MURINE TRANSGENE INSERTIONAL MUTATION INVOLVING *RUNX1T1* AND *GM11823* GENES AND THEIR CONTRIBUTION TO CLEFT PALATE AND RIB ANOMALIES

Melanie Fawaz Alazzam

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Dentistry (Oral Biology).

Chapel Hill
2014

Approved by:

Eric T. Everett

Sylvia Frazier-Bowers

P. Emile Rossouw

John van Aalst

John Timothy Wright

ABSTRACT

Melanie Fawaz Alazzam: Murine Transgene Insertional Mutation Involving *Runx1t1* and *Gm11823* Genes and their Contribution to Cleft Palate and Rib Anomalies
(Under the direction of Eric Everett)

Objective: Cleft Palate (CP) is a common birth defect in humans occurring in 6.35/10,000 live births and it has been repeatedly shown that animal models are useful in dissecting molecular etiologies of CP. The OVE1328 mouse line develops CP as a consequence of transgene (Tg) insertion mutagenesis. Preliminary data shows Tg complex integration at chromosome 4 band A2. The goal of this work was to characterize the mutation in the OVE1328 transgenic mouse line. Methods: Genotyping microarrays and RNA-seq were used to identify the Tg insertion site. The insertion site was further validated by conventional PCR. Histology and skeletal staining were used to phenotype CP (OVE1328 (Tg/Tg)) and Wt embryos. Results: In OVE1328 embryos the transgene disrupts *Runx1t1* gene at intron 12 (13,876,840 bp). The integration is associated with a deletion mutation of part of intron 12 and whole exons 13 and 14 of *Runx1t1*. *Runx1t1* (also known as *Cbfa2t1h, Eto, Mtg8*) is mainly studied for its role in acute myelogenous leukemia as a fusion gene in humans and gut development in mice. The Tg insertion disruption extends to the *Gm11823* gene which encodes a long non coding RNA. Little is known regarding the normal function and expression of *Gm11823*. The Tg integrates at intron 2 (13,949,713 bp) of *Gm11823* and induces the ectopic expression of an altered message. Homozygous disruption of both genes is associated with CP (100% of OVE1328 (Tg/Tg) embryos) and rib anomalies (supernumerary ribs, 86% of OVE1328 (Tg/Tg) embryos.

Furthermore, OVE1328 (Tg/Tg) embryos are ~13% smaller by weight than Wt and OVE1328 (Tg/+) littermates indicting growth delay. Supported by NIH/NIDCR DE015180.

To my lovely parents (Fawaz and Diana), who dedicated themselves to make me happy

To my sweethearts (Mohammad and Mariam), whose sacrifices made this dream come true

ACKNOWLWDGEMENTS

I would like to thank my mentor Dr. Eric Everett for his continuous support and his mentorship. I am so grateful to have such brilliant advisor with great capabilities. I would like as well to thank my committee members: Dr. Sylvia Frazier-Bowers, Dr. Emile Rossouw, Dr. John van Aalst, and Dr. Tim Wright for their valuable comments and contribution to putting the dissertation in its final outcome. Each of you was an inspiration to me. I would like to acknowledge Dr. Paul Overbeek for his precious gift of the OVE1328 transgenic line. I would like to thank Dr. Ceib Phillips for her help with some of the statistical analysis done. My acknowledgement also goes to National Institute of Dental and Craniofacial Research (NIDCR) and Jordan University of Science and Technology for supporting this project.

My deepest gratitude goes to my soul mates baba (Fawaz), mama (Diana), my husband (Mohammad) and (Mariam), sister (Nissreen) and brother (Adam)! Thank you so much for the tremendous unconditional love you surrounded me with. Thank you for believing in my abilities. Thank you for all the prayers you made, for the sacrifices you had to do to help me achieve my dream! Thanks for all the happiness you filled my heart with! You all are an inspiration to me and I would not have been able to achieve this without the help given by each of you!

My thanks go to my friends and colleagues at UNC, thank you for making me feel at home. I really had great moments with you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBRIVIATIONS AND SYMBOLS

µA	Microamperes
µm	Micrometers
AML	Acute myelogenous leukemia
bp	Base pair
cDNA	Complementary DNA
CL/P	Cleft lip with/without cleft palate
CLO	Cleft lip only
CLP	Cleft lip and palate
CO2	Carbon dioxide
CP	Cleft palate
DIG	Digoxigenin
E	Embryonic day
EDTA	Ethylenediaminetetraacetic acid
FISH	Fluorescence *in situ* hybridization
gDNA	Genomic DNA
H&E	Hematoxylin and eosin
IPD	Interpupillary distance

Kb	Kilobase
Kd	Kilodaltons
kV	Kilovolts
LncRNA	Long non coding RNA
Lt	Left
MgCl2	Magnesium chloride
Min	Minutes
mM	Millimolar
ms	Milliseconds
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NAT	Natural antisense transcripts
nmole	Nanomoles
Nt	nucleotide
PCR	Polymerase chain reaction
ROI	Region of interest
Rt	Right
SDS	Sodium Dodecyl Sulfate

SNPs	Single nucleotide polymorphisms
SOD	Snout occiput distance
SSC	Saline sodium citrate
SW	Snout width
Tg	Transgene
TMD	Tissue mineral density
TUNEL	Terminal deoxynucleotidyl transferase dUTP nick end labeling
Tyr	Tyrosinase
V	Volts
Wt	Wild type

CHAPTER 1: REVIEW OF THE LITRETURE

1.1 Development of the Oral Cavity; an Overview

The formation of the orofacial region in the developing embryo is demarcated by the appearance of the oropharyngeal membrane and subsequently the primary mouth (also known as the stomodeum). The **oropharyngeal membrane** seen at **stage 11 Carnegie, 24 days of gestation in human, and E9 mouse embryo**, is formed from a region that is cranial to the notochord where the ectoderm and endoderm fuse together in the trilaminar embryo. The ectoderm of the membrane will form the mucosal lining of the future oral cavity and the endoderm will form the future lining of the pharyngeal mucosa. The oropharyngeal membrane, a temporary weak membrane, separates the primary mouth from the foregut. The membrane becomes perforated at this stage where the oral cavity and the pharyngeal space become connected.[1-3]

The primary mouth is limited by 5 developing processes (or prominences). The processes are composed of a mesenchyme covered by an ectoderm. The mesenchyme is populated by cranial neural crest cells (CNC) (originating from the crest of the folding neural tube) which start migrating at **stage 10** towards their future destination. The CNC cells contribute to the formation of the facial skeleton. In addition to CNC cells, cells of mesodermal origin populate the mesenchyme which will contribute to the formation of facial musculature. These processes are; 1 median frontonasal process (rostrally), 2 maxillary processes (laterally) and 2 mandibular

processes (caudally) (the maxillary and mandibular processes develop from the first branchial arch).[2-4]

1.1.1 Development of the Upper Lip

At **stage 13-15 Carnegie**, **fourth to fifth week of gestation in human**, the formation of the paired telencephalic vesicles (future cerebral hemispheres) from the forebrain results in the widening of the frontonasal process with the formation of the median groove. Meanwhile, the 2 mandibular processes grow and merge together in a caudal to rostral aspect. The fusion of those processes results in the formation of lower jaw and the low lip.[2,4]

At **stage 14 Carnegie**, **around 32 days of gestation in human, and E10 mouse embryo**, two ectodermal thickenings form at the inferior lateral corners of the frontonasal processes. Those thickenings will form the nasal placodes. The placodes will develop into nasal pits due to the growth and bulging of the frontonasal processes and the formation of a horseshoe like medial and lateral nasal processes. The medial and lateral nasal processes are separated and the nasal pit is in continuity with the stomodeum. Furthermore, the median groove is seen in between the two medial nasal processes. At this stage the maxillary processes become discernable.[2,4]

At **stage 15 Carnegie, 35 days of gestation in human, and E10.5 mouse embryo,** further growth of the maxillary process pushes the medial nasal processes medially and results in the wedging of the lateral nasal process between the maxillary and medial nasal processes. By this stage, nasal pits are seen as distally pointed slits.[2,4]

At **stage 16 Carnegie, 38 days of gestation in human, and E11 mouse embryo,** rapid growth of the maxillary processes and the medial nasal processes results in two fusion events: 1.

