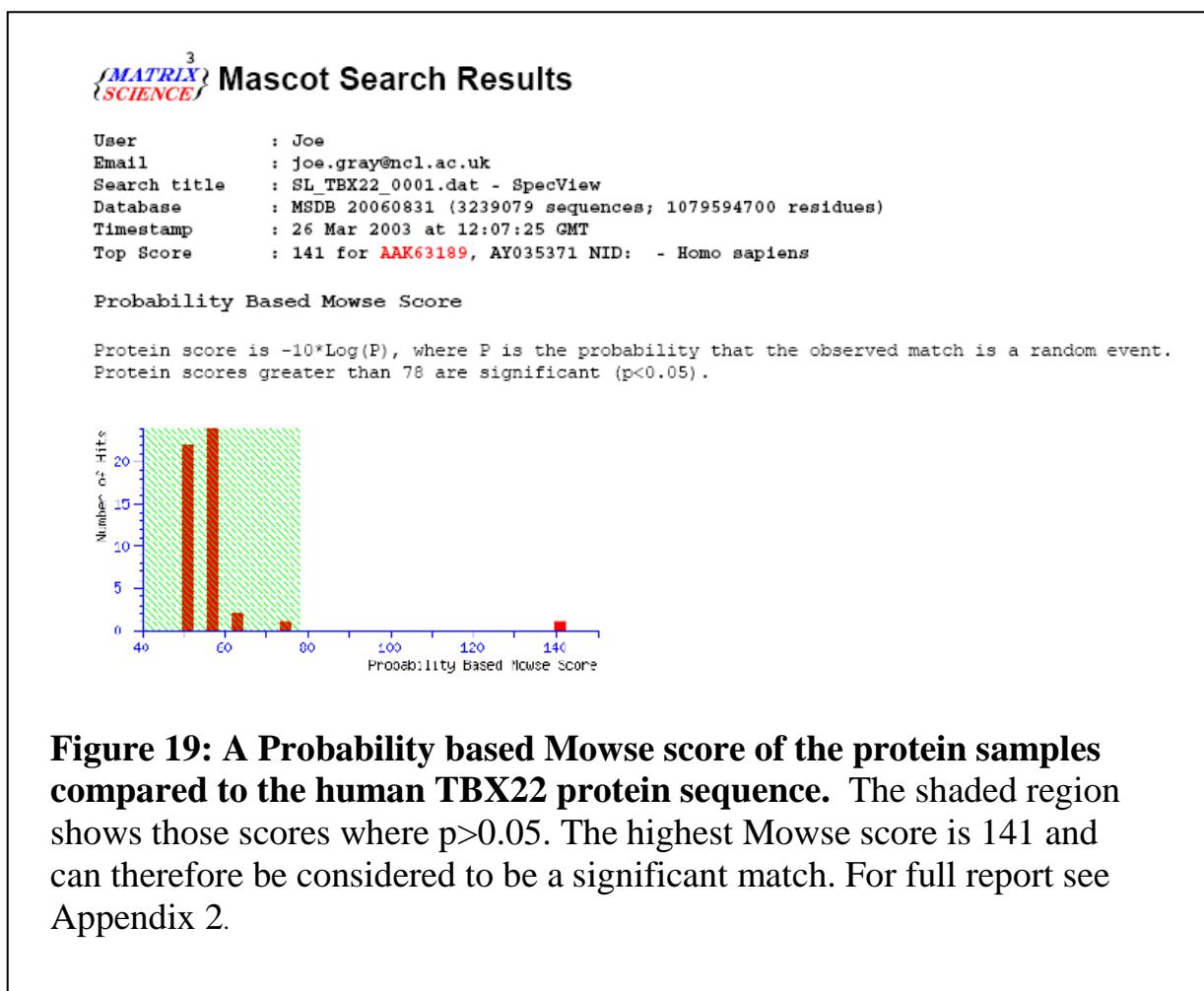
The optimum growth time for the bacterial cultures following induced expression by IPTG was determined by taking samples from induced (I0, I2, I4, I6) and uninduced control cultures (U0, U2, U4, U6) following the addition of IPTG to induce the expression of the recombinant 6xHis-TBX22 protein. Samples were taken from the growing induced and uninduced cultures immediately following the addition of IPTG and then after 2, 4 and 6 hours, lysed using a denaturing lysis buffer containing 8M urea and analysed by PAGE (Fig. 17).



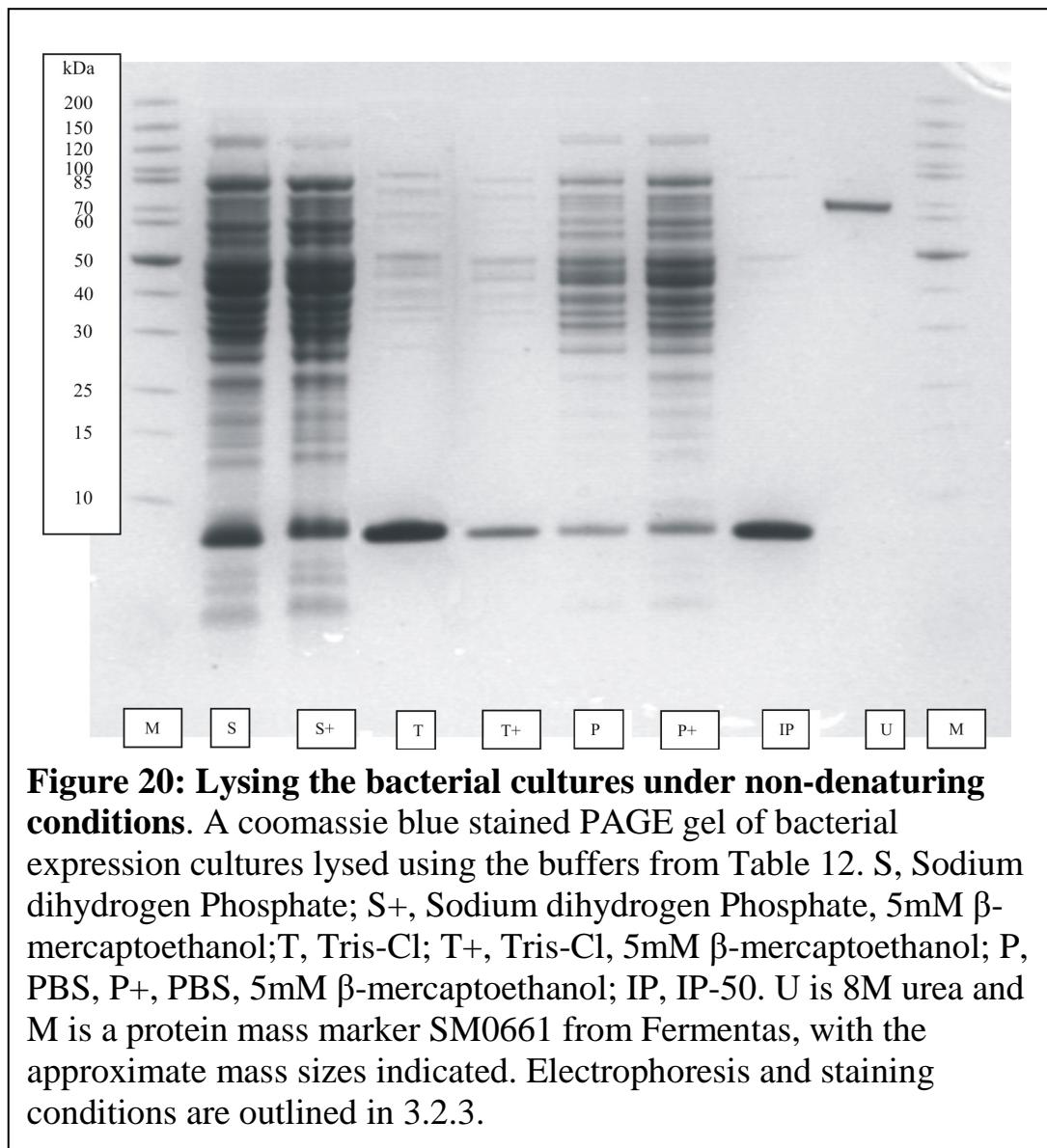
A strong band of approximately 70kDa is present on the PAGE gel in lanes I2, 14 and I6 but not lanes U0, U2, U4 and U6 (Fig. 17). In order to verify that the protein corresponding to this band was indeed the over expressed 6xHis-TBX22 tagged protein; a larger 50ml *6xHis-TBX22* plasmid culture was grown, harvested and purified using a Ni-NTA column (Fig. 18). The single band (lane E2 in Fig. 18) was excised and verified as TBX22 by MALDI-TOF (see 3.2.3.c, Fig. 19 and Appendix 2).



Using the Mowse scoring algorithm (Pappin *et al.* 1993), it was confirmed that the protein in the sample matches the human TBX22 protein sequence. The scores from this test show that the probability of detected peptides matching the human TBX22 protein sequence is much higher than would be expected by chance. A Mowse score of anything greater than 78 indicates that there is less than a 0.05% likelihood of the protein in the sample being the TBX22 protein sequence by chance alone. The highest Mowse score from the output of the MALDI-TOF experiment (Fig. 19, full report in Appendix 2) is 141 and matches to the protein sequence AAK63189, which is human TBX22. All of the other matches lie below the accepted threshold of significance.



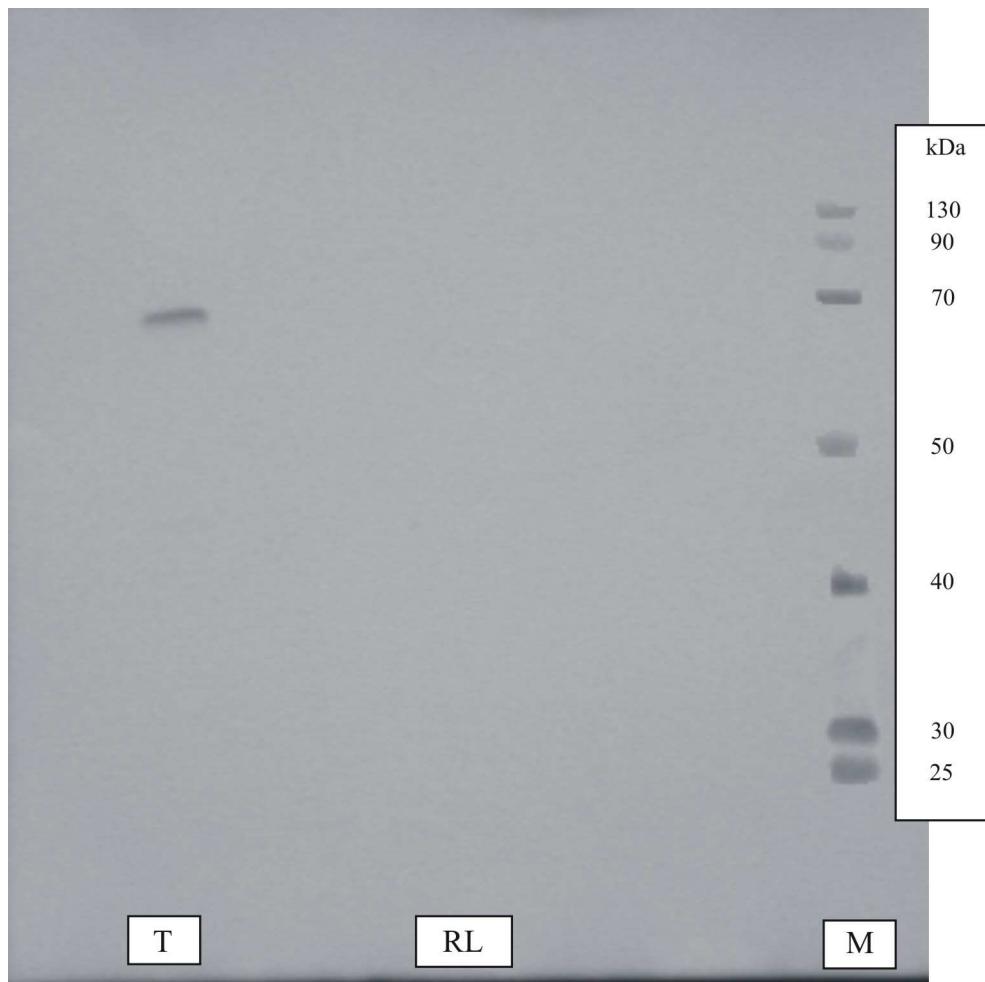
Having established that a human TBX22 protein could be biosynthesised using *E. coli*, a number of different buffers were used to lyse the cells and solubilise the protein under native conditions. Samples were taken from the cell cultures lysed using the different native buffers (see 3.2.1.I) and analysed on a denaturing PAGE gel for the presence of a 70kDa over-expressed protein (Fig. 20).



No band was detected on the PAGE gel from any of the different native buffer lysates employed (lanes S-IP in Fig. 20) corresponding to the over-expressed protein observed in the urea buffer lysate (lane U in Fig. 20). The protein had been produced, as could be seen from the 70kDa band on the PAGE gel from the urea lysed cells, but could not be solubilised using the native buffers and was therefore absent from the PAGE gel in these lanes.

As the TBX22 protein could be detected in the urea lysed cells, the denatured TBX22 protein from this sample was applied to a NI-NTA column; while the protein was attached to the column, urea was gradually removed from the buffer (see 3.2.1.A). However, when the urea concentration in the wash buffer became less than ~3.5M the protein came out of solution and blocked the pores in the spin column. Thus, this approach would not be suitable for the re-folding of the TBX22 recombinant protein.

The 6xHis-TBX22 pET100/D-TOPO plasmid was expressed in a Rabbit Reticulocyte Transcription/Translation System (see 3.2.2). The presence of the 6xHis-TBX22 protein in the cell lysate could be detected on a Western blot using an anti-6xHis tag antibody (see 3.3.3.B). Samples were taken from the rabbit reticulate lysate following the transcription/translation reaction and analysed against a control rabbit reticulate cell lysate that had not been incubated with the 6xHis-TBX22 expression vector (Fig. 21). Using an anti-6xHis antibody, the recombinant protein was detected in the cell lysate that had been incubated with the TBX22 expression plasmid (lane T in Fig. 21) but was absent from the cell lysate only (lane RL in Fig. 21).



**Figure 21: Detection of the 6xHis-TBX22 protein from a rabbit reticulocyte lysate.** An ~70kDa protein was detected by an anti-6xHis antibody by Western blot. Lane T is the lysate incubated with the 6xHis-TBX22 plasmid, RL is the lysate only, M is a pre-stained Protein marker with approximate sizes indicated. Visualisation of the antibody was achieved using HRP and DAB staining (see 3.2.3).

### 3.3.2 Determining a TBX22 preferential DNA binding sequence

Utilising the 6xHis-TBX22 protein produced by the transcription/translation system, an *in vitro* oligonucleotide selection assay using the double stranded random oligonucleotides was performed as described in section 3.2.4. Following five rounds of selection and amplification using the binding site selection assay, the resulting clones containing the oligonucleotide DNA to which TBX22 bound were sequenced. These sequences or their complement were aligned and a consensus sequence

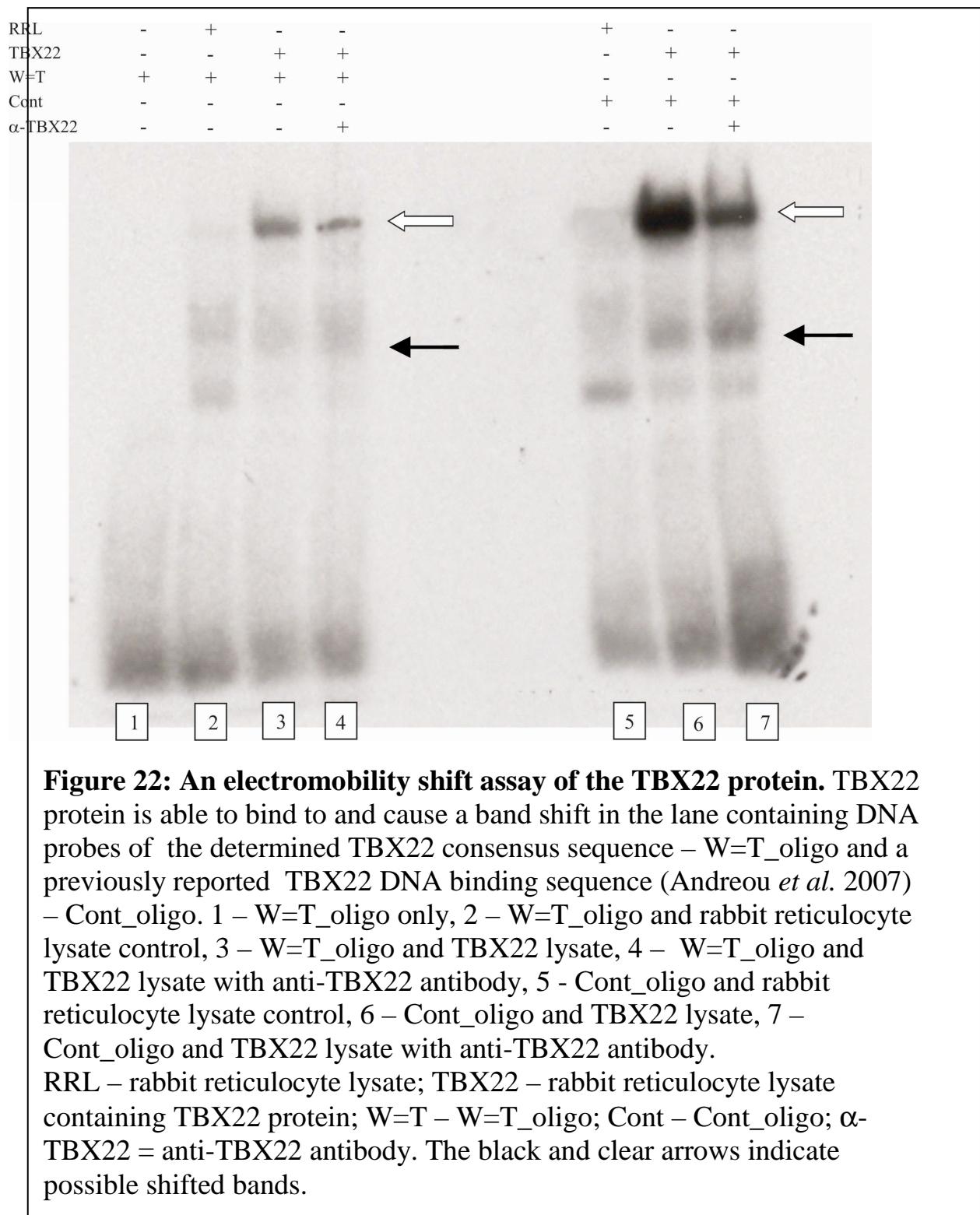
identified (see Table 14). An example chromatogram and complete alignment of the full 26 base random oligonucleotides are presented in Appendix 4.

Clone ID		Sequence									
		1	2	3	4	5	6	7	8	9	10
515-1		C	G	G	T	G	C	G	A	C	C
515-2		C	G	G	T	G	A	T	A	T	G
515-3		A	C	G	T	G	T	C	T	T	A
515-4		T	G	G	T	G	T	T	A	A	C
515-6		A	C	G	T	G	T	C	T	T	A
515-8		G	C	G	T	G	T	C	C	A	A
515-12		A	G	G	T	G	C	T	G	T	C
515-14		A	G	G	T	G	T	C	C	G	A
515-15		A	C	G	T	G	T	C	T	T	A
515-17		T	G	G	T	G	T	G	G	A	A
515-19		A	G	G	T	G	T	C	T	T	A
515-20		A	G	G	T	G	T	G	A	T	C
515-21		A	G	G	T	G	T	C	T	T	A
515-23		A	G	G	T	G	T	C	T	T	A
515-24		A	G	G	T	G	T	C	T	T	A
515-25		A	C	G	T	G	T	C	T	T	A
515-27		A	C	G	T	G	C	A	A	C	A
515-28		A	G	G	T	G	A	C	G	C	A
515-29		C	G	G	T	G	T	G	A	C	G
515-30		A	G	G	T	G	T	G	C	G	T

Base		Nucleotide Frequency									
		14	0	0	0	0	2	1	6	3	13
A		3	6	0	0	0	3	11	3	4	4
C		1	14	20	0	20	0	5	3	2	2
G		2	0	0	20	0	15	3	8	11	1
T											
<b>TBX22 Consensus</b>		A	G	G	T	G	T	C	W	T	A
<b>Brachyury 1/2 site</b>		A	G	G	T	G	T	G	A	A	A

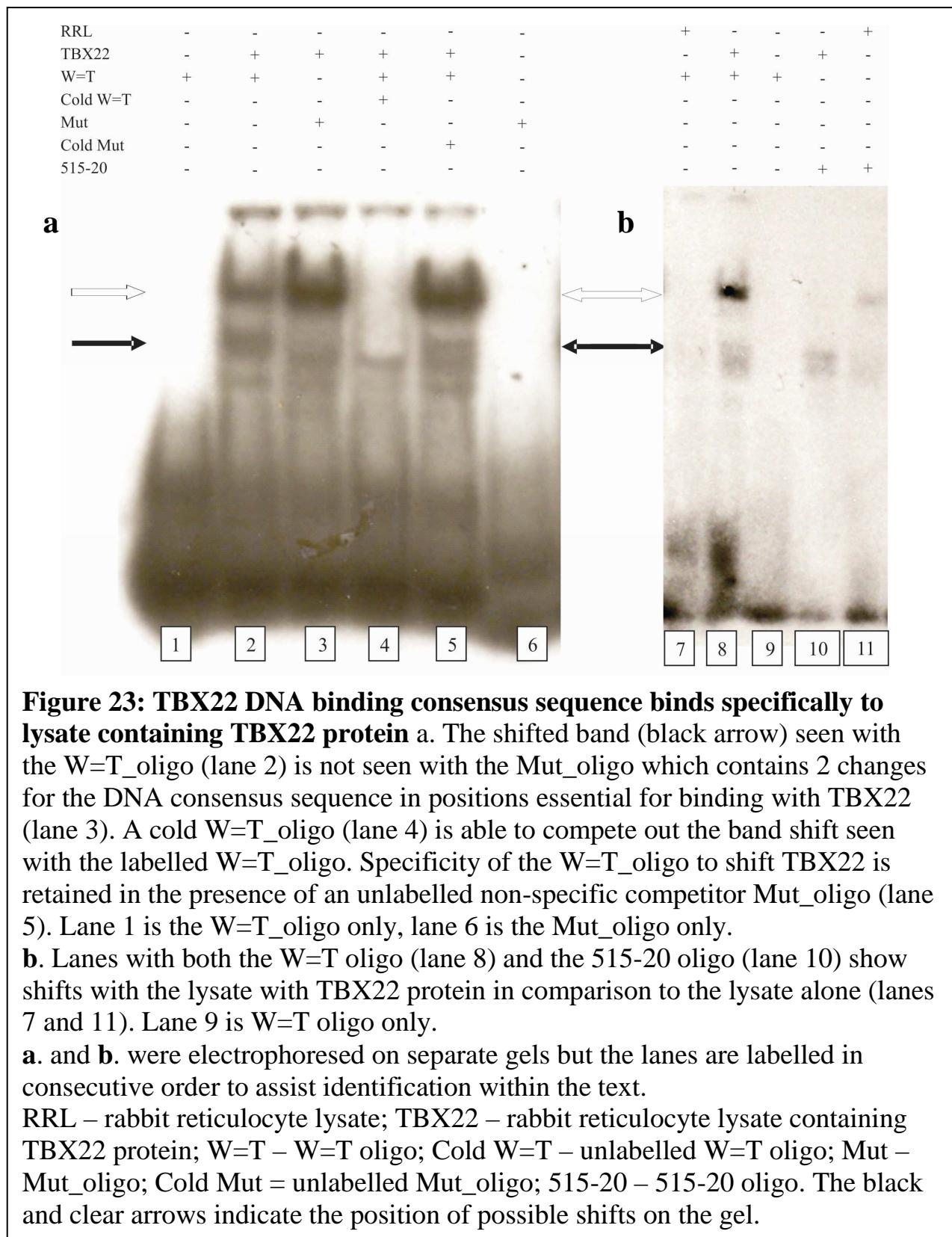
**Table 14: Determining the TBX22 binding site.** The aligned sequences of the cloned DNA isolated from the DNA binding site selection assay following five rounds of selection are shown, with red nucleotides highlighting a direct match with the Brachyury DNA T<sup>1/2</sup>-site. The frequency of each nucleotide at that position in the binding sequence is shown in the lower half of the table. The TBX22 binding site was determined from the consensus sequence of the aligned clones. W = A or T as per the standard IUPAC codes (IUPAC-IUB 1971). The position of each nucleotide within the aligned sequence is shown in the first row of the table.

The consensus TBX22 binding site – AGGTGTCWTA - was identified by aligning the DNA sequences of the 20 clones (Table 14). An electromobility shift assay (EMSA) was used to verify that TBX22 did indeed bind to this consensus DNA binding site (see 3.2.5), using the rabbit reticulocyte cell lysate and the pTNT-TBX22 protein. The resulting EMSA can be seen in Figure 22.



The derived TBX22 DNA binding sequence oligo (W=T\_oligo, lane 3 in Fig. 22) is shown to bind to a protein within the lysate containing the TBX22 protein which results in a shifted band on the gel (position indicated by the black arrow in Fig. 22). A similar shift can be seen in lane 6 of Fig. 22 (Cont\_oligo), which shows the shift by an oligonucleotide that has already been shown to bind TBX22 (Andreou *et al.* 2007) and serves as a positive control for this experiment. The cell lysate that did not contain the over expressed TBX22 protein (lane 2 in Fig. 22) did not display this shifted band illustrating the specificity of the W=T\_oligo for a target site that is present within the cell lysate containing the 6xHis-TBX22 protein and which is not present in the rabbit reticulocyte lysate alone. In addition to this shifted band, both the W=T\_oligo and the Cont\_oligo show a more intense higher molecular weight band on the EMSA (position indicated by clear arrow in Fig. 22) in the presence of the lysate containing the TBX22 protein (lanes 3 and 6 in Fig. 22) when compared to the cell lysate only lane (lanes 2 and 5 in Fig. 22). It is possible that this more intense band is actually caused by binding of the W=T\_oligo and Cont\_oligo to a full length TBX22 protein and that the lower shifted band is due to binding of the oligos to a smaller, incompletely transcribed or translated TBX22 polypeptide. An alternative explanation for the presence of the high molecular weight band is that the W=T\_oligo and Cont\_oligo are binding with additional proteins within the lysate containing the TBX22 protein. Further work would be required to discriminate between these two possibilities.

Further support of the specificity of the W=T\_oligo for a target site within the lysate containing the TBX22 protein is demonstrated by the ability of an unlabelled oligonucleotide, with the same sequence as that of the consensus TBX22 binding site, to completely compete out the shift of the labelled oligo (lane 4 in Fig. 23). The cold W=T\_oligo was also able to remove the higher molecular weight band (position indicated by a clear arrow) seen in lanes 2, 3 and 5.



Specificity is further confirmed by the failure to compete out this shift (lane 5 in Fig. 23a) by an unlabelled mutant oligonucleotide that has an identical sequence to the

consensus TBX22 binding site, except at bases 2 and 3 in which the guanine bases have been replaced with adenine bases. The Mut\_oligo also differs from the W=T\_oligo, but not the actual TBX22 consensus sequence, which can be either an adenine or thymine at the 8th base position, as this base is an adenine base as opposed to the thymine residue of the W=T oligo (see Table 13). Unfortunately the absence of the rabbit reticulocyte lysate only control lane in the first gel of Fig. 23 somewhat hinders direct comparison with the effects of the Mut\_oligo on the lysate containing TBX22 protein (lane 3). However, taken all together, this experiment suggests that a protein within the lysate containing TBX22 can bind to the derived consensus sequence (W=T oligo). As the only known difference between the rabbit reticulocyte lysate and the lysate containing TBX22 protein is the TBX22 protein, this supports the argument that the TBX22 protein can bind to the W=T oligo.

In order to confirm this, a super-shift experiment using an anti-TBX22 antibody was required. However, the anti-TBX22 antibody was unable to super-shift the “TBX22”/W=T\_oligo complex (lane 4 in Fig. 22) or the “TBX22”/Cont\_oligo (lane 7 in Fig. 22). Similarly the anti-His antibody also did not super-shift the complex under the same conditions (data not shown). Whilst the ability of the anti-TBX22 antibody to bind to 6xHis-TBX22 protein has not been tested, the anti-His antibody has been demonstrated to bind to this protein (Fig. 21). If the anti-His antibody had been employed using different experimental conditions, perhaps a super-shift confirming the specificity of the TBX22 DNA binding site to the TBX22 protein would have been achieved. However, in the absence of this data, it cannot be excluded that both the W=T\_oligo and Cont\_oligo are actually binding, and hence shifting, a protein other than the 6xHis-TBX22 protein.

The fact that shifts with the TBX22 DNA binding consensus sequence (W=T oligo) were seen in these experiments partially authenticates the co-immunoprecipitation experiments (3.2.4). As the consensus DNA binding sequence was also the most represented sequence in the list of clones sequenced from the *in vitro* oligonucleotide selection assay: six of the twenty clones sequenced contain a sequence that is identical to the TBX22 consensus binding site, whereas five of the clones had thymine at the eighth position (oligos 515-15, 515-19, 515-21, 515-23 and 515-14 (see Table 14). Further justification of the results of the *in vitro* oligonucleotide selection assay can be

seen in Fig. 23. The 515-20\_oligo, which contained a sequence identified in the *in vitro* oligonucleotide selection assay, but which differed from the TBX22 consensus sequence at two bases, was also capable of causing a shift with the TBXB2 protein (lane 10 in Fig. 23). It appears that the higher molecular weight band (clear arrow) seen with the W=T\_oligo (lane 8 in Fig 23) is not present in the lane containing the 515-20\_oligo (lane 10 in Fig. 23). The absence of this band may reflect a genuine difference in the binding capability of the 515-20\_oligo compared to the W=T\_oligo or Cont\_oligo, or may be due to an artefact of the experiment, perhaps due to radioactive incorporation variances between the different oligonucleotides. If the result reflects a genuine deficiency in the binding of the 515-20\_oligo to a protein within the lysate containing TBX22 then four possible explanations are put forward below:

The 515-20\_oligo binds more specifically to the 6xHis-TBX22 protein than either the W=T\_oligo or Cont\_oligo. If this is the case, then the higher molecular weight band that appears in the lanes with the lysate containing TBX22 protein and the W=T\_oligo and Cont\_oligo (indicated by the clear arrows in Fig. 22 and Fig. 23), could be due to the oligos binding to other proteins in the lysate. This is less likely because the higher bands are not seen with the rabbit reticulocyte lysate alone (compare lanes 2 & 3 and lanes 5 & 6, Fig. 22). However, there is perhaps partial support for this explanation in the fact that the 515-20\_oligo contains a single DNA sequence that binds to the 6xHis-TBX22 protein, as opposed to a consensus sequence composed of several combined sequences for the other two oligos which could account for the non-specificity of these oligos to the 6xHis-TBX22 protein. A second hypothesis is that the W=T\_oligo and Cont\_oligo are binding to the 6xHis-TBX22 protein which is complexed to other biomolecules, thus increasing the molecular weight and showing as a higher band on the EMSA. If this complex is changing the conformation of the 6xHis-TBX22 in such a way that the 515-20\_oligo can no longer bind, no corresponding band would be seen on an EMSA and only the lower shifted band (indicated by the black arrow) would be visible due to the binding to the “free” 6xHis-TBX22 protein. If either of these two possible explanations are correct, then this would give further credence that the shift in Figs. 22 and 23 indicated by the black arrow is due to binding of the TBX22 protein to the different oligonucleotides (W=T; Cont\_Oligo; 515-20).

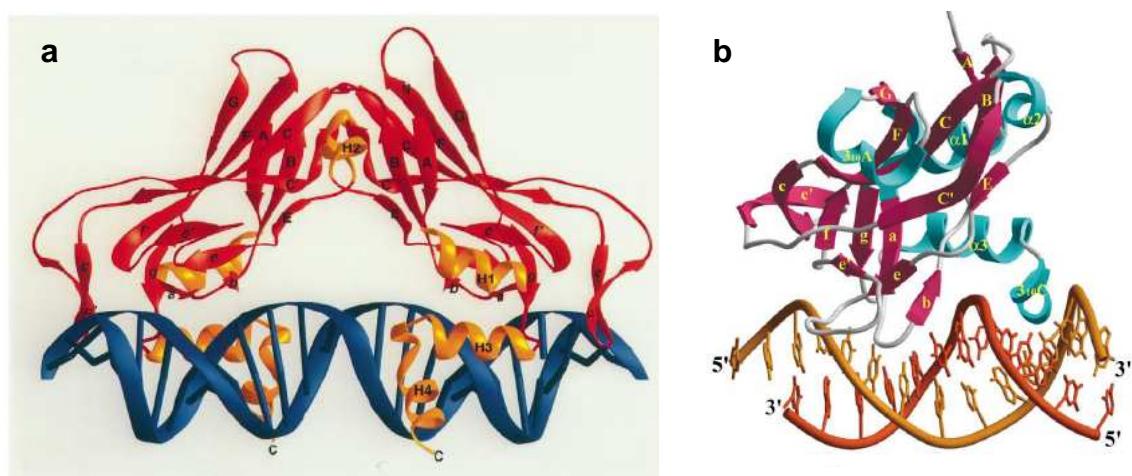
A third possible reason to explain an absence of a higher band in the 515-20\_oligo lane may be due to the fact that both the W=T\_oligo and Cont\_oligo are able to concatamerize and then bind multiple copies of TBX22 – which appears as a higher band on a gel - which the 515-20\_oligo is not capable of doing. A fourth explanation for this missing band may be that, as has been previously stated, it cannot be definitely concluded that the oligonucleotides are actually binding to the TBX22 protein as a super-shift was not detected with the antibodies and conditions tested. It could be argued then, that the different bands detected in the EMSA experiments when using the various oligonucleotides are not in fact due to binding with 6xHis-TBX22 protein, but result from DNA binding to very different proteins within the rabbit reticulocyte cell lysate.

### 3.4 DISCUSSION

X-ray crystallography has shown how T-box proteins bind to their DNA recognition sites (Muller and Herrmann 1997; Coll *et al.* 2002). Contact is made at specific sites between the T-box protein and the DNA molecule (Fig. 24), although X-ray crystallography studies have yet to detail the exact mechanism for the TBX22 protein and the DNA molecule.

The T-box protein binds to the DNA as either a dimer, with specific interaction between each dimer, as is the case with the Brachyury protein (Fig. 24a) or as a single monomer, as is the case with TBX3 (Fig. 24b). In either case, if the protein was folded such that these recognition sites were either not accessible to the DNA or the structure of the protein was physically altered, protein binding to the DNA would not be possible. Therefore it is essential that any production of recombinant protein that will be used for functional studies preserves the native protein structure.

One of the greatest challenges in isolating a relatively pure synthetic protein is maintaining the native protein conformation or where this can not be achieved easily, the re-folding of a denatured protein back to its native state.



**Figure 24: T-box proteins contact DNA at specific recognition sites.** X-ray crystallographic images of a) the *Xenopus* T protein (Muller and Herrmann 1997) and b) human TBX3 (Coll *et al.* 2002) bound to DNA.

Expression in *E.coli* has been successfully employed as a strategy for the production of other functionally active T-box proteins (Sinha *et al.* 2000; Ghosh *et al.* 2001) and was an obvious approach to produce TBX22. The addition of a fusion tag to the protein not only aids the purification (especially in the absence of a specific antibody to the protein of interest), but also can help to stabilise the expression of recombinant proteins and may also assist in re-folding the protein if required. The two most common fusion tags are glutathione S-transferase (GST) and 6 histidine residues (6xHis). 6xHis was chosen as the preferred tag as it has the advantage of being a small tag and therefore being less likely to interfere with the structure and function of the TBX22 protein and can be purified in a one-step recovery process (Hochuli 1988).

Induction of the 6xHis-TBX22 expression vector by IPTG drove the production of an over-expressed additional protein, absent in the uninduced cultures, which could clearly be seen in the cell lysates on a coomassie stained denaturing PAGE gel (Fig. 17). This band migrates to just below 70kDa on the PAGE gel when compared with the marker protein ladder. However, the predicted molecular weight of the 6xHis-TBX22 fusion protein using the online protein molecular weight prediction tool hosted by the bioinformatics organisation ([http://www.bioinformatics.org/sms/prot\\_mw.html](http://www.bioinformatics.org/sms/prot_mw.html)) is 62.05 kDa (the TBX22 protein and DNA sequences can be seen in Appendix 3). This molecular weight prediction does not account for any post-translational modification of the protein that may occur and therefore taking this into account the over-expressed band in the induced cultures is migrating to approximately the expected position on the PAGE gel.

Following Ni-NTA spin column purification, the over-expressed band seen in the induced cell lysate (Fig. 17) was isolated and appeared as a single band on a denaturing PAGE gel (Fig. 18), thus verifying that the over-expressed protein contained a 6xHis tag that was accessible by the Ni-NTA column and was therefore likely to be accessible to an anti-6xHis antibody in future experiments. MALDI-TOF mass fingerprinting analysis of the excised band after Ni-NTA purification revealed that the protein was unquestionably that of human TBX22 (Fig. 19 and Appendix 2). All of the other matches in the sample fell below the threshold level of significance showing that the 6xHis-TBX22 protein had been synthesised and successfully purified.

Although a 6xHis-TBX22 protein had been successfully produced and recovered by over-expression in bacterial cells and purification on a Ni-NTA column, doing so under non-denaturing conditions could not be achieved easily (Fig. 20). The 6xHis-TBX22 protein could not be solubilised using many different lysis buffers (Table 12). This is a common problem when producing recombinant proteins in this manner (Casey *et al.* 1998) as the over-expressed protein is often accumulated as insoluble protein aggregates known as inclusion bodies within the host bacterial cells (Villaverde and Carrio 2003; Ventura and Villaverde 2006). Whilst it is theoretically possible to re-fold the denatured protein once it is purified, it is not always achievable. Attempts to re-fold the 6xHis-TBX22 protein whilst bound to the NI-NTA column, a method used to re-fold a 6xHis-TBX5 protein (Ghosh *et al.* 2001), proved unsuccessful. Once the urea in the buffer solution fell below ~ 3.5M the protein was no-longer soluble and came out of solution, blocking the column. Another approach was investigated, partly due to the difficulties in re-folding the denatured protein and partly because re-folding a denatured protein does not always return the protein to its true native state (Shortle 1996).

An *in vitro* transcription/translation (TNT) system (see 3.2.2) was able to express the same plasmid used in the bacterial cell expression studies. This method uses pre-lysed cell extracts, meaning the recovery of the protein is not needed in applications where the whole cell lysate can be used. This is an appealing approach to use as it eliminates the need to solubilise or re-fold the protein. However, due to the limitations of the TNT technology, a much reduced protein yield is obtained. This has the consequence that there is insufficient sample to either see a band on a coomassie stained gel or, more importantly to use in a MALDI-TOF mass fingerprint experiment. Nevertheless, by using an anti-6xHis antibody, a protein can be detected by Western blot (lane T in Fig. 21). This protein is not present in the cell lysate that does not contain the 6xHis-TBX22 pET100/D-TOP0 expression vector (lane RL in Fig. 21), showing that this protein is being expressed by the plasmid in the TNT cell lysate system. This protein shows a similar migration when compared to a protein marker by Western blot following denaturing PAGE (Fig. 21) to the migration seen

by the 6xHis-TBX22 protein expressed by the bacterial cells when run on a denaturing PAGE gel (figs. 17 and 18).

Is the protein detected by Western blot (Fig. 21) the 6xHis-TBX22 protein, and is the protein in its native state? The evidence would suggest that both are likely to be the case. Although this protein cannot be verified by MALDI-TOF analysis due to the low sample yield, it is being expressed from the same plasmid that was used in the bacterial expression system and that protein was verified by MALDI-TOF analysis to be 6xHis-TBX22. As the protein synthesised by the TNT method was detected using an anti-6xHis antibody, one can be assured that the N-terminal 6xHis tag is present. Furthermore, the over-expressed proteins generated in both bacterial and mammalian cells have similar sizes as shown by their migrating to the same extent on a PAGE gel. As the proteins expressed by the TNT system are produced in a pre cell lysed environment and if the whole cell lysate can be used, no additional recovery process will be required which will make it more likely that the protein will retain its native conformation. Taken together there is satisfactory support that the protein produced by the TNT system is 6xHis-TBX22 protein and will be suitable for use in further functional studies, provided that using the whole cell lysate -which will include the 6xHis-TBX22 protein – will not compromise such studies.

There are additional benefits to using the TNT system over the bacterial expression method. As the TNT system employs mammalian cells, the post-translation protein modifications produced by this approach will be closer to the *in vivo* TBX22 protein than could be achieved by the prokaryote bacterial system, which has different post-translation mechanisms.

Using the protein expressed by the TNT system; an *in vitro* oligonucleotide selection assay yielded a TBX22 preferential DNA binding sequence: AGGTGTCWTA, where W is either an adenine or thymine (Table 14). The binding specificity of TBX22 to this sequence was confirmed when cold specific competitor reduced the binding of TBX22 to the DNA consensus oligonucleotide (lane 4 in Fig. 23a) and TBX22 was shown not to bind to an oligonucleotide with a similar sequence that differed at 2 key bases (Mut\_oligo, lane 3 in Fig. 23a). Although it was not possible to show a super-shift using a TBX22 antibody (lane 4 Fig. 22), this may be because the antibody used

was not actually binding to the protein at all. Antibodies are notoriously fragile in their ability to bind to their ligand when used under different environmental conditions. Although the antibody used in this study had been specifically designed for use in gel-shift experiments, the manufacturer could not guarantee that it would successfully bind to a TBX22 protein produced using a Rabbit Reticulocyte Transcription/Translation System or in cell lysate conditions. If an antibody that could bind TBX22 under the conditions used in this study were available, then a super-shift of consensus DNA binding sequence may well be observed. As no super-shift could be demonstrated in an EMSA experiment, the band shifts witnessed in the EMSA experiments could possibly be caused by binding of the oligos to proteins other than the TBX22 protein. Although an anti-His antibody demonstrated that a protein containing a 6xHis tag was being synthesised by the TNT system (Fig. 21), due to the lack of a super-shift even using this same antibody in an EMSA (results not shown), it is possible that the shifted bands may be due an experimental artefact.

The derived TBX22 DNA binding sequence is similar to the Brachyury  $T^{1/2}$ -site, except for slight deviations at the 3' end of the sequence (Table 14). In comparative binding studies three *Xenopus* proteins Xbra, VegT and Eomesodermin whilst all showing binding to the same core Brachyury  $T^{1/2}$ -site, all showed binding to various deviations from this sequence, most frequently at the flanking ends of this motif (Conlon *et al.*, 2001). A study by another group reported that TBX22 preferentially binds to the sequence AGGTGTGAAATTGTCACCT (Andreou *et al.*, 2007), which is orientated such that the second  $T^{1/2}$ -site is inverted compared to the Brachyury T-site. However, the authors found that binding to one  $T^{1/2}$ -site was possible, albeit in the presence of an antibody to help stabilise the protein/DNA complex. It has been previously noted that variations in binding assay conditions can influence T-box binding to DNA: one study reports that Xbra cannot bind to a  $T^{1/2}$ -site (Casey *et al.*, 1998), whereas two separate studies reported that indeed it can (Carreira *et al.*, 1998; Sinha *et al.*, 2000). Sinha and colleagues, point to salt and non-specific competitor concentrations as being a likely explanation for the differences between the reported findings of the three binding studies. These experimental differences could also be an explanation as to why the preferential TBX22 DNA binding site of this study and that of Andreou and colleagues do not exactly match. In addition, the authors of the latter study report that the full length inverted palindromic DNA binding sequence they

have derived for TBX22 is a sequence that is not naturally present in the human genome and propose it is more likely that the *in vivo* DNA binding site is a half site (Andreou *et al.*, 2007). Therefore, both that study and the data presented here suggest that the DNA binding element of potential target genes for TBX22 is likely to contain a sequence similar to a Brachyury T<sup>1/2</sup>-site.

The identified TBX22 binding sequence is similar to one half of the Brachyury DNA binding sequence, except that the seventh position, which is a guanine residue in the Brachyury target sequence, is replaced by a cytosine base in the TBX22 binding sequence. Also at the eighth and ninth positions, where the Brachyury protein preferentially binds to an adenine base, TBX22 seems capable of binding to either an adenine or a thymine residue at the eighth position and has a preference for a thymine residue at the ninth position (see Table 14).

## CHAPTER 4

### TBX22 DOWNSTREAM TARGETS

#### 4.1 INTRODUCTION

The *in vitro* techniques described in Chapter 3 identified a TBX22 binding sequence which can be used to search sequences of potential downstream target genes. The results of these searches can then be experimentally verified to reveal genuine T-box target genes. Such an approach has uncovered several Brachyury targets (Casey *et al.* 1998; Casey *et al.* 1999) as well as for other T-box proteins (Ghosh *et al.* 2001; Garnett *et al.* 2009).

Even after the identification of potential downstream target genes, validating them can still be an arduous task. Brachyury was the first T-box protein for which a preferential DNA binding sequence was identified (Kispert and Herrmann 1993) but it took a further five years before a downstream target gene was reported (Casey *et al.* 1998). Therefore, if TBX22 target genes were to be identified, careful consideration would have to be given to how the search was performed and also to exactly which genes would be included in the screen.

Although a DNA binding consensus sequence that was similar to the Brachyury T<sup>1/2</sup>-site had been determined for TBX22 (Chapter 3), it did not necessarily mean that this sequence would be the exact *in vitro* target for TBX22. For example, it had been shown that although the original T protein, Brachyury, would preferentially bind to the sequence – TTTCACACCTAGGTGTGAA, the DNA binding site in an *in vivo* target for the *Xenopus* Brachyury homologue Xbra, was – ATTCACACGT (Tada *et al.*, 1998). This sequence, whilst being very similar to the Brachyury T<sup>1/2</sup>-site, was not identical, showing that the actual *in vivo* T-box targets may show some degree of divergence from the preferential DNA binding site determined by an *in vitro* oligonucleotide selection assay.

A search strategy is needed that balances using a search sequence that is flexible enough not to miss potential targets, and using a sequence that is too encompassing which would lead to excessive numbers of matches to target sequences, potentially generating many false positives.

With this in mind, a more generic binding site than the consensus TBX22 DNA binding site was constructed by compiling the binding sequences of other known T-box target sequences and incorporating the preferential TBX22 binding site determined in Chapter 3. This sequence is referred to as the “generic T-box binding site”. The promiscuity of different T-box proteins to their DNA target sequences has been demonstrated several times (for review see Tada and Smith 2001). Indeed all T-box proteins tested have been reported to bind to the *Brachyury* target sequence, even if the experimentally derived DNA binding sequence and the preferred *Brachyury* DNA binding site were not exactly alike (Sinha *et al.* 2000; Conlon *et al.* 2001).

As *TBX22* is disrupted in CPX (Braybrook *et al.* 2001; Marcano *et al.* 2004; Chaabouni *et al.* 2005; Suphapeetiporn *et al.* 2007; Pauws *et al.* 2009b), one could reasonably assume that when TBX22 downstream target genes are disrupted, they too may give rise to a similar cleft palate phenotype. Therefore, all genes shown to underlie a cleft palate disorder would be candidates in which to screen for the presence of the identified preferential TBX22 binding site and the generic T-box binding site. Following the identification of a potential TBX22 target by such a screen, further evidence would be needed to determine whether the gene could be a TBX22 target *in vivo*: information gained from gene expression studies, animal models, gene function and involvement in human disease will be invaluable in supporting, or indeed dismissing, any of the genes identified from the screening procedure as having the potential to be an actual TBX22 downstream target.

## 4.2 MATERIALS AND METHODS

### 4.2.1 Electrophoretic Mobility Shift Assay of an *MSX1* oligonucleotide

An EMSA was performed in the same way as detailed in 3.2.5, with an additional double-stranded oligonucleotide containing the exact sequence found in the *MSX1* promoter (MSX1\_oligo). This was prepared by annealing the oligonucleotides 5'-CGAGAGGTGTTGAGCGAG-3' and 5'-CTCGCTAACACACT CTCG-3', which was 5' end labelled using [ $\gamma$ -<sup>32</sup>P]ATP as in 3.2.5.A. For competition studies, a cold MSX1\_oligo was also prepared by annealing the same oligonucleotides, but was not 5' end labelled. The cold MSX1\_oligo was used in an excess ratio of 100:1 with the labelled MSX1\_oligo added to the binding reaction (see 3.2.5). The sequences of the MSX1\_oligo and the W=T\_oligo also used are given along with all the other double-stranded oligonucleotide and PCR primers are given in Appendix 1.

### 4.2.2 Non-radioactive *in situ* Hybridisation

Paraffin sections of human embryonic tissue were used to determine the expression pattern of *MSX1*. The material was prepared in the same manner as outlined in 2.2.1.

#### A. DNA TEMPLATE PREPARATION AND RNA TRANSCRIPTION

Forward and reverse primers for *MSX1* containing a T7 RNA polymerase sequence and an SP6 RNA polymerase sequence respectively (see Appendix 1 for primer sequence) were used to generate a DNA template by PCR (see 3.2.1.B) using 25 cycles and an annealing temperature of 62°C. The resulting amplicon was extracted from a 1% agarose gel and purified by elution through a gel extraction column (Qiagen) following the manufacturer's instructions.

Using a DIG-labelling kit (Roche Applied-Science) sense and antisense riboprobes were synthesised using T7 and SP6 RNA polymerases respectively from the purified DNA template following the manufacturer's directions, except 75ng of PCR fragment, rather than 1 $\mu$ g of plasmid DNA was used in the reaction.

## B. PROBE VERIFICATION AND QUANTIFICATION

All probe concentrations were confirmed using a nanodrop instrument (Thermo) and the probe length was verified by electrophoresis on an agarose-formamide gel. Here, the agarose gel is prepared by adding 1g agarose to 41ml of DEPC treated H<sub>2</sub>O and dissolved by heating in a microwave and re-adjusted to 41ml. 9ml 37% formamide (Merk) is then added to the cooled mixture and the gel is poured into a casting mould containing a comb. RNA samples and an RNA ladder are heated to 70°C cooled on ice and RNA loading buffer (Fermentas) is added to the samples to a final concentration of 1X. 6µl of ladder and sample are electrophoresed on the gel in 1X MOPS-EDTA (Sigma) running buffer for 2 hours at 50V. The gel is then removed from the tank and cassette and transferred to SYBR Green (Molecular Probes) for 15 minutes with agitation, rinsed in H<sub>2</sub>O and the migration of the samples compared to the RNA ladder under UV illumination.

## C. PROBE HYBRIDISATION

All solutions used were made using DEPC treated H<sub>2</sub>O or PBS and all glassware used to contain these solutions was baked at 180°C for 4 hours before use to denature RNases. To remove the paraffin wax the slides containing the sectioned material were transferred to metal racks and taken through three 5 minute changes of xylene. The xylene was then removed from the sections by 3 minute washes in 1:1 xylene/ethanol, two changes in 100% and then 90%, 70% and 50% ethanol washes. The ethanol was removed from the sections by two separate 2 minute washes in PBS and then the sections were subjected to proteolytic digestion by 20µg/ml Proteinase K (Sigma) in PBS at 37°C. The sections were then rinsed for 30 seconds in PBS and then fixed in 4% PFA for 20 minutes at room temperature. Following fixation, the sections were rinsed in 2 changes of PBS and then were placed in a 0.1M triethanolamine pH 8.0, 0.25% acetic anhydride/PBS solution for 10 minutes. The sections were then given two more rinses in PBS before being dehydrated in 50%, 70%, 90% and finally two 100% ethanol washes for 2 mins each. The sections were then air dried under a filtered air stream.

A hybridisation mix is prepared by first heating the labelled probe (4.2.2.A) to 70°C for five minutes and then cooling on ice. For each slide, 100µl Dig Easy Hyb Mix (Roche Applied Sciences) containing 300ng of the prepared RNA probe is used to cover the slide and a glass cover-slip is placed carefully on top. The slides are then placed inside plastic trays and the plastic trays placed inside a hybridisation chamber with a paper towel soaked in 2X SSC. The hybridisation is performed overnight at 68°C.

#### D. POST HYBRIDISATION WASHES

Following hybridisation, the slides were taken out of the hybridisation oven and removed from the trays. The cover-slips were removed by rinsing the slides in 5X SSC, pre warmed to 60°C, and the slides placed in plastic slide racks. The slides were then washed in two 10 minute washes of 5X SSC at 60°C, followed two 10 minute washes in 2X SSC first at 60°C and then at room temperature.

#### E. ANTIBODY DETECTION

The slides were removed from the hybridisation wash and given three 10 minute washes in wash buffer (0.1M Tris (pH 7.6), 0.15M NaCl). The slides were then covered in a blocking solution of 10% fetal calf serum (Sigma; previously heat inactivated at 58°C for 30mins) diluted in the antibody detection buffer (0.1M Tris (pH 7.6), 0.15M NaCl, 2% fetal calf serum) and left for 1 hour at room temperature. Next, the blocking solution was poured from the slides and 150µl anti-DIG antibody (Roche Applied-Science) diluted 1:1000 in 2% fetal calf serum/antibody detection buffer was placed directly onto the slide and covered with pieces of laboratory parafilm (Pechiney Plastic Packaging Company). The slides were then incubated overnight at 4°C.

## F. Signal Detection

The parafilm covers were removed from the slides by rinsing in the antibody detection buffer and performing three 10 minute washes in this buffer. The slides were then transferred to a signal detection buffer (0.1M Tris (pH 9.5), 0.1M NaCl) and three further 5 minute washes were performed in this buffer. NBT/BCIP (Roche Applied Science) was diluted to 20 $\mu$ l/ml in the signal detection buffer and the reaction left to develop in the dark overnight. The following day the slides were rinsed in several changes of H<sub>2</sub>O and mounted using Aquamount (VWR) and cover slips added. Images were taken as described in 2.2.2.G

## 4.3 RESULTS

### 4.3.1 Computational Search for Potential TBX22 Target Genes

#### A. GENERATING A GENERIC T-BOX BINDING SITE

In an attempt not to exclude possible *bona fide* TBX22 downstream target genes that did not exactly match the derived TBX22 DNA binding site, a generic T-box binding site was compiled by combining the TBX22 DNA binding consensus sequence (Chapter 3) with the sequences, or reverse complement of T-box half sites with which it showed similarity (see Table 15 below and Table 10, Chapter 3). All of the nucleotides that occurred more than once in the same position when these sequences were aligned were included in the generic T-box DNA binding sequence. However, in the binding sequences identified for human TBX5 (Ghosh *et al.* 2001) and mouse Tbx6 (White and Chapman 2005) any nucleotide may be positioned at the final two nucleotide positions (RGGTGTBRNN and AGGTGTBRNN respectively).

Incorporating N at positions 9 and 10 in the generic T-box binding site would effectively have given a search sequence of just eight nucleotides. Therefore, the information from human TBX5 and mouse Tbx6 at positions 9 and 10 was not included when compiling the sequences used to determine the generic T-box binding site.

This gave a final generic T-box binding sequence of AGGTGTBDWR, which provided a more flexible binding site target sequence than the identified TBX22 binding site alone (Chapter 3), but one that should still be specific enough not to generate large numbers of false positive results.

Species	Protein	DNA Binding Site										References
Human	Brachyury	A	G	G	T	G	T	G	A	A	A	(Papapetrou <i>et al.</i> 1997)
Human	TBX1	A	G	G	T	G	T	G	A	A	A	(Sinha <i>et al.</i> 2000)
Human	TBX2	A	G	G	T	G	T	G	A	A	A	(Sinha <i>et al.</i> 2000; Lingbeek <i>et al.</i> 2002)
Human	TBX3	T	G	G	T	G	C	C	A	A	A	(Lingbeek <i>et al.</i> 2002)
Human	TBX5	R	G	G	T	G	T	B	R			(Ghosh <i>et al.</i> 2001)
Human	TBX22	A	G	G	T	G	T	G	A	A	A	(Andreou <i>et al.</i> 2007)
Mouse	Brachyury	A	G	G	T	G	T	G	A	A	A	(Kispert <i>et al.</i> 1995)
Mouse	Tbx6	A	G	G	T	G	T	B	R			(White and Chapman 2005)
Xenopus	VegT	A	G	G	T	G	T	C	T	T	G	(Casey <i>et al.</i> 1999)
Xenopus	VegT	A	G	G	T	G	T	G	A	A	G	(Hyde and Old 2000)
Xenopus	Xbra	A	G	G	T	G	T	G	A	A	A	(Casey <i>et al.</i> 1998)
Xenopus	Xbra	A	G	G	T	G	T	C	T	T	G	(Casey <i>et al.</i> 1999)
Ciona	Ci-Bra	A	G	G	A	G	G	T	G	C	C	(Di Gregorio and Levine 1999)
Human	TBX22	A	G	G	T	G	T	C	W	T	A	See Chapter 3
Generic T-box Binding Site		A	G	G	T	G	T	B	D	W	R	

**Table 15: Sequences of T-box protein DNA binding sites included when determining the generic T-box binding site.** The DNA binding sequences of the different T-box proteins shown in this table include both the preferred *in vitro* DNA binding sites and known *in vivo* target sequences. Standard IUPAC abbreviations have been used to represent ambiguous bases (see Table 11, Chapter 3). One of the half-sites of the TBX22 binding site identified by Andreou *et al.* 2007 is shown. The other has a similar sequence but is in the opposite orientation.

## B. SEARCHING HUMAN CLEFT PALATE GENES FOR A GENERIC T-BOX BINDING SITE

Single genes that had been shown to underlie a disease phenotype that included a cleft palate were identified from OMIM (Online Mendelian Inheritance in Man; <http://www.ncbi.nlm.nih.gov/omim/>). Genes shown to be implicated in a cleft palate phenotype by linkage analysis, but were not specifically shown to be disrupted in an individual with a cleft palate, were not included. Similarly, instances where large DNA regions or chromosome abnormalities were implicated in cleft palate disorders were also not included. Appendix 5 gives the full list of genes screened. As of the tenth of February 2010, data mining the OMIM database revealed 132 genes that had been shown to cause a human genetic abnormality that included a cleft palate (first column of Appendix 5).

Having compiled a list of known human cleft palate causing genes, the 2kb sequence upstream from the start of transcription of each of the genes identified was used to search for the presence of the TBX22 consensus sequence and the generic T-box binding sequence. 2kb were selected as the size of the sequence to screen, following the protocol of a previous T-box gene study (Ghosh *et al.* 2001). The DNA sequence used for the search and also the position of the start of transcription was based upon that given by the gene reference sequence archived in the RefSeq depository at the NCBI (<http://www.ncbi.nlm.nih.gov/RefSeq/>). The current RefSeq sequence (Pruitt *et al.* 2007) of these genes and the position of the gene within the GenBank accession numbers (Benson *et al.* 2008) used are highlighted in the second column of Appendix 5. Entries with a suffix “complement” denote that the gene is transcribed from the other strand to that given by the GenBank accession number stated. The third column of Appendix 5 shows the 2kb sequence which was searched for the presence of the derived TBX22 DNA binding consensus sequence (AGGTGTCWTA) and the generic T-box binding site (AGGTGTBDWR).

A web based version of Fuzznuc, housed by Anabench online molecular biology tools (<http://anabench.bcm.umontreal.ca/anabench/index.jsp>), was used to perform the search for the presence of potential *in vitro* binding sites in the isolated 2kb region upstream of the cleft palate causing genes. Fuzznuc was developed as part of the EMBOSS package (Rice *et al.* 2000) and is a bioinformatics tool which allows

the user to search longer sequences for the presence of short sequences containing ambiguous nucleotides at the same position in the search sequence.

A screen of the 2kb promoter regions of these 132 identified genes for the presence of the derived TBX22 DNA binding consensus sequence did not uncover any matches. The list included the *TBX22* gene itself, to which TBX22 is known to bind (Andreou *et al.* 2007). A search for the generic T-box binding site - AGGTGTBDWR - was performed both with a limit of no mis-matches and also allowing for 1 mismatch to the search sequence.

Twelve genes were identified as having the generic T-box binding sequence (allowing no mis-matches) present in the 2kb search region (column 4 in Appendix 5 and summarised in Table 16).

Gene	Position from start of transcription	Sequence of match to the generic T-box site and nucleotide position within binding site									
		1	2	3	4	5	6	7	8	9	10
<i>ESCO2</i>	1501-1510	A	G	G	T	G	T	T	A	A	G
<i>FOXC2</i>	124-133	A	G	G	T	G	T	G	G	A	A
<i>FTO</i>	1677-1687	A	G	G	T	G	T	C	T	T	G
<i>GPC3</i>	473-482	A	G	G	T	G	T	T	A	A	G
<i>KIAA1279</i>	196-205	A	G	G	T	G	T	G	G	T	G
<i>MSX1</i>	1166-1175	A	G	G	T	G	T	T	G	T	G
<i>POMT2</i>	1648-1657	A	G	G	T	G	T	C	T	A	A
<i>PTCH2</i>	383-392	A	G	G	T	G	T	G	G	T	G
<i>RAPSN</i>	53-62	A	G	G	T	G	T	G	G	T	G
<i>RPS19</i>	731-720	A	G	G	T	G	T	G	G	T	G
<i>SPINT2</i>	870-879	A	G	G	T	G	T	G	G	T	G
<i>SUMO1</i>	71-80	A	G	G	T	G	T	G	A	A	G
Sequence of generic T-box binding site		A	G	G	T	G	T	B	D	W	R

**Table 16: A list of the human cleft palate genes containing a hit to the generic T-box binding site.** The numbers show the nucleotide position upstream from the start of transcription and the specific sequence at that site is shown. B = C or G or T; D = A or G or T; R = A or G; W = T or A (see Table 11).

However, none of these identified genes had a sequence that matched directly with any of the clones sequenced from the oligonucleotide selection assay (see Table 14). This is perhaps due to the differences between the “perfect” target site identified from every single possible DNA sequence using such an assay under experimental conditions and the true gene target sites that exist *in vivo*. Similarly, none of these sequences were an exact match to the TBX22 DNA binding consensus sequence identified in Chapter 3. *ESCO2* and *POMT2* had the closest matches to this sequence, differing by only one base in each case; the remaining ten genes all differed from the derived TBX22 DNA binding sequence at three bases. However, all of these base differences were outside of the core sequence that is common amongst all of the known T-box DNA binding sequences (see Table 15).

The results when this search was performed allowing one mis-match can be seen in column 5 of Appendix 5. This less stringent search, as may have been expected, produced many hits. Indeed, all but 30 of the 132 genes showed a hit to the search sequence. Several genes showed more than one potential generic T-box binding site and sixteen had three or more sites detected using the 1 mis-match search criterion. These genes are shown in Table 17 and include *TBX22* for which four binding sites were identified.

Further evidence that a gene identified from the *in silico* search might be a TBX22 downstream target was sought from mouse models. The human cleft palate genes identified in 4.3.1.B were used to search the Mouse Genome Database (MGD) database (<http://www.informatics.jax.org>) for the presence of a cleft palate phenotype in the mutant mouse with a disruption in those genes. The MGD (Bult *et al.* 2008; Blake *et al.* 2009) is an electronic depository for the genomic, phenotypic and gene expression information gathered from mutant mouse models for the purpose of investigation of human genetic disease. Of the 132 human cleft palate genes, a mouse model for that gene had been reported in the literature for 108 and of these 108 mutants, 39 were reported to have a cleft palate phenotype (final column, Appendix 5).

Gene	Position from start of transcription	Sequence of 1 mis-match to the generic T-box site
		1 2 3 4 5 6 7 8 9 10
<i>BMP4</i>	617-626	A G G T G T A T A A
	705-714	A A G T G T C A T G
	1390-1399	A G G T G T T C T A
	1583-1592	A G G T G T T T C A
<i>COL2A1</i>	260-269	A T G T G T G G A A
	586-595	A A G T G T C A A A
	927-936	A G G T G A G G T G
	1383-1392	A G G G G T G G A G
	1820-1829	A G G G G T G G A G
<i>EFNB1</i>	178-187	A G G A G T C A A A
	914-923	A G G T C T T G T G
	1239-1248	A G G T G T C T T T
<i>GDF1</i>	355-364	A G G T G T C G A T
	926-935	G G G T G T G G T G
	1599-1608	G G G T G T G A A A
<i>GLI2</i>	674-683	A G G T G T T A C G
	1196-1205	A G C T G T C A T G
	1472-1481	A T G T G T G A T G
<i>HOXA2</i>	61-70	A G C T G T C A A G
	550-559	A G G T G T T A T C
	740-749	A G G T G T G G G A
	1597-1606	G G G T G T C A A A
	1657-1666	A G G T G C T T A A
<i>MKS1</i>	27-36	A G G T A T C T A A
	441-450	A G C T G T T G T G
	930-939	A G G T G G T T A G
	525-534	A G G T C T G T A G
<i>PEX7</i>	26-35	A G G T G G G A T G
	569-578	A G G T G T G A A A C
	1804-1813	A G C T G T C G T A

continued overleaf ...

Gene	Position from start of transcription	Sequence of 1 mis-match to the generic T-box site
		1 2 3 4 5 6 7 8 9 10
<i>PROK2</i>	249-258	A G C T G T G T A A
	615- 624	A C G T G T G A A A
	626-635	A G G T T T T T A A
	669-678	A G C T G T T T A A
<i>PROKR2</i>	834-843	A G G T G T G C T G
	895-904	A G G T G T T G G G
	1406-1415	A G G A G T G G T G
<i>RAPSN</i>	355-364	A G G C G T G G T G
	977-986	A G G T G G G G G A A
	1242-1251	T G G T G T T A A A
	1678-1687	G G G T G T C T T G
	1862-1871	A G G T G C C T A A
	1893-1902	A G G T G G C A A G
<i>RPL5</i>	170-179	A G G T G T G A A C
	1124-1133	A G G T T T T T T A A
	1220-1229	A A G T G T T T T A
<i>SLC26A2</i>	534-543	A G C T G T T G T G
	1015-1024	A G G T G A G G G A A
	1157-1166	T G G T G T T T T G
<i>TBX22</i>	991-1000	A G G G G T T T T A
	1044-1053	C G G T G T G G T A
	1094-1103	A G G T A T G G G T G
	1461-1470	A G G T G T A A A G
<i>TNNT3</i>	631-640	A G G C G T G G A A
	853-852	A G G G G T G G G A A
	1519-1528	C G G T G T C T A G
<i>ZEB2</i>	284-293	A G G T G A G T A A
	1057-1066	A G A T G T G G A A
	1410-1419	A G G T G T A G A G

**Table 17: Human cleft palate genes showing three or more generic T-box binding sites with a one base mis-match.** The position upstream from the start of transcription and the specific sequence at that site is shown along with the nucleotide position of each base.

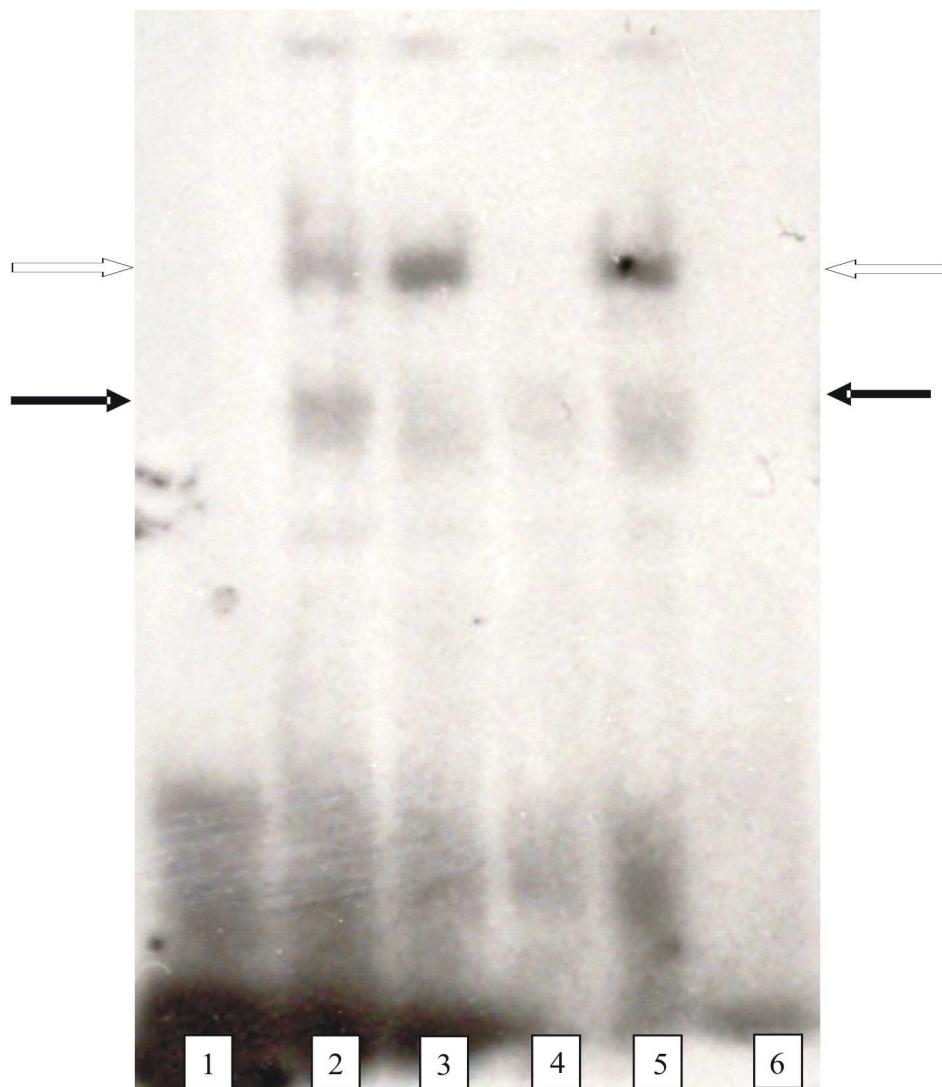
Combining the information from the genes that contained a hit to the generic T-box sequence or to a minimum of three hits to the 1 mis-match sequence within their promoter sequence with the information from the mouse mutants increased the likelihood of identifying biologically relevant candidates for downstream targets of TBX22. Of the 132 identified human cleft palate causing genes, 10 fell into this category and are summarised in Table 18.

GENE
<i>FOXC2</i>
<i>MSX1</i>
<i>SUMO1</i>
<i>COL2A1</i>
<i>EFNB1</i>
<i>GLI2</i>
<i>HOXA2</i>
<i>MKS1</i>
<i>TBX22</i>
<i>ZEB2</i>

**Table 18: The known human cleft palate causing genes that contain either an exact match to the generic T-box binding site, or at least 3 copies of a 1 base mis-match to it, in a 2kb region from the start of transcription and also give rise to a cleft palate phenotype when disrupted in mouse.**

One of the genes from Table 18 - *MSX1* - was then examined further as a candidate TBX22 downstream target. The reasoning behind the decision to examine *MSX1* as a potential candidate and the exclusion of the other genes is expanded upon in the Discussion at the end of this chapter. Briefly, when information from known gene expression patterns, function and available evidence from mouse mutants was examined two genes remained as strong candidates- *MSX1* and *SUMO1*. TBX22 had already been shown to undergo sumoylation (Andreou *et al.* 2007) and so it was decided to pursue *MSX1*. To verify whether the generic T-box binding sequence identified within the *MSX1* promoter (see Table 16) could bind the TBX22 protein, an EMSA (section 3.2.5) was performed using an oligonucleotide containing this sequence – *MSX1*\_oligo (section 4.2.1). A band shift in an EMSA was seen with the *MSX1*\_oligo (lane 2) similar to that seen with the *W=T*\_oligo (lane 5). Furthermore, specificity of the binding of the *MSX1*\_oligo to a protein within the lysate containing the TBX22 protein is confirmed by a cold *MSX1*\_oligo competing out the shifted band (lane 4). This cold competitor can also remove the higher band (indicated by the clear arrow) seen in lanes 2, 3 and 4 as did the *W=T* competitor oligo (see lane 4 Fig. 23a in Chapter 3).

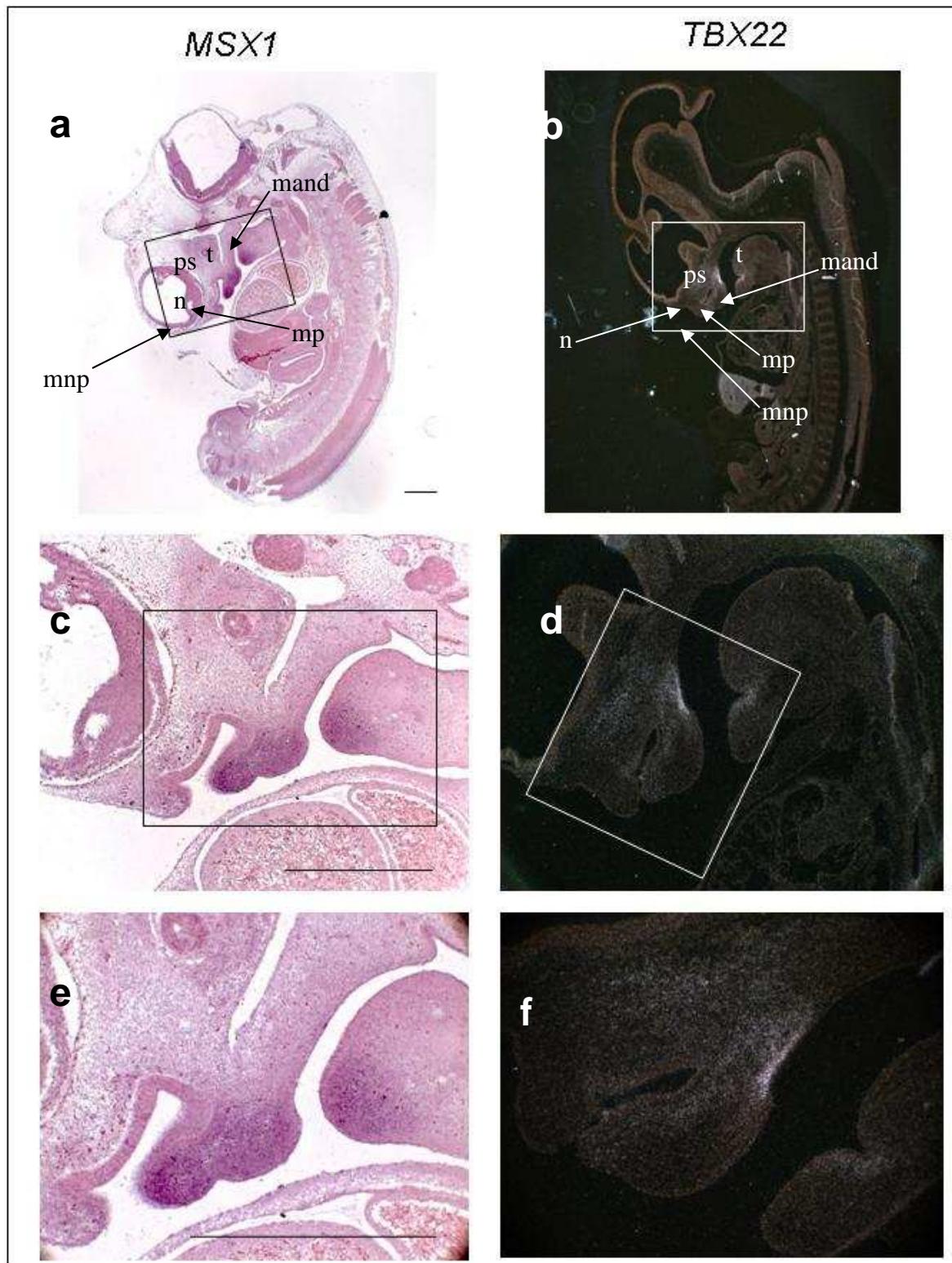
RRL	-	-	+	-	-	-
TBX22	-	+	-	+	+	-
W=T	-	-	-	-	+	+
MSX1	+	+	+	+	-	-
Cold MSX1	-	-	-	+	-	-



**Figure 25: The cell lysate containing TBX22 protein causes a band shift with an oligo to the TBX22 binding site identified by the *in silico* screen of 2kb of the MSX1 promoter (MSX1\_oligo). The MSX1\_oligo (lane 2) and W=T\_oligo (lane 5) bind to a protein in the lysate containing the TBX22 protein and cause a shift on the gel when compared to the rabbit reticulocyte lysate control (lane 3). A cold MSX1\_oligo is able to compete with the labelled MSX1\_oligo and eliminate the shift (lane 4). Lane 1 is the labelled MSX1\_oligo only and lane 6 is the W=T\_oligo only.**

RRL – rabbit reticulocyte lysate; TBX22 – lysate containing the TBX22 protein; W=T – W=T\_oligo; MSX1 – MSX1\_oligo; Cold MSX1 – unlabelled MSX1\_oligo. The black and clear arrows indicate the position of potential shifts.

In common with the EMSA studies in Chapter 3, experiments using ant-TBX22 antibodies did not produce a super-shift (data not shown) and so it could not be confirmed directly that the MSX1\_oligo was indeed binding to the TBX22 protein. However, the results from this EMSA experiment are consistent with the possibility of *MSX1* being a candidate as a downstream TBX22 target gene. The expression of *MSX1* was compared to that of *TBX22* during human palatogenesis.



**Figure 26: mRNA expression of *MSX1* and *TBX22* in the developing palate.** The expression of *MSX1* is revealed by non-radioactive *in situ* hybridisation (a, c and d) and is shown in relation to the expression of *TBX22* by radioactive *in situ* hybridisation (b, d and f). The images show the expression detected by the antisense probe; the sense probe gave no detectable signal for either gene (data not shown). Hybridisation of the antisense probe is to a CS18 (a, c and d) and CS19 (b, d and f) human embryo sectioned in the sagittal plane through comparable regions in both embryos. The rectangles shown on images a, b, c and d depict the approximate area highlighted in the image directly below it.

t – tongue, mand – mandible, n – nasal pit, mnp – medial nasal process, ps – palatal shelf.

Bar =1000µm in that and the image adjacent to it.

From the low power magnification images of the expression of *MSX1* and *TBX22* (**a** and **b** in Fig. 26) it can be seen that both genes have discrete regions of expression in the developing embryonic face. In the higher magnification images (**d** and **f** in Fig. 26) *TBX22*, as was seen in the transverse sections at CS19 (Fig. 11, Chapter 2), is weakly expressed at the base of the tongue. In contrast to this, *MSX1* is expressed more strongly in the mesenchyme of the mandible (**c** and **e** in Fig. 26), as can be seen from the position of the signal in relation to that of *TBX22*.

Stronger *TBX22* expression is seen within the mesenchyme surrounding the nasal pit with the most intense signal being restricted to mesenchyme immediately adjacent to the epithelial cell layer posterior to the nasal pit. The expression of *MSX1* is restricted to the mesenchyme within the most anterior region of the palate. Close comparison of this same region hybridised with the *TBX22* antisense probe in the CS19 embryo shows that *TBX22* is not being expressed (**f** in Fig. 26). Similarly, in the mesenchyme in the region below the forebrain and above the nasal pit, where lower levels of *TBX22* are expressed, no *MSX1* expression is detected. These data indicate that *MSX1* and *TBX22* may be expressed in a complementary pattern within the developing palate and other facial structures; where *TBX22* is detected *MSX1* is absent and vice versa.

The experiments for the two genes have been conducted in different human embryos which are of slightly different ages (CS18- *MSX1* and CS19- *TBX22*) and which have been sectioned at slightly different angles. As both the morphology of the facial structures (see 1.2) and the expression pattern of *TBX22* (see Chapter 2) are changing rapidly during this stage of development, the definitive conclusion that *TBX22* and *MSX1* are indeed expressed in a complementary pattern cannot not be made.

Additionally, the two experiments have been conducted using different techniques – *TBX22* by radioactive *in situ* and *MSX1* using a non-radioactive labelling method. To increase confidence when drawing such a conclusion, the same *in situ* method performed on consecutive sections should be employed. However, to definitively show that *TBX22* protein is present in the same cells, or those adjacent to those that are transcribing *MSX1*, an immunohistochemistry methodology using an anti-*TBX22* antibody in the same embryo section as and *in situ* probe for *MSX1* would be required. However, this is difficult due to the previously discussed problems with antibodies for

TBX22 (section 3.3.2 and 3.4).

#### 4.4 DISCUSSION

Searching promoter regions of 132 known human cleft palate causing genes for the presence of the TBX22 DNA consensus sequence did not yield any hits. However screening the same set of genes with a generic T-box binding sequence produced several potential matches. This is perhaps not surprising given that a search for targets of the T-box gene, T-bet, by chromatin immunoprecipitation, identified 832 protein coding target genes in human Th1 cells (Jenner *et al.* 2009). However, only the first 2kb region upstream relative to the start of transcription was examined for the presence of potential TBX22 binding sites. It is therefore possible that potential TBX22 binding sites present in areas that extend beyond this region have been missed, as the promoter can extend for several kilobases upstream of the start of transcription (for review see Levine and Tjian 2003). Similarly, any intronic regulatory regions within the gene which act as binding sites would also not be detected.

Whilst ten genes were found to contain a direct match to the generic T-box binding sequence within the 2kb region examined (Table 16), relaxing the conditions to allow a 1 base mis-match to the generic binding sequence produced more gene hits. Relaxing the stringency of the binding site search sequence in this manner produced a hit in 102 of the 132 genes investigated. While it is possible that all of these genes are downstream targets of TBX22, for this study two further criteria were used to refine the search. Firstly, as it has been shown that functional T-box binding sites are often found together in clusters (Barron *et al.* 2005; Garnett *et al.* 2009), genes that contained 3 or more of these hits (Table 17) were also considered with those genes that contained a direct match.

Secondly, further evidence of possible TBX22 target genes was sought from mouse mutants. Of the 132 genes examined only 39 had been reported to display a cleft plate when disrupted in the mouse mutant. It has been shown several times that oral clefts are not always present when a single gene is disrupted, but when a compound mutant is generated with another developmental craniofacial gene a cleft palate phenotype arises (Beverdam *et al.* 2001; Ding *et al.* 2004; Alkuraya *et al.* 2006). In cases where

a single gene disruption does cause a cleft palate mutant phenotype, generating compound mutants with other clefting genes often increases the frequency and/or the severity of the cleft (Alkuraya *et al.* 2006; Nakatomi *et al.* 2010).