Fusion of the medial nasal process and the lateral nasal process 2. Fusion of the medial nasal process and the maxillary process. These fusion events continue through **stage 16-18 Carnegie, beginning in the seventh week of human gestation, and to E11.5 to E12 in a mouse embryo.** The aforementioned fusion events require the involvement of the epithelial cells in several cellular processes represented by apoptosis, formation of filopodia and epithelial mesenchymal transformation (EMT). However, for the median groove, growth and confluence of the medial nasal and maxillary processes will fill and smoothen the median groove in between the medial nasal processes.[2,4,5]

At **Stage 19, around 48 days of gestation in human, and in E12.5 mouse embryo,** the upper lip development is complete. The distal part of the medial nasal processes, the intermaxillary segment, forms the central part of the lip. Further growth of the intermaxillary segment into the oral cavity results in the formation of the primary palate which later fuses with the secondary palate. [2]

1.1.2 Development of the Secondary Palate (Palatogenesis)

At 6 weeks of gestation in human and E11 mouse embryo, palatogenesis starts as an oral outgrowth of the developing maxillary processes known as palatal shelf primordia. Palatal shelf primordia, as other facial primordia, is composed of a mesenchyme (mainly of cranial neural crest (CNC) origin) with an overlying epithelium of an ectodermal origin. [6-8]

In the same week (6 weeks of gestation) in human and at E12.5-14 mouse embryo, the palatal shelves will grow downward (however, it should be noted that there is growth in the antero-posterior and medio-lateral aspects as well) around the developing tongue.[6-8]

At 7-8 weeks of gestation in human and E14.5-15 in mouse embryo, palatal shelves will undergo rapid morphological changes that will eventually result in reorienting the growth pattern of the shelves from vertical to horizontal plane. It was proposed that the elevation of palatal shelves is due to remodeling in which a protrusion of the medial wall and a regression in the ventral end of the shelves results in their elevation.[6] Recent evidence has indeed revealed that cells expressing medial edge epithelium (MEE) markers were detected along the medial side of the developing palatal shelves.[6,9] Furthermore, a recent study supported the tissue remodeling hypothesis for certain regions of the elevating shelves (mid-posterior region) but suggested that there might be other mechanisms involved in the elevation of the anterior and most posterior end of the shelves.[9] It was noticed that the elevation starts at the mid-posterior region of the palatal shelves and then moves to the anterior region, and once elevated, the elevation of the most posterior end region of the shelves is achieved.[10] The elevation process is associated with other morphological changes in the craniofacial region where the head is extended and there is an increase in the vertical dimension of the head with the downward positioning of the tongue.[11]

By this time the palatal shelves are in the horizontal plane. The two layered epithelium on the medial side of the pre-fusion palatal shelves is known as the medial edge epithelium (MEE) and is composed of a basal layer of cuboidal cells lying on a basal lamina and an outer layer of flat cells facing the amniotic fluid, known as the periderm.[7,12]It is agreed that the fusion of the basal layer of both shelves contributes to the formation of the midline epithelial seam (MES).[12] On the other hand, the fate of the periderm layer is still not ascertained. Several mechanisms were suggested: the periderm layer undergoes desquamation (peeling off), can be trapped in the MES, and undergo apoptosis or migrates to the oral and nasal epithelial triangles.[7,12,13] Once the MES forms it will undergo degeneration to allow for the confluence of the mesenchyme in order for the fusion process to be completed.[6,7]However, there is still inconclusive evidence of how the MES is removed. There has been evidence supporting different fates of the MES. The MES can undergo apoptosis as some studies demonstrated that many MES cells were terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) positive and Caspase 3 positive.[6,14] Other studies showed the MES cells in the middle palatal region had strong TUNEL positive staining while in the anterior region (where the primary and the secondary palate fuse in the anterior palate region) had few TUNEL positive cells.[15] In addition to apoptosis, it was proposed that MES cells can also undergo migration.[7,16] or epithelial mesenchymal transformation. However, the available evidence regarding the latter fate is still controversial.[6]

By 8 weeks of gestation in human and E15.5 in mouse embryo, the MES has disappeared and the palatal mesenchyme is confluent with the initiation of palatal bone formation by intramembranous ossification. The adult hard palate (of secondary palate origin) is formed of the palatal process of the maxilla and the horizontal plate of the palatine.[6,17] In the developing embryo bone formation of the palatal process of the maxilla starts at a new

ossification center which is initially separate from the maxillary ossification center. The horizontal plate of palatine bone formation occurs as an extension of the osteogenic front of the palatine[6,18]. **By 10 weeks of gestation in human and E16.5 in mice,** palate formation has been completed and successful separation of the developing oral and nasal cavities takes place.[8,19,20]

Though the formation of the oral cavity seems straightforward, development and morphogenesis of oral structures depend on successful spatiotemporal regulation of different cellular processes. A disruption or alteration at any of the stages explained above will prevent the normal fusion of the lip and/or the palate resulting in a cleft.

1.1.3 Human vs Mouse; A Comparison of Lip and Palate Structures

In spite of the prominent external differences between humans and mice, lip and palate structures demonstrate many similarities. The next section will focus on comparing humans and mice in 3 aspects: lip and palate development, lip and palate anatomy and lip and palate cleft types.

Lip and Palate Development

During early craniofacial development, 32 day human and E10 mouse embryos, not only are human and mouse embryos similar in their shape and size but the formation of cleft and palate structures is basically the same.[20] However, lip and palate formation in humans (lip: ~4 weeks and palate: ~4-6 weeks) requires more time to complete when compared to that of mice (lip: ~2.5 days and palate: ~4-4.5 days).[20] Other developmental differences are seen during palatal shelf elevation. There is evidence that the site of palatal shelf elevation might differ depending on the species studied.[10] For example, in rats and humans, palatal shelves start to elevate in an anterior to posterior manner.[10,21] On the other hand, recent histomorphological

analysis of palatal shelf elevation in mice has shown that the shelves start to elevate at the mid-posterior region and then move to the anterior region. Once this takes place the most posterior end of the palatal shelves (presumptive soft palate) elevates.[10]

In addition to palatal shelf elevation, the fusion process might also differ between humans and mice. Evidence demonstrates that palatal shelves fuse first anteriorly (at the incisive foramen region) and then proceed posteriorly.[21] Following fusion, the epithelial nests, or remnants formed in the midline (due to the generation of the MES), remain till birth after which they completely degenerate.[21,22] On the other hand, there is conflicting evidence for the site of palatal shelf fusion in mice. Some evidence supports that fusion takes place anteriorly (at the incisive foramen region) and then proceeds posteriorly in a zipper like mechanism as in humans.[23-25] Other evidence supports the notion that the fusion process posteriorly can occur independently of the fusion process anteriorly as seen in *Shox2* knock out mice.[26] Furthermore, the epithelial seam disappears completely and no epithelial remnants persist until birth.[22]

Lip and Palate Anatomy (Selected Differences)

Despite the fact that lip embryogenesis in mice and humans is similar at early stages, the gross anatomy is clearly different. In contrast to our lips, close examination of the lips in mice reveal that their upper lips are normally fissured, exposing the upper incisors. Furthermore the lips are devoid of the vermillion border seen in humans. This border separates the skin from the oral mucosa and is rich in vascular supply. [27]

As far as the anatomy of the hard palate, both species are anatomically similar. The anterior two thirds of the hard palate are composed of the palatal process of the maxilla and the posterior one third is formed by the horizontal plate of the palatine bone. The two bones are

joined by the transverse palatine suture. In the sagittal plane the bones of the hard palate are joined by the mid palatal and the interpalatine sutures, respectively. However, the midpalatal suture in mice is straight while in humans is interdigitated.[27,28]

Another difference is seen in the soft palate- in mice no uvula is seen at the midline whereas in humans, the uvula forms at the midline of the soft palate.[27]

Lip and Palate Cleft Types

Oral clefts in humans are diverse and present with varying degrees of severity: unilateral or bilateral cleft lip only (CLO), unilateral or bilateral cleft lip and palate (CL/P) or cleft palate only (CPO). In CL/P, the cleft can involve the alveolar ridge and the primary palate or can be more extensive, involving the alveolar ridge, primary palate and secondary palate as well. In the case CPO, the cleft can involve the hard palate or the soft palate (bifid uvula) or involve both the hard and soft palate.[20,29-32] Closer examination of the type of clefts seen in mice reveals that mice do have similar phenotypes to humans. For example, CPO and CL/P are seen in mice. However, in contrast to humans, CPO is more commonly seen in mice than CL/P.[20] Furthermore, CLO is extremely rare in mice. Only one mouse model was found to have CLO where the cleft did not involve the primary palate and there was no evidence that the alveolus was involved. Though the cleft did not involve the primary palate, a cleft of the secondary palate was seen[33]. Other examples of CL in mice are more commonly seen involving the lip and both the primary and secondary palates[20]. Clefts of both the lip and the primary palate are seen less commonly.[34,35] It is crucial to distinguish the fact that mouse models with midline clefts of the lip and the midfacial region[36,37] are considered a different entity than the cleft lip mouse models discussed here[38]. The pathogenesis of a midline cleft lip differs from what was previously discussed regarding cleft lip. Midline cleft lip is seen when both medial nasal processes fail to fuse.[38,39]

1.1.4 Lip and Palate Development: Interspecies Variation

The basic steps involved in palatal shelf formation vary across different species. In some species CP is considered as being the norm such as in birds and some reptiles. Tables (1.1 and 1.2) address some of the known variations. The content of the tables 1.1 and 1.2 were adapted from Ferguson et al.[40]

Vertebrate	Hard palate	Embryonic origin
Mammals	Primary and secondary palate	Medial nasal processes (frontonasal origin) & maxillary processes
Birds	Primary and secondary palate	Medial nasal processes & maxillary processes
Amphibians and some reptiles (certain species of snakes & turtiles)	Posterior growth of primary palate	Frontonasal origin
Reptiles (lizards)	Primary and secondary palate	Maxillary processes
Reptiles	Primary and secondary palate	Maxillary processes
Crocodilians (alligators & crocodiles)	Primary and secondary palate	Maxillary processes

Table 1.1 Interspecies variation in the embryonic origin of the secondary palate. Content adapted from Ferguson et al.[40]

Species	Vertical growth	Horizontal growth	MEE	Keratinization of MEE	MES formation	Natural CP
Mammals	Yes	Yes	Yes	No	Yes	No
Birds	No	Yes	No	Yes	No	Yes
Amphibians and some reptiles (certain species of snakes and turtiles)	No	Yes	No	No	No	No
Reptiles (lizards)	No	Yes	No	No	No	Choanal opening
Reptiles	No	Yes	No	Yes	No	Yes
Crocodilians (alligators and crocodiles)	Posterior 1/5th of the shelf, remodels later in development to become horizontal. The most posterior end fuses to the floor of the mouth forming the basihyal valve	Anterior 4/5th of the shelf	Restricted region, the rest of the palatal shelves are fused by migration of MEE to the nasal epithelium and mesenchymal confluence	No	No	No

Table 1.2 Interspecies variation in palatogenesis. Content adapted from Ferguson et al.[40]

1.2 Oral Clefts in Humans

Oral clefts are among the most common congenital anomalies in humans with a prevalence of 1/700 live births.[41-44] Oral clefts include a group of anomalies in which a cleavage of the lip and/or the palate is seen due to the abnormal embryonic development of these structures. Oral clefts can be divided based on the embryological origin of the affected structures into: 1. Cleft lip with/without cleft palate (CL/P) 2. Cleft palate (CP). There is a great variation in the prevalence of CL/P and CP among different geographical regions, races, ethnic groups and socioeconomic status. For example, higher prevalence of CL/P was seen in Latin America and Asia compared with lower prevalence in Israel, South Africa and Southern Europe[44]. On the other hand, higher prevalence rates of CP were seen in Canada and parts of Northern Europe and lower rates were seen in pats of Latin America and South Africa. Furthermore, the prevalence of CL/P was the highest among Asians and Native Americans (1/500), intermediate among Europeans (1:1000) and the lowest among African American populations (1:2500).[30,32,44,45]

Depending on the severity/ extent of the defect, CL/P can be divided into cleft lip only (CLO) and Cleft lip and Palate (CLP) (Fig 1.2). In CLO, the cleft can be complete (extending to the nostril and loss of Simonarts band) or incomplete (in which the nostril is not affected and Simonarts band is seen).[46-48] As CLO, cleft lip and palate (CLP) has variable manifestations based on the severity: cleft lip combined with clefting of the alveolar ridge, cleft lip combined with clefting of both alveolar ridge and primary palate, cleft lip combined with clefting of the alveolar ridge, primary palate and the secondary palate (Fig 1.1). CP shows such variations in severity in which the cleft affects both the hard and soft palate or in which the hard palate only or the soft palate only are affected (Fig 1.1). CL/P is further categorized, based on the side being affected, into unilateral and bilateral clefts. Unilateral clefts are seen in 90% of CLO with the left

side being mostly affected (2/3 of the cases). However, a greater proportion of CLP patients (30.2%) show a bilateral involvement of the lips compared to the bilateral involvement seen in CLO patients (10.3%).[20,41,44,49] CL/P is seen more in males compared to females at a 2:1 ratio. In contrast, CP tends to be more frequently seen in females compared to males.[44]