Possible TBX22 downstream targets, therefore, were identified if they met the following criteria:

a) the promoter region screened contained either the generic T-box binding sequence or at least three copies of the “1 mis-match” sequence and b) a mouse mutant for that gene displayed a cleft palate phenotype. This generated 10 candidate genes for consideration, one of which was already known to be a TBX22 target – *TBX22* (Andreou *et al.* 2007). From the remaining nine genes (Table 18), *MSX1* was chosen for further study as being the most likely to be a TBX22 target gene for the following reasons:

Mutations in the *ZEB2* gene give rise to Mowat-Wilson syndrome (Cacheux *et al.* 2001). This is a complex syndrome where mental retardation and impaired motor development are prominent features (Mowat *et al.* 2003). Whilst there has been little to suggest that *ZEB2* has a significant role in craniofacial development, interactions with *TGF $\beta$ 1* have been demonstrated (Gregory *et al.* 2008) and associations to CL/P and *TGF $\beta$ 1* have been made (Stoll *et al.* 2004). The *Zeb1*<sup>+/−</sup>; *Zeb2*<sup>−/−</sup> compound mutant was reported as having a cleft of the maxilla (Miyoshi *et al.* 2006). However, as the cleft seen in CPX is of the secondary palate, *ZEB2* was not considered as being a good candidate as a TBX22 target gene. *MKS1* is the gene disrupted in Meckel syndrome (Kyttala *et al.* 2006). The protein encoded by *MKS1* has been shown to be associated with cilia formation in most tissues (Weatherbee *et al.* 2009) and as such was not considered to have a specific role in palatogenesis and therefore interaction with TBX22 was considered unlikely. Whilst a *Hoxa2* mutant has a cleft plate phenotype (Barrow and Capecchi 1999) the expression of *Hoxa2* in developing pharyngeal arches (Barrow *et al.* 2000; Trainor and Krumlauf 2001; Creuzet *et al.* 2005) precedes that of *TBX22* and is therefore unlikely to be downstream of TBX22.

Three of the ten genes identified as potential TBX22 candidate targets (see Table 18) have been shown to have a role in cartilage or bone differentiation - *FOXC2*, *COL2A1* and *EFNB1* (Vandenberg *et al.* 1991; Mo *et al.* 1997; Barbieri *et al.* 2003; Compagni

*et al.* 2003; Kim *et al.* 2009a). This is interesting given that the major role of *TBX22* based on phenotypic observation of the *TBX22* null mutant mouse was adjudged to be in osteoblast formation (Pauws *et al.* 2009a). *EFNB1* is the gene mutated in craniofrontonasal syndrome (Twigg *et al.* 2004; Wieland *et al.* 2005; Torii *et al.* 2007). *EphB2/EphB3* receptor mutants have a cleft palate phenotype and other abnormal ossification patterns (Compagni *et al.* 2003). Mouse *Efnb1* is expressed in neural crest cells that have migrated to the frontonasal process (Twigg *et al.* 2004). As this expression would precede that of *TBX22*, *EFNB1* was not considered further as a target gene.

A *FOXC2* mutation was reported in a patient with Lymphedema-distichiasis syndrome (Tanpaiboon *et al.* 2010), one of the characteristics of which being a cleft palate. Moreover, the palatal shelves of the *FOXC2* null mutants failed to fuse (De Felice *et al.* 1998). However as *FOXC2* has been shown to be involved in mesenchymal-epithelial transformation (Hader *et al.* 2010) it could be postulated that the failure of the palatal shelves to fuse was due to incorrect mesenchymal-epithelial transformation of the MEE (see 1.2.5). However, the expression pattern of *TBX22* seen in Chapter 2, suggests that *TBX22* has a role in the proliferation of the mesenchyme within the palatal shelves and on this basis *FOXC2* was not taken forward for further investigation as a *TBX22* target.

Although mutations in *COL2A1* cause Stickler syndrome (Vintiner *et al.* 1991), an association between *COL2A1* and non-syndromic cleft palate has also been established (Melkoniemi *et al.* 2003). The finding that *COL2A1* has a role in chondrocyte differentiation (Vandenberg *et al.* 1991; Barbieri *et al.* 2003) made this gene less likely to be a *TBX22* target for the same reasons as for *FOXC2*, namely that the *TBX22* expression patterns suggests that it is likely to have a role in mesenchyme proliferation.

Point mutations in *GLI2* have been suggested to cause non-syndromic CL/P (Vieira *et al.* 2005). Additionally, *Gli2* has been shown to be expressed in mouse palatal shelves and *Gli2* mutant mice have a severe cleft secondary palate (Mo *et al.* 1997). In these mice, the palatal shelves are sometimes completely missing or else palatal shelf elevation was delayed. As mesenchyme-epithelial transition was unaffected (Mo *et al.*

1997) this points towards a role in proliferation – similar to the expected role of TBX22 based upon the gene expression results (Chapter 2). However the *Gli2* mutant has other skeletal abnormalities indicating that *Gli2* has a more general role in skeletal development, rather than in palatogenesis specifically, so *Gli2* was not chosen for further investigation.

Of the ten genes identified from the *in silico* screen (Table 18) two were known to cause non-syndromic CL/P causing (see Table 4, Chapter 1), namely *SUMO1* and *MSX1*. These genes therefore were the strongest candidates for being authentic TBX22 targets. While it can be hypothesised that TBX22 may target *SUMO1* in auto-regulatory manner, an interaction between TBX22 and SUMO1 had already been demonstrated: TBX22 undergoes sumoylation; (Andreou *et al.* 2007). The remaining gene, *MSX1*, was therefore chosen for further investigation.

The heterozygous *Msx1*<sup>+/−</sup> knockout mouse has a cleft of the secondary palate (Satokata and Maas 1994). The palatal shelves of this mutant are correctly elevated, but have not fused with each other or with the nasal septum, correlating with the sites of TBX22 expression (see Chapter 2). The authors postulate that the failure of the palatal shelves to fuse in the *Msx1*<sup>+/−</sup>/*Msx1*<sup>+/−</sup> mutant is a consequence of insufficient accumulation of mesenchymal tissue within the palatal shelves.

*Msx1* has been shown to be expressed in the mandibular arch of the developing mouse embryo by *in situ* hybridisation (Robert *et al.* 1989) and was shown by northern blot to be expressed in the developing palate and in primary cultures of murine embryonic palate mesenchymal cells (Nugent and Greene 1998). Later it was reported that *Msx1* in the developing mouse face was exclusively expressed in the mesenchyme of the medial nasal, lateral nasal, maxillary and mandibular processes (Mackenzie *et al.* 1991; Jowett *et al.* 1993). The evidence from tissue *in situ* hybridisation to human sections is consistent with this (Fig. 26), and moreover the expression would appear to be complementary to that of TBX22 (see Chapter 2 and Fig. 26). The expression profile of the TBX22 and *MSX1* genes was examined in separate embryos and at slightly different developmental stages which obviously hinders the extent to which conclusions concerning the relative localisation of these genes can be made. These problems are always more difficult to overcome when working directly in human

rather than animal models due to the restricted availability of the material to perform such experiments. As a consequence of the limited tissue availability, repeating the same gene expression patterns in similar stage embryos is not always possible. International coordination in creating and populating human gene expression databases has long since been identified as one important way to try and address these issues (Strachan *et al.* 1997). The aim would be to provide the comparison of gene expression patterns from as large a data set as possible, whilst making the most effective use of this limited and valuable material.

Despite the limitations of working with human material and the differences in embryo sections and stages used in the experiments here, the gene expression patterns of human *TBX22* and *MSX1* would appear to correlate with what has been shown in mouse. In the mouse, *Msx1* is restricted to the anterior region of the palatal shelves (Welsh and O'Brien 2009) and *Tbx22* is restricted to the posterior region of the palatal shelves (Pantalacci *et al.* 2008; Welsh and O'Brien 2009). An equivalent expression pattern of these genes is also seen in human palate development (Fig. 26 and, for *TBX22* only, Baybrook *et al.* 2001) and it could be possible that *MSX1* expression was being restricted to the anterior palate by *TBX22* regulation.

Interaction of *MSX1* and T-box proteins at a protein-protein level has already been demonstrated (Boogerd *et al.* 2008). In addition, although *TBX22* itself is not the causative gene in a Brazilian family displaying ankyloglossia and tooth anomalies (Acevedo *et al.* 2010), the authors do suggest that a downstream gene of *TBX22* could be responsible. Mutations in *MSX1* have been shown to cause tooth abnormalities (van den Boogaard *et al.* 2000) and tooth deformities are also seen, in addition to a cleft palate, in the *Pax9<sup>+/−</sup>*; *Msx1<sup>+/−</sup>* compound mutant mouse (Nakatomi *et al.* 2010), perhaps suggesting a combined role for these genes in tooth formation as well as palatogenesis.

To verify that the *MSX1* gene had the potential to be regulated by *TBX22*, an EMSA was performed using an oligonucleotide containing the generic T-box binding sequence as found in the *MSX1* promoter (Table 16). The results of this EMSA showed that, *in vitro*, this sequence could bind to the lysate containing *TBX22* protein (Fig. 25) and therefore supported the idea that the *MSX1* gene could be a *TBX22*

target. The promoter region of the *MSX1* gene in which the T-box binding site was found is known to contain several transcription factor binding elements and to show extensive conservation with the mouse *Msx1* promoter, including many of the consensus binding sites (Shen et al. 1994; Gonzalez et al. 1998). The identified generic T-box binding site in the *MSX1* promoter (AGGTGTTGTG) is positioned 834 bases upstream from start of transcription. A similar sequence (AGGTCTTCTG) was found 1103 bases upstream from the start of transcription of the mouse *Msx1* gene. In the 5<sup>th</sup> and 8<sup>th</sup> positions in the site there is a change from G>C. None of the binding sites shown in Table 15 have a C in either the 5<sup>th</sup> or 8<sup>th</sup> position of the binding sites (*in vitro* established or *in vivo* where known) for a number of T-box genes. *In vivo* binding sites have been shown to be different to those established *in vitro* in several cases (Casey et al. 1998; Casey et al. 1999) so the significance of the G>C change at two positions is not clear. There are differences in expression and regulation of orthologues in mouse and human (for example see Fougerousse et al. 2000) and presumably such differences, in at least some cases, explain why no cleft palate is seen in 69 of the 108 mouse mutants for genes underlying human cleft palate disorders (Appendix 5). The presence of a similar site to the generic T-box binding site within the *Msx1* promoter, however, supports the possibility that *MSX1* is a downstream target gene.

The results from the EMSA studies using the *MSX1*\_oligo, together with the expression data and the malformations seen in the *Msx1* mutant mouse (Satokata and Maas 1994) go some way in providing evidence for the possibility that *MSX1* is regulated by TBX22. Therefore, further *in vitro* experiments were carried out to see if *MSX1* were indeed a true TBX22 downstream target (Chapter 5).

The generic T-box binding sequence was generated by aligning all of the known T-box DNA binding sites. The TBX22 preferential DNA binding sequence, identified in Chapter 3, was also included when compiling the generic T-box binding sequence. However, even if this sequence was not known the same generic T-box binding site would have been compiled (see Table 15). On the basis of this and the fact that all T-box proteins are known to bind to a similar DNA sequence (Sinha et al. 2000; Tada and Smith 2001), the identification of the *in vitro* preferential DNA binding site may well be redundant. Indeed, a previous study identified a preferential 20 base DNA

binding sequence for TBX22 (Andreou *et al.* 2007). This sequence consisted of 2 half binding sites in opposite orientations separated by 3 nucleotides (Table 15), although this sequence as a whole is not actually present within the genome. Including the half-site does not change the generic T-box binding site generated (Table 15). In fact, with advances in chromatin immunoprecipitation techniques – ChIP, reviewed in (Horak and Snyder 2002; Wong and Wei 2009), the DNA binding sequence has, to some extent, become incidental as protein-DNA interactions can be investigated directly within the cell. Using this approach would identify actual transcription factor target genes directly, although confirmation of these interactions is still needed by other *in vitro* methods (Garnett *et al.* 2009).

## CHAPTER 5

### **MSX1: A POTENTIAL TBX22 TARGET GENE**

#### 5.1 INTRODUCTION

As has previously been described, the possibility of an interaction between TBX22 and *MSX1* is supported by their abutting expression patterns in the palatal shelves in human (Chapter 4); a similar restricted expression to either the anterior or posterior poles of the palatal shelves in mouse (Li and Ding 2007; Pantalacci *et al.* 2008); both *Tbx22* (Satokata and Maas 1994; Pauws *et al.* 2009a) *Msx1* mutant mice displaying a cleft palate, and the fact that mutations in both of these genes result in non-syndromic cleft palate in human (van den Boogaard *et al.* 2000; Braybrook *et al.* 2001). Also of note, is the fact that both MSX1 and TBX22 proteins have been shown to undergo post-translational modification by SUMO1 (Gupta and Bei 2006; Andreou *et al.* 2007), which is interesting as haploinsufficiency of the *SUMO1* gene causes non-syndromic CL/P (Alkuraya *et al.* 2006). Direct protein-protein interaction of MSX1 and T-box proteins has also been reported to exist via their respective homeobox and T-box domains, (Boogerd *et al.* 2008).

With the results from the EMSA study in Chapter 4 give encouragement that TBX22 may bind to an oligonucleotide with a sequence found in the *MSX1* promoter, this increased support for *MSX1* being a possible downstream target gene of TBX22. However, further investigation as to whether *MSX1* was indeed an *in vivo* TBX22 target was still needed. Therefore, the effect of over-expressed TBX22 on a reporter gene, fused downstream of the *MSX1* promoter region (containing the potential TBX22 binding site), was explored in human cell lines.

Initially, it had been planned to perform the transfection studies using the human embryonic palatal mesenchyme (HEPM) cell line (ECACC No. 90120505). The

rationale behind this being that if TBX22 were indeed capable of *trans* regulation of the *MSX1* gene during palatogenesis, then the best model to explore this would be to utilise those cells where this would take place *in vivo* - the embryonic palate mesenchyme. However, following several unsuccessful attempts using several different transfection reagents, the decision was made to change to another cell line that was known to transfect easily.

Previous studies had shown that TBX22 was endogenously expressed in several cell lines that could easily be transfected (Andreou *et al.* 2007) meaning that these cells should be capable of transcribing and translating a transfected TBX22 plasmid. However, in the same study, it was reported that TBX22 repressed the transcription of *TBX22* in an autoregulatory manner. This posed a potential problem: if one was to over-express a TBX22 protein in a cell that had endogenous TBX22 expression, due to transcriptional repression of the endogenous *TBX22* by the transfected TBX22, the net effect could in fact be a reduction in the total TBX22 protein in the cell. To account for this, experiments were performed in two cell lines – one which had been shown to endogenously express TBX22 – HeLa, and one which did not – 293T.

The HeLa cell line is derived from cervical carcinoma epithelial cells (Gey *et al.* 1952) from a woman named Henrietta Lacks, from whom the cell line name is taken (Jones *et al.* 1971). They are an immortal aneuploid cell line and are widely used in functional studies, including the study of the *trans* regulation of reporter genes by the Brachyury protein (Kispert *et al.* 1995). HEK 293T cells are a highly transfectable derivative of 293 cells, a hypo-triploid human embryonic kidney cell line, into which the gene for SV40 T-antigen has been inserted.

A Dual Reporter system was used to visualize the effect of TBX22 on an *MSX1* promoter. This system employs two separate luciferase reporter genes: one as an experimental reporter, the other being used as an internal control. The benefit of having an internal control reporter means that experimental variance, arising from differences in cell density, cell viability and transfection efficiency can be eliminated. The experimental reporter is a firefly (*Photinus pyralis*) luciferase which is cloned downstream of the *MSX1* promoter sequence. The internal control is a *Renilla reniformis* luciferase under the control of a Herpes Simplex Virus thymidine kinase

(HSV-TK) promoter. Both luciferase gene constructs are co-transfected and the resultant protein for which they code can be assayed sequentially. This is possible as each luciferase enzyme can be activated by separate substrates to emit a bioluminescent signal. The amount of luminescence generated is proportional to the amount of luciferase enzyme present. To account for inter-experimental differences, the assay is normalised using the ratio of the luminescence generated by the firefly to that of the *Renilla* luciferases.

The effect of TBX22 on the *MSX1* reporter plasmid was assessed by comparing the normalised result of the luciferase assay for cells co-transfected with either 26.7ng or 53.4ng of the TBX22 expression plasmid, or the same plasmid lacking the TBX22 coding sequence. This plasmid was co-transfected with a firefly luciferase gene under the control of an *MSX1* promoter fragment and the *Renilla* reporter control. The *MSX1* reporter contained the TBX22 binding site and also covered the region of the *MSX1* promoter which had previously been shown to contain the minimal elements required to drive expression of the *MSX1* gene (Gonzalez *et al.* 1998). An overview of the experiment and the different plasmids used in each transfection is shown in Table 19, section 5.2.2.B.

## 5.2 MATERIALS AND METHODS

### 5.2.1 Preparation of the Expression and Reporter constructs

#### A. PREPARATION OF THE PCR3.1\_TBX22 EXPRESSION PLASMID

A 1566bp fragment (bases 132-1697 of GenBank accession No. NM\_001109878) which included the entire *TBX22* coding sequence and Kozac initiation sequence was PCR amplified (see 3.2.1.B), using the Pfu DNA polymerase and forward primer 5'-GGGATGGCTCTGAGCTCTC-3' and reverse primer 5'-CATAAGGTAATGGTTAATTGCTGAA-3' with 50ng of I.M.A.G.E. clone BC014194 (Geneservice) as a starting template for 25 cycles with an annealing temperature 55°C. After the PCR reaction, the tubes were placed on ice and 1 unit of Taq polymerase was added to the reaction and placed in hot block at 72°C for 10 minutes. This was necessary in order to produce a 3' adenosine residue to the fragment to assist with the ligation of the insert into the plasmid vector. The DNA was then phenol-chloroform extracted/ethanol precipitated (see 2.2.2.A) and re-suspended in 30μl TE buffer. An approximate 2 molar excess ratio of insert:vector was used for the ligation (see 3.2.1.C). A 10μl ligation reaction containing 1μl 10X ligation buffer (60mM Tris-Cl, pH 7.5, 500mM KCl, 25 mM MgCl<sub>2</sub>, 0.01% gelatin), 60ng pCR3.1 cloning vector (Invitrogen), 40ng DNA fragment and 4 units of T4 DNA ligase (Invitrogen) was incubated overnight at 15°C. 2μl of the ligation reaction was transformed into One Shot TOP10F' competent cells (Invitrogen) as per the heat shock method described in 3.2.1.D.

The pCR3.1 cloning vector is an eukaryotic expression vector that drives expression of a cloned gene product using the promoter/enhancer region from the human cytomegalovirus, CMV (Thomsen *et al.* 1984). In addition, the cloning vector confers a polyadenylation signal to produce a 3' poly-A tail to the transcribed gene insert to reduce degradation of the recombinant protein mRNA and facilitates translation (Mangus *et al.* 2003).

For use as a control plasmid, the cloning site of the pCR3.1 vector was removed by enzymatic digestion using *EcoRI* (Promega) following the manufacturer's guidelines.

The DNA was then gel purified using a Qiagen Gel Extraction Kit following the manufacturer's directions and the plasmid was self-ligated using 4 units of T4 DNA ligase and 1 $\mu$ l 10X ligation buffer (see above) with 20ng of the digested vector in a 10 $\mu$ l reaction at 15°C overnight. The ligation reaction was then transformed as per the TBX22 expression plasmid above.

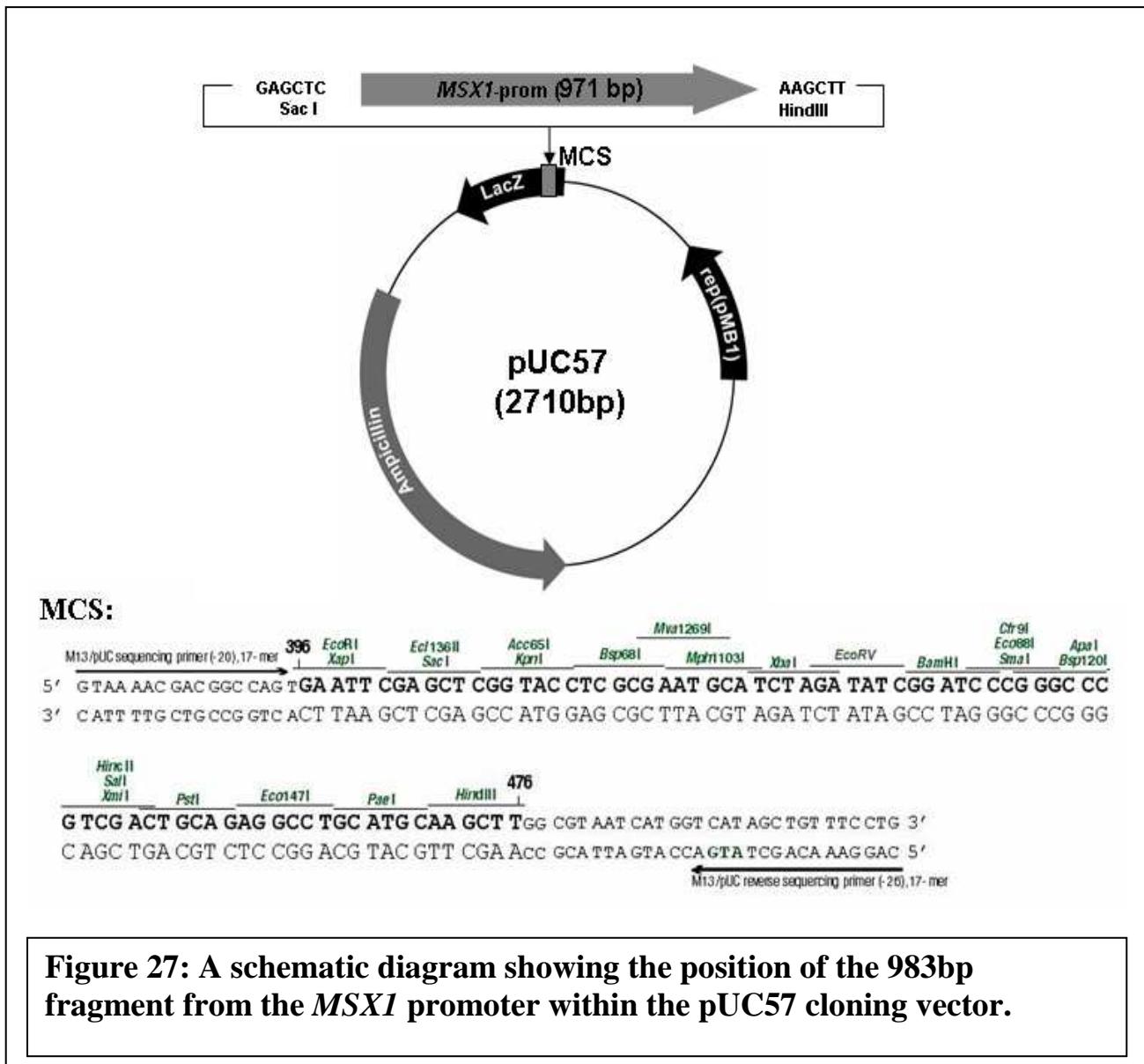
#### B. PREPARATION OF THE *MSX1* REPORTER PLASMID (M-PROM)

983 bp of the *MSX1* promoter (947860-948820 of GenBank accession No. NT\_006051.17) was cloned commercially by the GeneScript Corporation into a pUC57 vector between the *SacI* and *HindIII* restriction enzyme sites of the multiple cloning site (see Fig. 27). The fragment was excised from the plasmid by double digestion with these enzymes in the Multi Core buffer from Promega according to the manufacturer's recommended guidelines. The pGL4.17[*luc2*/Neo] vector (Promega) was similarly double digested with both of these enzymes and products of both reactions were electrophoresed on a 1% agarose gel. The resulting lower molecular weight band from the pUC57 vector and the ~ 5.5 kb band from the pGL4.17[*luc2*/Neo] vector were excised from the gel and purified using a Qiagen gel extraction kit following the manufacturer's instructions. A ligation reaction and subsequent transformation into TOP10F' competent cells was then performed using the digested *MSX1* promoter fragment and pGL4.17[*luc2*/Neo] vector, as described for the TBX22 expression construct above (5.2.1.A), to generate the *MSX1* promoter reporter plasmid M\_prom: the pGL4.17[*luc2*/Neo] vector carrying the firefly luciferase gene *luc2* cloned downstream of the *MSX1* promoter fragment.

#### C. PREPARATION OF THE *RENNILLA* INTERNAL CONTROL REPORTER PLASMID (PGL4.74[hRLUC/TK]).

An internal control reporter plasmid pGL4.74[*hRluc*/TK] was supplied by the manufacturer (Promega) as pure plasmid DNA and could be used directly in the transfection studies. The pGL4.74[*hRluc*/TK] plasmid carries the *Renilla hRluc*

luciferase gene, which is driven by the Herpes Simplex Virus thymidine kinase (HSV-TK) promoter (Wigler *et al.* 1979).



**Figure 27: A schematic diagram showing the position of the 983bp fragment from the *MSX1* promoter within the pUC57 cloning vector.**

## 5.2.2 Co-transfection of the Expression and Reporter constructs

### A. CELL CULTURE PROPAGATION

HeLa cells ECACC No. 93021013 and 239T cells ATCC No. CRL-11268 were propagated in culture medium containing minimum essential medium (MEM) with non-essential amino acids (NEAA) and Earle's Salts (Gibco) with 2mM L-glutamine (Gibco) and 10% fetal bovine serum (FBS, Sigma) in a 5% CO<sub>2</sub> incubator at 37°C. The cultures were subcultured by splitting them 1:3 when they reached 70-80% confluence by removing the culture media and washing the cells in PBS (Gibco) and then adding 0.5g/L of trypsin and 0.2g/L of EDTA (Gibco) and allowing the cells to detach from the culture vessel. Fresh culture medium was then added and the cultures were aspirated into new culture flasks at the appropriate dilution.

For transfection, the HeLa cells were seeded at a density of 1 X 10<sup>4</sup> cells per well and the 293T cells at 2 X 10<sup>4</sup> cells per well, as per the directions given in the Promega transfection database for each cell line (<http://www.promega.com/techserv/tools/fugenehdtool/>), in a clear bottom 96-well BD Optilux plates (BD Falcon) in a 100µl volume. The resultant cultures obtained 50-70% confluence in 24 hours.

### B. TRANSFECTION PROCEDURE

Transfection grade DNA was obtained by culturing the transformed TOP10F' competent cells containing each of the 4 separate plasmids (pCR3.1\_TBX22; pCR3.1\_Null; M\_prom or pGL4.74[hRluc/TK]) using the method described in 5.2.1, and preparing plasmid DNA using a plasmid mini kit (Qiagen), following the manufacturer's instructions and eluting in 20µl volume of Milli-Q H<sub>2</sub>O. The concentration and purity were verified by a Nanodrop spectrophotometer. The insert sequence and orientation of the TBX22 expression plasmid was verified by sequencing analysis using 5'-TAATACGACTCACTATAGGG- 3' and 5'-TAGAAGGCACAGTCGAGG-3'. The sequence of the *MSX1* reporter insert was verified using sequencing primers 5' -CTAGCAAAATAGGCTGTCCC-3' and 5' -TCGATATGTGCGTCGGTAAA-3'. (see 3.2.1.F).

A third internal reporter was transfected into the cells alongside the TBX22 expression and *MSX1* promoter reporter plasmids. This control plasmid contained the *Renilla hRluc* gene which is driven by the HSV-TK promoter.

The FuGENE HD transfection reagent (Promega) was used to co-transfect the cells with either: a no TBX22 control (53.4ng pCR3.1\_Null), 26.7ng pCR3.1\_TBX22, or 53.4ng pCR3.1\_TBX22 expression plasmid, together with 26.7ng of the M\_prom and 26.7ng of the pGL4.74[*hRluc*/TK] reporter plasmids per well. In each case the required amount of each plasmid DNA was added to diluted FuGENE HD as per the manufacturer's directions, the tubes were tapped gently to mix the contents and incubated at room temperature for 20 minutes before 5 $\mu$ l of the transfection reagent and DNA were pipetted directly into the wells of the HeLa or 293T cells. Each of the transfections is detailed below and an overview is given in Table 19:-

#### No TBX22 Control

1100ng pCR3.1\_Null, 550ng M\_prom and 550ng pGL4.2[*hRluc*/TK] was added to 103 $\mu$ l H<sub>2</sub>O and 6.6 $\mu$ l of FuGENE HD reagent.

5 $\mu$ l of the transfection mix was added to each well.

#### 26.7ng TBX22

550ng pCR3.1\_TBX22, 550ng M\_prom and 550ng pGL4.2[*hRluc*/TK] was added to 103 $\mu$ l H<sub>2</sub>O and 6.6 $\mu$ l of FuGENE HD reagent.

4.95 $\mu$ l of the transfection mix was added to each well.

#### 53.4ng TBX22

1100ng pCR3.1\_TBX22, 550ng M\_prom and 550ng pGL4.2[*hRluc*/TK] was added to 103 $\mu$ l H<sub>2</sub>O and 6.6 $\mu$ l of FuGENE HD reagent.

5 $\mu$ l of the transfection mix was added to each well.

Experiment	Plasmids transfected	Luciferase	Expected Result
<b>Experimental Transfection (26.7ng TBX22 or 53.4ng TBX22)</b>	pCR3.1_TBX22  M_prom (pGL4.17[ <i>luc2</i> /Neo] vector with <i>MSX1</i> promoter)  pGL4.2[ <i>hRluc</i> /TK]	None  Firefly luciferase  <i>Renilla</i> luciferase	Over-expression of TBX22  Expression of firefly reporter gene  Expression of <i>Renilla</i> control reporter gene
<b>Control Transfection</b>	pCR3.1_Null  M_prom (pGL4.17[ <i>luc2</i> /Neo] vector with <i>MSX1</i> promoter)  pGL4.2[ <i>hRluc</i> /TK]	None  Firefly luciferase  <i>Renilla</i> luciferase	No recombinant expression of TBX22  Expression of firefly reporter gene  Expression of <i>Renilla</i> control reporter gene
<b>Blank Transfection</b>	No Plasmids	None	No luciferase expression. Value to which lumenesence assays can be blanked

**Table 19: A summary of the plasmids used in the transfection experiments.**

The cells had been seeded the previous day at the appropriate cell density described above, so that they were 50-70% confluent. 5 $\mu$ l of each of the transfection reagent:DNA complexes was added to separate wells of a 96-well plate. For both cell lines, three independent experiments were performed using different 96-well plates with four-to-six replicates in each of the transfection conditions (ie control

transfection with pCR3.1\_Null; transfection with 26.7ng pCR3.1\_TBX22 or transfection with 53.4ng pCR3.1\_TBX22 together with co-transfection with 26.7ng M\_prom and 26.7ng pGL4.2[hRluc/TK]). Once the transfection reagents had been added to the wells containing the cells, the culture plates were returned to the incubator and incubated for a further 24 hours before being lysed and luciferase levels determined.

The six 96-well plates used for these experiments were arranged as described in Fig. 28. Cells were not grown in the outside wells (columns 1 and 12 and rows A and H) so as to reduce the light interference from within the luminometer. The cells were also spaced so that a row of blank wells was positioned between each of the 3 transfection conditions so as to reduce the effects of cross-talk (Irawan *et al.* 2005) between wells. The cultures in one of the wells of each 96-well plate was left free from transfection reagents and served as a blank in the luciferase assay (5.2.2.D.) Following luciferase detection using a luminometer (see 5.2.2.C), if an experimental *Renilla* luciferase value was not statistically different from the *Renilla* luciferase of this blank assay, then it was considered that the transfection had failed and hence that whole column (including results for the other experimental conditions) were excluded. This was done in order to retain the same number of samples for each condition.

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B			●	●	●	●	●	●				
C												
D			●	●	●	●	●	●			●	
E												
F												
G			●	●	●	●	●	●				
H												

**Figure 28: A representation of the layout of the 96-well plate used for each transfection experiment.** For both the HeLa and 293T cells, 3 separate experiments were performed, each using a different 96-well plate, setup similar to that above. For every experiment, three rows of cells containing 4-6 wells of growing cells were transfected with one of the transfection conditions: 26.7ng pCR3.1\_Null (black dots); 26.7ng pCR3.1\_TBX22 (red dots) or 53.4ng pCR3.1\_TBX22 (blue dots). All three transfections were co-transfected with 26.7ng M\_prom and 26.7ng pGL4.2[hRluc/TK]. A blank transfection was also prepared for each experiment (yellow dot).

### C. LUCIFERASE ASSAY

The culture medium was aspirated from the wells and the cells rinsed with PBS to remove detached cells and residual growth media. The PBS was then completely removed and 20 $\mu$ l of 1X Passive Lysis buffer (PLB, Promega) was added to each well and the plate agitated on a gentle rocker at room temperature for 15 minutes.

The injector tips of a Fluoroskan Ascent FL luminometer (Thermo) were cleaned by rinsing in Milli-Q H<sub>2</sub>O and then soaked in 70% ethanol overnight, before being thoroughly rinsed with Milli-Q H<sub>2</sub>O. Luciferase assay reagent II (LARII, Promega) and the Stop & Glo reagent (Promega) were prepared following the manufacturer's instructions and placed in separate injectors, which were primed before use. For each well that was to be measured, the luminometer was programmed to perform a 2 second premeasurement delay; inject 100 $\mu$ l LARII and take a 10 second measurement of the luminescence, wait 2 seconds, inject 100 $\mu$ l of the Stop & Glo reagent and then take a 10 second reading of the luminescence, before repeating this cycle at the next marked well. Statistical analysis of the results was performed using Minitab 15.1.0.0 (Microsoft).

Pilot data generated from previous experiments using HeLa cells only is shown in Appendix 7.

### 5.2.3 Detection of *TBX22* and *MSX1* expression in HeLa and 293T cells

#### A. RNA ISOLATION

HeLa and 293T cells growing in a T75 culture flask were lysed directly in the culture vessel by the addition of 7.5ml TRIzol reagent (Invitrogen) and passing the cell lysate through the pipette tip several times to homogenise the cells. The cells were incubated in this reagent for 5 minutes and then transferred to a 15ml Falcon tube. 1.5ml of chloroform was added to the solution and mixed thoroughly by shaking the tube and left to incubate at room temperature for 3 minutes. A Micro Centaur (Sanyo) bench top centrifuge was used throughout this procedure and the sample was phase separated by centrifugation at 12,000g for 15 minutes at 4°C in this instrument. Following centrifugation, the RNA was collected from the upper aqueous phase and transferred to a separate tube. 3.75ml isopropanol was added to the aqueous phase and incubated at room temperature for 10 minutes to precipitate the RNA. The samples were then centrifuged at 12,000g for 15 minutes at 4°C and the supernatant removed. 8ml 75% ethanol was added to the RNA pellet, mixed with a vortex and centrifuged once more at 7,500g for 5 minutes at 4°C. The supernatant was again removed from the sample and the pellet was left to air dry for 5 minutes. The RNA was redissolved using 100μl of RNase free Milli-Q H<sub>2</sub>O.

#### B. FIRST STRAND cDNA SYNTHESIS

First strand cDNA was synthesised by adding 1μg RNA from the lysed cells (5.2.3.A) to 2.5μM anchored oligo(dT)<sub>18</sub> primer (Roche), 60μM random hexamer primer (Roche) and brought to a total of 13μl with RNase free Milli-Q H<sub>2</sub>O. The reaction was heated in a hot-block for 10 minutes at 65°C and then immediately placed on ice. 4μl 5X Transcriptor Reverse Transcriptase reaction buffer (5X buffer: 250mM Tris-Cl, 150mM KCl, 40 mM MgCl<sub>2</sub>, pH 8.5; Roche), 0.5μl Protector RNase Inhibitor (40U/μl; Roche), 2μl dNTP's mix (10mM ATP, CTP, GCP, TCP; Roche) and 0.5μl Transcriptor Reverse Transcriptase (20U/μl; Roche) were then added to the vial and the contents mixed gently. The tubes were then transferred to a Dyad MJ Research thermocycler and incubated for 10 minutes at 25°C followed by incubation at 50°C for 1 hour. The Reverse Transcriptase enzyme was inactivated by a final incubation at

85°C for 5 minutes. The tubes were then cooled on ice and stored at -25°C. A control reaction using no Transcriptor Reverse Transcriptase enzyme was also performed.

### C. DETECTION BY RT-PCR

The presence of a *TBX22* and *MSX1* cDNA fragment was determined using a PCR reaction (3.2.1.B), with 2µl of the first strand synthesis and reverse transcriptase negative reactions (5.2.3.B) as template DNA. 5'- AAGCGGGCAGGCAGGATGTTC -3' and 5'- AGGTCTCTCCGAGCAGGGT -3' primers were used to detect *TBX22* with an annealing temperature of 61°C; primers 5'- AGAAGATGCGCTCGTCAAAG -3' and 5'- CCATATCTTCACCTGCGTCTC -3' were used for *MSX1* with an annealing temperature of 54°C and primers 5'- TGCACCACCAACTGCTTAGC -3' and 5'- GGCATGGACTGTGGTCATGAG -3' were used for *GAPDH* with an annealing temperature of 53°C. Details of these primers are given in Appendix 1. Electrophoresis was then performed using 5µl of the RT-PCR reaction on a 1.2% agarose gel and visualised under UV light. The resulting image was captured using a Gene Genius (Syngen) system.

#### 5.2.4 Detection of luciferase by Western blot

In order to determine that the *MSX1* promoter was capable of driving the expression of the luciferase gene, the cells transfected under the no *TBX22* conditions (i.e. with pCR3.1\_Null , M\_prom andpGL4.2[*hRluc*/TK]) were lysed as in 5.2.2.C and then assayed for the presence of the luciferase protein by Western blot (see 3.2.3.B) using an anti-luciferase antibody (Promega), at a concentration of 1µg/ml, followed by a secondary anti-goat HRP antibody (Vector Laboratories) used at a dilution of 1:200.

### 5.3 RESULTS

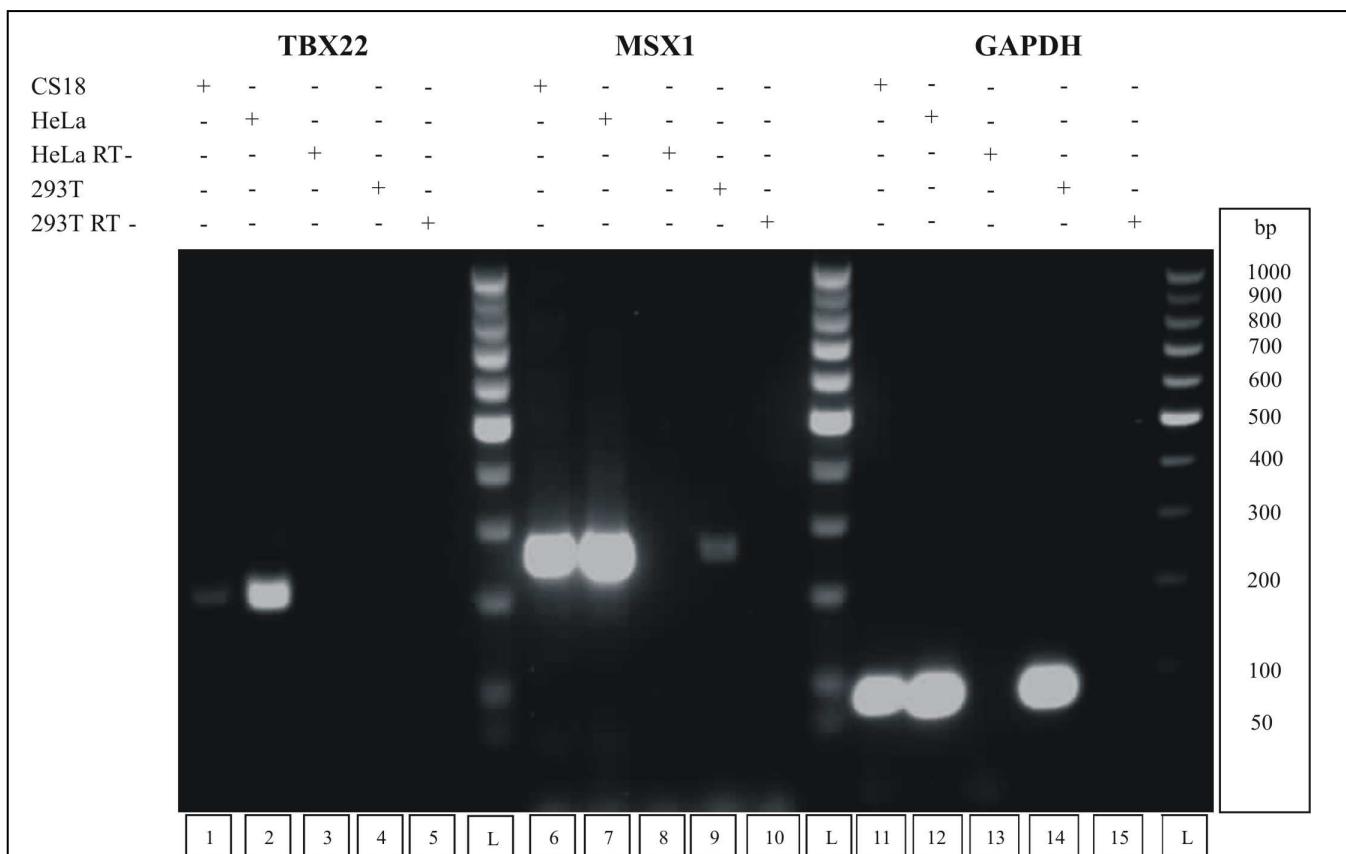
All the sets of primer pairs used amplify regions in their respective genes that bridge exon/intron boundaries, ensuring that the amplification of *MSX1* and *TBX22* cannot be from genomic DNA contamination. Fragments amplified from genomic DNA using these primers would be unlikely to result in the generation of amplicons, due to the significant distance between the two primer annealing sites (2479 bp for *TBX22* pair and 2608bp for *MSX1* pair). Lanes 2, 7 and 12 in Fig. 29 are a PCR reaction using cDNA samples prepared from the RNA of HeLa cells and lanes 3, 8 and 13 from RNA samples from HeLa cells without the addition of the Reverse Transcriptase enzyme which serve as a negative control for the experiment. Similarly, lanes 4, 9 and 14 shown the result of a PCR reaction using cDNA samples prepared from the RNA of 293T cells and lanes 5, 10 and 15 from RNA samples from 293T cells without the addition of the Reverse Transcriptase enzyme.

As *MSX1* and *TBX22* are both expressed around CS18 (see Fig. 26, Chapter 4 and Chapter 2 Fig 13 and 14 for *TBX22*), whole embryonic RNA prepared from a CS18 embryo (supplied by the HDBR, see 1.2) was used as a positive control for both primer pairs (Fig. 29 lanes 1 and 6). Additionally, GAPDH primers were also used (see Appendix 1 for sequence) as a positive control for the RNA from each of the samples (Fig. 29 lane 11 - CS18 cDNA; lane 12 – HeLa cDNA; lane 14 – 293T cDNA) which generates an amplicon of 86bp.

The RT-PCR using primers to *TBX22* gave a positive band for the HeLa cell line (lane 2 in Fig. 29), but did not produce a product from the 293T cells (lane 4 in Fig. 29). The migration of the fragment on the agarose gel corresponds to that expected by amplification of the cDNA of 219bp, showing that, as has been previously reported (Andreou *et al.* 2007), HeLa cells endogenously express *TBX22* and 293T cells do not.

RT-PCR using primers to *MSX1* gave a positive band for both HeLa and 293T cell lines (lanes 7 and 9 respectively in Fig. 29), amplifying a product of the expected

271bp. Thus, it may be concluded that both the HeLa and 293T cell lines used in these experiments express *MSX1* mRNA transcripts.



**Figure 29: HeLa cells express endogenous *TBX22* and *MSX1*, 239T cells express *MSX1* only.** The *TBX22* primers amplified a 219 bp product in the control CS18 cDNA (lane 1) and HeLa cDNA (lane 2), but no amplicon is detected in cDNA from 239T cells (lane 4). No product is detected in the Reverse Transcriptase negative samples from either HeLa (lane 3) or 239T cells (lane 5).

A 271 bp product is amplified by the *MSX1* primers from the cDNA of the control CS18 (lane 6), HeLa cells (lane 7) and the 239T cells (lane 9). No amplification was seen in the Reverse Transcriptase negative samples from HeLa (lane 8) or 239T cells (lane 10).

A positive control *GAPDH* primer pair amplified a product of the expected size of 86bp in the cDNA of the CS18 (lane 11), HeLa (lane 12) and 239T cells (lane 14). No product was detected in the Reverse Transcriptase negative samples from HeLa (lane 13) or 239T cells (lane 15).

**TBX22** – amplified using *TBX22* primers (lanes 1-5); **MSX1** – amplified using *MSX1* primers (lanes 6-10); **GAPDH** – amplified using *GAPDH* primers (lanes 11-15).

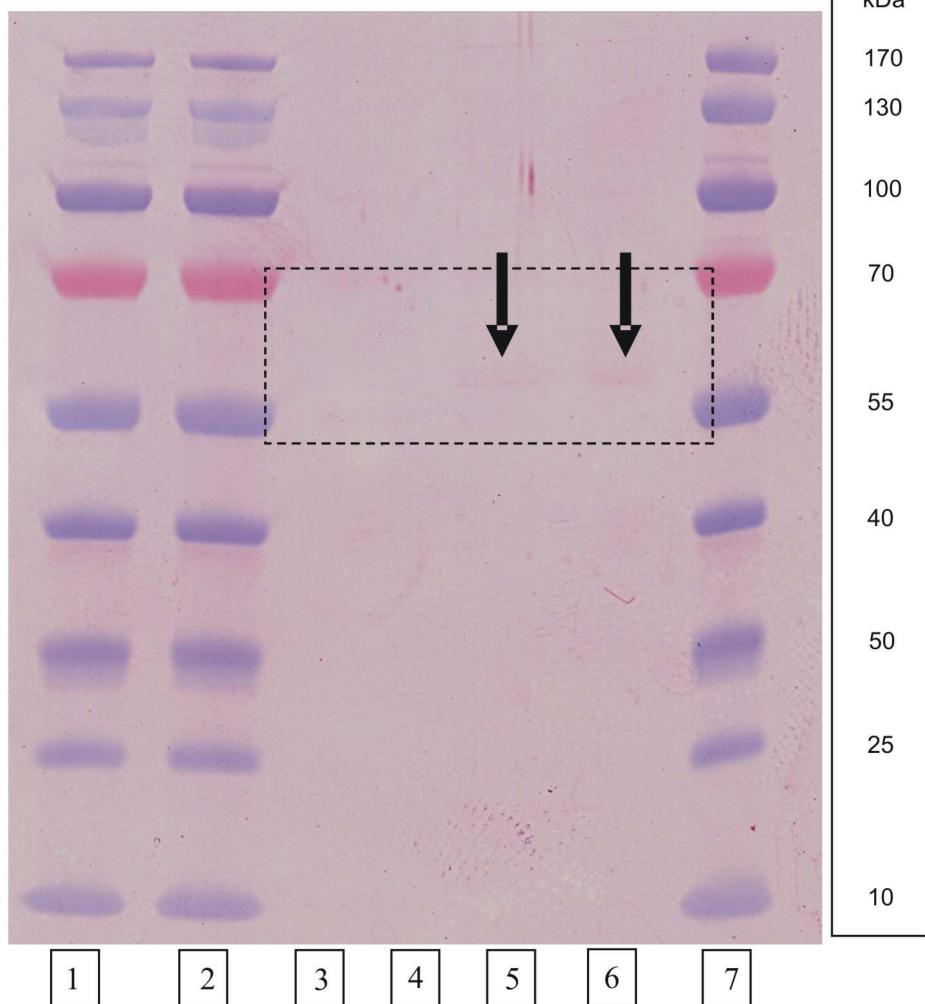
CS18 – cDNA from whole CS18 embryo; HeLa – cDNA from HeLa cells; HeLa RT- – Reverse Transcriptase negative samples prepared from HeLa RNA; 239T – cDNA from 239T cells; 239T RT- –Reverse Transcriptase negative samples prepared from 239T cell RNA. L is a 100bp DNA mass marker (Promega) with approximate sizes indicated.

If a luciferase assay was going to provide meaningful data to compare the effects of over-expressing TBX22, it was important to determine that the *MSX1* promoter reporter plasmid (M\_prom) was capable of driving expression of the luciferase protein. Figure 30 shows the presence, albeit very weakly, of luciferase protein in transfected HeLa (lane 5 in Fig. 30a) and transfected 293T cells (lane 6 in Fig. 30a) compared to the untransfected cells (lanes 3 and 4 in Fig. 30a). The predicted protein of 550 amino acid residues has a molecular weight of 60.64 kDa (sequence information from Promega website - <http://www.promega.com/vectors/pgl4.17.txt>; *luc2* gene nucleotides 100-1752 were translated using the ExPASy online translation tool hosted by the Swiss Institute of Bioinformatics - <http://www.expasy.ch/tools/dna.html>).

Although the luciferase protein seen in the transfected HeLa cells (lane 5 Fig. 30) appears weaker than that seen in the 293T cells (lane 6 Fig. 30), this may be due to the difference in seeding densities: HeLa cells seeded at  $1 \times 10^4$  cells per well and the 293T cells at  $2 \times 10^4$  cells per well (see 5.2.2).

As the M\_prom plasmid was expressing the luciferase protein in both cell types, the effect of over-expressing TBX22 protein in these cells on the amount of luciferase produced would be predicted to indicate any regulatory effect of TBX22 on the *MSX1* promoter fragment. This experiment was done by using a luminometer to measure the amount of luminescence emitted by the transfected cells after the addition of a suitable substrate for the luciferase protein. The report from the Fluoroskan luminometer showing the absolute values for all of the replicates from each of the 3 independent experiments at each transfection condition (no TBX22, 26.7ng TBX22 or 53.4ng of TBX22 plasmid) is shown in Appendix 6 and the successful transfections are summarised in Tables 20 and 21.

a



b



**Figure 30: Luciferase expression driven by the *MSX1* promoter in HeLa and 293T cells. a.** Luciferase protein is detected in HeLa cells (lane 5) and 239T cells (lane 6) transfected with the 1100ng pCR3.1\_Null, 550ng M\_prom and 550ng pGL4.2[*hRluc*/TK] using an anti-luciferase antibody (see 5.2.4). The dotted-box shows the region shown in **b**, which shows lanes 3 - 6 at higher magnification and increased contrast to aid identification. The ladder (lane 7) is a PageRule Prestained Protein Ladder (SM0671; Fermentas) with the approximate migration of protein molecular weights indicated (kDa). Lane 3 is from untransfected HeLa cells and lane 4 is untransfected 293T cells. The luciferase protein is detected using an anti-luciferase antibody (see 5.2.4), which is then visualised using HRP/DAB (see 3.2.3).

	M_prom + No TBX22			M_prom + 26.7ng TBX22			M_prom + 53.4ng TBX22			Cells alone
Experiment	Firefly Luciferase	<i>Renilla</i> Luciferase	Ratio	Firefly Luciferase	<i>Renilla</i> Luciferase	Ratio	Firefly Luciferase	<i>Renilla</i> Luciferase	Ratio	<i>Renilla</i> Luciferase
1	5.945508	1.276604	4.657284	3.431129	2.157015	1.590684	7.038646	2.153371	3.268664	0.294113
	5.766866	0.760457	7.583426	2.500051	1.420697	1.759736	5.615232	1.727434	3.250620	
	3.494552	0.773907	4.515471	1.771932	0.908001	1.951466	10.131340	2.766837	3.661705	
2	10.354900	1.319740	7.846167	3.784751	2.534476	1.493307	3.903111	1.189683	3.280799	0.345382
	14.131780	1.022790	13.816893	5.896039	1.648178	3.577307	3.326693	1.238724	2.685580	
	12.511860	2.583470	4.843044	5.253373	1.932231	2.718812	3.420082	1.610355	2.123806	
3	14.678420	1.317048	11.144939	5.040354	1.188119	4.242297	2.773540	1.432939	1.935560	0.108811
	7.049310	1.650263	4.271628	6.646171	2.227520	2.983664	5.246964	1.698936	3.088382	
	5.330808	1.173110	4.544167	3.344198	0.753062	4.440801	7.771098	1.403407	5.537309	

**Table 20: Expression of Luciferase from the M\_prom promoter in HeLa cells.** The relative light units from the firefly luciferase, are normalised to the *Renilla* luciferase transfection control values (Ratio) following transfection with either 26.7ng pCR3.1\_Null plasmid (No TBX22), 26.7ng of pCR3.1\_TBX22 or 53.4ng pCR3.1\_TBX22, together with the 26.7ng M\_prom and 26.7ng pGL4.2[*hRluc*/TK] control plasmid. The values for the non transfected cell (Control *Renilla*) for each experiment are also shown.

M_prom + No TBX22			M_prom + 26.7ng TBX22			M_prom + 53.4ng TBX22			Cells alone	
	Firefly Luciferase	Renilla Luciferase	Ratio	Firefly Luciferase	Renilla Luciferase	Ratio	Firefly Luciferase	Renilla Luciferase	Ratio	Renilla Luciferase
1	2.401089	2.922685	0.821535	1.01768	0.476687	2.1349	2.207481	1.834464	1.203338	0.408504
	2.721097	1.302751	2.088731	0.756599	1.392531	0.543326	0.875337	1.267416	0.690647	
	1.165447	1.693124	0.688341	0.127379	1.340891	0.094995	1.368239	1.181022	1.158521	
2	1.45233	1.45233	1.03901	2.505352	2.384261	1.050788	2.343781	1.584134	1.479535	0.370679
	0.658094	0.658094	2.745829	2.052036	0.344296	5.960101	3.325981	1.246932	2.667331	
	2.29725	2.29725	2.116814	0.294337	2.399177	0.122682	2.070588	2.344638	0.883116	
	1.12741	1.12741	0.679699	1.42744	0.870507	1.639781	1.351431	2.06572	0.654218	
3	1.4602	1.026468	1.422548	1.462127	0.629199	2.323792	1.317282	2.435744	0.540813	0.346546
	2.296965	1.202742	1.909774	2.79639	2.114703	1.322356	1.817282	0.536731	3.385835	
	0.560777	0.213527	2.626264	1.443196	0.800868	1.802039	3.217748	0.827336	3.889287	
	2.655885	1.371096	1.937053	2.067135	1.280424	1.614414	1.946758	2.025865	0.960951	
	1.558073	1.77942	0.875607	1.385782	1.087407	1.274391	1.93726	1.66293	1.164968	

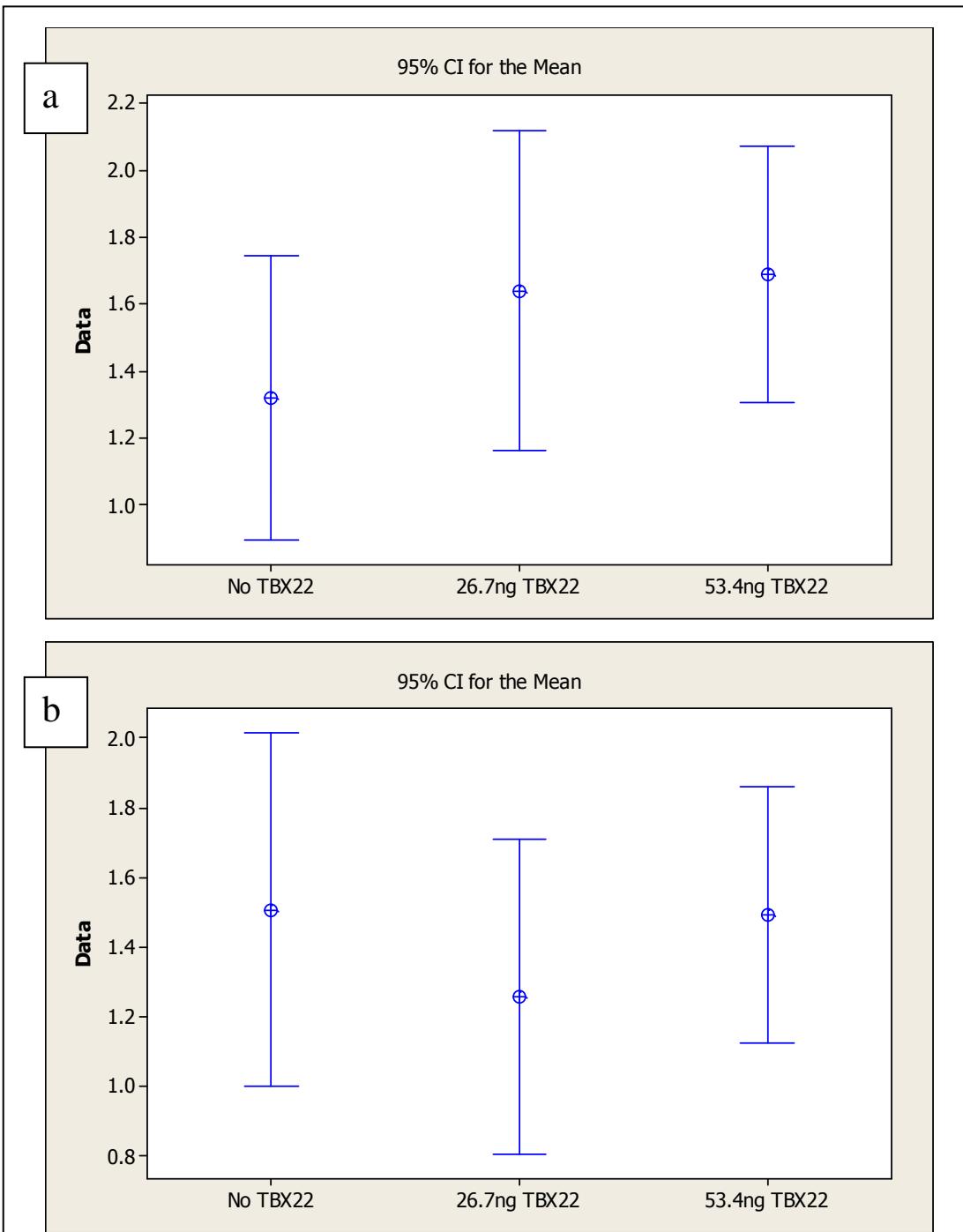
**Table 21: Expression of Luciferase from the M\_prom promoter in 239T cells.** The relative light units from the firefly luciferase, are normalised to the *Renilla* luciferase transfection control values (Ratio) following transfection with either 26.7ng pCR3.1\_Null plasmid (No TBX22), 26.7ng of pCR3.1\_TBX22 or 53.4ng pCR3.1\_TBX22, together with the 26.7ng M\_prom and 26.7ng pGL4.2[hRluc/TK] control plasmid. The values for the non transfected cell (Control *Renilla*) for each experiment are also shown.

To verify that the transfections used in the experiments were reliable, a student's t-test was employed to confirm that there was no statistical significance in the luminescence emitted from the control reporter plasmid (pGL4.2[hRluc/TK]) and each of the three different transfection conditions.

For the 239T cells at a 95% confidence interval, no statistical significance was found between the pCR3.1\_Null plasmid and 26.7ng of pCR3.1\_TBX22 plasmid (p-value = 0.430); between pCR3.1\_Null plasmid and 53.4ng of pCR3.1\_TBX22 plasmid (p-value = 0.959); or between the 26.7ng of pCR3.1\_TBX22 plasmid and 53.4ng of pCR3.1\_TBX22 plasmid (p-value = 0.387). The distribution about the mean of these values is represented in Fig. 31a.

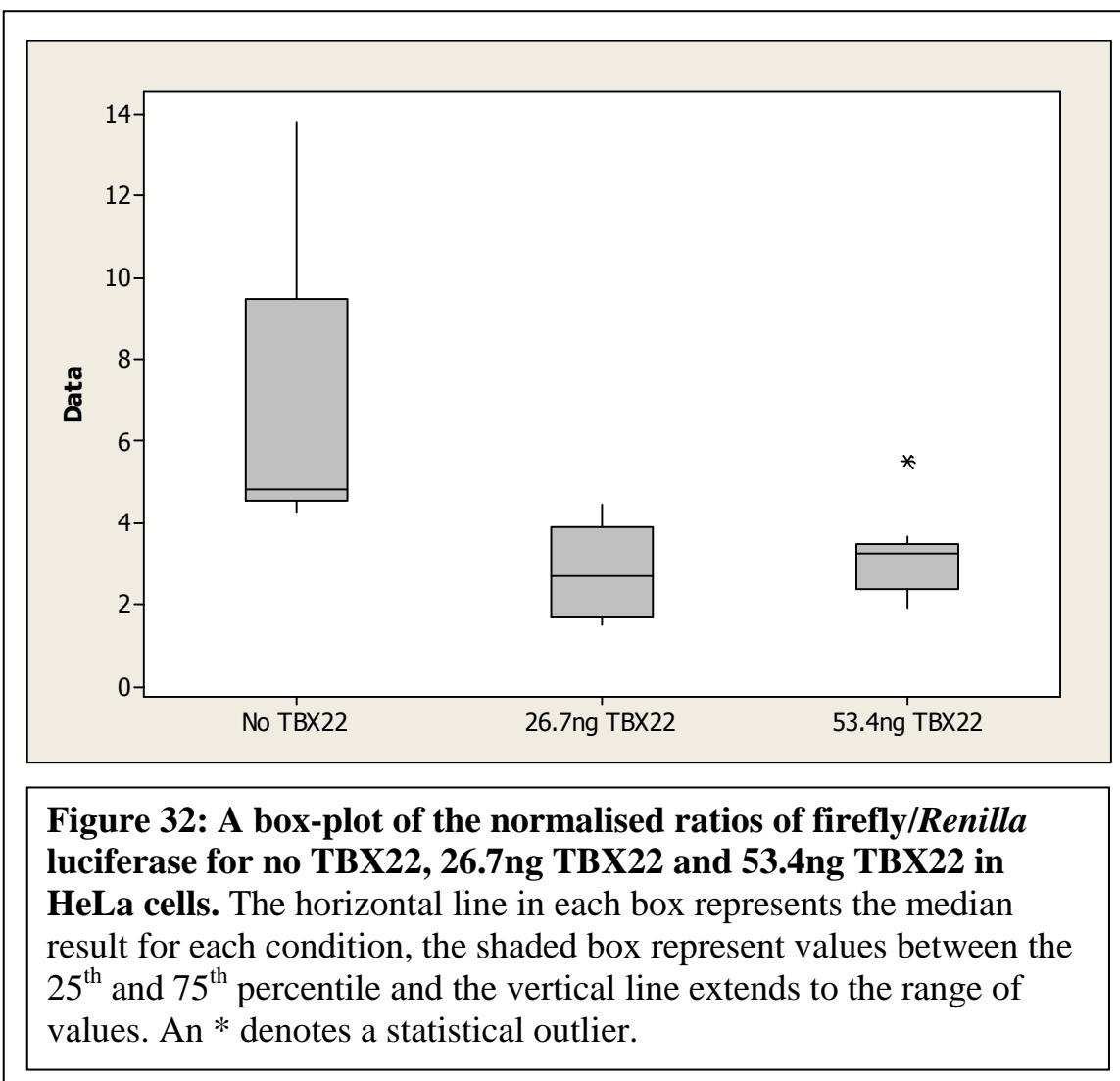
For the HeLa cells at a 95% confidence interval, no statistical significance was found between the pCR3.1\_Null plasmid and 26.7ng of pCR3.1\_TBX22 plasmid (p-value = 0.262); between pCR3.1\_Null plasmid and 53.4ng of pCR3.1\_TBX22 plasmid (p-value = 0.153); or between the 26.7ng of pCR3.1\_TBX22 plasmid and 53.4ng of pCR3.1\_TBX22 plasmid (p-value = 0.852). The distribution about the mean of these values is represented in Fig. 31b.

As no statistical difference was found between any transfection conditions and the luminescence emitted from the pGL4.2[hRluc/TK] plasmid, the experiments analysed were considered successful. Therefore the ratios for the firefly/*Renilla* luciferase assays were examined for statistical differences between the no TBX22, 26.7ng TBX22 and 53.4ng TBX22 conditions. As the comparison was between ratios, a Mann-Whitney *U* test (Wilcoxon 1945), which is a test of medians is preferable to a student's t-test, which is a test of means, but should only be used when comparing parametric data.

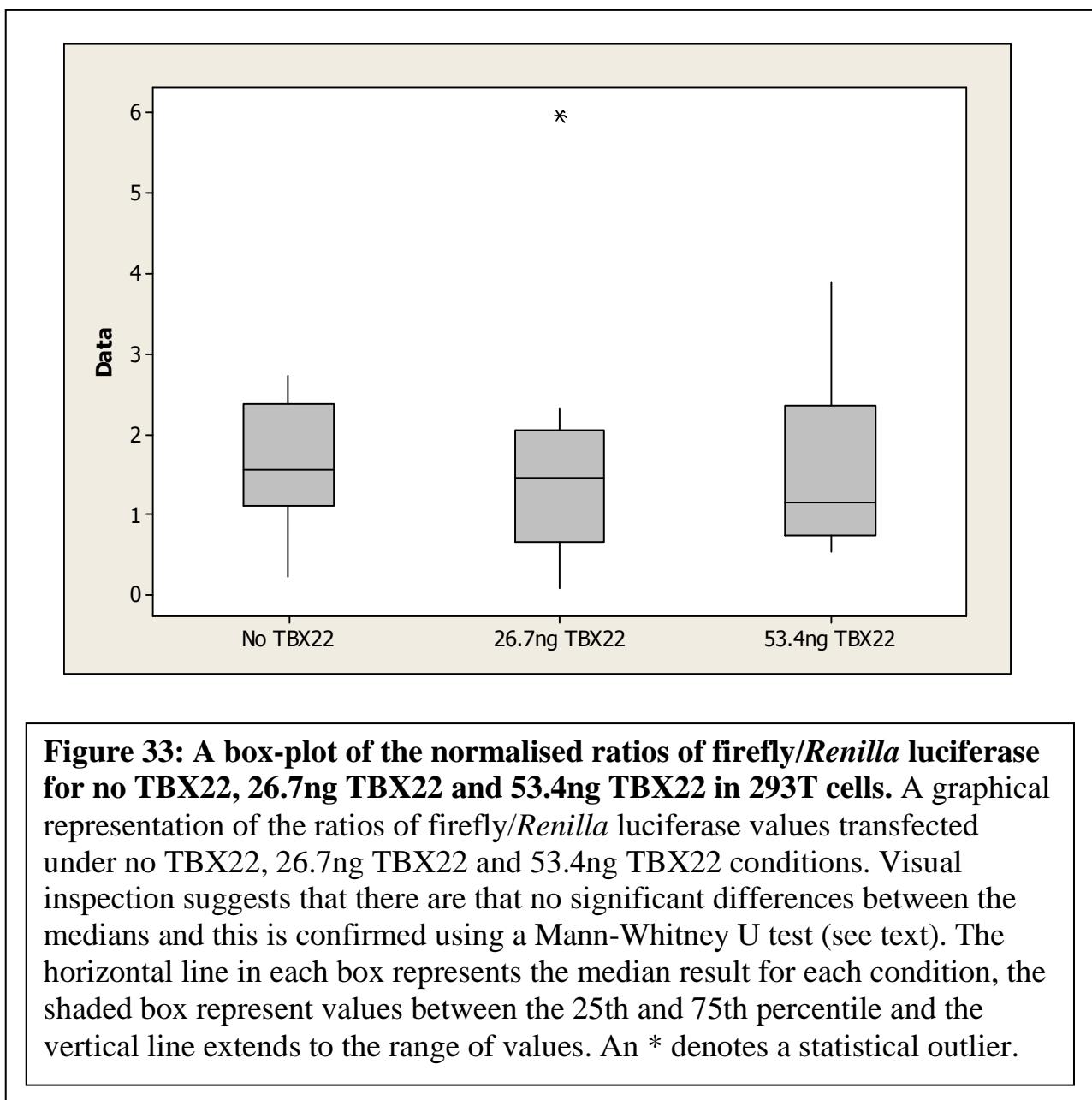


**Figure 31: An interval plot showing there was no significant difference between absolute relative light units emitted from the control *Renilla* luciferase.** The results of the 293T (a) and HeLa (b) cell transfections showing the distribution about the mean for each of the *Renilla* luciferase values for the three different transfection conditions: No TBX22 (pCR3.1\_Null plasmid), 26.7ng TBX22 or 53.4ng TBX22.

When the results in the HeLa cells are plotted (Fig. 32) it appears that there is a difference between the no TBX22 and both the 26.7ng or 53.4ng TBX22 conditions. The median result for the no TBX22 condition is 4.843, 26.7ng TBX22 - 2.719 and 53.4ng TBX22 - 3.251 (Fig. 32). The comparison of ratios of firefly to *Renilla* luciferase, between the no TBX22 and 26.7ng TBX22 experimental conditions is statistically significant using a Mann-Whitney U test at a 95% confidence interval with a p-value = 0.0006 and the comparison between no TBX22 and 53.4ng TBX22 conditions is significant with a p-value = 0.002. There was no statistical significance between the 26.7ng TBX22 and the 53.4ng TBX22 conditions (p-value = 0.3772).



The same Mann-Whitney U test was employed to test whether there was a statistically significant difference between the medians of the no TBX22, 26.7ng TBX22 and 53.4ng TBX22 transfection conditions in the 293T cells. Using a 95% confidence interval the difference between the no TBX22 and 26.7ng TBX22 conditions has a p-value = 0.7075 and between no TBX22 and 53.4ng TBX22 the p-value = 0.7508. Therefore, one must conclude that there is not a statistically significant difference between the no TBX22 and either the 53.4ng or 26.7ng TBX22 conditions. As above, the data have been represented in a box plot, the median values are no TBX22 – 1.666; 26.7ng TBX22 – 1.468 and 53.4ng TBX22 – 1.162; (Fig. 33). Figure 33 indicates visually that there are no differences between these three experimental conditions.



## 5.4 DISCUSSION

Initially it had been intended to study the effect of TBX22 on the *MSX1* promoter in a cell line derived from human embryonic palatal mesenchyme cells (HEPM – [ECACC# 90120505] (Yoneda and Pratt 1981)). This cell line would have been preferable to the HeLa cell line used, as they would have been a closer model to the environment of potential TBX22/*MSX1* interactions in the developing palate *in vivo*. However, these cells proved difficult to transfect using several transfection reagents (FuGENE6 (Roche), FuGENE HD (Promega), TransPass D1 (New England Biolabs) - data not shown. Although mammalian non-dividing cells are notoriously difficult cell lines to transfect (Subramanian *et al.* 1999), there are several dividing cell lines that are also known to be difficult to transfect using lipid-cation based transfer, for example the monocytic cell lines U937 and THP-1 (Martinet *et al.* 2003). A transfection method that was not tested in the present study, however, was electroporation; which has been shown to increase the transfection efficiency of cell lines that have proven challenging to transfect (Melkonyan *et al.* 1996). Although the technological advancements in chemical transfection are now generally considered to yield superior transfection efficiency to that achieved by electroporation, it is a method that may be capable of transfecting HEPM cells and could be employed in future studies.

The transfection strategy used in this study attempted to co-transfect three separate plasmids: the *MSX1* reporter plasmid (M\_prom; 6530bp), the internal control plasmid (pGL4.2[hRluc/Tk]; 4237bp) and either the TBX22 expression plasmid (pCR3.1\_TBX22; 6623bp) or expression control plasmid (pCR3.1\_Null;5045bp). Transfection of three plasmids of such size is a technically challenging task. However, as the M\_prom plasmid was successfully transfected and produced a protein which could be detected by Western blot (Fig. 30), then it is likely that the other plasmids were also transfected successfully.

Detection of TBX22 in the transfected cells was also attempted by Western blot using the same antibody used in Chapter 3 (section 3.2.5.B); however this did not result in any detectable signal. As TBX22 was detected by RT-PCR (Fig. 29) in the non-

transfected HeLa cells, at least some TBX22 protein would be expected to be detected in these cells. However, as this was not seen, it was assumed that either the antibody did not function in Western blots or there were problems with antibody specificity as discussed in Chapter 3. To circumvent the difficulty of using an unreliable anti-TBX22 antibody, the expression of a tagged TBX22 protein would have perhaps been a more suitable approach. There are several antibodies available to reliably detect protein tags, for example the 6xHis employed in the earlier bacterial expression experiments (see Chapter 3).

RT-PCR was considered to at least determine the presence of transfected TBX22 mRNA, however as the pCR3.1\_TBX22 plasmid contains only the TBX22 cDNA, then unless the cDNA could be prepared from RNA with absolutely no DNA contamination, it would have been impossible to tell whether the primers were annealing and amplifying the cDNA from the reverse transcribed RNA, or if in fact they would be amplifying contaminating pCR3.1\_TBX22 plasmid DNA. As it was considered that the preparation of cDNA could not be guaranteed to be completely devoid of DNA contamination, unfortunately, the presence of the transfected TBX22 protein in the 293T cells could not be verified and the possibility exists that no TBX22 protein is being produced by the pCR3.1\_TBX22 expression vector. Obviously if this plasmid is not driving expression of TBX22, then one cannot draw any conclusions from the effect TBX22 has upon the transcription levels of *MSX1* in this system.

The Dual Reporter system is based upon the assumption that the transfection efficiency of the internal control plasmid is proportional to that of the experimental reporter. As the *Renilla* control plasmid (pGL4.2[*hRluc*/Tk]) did not show any significant variance between the experiments (see Fig. 31), a reliable and reproducible transfection between experiments is occurring. Furthermore, by normalising the results to this internal control reporter (pGL4.2[*hRluc*/TK]) differences in transfection efficiency between assays should be accounted for.

There are various assumptions that have been made during the interpretation of the transfection study results, which have not been accounted for as the appropriate

controls were not incorporated into the experimental design. These assumptions and those controls that ought to have been employed to test these assumptions are shown in Table 22.

Assumption being made	Control to test for this assumption and expected outcome	Consequence to interpretation of results if assumption is not accurate
TBX22 protein is being produced by the pCR3.1_TBX22 expression plasmid.	Western blot using either anti-TBX22 antibody or an antibody to a tagged TBX22 protein construct to ensure TBX22 protein is being expressed.	If TBX22 is not being produced by the expression vector, the effects of TBX22 on the transcription of <i>MSX1</i> cannot be measured using this experiment.
TBX22 is binding to the <i>MSX1</i> promoter and not to some other region of the M_prom vector.	Replace the <i>MSX1</i> promoter with a promoter that will drive expression of the luciferase but does not contain a T-box binding site. Over expressing TBX22 should have no affect upon the luciferase levels of such a construct.	If TBX22 is binding to a region on the M_prom vector other than the <i>MSX1</i> promoter, then TBX22 does not regulate transcription of <i>MSX1</i> .
TBX22 is binding to the <i>MSX1</i> promoter at the identified T-box DNA binding domain and not at another region outside of this site.	Replace the T-box binding site of the <i>MSX1</i> promoter with the sequence of the Mut_oligo that did not show binding to TBX22 in the EMSA studies. Over expression of TBX22 should not alter luciferase levels of the <i>MSX1</i> promoter containing the Mut_oligo sequence.	TBX22 does not regulate transcription via the identified T-box binding domain (i.e. it is acting non-directly via some intermediary protein(s)), or the T-box binding domain is positioned at a different location on the <i>MSX1</i> promoter.
TBX22 is binding to the <i>MSX1</i> promoter and altering the level of transcription of a reporter vector.	Replace the TBX22 DNA binding domain with a mutated sequence that is known not to bind DNA. Over expressing this protein would not be expected to have an affect on the expression of luciferase.	TBX22 does not regulate transcription of <i>MSX1</i> via the identified T-box binding domain (i.e. it is acting non-directly via some intermediary protein(s)).
The M_prom construct is driving the expression of the luciferase.	Measure the amount of luciferase produced by a luciferase reporter construct containing the reverse complement of the <i>MSX1</i> promoter sequence. The level of luciferase expressed by such a construct would be expected to be significantly less than that seen with the M_prom vector with the <i>MSX1</i> promoter in the correct orientation. TBX22 would not be anticipated to bind to such a sequence.	The apparent expression of luciferase by the M_prom construct is actually due to low level expression of luciferase produced solely by the pGL4.17 vector. In which case altering the level of TBX22 expression will have no detectable effect upon transcription acting via the <i>MSX1</i> promoter.

**Table 22: A summary of the assumptions made in the interpretation of results for the transfection assay and the controls that should be included to verify that they hold true.**

Clearly there are several further controls needed to in order to ascertain whether the assumptions being made in these experiments are valid. The lack of these controls means that that definitive conclusions cannot be drawn from these studies. However, providing that the assumptions made in Table 22 are correct, the results of the transfection studies would suggest that over expression of TBX22 in the HeLa cell line reduced the expression from the *MSX1* reporter construct. Figure 32 shows an overall reduction in *M\_prom* reporter activity when co- transfected with the pCR3.1\_TBX22 plasmid, compared with co-transfection with the pCR31.\_Null plasmid. However, at the higher levels of TBX22 transfection, the reduction of expression appears more variable and with a slightly higher median than is seen in 26.7ng TBX22 conditions. Whilst the Mann-Whitney *U* test shows that differences seen in the medians between the 26.7ng pCR3.1\_TBX22 and 53.4ng pCR3.1\_TBX22 plasmids are not statistically significant, the small increase seen in the median of the 53.4ng pCR3.1\_TBX22 plasmid transfection may be due to an auto-regulatory effect of the endogenous TBX22 (Andreou *et al.* 2007) resulting in an overall net reduction in the amount of TBX22 protein in the cell.

The addition of TBX22 had no effect on the *MSX1* reporter in the 293T cells. As these cells were shown not to express *TBX22*, one could postulate that other factors required for the action of TBX22 may also be absent. SUMO1, for example, as sumoylation of TBX22 has been shown to be a necessary requirement for it to function as a repressor (Andreou *et al.* 2007).

The results of the transfection experiments hold up the possibility that TBX22 can regulate the expression of *MSX1* in HeLa cells. Evidence that TBX22 may bind to a DNA-binding site found in the *MSX1* promoter (Chapter 4), together with these results suggest that regulation of *MSX1* by TBX22 may be due to direct protein-DNA interaction. However, the *Tbx22* null mutant mouse did not display a significant difference in the distribution of expression of *Msx1* within the developing palatal shelves, although the authors do suggest that an increase in *Msx2* was seen in the posterior tongue (Pauws *et al.* 2009a). It is possible that there may be functional compensation by other T-box genes in the repression of *Msx1* in the palatal shelves *in vivo* in the absence of *Tbx22*, for example by *Tbx1* which has been shown to be necessary for secondary palate elongation (Goudy *et al.* 2010). Investigations in

development of the chick frontonasal process identified a role for *TBX22* in the regulation of proliferation in this region (Higashihori *et al.* 2010) and it is possible that *TBX22* exerts its effect over *MSX1* in the developing lip where both genes are expressed (Figs.13, 14 in Chapter 2 and Fig. 26 in Chapter 4). The authors also suggest that as *TBX22* has an involvement in proliferation within the frontonasal prominence that subtle changes to proliferation within the palatal shelves may also be occurring in the *Tbx22* null mouse. Therefore, the possibility exists that subtle changes in *Msx1* expression may also be seen within the palatal shelves of the *Tbx22*<sup>-/-</sup> mutant.

Evidence already exists for transcriptional repression by T-box genes. It has been shown that *TBX22* functions to repress expression of itself (Andreou *et al.* 2007) and both *TBX2* and *TBX3* have been shown to repress expression of target genes (Brummelkamp *et al.* 2002; Paxton *et al.* 2002; Prince *et al.* 2004). Interestingly, a conserved *MSX1* binding site has been identified within the 5'-region of the *TBX22* gene (Herr *et al.* 2003), raising the possibility of a regulatory feedback loop between *TBX22* and *MSX1*. Indeed it has been proposed that a similar regulation both upstream and downstream exists between *TBX22* and *MSX2* (Higashihori *et al.* 2010).

The fact that a difference in the effect of *TBX22* upon the repression on the *MSX1* reporter construct was seen between the HeLa and 293T cell lines supports the sensitivity and effectiveness of the Dual Reporter assay and it has been employed to study the effects of over expressing proteins on the transcription of luciferase promoter reporters in several studies (Liu *et al.* 2008; Ma *et al.* 2008; Zhang and Nohturfft 2008; Hou *et al.* 2009). The repression of *MSX1* was only seen in HeLa cells and whilst this is derived from a human cell line, it could still be considered that this is an artificial environment as they are aneuploid cells and as such are abnormal. However, as many functional studies have employed HeLa cells in the analysis of protein function, including that of Brachyury (Kispert *et al.* 1995) and also *TBX22* previously (Andreou *et al.* 2007), they remain a useful tool for understanding what is occurring *in vivo* under normal conditions.

## CHAPTER 6

### CONCLUSION AND FUTURE DIRECTIONS

#### 6.1 CONCLUSION

This thesis aimed to uncover the developmental expression pattern and protein-DNA binding properties of human *TBX22*, identify potential downstream target genes and evaluate one such gene. This has revealed several new insights in this field.

The spatiotemporal expression of human *TBX22* during the formation of the lip and palate was characterised and was shown to correlate with features disrupted in CPX. *TBX22*, in addition, is strongly expressed in the medial and lateral nasal processes during the formation of lip. This expression pattern has now also been demonstrated in mouse and chick (Braybrook *et al.* 2002; Bush *et al.* 2002; Haenig *et al.* 2002; Herr *et al.* 2003) and gene expression similarities within the formation of the oral cavity in zebrafish have also been uncovered (Jezewski *et al.* 2009). As a result of the common gene expression pattern seen in human and these animal models, the results of function studies of *TBX22* undertaken in these animal models (Pauws *et al.* 2009a; Higashihori *et al.* 2010) can be reliably extrapolated to man.

Several genes expressed in the palate during extension of the palatal shelves in mouse have expression domains restricted to either the anterior or posterior palatal shelves. *Tbx22* being one such gene with restricted expression to the posterior palate. This has exposed distinct regulatory pathways for the anterior and posterior palate formation (Li and Ding 2007; Pantalacci *et al.* 2008; Welsh and O'Brien 2009). In human the anterior palate is proportionately smaller in comparison to the posterior palate than in mouse. This makes it more difficult to distinguish between expression in the primary palate and expression in the anterior secondary palate in human. *TBX22* is clearly expressed in posterior secondary palate in human, similar to that reported for mouse and chick (Bush *et al.* 2002; Haenig *et al.* 2002; Herr *et al.* 2003)

A mouse lacking *Tbx22* has now been created (Pauws *et al.* 2009a) and this mutant displays many of the phenotypic features seen in CPX. The attachment of the tongue to the mandible at a more anterior position than in the wild type mouse was seen in all mutants, representing a mild form of ankyloglossia; a disruption to the formation of the palatal rugae was evident; in some mutants there was incomplete disruption to the oro-nasal membrane leading to choanal atresia ultimately resulting in post natal lethality; in surviving mutants a notch was seen in the posterior hard palate akin to submucous cleft palate and a few of the mutants analysed were born with an overt cleft palate. Further investigation of the secondary palate revealed a lack of mineralisation of the palatal bone and under development of the vomer, leading the authors to conclude that the major role of *Tbx22* is in the regulation of palatal osteoblast formation (Paws *et al.* 2009). However, as *TBX22* was seen to be expressed at the sites of the vertically elongating palatal shelves (Chapter 2), a further role for *TBX22* in the regulation the proliferation of the mesenchyme within the palatal shelves seems probable.