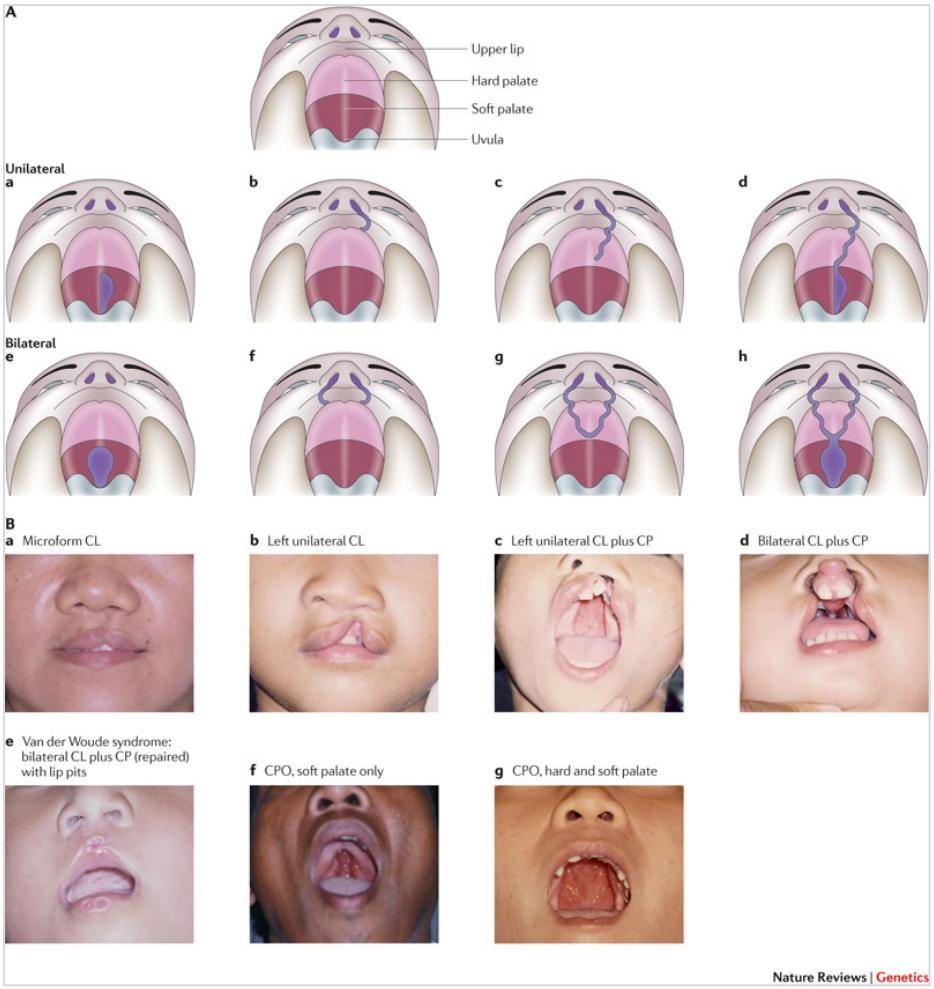


Figure 1.1. Types of oral clefts in humans. A. a schematic representation of oral clefts in humans. B. Clinical cases representing variable severity of oral clefts in humans. *Figure reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Genetics (Dixon, M. J., Marazita, M. L., Beaty, T. H. & Murray, J. C. Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet **12**, 167-178, doi:10.1038/nrg2933 (2011), copyright 2011.*

CL/P and CP can be associated with other anomalies and can be classified accordingly. The greatest prevalence of associated anomalies are seen with CP, followed by CLP, and the least prevalence is seen in CL. Congenital cardiac defects, limb and vertebral anomalies were the anomalies often seen in CL/P and CP.[44] If no other anomalies are associated with CL/P and CP, then the cleft is known as isolated-non syndromic CL/P or CP.[30,45,50] The frequency of isolated CL/P was around (76.8%). The rest of CL/P cases were either associated with other anomalies (15.9%) or were recognized as part of a syndrome (7.3%) in such case known as syndromic CL/P. On the other hand, the frequency of isolated-non syndromic CP was (54.8%), while the frequency of CP with other anomalies was (27.2%) and CP with anomalies as part of a recognizable syndrome (syndromic CP) were (18%).[44]

1.3 Common Problems Associated with Oral Clefts

Patients born with oral clefts suffer from different problems and complications. This section comes across the common problems associated with oral clefts in humans. These are: **Feeding problems:** Proper feeding in a newborn requires normal sucking and swallowing mechanisms. A negative intraoral pressure is required to produce efficient sucking. The negative pressure is formed by a lip seal, elevation of the soft palate to close the nasopharynx, and expanding the intraoral cavity through contraction of the tongue or the movement of the mandible. A baby with CL/P will have an inefficient sucking mechanism because this negative intra oral pressure is lost. CL/P baby, depending on the type and affected structures, will not have a lip seal and/or will have abnormal anatomy of the soft palate muscles, the most important ones are levator veli palatani and tensor veli palatani. Because of the abnormal anatomy of the lip and palate the baby can have inefficient negative pressure and sucking, excessive air intake, regurgitation of milk

into the nasal cavity, choking, inadequate milk intake, failure to gain weight, prolonged feeding time and fatigue.[51,52]

Upper respiratory tract infections, recurrent ear infections and hearing loss: the lack of separation between the oral and nasal cavities in CLP and CP patients results in the regurgitation of food and milk into to the nasal cavity and aspiration, inducing upper and lower respiratory tract infections.[53,54] Furthermore, CLP and CP patients suffer continuous or recurrent otitis media with effusion (OME) that can lead to conductive hearing loss. In OME, the fluids build up in the middle ear for 3 months or more results in recurrent ear infections and damage to the ear drum. Patients with CLP and CP are more prone to OME due to A. regurgitation of food and milk into the nasal cavity that induces edema and inflammation of the orifice of the tube causing its blockage. B. the structure of the Eustachian tube due to multiple reasons one of which is the abnormal anatomy of the levator veli palatani and tensor veli palatani.[55,56]

Dental anomalies: CL/P patients are known to have a higher prevalence of dental anomalies compared to non-cleft population.[57-60] Some of the dental anomalies encountered in CL/P patients are missing teeth (with the lateral incisor to be the most common tooth to be absent), supernumerary teeth, microdontia (such as peg shaped lateral incisors, malformed teeth (enamel hypoplasia), taurodontism, dilacerations and others.[57-60] It is evident that the prevalence of those anomalies varies between different studies in different populations. For example, several studies demonstrate that that agenesis of lateral incisors at the cleft side and supernumerary teeth are the most common and the second most common anomalies encountered in CL/P patients, respectively.[57,58] On the other hand, the developmentally missing lateral incisor at the cleft side and taurodontism were the most common and the second most common dental anomalies encountered in a group of CL/P patients.[61] Furthermore, there is evidence that the type of

anomalies seen can vary among different types of oral clefts.[57] For example, Ackam et al found that dilaceration, taurodontism and *dens evaginatus* was only seen in unilateral left cleft lip and palate patients (UCLP).[57]

Dental malocclusion: CL/P patients suffer abnormal growth of the craniofacial structures. However, there are few differences seen in the type of malocclusion among operated and non-operated CP patients. In general, operated CL/P patients have concave faces, midfacial deficiency and Class III skeletal relation with varying degrees in severity (mild to severe). The antero-posterior, transverse and vertical dimensions of the maxilla are deficient. Dentally, anterior and posterior cross bites are seen as well[50,62]. It is believed that the maxillary deficiency in operated cleft patients is due to: the tissue deficiency associated with cleft itself, the scarring associated with the corrective surgery, or the inherited genetic makeup of the patient.[62] On the other hand, the maxilla in non-operated cleft patients (specifically UCLP) is normal or prognathic. The prognathism of the maxilla is seen on the non-cleft side therefore resulting in a hemifacial maxillary prognathism with the cleft side being retruded. Posterior molar relation tend to be normal with less frequency of posterior cross bites.[62] In addition, there is evidence that the severity of malocclusion increases with the severity of the cleft.[63] For example, in one study the frequency of Class III malocclusion with CLP (5.5 times) CP (3.5 times) higher than that seen in CL patients.[64]