Experiments in chick have begun to uncover a function for *TBX22* during formation of the lip (Higashihori *et al.* 2010), where over-expression of *TBX22* in the chick frontonasal mass leads to a cleft lip. Although cleft lip is not a feature seen in the classic CPX phenotype, where mutations both to the coding and promoter regions in the affected individuals lead to an overall reduction of functional *TBX22* (Braybrook *et al.* 2001; Braybrook *et al.* 2002; Marcano *et al.* 2004; Andreou *et al.* 2007; Suphapeetiporn *et al.* 2007; Pauws *et al.* 2009b), Higashihori and colleagues postulate that genetic or environmental influences affecting the regulation of *TBX22*, resulting in a gain of functional *TBX22*, may indeed give rise to a cleft lip in humans. The expression data from human would support this, given the strong expression seen in the nasal processes during formation of the lip (Chapter 2).

The third Chapter of this thesis identified a *TBX22* DNA binding sequence that is similar to the Brachyury half site (Kispert and Herrmann 1993) and to half of the DNA sequence previously shown to bind to *TBX22* (Andreou *et al.* 2007). Different T-box proteins have been shown to bind to the same DNA sequence: *Xbra*, *TBX1* and

TBX2 were all shown to bind to the same DNA sequence - AGGTGTGAAAT (Sinha et al. 2000), although the mechanism by which this was achieved was different.

Whilst TBX1 bound to the DNA as a dimer in a similar manner to that reported for Xbra, TBX2 bound as a monomer. As the TBX22 DNA binding site reported in Chapter 3 resembles a Brachyury half site, it is likely that TBX22 binds to the DNA as a monomer, although confirmation of this would need to be provided using X-ray crystallography techniques, as has been shown for the human Brachyury protein (Papapetrou *et al.* 1997). The targeting of the same DNA binding site by different T-box proteins is postulated as one mechanism by which they regulate gene expression. During cardiac development, for example, it has been proposed that Tbx18 competes with Tbx5 for the same binding site on the *Nppa* promoter (Farin *et al.* 2007). Once bound Tbx18 acts as a repressor, whereas Tbx5 functions as a transcriptional activator. It would be enlightening to investigate whether other T-box genes, particularly the other TBX1 family genes which share over-lapping gene expression patterns during craniofacial development – TBX1/10/15 and 18 (Chapman *et al.* 1996; Agulnik *et al.* 1998; Kraus *et al.* 2001a; Bush *et al.* 2003), have a similar antagonistic regulatory mechanism in some or all cases.

Incorporating the TBX22 DNA binding site identified above with all of the previously determined T-box binding sites, a generic T-box binding site was devised. An *in silico* screen for the presence of this site within the promoter regions of 132 potential TBX22 target gene candidates is described in Chapter 4. This resulted in the identification of 102 genes, with at least one copy of a sequence that exactly matched, or had only one mis-match to the generic T-box binding site in the 2kb promoter region screened. Although only one gene was selected for further analysis, it is conceivable that all of these genes could be TBX22 targets. The actual number of TBX22 gene targets may well turn out to be much higher: a chromatin immunoprecipitation (ChIP) study aiming to identify T-bet targets, uncovered 832 protein coding target genes in human Th1 cells (Jenner *et al.* 2009) and a similar whole genome approach at identifying TBX22 targets may well uncover a large number of potential target genes. However, the *in silico* approach employed in this study focused upon a sub-set of candidate genes – those genes which when disrupted had been shown to cause a human cleft palate phenotype. This approach whilst

identifying a number of likely TBX22 target gene candidates, at the same time also excluded some genes that one could consider as potential TBX22 targets based upon known gene expression patterns, their presence in regions linked to cleft palate disorders or that give rise to cleft palate in mouse but have not yet been shown to have a role in human cleft palate disorders (Gritli-Linde 2007; Gritli-Linde 2008; Jugessur *et al.* 2009). Similarly, only the first 2kb of the promoter upstream for the transcription start site was screened for the presence of the binding site and regions outside of this region either further upstream or indeed within intronic gene regions would have been missed. Although it is not impossible that all of the human cleft palate causing genes may contain a TBX22 binding site, it would seem rather unlikely that all of these are indeed true *in vivo* TBX22 target genes. The *in silico* search employed in this study was probably always going to produce many false positive target genes. If one assumes that each of the 4 bases is equally represented throughout the genome then the generic T-box binding sequence used to search the promoter regions – AGGTGTBDWR - has an approximate 1:30,000 chance of appearing by chance alone. If the human genome has approximately 3 billion bases, then this would mean that there would be around 100,000 copies of this sequence occurring by chance in the genome. It is fanciful to imagine that TBX22 actually has anywhere near this number of target genes and therefore applying a similar search strategy to the entire genome would be unproductive and a more direct targeting approach, such as ChIP (Horak and Snyder 2002; Weinmann 2004; Wong and Wei 2009), is required to extend the search genome wide for TBX22 targets.

Although many genes were identified as potential TBX22 target gene candidates using the *in silico* screen, only 10 of these also display a cleft palate in the known mouse mutants for these genes (Table 18 Chapter 4 and Appendix 5). From this subset only two were known to be involved in non-syndromic cleft palate and as protein-protein interaction had already been demonstrated between one of these and TBX22 (Andreou *et al.* 2007), the other - *MSX1*, was selected for further investigation.

There was evidence that the level of transcription from an *MSX1* gene reporter (M\_prom, Chapter 5) was repressed in HeLa cells in which a TBX22 expression construct (pCR3.1\_TBX22, Chapter 5) had also been transfected. However, as several controls were absent from the transfection study experiments (see Table 22), the

mechanism responsible for the repression could not be identified with certainty. This repression was seen when both 26.7ng and 53.4ng of the TBX22 expression construct (pCR3.1\_TBX22) were transfected into the HeLa cells, although no dosage effect with the different amounts used was observed. Indeed, the repression was statistically equal in both cases. Repression was only observed in HeLa cells - a cell line known to endogenously express TBX22 (Andreou *et al.* 2007) and Fig. 29, Chapter 5). There was no difference in expression of the *MSX1* reporter in the presence or absence of TBX22 in the 293T cells, a cell-line that did not show endogenous *TBX22* expression (Fig. 29 in Chapter 5). Although it was not experimentally confirmed, the cells lacking endogenous TBX22 expression were postulated to also lack the necessary co-factors needed for TBX22 to repress expression from the *MSX1* reporter. The essential requirement for co-factors, such as YAP and TAZ, as well as other interacting T-box proteins for correct T-box protein *trans* regulation of target genes has previously reported (Murakami *et al.* 2005; Boogerd *et al.* 2008; Boogerd *et al.* 2009). TBX22 specifically has been shown to need modification by SUMO1 in order to be able to repress activity of downstream target genes (Andreou *et al.* 2007).

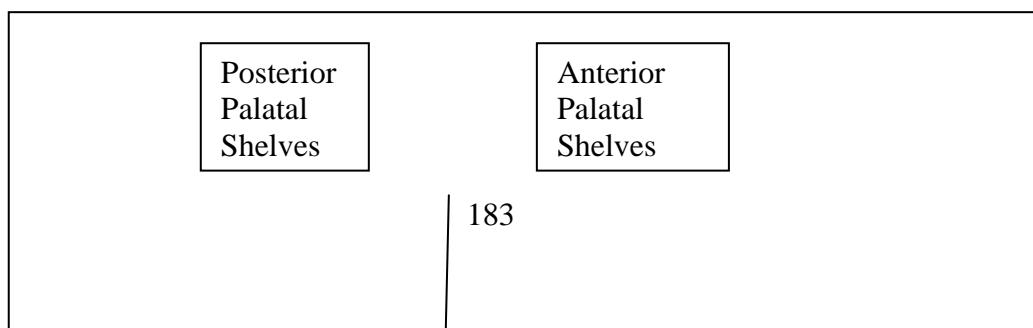
When taken together, the data presented within this thesis (the characterisation of the expression pattern of *TBX22* and of *MSX1* in the palate, the *in vitro* studies showing TBX22 probably binding to a DNA sequence found within the *MSX1* promoter and then the subsequent repression of an *MSX1* gene reporter in HeLa cells) suggest that TBX22 may have a role in the regulation of *MSX1* transcription. However, the lack of a suitable antibody has somewhat hampered the degree to which this conclusion can be made: a super-shift was not detected which meant that the EMSA studies could not convincingly show that the identified T-box binding site within the *MSX1* promoter was definitely binding to TBX22. Also, as the expression of TBX22 protein from the pCR3.1 expression vector employed in the transfection studies (Chapter 5) could not be verified, a completely conclusive statement that TBX22 does indeed negatively regulate transcription of *MSX1* cannot be made.

The work of others suggest that TBX22 regulates mesenchyme cell proliferation in the palatal shelves (see section 1.2.3) and studies in the chick frontonasal mass, suggest that TBX22 acts to negatively regulate proliferation in these cells

(Higashihori *et al.* 2010). The *in vitro* binding studies presented in this study point towards a possible repression of *MSX1* by TBX22.

A model of how this interaction could be acting during palatogenesis is now discussed. However, as the DNA binding site located within the *MSX1* promoter was identified from a generic T-box binding site, rather than matching the specific TBX22 DNA binding site consensus site, it is possible that any T-box gene maybe influencing *MSX1* transcription. This would be perhaps even more plausible if the expression of any or all of the human T-box gene family proteins are shown to overlap with that of TBX22. However, in the model discussed below it is assumed that the T-box gene in question is indeed TBX22.

*MSX1* and *TBX22* display complementary expression patterns in the palatal shelves: *MSX1* is restricted to the anterior palate, *TBX22* to the posterior (Fig. 26, Chapter 4). These restricted anterior/posterior expression patterns have been suggested to underlie the future hard and soft palate respectively (Welsh and O'Brien 2009). It is possible that TBX22, by down regulating *MSX1*, assists in maintaining the distinction between these two regions. During palatogenesis in the mouse, the anterior palate is extended (see 1.2.4) by proliferation of the mesenchyme regulated by *Bmp4* (Zhang *et al.* 2002). *Bmp* signalling, through ligand binding with various *Bmp*-receptors, has been shown to induce the activation of *Smad8*; which in turn promotes mitogenesis (Kawai *et al.* 2000). The mouse null mutant *Msx1*<sup>-/-</sup> has a cleft palate (Satokata and Maas 1994) that was shown to arise due to insufficient cell proliferation in the palate mesenchyme (Zhang *et al.* 2002). By ectopically expressing *Bmp4* in the palatal mesenchyme of the *Msx1*<sup>-/-</sup> mouse, the cleft palate phenotype was rescued. This demonstrated *Bmp4* to be downstream of *Msx1* (Zhang *et al.* 2002). Furthermore, the same authors showed that *Msx1* was required for the expression of *Bmp4*. Thus, by repressing *MSX1* expression in the posterior palate, TBX22 restricts *MSX1* expression, and therefore *BMP4* mediated proliferation to the anterior palate. A diagrammatic representation of this is shown in Figure 34 below.



TBX22

$\perp$   
MSX1

**Figure 34: A proposed model for repression of *MSX1* by *TBX22* in the formation of the palatal shelves.** *TBX22* represses *MSX1* in the posterior palatal shelves, thus limiting *MSX1* to the anterior palate only. *MSX1* is able to induce proliferation via *BMP4*, leading to extension of the anterior palate.

Identifying human transcription factor target genes is a valuable, but challenging task given the obvious limitations that exist compared to the gene targeted mouse models that can be created or transplantation experiments that can be performed in chick. This means that there may always be a reliance upon such animal models. However the *in vitro* binding studies in the cell culture based experiments used in this thesis are necessary to confirm that functional effects witnessed in animal models are applicable in human and to verify *in vitro* findings. Indeed the T-box targets identified by a chromatin immunoprecipitation study in developing zebrafish mesoderm (Garnett et al. 2009), were still verified using *in vitro* DNA binding assays similar to that employed in this study.

This present study took a candidate TBX22 target gene, based on gene expression and other reported evidence, and tested it for possible interaction with TBX22. However, an alternative, and perhaps more direct strategy may have been to first uncover possible target genes through wider reaching whole transcriptome or proteome studies. This could have been achieved by employing a ChIP (Jenner *et al.* 2009) or by a microarray approach (Bhattacherjee *et al.* 2007) – for example comparing the gene expression profile of palate cells in which TBX22 was over expressed to the gene expression profile of an unmodified cell, or conversely silenced using an RNAi, - using a technique similar to that employed to silence the T-box domain containing genes *Doc1*, 2, and 3 genes in drosophila (Hamaguchi *et al.* 2004). One or all of these approaches would have been employed had the studies described in this thesis commenced now.

Clearly the regulatory pathways involved in craniofacial development are only just beginning to emerge, as is the role of TBX22. However, since the identification of the *TBX22* gene in 2000 (Laugier-Anfossi and Villard 2000), there has been considerable progress as to the function of TBX22. It is now known that mutations to the coding region of TBX22 gave rise to CPX (Braybrook *et al.* 2001) and that alleles containing risk SNP's within the promoter region of TBX22 are associated with cleft palate and ankyloglossia (Pauws *et al.* 2009b). The spatiotemporal expression pattern of *TBX22* during craniofacial development in human and model animals has also been uncovered (Braybrook *et al.* 2002; Bush *et al.* 2002; Haenig *et al.* 2002; Herr *et al.* 2003). *Mnl* has been shown to be a transcriptional activator of *Tbx22* (Liu *et al.* 2008)

and it can be induced by Fgf8, Fgf2 and Noggin, whilst it is repressed by Bmp4 and Bmp7 (Higashihori *et al.* 2010). Human TBX22 has been demonstrated to be targeted for sumoylation and that this is necessary in order for it to exert repression on target genes (Andreou *et al.* 2007). To date the only downstream targets known are *TBX22* itself (Andreou *et al.* 2007) and possibly *MSX1* as identified from the outcome of this study.

Elucidating how *TBX22*, its target genes and the many other genes involved in the complex process of craniofacial development, fit together in the signalling and regulatory networks that are required for normal development is an admirable goal. Not only will these interacting pathways reveal how normal palatogenesis proceeds, but also how the effects of disruptions to this process can begin to be understood. With this understanding, the ultimate aim of prevention, treatment and counselling of patients with craniofacial disorders can be more fully realised. *TBX22* clearly has an important role in craniofacial development, the findings that many T-box genes, including *TBX22* are expressed in both overlapping and complementary patterns in the very early developing somites (Wardle and Papaioannou 2008) and the identification that *TBX22* was often deleted in a genome-wide screen of colorectal cancer biopsy samples (Ashktorab *et al.* 2010) points to other as yet unknown roles for *TBX22* outside of craniofacial development.

## 6.2 FUTURE DIRECTIONS

During the formation of the palate it is becoming apparent that the palatal rugae play an important role in both defining the anterior/posterior boundary and that they act as signalling centres regulating the extension of the anterior palate (Pantalacci *et al.* 2008; Welsh and O'Brien 2009). As with many of the models of human development our interpretation of this formation is based upon information gained from model organisms. It would be crucial to confirm or otherwise, whether the formation of the human palatal rugae are formed in a similar orchestrated manner to that observed in mice and that the timing of palatal rugae formation is comparable to mouse.

As the search for potential TBX22 target genes continues, a microarray based study would certainly enhance progress in this field, as would the information gained through chromatin immunoprecipitation genome-wide location analysis or ChIP-on-Chip (Horak and Snyder 2002; Weinmann 2004; Wong and Wei 2009). In this technique cells are cross-linked with formaldehyde to cross-link transcription factors whilst bound to DNA. An antibody to the transcription factor is then used to immunoprecipitate the protein-DNA and then the cross-links are reversed and the DNA purified. The purified DNA can then be labelled and used as probes to a genomic microarray containing selected enhancers and promoters. This technique has two major advantages over traditional microarray studies which rely on either the overexpression or silencing of a transcription factor. The first is that, as the technique is performed without altering the environment by artificially expressing or silencing a gene of interest, stoichiometry of potentially biological relevant proteins will remain. The second advantage is that ChIP-on-chip can determine direct targets of the transcription factor of interest, as opposed to that of the traditional methods which cannot discriminate between direct and indirect target genes. As ChIP-on-chip has been successfully employed in revealing several targets of the T-box protein T-bet (Townsend *et al.* 2004), ChIP-on-chip would seem ideally suited in identifying downstream TBX22 target genes. However, one essential requirement for a ChIP method is a specific and reliable TBX22 antibody, which was unfortunately lacking during the investigations performed in this thesis. Although efforts such as the Human Antibody Initiative (<http://www.hupo.org/research/hai/>), which aims to produce

antibodies against all human proteins and thoroughly test each one for its ability to perform immunoprecipitation, immunoblotting and immunohistochemistry, should address this difficulty in the future.

As well as uncovering target genes of TBX22, it is important to uncover protein-protein interactions between TBX22 and other functional protein partners which may affect transcription regulation. Recently, Msx1 and Msx2 have both been shown to bind through the homeodomain to the T-box of Tbx2 and these proteins function in concert to suppress the expression of Cx43 (Boogerd *et al.* 2008). It would obviously be very interesting to investigate whether there is any protein-protein interaction between TBX22 and MSX1. Investigation into possible interactions between TBX22 and other T-box genes in the regulation of the palate are also likely to be revealing as has been described during cardiac formation (Boogerd *et al.* 2009).

## BIBLIOGRAPHY

Acevedo, A. C., J. A. da Fonseca, J. Grinham, K. Doudney, R. R. Gomes, L. M. de Paula and P. Stanier (2010). "Autosomal-dominant ankyloglossia and tooth number anomalies." *J Dent Res* **89**(2): 128-32.

Agulnik, S. I., V. E. Papaioannou and L. M. Silver (1998). "Cloning, mapping, and expression analysis of TBX15, a new member of the T-Box gene family." *Genomics* **51**(1): 68-75.

Akahoshi, M., K. Obara, T. Hirota, A. Matsuda, K. Hasegawa, N. Takahashi, M. Shimizu, K. Nakashima, L. Cheng, S. Doi, H. Fujiwara, A. Miyatake, K. Fujita, N. Higashi, M. Taniguchi, T. Enomoto, X. Q. Mao, H. Nakashima, C. N. Adra, Y. Nakamura, M. Tamari and T. Shirakawa (2005). "Functional promoter polymorphism in the TBX21 gene associated with aspirin-induced asthma." *Hum Genet* **117**(1): 16-26.

Alappat, S. R., Z. Zhang, K. Suzuki, X. Zhang, H. Liu, R. Jiang, G. Yamada and Y. Chen (2005). "The cellular and molecular etiology of the cleft secondary palate in Fgf10 mutant mice." *Dev Biol* **277**(1): 102-13.

Al-Dajani, M. (2009). "Comparison of dental caries prevalence in patients with cleft lip and/or palate and their sibling controls." *Cleft Palate Craniofac J* **46**(5): 529-31.

Alkuraya, F. S., I. Saadi, J. J. Lund, A. Turbe-Doan, C. C. Morton and R. L. Maas (2006). "SUMO1 haploinsufficiency leads to cleft lip and palate." *Science* **313**(5794): 1751.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." *J Mol Biol* **215**(3): 403-10.

Alwine, J. C., D. J. Kemp and G. R. Stark (1977). "Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes." *Proc Natl Acad Sci U S A* **74**(12): 5350-4.

Andreou, A. M., E. Pauws, M. C. Jones, M. K. Singh, M. Bussen, K. Doudney, G. E. Moore, A. Kispert, J. J. Brosens and P. Stanier (2007). "TBX22 missense mutations found in patients with X-linked cleft palate affect DNA binding, sumoylation, and transcriptional repression." *Am J Hum Genet* **81**(4): 700-12.

Annweiler, A., R. A. Hipskind and T. Wirth (1991). "A strategy for efficient in vitro translation of cDNAs using the rabbit beta-globin leader sequence." *Nucleic Acids Res* **19**(13): 3750.

Ashktorab, H., A. A. Schaffer, M. Daramipouran, D. T. Smoot, E. Lee and H. Brim (2010). "Distinct genetic alterations in colorectal cancer." *PLoS One* **5**(1): e8879.

Austin, A. A., C. M. Druschel, M. C. Tyler, P. A. Romitti, West, II, P. C. Damiano, J. M. Robbins and W. Burnett (2010). "Interdisciplinary craniofacial teams compared with individual providers: is orofacial cleft care more comprehensive and do parents perceive better outcomes?" *Cleft Palate Craniofac J* **47**(1): 1-8.

Baala, L., S. Briault, H. C. Etchevers, F. Laumonnier, A. Natiq, J. Amiel, N. Boddaert, C. Picard, A. Sbiti, A. Asermouh, T. Attie-Bitach, F. Encha-Razavi, A. Munnoch, A. Sefiani and S. Lyonnet (2007). "Homozygous silencing of T-box transcription factor EOMES leads to microcephaly with polymicrogyria and corpus callosum agenesis." *Nat Genet* **39**(4): 454-6.

Bamshad, M., R. C. Lin, D. J. Law, W. C. Watkins, P. A. Krakowiak, M. E. Moore, P. Franceschini, R. Lala, L. B. Holmes, T. C. Gebuhr, B. G. Bruneau, A. Schinzel, J. G. Seidman, C. E. Seidman and L. B. Jorde (1997). "Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome." *Nat Genet* **16**(3): 311-5.

Barbieri, O., S. Astigiano, M. Morini, S. Tavella, A. Schito, A. Corsi, D. Di Martino, P. Bianco, R. Cancedda and S. Garofalo (2003). "Depletion of cartilage collagen fibrils in mice carrying a dominant negative Col2a1 transgene affects chondrocyte differentiation." *Am J Physiol Cell Physiol* **285**(6): C1504-12.

Barron, M. R., N. S. Belaguli, S. X. Zhang, M. Trinh, D. Iyer, X. Merlo, J. W. Lough, M. S. Parmacek, B. G. Bruneau and R. J. Schwartz (2005). "Serum response factor, an enriched cardiac mesoderm obligatory factor, is a downstream gene target for Tbx genes." *J Biol Chem* **280**(12): 11816-28.

Barrow, J. R. and M. R. Capecchi (1999). "Compensatory defects associated with mutations in Hoxa1 restore normal palatogenesis to Hoxa2 mutants." *Development* **126**(22): 5011-26.

Barrow, J. R., H. S. Stadler and M. R. Capecchi (2000). "Roles of Hoxa1 and Hoxa2 in patterning the early hindbrain of the mouse." *Development* **127**(5): 933-44.

Basson, C. T., D. R. Bachinsky, R. C. Lin, T. Levi, J. A. Elkins, J. Soultz, D. Grayzel, E. Kroumpouzou, T. A. Traill, J. Leblanc-Straceski, B. Renault, R. Kucherlapati, J. G. Seidman and C. E. Seidman (1997). "Mutations in human TBX5 [corrected] cause limb and cardiac malformation in Holt-Oram syndrome." *Nat Genet* **15**(1): 30-5.

Beaty, T. H., N. E. Maestri, J. B. Hetmanski, D. F. Wyszynski, C. A. Vanderkolk, J. C. Simpson, I. McIntosh, E. A. Smith, J. S. Zeiger, G. V. Raymond, S. R. Panny, C. J. Tiffet, A. F. Lewanda, C. A. Cristion and E. A. Wulfsberg (1997). "Testing for interaction between maternal smoking and TGFA genotype among oral cleft cases born in Maryland 1992-1996." *Cleft Palate Craniofac J* **34**(5): 447-54.

Beaumont, D. (2008). "A study into weight gain in infants with cleft lip/palate." *Paediatr Nurs* **20**(6): 20-3.

Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell and D. L. Wheeler (2008). "GenBank." *Nucleic Acids Res* **36**(Database issue): D25-30.

Berardi, M. J., C. Sun, M. Zehr, F. Abildgaard, J. Peng, N. A. Speck and J. H. Bushweller (1999). "The Ig fold of the core binding factor alpha Runt domain is a member of a family of structurally and functionally related Ig-fold DNA-binding domains." *Structure* **7**(10): 1247-56.

Beverdam, A., A. Brouwer, M. Reijnen, J. Korving and F. Meijlink (2001). "Severe nasal clefting and abnormal embryonic apoptosis in Alx3/Alx4 double mutant mice." *Development* **128**(20): 3975-86.

Bhattacherjee, V., P. Mukhopadhyay, S. Singh, C. Johnson, J. T. Philipose, C. P. Warner, R. M. Greene and M. M. Pisano (2007). "Neural crest and mesoderm lineage-dependent gene expression in orofacial development." *Differentiation*.

Bille, C., A. Skytthe, W. Vach, L. B. Knudsen, A. M. Andersen, J. C. Murray and K. Christensen (2005). "Parent's age and the risk of oral clefts." *Epidemiology* **16**(3): 311-6.

Bjornsson, A., A. Arnason and P. Tippet (1989). "X-linked cleft palate and ankyloglossia in an Icelandic family." *Cleft Palate J* **26**(1): 3-8.

Blake, J. A., C. J. Bult, J. T. Eppig, J. A. Kadin and J. E. Richardson (2009). "The Mouse Genome Database genotypes::phenotypes." *Nucleic Acids Res* **37**(Database issue): D712-9.

Blanton, S. H., A. Cortez, S. Stal, J. B. Mulliken, R. H. Finnell and J. T. Hecht (2005). "Variation in IRF6 contributes to nonsyndromic cleft lip and palate." *Am J Med Genet A* **137**(3): 259-62.

Bohren, K. M., V. Nadkarni, J. H. Song, K. H. Gabbay and D. Owerbach (2004). "A M55V polymorphism in a novel SUMO gene (SUMO-4) differentially activates heat shock transcription factors and is associated with susceptibility to type I diabetes mellitus." *J Biol Chem* **279**(26): 27233-8.

Bongers, E. M., P. H. Duijf, S. E. van Beersum, J. Schoots, A. Van Kampen, A. Burckhardt, B. C. Hamel, F. Losan, L. H. Hoefsloot, H. G. Yntema, N. V. Knoers and H. van Bokhoven (2004). "Mutations in the human TBX4 gene cause small patella syndrome." *Am J Hum Genet* **74**(6): 1239-48.

Boogerd, C. J., A. F. Moorman and P. Barnett (2009). "Protein interactions at the heart of cardiac chamber formation." *Ann Anat* **191**(6): 505-17.

Boogerd, K. J., L. Y. Wong, V. M. Christoffels, M. Klarenbeek, J. M. Ruijter, A. F. Moorman and P. Barnett (2008). "Msx1 and Msx2 are functional interacting partners of T-box factors in the regulation of Connexin43." *Cardiovasc Res*.

Braybrook, C., K. Doudney, A. C. Marcano, A. Arnason, A. Bjornsson, M. A. Patton, P. J. Goodfellow, G. E. Moore and P. Stanier (2001). "The T-box transcription factor gene TBX22 is mutated in X-linked cleft palate and ankyloglossia." *Nat Genet* **29**(2): 179-83.

Braybrook, C., S. Lisgo, K. Doudney, D. Henderson, A. C. Marcano, T. Strachan, M. A. Patton, L. Villard, G. E. Moore, P. Stanier and S. Lindsay (2002). "Craniofacial expression of human and murine TBX22 correlates with the cleft palate and ankyloglossia phenotype observed in CPX patients." *Hum Mol Genet* **11**(22): 2793-804.

Brewer, C. M., J. P. Leek, A. J. Green, S. Holloway, D. T. Bonthron, A. F. Markham and D. R. FitzPatrick (1999). "A locus for isolated cleft palate, located on human chromosome 2q32." *Am J Hum Genet* **65**(2): 387-96.

Brinkley, L., J. Morris-Wiman and M. Bernfield (1992). "The distribution of syndecan during murine secondary palate morphogenesis." *J Craniofac Genet Dev Biol* **12**(2): 82-9.

Brinkley, L. L. (1984). "Changes in cell distribution during mouse secondary palate closure in vivo and in vitro. I. Epithelial cells." *Dev Biol* **102**(1): 216-27.

Brinkley, L. L. and F. L. Bookstein (1986). "Cell distribution during mouse secondary palate closure. II. Mesenchymal cells." *J Embryol Exp Morphol* **96**: 111-30.

Brinkley, L. L. and J. Morris-Wiman (1987). "Computer-assisted analysis of hyaluronate distribution during morphogenesis of the mouse secondary palate." *Development* **100**(4): 629-35.

Bronner-Fraser, M. (1995). "Origins and developmental potential of the neural crest." *Exp Cell Res* **218**(2): 405-17.

Brummelkamp, T. R., R. M. Kortlever, M. Lingbeek, F. Trettel, M. E. MacDonald, M. van Lohuizen and R. Bernards (2002). "TBX-3, the gene mutated in Ulnar-Mammary Syndrome, is a negative regulator of p19ARF and inhibits senescence." *J Biol Chem* **277**(8): 6567-72.

Bruneau, B. G., M. Logan, N. Davis, T. Levi, C. J. Tabin, J. G. Seidman and C. E. Seidman (1999). "Chamber-specific cardiac expression of Tbx5 and heart defects in Holt-Oram syndrome." *Dev Biol* **211**(1): 100-8.

Bruneau, B. G., G. Nemer, J. P. Schmitt, F. Charron, L. Robitaille, S. Caron, D. A. Conner, M. Gessler, M. Nemer, C. E. Seidman and J. G. Seidman (2001). "A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease." *Cell* **106**(6): 709-21.

Bullen, P. and D. Wilson (1997). The Carnegie staging of human embryos: a practical guide. *Molecular Genetics of Early Human Development*. T. Strachan, S. Lindsay and D. I. Wilson. Oxford, Bios Scientific Publishers: 27-35.

Bullen, P. J., S. C. Robson and T. Strachan (1998). "Human post-implantation embryo collection: medical and surgical techniques." *Early Hum Dev* **51**(3): 213-21.

Bult, C. J., J. T. Eppig, J. A. Kadin, J. E. Richardson and J. A. Blake (2008). "The Mouse Genome Database (MGD): mouse biology and model systems." *Nucleic Acids Res* **36**(Database issue): D724-8.

Bush, J. O., Y. Lan, K. M. Maltby and R. Jiang (2002). "Isolation and developmental expression analysis of Tbx22, the mouse homolog of the human X-linked cleft palate gene." *Dev Dyn* **225**(3): 322-6.

Bush, J. O., K. M. Maltby, E. S. Cho and R. Jiang (2003). "The T-box gene Tbx10 exhibits a uniquely restricted expression pattern during mouse embryogenesis." *Gene Expr Patterns* **3**(4): 533-8.

Bussen, M., M. Petry, K. Schuster-Gossler, M. Leitges, A. Gossler and A. Kispert (2004). "The T-box transcription factor Tbx18 maintains the separation of anterior and posterior somite compartments." *Genes Dev* **18**(10): 1209-21.

Butler, H. and B. H. J. Juurlink (1987). *An atlas for staging mammalian and chick embryos*. Florida, CRC Press Inc.

C.D.C., C. f. D. C. a. P. (2006). "Improved national prevalence estimates for 18 selected major birth defects--United States, 1999-2001." *MMWR Morb Mortal Wkly Rep* **54**(51): 1301-5.

Cacheux, V., F. Dastot-Le Moal, H. Kaariainen, N. Bondurand, R. Rintala, B. Boissier, M. Wilson, D. Mowat and M. Goossens (2001). "Loss-of-function mutations in SIP1 Smad interacting protein 1 result in a syndromic Hirschsprung disease." *Hum Mol Genet* **10**(14): 1503-10.

Carinci, F., L. Scapoli, A. Palmieri, I. Zollino and F. Pezzetti (2007). "Human genetic factors in nonsyndromic cleft lip and palate: an update." *Int J Pediatr Otorhinolaryngol* **71**(10): 1509-19.

Carlson, H., S. Ota, C. E. Campbell and P. J. Hurlin (2001). "A dominant repression domain in Tbx3 mediates transcriptional repression and cell immortalization: relevance to mutations in Tbx3 that cause ulnar-mammary syndrome." *Hum Mol Genet* **10**(21): 2403-13.

Carstens, M. H. (2002). "Development of the facial midline." *J Craniofac Surg* **13**(1): 129-87; discussion 188-90.

Carter, T. C., A. M. Molloy, F. Pangilinan, J. F. Troendle, P. N. Kirke, M. R. Conley, D. J. Orr, M. Earley, E. McKiernan, E. C. Lynn, A. Doyle, J. M. Scott, L. C. Brody and J. L. Mills (2010). "Testing reported associations of genetic risk factors for oral clefts in a large Irish study population." *Birth Defects Res A Clin Mol Teratol* **88**(2): 84-93.

Casey, E. S., M. A. O'Reilly, F. L. Conlon and J. C. Smith (1998). "The T-box transcription factor Brachyury regulates expression of eFGF through binding to a non-palindromic response element." *Development* **125**(19): 3887-94.

Casey, E. S., M. Tada, L. Fairclough, C. C. Wylie, J. Heasman and J. C. Smith (1999). "Bix4 is activated directly by VegT and mediates endoderm formation in Xenopus development." *Development* **126**(19): 4193-200.

Cedergren, M. and B. Kallen (2005). "Maternal obesity and the risk for orofacial clefts in the offspring." *Cleft Palate Craniofac J* **42**(4): 367-71.

Celli, J., P. Duijf, B. C. Hamel, M. Bamshad, B. Kramer, A. P. Smits, R. Newbury-Ecob, R. C. Hennekam, G. Van Buggenhout, A. van Haeringen, C. G. Woods, A. J. van Essen, R. de Waal, G. Vriend, D. A. Haber, A. Yang, F. McKeon, H. G. Brunner and H. van Bokhoven (1999). "Heterozygous germline mutations in the p53 homolog p63 are the cause of EEC syndrome." *Cell* **99**(2): 143-53.

Chaabouni, M., N. Smaoui, N. Benneji, R. M'Rad, L. B. Jemaa, S. Hachicha and H. Chaabouni (2005). "Mutation analysis of TBX22 reveals new mutation in Tunisian CPX family." *Clin Dysmorphol* **14**(1): 23-5.

Chai, Y. and R. E. Maxson, Jr. (2006). "Recent advances in craniofacial morphogenesis." *Dev Dyn* **235**(9): 2353-75.

Chapman, D. L., N. Garvey, S. Hancock, M. Alexiou, S. I. Agulnik, J. J. Gibson-Brown, J. Cebra-Thomas, R. J. Bollag, L. M. Silver and V. E. Papaioannou (1996). "Expression of the T-box family genes, Tbx1-Tbx5, during early mouse development." *Dev Dyn* **206**(4): 379-90.

Chapman, D. L. and V. E. Papaioannou (1998). "Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6." *Nature* **391**(6668): 695-7.

Chenevix-Trench, G., S. Healey and N. G. Martin (1993). "Reproductive hormone genes in mothers of spontaneous dizygotic twins: an association study." *Hum Genet* **91**(2): 118-20.

Chiang, P. W., W. J. Song, K. Y. Wu, J. R. Korenberg, E. J. Fogel, M. L. Van Keuren, D. Lashkari and D. M. Kurnit (1996). "Use of a fluorescent-PCR reaction to detect genomic sequence copy number and transcriptional abundance." *Genome Res* **6**(10): 1013-26.

Chiquet, B. T., S. H. Blanton, A. Burt, D. Ma, S. Stal, J. B. Mulliken and J. T. Hecht (2008). "Variation in WNT genes is associated with non-syndromic cleft lip with or without cleft palate." *Hum Mol Genet* **17**(14): 2212-8.

Chiquet, B. T., A. C. Lidral, S. Stal, J. B. Mulliken, L. M. Moreno, M. Arcos-Burgos, C. Valencia-Ramirez, S. H. Blanton and J. T. Hecht (2007). "CRISPLD2: a novel NSCLP candidate gene." *Hum Mol Genet* **16**(18): 2241-8.

Christensen, K., J. Olsen, B. Norgaard-Pedersen, O. Basso, H. Stovring, L. Milhollin-Johnson and J. C. Murray (1999). "Oral clefts, transforming growth factor alpha gene variants, and maternal smoking: a population-based case-control study in Denmark, 1991-1994." *Am J Epidemiol* **149**(3): 248-55.

Clancy, B., R. B. Darlington and B. L. Finlay (2001). "Translating developmental time across mammalian species." *Neuroscience* **105**(1): 7-17.

Clancy, B., B. L. Finlay, R. B. Darlington and K. J. Anand (2007). "Extrapolating brain development from experimental species to humans." *Neurotoxicology* **28**(5): 931-7.

Cole, A., J. Tomlinson, R. Slator and J. Reading (2009). "Understanding cleft lip and palate. 3: feeding the baby." *J Fam Health Care* **19**(5): 157-8.

Coll, M., J. G. Seidman and C. W. Muller (2002). "Structure of the DNA-bound T-box domain of human TBX3, a transcription factor responsible for ulnar-mammary syndrome." *Structure* **10**(3): 343-56.

Compagni, A., M. Logan, R. Klein and R. H. Adams (2003). "Control of skeletal patterning by ephrinB1-EphB interactions." *Dev Cell* **5**(2): 217-30.

Condie, B. G. and M. R. Capecchi (1993). "Mice homozygous for a targeted disruption of Hoxd-3 (Hox-4.1) exhibit anterior transformations of the first

and second cervical vertebrae, the atlas and the axis." *Development* **119**(3): 579-95.

Conlon, F. L., L. Fairclough, B. M. Price, E. S. Casey and J. C. Smith (2001). "Determinants of T box protein specificity." *Development* **128**(19): 3749-58.

Couly, G. and N. M. Le Douarin (1990). "Head morphogenesis in embryonic avian chimeras: evidence for a segmental pattern in the ectoderm corresponding to the neuromeres." *Development* **108**(4): 543-58.

Creuzet, S., G. Couly and N. M. Le Douarin (2005). "Patterning the neural crest derivatives during development of the vertebrate head: insights from avian studies." *J Anat* **207**(5): 447-59.

D'Angelo, M. and R. M. Greene (1991). "Transforming growth factor-beta modulation of glycosaminoglycan production by mesenchymal cells of the developing murine secondary palate." *Dev Biol* **145**(2): 374-8.

Darby, I. A., A. Desmouliere and T. D. Hewitson (2009). "Using *in situ* hybridisation to localize renal gene expression in tissue sections." *Methods Mol Biol* **466**: 119-132.

Davenport, T. G., L. A. Jerome-Majewska and V. E. Papaioannou (2003). "Mammary gland, limb and yolk sac defects in mice lacking Tbx3, the gene mutated in human ulnar mammary syndrome." *Development* **130**(10): 2263-73.

De Felice, M., C. Ovitt, E. Biffali, A. Rodriguez-Mallon, C. Arra, K. Anastassiadis, P. E. Macchia, M. G. Mattei, A. Mariano, H. Scholer, V. Macchia and R. Di Lauro (1998). "A mouse model for hereditary thyroid dysgenesis and cleft palate." *Nat Genet* **19**(4): 395-8.

De Mey, A., D. Franck, N. Cuylits, G. Swennen, C. Malevez and M. Lejour (2009). "Early one-stage repair of complete unilateral cleft lip and palate." *J Craniofac Surg* **20 Suppl 2**: 1723-8.

DeAngelis, V. and J. Nalbandian (1968). "Ultrastructure of mouse and rat palatal processes prior to and during secondary palate formation." *Arch Oral Biol* **13**(6): 601-8.

Degitz, S. J., B. M. Francis and G. L. Foley (1998). "Mesenchymal changes associated with retinoic acid induced cleft palate in CD-1 mice." *J Craniofac Genet Dev Biol* **18**(2): 88-99.

Di Gregorio, A. and M. Levine (1999). "Regulation of Ci-tropomyosin-like, a Brachury target gene in the ascidian, *Ciona intestinalis*." *Development* **126**(24): 5599-609.

Diewert, V. M. (1983). "A morphometric analysis of craniofacial growth and changes in spatial relations during secondary palatal development in human embryos and fetuses." *Am J Anat* **167**(4): 495-522.

Ding, H., X. Wu, H. Bostrom, I. Kim, N. Wong, B. Tsoi, M. O'Rourke, G. Y. Koh, P. Soriano, C. Betsholtz, T. C. Hart, M. L. Marazita, L. L. Field, P. P. Tam and A. Nagy (2004). "A specific requirement for PDGF-C in palate formation and PDGFR-alpha signaling." *Nat Genet* **36**(10): 1111-6.

Dobreva, G., J. Dambacher and R. Grosschedl (2003). "SUMO modification of a novel MAR-binding protein, SATB2, modulates immunoglobulin mu gene expression." *Genes Dev* **17**(24): 3048-61.

Dobrovolskia-Zavadskaa, N. (1927). "La spontanée de la queue chez la souris nouveau-née et sur l'existence d'un caractère héréditaire "non-viable" {Translation: on the spontaneous necrosis of the tail in the newborn mouse and on the existence of a hereditary phenotype "non-viable"}." *C R Soc Biol* **97**: 114-116.

Drake, R., W. A. Vogl and A. W. M. Mitchell (2009). Gray's Anatomy for Students. London, Churchill Livingstone.

Dudas, M., W. Y. Li, J. Kim, A. Yang and V. Kaartinen (2007). "Palatal fusion - where do the midline cells go? A review on cleft palate, a major human birth defect." Acta Histochem **109**(1): 1-14.

Evdokimov, E., P. Sharma, S. J. Lockett, M. Lualdi and M. R. Kuehn (2008). "Loss of SUMO1 in mice affects RanGAP1 localization and formation of PML nuclear bodies, but is not lethal as it can be compensated by SUMO2 or SUMO3." J Cell Sci **121**(Pt 24): 4106-13.

Farin, H. F., M. Bussen, M. K. Schmidt, M. K. Singh, K. Schuster-Gossler and A. Kispert (2007). "Transcriptional repression by the T-box proteins Tbx18 and Tbx15 depends on Groucho corepressors." J Biol Chem **282**(35): 25748-59.

Felix, T. M., B. C. Hanshaw, R. Mueller, P. Bitoun and J. C. Murray (2006). "CHD7 gene and non-syndromic cleft lip and palate." Am J Med Genet A **140**(19): 2110-4.

Ferguson, M. W. (1981). "Developmental mechanisms in normal and abnormal palate formation with particular reference to the aetiology, pathogenesis and prevention of cleft palate." Br J Orthod **8**(3): 115-37.

Ferguson, M. W. (1988). "Palate development." Development **103 Suppl**: 41-60.

Ferguson, M. W., L. S. Honig and H. C. Slavkin (1984). "Differentiation of cultured palatal shelves from alligator, chick, and mouse embryos." Anat Rec **209**(2): 231-49.

Fitchett, J. E. and E. D. Hay (1989). "Medial edge epithelium transforms to mesenchyme after embryonic palatal shelves fuse." Dev Biol **131**(2): 455-74.

FitzPatrick, D. R., I. M. Carr, L. McLaren, J. P. Leek, P. Wightman, K. Williamson, P. Gautier, N. McGill, C. Hayward, H. Firth, A. F. Markham, J. A. Fantes and D. T. Bonthron (2003). "Identification of SATB2 as the cleft palate gene on 2q32-q33." Hum Mol Genet **12**(19): 2491-501.

Flynn, T., C. Moller, R. Jonsson and A. Lohmander (2009). "The high prevalence of otitis media with effusion in children with cleft lip and palate as compared to children without clefts." Int J Pediatr Otorhinolaryngol.

Fogh-Anderson, P. (1942). Inheritance of a harelip and cleft palate. Nyt Nordisk Forlag. Copenhagen.

Fougerousse, F., P. Bullen, M. Herasse, S. Lindsay, I. Richard, D. Wilson, L. Suel, M. Durand, S. Robson, M. Abitbol, J. S. Beckmann and T. Strachan (2000). "Human-mouse differences in the embryonic expression patterns of developmental control genes and disease genes." Hum Mol Genet **9**(2): 165-73.

Fraser, F. C. (1970). "The genetics of cleft lip and cleft palate." Am J Hum Genet **22**(3): 336-52.

Fraser, F. C. (1976). "The multifactorial/threshold concept - uses and misuses." Teratology **14**(3): 267-80.

Fraser, F. C. (1996). "Liability, thresholds, malformations, and syndromes." Am J Med Genet **66**(1): 75-6.

Fraser, F. C. (1998). "Some overlooked properties of the multifactorial/threshold model." Am J Hum Genet **62**(5): 1262-5.

Gagnon, A., R. D. Wilson, V. M. Allen, F. Audibert, C. Blight, J. A. Brock, V. A. Desilets, J. A. Johnson, S. Langlois, L. Murphy-Kaulbeck and P. Wyatt (2009). "Evaluation of prenatally diagnosed structural congenital anomalies." J Obstet Gynaecol Can **31**(9): 875-81, 882-9.

Gall, J. G. and M. L. Pardue (1969). "Formation and detection of RNA-DNA hybrid molecules in cytological preparations." *Proc Natl Acad Sci U S A* **63**(2): 378-83.

Garnett, A. T., T. M. Han, M. J. Gilchrist, J. C. Smith, M. B. Eisen, F. C. Wardle and S. L. Amacher (2009). "Identification of direct T-box target genes in the developing zebrafish mesoderm." *Development* **136**(5): 749-60.

Gendron-Maguire, M., M. Mallo, M. Zhang and T. Gridley (1993). "Hoxa-2 mutant mice exhibit homeotic transformation of skeletal elements derived from cranial neural crest." *Cell* **75**(7): 1317-31.

Gey, G. O., W. D. Coffman and M. T. Kubicek (1952). "SCIENTIFIC proceedings: American Association for Cancer Research, Inc., New York, N.Y., April 11-13, 1952." *Cancer Res* **12**(4): 264-265.

Ghosh, T. K., E. A. Packham, A. J. Bonser, T. E. Robinson, S. J. Cross and J. D. Brook (2001). "Characterization of the TBX5 binding site and analysis of mutations that cause Holt-Oram syndrome." *Hum Mol Genet* **10**(18): 1983-94.

Gibson, U. E., C. A. Heid and P. M. Williams (1996). "A novel method for real time quantitative RT-PCR." *Genome Res* **6**(10): 995-1001.

Gonzalez, S. M., L. H. Ferland, B. Robert and E. Abdelhay (1998). "Structural and functional analysis of mouse Msx1 gene promoter: sequence conservation with human MSX1 promoter points at potential regulatory elements." *DNA Cell Biol* **17**(6): 561-72.

Gorski, S. M., K. J. Adams, P. H. Birch, B. N. Chodirker, C. R. Greenberg and P. J. Goodfellow (1994). "Linkage analysis of X-linked cleft palate and ankyloglossia in Manitoba Mennonite and British Columbia Native kindreds." *Hum Genet* **94**(2): 141-8.

Gorski, S. M., K. J. Adams, P. H. Birch, J. M. Friedman and P. J. Goodfellow (1992). "The gene responsible for X-linked cleft palate (CPX) in a British Columbia native kindred is localized between PGK1 and DXYS1." *Am J Hum Genet* **50**(5): 1129-36.

Goudy, S., A. Law, G. Sanchez, H. S. Baldwin and C. Brown (2010). "Tbx1 is necessary for palatal elongation and elevation." *Mech Dev*.

Graham, A., I. Heyman and A. Lumsden (1993). "Even-numbered rhombomeres control the apoptotic elimination of neural crest cells from odd-numbered rhombomeres in the chick hindbrain." *Development* **119**(1): 233-45.

Graham, A. and A. Smith (2001). "Patterning the pharyngeal arches." *Bioessays* **23**(1): 54-61.

Gray, P. A., H. Fu, P. Luo, Q. Zhao, J. Yu, A. Ferrari, T. Tenzen, D. I. Yuk, E. F. Tsung, Z. Cai, J. A. Alberta, L. P. Cheng, Y. Liu, J. M. Stenman, M. T. Valerius, N. Billings, H. A. Kim, M. E. Greenberg, A. P. McMahon, D. H. Rowitch, C. D. Stiles and Q. Ma (2004). "Mouse brain organization revealed through direct genome-scale TF expression analysis." *Science* **306**(5705): 2255-7.

Greene, R. M. and D. M. Kochhar (1974). "Surface coat on the epithelium of developing palatine shelves in the mouse as revealed by electron microscopy." *J Embryol Exp Morphol* **31**(3): 683-92.

Gregory, P. A., A. G. Bert, E. L. Paterson, S. C. Barry, A. Tsykin, G. Farshid, M. A. Vadas, Y. Khew-Goodall and G. J. Goodall (2008). "The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1." *Nat Cell Biol* **10**(5): 593-601.

Gritli-Linde, A. (2007). "Molecular control of secondary palate development." *Dev Biol* **301**(2): 309-26.

Gritli-Linde, A. (2008). "The etiopathogenesis of cleft lip and cleft palate: usefulness and caveats of mouse models." *Curr Top Dev Biol* **84**: 37-138.

Grosen, D., C. Chevrier, A. Skytthe, C. Bille, K. Molsted, A. Sivertsen, J. C. Murray and K. Christensen (2009). "A cohort study of recurrence patterns among more than 54,000 relatives of oral cleft cases in Denmark: support for the multifactorial threshold model of inheritance." *J Med Genet*.

Gupta, V. and M. Bei (2006). "Modification of Msx1 by SUMO-1." *Biochem Biophys Res Commun* **345**(1): 74-7.

Hader, C., A. Marlier and L. Cantley (2010). "Mesenchymal-epithelial transition in epithelial response to injury: the role of Foxc2." *Oncogene* **29**(7): 1031-40.

Haenig, B., C. Schmidt, F. Kraus, M. Pfordt and A. Kispert (2002). "Cloning and expression analysis of the chick ortholog of TBX22, the gene mutated in X-linked cleft palate and ankyloglossia." *Mech Dev* **117**(1-2): 321-5.

Hamaguchi, T., S. Yabe, H. Uchiyama and R. Murakami (2004). "Drosophila Tbx6-related gene, Dorsocross, mediates high levels of Dpp and Scw signal required for the development of amnioserosa and wing disc primordium." *Dev Biol* **265**(2): 355-68.

Hamburger, V. and H. L. Hamilton (1992). "A series of normal stages in the development of the chick embryo. 1951." *Dev Dyn* **195**(4): 231-72.

Hargrave, M., J. Bowles and P. Koopman (2006). "In situ hybridization of whole-mount embryos." *Methods Mol Biol* **326**: 103-13.

Harrelson, Z., R. G. Kelly, S. N. Goldin, J. J. Gibson-Brown, R. J. Bollag, L. M. Silver and V. E. Papaioannou (2004). "Tbx2 is essential for patterning the atrioventricular canal and for morphogenesis of the outflow tract during heart development." *Development* **131**(20): 5041-52.

Haworth, K. E., C. Healy, P. Morgan and P. T. Sharpe (2004). "Regionalisation of early head ectoderm is regulated by endoderm and prepatterns the orofacial epithelium." *Development* **131**(19): 4797-806.

Heid, C. A., J. Stevens, K. J. Livak and P. M. Williams (1996). "Real time quantitative PCR." *Genome Res* **6**(10): 986-94.

Heilbronner, C. (2005). "[What are the short, mid, and long term consequences of smoking during pregnancy?]." *J Gynecol Obstet Biol Reprod (Paris)* **34 Spec No 1**: 3S390-446.

Hern, W. M. (1984). "Correlation of fetal age and measurements between 10 and 26 weeks of gestation." *Obstet Gynecol* **63**(1): 26-32.

Herr, A., D. Meunier, I. Muller, A. Rump, R. Fundele, H. H. Ropers and U. A. Nuber (2003). "Expression of mouse Tbx22 supports its role in palatogenesis and glossogenesis." *Dev Dyn* **226**(4): 579-86.

Herrmann, B. G., S. Labeit, A. Poustka, T. R. King and H. Lehrach (1990). "Cloning of the T gene required in mesoderm formation in the mouse." *Nature* **343**(6259): 617-22.

Higashihori, N., M. Buchtova and J. M. Richman (2010). "The function and regulation of TBX22 in avian frontonasal morphogenesis." *Dev Dyn* **239**(2): 458-73.

Hilliard, S. A., L. Yu, S. Gu, Z. Zhang and Y. P. Chen (2005). "Regional regulation of palatal growth and patterning along the anterior-posterior axis in mice." *J Anat* **207**(5): 655-67.

Hochuli, E. (1988). "Large-scale chromatography of recombinant proteins." *J Chromatogr* **444**: 293-302.

Holland, L. Z., R. Albalat, K. Azumi, E. Benito-Gutierrez, M. J. Blow, M. Bronner-Fraser, F. Brunet, T. Butts, S. Candiani, L. J. Dishaw, D. E. Ferrier, J. Garcia-Fernandez, J. J. Gibson-Brown, C. Gissi, A. Godzik, F. Hallbook, D. Hirose, K. Hosomichi, T. Ikuta, H. Inoko, M. Kasahara, J. Kasamatsu, T. Kawashima, A. Kimura, M. Kobayashi, Z. Kozmik, K. Kubokawa, V. Laudet, G. W. Litman, A. C. McHardy, D. Meulemans, M. Nonaka, R. P. Olinski, Z. Pancer, L. A. Pennacchio, M. Pestarino, J. P. Rast, I. Rigoutsos, M. Robinson-Rechavi, G. Roch, H. Saiga, Y. Sasakura, M. Satake, Y. Satou, M. Schubert, N. Sherwood, T. Shiina, N. Takatori, J. Tello, P. Vopalensky, S. Wada, A. Xu, Y. Ye, K. Yoshida, F. Yoshizaki, J. K. Yu, Q. Zhang, C. M. Zmasek, P. J. de Jong, K. Osoegawa, N. H. Putnam, D. S. Rokhsar, N. Satoh and P. W. Holland (2008). "The amphioxus genome illuminates vertebrate origins and cephalochordate biology." *Genome Res* **18**(7): 1100-11.

Holland, L. Z., V. Laudet and M. Schubert (2004). "The chordate amphioxus: an emerging model organism for developmental biology." *Cell Mol Life Sci* **61**(18): 2290-308.

Holt, M. and S. Oram (1960). "Familial heart disease with skeletal malformations." *Br Heart J* **22**: 236-42.

Horak, C. E. and M. Snyder (2002). "ChIP-chip: a genomic approach for identifying transcription factor binding sites." *Methods Enzymol* **350**: 469-83.

Hou, W., Q. Tian, J. Zheng and H. L. Bonkovsky (2009). "MicroRNA-196 represses Bach1 protein and hepatitis C virus gene expression in human hepatoma cells expressing hepatitis C viral proteins." *Hepatology*.

Huang, R., Q. Zhi, J. C. Izpisua-Belmonte, B. Christ and K. Patel (1999). "Origin and development of the avian tongue muscles." *Anat Embryol (Berl)* **200**(2): 137-52.

Huang, Y. P., G. Wu, Z. Guo, M. Osada, T. Fomenkov, H. L. Park, B. Trink, D. Sidransky, A. Fomenkov and E. A. Ratovitski (2004). "Altered sumoylation of p63alpha contributes to the split-hand/foot malformation phenotype." *Cell Cycle* **3**(12): 1587-96.

Hunt, P., J. Whiting, S. Nonchev, M. H. Sham, H. Marshall, A. Graham, M. Cook, R. Allemann, P. W. Rigby, M. Gulisano and et al. (1991). "The branchial Hox code and its implications for gene regulation, patterning of the nervous system and head evolution." *Development Suppl* **2**: 63-77.

Hurlin, P. J., E. Steingrimsson, N. G. Copeland, N. A. Jenkins and R. N. Eisenman (1999). "Mga, a dual-specificity transcription factor that interacts with Max and contains a T-domain DNA-binding motif." *Embo J* **18**(24): 7019-28.

Hurst, R., M. Maffit, E. Murray, J. Kappelmann, Q. Xu, B. Butler, A. Ryan and G. Beckler (1996). "The TNT T7 quick coupled Transcription/Translation system." *Promega Notes*(58): 8-11.

Hyde, C. E. and R. W. Old (2000). "Regulation of the early expression of the *Xenopus* nodal-related 1 gene, Xnr1." *Development* **127**(6): 1221-9.

Ichikawa, E., A. Watanabe, Y. Nakano, S. Akita, A. Hirano, A. Kinoshita, S. Kondo, T. Kishino, T. Uchiyama, N. Niikawa and K. Yoshiura (2006). "PAX9 and TGFB3 are linked to susceptibility to nonsyndromic cleft lip with or without cleft palate in the Japanese: population-based and family-based candidate gene analyses." *J Hum Genet* **51**(1): 38-46.

Iimura, T. and O. Pourquie (2007). "Hox genes in time and space during vertebrate body formation." *Dev Growth Differ* **49**(4): 265-75.

Ingersoll, R. G., J. Hetmanski, J. W. Park, M. D. Fallin, I. McIntosh, Y. H. Wu-Chou, P. K. Chen, V. Yeow, S. S. Chong, F. Cheah, J. W. Sull, S. H. Jee, H. Wang, T. Wu, T. Murray, S. Huang, X. Ye, E. W. Jabs, R. Redett, G. Raymond, A. F. Scott and T. H. Beaty (2010). "Association between genes on chromosome 4p16 and non-syndromic oral clefts in four populations." *Eur J Hum Genet*.

Innan, H. and F. Kondrashov (2010). "The evolution of gene duplications: classifying and distinguishing between models." *Nat Rev Genet* **11**(2): 97-108.

Inoue, H., S. Kayano, Y. Aoki, S. Kure, A. Yamada, A. Hata, Y. Matsubara and Y. Suzuki (2008). "Association of the GABRB3 gene with nonsyndromic oral clefts." *Cleft Palate Craniofac J* **45**(3): 261-6.

Irawan, R., S. C. Tjin, P. Yager and D. Zhang (2005). "Cross-talk problem on a fluorescence multi-channel microfluidic chip system." *Biomed Microdevices* **7**(3): 205-11.

IUPAC-IUB (1971). "IUPAC-IUB commission on biochemical nomenclature (CBN). Abbreviations and symbols for nucleic acids, polynucleotides and their constituents." *J Mol Biol* **55**(3): 299-310.

Jenner, R. G., M. J. Townsend, I. Jackson, K. Sun, R. D. Bouwman, R. A. Young, L. H. Glimcher and G. M. Lord (2009). "The transcription factors T-bet and GATA-3 control alternative pathways of T-cell differentiation through a shared set of target genes." *Proc Natl Acad Sci U S A* **106**(42): 17876-81.

Jensen, L. E., S. Barbaux, K. Hoess, S. Fraterman, A. S. Whitehead and L. E. Mitchell (2004). "The human T locus and spina bifida risk." *Hum Genet* **115**(6): 475-82.

Jerome, L. A. and V. E. Papaioannou (2001). "DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1." *Nat Genet* **27**(3): 286-91.

Jezewski, P. A., P. K. Fang, T. L. Payne-Ferreira and P. C. Yelick (2009). "Alternative splicing, phylogenetic analysis, and craniofacial expression of zebrafish tbx22." *Dev Dyn* **238**(6): 1605-12.

Jheon, A. H. and R. A. Schneider (2009). "The cells that fill the bill: neural crest and the evolution of craniofacial development." *J Dent Res* **88**(1): 12-21.

Jin, J. Z. and J. Ding (2006). "Analysis of Meox-2 mutant mice reveals a novel postfusion-based cleft palate." *Dev Dyn* **235**(2): 539-46.

Jones, H. W., Jr., V. A. McKusick, P. S. Harper and K. D. Wuu (1971). "George Otto Gey. (1899-1970). The HeLa cell and a reappraisal of its origin." *Obstet Gynecol* **38**(6): 945-9.

Jowett, A. K., S. Vainio, M. W. Ferguson, P. T. Sharpe and I. Thesleff (1993). "Epithelial-mesenchymal interactions are required for msx 1 and msx 2 gene expression in the developing murine molar tooth." *Development* **117**(2): 461-70.

Jugessur, A., P. G. Farlie and N. Kilpatrick (2009). "The genetics of isolated orofacial clefts: from genotypes to subphenotypes." *Oral Dis* **15**(7): 437-53.

Juriloff, D. M., M. J. Harris, A. P. McMahon, T. J. Carroll and A. C. Lidral (2006). "Wnt9b is the mutated gene involved in multifactorial nonsyndromic cleft lip with or without cleft palate in A/WySn mice, as confirmed by a genetic complementation test." *Birth Defects Res A Clin Mol Teratol* **76**(8): 574-9.

Kanno, K., Y. Suzuki, A. Yamada, Y. Aoki, S. Kure and Y. Matsubara (2004). "Association between nonsyndromic cleft lip with or without cleft palate and

the glutamic acid decarboxylase 67 gene in the Japanese population." *Am J Med Genet A* **127A**(1): 11-6.

Karayazgan, B., Y. Gunay, B. Gurbuzer, M. Erkan and A. Atay (2009). "A preoperative appliance for a newborn with cleft palate." *Cleft Palate Craniofac J* **46**(1): 53-7.

Kawai, S., C. Faucheu, S. Gallea, S. Spinella-Jaegle, A. Atfi, R. Baron and S. R. Roman (2000). "Mouse smad8 phosphorylation downstream of BMP receptors ALK-2, ALK-3, and ALK-6 induces its association with Smad4 and transcriptional activity." *Biochem Biophys Res Commun* **271**(3): 682-7.

Kelley, R. I., E. H. Zackai, B. S. Emanuel, M. Kistenmacher, F. Greenberg and H. H. Punnett (1982). "The association of the DiGeorge anomalad with partial monosomy of chromosome 22." *J Pediatr* **101**(2): 197-200.

Kerrigan, J. J., J. P. Mansell, A. Sengupta, N. Brown and J. R. Sandy (2000). "Palatogenesis and potential mechanisms for clefting." *J R Coll Surg Edinb* **45**(6): 351-8.

Kim, S. H., K. W. Cho, H. S. Choi, S. J. Park, Y. Rhee, H. S. Jung and S. K. Lim (2009a). "The forkhead transcription factor Foxc2 stimulates osteoblast differentiation." *Biochem Biophys Res Commun* **386**(3): 532-6.

Kim, S. M., J. H. Lee, S. Jabaiti, S. K. Lee and J. Y. Choi (2009b). "Tbx22 expressions during palatal development in fetuses with glucocorticoid-alcohol-induced C57BL/6N cleft palates." *J Craniofac Surg* **20**(5): 1316-26.

Kispert, A. and B. G. Herrmann (1993). "The Brachyury gene encodes a novel DNA binding protein." *Embo J* **12**(8): 3211-20.

Kispert, A., B. Koschorz and B. G. Herrmann (1995). "The T protein encoded by Brachyury is a tissue-specific transcription factor." *Embo J* **14**(19): 4763-72.

Kitamura, H. (1989). *Embryology of the mouth and related structures*. Tokyo, Maruzen Co. Ltd.

Koch, W. E. and G. R. Smiley (1981). "In-vivo and in-vitro studies of the development of the avian secondary palate." *Arch Oral Biol* **26**(3): 181-7.

Kondo, S., B. C. Schutte, R. J. Richardson, B. C. Bjork, A. S. Knight, Y. Watanabe, E. Howard, R. L. de Lima, S. Daack-Hirsch, A. Sander, D. M. McDonald-McGinn, E. H. Zackai, E. J. Lammer, A. S. Aylsworth, H. H. Ardinger, A. C. Lidral, B. R. Pober, L. Moreno, M. Arcos-Burgos, C. Valencia, C. Houdayer, M. Bahauau, D. Moretti-Ferreira, A. Richieri-Costa, M. J. Dixon and J. C. Murray (2002). "Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes." *Nat Genet* **32**(2): 285-9.

Kontges, G. and A. Lumsden (1996). "Rhombencephalic neural crest segmentation is preserved throughout craniofacial ontogeny." *Development* **122**(10): 3229-42.

Kraus, F., B. Haenig and A. Kispert (2001a). "Cloning and expression analysis of the mouse T-box gene Tbx18." *Mech Dev* **100**(1): 83-6.

Kraus, F., B. Haenig and A. Kispert (2001b). "Cloning and expression analysis of the mouse T-box gene tbx20." *Mech Dev* **100**(1): 87-91.

Krost, B. and J. Schubert (2006). "Influence of season on prevalence of cleft lip and palate." *Int J Oral Maxillofac Surg* **35**(3): 215-8.

Kyttala, M., J. Tallila, R. Salonen, O. Kopra, N. Kohlschmidt, P. Paavola-Sakki, L. Peltonen and M. Kestila (2006). "MKS1, encoding a component of the flagellar apparatus basal body proteome, is mutated in Meckel syndrome." *Nat Genet* **38**(2): 155-7.

Lamolet, B., A. M. Pulichino, T. Lamonerie, Y. Gauthier, T. Brue, A. Enjalbert and J. Drouin (2001). "A pituitary cell-restricted T box factor, Tpit, activates POMC transcription in cooperation with Pitx homeoproteins." *Cell* **104**(6): 849-59.

Larroux, C., G. N. Luke, P. Koopman, D. S. Rokhsar, S. M. Shimeld and B. M. Degnan (2008). "Genesis and expansion of metazoan transcription factor gene classes." *Mol Biol Evol* **25**(5): 980-96.

Larsen, W. J. (1997). *Essentials of Human Embryology*. UK, Elsevier Health Sciences.

Laugier-Anfossi, F. and L. Villard (2000). "Molecular characterization of a new human T-box gene (TBX22) located in xq21.1 encoding a protein containing a truncated T-domain." *Gene* **255**(2): 289-96.

Lausch, E., P. Hermanns, H. F. Farin, Y. Alanay, S. Unger, S. Nikkel, C. Steinwender, G. Scherer, J. Spranger, B. Zabel, A. Kispert and A. Superti-Furga (2008). "TBX15 mutations cause craniofacial dysmorphism, hypoplasia of scapula and pelvis, and short stature in Cousin syndrome." *Am J Hum Genet* **83**(5): 649-55.

Law, D. J., N. Garvey, S. I. Agulnik, V. Perlroth, O. M. Hahn, R. E. Rhinehart, T. C. Gebuhr and L. M. Silver (1998). "TBX10, a member of the Tbx1-subfamily of conserved developmental genes, is located at human chromosome 11q13 and proximal mouse chromosome 19." *Mamm Genome* **9**(5): 397-9.

Leoyklang, P., K. Suphapeetiporn, P. Siriwan, T. Desudchit, P. Chaowanapanja, W. A. Gahl and V. Shotelersuk (2007). "Heterozygous nonsense mutation SATB2 associated with cleft palate, osteoporosis, and cognitive defects." *Hum Mutat* **28**(7): 732-8.

Li, Q. and J. Ding (2007). "Gene expression analysis reveals that formation of the mouse anterior secondary palate involves recruitment of cells from the posterior side." *Int J Dev Biol* **51**(2): 167-72.

Li, Q. Y., R. A. Newbury-Ecob, J. A. Terrett, D. I. Wilson, A. R. Curtis, C. H. Yi, T. Gebuhr, P. J. Bullen, S. C. Robson, T. Strachan, D. Bonnet, S. Lyonnet, I. D. Young, J. A. Raeburn, A. J. Buckler, D. J. Law and J. D. Brook (1997). "Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family." *Nat Genet* **15**(1): 21-9.

Lidral, A. C., P. A. Romitti, A. M. Basart, T. Doetschman, N. J. Leysens, S. Daack-Hirsch, E. V. Semina, L. R. Johnson, J. Machida, A. Burds, T. J. Parnell, J. L. Rubenstein and J. C. Murray (1998). "Association of MSX1 and TGFB3 with nonsyndromic clefting in humans." *Am J Hum Genet* **63**(2): 557-68.

Lin, X., M. Liang, Y. Y. Liang, F. C. Brunicardi, F. Melchior and X. H. Feng (2003). "Activation of transforming growth factor-beta signaling by SUMO-1 modification of tumor suppressor Smad4/DPC4." *J Biol Chem* **278**(21): 18714-9.

Lin, Y. C., L. J. Lo, M. S. Noordhoff and Y. R. Chen (1999). "Cleft of the lip and palate in twins." *Changgeng Yi Xue Za Zhi* **22**(1): 61-7.

Lindsay, S. and A. J. Copp (2005). "MRC-Wellcome Trust Human Developmental Biology Resource: enabling studies of human developmental gene expression." *Trends Genet* **21**(11): 586-90.

Lingbeek, M. E., J. J. Jacobs and M. van Lohuizen (2002). "The T-box repressors TBX2 and TBX3 specifically regulate the tumor suppressor gene p14ARF via a variant T-site in the initiator." *J Biol Chem* **277**(29): 26120-7.

Liu, W., Y. Lan, E. Pauws, M. A. Meester-Smoor, P. Stanier, E. C. Zwarthoff and R. Jiang (2008). "The Mn1 transcription factor acts upstream of Tbx22 and

preferentially regulates posterior palate growth in mice." *Development* **135**(23): 3959-68.

Lowry, R. B. (1970). "Sex-linked cleft palate in a British Columbia Indian family." *Pediatrics* **46**(1): 123-8.

Luke, D. A. (1976). "Development of the secondary palate in man." *Acta Anat (Basel)* **94**(4): 596-608.

Lumsden, A., N. Sprawson and A. Graham (1991). "Segmental origin and migration of neural crest cells in the hindbrain region of the chick embryo." *Development* **113**(4): 1281-91.

Ma, Y., L. Xu, D. Rodriguez-Agudo, X. Li, D. M. Heuman, P. B. Hylemon, W. M. Pandak and S. Ren (2008). "25-Hydroxycholesterol-3-sulfate regulates macrophage lipid metabolism via the LXR/SREBP-1 signaling pathway." *Am J Physiol Endocrinol Metab* **295**(6): E1369-79.

Maarse, W., S. J. Berge, L. Pistorius, T. van Barneveld, M. Kon, C. Breugem and A. B. Mink van der Molen (2010). "Diagnostic accuracy of transabdominal ultrasound in detecting prenatal cleft lip and palate: a systematic review." *Ultrasound Obstet Gynecol*.

Mackenzie, A., G. L. Leeming, A. K. Jowett, M. W. Ferguson and P. T. Sharpe (1991). "The homeobox gene Hox 7.1 has specific regional and temporal expression patterns during early murine craniofacial embryogenesis, especially tooth development in vivo and in vitro." *Development* **111**(2): 269-85.

Mangus, D. A., M. C. Evans and A. Jacobson (2003). Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol.* **4**: 223.

Mani, M., M. Carlsson and A. Marcusson (2010). "Quality of Life Varies with Gender and Age among Adults Treated for Unilateral Cleft Lip and Palate." *Cleft Palate Craniofac J.*

Mansilla, M. A., M. E. Cooper, T. Goldstein, E. E. Castilla, J. S. Lopez Camelo, M. L. Marazita and J. C. Murray (2006). "Contributions of PTCH gene variants to isolated cleft lip and palate." *Cleft Palate Craniofac J* **43**(1): 21-9.

Marcano, A. C., K. Doudney, C. Braybrook, R. Squires, M. A. Patton, M. M. Lees, A. Richieri-Costa, A. C. Lidral, J. C. Murray, G. E. Moore and P. Stanier (2004). "TBX22 mutations are a frequent cause of cleft palate." *J Med Genet* **41**(1): 68-74.

Martinelli, M., M. Di Stazio, L. Scapoli, J. Marchesini, F. Di Bari, F. Pezzetti, F. Carinci, A. Palmieri, P. Carinci and A. Savoia (2007). "Cleft lip with or without cleft palate: implication of the heavy chain of non-muscle myosin IIA." *J Med Genet* **44**(6): 387-92.

Martinet, W., D. M. Schrijvers and M. M. Kockx (2003). "Nucleofection as an efficient nonviral transfection method for human monocytic cells." *Biotechnol Lett* **25**(13): 1025-9.

McGrath, J. A., P. H. Duijf, V. Doetsch, A. D. Irvine, R. de Waal, K. R. Vanmolkot, V. Wessagowit, A. Kelly, D. J. Atherton, W. A. Griffiths, S. J. Orlow, A. van Haeringen, M. G. Ausems, A. Yang, F. McKeon, M. A. Bamshad, H. G. Brunner, B. C. Hamel and H. van Bokhoven (2001). "Hay-Wells syndrome is caused by heterozygous missense mutations in the SAM domain of p63." *Hum Mol Genet* **10**(3): 221-9.

McLachlan, J. (1994). *Medical Emryology*. Wokingham, Addison-Wesley.

Melkonniemi, M., H. Koillinen, M. Mannikko, M. L. Warman, T. Pihlajamaa, H. Kaariainen, J. Rautio, J. Hukki, J. A. Stofko, G. J. Cisneros, D. Krakow, D. H.

Cohn, J. Kere and L. Ala-Kokko (2003). "Collagen XI sequence variations in nonsyndromic cleft palate, Robin sequence and micrognathia." *Eur J Hum Genet* **11**(3): 265-70.

Melkonyan, H., C. Sorg and M. Klemp (1996). "Electroporation efficiency in mammalian cells is increased by dimethyl sulfoxide (DMSO)." *Nucleic Acids Res* **24**(21): 4356-7.

Mendoza, R. L. (2009). "Public health policy and medical missions in the Philippines: the case of oral--facial clefting." *Asia Pac J Public Health* **21**(1): 94-103.

Meng, L., Z. Bian, R. Torensma and J. W. Von den Hoff (2009). "Biological mechanisms in palatogenesis and cleft palate." *J Dent Res* **88**(1): 22-33.

Meulmeester, E. and F. Melchior (2008). "Cell biology: SUMO." *Nature* **452**(7188): 709-11.

Milunsky, J. M., T. A. Maher, G. Zhao, A. E. Roberts, H. J. Stalker, R. T. Zori, M. N. Burch, M. Clemens, J. B. Mulliken, R. Smith and A. E. Lin (2008). "TFAP2A mutations result in branchio-oculo-facial syndrome." *Am J Hum Genet* **82**(5): 1171-7.

Minguillon, C. and M. Logan (2003). "The comparative genomics of T-box genes." *Brief Funct Genomic Proteomic* **2**(3): 224-33.

Miyoshi, T., M. Maruhashi, T. Van De Putte, H. Kondoh, D. Huylebroeck and Y. Higashi (2006). "Complementary expression pattern of Zfhx1 genes Sip1 and deltaEF1 in the mouse embryo and their genetic interaction revealed by compound mutants." *Dev Dyn* **235**(7): 1941-52.

Mo, R., A. M. Freer, D. L. Zinyk, M. A. Crackower, J. Michaud, H. H. Heng, K. W. Chik, X. M. Shi, L. C. Tsui, S. H. Cheng, A. L. Joyner and C. Hui (1997). "Specific and redundant functions of Gli2 and Gli3 zinc finger genes in skeletal patterning and development." *Development* **124**(1): 113-23.

Mogass, M., P. Bringas, Jr. and C. F. Shuler (2000). "Characterization of desmosomal component expression during palatogenesis." *Int J Dev Biol* **44**(3): 317-22.

Moore, G. E., A. Ivens, J. Chambers, M. Farrall, R. Williamson, D. C. Page, A. Bjornsson, A. Arnason and O. Jansson (1987). "Linkage of an X-chromosome cleft palate gene." *Nature* **326**(6108): 91-2.

Moore, G. E., R. Williamson, O. Jansson, J. Chambers, F. Takakubo, R. Newton, M. A. Balacs and A. Ivens (1991). "Localization of a mutant gene for cleft palate and ankyloglossia in an X-linked Icelandic family." *J Craniofac Genet Dev Biol* **11**(4): 372-6.

Moore, K. L. and T. V. N. Persaud (2003). *The Developing Human*. London, W.B. Saunders.

Moorman, A. F., P. A. De Boer, J. L. Vermeulen and W. H. Lamers (1993). "Practical aspects of radio-isotopic in situ hybridization on RNA." *Histochem J* **25**(4): 251-66.

Moreno, L. M., M. A. Mansilla, S. A. Bullard, M. E. Cooper, T. D. Busch, J. Machida, M. K. Johnson, D. Brauer, K. Krahn, S. Daack-Hirsch, J. L'Heureux, C. Valencia-Ramirez, D. Rivera, A. M. Lopez, M. A. Moreno, A. Hing, E. J. Lammer, M. Jones, K. Christensen, R. T. Lie, A. Jugessur, A. J. Wilcox, P. Chines, E. Pugh, K. Doheny, M. Arcos-Burgos, M. L. Marazita, J. C. Murray and A. C. Lidral (2009). "FOXE1 association with both isolated cleft lip with or without cleft palate, and isolated cleft palate." *Hum Mol Genet* **18**(24): 4879-96.

Morrison, K., C. Papapetrou, J. Attwood, F. Hol, S. A. Lynch, A. Sampath, B. Hamel, J. Burn, J. Sowden, D. Stott, E. Mariman and Y. H. Edwards (1996). "Genetic

mapping of the human homologue (T) of mouse T(Brachyury) and a search for allele association between human T and spina bifida." *Hum Mol Genet* **5**(5): 669-74.

Mowat, D. R., M. J. Wilson and M. Goossens (2003). "Mowat-Wilson syndrome." *J Med Genet* **40**(5): 305-10.

Muller, C. W. and B. G. Herrmann (1997). "Crystallographic structure of the T domain-DNA complex of the Brachyury transcription factor." *Nature* **389**(6653): 884-8.

Muller, F. and R. O'Rahilly (1997). "The timing and sequence of appearance of neuromeres and their derivatives in staged human embryos." *Acta Anat (Basel)* **158**(2): 83-99.

Mullis, K., F. Falloona, S. Scharf, R. Saiki, G. Horn and H. Erlich (1986). "Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction." *Cold Spring Harb Symp Quant Biol* **51 Pt 1**: 263-73.

Murakami, M., M. Nakagawa, E. N. Olson and O. Nakagawa (2005). "A WW domain protein TAZ is a critical coactivator for TBX5, a transcription factor implicated in Holt-Oram syndrome." *Proc Natl Acad Sci U S A* **102**(50): 18034-9.

Murray, J. C. and B. C. Schutte (2004). "Cleft palate: players, pathways, and pursuits." *J Clin Invest* **113**(12): 1676-8.

Naiche, L. A. and V. E. Papaioannou (2003). "Loss of Tbx4 blocks hindlimb development and affects vascularization and fusion of the allantois." *Development* **130**(12): 2681-93.

Nakatomi, M., X. P. Wang, D. Key, J. J. Lund, A. Turbe-Doan, R. Kist, A. Aw, Y. Chen, R. L. Maas and H. Peters (2010). "Genetic interactions between Pax9 and Msx1 regulate lip development and several stages of tooth morphogenesis." *Dev Biol*.

Nawshad, A., D. LaGamba and E. D. Hay (2004). "Transforming growth factor beta (TGFbeta) signalling in palatal growth, apoptosis and epithelial mesenchymal transformation (EMT)." *Arch Oral Biol* **49**(9): 675-89.

Nieto, M. A., L. C. Bradley, P. Hunt, R. Das Gupta, R. Krumlauf and D. G. Wilkinson (1992). "Molecular mechanisms of pattern formation in the vertebrate hindbrain." *Ciba Found Symp* **165**: 92-102; discussion 102-7.

Noden, D. M. (1983). "The role of the neural crest in patterning of avian cranial skeletal, connective, and muscle tissues." *Dev Biol* **96**(1): 144-65.

Nugent, P. and R. M. Greene (1998). "MSX-1 gene expression and regulation in embryonic palatal tissue." *In Vitro Cell Dev Biol Anim* **34**(10): 831-5.

Oganesyan, N., S. H. Kim and R. Kim (2005). "On-column protein refolding for crystallization." *J Struct Funct Genomics* **6**(2-3): 177-82.

Oliver, R. G. and G. Jones (1997). "Neonatal feeding of infants born with cleft lip and/or palate: parental perceptions of their experience in south Wales." *Cleft Palate Craniofac J* **34**(6): 526-32.

O'Rahilly, R. and F. Muller (1987). *Developmental stages in human embryos*, Carnegie Institute of Washington.

Osoegawa, K., G. M. Vessere, K. H. Utami, M. A. Mansilla, M. K. Johnson, B. M. Riley, J. L'Heureux, R. Pfundt, J. Staaf, W. A. van der Vliet, A. C. Lidral, E. F. Schoenmakers, A. Borg, B. C. Schutte, E. J. Lammer, J. C. Murray and P. J. de Jong (2008). "Identification of novel candidate genes associated with cleft lip and palate using array comparative genomic hybridisation." *J Med Genet* **45**(2): 81-6.

Pantalacci, S., J. Prochazka, A. Martin, M. Rothova, A. Lambert, L. Bernard, C. Charles, L. Viriot, R. Peterkova and V. Laudet (2008). "Patterning of palatal rugae through sequential addition reveals an anterior/posterior boundary in palatal development." *BMC Dev Biol* **8**: 116.

Papaioannou, V. E. (2001). "T-box genes in development: from hydra to humans." *Int Rev Cytol* **207**: 1-70.

Papapetrou, C., Y. H. Edwards and J. C. Sowden (1997). "The T transcription factor functions as a dimer and exhibits a common human polymorphism Gly-177-Asp in the conserved DNA-binding domain." *FEBS Lett* **409**(2): 201-6.

Pappin, D. J., P. Hojrup and A. J. Bleasby (1993). "Rapid identification of proteins by peptide-mass fingerprinting." *Curr Biol* **3**(6): 327-32.

Pardue, M. L. and J. G. Gall (1969). "Molecular hybridization of radioactive DNA to the DNA of cytological preparations." *Proc Natl Acad Sci U S A* **64**(2): 600-4.

Patil, M. S., S. B. Patil and A. B. Acharya (2008). "Palatine rugae and their significance in clinical dentistry: a review of the literature." *J Am Dent Assoc* **139**(11): 1471-8.

Pauws, E., A. Hoshino, L. Bentley, S. Prajapati, C. Keller, P. Hammond, J. P. Martinez-Barbera, G. E. Moore and P. Stanier (2009a). "Tbx22null mice have a submucous cleft palate due to reduced palatal bone formation and also display ankyloglossia and choanal atresia phenotypes." *Hum Mol Genet* **18**(21): 4171-9.

Pauws, E., G. E. Moore and P. Stanier (2009b). "A functional haplotype variant in the TBX22 promoter is associated with cleft palate and ankyloglossia." *J Med Genet* **46**(8): 555-61.

Pauws, E. and P. Stanier (2007). "FGF signalling and SUMO modification: new players in the aetiology of cleft lip and/or palate." *Trends Genet* **23**(12): 631-40.