Speech problems: the anatomic abnormalities affecting the velopharyngeal function, the chronic/recurrent ear infections and hearing loss, dental abnormalities and malocclusion in CL/P patients contribute to a wide variety of speech problems. Some of the speech problems encountered in CL/P patients are resonance disorders (hyper and/or hyponasality), dentalizing alveolars (s,t).[65,66]

1.4 Etiology of CP

It is agreed that the etiology of oral clefts (CL/P and CP) is multifactorial involving genetic and environmental factors. The etiology is even more complex as both genetic and environmental factors interact to modify the risk for developing oral clefts.[20,30,45,67,68] Furthermore, CL/P and CP are not only developmentally distinct anomalies but also etiologically distinct.[30,68] This belief originates from the observation that CL/P and CP do not segregate in the same family indicating that each of them has distinct etiologic factors. However, exceptions are seen where CL/P and CP are segregating in the same family with the same mutation contributing to CL/P or CP (known as mixed clefting).[30,43,68] Mixed clefting is seen in cases involving mutations of *IRF6, Msx1,* and *FGFR1*.[30,68] Since CP and the genetic factors regulating palatogenesis are within the scope of interest of this thesis, I will be focusing on the contribution of genetic factors to CP. I will touch on some of the environmental factors associated with increased risk of CP in humans.

1.4.1. Genetic Factors

It was observed that the risk of CP is 56 times higher in the first degree relatives of a patient with CP compared to individuals with no family history of CP.[41] There is a higher concordance rate in the CP phenotype in monozygotic twins (33%) compared with dizygotic twins (7%).[69] These observations indicate that the genetic make-up of humans contributes greatly, but not solely, to the CP phenotype in humans.[41,69]

CP in humans is seen with other developmental or cognitive defects where it is known as syndromic CP.[30,67] More than 275 syndromes have oral clefting as a primary feature and are due to single gene mutations, chromosomal abnormalities or teratogens. About 75% of these

yndromes have an identified genetic component.[67,70] Some examples of human syndromes with CP are presented in (Table 1.3) which was adapted and modified from Gritli- Linde[20]

Gene	Syndromic oral clefting, oral cleft type	Gene product	Non-syndromic	Mouse model
DHCR	Smith–Lemli–Opitz syndrome, CL/P or CPO	3-β-hydroxysterol-Delta 7-reductase	No	Yes, CPO
FGFR1	Kallmann syndrome (KAL2) , CL/P or CPO,	TK receptor	Yes	Yes, CPO
FGFR2	Apert syndrome , CPO (in 44% of cases)	TK receptor	Yes	Yes, no clefting
FOXE1	Bamforth–Lazarus syndrome , CPO	TF	Yes	Yes, CPO
IRF6	Van der Woude (VWS), CL/P or CPO. Popliteal Pterygium syndromes (PPS) CL/P,	TF	Yes	Yes, CPO
P63	Autosomal dominant Ectrodactyly–ectodermal dysplasia–cleft syndrome (EEC) syndrome, CL/P. Hay-Wells syndrome, CL/P. limb-mammary syndrome, CPO.	TF	Yes	Yes, truncated palate
TCOF1	Treacher–Collins syndrome, CPO	Treacle (nucleolar protein)	No	Yes, CPO
TFAP2A	Branchio- oculo-facial syndrome, CL and/or CP	TF	No	Yes, midline clefting

Table 1.3. Examples of syndromic CP in humans. Note that the mouse model in some examples has either no cleft phenotype or different phenotype than human species. The table demonstrates that some syndromic CP genes can contribute to non-syndromic forms of CP. Content adapted from Gritli-Linde.[20]

CP can be seen as an isolated anomaly i.e. with no other structural, cognitive abnormalities where it is known as non-syndromic CP. (Table 1.4) shows genes that are implicated in non-syndromic CP in humans.

As shown in (Table 1.3) and (Table 1.4), the products of the genes implicated in syndromic and non-syndromic CP can be subdivided into 4 categories (Fig 1.2): 1. Signaling proteins and receptors. 2. Transcription factors and nuclear proteins. 3. Cytoplasmic and membrane bound proteins.4. Extracellular matrix proteins.[20,71] An example of each subcategory is given in the following section. The genes that are chosen represent those associated with CP in humans and mice.

Gene	Human (Oral cleft type)	Cytogenetic location	Gene function	Available mouse model, (Oral cleft type)
FGFR1	CL and CP or CP	8p11.23-p11.22	Receptor	Yes, CP
FGFR2	CL and CP or CP	10q26	Receptor	Yes, CP
IRF6	CL or CP	1q32.3-q41	Transcription factor	Yes, CP
MSX1	CL or CP	4p16.2	Transcription factor	Yes, CP
PTCH	CL/P or CP	9q22.3	Receptor	Yes, CL*
SATB2	CL/P or CP	2q33	Transcription factor	Yes, CP
TBX22	CP	Xq21.1	Transcription factor	Yes, SMCP and ~7% CP
TGFβ3	CP	14q24	Signaling protein (growth factor)	Yes, CP
TGFA	CP	2p13	Signaling protein (growth factor)	Yes, no clefting defect
SUMO1	CLP or CP	2q33	cytoplasmic and membrane bound protein	Yes, CLP*
ESR1	CL or CP	6q25.1	Transcription factor (estrogen receptor)	Yes, no clefting defect
PVR	CL/P or CP	19q13.2	Poliovirus receptor	Yes, no clefting defect

Table 1.4 Genes associated with non-syndromic CP in humans. Note that some of those genes are associated with syndromic forms of CP in humans.* PTCH associated mouse model is a conditional knock out. *SUMO1 associated mouse model is a haploinsufficient. Content adapted and updated from Gritli-Linde.[20]

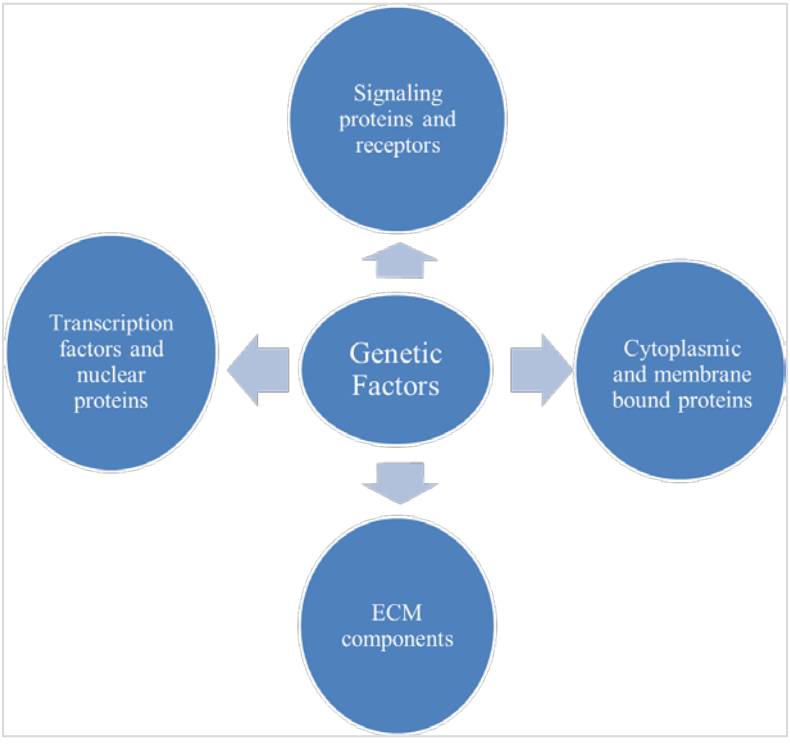


Figure 1.2 Functional classification of genetic factors contributing to oral clefts (including CP) in humans and mice. An example for each sub-category and its contribution to CP in humans and mice is provided in the following section.

Cell Surface Receptors: FGFR1 and FGFR2 Mutations

In humans, loss of function mutations in *FGFR1* are associated with Kallman syndrome in which CP or CLP is one of its features. Gain of function mutation in FGFR1 is associated with craniosynostosis features in humans.[72] Mutations in *FGFR2* specifically (S252W, S252F, and P253R) are seen in most cases of Apert syndrome.[72] It is proposed that syndromes associated genetic mutations could contribute to non-syndromic CP by having hypomophic variants of these genes.[72] Mutations in FGFR1 and FGFR2 were shown to contribute to non-syndromic CP and CLP in humans. [73,74] The role of these genes in human CP can possibly be better understood from *Fgfr2b* (-/-) and *Fgf10 (-/-)* mouse models which demonstrate CP phenotype beside other craniofacial and general abnormalities.[75,76] In these murine models, it appears that the FGF signaling pathway plays an important role in palatal shelf growth. A significant reduction in the proliferation of epithelial (where *Fgfr2* is highly expressed) and mesenchymal cells (where *Fgf10* expression is limited) of palatal shelves contributes to the CP phenotype. Another observation was the reduced expression of *Shh* in the epithelium of both *Fgfr2b (-/-)* and *Fgf10 (-/-)* mice. This observation revealed a possible role for FGF-SHH signaling network in palatal shelf growth. The positive feedback between FGF and SHH signaling is also demonstrated in mesenchymal knockout of SMO. In this case it was shown that *Fgf10* expression was significantly reduced in the mesenchyme.[6]

Transcription Factors and Nuclear Proteins: IRF6 Mutations

Mutations in *IRF6* gene are associated with van der Woude syndrome (VWS) and popliteal pterygium (PPS) syndromes in humans in which CL/P or CP is a major characteristic feature in humans. Mutations in *IRF6* were also found to be associated with non-syndromic forms of CP in humans.[77] The possible role of *Irf6* in palatogenesis was discovered by studying

Irf6 (-/-) and *Irf6* (R84C) knock in mouse models.[6,78] *Irf6* mutant mice showed an abnormally hyperproliferative epithelium which did not differentiate. The associated epithelial defect resulted in abnormal oral adhesions, CP and other associated abnormalities (abnormal limbs and skin).[78] In addition, studies demonstrated that *Irf6* (epithelial expression) is regulated by *p63* transcription factor (epithelial expression). In humans, mutations in *p63* are associated with syndromic CP such as ectrodactyly ectodermal dysplasia-cleft lip/palate syndrome (EEC), ankyloblepharon ectodermal dysplasia clefting (AEC), and nonsyndromic split-hand/foot malformation (SHFM).[79] In Mice, a null mutation for *p63* results in abnormally thin, undifferentiated epidermis and CP. In *p63* (-/-) mice, the expression of *Irf6* in the palatal epithelium is reduced and *p63* is found to bind to an enhancer sequence upstream of *Irf6* and to promote a luciferase reporter expression driven by an *Irf6* enhancer.[6,77]

Nuclear Membrane Bound Proteins: SUMO1

SUMO1 (small ubiquitin-related modifier) is part of the SUMO family (SUMO1-4). The SUMO family are responsible for the reversible posttranslational modification of proteins by SUMOlyation. As in ubiquitination, SUMOlyations requires will require other enzymes that results in the activation of SUMO and its addition to its target protein. In most of the cases a single SUMO is added or sometimes a poly-SUMO chain is added.[20,80] In humans mutations in SUMO1 were found to be associated with both non-syndromic CL/P and CP.[81,82] Interestingly, haploinsufficiency of SUMO1 in a mouse model (*Sumo1*$^{Gt/+}$ heterozygous) resulted in an incompletely penetrant CP or oblique facial cleft (8.7%).[83] Furthermore, SUMO1 is expressed in both the mesenchyme and the epithelium of mouse palatal sheleves.[20,83] There is evidence that several proteins important for palatogenesis are being SUMOylated.[20]

Extracellular Matrix Proteins: Collagens (COL2A1 and COL11A2)

In humans, mutations in collagens (COL2A1 or COL11A2) are associated with Robin sequence. Robin sequence has a combination of 3 out of 4 features: micrognathia, glossoptosis, obstructive apnea, cleft palate. Robin sequence is commonly seen as part of a syndrome (80%) such as Stickler syndrome, velocardiofacial syndrome. Non-syndromic forms of Robin sequence are less frequent (20%).[84]The role of collagens and their relevance to palatogenesis is demonstrated by examining the (*cho/cho*) homozygous mice which carry an autosomal recessive mutation in collagen (*Col11a1*). These mice demonstrate shortened heads and mandibles, U shaped cleft palates, short limbs. The cleft phenotype in these mice is believed to be due to the abnormal skeletal growth leading to a smaller mandible and a higher tongue position which prevents the palatal shelves from adhering. It was shown that cultured palatal shelves of (*cho/cho*) mice were able to adhere and fuse normally (no remnants of MES).[71,85]

1.4.2. Environmental Factors

The finding that the concordance rate of CP and CL/P phenotypes did not reach 100% in monozygotic twins demonstrates that non-genetic factors (such as environmental factors) contribute to such birth defects.[45,68,69] This does not exclude the possibility that such dis-concordance can be due to genetic and cytogenetic and epigenetic differences in one of the twins.[69] However, the finding that the concordance in dizygotic twins is higher than the concordance rate among singleton siblings further supports the role of environmental factors in such birth defects.[69]

Multiple environmental factors are associated with higher risk of CP or had a potential association with CP in the population studied such as smoking,[49,86,87] alcohol (well established

with oral clefts in fetal alcohol syndrome),[49,86] vitamin A,[88] maternal hyperthermia and viral infections during the first trimester[49], anticonvulsant drugs such as diazepam, phenobarbital and phenytoin,[49,86] corticosteroid therapy during pregnancy,[49,86] nutritional deficiencies such vitamin B6 and zinc.[49] Other environmental factors are associated with a lower risk of developing CP or possibly having a protective effect against CP such as folic acid supplements, multivitamin supplements.[49]

Identifying environmental factors contributing to CP helps to identify the metabolic pathways that could be possibly altered and could be contributing to the formation of such defects.[45]

1.4.3. Gene-Environment Interaction

The cross talk between genetic and environmental factors and their effect on oral clefting was demonstrated in humans and mice. For example, evidence suggested that mothers who smoke have an increased risk (2 fold) of having babies with CP.[86,89] However, if the infant had TGFα *Taq*I C2 allele variant and the mother was a smoker during pregnancy (≤10 cig/day) there was an increased risk of (6.16 fold) to have CP.[86,90] Furthermore in the same study, if this specific polymorphism in TGFα was present but the mother was not a smoker the risk for having CP (0.9) fold higher therefore showing how the presence of an environmental factor can modify the risk for developing CP with specific genotype.[90]