Paxton, C., H. Zhao, Y. Chin, K. Langner and J. Reecy (2002). "Murine Tbx2 contains domains that activate and repress gene transcription." *Gene* **283**(1-2): 117-24.

Paylor, R., B. Glaser, A. Mupo, P. Ataliotis, C. Spencer, A. Sobotka, C. Sparks, C. H. Choi, J. Oghalai, S. Curran, K. C. Murphy, S. Monks, N. Williams, M. C. O'Donovan, M. J. Owen, P. J. Scambler and E. Lindsay (2006). "Tbx1 haploinsufficiency is linked to behavioral disorders in mice and humans: implications for 22q11 deletion syndrome." *Proc Natl Acad Sci U S A* **103**(20): 7729-34.

Peterkova, R., I. Klepacek and M. Peterka (1987). "Prenatal development of rugae palatinæ in mice: scanning electron microscopic and histologic studies." *J Craniofac Genet Dev Biol* **7**(2): 169-89.

Peters, H., A. Neubuser, K. Kratochwil and R. Balling (1998). "Pax9-deficient mice lack pharyngeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities." *Genes Dev* **12**(17): 2735-47.

Pflugfelder, G. O., H. Roth and B. Poeck (1992). "A homology domain shared between Drosophila optomotor-blind and mouse Brachyury is involved in DNA binding." *Biochem Biophys Res Commun* **186**(2): 918-25.

Poletta, F. A., E. E. Castilla, I. M. Orioli and J. S. Lopez-Camelo (2007). "Regional analysis on the occurrence of oral clefts in South America." *Am J Med Genet A* **143**(24): 3216-27.

Pollock, R. and R. Treisman (1990). "A sensitive method for the determination of protein-DNA binding specificities." *Nucleic Acids Res* **18**(21): 6197-204.

Pratt, R. M., J. F. Goggins, A. L. Wilk and C. T. King (1973). "Acid mucopolysaccharide synthesis in the secondary palate of the developing rat at the time of rotation and fusion." *Dev Biol* **32**(1): 230-7.

Prince, S., S. Carreira, K. W. Vance, A. Abrahams and C. R. Goding (2004). "Tbx2 directly represses the expression of the p21(WAF1) cyclin-dependent kinase inhibitor." *Cancer Res* **64**(5): 1669-74.

Pruitt, K. D., T. Tatusova and D. R. Maglott (2007). "NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins." *Nucleic Acids Res* **35**(Database issue): D61-5.

Puelles, L. and J. L. Rubenstein (2003). "Forebrain gene expression domains and the evolving prosomeric model." *Trends Neurosci* **26**(9): 469-76.

Quaderi, N. A., S. Schweiger, K. Gaudenz, B. Franco, E. I. Rugarli, W. Berger, G. J. Feldman, M. Volta, G. Andolfi, S. Gilgenkrantz, R. W. Marion, R. C. Hennekam, J. M. Opitz, M. Muenke, H. H. Ropers and A. Ballabio (1997). "Opitz G/BBB syndrome, a defect of midline development, is due to mutations in a new RING finger gene on Xp22." *Nat Genet* **17**(3): 285-91.

Rice, D. P. (2005). "Craniofacial anomalies: from development to molecular pathogenesis." *Curr Mol Med* **5**(7): 699-722.

Rice, P., I. Longden and A. Bleasby (2000). "EMBOSS: the European Molecular Biology Open Software Suite." *Trends Genet* **16**(6): 276-7.

Rijli, F. M., M. Mark, S. Lakkaraju, A. Dierich, P. Dolle and P. Chambon (1993). "A homeotic transformation is generated in the rostral branchial region of the head by disruption of Hoxa-2, which acts as a selector gene." *Cell* **75**(7): 1333-49.

Riley, B. M., M. A. Mansilla, J. Ma, S. Daack-Hirsch, B. S. Maher, L. M. Raffensperger, E. T. Russo, A. R. Vieira, C. Dode, M. Mohammadi, M. L. Marazita and J. C. Murray (2007a). "Impaired FGF signaling contributes to cleft lip and palate." *Proc Natl Acad Sci U S A* **104**(11): 4512-7.

Riley, B. M., R. E. Schultz, M. E. Cooper, T. Goldstein-McHenry, S. Daack-Hirsch, K. T. Lee, E. Dragan, A. R. Vieira, A. C. Lidral, M. L. Marazita and J. C. Murray (2007b). "A genome-wide linkage scan for cleft lip and cleft palate identifies a novel locus on 8p11-23." *Am J Med Genet A* **143A**(8): 846-52.

Robert, B., D. Sasseon, B. Jacq, W. Gehring and M. Buckingham (1989). "Hox-7, a mouse homeobox gene with a novel pattern of expression during embryogenesis." *Embo J* **8**(1): 91-100.

Roberts, J. A. (1961). "Multifactorial inheritance in relation to normal and abnormal human traits." *Br Med Bull* **17**: 241-6.

Roberts, J. A. (1964). "Multifactorial Inheritance and Human Disease." *Prog Med Genet* **23**: 178-216.

Rodriguez, M. S., J. M. Desterro, S. Lain, C. A. Midgley, D. P. Lane and R. T. Hay (1999). "SUMO-1 modification activates the transcriptional response of p53." *Embo J* **18**(22): 6455-61.

Rubenstein, J. L., S. Martinez, K. Shimamura and L. Puelles (1994). "The embryonic vertebrate forebrain: the prosomeric model." *Science* **266**(5185): 578-80.

Rullo, R., D. Di Maggio, V. M. Festa and N. Mazzarella (2009). "Speech assessment in cleft palate patients: a descriptive study." *Int J Pediatr Otorhinolaryngol* **73**(5): 641-4.

Ruvinsky, I., L. M. Silver and J. J. Gibson-Brown (2000). "Phylogenetic analysis of T-Box genes demonstrates the importance of amphioxus for understanding evolution of the vertebrate genome." *Genetics* **156**(3): 1249-57.

Saka, Y., M. Tada and J. C. Smith (2000). "A screen for targets of the Xenopus T-box gene Xbra." *Mech Dev* **93**(1-2): 27-39.

Satokata, I. and R. Maas (1994). "Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development." *Nat Genet* **6**(4): 348-56.

Scapoli, L., M. Martinelli, F. Pezzetti, F. Carinci, M. Bodo, M. Tognon and P. Carinci (2002). "Linkage disequilibrium between GABRB3 gene and nonsyndromic familial cleft lip with or without cleft palate." *Hum Genet* **110**(1): 15-20.

Scapoli, L., A. Palmieri, M. Martinelli, F. Pezzetti, P. Carinci, M. Tognon and F. Carinci (2005). "Strong evidence of linkage disequilibrium between polymorphisms at the IRF6 locus and nonsyndromic cleft lip with or without cleft palate, in an Italian population." *Am J Hum Genet* **76**(1): 180-3.

Schilling, T. F., V. Prince and P. W. Ingham (2001). "Plasticity in zebrafish hox expression in the hindbrain and cranial neural crest." *Dev Biol* **231**(1): 201-16.

Schinzel, A., R. Illig and A. Prader (1987). "The ulnar-mammary syndrome: an autosomal dominant pleiotropic gene." *Clin Genet* **32**(3): 160-8.

Sharma, R. K. and V. Nanda (2009). "Problems of middle ear and hearing in cleft children." *Indian J Plast Surg* **42 Suppl**: S144-8.

Sharpe, J., U. Ahlgren, P. Perry, B. Hill, A. Ross, J. Hecksher-Sorensen, R. Baldeck and D. Davidson (2002). "Optical projection tomography as a tool for 3D microscopy and gene expression studies." *Science* **296**(5567): 541-5.

Shaw, G. M. and E. J. Lammer (1999). "Maternal periconceptional alcohol consumption and risk for orofacial clefts." *J Pediatr* **134**(3): 298-303.

Shaw, G. M., C. R. Wasserman, E. J. Lammer, C. D. O'Malley, J. C. Murray, A. M. Basart and M. M. Tolarova (1996). "Orofacial clefts, parental cigarette smoking, and transforming growth factor-alpha gene variants." *Am J Hum Genet* **58**(3): 551-61.

Shen, R., Y. Chen, L. Huang, E. Vitale and M. Solursh (1994). "Characterization of the human MSX-1 promoter and an enhancer responsible for retinoic acid induction." *Cell Mol Biol Res* **40**(4): 297-312.

Shortle, D. (1996). "The denatured state (the other half of the folding equation) and its role in protein stability." *Faseb J* **10**(1): 27-34.

Showell, C., O. Binder and F. L. Conlon (2004). "T-box genes in early embryogenesis." *Dev Dyn* **229**(1): 201-18.

Shuler, C. F. (1995). "Programmed cell death and cell transformation in craniofacial development." *Crit Rev Oral Biol Med* **6**(3): 202-17.

Silva, R. G., L. P. Carvalho, J. S. Oliveira, C. A. Pinto, M. A. Mendes, M. S. Palma, L. A. Basso and D. S. Santos (2003). "Cloning, overexpression, and purification of functional human purine nucleoside phosphorylase." *Protein Expr Purif* **27**(1): 158-64.

Simon, H. (1999). "T-box genes and the formation of vertebrate forelimb- and hindlimb specific pattern." *Cell Tissue Res* **296**(1): 57-66.

Singh, M. K., V. M. Christoffels, J. M. Dias, M. O. Trowe, M. Petry, K. Schuster-Gossler, A. Burger, J. Ericson and A. Kispert (2005a). "Tbx20 is essential for cardiac chamber differentiation and repression of Tbx2." *Development* **132**(12): 2697-707.

Singh, M. K., M. Petry, B. Haenig, B. Lescher, M. Leitges and A. Kispert (2005b). "The T-box transcription factor Tbx15 is required for skeletal development." *Mech Dev* **122**(2): 131-44.

Sinha, S., S. Abraham, R. M. Gronostajski and C. E. Campbell (2000). "Differential DNA binding and transcription modulation by three T-box proteins, T, TBX1 and TBX2." *Gene* **258**(1-2): 15-29.

Sivertsen, A., A. J. Wilcox, R. Skjaerven, H. A. Vindenes, F. Abyholm, E. Harville and R. T. Lie (2008). "Familial risk of oral clefts by morphological type and severity: population based cohort study of first degree relatives." *Bmj* **336**(7641): 432-4.

Song, Y., J. N. Hui, K. K. Fu and J. M. Richman (2004). "Control of retinoic acid synthesis and FGF expression in the nasal pit is required to pattern the craniofacial skeleton." *Dev Biol* **276**(2): 313-29.

Stanier, P., S. A. Forbes, A. Arnason, A. Bjornsson, E. Sveinbjornsdottir, R. Williamson and G. Moore (1993). "The localization of a gene causing X-linked cleft palate and ankyloglossia (CPX) in an Icelandic kindred is between DDX326 and DDX51X." *Genomics* **17**(3): 549-55.

Stanier, P. and G. E. Moore (2004). "Genetics of cleft lip and palate: syndromic genes contribute to the incidence of non-syndromic clefts." *Hum Mol Genet* **13 Spec No 1**: R73-81.

Stedman, T. L. (2005). *Stedman's Medical Dictionary*. Philadelphia, Lippincott Williams and Wilkins.

Stoll, C., S. Mengsteab, D. Stoll, D. Riediger, A. M. Gressner and R. Weiskirchen (2004). "Analysis of polymorphic TGFB1 codons 10, 25, and 263 in a German patient group with non-syndromic cleft lip, alveolus, and palate compared with healthy adults." *BMC Med Genet* **5**: 15.

Stoller, J. Z. and J. A. Epstein (2005). "Identification of a novel nuclear localization signal in Tbx1 that is deleted in DiGeorge syndrome patients harboring the 1223delC mutation." *Hum Mol Genet* **14**(7): 885-92.

Strachan, T., M. Abitbol, D. Davidson and J. S. Beckmann (1997). "A new dimension for the human genome project: towards comprehensive expression maps." *Nat Genet* **16**(2): 126-32.

Su, H. L. and S. S. Li (2002). "Molecular features of human ubiquitin-like SUMO genes and their encoded proteins." *Gene* **296**(1-2): 65-73.

Subramanian, A., P. Ranganathan and S. L. Diamond (1999). "Nuclear targeting peptide scaffolds for lipofection of nondividing mammalian cells." *Nat Biotechnol* **17**(9): 873-7.

Sull, J. W., K. Y. Liang, J. B. Hetmanski, M. D. Fallin, R. G. Ingersoll, J. Park, Y. H. Wu-Chou, P. K. Chen, S. S. Chong, F. Cheah, V. Yeow, B. Y. Park, S. H. Jee, E. W. Jabs, R. Redett, E. Jung, I. Ruczinski, A. F. Scott and T. H. Beaty (2008a). "Differential parental transmission of markers in RUNX2 among cleft case-parent trios from four populations." *Genet Epidemiol* **32**(6): 505-12.

Sull, J. W., K. Y. Liang, J. B. Hetmanski, M. D. Fallin, R. G. Ingersoll, J. W. Park, Y. H. Wu-Chou, P. K. Chen, S. S. Chong, F. Cheah, V. Yeow, B. Y. Park, S. H. Jee, E. W. Jabs, R. Redett, A. F. Scott and T. H. Beaty (2008b). "Excess maternal transmission of markers in TCOF1 among cleft palate case-parent trios from three populations." *Am J Med Genet A* **146A**(18): 2327-31.

Suphapeetiporn, K., S. Tongkobpatch, P. Siriwan and V. Shotelersuk (2007). "TBX22 mutations are a frequent cause of non-syndromic cleft palate in the Thai population." *Clin Genet* **72**(5): 478-83.

Suzuki, K., D. Hu, T. Bustos, J. Zlotogora, A. Richieri-Costa, J. A. Helms and R. A. Spritz (2000). "Mutations of PVRL1, encoding a cell-cell adhesion

molecule/herpesvirus receptor, in cleft lip/palate-ectodermal dysplasia." *Nat Genet* **25**(4): 427-30.

Szabo, S. J., B. M. Sullivan, C. Stemmann, A. R. Satoskar, B. P. Sleckman and L. H. Glimcher (2002). "Distinct effects of T-bet in TH1 lineage commitment and IFN-gamma production in CD4 and CD8 T cells." *Science* **295**(5553): 338-42.

Tada, M., E. S. Casey, L. Fairclough and J. C. Smith (1998). "Bix1, a direct target of Xenopus T-box genes, causes formation of ventral mesoderm and endoderm." *Development* **125**(20): 3997-4006.

Tada, M. and J. C. Smith (2001). "T-targets: clues to understanding the functions of T-box proteins." *Dev Growth Differ* **43**(1): 1-11.

Talmant, J. C. (2006). "Evolution of the functional repair concept for cleft lip and palate patients." *Indian J Plast Surg* **39**(2): 196-209.

Tanpaiboon, P., P. Kantaputra, K. Wejathikul and W. Piaymongkol (2010). "c. 595-596 insC of FOXC2 underlies lymphedema, distichiasis, ptosis, ankyloglossia, and Robin sequence in a Thai patient." *Am J Med Genet A* **152A**(3): 737-40.

Taylor, K. M. and C. Labonne (2005). "SoxE factors function equivalently during neural crest and inner ear development and their activity is regulated by SUMOylation." *Dev Cell* **9**(5): 593-603.

Thomsen, D. R., R. M. Stenberg, W. F. Goins and M. F. Stinski (1984). Promoter-regulatory region of the major immediate early gene of human cytomegalovirus. *Proc Natl Acad Sci U S A* **81**: 659-63.

Torii, C., K. Izumi, H. Nakajima, T. Takahashi and K. Kosaki (2007). "EFNB1 mutation at the ephrin ligand-receptor dimerization interface in a patient with craniofrontonasal syndrome." *Congenit Anom (Kyoto)* **47**(1): 49-52.

Townsend, M. J., A. S. Weinmann, J. L. Matsuda, R. Salomon, P. J. Farnham, C. A. Biron, L. Gapin and L. H. Glimcher (2004). "T-bet regulates the terminal maturation and homeostasis of NK and Valpha14i NKT cells." *Immunity* **20**(4): 477-94.

Trainor, P. and R. Krumlauf (2000). "Plasticity in mouse neural crest cells reveals a new patterning role for cranial mesoderm." *Nat Cell Biol* **2**(2): 96-102.

Trainor, P. A. and R. Krumlauf (2001). "Hox genes, neural crest cells and branchial arch patterning." *Curr Opin Cell Biol* **13**(6): 698-705.

Turley, E. A., M. D. Hollenberg and R. M. Pratt (1985). "Effect of epidermal growth factor/urogastrone on glycosaminoglycan synthesis and accumulation in vitro in the developing mouse palate." *Differentiation* **28**(3): 279-85.

Twigg, S. R., R. Kan, C. Babbs, E. G. Bochukova, S. P. Robertson, S. A. Wall, G. M. Morriss-Kay and A. O. Wilkie (2004). "Mutations of ephrin-B1 (EFNB1), a marker of tissue boundary formation, cause craniofrontonasal syndrome." *Proc Natl Acad Sci U S A* **101**(23): 8652-7.

van den Boogaard, M. J., M. Dorland, F. A. Beemer and H. K. van Amstel (2000). "MSX1 mutation is associated with orofacial clefting and tooth agenesis in humans." *Nat Genet* **24**(4): 342-3.

Vandenberg, P., J. S. Khillan, D. J. Prockop, H. Helminen, S. Kontusaari and L. Ala-Kokko (1991). "Expression of a partially deleted gene of human type II procollagen (COL2A1) in transgenic mice produces a chondrodysplasia." *Proc Natl Acad Sci U S A* **88**(17): 7640-4.

Vega, H., Q. Waisfisz, M. Gordillo, N. Sakai, I. Yanagihara, M. Yamada, D. van Gosliga, H. Kayserili, C. Xu, K. Ozono, E. W. Jabs, K. Inui and H. Joenje (2005). "Roberts syndrome is caused by mutations in ESCO2, a human

homolog of yeast ECO1 that is essential for the establishment of sister chromatid cohesion." *Nat Genet* **37**(5): 468-70.

Veitch, E., J. Begbie, T. F. Schilling, M. M. Smith and A. Graham (1999). "Pharyngeal arch patterning in the absence of neural crest." *Curr Biol* **9**(24): 1481-4.

Ventura, S. and A. Villaverde (2006). "Protein quality in bacterial inclusion bodies." *Trends Biotechnol* **24**(4): 179-85.

Vieira, A. R. (2006). "Association between the transforming growth factor alpha gene and nonsyndromic oral clefts: a HuGE review." *Am J Epidemiol* **163**(9): 790-810.

Vieira, A. R., J. R. Avila, S. Daack-Hirsch, E. Dragan, T. M. Felix, F. Rahimov, J. Harrington, R. R. Schultz, Y. Watanabe, M. Johnson, J. Fang, S. E. O'Brien, I. M. Orioli, E. E. Castilla, D. R. Fitzpatrick, R. Jiang, M. L. Marazita and J. C. Murray (2005). "Medical sequencing of candidate genes for nonsyndromic cleft lip and palate." *PLoS Genet* **1**(6): e64.

Villaverde, A. and M. M. Carrio (2003). "Protein aggregation in recombinant bacteria: biological role of inclusion bodies." *Biotechnol Lett* **25**(17): 1385-95.

Vintiner, G. M., I. K. Temple, H. R. Middleton-Price, M. Baraitser and S. Malcolm (1991). "Genetic and clinical heterogeneity of Stickler syndrome." *Am J Med Genet* **41**(1): 44-8.

Wakiyama, M., T. Futami and K. Miura (1997). "Poly(A) dependent translation in rabbit reticulocyte lysate." *Biochimie* **79**(12): 781-5.

Wales, C. J., K. Corsar and M. F. Devlin (2009). "Submucous cleft palate." *Br Dent J* **207**(6): 254.

Wardle, F. C. and V. E. Papaioannou (2008). "Teasing out T-box targets in early mesoderm." *Curr Opin Genet Dev* **18**(5): 418-25.

Warrington, A., A. R. Vieira, K. Christensen, I. M. Orioli, E. E. Castilla, P. A. Romitti and J. C. Murray (2006). "Genetic evidence for the role of loci at 19q13 in cleft lip and palate." *J Med Genet* **43**(6): e26.

Weatherbee, S. D., L. A. Niswander and K. V. Anderson (2009). "A mouse model for Meckel syndrome reveals Mks1 is required for ciliogenesis and Hedgehog signaling." *Hum Mol Genet* **18**(23): 4565-75.

Weinmann, A. S. (2004). "Novel ChIP-based strategies to uncover transcription factor target genes in the immune system." *Nat Rev Immunol* **4**(5): 381-6.

Weinstein, E. D. and M. M. Cohen (1966). "Sex-linked cleft palate. Report of a family and review of 77 kindreds." *J Med Genet* **3**(1): 17-22.

Wellik, D. M. (2007). "Hox patterning of the vertebrate axial skeleton." *Dev Dyn* **236**(9): 2454-63.

Welsh, I. C. and T. P. O'Brien (2009). "Signaling integration in the rugae growth zone directs sequential SHH signaling center formation during the rostral outgrowth of the palate." *Dev Biol* **336**(1): 53-67.

White, P. H. and D. L. Chapman (2005). "Dlx1 is a downstream target of Tbx6 in the paraxial mesoderm." *Genesis* **42**(3): 193-202.

Wieland, I., W. Reardon, S. Jakubiczka, B. Franco, W. Kress, C. Vincent-Delorme, P. Thierry, M. Edwards, R. Konig, C. Rusu, S. Schweiger, E. Thompson, S. Tinschert, F. Stewart and P. Wieacker (2005). "Twenty-six novel EFNB1 mutations in familial and sporadic craniofrontonasal syndrome (CFNS)." *Hum Mutat* **26**(2): 113-8.

Wigler, M., R. Sweet, G. K. Sim, B. Wold, A. Pellicer, E. Lacy, T. Maniatis, S. Silverstein and R. Axel (1979). Transformation of mammalian cells with genes from prokaryotes and eukaryotes. *Cell* **16**: 777-85.

Wilcoxon, F. (1945). "Individual Comparisons by Ranking Methods." *Biometrics Bulletin* **1**: 80-83.

Wilson, I. and M. Gamble (2002). Section 8: The haematoxylin stains. *Theory and practice of histological techniques*. J. D. Bancroft and M. Gamble. London, Elsevier Science Ltd.

Wong, E. and C. L. Wei (2009). "ChIP'ing the mammalian genome: technical advances and insights into functional elements." *Genome Med* **1**(9): 89.

Wyszynski, D. F. and T. H. Beaty (1996). "Review of the role of potential teratogens in the origin of human nonsyndromic oral clefts." *Teratology* **53**(5): 309-17.

Yagi, H., Y. Furutani, H. Hamada, T. Sasaki, S. Asakawa, S. Minoshima, F. Ichida, K. Joo, M. Kimura, S. Imamura, N. Kamatani, K. Momma, A. Takao, M. Nakazawa, N. Shimizu and R. Matsuoka (2003). "Role of TBX1 in human del22q11.2 syndrome." *Lancet* **362**(9393): 1366-73.

Yoneda, T. and R. M. Pratt (1981). "Mesenchymal cells from the human embryonic palate are highly responsive to epidermal growth factor." *Science* **213**(4507): 563-5.

Yoon, H., I. S. Chung, E. Y. Seol, B. Y. Park and H. W. Park (2000). "Development of the lip and palate in staged human embryos and early fetuses." *Yonsei Med J* **41**(4): 477-84.

Yoshikawa, H., T. Kukita, K. Kurisu and H. Tashiro (1987). "Effect of retinoic acid on in vitro proliferation activity and glycosaminoglycan synthesis of mesenchymal cells from palatal shelves of mouse fetuses." *J Craniofac Genet Dev Biol* **7**(1): 45-51.

Zambonato, T. C., M. R. Feniman, W. Q. Blasca, J. R. Lauris and L. P. Maximino (2009). "Profile of patients with cleft palate fitted with hearing AIDS." *Braz J Otorhinolaryngol* **75**(6): 888-92.

Zeiger, J. S., T. H. Beaty and K. Y. Liang (2005). "Oral clefts, maternal smoking, and TGFA: a meta-analysis of gene-environment interaction." *Cleft Palate Craniofac J* **42**(1): 58-63.

Zhang, S. and A. Nohturfft (2008). "Studying Membrane Biogenesis with a Luciferase-Based Reporter Gene Assay." *J Vis Exp*(19).

Zhang, Z., Y. Song, X. Zhao, X. Zhang, C. Fermin and Y. Chen (2002). "Rescue of cleft palate in Msx1-deficient mice by transgenic Bmp4 reveals a network of BMP and Shh signaling in the regulation of mammalian palatogenesis." *Development* **129**(17): 4135-46.

Zuccheri, T. M., M. E. Cooper, B. S. Maher, S. Daack-Hirsch, B. Nepomuceno, L. Ribeiro, D. Caprau, K. Christensen, Y. Suzuki, J. Machida, N. Natsume, K. Yoshiura, A. R. Vieira, I. M. Orioli, E. E. Castilla, L. Moreno, M. Arcos-Burgos, A. C. Lidral, L. L. Field, Y. E. Liu, A. Ray, T. H. Goldstein, R. E. Schultz, M. Shi, M. K. Johnson, S. Kondo, B. C. Schutte, M. L. Marazita and J. C. Murray (2004). "Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate." *N Engl J Med* **351**(8): 769-80.

Zweier, C., H. Sticht, I. Aydin-Yaylagul, C. E. Campbell and A. Rauch (2007). "Human TBX1 missense mutations cause gain of function resulting in the same phenotype as 22q11.2 deletions." *Am J Hum Genet* **80**(3): 510-7.

## APPENDICES

### APPENDIX 1: OLIGONUCLEOTIDES AND ANTIBODIES

Primers used to generate DNA fragments, which were ligated into cloning vectors

#### **Primers used to amplify TBX22 coding sequence for ligation in to the pET100/D-TOPO vector**

5'-(CACCATGGCTCTGAGCTCTCGGG)-3'  
5'-(CTAAAGGTAATGGTTAATTGCTGG)-3'

#### **Primers used to generate PCR product for ligation in to pTNT**

XhoI-HisTBX22(F)  
5'-(CAGCACTCGAGTCACAATGGGGGGTCTCATCATCATC)-3'  
SalI-HisTBX22(R)  
5'-(CCATGGTCGACTCGCGCTAAAGGTAATGGT TAATTGCTGGATAC)-3'

#### **Primers used to generate the pCR3.1\_TBX22 plasmid**

5'-(GGGATGGCTCTGAGCTCTC)-3'  
5'-(CATAGGTAATGGTTAATTGCTGAA)-3'

Primers used for RT-PCR

#### ***TBX22***

5'-(AAGCGGGCAGGCGGATGTT)-3' (481-500 NM\_001109878)  
5'-(AGGTCTCTCCGAGCAG GGT)-3' (680-700 NM\_001109878)

#### ***MSX1***

5'-(CTGGAGCGCAAGTT CCGCCA)-3' (616-637 NM\_002448)  
5'-(AGGCACCGTAGAG CGAGGCA)-3'' (872-892 NM\_002448)

#### ***GAPDH***

5'-(TGCACCACCAACTGCTTAGC)-3' (556-572 NM\_002046)  
5'-(GGCATGGACTGTGGTCATGAG)-3' (NM\_622-642 NM\_002046)

Sequencing primers

T7 Forward 5' (TAATACGACTCACTATAGGG) 3'  
T7 Reverse 5' (TATGCTAGTTATTGCTCAG) 3'

M13 Forward 5'-(CGCCAGGGTTTCCCAGTCACGAC)-3'  
M13R Reverse 5'-(TCACACAGGAAACAGCTATGAC)-3'

## TBX22 coding sequence

SEQ\_1 5'-(GGCACGAGGCAAAGAAC)-3'  
SEQ\_2 5'-(CTGAAAGTCTGGAAGAGAAAG)-3'  
SEQ\_3 5'-(CGCTATAGGTACGTCTAT)-3'  
SEQ\_4 5'-(AGGCCTTCTTCACCTCTC)-3'  
SEQ\_5 5'-(CAGTCTTAGCCCCACTC)-3'  
SEQ\_6 5'-(TAAAGAAACTGAGTCAC)-3'

## **pCR3.1\_TBX22 Sequencing**

5'-TAATACGACTCACTATAGGG-3'  
5'-TAGAAGGCACAGTCGAGG-3'

### M\_prom sequencing primer

5'-CTAGAAAAATAGGCTGTCCC-3'  
5'-TCGATATGTGCGTCGGTAAA-3'

### Oligonucleotides used in the *in vitro* binding selection

## Random Oligo

## Randon oligo primers

5'-(GCTGCAGTTGCACTGAATTGCGCCT)-3'  
5'-(CAGGTCAAGTCAGCGGATCCTGTCG)-3'

### Oligonucleotides used in the EMSA experiments

W=T\_oligo

5'-(CGAGAGGTGTCTTACGAG)-3'  
5'-(CTCGTAAGACACCTCTCG)-3'

515-20\_oligo

5'-(CGAGAGGTGTCATACGAG)-3'  
5'-(CTGGTATGACACCTCTCG)-3'

## Mut\_oligo

5'-CGAGAAATGTCATACGAG-3'  
5'-CTCGTATGACATTCTCG-3'

### Cont. oligo

5'-(CTAGCAAGGTGTGAAATTGTCACCTCAA)-3'  
5'-(GTTCCACACTTAACAGTGGAGTTTCGA)-3'

## 5'-GTCCTA MSX1 oligo

**MSXI<sub>1</sub>\_org**  
5'-(CGAGAGGTGTTGAGCGAG)-3'  
5'-(CTCGCTAACACCTCTCG)-3'

Antibodies

**Anti-Luciferase pAb - Promega**

**Anti-HisG Antibody – Invitrogen**

**Anti-HisG-HRP Antibody Invitrogen**

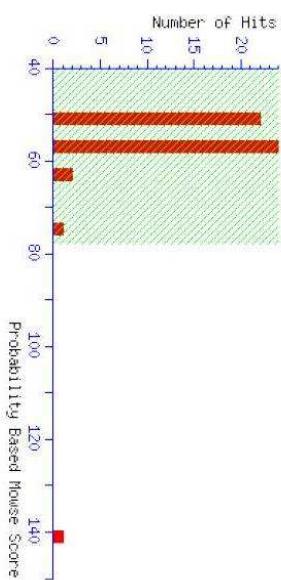
**TBX22 (K-20) X sc-17862 X – Santa Cruz Biotechnology**

## {MATERIAL} {SCIENCE} Mascot Search Results

User : Joe  
 Email : joe.gray@ncl.ac.uk  
 Search title : SL\_TBX22\_0001.dat - SpecView  
 Database : MSDB 20060831 (3239079 sequences; 107954700 residues)  
 Timestamp : 26 Mar 2008 at 12:07:25 GMT  
 Top Score : 141 for **AAK63189**, AY035371 NID: - Homo sapiens

### Probability Based Mowse Score

Protein score is  $-10 \cdot \log(P)$ , where  $P$  is the probability that the observed match is a random event. Protein scores greater than 7.8 are significant ( $p < 0.05$ ).



### Protein Summary Report

<a href="#">Format As</a>	<a href="#">Protein Summary</a>	<a href="#">Help</a>
Significance threshold p<		Max. number of hits
<input type="text" value="0.05"/>		<input type="text" value="50"/>
<a href="#">Re-Search All</a> <a href="#">Search Unmatched</a>		

### Index

Accession	Mass	Score	Description
1. <a href="#">AAK63189</a>	578.3	141	AY035371 NID: - Homo sapiens
2. <a href="#">Q5JZ06</a> HUMAN	446.9	74	T-box 22 - Homo sapiens (Human).
3. <a href="#">Q4DQC5</a> TRYCR	544.8	64	Hypothetical protein. - Trypanosoma cruzi.
4. <a href="#">Q2FNB7</a> METHJ	367.95	63	Hypothetical protein. - Methanospirillum hungatei (strain JF-1 / DSM 864).

**Appendix 2.1: The protein match to TBX22 is statistically significant.** The only matched protein in the sample which lies beyond the level of statistical significance using a Mowes probability score is TBX22, which has a probability based Mowes core of 140.

## {MATRIX} Mascot Search Results

### Protein View

Match to: **AAK63189** Score: 141 Expect: 2.6e-08  
**AY035371** MID: - Homo sapiens

Nominal mass (M<sub>r</sub>): 57873; Calculated pI value: 6.92  
 NCBI BLAST search of AAK63189 against nr  
 Unformatted sequence string for pasting into other applications

Taxonomy: Homo sapiens  
 Links to retrieve other entries containing this sequence from NCBI Entrez:  
AAH14194 from Homo sapiens  
CAI43070 from Homo sapiens  
TBX22 HUMAN from Homo sapiens

Variable modifications: Carbamidomethyl (C), Gln->pyro-Glu (N-term Q), Oxidation (M)  
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
 Number of mass values searched: 39  
 Number of mass values matched: 20  
 Sequence Coverage: 40%

Matched peptides shown in **Bold Red**

1 **MA**SSRAR**A**F **S**VEALVGRPS **K**RKL**Q**D**P**I**Q**A **E**QPELREKKG GEEEEERRSS  
 51 AAGKSEPLEK QPKTEPSTSA SSGCGSDSGY GNSSSESLEEK **DIQ**M**E**L**Q**G**S**E  
 101 **L**WKR**F**H**D**I**G**T **E**MI**I**T**K**AGR**R** MFP**S**VR**V**K**V**K GLDPG**K**Q**I**H**V** AID**V**V**P**V**D**SK  
 151 RYRYV**V**HSSQ **W**M**V**AGNT**D**H**L** C**I**I**P**RF**V**Y**V**H**D** SP**C**SG**E**T**W**M**R** RQ**I**I**S**FD**R**MK  
 201 LT**M**NEMDD**K** H**I**LLQSM**H**KY KPR**V**H**V**IE**Q**G SS**V**D**L**S**Q**I**Q**S LP**T**E**G**W**K**T**F**S  
 251 F**K**E**T**E**F**T**T**V**T** **A**Y**Q**N**Q**Q**I**T**K**L KIERNPFA**K**G FRDTG**R**NR**G**V LD**G**LL**E**T**P**W  
 301 **R**PS**F**T**L**D**F**K**T** FG**A**D**T**OS**G**SS GS**S**SP**V**T**S**GG AP**S**PL**N**SL**L**S PLC**F**SP**M**FL  
 351 PT**S**SL**G**MP**C** E**A**Y**L**PN**V**N**L**P LC**V**K**I**C**P**T**N**E **W**Q**Q**Q**P**L**V**L**P**A PER**L**ASS**N**SS  
 401 QSLAP**L**MM**V** PNLSS**L**GV**V**N SK**G**S**S**ED**S**SS DQ**Y**L**Q**AP**N**ST NQ**M**LY**G**Q**S**P  
 451 GNIFLP**N**M**V** PEALSCSF**H**P SYDFY**R**Y**N**FS **M**PS**R**L**I**SG**S**N HL**K**V**N**DD**S**QV  
 501 SF**G**E**G**K**C****H**N**V** **W**Y**P**A**I**N**H****Y**

Show predicted peptides also

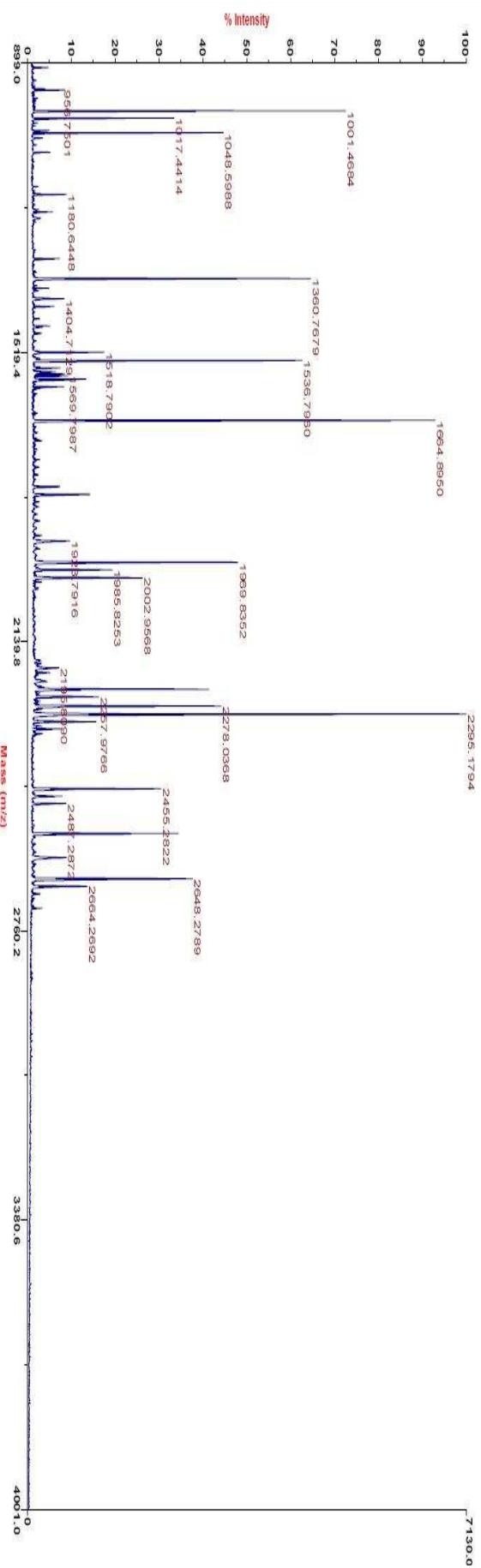
Sort Peptides By  Residue Number  Increasing Mass  Decreasing Mass

Start - End	Observed	Mr (expt)	Mr (calc)	ppm	Miss	Sequence
9 - 21	<b>1360.7682</b>	<b>1359.7609</b>	<b>1359.7510</b>	7	<b>0</b>	<b>R.A</b> F <b>S</b> VEALVGRPS <b>K</b> R
23 - 36	<b>1664.8952</b>	<b>1663.8879</b>	<b>1663.8893</b>	-1	<b>1</b>	<b>R.K</b> L <b>Q</b> D <b>P</b> I <b>Q</b> A <b>E</b> QPEL <b>R</b>
24 - 36	<b>1536.7963</b>	<b>1535.7891</b>	<b>1535.7943</b>	-3	<b>0</b>	<b>K.L</b> Q <b>D</b> P <b>I</b> Q <b>A</b> E <b>Q</b> PEL <b>R</b>
91 - 103	<b>1576.7610</b>	<b>1575.7537</b>	<b>1575.7603</b>	-4	<b>0</b>	<b>K.D</b> I <b>Q</b> M <b>E</b> L <b>Q</b> G <b>S</b> EL <b>W</b> K <b>R</b>
91 - 103	<b>1592.7451</b>	<b>1591.7378</b>	<b>1591.7552</b>	-11	<b>0</b>	<b>K.D</b> I <b>Q</b> M <b>E</b> L <b>Q</b> G <b>S</b> EL <b>W</b> K <b>R</b> Oxidation (M)
104 - 116	<b>1560.8278</b>	<b>1559.8205</b>	<b>1559.8130</b>	5	<b>1</b>	<b>K.R</b> F <b>H</b> D <b>I</b> G <b>T</b> E <b>M</b> I <b>I</b> T <b>K</b> <b>A</b>
105 - 116	<b>1404.7124</b>	<b>1403.7051</b>	<b>1403.7119</b>	-5	<b>0</b>	<b>R.F</b> H <b>D</b> I <b>G</b> T <b>E</b> M <b>I</b> I <b>T</b> <b>K</b> <b>A</b>
137 - 150	<b>1552.7799</b>	<b>1551.7727</b>	<b>1551.7933</b>	-13	<b>0</b>	<b>K.Q</b> Y <b>H</b> V <b>A</b> l <b>D</b> V <b>V</b> P <b>V</b> D <b>S</b> <b>K</b> <b>R</b> Gln->pyro-Glu (N-term Q)
137 - 150	<b>1569.7993</b>	<b>1568.7920</b>	<b>1568.8199</b>	-18	<b>0</b>	<b>K.Q</b> Y <b>H</b> V <b>A</b> l <b>D</b> V <b>V</b> P <b>V</b> D <b>S</b> <b>K</b> <b>R</b>
154 - 175	<b>2647.2732</b>	<b>2646.2659</b>	<b>2646.2471</b>	7	<b>0</b>	<b>R.Y</b> Y <b>H</b> S <b>S</b> Q <b>W</b> M <b>V</b> AGNT <b>D</b> H <b>L</b> C <b>I</b> I <b>P</b> <b>R</b> <b>F</b> Carbamidomethyl (C)
154 - 175	<b>2663.2730</b>	<b>2662.2657</b>	<b>2662.2421</b>	9	<b>0</b>	<b>R.Y</b> Y <b>H</b> S <b>S</b> Q <b>W</b> M <b>V</b> AGNT <b>D</b> H <b>L</b> C <b>I</b> I <b>P</b> <b>R</b> <b>F</b> Carbamidomethyl (C); Oxidation (M)
176 - 191	<b>1968.8306</b>	<b>1967.8233</b>	<b>1967.8295</b>	-3	<b>0</b>	<b>R.F</b> Y <b>V</b> H <b>D</b> SP <b>C</b> SG <b>E</b> T <b>W</b> M <b>R</b> <b>Q</b> Carbamidomethyl (C)
176 - 191	<b>1984.8228</b>	<b>1983.8155</b>	<b>1983.8244</b>	-4	<b>0</b>	<b>R.F</b> Y <b>V</b> H <b>D</b> SP <b>C</b> SG <b>E</b> T <b>W</b> M <b>R</b> <b>Q</b> Carbamidomethyl (C); Oxidation (M)
224 - 247	<b>2550.3483</b>	<b>2549.3410</b>	<b>2549.3337</b>	3	<b>0</b>	<b>R.V</b> H <b>V</b> I <b>E</b> F <b>T</b> T <b>V</b> T <b>A</b> Y <b>Q</b> N <b>Q</b> I <b>T</b> <b>K</b> <b>L</b>
253 - 269	<b>2001.9515</b>	<b>2000.9442</b>	<b>2000.9691</b>	-12	<b>0</b>	<b>K.E</b> TE <b>F</b> T <b>T</b> V <b>A</b> Y <b>Q</b> N <b>Q</b> I <b>T</b> <b>K</b> <b>L</b>
289 - 309	<b>2454.2861</b>	<b>2453.2788</b>	<b>2453.2631</b>	6	<b>0</b>	<b>R.G</b> W <b>D</b> G <b>L</b> E <b>T</b> <b>P</b> W <b>R</b> P <b>S</b> F <b>T</b> <b>L</b> <b>D</b> <b>F</b> <b>K</b> <b>T</b>
375 - 393	<b>2294.1781</b>	<b>2293.1708</b>	<b>2293.1678</b>	1	<b>0</b>	<b>K.I</b> C <b>P</b> T <b>N</b> FW <b>Q</b> Q <b>P</b> L <b>V</b> P <b>A</b> <b>R</b> <b>E</b> Carbamidomethyl (C)
477 - 484	<b>1001.4683</b>	<b>1000.4610</b>	<b>1000.4436</b>	17	<b>0</b>	<b>R.Y</b> HF <b>S</b> M <b>P</b> S <b>R</b> <b>L</b> Oxidation (M)
507 - 520	<b>1823.8235</b>	<b>1822.8163</b>	<b>1822.8362</b>	-11	<b>0</b>	<b>K.C</b> N <b>H</b> V <b>W</b> Y <b>P</b> A <b>I</b> N <b>H</b> <b>Y</b> Carbamidomethyl (C)

No match to: 956.9027, 1048.5985, 1180.6461, 1318.5959, 1518.7909, 1565.6542, 1806.7971, 1923.8136, 2195.8660, 2240.9871, 2256.9902, 2277.0368, 2292.1419, 2310.1749, 2326.1808, 2470.2774, 2486.2894, 2602.1761, 2604.0537

**Appendix 2.2: The peptides detected by MALDI analysis.** The peptides detected by MALDI analysis are aligned to the TBX22 protein and provide 40% coverage of the protein. The peptide identification numbers correspond to those in appendix 2.3.

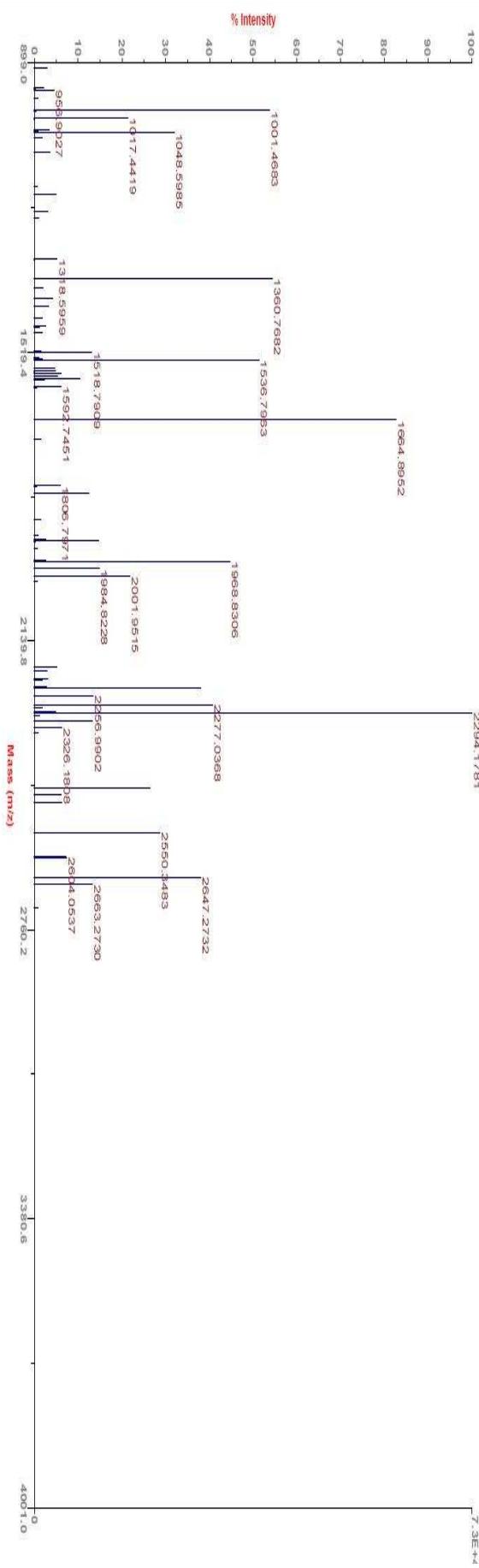
Voyager Spec #1[BP = 2295.2, 7130]



Voyager Spec #1=>AdvBC(32,0,5,0,1)=>INFO.7=>DI[BP = 2294.2, 7307]

Sample: SL\_TBx22

7.3E+42



**Appendix 2.3: A Graphical representation of the peptides detected by MALDI analysis.** The molecular mass of the peptides detected in the protein sample by MALDI are represented by their identification number which can be found in appendix 2.2

## APPENDIX 3: THE TBX22 DNA CODING AND PROTEIN SEQUENCE

-3 GGGatggctctgagctctcgggcgcgtgccttcgttggaaagccttggggagaccc  
M A L S S R A R A F S V E A L V G R P  
-----|-----|-----|-----|-----|-----|-----|-----|  
61 agcaaaaagaaaactccaagacccaatacaggcggagcagcctgagctgcgggagaaaaag  
20 S K R K L Q D P I Q A E Q P E L R E K K  
-----|-----|-----|-----|-----|-----|-----|-----|  
121 ggcggagagaagaggaggagagaaggagcagcgtgcaggaaagagcgagccgcttgcggaa  
40 G G E E E E R R S S A A G K S E P L E  
-----|-----|-----|-----|-----|-----|-----|-----|  
181 aaacaacctaagacagacagccctcaacatctgccttcgtggctgcggcagcgcacagcggc  
60 K Q P K T E P S T S A S S G C G S D S G  
-----|-----|-----|-----|-----|-----|-----|-----|  
241 tacggcaacacagctctgaaagtcttgcggaaagagaagatattcaaatggagcttcaaggatct  
80 Y G N S S E S L E E K D I Q M E L Q G S  
-----|-----|-----|-----|-----|-----|-----|-----|  
301 gaactgtggaaaagatccatgacatcggaactgagatgatcattactaaagcgggcagg  
100 E L W K R F H D I G T E M I I T K A G R  
-----|-----|-----|-----|-----|-----|-----|-----|  
361 cgatgttcccctctgttcgggtcaagggtgaaagggtggatccaggaaagcagtaccat  
120 R M F P S V R V K V K G L D P G K Q Y H  
-----|-----|-----|-----|-----|-----|-----|-----|  
421 gtggccatcgatgtggccgggtggattccaaacgcgtataaggtagtacgtctatcacagctca  
140 V A I D V V P V D S K R Y R Y V Y H S S  
-----|-----|-----|-----|-----|-----|-----|-----|  
481 cagtggatggtagctggaaatacagaccattgtgcatttccttagatttatgttac  
160 Q W M V A G N T D H L C I I P R F Y V H  
-----|-----|-----|-----|-----|-----|-----|-----|  
541 ccggactcaccctgctcggaagagacctggatgcggcagatcatcagcttgcgcatt  
180 P D S P C S G E T W M R Q I I S F D R M  
-----|-----|-----|-----|-----|-----|-----|-----|  
601 aaactcaccacaaatgagatggatgacaaaggccacatcattctgcaatccatgcataag  
200 K L T N N E M D D K G H I I L Q S M H K  
-----|-----|-----|-----|-----|-----|-----|-----|  
661 tacaaaccccgagtgcacgtgatagagcaaggcagcagtgttgcacccatccatcgattcag  
220 Y K P R V H V I E Q G S S V D L S Q I Q  
-----|-----|-----|-----|-----|-----|-----|-----|  
721 tccttgccactgaagggttaaaacatttcctttaaagaaaacttagttaccacagta  
240 S L P T E G V K T F S F K E T E F T T V  
-----|-----|-----|-----|-----|-----|-----|-----|  
781 acggcttacaaaaccaacagattacgaaactaaaaatagaaaagaaaatcctttgtcaaaa  
260 T A Y Q N Q Q I T K L K I E R N P F A K  
-----|-----|-----|-----|-----|-----|-----|-----|  
841 ggatttagagatactgaaagaaaacagggtgtattggatggcttttagagacccatcca  
280 G F R D T G R N R G V L D G L L E T Y P  
-----|-----|-----|-----|-----|-----|-----|-----|  
901 tggaggcctctttcaactctcgatttaaaaccttggcgcagacacacaaaatgtggaaac  
300 W R P S F T L D F K T F G A D T Q S G S  
-----|-----|-----|-----|-----|-----|-----|-----|  
961 agtggctcatctccagtgacctcttagtgagggggccccctctcctttgaactccttactt  
320 S G S S P V T S S G G A P S P L N S L L  
-----|-----|-----|-----|-----|-----|-----|-----|  
1021 tctccacttgccttcacctatgtttcaccttacatcacaagcgtcccttggaaatgcctgt  
340 S P L C F S P M F H L P T S S L G M P C  
-----|-----|-----|-----|-----|-----|-----|-----|  
1081 ccagaggcatacctggccaatgtcaacctgcctcatgttacaagatgttgcactaat  
360 P E A Y L P N V N L P L C Y K I C P T N  
-----|-----|-----|-----|-----|-----|-----|-----|  
1141 ttttggcaacagcaacccatttttaccggcgttgcggaaagactagcaagcagcaacagt  
380 F W Q Q Q P L V L P A P E R L A S S N S  
-----|-----|-----|-----|-----|-----|-----|-----|  
1201 tctcagtctttagccccactcatgtatggaaagtgccatgttaccccttgggggttgcacc  
400 S Q S L A P L M M E V P M L S S L G V T  
-----|-----|-----|-----|-----|-----|-----|-----|  
1261 aattcaaaaqcgqgtcatctqaaqactccagtgatcgtatctacaacqcacctaattct

```

420  N  S  K  S  G  S  S  E  D  S  S  D  Q  Y  L  Q  A  P  N  S
-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1321 accaatcaaatgttatatggattacagtcacctggaaatattttctgccaaactccatc
440  T  N  Q  M  L  Y  G  L  Q  S  P  G  N  I  F  L  P  N  S  I
-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1381 accccagaaggacttagttgctccttcattccttatgactttatagataacaatttc
460  T  P  E  A  L  S  C  S  F  H  P  S  Y  D  F  Y  R  Y  N  F
-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1441 tctatgccatctagactgataagtgggttccaaccatcttaaagtgaatgacgacagtcaa
480  S  M  P  S  R  L  I  S  G  S  N  H  L  K  V  N  D  D  S  Q
-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1501 gtttctttggagaaggcaaatgtaatcatgttcattggtatccagcaattaaccattac
500  V  S  F  G  E  G  K  C  N  H  V  H  W  Y  P  A  I  N  H  Y
-----|
1561 ctttag 1566
520  L  *  520

```

**Appendix 3: The TBX22 DNA coding and protein sequence.** The initiation and stop codons are highlighted in bold and the three bases upstream of the translation start site which form part of the Kozac sequence are shown in capitals. Single letter amino acid abbreviations are shown below their respective codons.

APPENDIX 4: ALIGNMENT OF THE CLONES FROM THE *IN VITRO* DNA  
BINDING ASSAY AND AN EXAMPLE CHROMATOGRAM OF CLONED DNA  
FROM *IN VITRO* OLIGONUCLEOTIDE BINDING STUDY

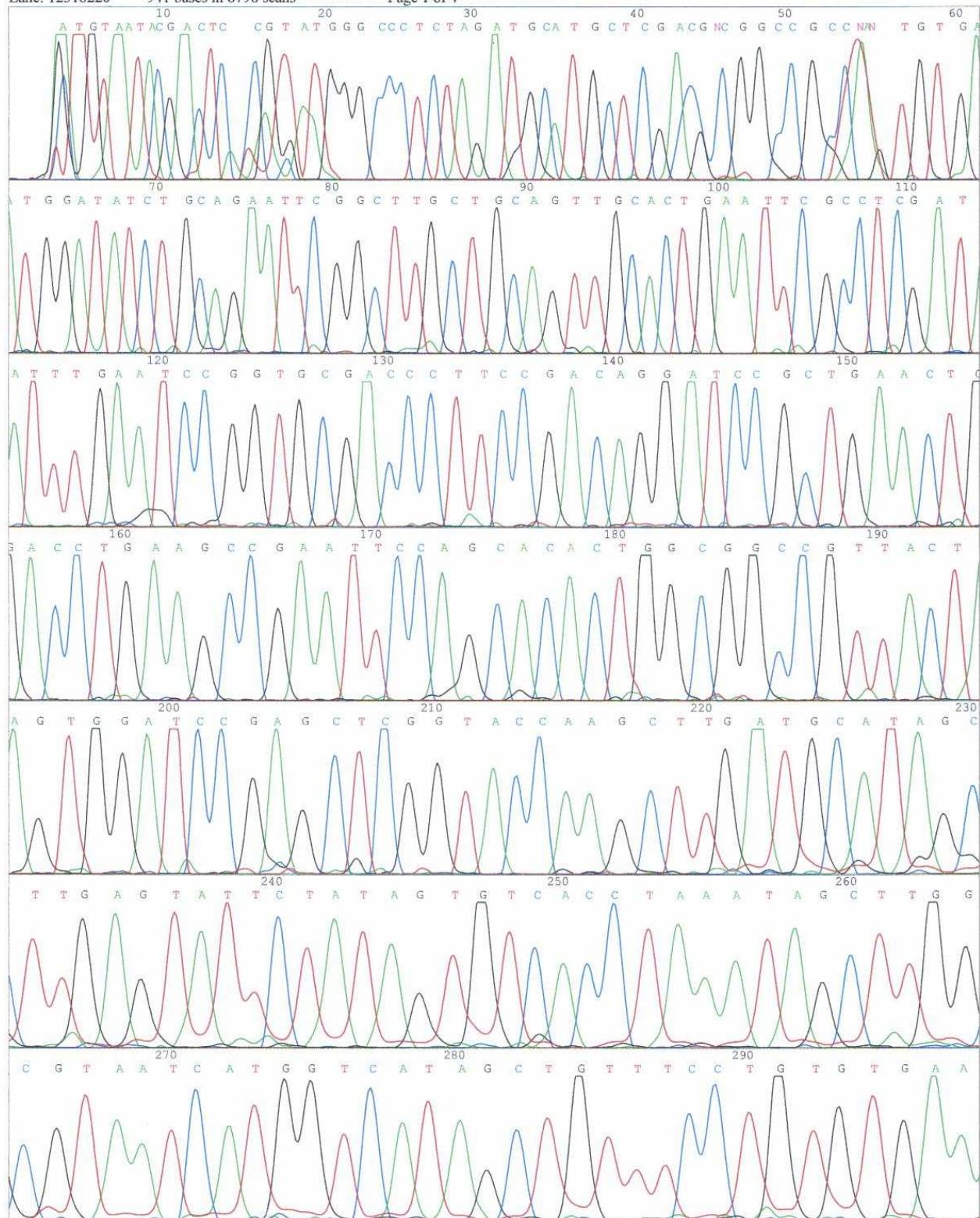
Clone ID	Sequence
515-1	A T A T T G A A T C C G C G A C C C C T T C C
515-2	T G G A C T C C G C C G G T A T G A C T A C T
515-3	A C A T A T A G G C T A C C G T C T T A G G G T
515-4	A C C A T C C G A G C T G G T G T A A C C A A T C
515-6	G A C C G T G T C T T A T A T A C G A A G T C G A A C
515-8	G C T A G C G T C A G G T G C C A A T G C A A C T G C A G C
515-12	A C G A C G T C C C A G G T G T C C G A A C
515-14	C C C A C G G T T C C C A G G T G T C C G A A C
515-15	A A C T T A C G G T T C A C G T G T G G A A G C A
515-17	A T A C G C A G G C A G G T G T G G A A G C A
515-19	T A C C T G C A G G T G T C T T A T C A G C G A
515-20	C G A C T T C G T A T A G G T G T G A T C G A C G T
515-21	C C G C T C C A G G A A G G T G T C T T A C T C C
515-23	T C T C A G G T G T C T T A A G G C C A A
515-24	A A C G T A T A C A A G G T G T C T T A A C T C
515-25	G T C G C C C G C A C G T G T C T G C A A C A A T A A C G A C T G
515-27	C T G T A C C G T G C A A C A A T A A C G A C T G
515-28	T T C G G T G T G A C A G G T G A C G C A C G A G C
515-29	T C A A G A T G G C G G T G T G A C G G A T G G A C C G
515-30	G T T C G A C A G A A G G T G T G C G T G A T C T G

File: 515-1.abd  
Lane: 12318220

Run Ended: Dec 2, 2005, 18:15:02  
941 bases in 8798 scans

Comment: Standard Terminator  
Page 1 of 4

Sample: 515-1



**Appendix 4.1: An example of a chromatogram of the DNA from the *in vitro* oligonucleotide binding assay.**  
The sequencing of the clone DNA was performed using the M13 forward and reverse sequencing primers. This chromatogram is from the clone 515-1. A Genbank formatted sequence of this trace is shown in appendix 3.2

```
1 ATGTAATACG ACTCCGTATG GGCCCTCTAG ATGCATGCTC GACGNCGGCC GCCNANTGTG
 61 ATGGATATCT GCAGAATTG GCTTGGCTGCA GTTGCACTGA ATTCGCCTCG ATATTTGAAT
121 CCGGTGCGAC CCCTCCGACA GGATCCGCTG AACTGACCTG AAGCCGAATT CCAGCACACT
181 GGCGGCCGTT ACTAGTGGAT CCGAGCTCGG TACCAAGCTT GATGCATAGC TTGAGTATTG
241 TATAGTGTCA CCTAAATAGC TTGGCGTAAT CATGGTCATA GCTGTTCCCT GTGTGAAATT
301 GTTATCGCT CACAATTCCA CACAACATAC GAGCCGGAAG CATAAAAGTGT AAAGCCTGGG
361 GTGCCTAATG AGTGAGCTAA CTCACATTAA TTGGCGTTGCC CTCACTGCC GCTTCCAGT
421 CGGGAAACCT GTCGTGCCAG CTGCATTAAT GAATCGGCCA ACGCGGGGG AGAGGCAGTT
481 TGCGTATTGG GCGCTCTTCC GCTTCCTCGC TCACTGACTC GCTGCGCTCG GTCGTTCGGC
541 TGCGGCGAGC GGTATCAGCT CACTCAAAGG CGGGTAATTAA CGGTTTATCC ACAGAATTCA
601 GGGGATAACG CAGGAAAGAA CATGTGAGCC AAAGGCCCGC AAAGGGGCCA AGGAACCCGT
661 AAAAAAGGGG CCGGGTTTGC TGGGCGNTT TTTCCCATTA AGGCTTCCCG GCCCCCCTGA
721 ACGAAGCCATC ACAAAAAAATC CGNACGTCCA AGTCAGAAGN TGGCGAACCN GACAGGNNN
781 TTAAAGAAAA CAGGGTCCCC CGAANAAACA CACCGCGCNC TTTTNGACAA CGGAGAAAAA
841 AAAGGGGACC ACACGAAAGN GAGAGNACAC AAACACCNGA TGATNAAAGA GCGTACACGG
901 GGTGTGACGC ACCGGGTATA TATTACACAC AACCCGAGGG T
```

**Appendix 4.2: The Genbank formatted sequence of clone 515-1.** The known 5' and 3' flanking regions of the random oligonucleotide used in the in vitro oligonucleotide binding assay are highlighted in bold; the sequence used in the alignment of the cloned DNA which generated the TBX22 binding sequence is highlighted in red.

APPENDIX 5: THE GENES SHOWN TO CAUSE A CLEFT PLATE IN HUMAN (TAKEN FROM OMIM 10 FEBRUARY 2010)

Human Disease Gene	OMIM	Gene Ref Seq Nucleotides	Region Searched	Allowing 1 mis-matches from generic T-box sequence Found & Position	Mouse Mutant has a cleft palate abnormality
ACTB	#607371	NC_000007.12 5533312..5536747, complement	5536747..5538747	1238-1247 aggtttgtag	No Cleft palate
ALX3	#136760	NG_012039 5000..15325	3000..5000	584-593 tggtgttcaa 1524-1533 aggtgtttt	No Cleft palate
ATR	#210600	NC_000003 143650767..143780358, complement	143780358..143782358	710-719 gggtgtggtg 840-849 aggcgtggtg	No Cleft palate
ATRX	#301040	NC_000023.9 76928375..76647012, complement	76928375..76930375	None	No Cleft palate
B3GALT1	#261540	NC_000013.9 30672112..30804413	30670112..30672112	1770-1779 aggggtggag	Not Reported
BCOR	#300166	NC_000023.9 39846998..39795561, complement	39846998..39848998	1757-1766 aggtatggag 1961-1970 aggtatttag	Not Reported
BMP4	#600625	NC_000014.7 53493304..53486204, complement	53493304..53495304	617-626 aggtgtataa 705-714 aagtgtcatg 1390-1399 aggtgttcta 1583-1592 aggtgttca	No Cleft palate
BRAF	#115150	NG_007873.1 5001..195753	3001..5001	53-62 aggtccatg 1162-1171 agatgtctag	No Cleft palate
BUB1B	#257300	NC_000015.8 38240530..38300629	38238530..38240530	260-269 aggggtttag 679-688 aggggttaag 1121-1130 aggtggctaa	No Cleft palate
CD96	#211750	NC_000003.10 112743616..112853896	112741616..112743616	637-646 aggggtgata	Not Reported
CDH1	#137215	NG_008021.1 5000..103249	3000..5000	1152-1161 cggtgtggtg	No Cleft palate
CHD7	#612370	NG_007009.1 5000..193126	3000..5000	1319-1328 aggtggaaa	Cleft palate
CHRN	#265000	NC_000002.10 233112681..233119282	233110681..233112681	None	No Cleft palate
CHST14	#601776	NC_000015.9 40763212..40765353	40761212..40763212	569-578 agatgtggag 1303-1302 atgtgtgttgc	Not Reported
COL11A1	#604841	NG_008033.1 5001..237030	3001..5001	None	Cleft palate

COL11A2	#184840	NC_000006.10 33268223..33238447, complement	33268223..33270223	772-781 aggtgggtg	Cleft palate
COL2A1	#108300	NG_008072.1 5001..36538	3001..5001	260-269 atgtgtggaa 586-595 aagtgtcaaa 927-936 aggtgaggtg 1383-1392 aggggtggag 1820-1829 aggggtggag	Cleft palate
DHCR24	#602398	NC_000001.9 55125509..55087888, complement	55125509..55127509	578-587 agctgtgtg	No Cleft palate
DHCR7	#270400	NC_000011.8 70837125..70823105, complement	70837125..70839125	248-257 agatgtaaa 1718-1727 aggtgatgag	Cleft palate
EFNB1	#304110	NC_000023.9 67965556..67978728	67963556..67965556	178-187 aggagtcaaa 914-923 aggtctgtg 1239-1248 aggtgtctt	Cleft palate
DOK7	#254300	NG_013072.1 5000..36176	3000..5000	1623-1632 ggggtctag	No Cleft palate
ERCC5	#278780	NG_007146.1 5001..35172	3001..5001	None	No Cleft palate
ESCO2	#268300	NG_008117.1 5001..35368	3001..5001	609-618 aggtctggag	Not Reported
EVC	#225500	NC_000004.10 5763825..5866932	5761825..5763825	1646-1655 aggcgtggag 1870-1879 aggggtggaa	No Cleft palate
EXT1	#215300	NG_007455.1 5001..317457	3001..5001	529-538 aggtctgaaa 748-757 aggtgtttc	No Cleft palate
EYA1	#166780	NC_000008.9 72437021..72272222, complement	72437021..72439021	968-977 atgtgttta	Cleft palate
FAM20C	#259775	NC_000007.12 288052..304059	286052..288052	155-164 aggggtgaag	Not Reported
FANCB	#300514	NG_007310.1 5001..34656	3001..5001	241-250 ggggtgtgaag 264-273 aggtatgaaa	No Cleft palate
FBN1	#154700	NC_000015.8 46725210..46487797, complement	46725210..46727210	1978-1987 aggtggggaa	No Cleft palate
FGD1	#305400	NG_008054.1 5001..55713	3001..5001	1043-1052 aggtggggaaa 1337-1346 aggttgaaaa	No Cleft palate
FGFR1	#147950	NG_007729.1 5001..62697 NC_000010.9 123347962..123227845,	3001..5001	None	Cleft palate
FGFR2	#101200	complement	123347962..123349962	None	Cleft palate
FKRP	#236670	NC_000019.8 51941143..51953582	51939143..51941143	None	Not Reported
FKTN	#613152	NC_000009.10 107360232..107443220	107358232..107360232	None	No Cleft palate
FLNA	#304120	NC_000023.9 153252845..153230091,	153252845..153254845	388-397 acgtgtcatg	Cleft palate

complement						
FLNB	#272460	NC_000003.10 57969167..58133018	57967167..57969167	1334-1343 aggtgatgag		
FOXC2	#153400	NC_000016.8 85158443..85159948	85156443..85158443	463-472 aggtgggg	No Cleft palate	
FOXE1	#241850	NC_000009.10 99655358..99658818	99653358..99655358	18-27 aggttttag	Cleft palate	
				365-374 aggaggtttgc	Cleft palate	
FRAS1	#219000	NC_000004.10 79198120..79684447	79196120..79198120	1432-1441 cggtgtggtg	No Cleft palate	
FREM2	#219000	NG_008125.1 5001..205096	3001..5001	519-528 aggttcatg	No Cleft palate	
FTO	#612938	NG_012969.1 5000..415504	3000..5000	1165-1174 gggtgtggtg	No Cleft palate	
GDF1	#217095	NC_000019.8 18842116..18840354, complement	18842116..18844116	355-364 aggtgtcgat	No Cleft palate	
				926-935 gggtgtggtg		
				1599-1608 gggtgtgaaa		
GJA1	#164200	NG_008308.1 5001..19129	3001..5001	756-765 agatgttata	No Cleft palate	
				1454-1463 aggggttgaa		
GJB2	#124500	NC_000013.9 19665114..19659605, complement	19659605..19661605	998-1007 aggtgtcag	No Cleft palate	
GLI2	#610829	NC_000002.10 121266327..121466321	121264327..121266327	674-683 aggtgttacg	Cleft palate	
				1196-1205 agctgtcatg		
				1472-1481 atgtgtgatg		
GLI3	#146510	NC_000007.12 42243137..41967072, complement	41967072..41969072	805-814 cggtgtggag	Cleft palate	
		NC_000023.9 132947332..132497439,				
GPC3	#312870	complement	132947332..132949332	526-535 gggtgtggag	No Cleft palate	
HOXA2	#612290	NG_012078 5000..7421	3000..5000	61-70 agctgtcaag	Cleft palate	
				550559 aggtgttac		
				740-749 aggtgtggga		
				1597-1606 gggtgtcaaa		
				1657-1666 aggtgcttaa		
HYAL1	#601492	NC_000003.10 50324816..50312324, complement	50324816..50326816	938-947 aggtatgaaa	No Cleft palate	
HYLS1	#236680	NC_000011.8 125258719..125275749	125256719..125258719	1934-1943 aggtgaggta	Not Reported	
ICK	#612651	NG_012159.1 5000..65502	3000..5000	1057-1066 aagtgtcaaa	Not Reported	
				1342-1341 atgtgttata		
IRF6	#608864	NG_007081.1 5001..23218	3001..5001	1129-1138 aggtgggata	Cleft palate	
				1309-1318 aggtgttgca		

KAL1	+308700	NG_007088.1 5001..208313	3001..5001	1341-1350 aggtggtag	Not Reported
KCNJ2	#170390	NC_000017.9 65677271..65687780	65675271..65677271	410-419 agatgtaaa	Cleft palate
KIAA1279	#235730	NC_000010.9 70418499..70446744	70416499..70418499	602-611 aggtgtatt 1286-1295 aggtttag	Not Reported
KRAS	#115150	NG_007524.1 5001..50675	3001..5001	113-122 aggtgatgag	No Cleft palate
LARGE	#119540	NC_000022.9 32646410..31999063, complement	32646410..32648410	None	Not Reported
LMNA	#275210	NC_000001.9 154351085..154376502	154349085..154351085	1544-1553 acgtgtggag	No Cleft palate
LMX1B	#161200	NC_000009.10 128416619..128498551	128414619..128416619	None	No Cleft palate
MAP2K1	#115150	NG_008305.1 5001..109672	3001..5001	339-348 aggtgtcaa	No Cleft palate
MAP2K2	#115150	NG_007996.1 5001..38808	3001..5001	649-658 aagtgtctag 1315-1324 agtttcata	No Cleft palate
MED12	#305450	NC_000023.9 70255131..70279029	70253131..70255131	None	Not Reported
MID1	*300552	NG_008197.1 5001..393135	3001..5001	1205-1214 agttgttaa 1932-1941 aggcgtgtaa	No Cleft palate
MKS1	#249000	NC_000017.9 53651665..53637797, complement	53651665..53653665	27-36 agtatctaa 441-450 agctgtgtg 930-939 aggtggtag	Cleft palate
MKX	*601332	NC_000010.10 28034777..27961802, complement	28036777..28034777	525-534 aggtctgtag	
MSX1	#608874	NG_008121.1 5001..9272	3001..5001		Cleft palate
MSX2	#168500	NG_008124.1 5001..11328	3001..5001	586-595 aggggtcaag 989-998 aggtccatg	Cleft palate
MTR	1195.3	NG_008959.1 5000..113700	3000..5000	540-549 aggagtttag	No Cleft palate
MYH3	#601680	NC_000017.9 10501340..10472568, complement		1665-1674 gggtgtatg 1866-1875 acgtgtggaa	Not Reported
NBN	#251260	NC_000008.9 91066075..91014740, complement NC_000002.10 152299235..152050099,	91066075..91068075	None	No Cleft palate
NEB	#256030	complement	152299235..152301235	1385-1394 aggtgctaaa	No Cleft palate
NIPBL	#122470	NG_006987.1 4877..193937	2877..4877	192-201 tggtgttg 254-263 acgtgttta	No Cleft palate
NKX2-6	#217095	NC_000008.9 23621070..23615909, complement	23621070..23623070	None	No Cleft palate
OFD1	#311200	NC_000023.9 13662801..13697393	13660801..13662801	228-237 aggtcttaa 1663-1672 aggggtggag	Cleft palate

PAX3	#193500	NC_000002.10 222871944..222772851, complement	222871944..222873944	None	No Cleft palate
PEX7	#215100	NC_000006.10 137185416..137276752	137183416..137185416	26-35 aggtggatg 569-578 aggtgtaaac 1804-1813 agctgtcgta	No Cleft palate
PHF8	#300263	NC_000023.9 54087036..53979838, complement	53979838..53981838	761-770 agttgtggag	Not Reported
POMT1	#236670	NC_000009.10 133368110..133389014	133366110..133368110	None	No Cleft palate
POMT2	#236670	NC_000014.7 76856970..76811052, complement	76811052..76813052		Not Reported
PORCN	#305600	NC_000023.9 48252315..48264146	48250315..48252315	29-38 aggtgtgtg 765-774 aggagtttg	Not Reported
PQBP1	#309500	NC_000023.9 48640139..48645364	48638139..48640139	406-415 aggtctggtg 1264-1273 aggcgtgtta	Not Reported
PROK2	#610628	NG_008275.1 5001..18552	3001..5001	249-258 agctgtgtaa 615-624 acgtgtaaaa 626-635 aggttttaa 669-678 agctgtttaa	No Cleft palate
PROKR2	#244200	NG_008132.1 5001..17330	3001..5001	834-843 aggtgtgctg 895-904 aggtgtggg 1406-1415 aggagtggtg	No Cleft palate
PTCH1	#109400	NG_007664.1 5001..78984	3001..5001	None	No Cleft palate
PTCH2	#109400	NC_000001.9 45081203..45060674, complement	45081203..45083203	166-175 agttgttatg 226-235 aggtttcaaa	No Cleft palate
PTEN	#276950	NG_007466.1 5001..110337	3001..5001	None	No Cleft palate
PTPN11	#151100	NG_007459.1 5000..96181	3000..5000	None	No Cleft palate
PVRL1	#225060	NC_000011.8 119104645..119014018, complement	119104645..119106645	None	No Cleft palate
RAI1	#182290	NG_007101.2 5001..134981	3001..5001	None	No Cleft palate
RAPSN	#208150	NG_008312.1 5001..16416	3001..5001	355-364 aggcgtggtg 977-986 aggtggggaa 1242-1251 tggtgtaaaa 1678-1687 gggtgtcttg 1862-1871 aggtgcctaa	No Cleft palate

				1893-1902 aggtggcaag	
RECQL4	#268400	NC_000008.9 145714008..145707479, complement	145714008..145716008	555-564 aggagtcaag	Cleft palate
ROR2	#268310	NG_008089.1 5001..232561	3001..5001	577-586 gggtgtggaa	Cleft palate
RPGRIPL1	#611561	NC_000016.8 52295272..52191319, complement	52295272..52297272	277-286 aggtgtattg	Cleft palate
RPL5	#612561	NG_011779.1 5000..14887	3000..5000	170-179 aggtgtgaac 1124-1133 aggttttaa 1220-1229 aagtgttta	Not Reported
RPL11	#612562	NG_011741.1 5000..9621	3000..5000	371-380 gggtgtggtg	Not Reported
RPS17	#612527	NG_009890.1 4780..8484	2780..4780	1029-1038 aggtgcgttg 1275-1284 aggtgtggg	Not Reported
RPS7	#612563	NG_011744.1 5000..10656	3000..5000	None	Not Reported
RPS19	#105650	NG_007080.2 5000..16496	3000..5000		No Cleft palate
RUNX2	#119600	NG_008020.1 5001..227766	3001..5001	1970-1979 aggagttta	Cleft palate
SATB2	#119540	NC_000002.10 200033446..199842469, complement	200033446-200033446	1617-1626 aggggttaa	Cleft palate
SC5DL	#607330	NC_000011.8 120668598..120689329	120666598..120668598	1031-1040 atgtgttaaa	Cleft palate (Scd5)
SEMA3E	#214800	NC_000007.12 83116260..82831158, complement	82831158..82851158	600 609 1 aggtctgttag	No Cleft palate
SEPT9	#162100	NC_000017.9 72827197..73008273	72825197..72827197	641-650 aggtccaaa 1435-1444 aagtgtttg	Not Reported
SHH	#162100	NG_007504.1 5001..14410	3001..5001	None	Cleft palate
SHOX	#127300	NC_000023.9 505079..540146	503079..505079	None	Cleft palate (SHOX2)
SHOXY	#127300	NC_000024.8 505079..540146	503079..505079	None	Cleft palate (SHOX2)
SIX3	#157170	NC_000002.10 45022541..45025894	45020541..45022541	None	No Cleft palate
SIX5	#113650	NC_000019.8 50964152..50959884, complement	50964152..50966152	None	No Cleft palate
SLC26A2	#222600	NC_000005.8 149320493..149347156	149318493..149320493	534-543 agctgtgtg 1015-1024 aggtgaggaa 1157-1166 tggtgtttg	No Cleft palate
SLC35D1	#269250	NC_000001.9 67292316..67242439, complement	67292316..67294316	839-848 aggtgtgaac	No Cleft palate
SMS	#309583	NC_000023.9 21868763..21922876	21866763..21868763	1698-1707 aggtgtgtgg	Not Reported
SNX3	%601349	NC_000006.10 108689156..108639410, complement	108689156..108691156	520-529 gggtgtcatg	No Cleft palate
SOX2	#206900	NC_000003.10 182912416..182914918	182910416..182912416	1350-1359 aggttttgag	No Cleft palate

SOX9	#114290	NC_000017.9 67628756..67634156	67626756..67628756	121-130 aggtgtctga 915-924 aggtgttcg	Cleft palate
SPINT2	#270420	NG_013372.1 5000..33156	3000..5000	1325-1334 aggtgtcttc	No Cleft palate
STRA6	#601186	NC_000015.8 72282245..72258860, complement NC_000002.10 202811567..202779148,	72282245..72284245	1260-1269 agctgtgaa	No Cleft palate
SUMO1	%119530	complement	202811567..202813567	952-961 aggttggta	Cleft palate
TBX1	#188400	NC_000022.9 18124226..18151116	18122226..18124226	1375-1384 aggtgctgtg 1985-1994 ggggtgtggag	Cleft palate
TBX15	#260660	NG_013361.1 5000..111513	3000..5000	556-565 ggggtgtggaa 1917-1926 aggttggag	No Cleft palate
TBX22	#303400	NC_000023.9 79156911..79173924	79154911..79156911	991-1000 aggggttta 1044-1053 cggtgtggta 1094-1103 aggtatggtg 1461-1470 aggtgtaaag	Cleft palate
TCOF1	#154500	NC_000005.9 149737201..149779870	149735201..149737201	None	Cleft palate
TFAP2A	#113620	NG_016151.1 5000..27881	3000-5000	456-465 aggagttaaa	Cleft palate (Tcfap2a)
TGFBR1	#609192	NC_000009.11 101867411..101916473	101865411..101867411	935 ggggtgtggtg	No Cleft palate
TGFBR2	#610380	NC_000003.11 30647993..30735633	30645993..30647993	96-105 agctgttaaa	No Cleft palate
TGIF1	#142946	NG_007447.1 5000..51337	3000..5000	None	No Cleft palate
TNNI2	#601680	NG_011621.1 5000..7677	3000..5000	1637-1646 aggtgtggat	Not Reported
TNNT3	#601680	NG_013085.1 5000..24137	3000..5000	631 aggcgtggaa 853 aggggtggaa 1519 cggtgtctag	No Cleft palate
TP63	#129400	NG_007550.1 5000..270852	3000..5000	1614-1628 aagtgttatg	Cleft palate (Trp63)
TRPS1	#190350	NG_012383.1 5000..265504	3000..5000	1433-1442 aggtgtccta	No Cleft palate
TWIST1	#101400	NG_008114 5000..7204	3000..5000	None	Cleft palate
WNT3	#273395	NC_000017.10 44896081..44841685	44896081..44898081	None	No Cleft palate
WNT7A	#276820	NC_000003.11 13921617..13860081	13923617..13921617	None	No Cleft palate
ZEB2	#235730	NW_001838859 4419110..4551938	4417110..4419110	284-293 aggtgagtaa 1057-1066 agatgtggaa 1410-1419 aggtgttagag	Cleft palate (with zeb1)

**APPENDIX 6: LUCIFERASE ASSAY REPORT FROM FLUOROSCAN ASCENT FL  
LUMINOMETER**

**The following reports contain the read-outs from the Fluroskan Ascent FL Luminometer for each of the separate transfection experiments. The first three reports are from the three independent transfection experiments in HeLa cells and the final three reports shown the data from the 293T cell transfections.**

**No T – is from the cells transfected with the pCR3.1\_Null plasmid**

**26.7 - is from cells transfected with 26.7ng of pCR3.1\_TBX22**

**53.4 – is from the cells transfected with 53.4ng of pCR31.TBX22**

**All of these cells were co-transfected with 50ng M\_Prom plasmid and 50ng of the pGL4.17[*luc2*/Neo] plasmid.**

**Ctrl is a blank, untransfected cell.**













## Appendix 7: Luciferase Assay Report from Fluoroskan Ascent FL luminometer from original pilot data.

Ascent Software

Expt1.XLS

28/04/2008 10:20

## Promega Dual Luciferase Assay

### Report Sheet

Plate template: Thermo Labsystems Microtiter 96 Plate  
User name: JL  
Comment: Promega DLR assay using Thermo Labsystems Microtiter 96 Plate  
LAR II Reagent in Dispenser 1  
Stop&Glo Reagent in Dispenser 2

### Layout Map:

### Raw Data (BL1D):

### Luciferase Assay

### Raw Data (BLU):

### Renilla Assay

### Normalized results:

### Luciferase/Renilla Ratio

## Promega Dual Luciferase Assay

## Report Sheet

Plate template: Thermo Labsystems Microtiter 96 Plate  
User name: JL  
Comment: Promega DLR assay using Thermo Labsystems Microtiter 96 Plate  
LAR II Reagent in Dispenser 1  
Stop&Glo Reagent in Dispenser 2

### Layout Map:

**Raw Data (RLU):**

## Luciferase Assay

### Raw Data (BLU):

## Renilla Assay

### Normalized results:

### Luciferase/Renilla Ratio

## Promega Dual Luciferase Assay

## Report Sheet

Plate template: Thermo Labsystems Microtiter 96 Plate  
User name: JL  
Comment: Promega DLR assay using Thermo Lasbsystems Microtiter 96 Plate  
LAR II Reagent in Dispenser 1  
Stop&Glo Reagent in Dispenser 2

### Layout Map:

### Raw Data (RLU):

## Luciferase Assay

#### Raw Data (RLU):

## Renilla Assay

### Normalized results:

### Luciferase/Renilla Ratio