Moreover, experiments using mouse models demonstrated such cross talk. For example, upon examining the effect of cortisone on CP frequency in two mouse strains (different genetic backgrounds) A/J strain and C57BL/6, 100% of A/J mice offspring had a CP phenotype while

only 17% of C57BL/6 had CP. The two mouse strains obviously had different responses to the same teratogen under similar experimental conditions.[20,91]

1.5 Types of Cleft Palate Mouse Models

Mouse models are very useful for dissecting the genetic basis of palatogenesis.[20,92] Our knowledge of how genetic factors contribute to oral clefts in humans was greatly advanced by studying CP murine models.[20] Mouse models used to study CP can be divided into two main categories: Spontaneous CP mouse models and Induced CP mouse models (Fig 1.3).[93]

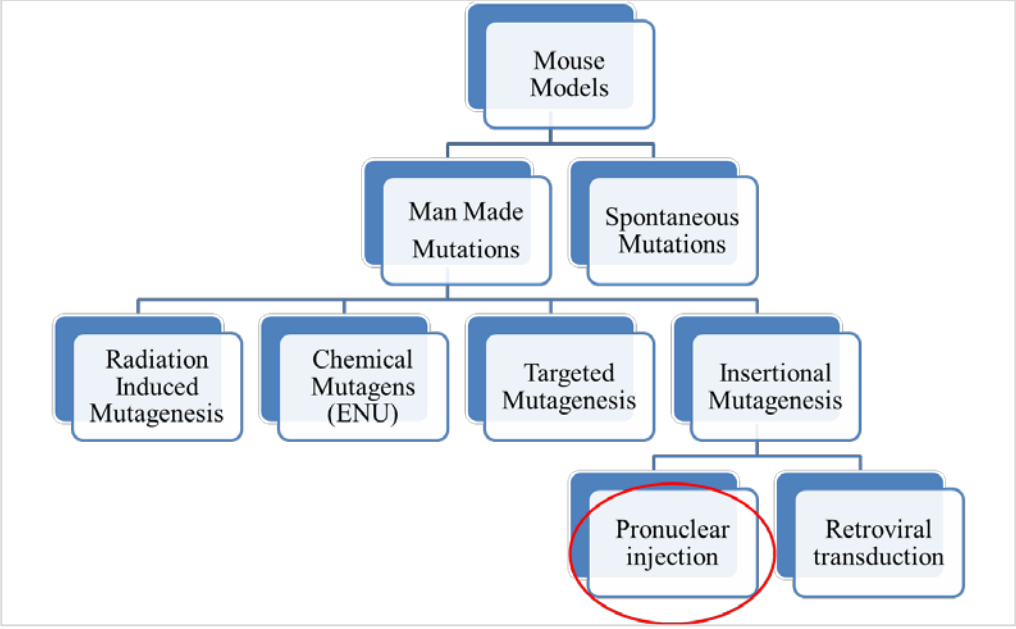


Figure 1.3. Types of CP murine models. CP mutations studied in mice are either spontaneous (no intervention) or Induced (i.e. manmade mutations).

1.5.1 Spontaneous CP Mouse Models

Mouse models in which CP occurs spontaneously (without human intervention) at a stable frequency are known as spontaneous CP mouse model.[94] Examples of CP mouse models with the frequency of CP are: SW/Fr (6%), CF1 (3%), J/Glw (25%). The genetic cause of CP in theses strains has not been studied yet.[50,95]

1.5.2 Induced CP Mouse Models

Induced CP mouse models are produced through human intervention. Induced mutations are further subdivided into 4 main subcategories based on the strategy used: radiation induced mutation, chemical mutagenesis, targeted mutagenesis and insertional mutagenesis.

Ionizing Radiation Induced Mutations

Ionizing radiation is known to induce DNA damage (base pair damage, single and double strand breaks). Mutations induced by ionizing radiation can vary from point mutations to small deletion mutations, and can be more severe and result in major genomic rearrangement.[96-98] An example of ionizing radiation induced CP mouse model is (p4THO-I/ p4THO-I) line which has a radiation induced deletion mutation at pink eyed dilution locus (p). These mice had pink eyes and CP with no other craniofacial abnormality (95%). The mutation was mapped *Gabrb3* locus (β 3 subunit of $GABA_A$ receptor). [20,99]

Chemical Mutagen Induced Mutations

The *N*-ethyl-*N*-nitrosourea (ENU) is the most potent chemical mutagen in mice. ENU is an alkylating agent which can add an alkyl group to different atoms in a DNA molecule such as N-1, N-3, and N-7 groups of adenine; the O^2 and N-3 groups of cytosine; the N-3, O^6, and N-7

groups of guanine; the O^2, N-3, and O^4 groups of thymine; and the phosphate groups of the DNA backbone.[100,101] ENU is known to induce point mutations, rarely small deletion mutations, randomly all around the genome.[100,101] An optimal ENU dose can result in one mutation every 1-1.5 Mb.[100] The point mutations are mostly seen at A-T base pairs resulting in a transversion mutations (A→T or T→A) or transition mutations (A-T→G-C).[100,101] An example of a CP mouse model due to ENU mutagenesis is the cleft secondary palate 1 (csp1) mutant mice due to a point mutation in *Prdm16* gene. The CP is due to the failure of palatal shelves elevation as a consequence to tongue obstruction and micrognathia.[102]

Targeted Gene Mutations

Gene targeting involves modifying a specific genetic locus by the use of homologous recombination between a targeting vector and an endogenous gene in mouse embryonic stem (ES) cells.[103] The genetically modified ES cells are then microinjected into embryos at the blastocyst stage (8 cell stage) and are then transferred to a pseudopregnant female mouse. The resulting chimeric progeny (with the targeted mutation) are then crossed with mice of the same genetic background from which ES cells were derived.[104] If the gene of interest is no more functional, i.e. not expressed or altered then the mouse model is known as a knockout model. On the other hand, if the line has a duplication of a target gene or a tailored mutation (such as point mutations) then the mouse model is known as a knockin model.[104,105] The *TGFβ3*[106] and *Msx1*[107] mouse models and $Irf6^{R84C}$ [108]mouse model are examples of knockout and knockin CP mouse models, respectively.

Ubiquitous gene targeting is sometimes associated with early embryonic lethality, therefore preventing researchers from examining the effects of gene targeting on the developing palatal shelves.[93,104] A solution to this problem was the advent of conditional gene targeting in

which the target gene is modified only in a specific tissue and at a certain developmental stage. An example of a conditional gene targeting system is the *Cre-loxP* system. In this system, mice expressing the Cre recombinase enzyme in a specific tissue or cell line (i.e. driven by tissue specific promoter) are bred to another *loxP* mouse harboring *loxP* sites flanking the gene of interest. Breeding a Cre-mouse and a loxP mouse results in a progeny expressing Cre and having a target gene flanked with *loxP* sites. If the *loxP* sites are oriented in the same direction, the Cre recombinase would result in deleting the DNA segment. If the *loxP* sites are directed in opposite directions, the Cre recombinase would result in inverting the DNA segment flanked by the *loxP* sites.[104] *Tgfβr2* (K14- Cre)[109] and *Shox2* (Wnt1-Cre).[110]

Insertional Mutagenesis

Insertional mutations take place when the integration of an exogenous DNA results in the disruption or alteration of a functional gene; the mutation seen is due to the physical damage at the integration site rather than the expression of the transgene itself.[111,112] Random insertions can be produced by microinjection of the transgene or viral infection (retroviral/lentiviral infection) into a zygote (single cell embryo).[111] One example of an insertional mutant CP mouse model, generated by microinjection technique, is OVE427[113]. A detailed explanation of insertional mutations produced by this microinjection technique is present in chapter 2.

Random insertions can be produced by the gene trap approach in embryonic stem cells which uses a trapping vector that has promoter-less marker/reporter with an upstream splice acceptor site and downstream poly-A signal.[114,115] The trapping vector can integrate into an exon/ intron of a gene.[116] Transcriptional activation of the trapped gene results in a fusion transcript containing part of the original transcript fused to the reporter. The resulting fusion transcript represents an alteration of the original gene and at the same time a readout for the pattern of this

altered expression (due to the expression of the reporter).[114,115] An example of a CP mouse model generated by gene trap approach is Sumo1(Gt/+) mutants.[83]

Another approach for the generation of random insertions (random to a certain extent) is the use of DNA transposons, as the Sleeping Beauty (SB) transposon system.[93,114] The system is composed of the transposase enzyme and the transposon substrate. The transposase identifies inverted repeats that flank the transposable element and then cuts the transposable element. This is followed by pasting the element (inserting) at another site with a TA dinucleotide. Therefore the transposition is done in a "cut-paste".[93,114] An example of a CP mouse model generated by such approach is the Smoc1 mouse model.[117]

1.6 Development of the Vertebral Column and the Ribs

The term axial skeleton refers to head (skull and facial bones), vertebral column, and rib cage (ribs and sternum).[118] This section focuses on the development of the vertebral column and the ribs due to their relevancy to the thesis. During gastrulation, paraxial mesoderm forms on each side of the axial structures (notochord and neural tube) due to the ingression of the epiblastic cells in the developing embryo.[119,120] Paraxial mesoderm undergoes segmentation (is chopped off) to form blocks of epithelial cells known as somites.[120,121] The pace of the segmentation process is unique across different species. For example in humans a new pair of somites is formed every 4-5 hours while in mice every 120 min.[120,122,123] Somitogenesis (the formation of somites) is seen between 25-35 days post conception in humans and E8-E13 in mice.[121,124] After the formation of the somites, 42 pairs in humans and around 65 pairs in mice[125], they undergo differentiation. The cells on the ventromedial aspect of the somite will differentiate into the sclerotome while cells on the dorso-lateral part of the somite will differentiate in to the dermomyotome. The dermomyotome further differentiates into dermatome (produces the dermis of the back) and the myotome (gives rise to the body and limb musculature). The sclerotomal cells delaminate then surround the neural tube and the notochord. These cells will condense and differentiate into chondrocytes that form the cartilaginous skeleton which is then replaced by bony tissue i.e. the vertebral skeleton and the ribs which are formed by endochondral ossification.[120,126]

The development of vertebra does not follow one somite one vertebra relation, but rather each vertebra is formed by a caudal half of one somite and the rostral half of the following somite.[120] This takes place due to another segmentation process known as re-segmentation which happens in the sclerotomal compartment. The sclerotomal cells will rearrange into two cellular

compartments, rostral and caudal cells. The caudal cells from one somite will fuse with the rostral cells from the following somite and both contribute to the formation of the developing vertebra.[121]

The resulting vertebrae differ in their shape and size and can be generally categorized based on their anatomical position to the cervical, thoracic, lumbar, sacral and coccygeal (caudal) vertebrae. The formula in humans is 7 cervical, 12 thoracic, 5 lumbar, 5 sacral and 4 coccygeal. The ribs are attached to the thoracic vertebrae and are classified as true ribs (1-7), false ribs (8-10) and floating ribs (11-12).[127,128] On the other hand, the mouse vertebral formula is 7 cervical, 13 thoracic, 6 lumbar, 4 sacral and 30 caudal vertebrae. As in humans, ribs attach to the thoracic vertebrae whereas only the first 7 ribs are attached to the sternum.[120,129]

1.6.1 Patterning the Axial Skeleton

During somitogenesis, the somites that are formed look similar. However, the somites eventually contribute to the formation of the vertebral column with characteristic morphological features.[130] Genetic factors are known to regulate the segmental identity i.e. the type of vertebra each somite pair will form and its position along the anterior-posterior axis. Genetic factors regulating segmental identity and therefore the patterning of the axial skeleton can be subdivided mainly to transcription factors and signaling molecules.[120]

Transcription Factors

The classical example of a transcription factor regulating the identity of the developing vertebra is the *Hox* genes.[120] These homeo-box containing genes are considered of the major players regulating the patterning of the developing vertebral column.[129] In humans and mice, there are 39 *Hox* genes clustered in 4 groups (A, B, C, D) found on chromosomes (6, 11, 15, 2),

respectively. There are 13 paralogous groups with 2-4 members in each group. [131-133] *Hox* genes generally demonstrate a spatial-temporal colinearity in terms of their expression; the 3' genes in each cluster are expressed earlier and contribute to determining the identity of anterior structures and the 5' genes are expressed later and contribute to identity formation of more posterior structures.[120,133] Little is known about how exactly *Hox* genes regulate the identity of the developing vertebral column.[120,133]

Targeting *Hox* genes in mice revealed the significance of Hox3-Hox11 to the normal development of the axial skeleton. Some single *Hox* genes mutants in mice had phenotypes consistent with their colinear expression such as *Hoxd3* which had defects in C1 and C2 vertebra and *Hoxd11* had defects in patterning the sacral vertebra. On the other hand, other single and multiple mutants demonstrated a phenotype not consistent with the colinear expression pattern of the *Hox* genes. for example, single and multiple mutants of *Hox5, Hox6, Hox7, Hox8, and Hox9* group genes had defects in the rib on T1.[133]

Loss of function mutations of single or multiple mutants are sometimes associated with anterior homeotic transformation. An example of such anterior homeotic transformation is seen in the case of *Hoxb9* mutants (14th rib on first lumbar vertebra) and in *Hoxa7/b7* double mutants (T1 to C7 transformation). Posterior homeotic transformations are also seen due to the loss of function of single mutants such as the extra rib seen on C7 in Hoxa5 and Hoxa6 mutants.[133]

Evidence revealed that there is functional redundancy among paralogous *Hox* genes. It was also demonstrated that a synergistic phenotype is seen when more than one gene in a paralogous group is mutated.[133] For example, single mutants of *Hox4a, Hox4b* and *Hox4d* had incompletely penetrant defects in C1 and C2 vertebrae. However, a combined triple mutant of *Hox4a, Hox4b* and *Hox4d* result in a fully penetrant anterior homeotic transformation in C2-C5.

By now, complete paralogous mutants have been generated for *Hox5, Hox6, Hox7, Hox8, Hox9, Hox10 and Hox11* paralogous gene groups. In these mutants anterior homeotic transformation is seen.[133]

Another example of transcription factors regulating the identity of the developing vertebral column is Caudal type homeobox gene family composed of *Cdx1,2*and *4*.[120] *Cdx1* null mutants show anterior homeotic transformations, not as severe as those with *Hox* mutants, in the cervical and upper thoracic vertebrae. *Cdx2* null mutation is embryonically lethal but heterozygote null mutants demonstrate anterior homeotic transformation in the cervical to thoracic vertebra transition. *Cdx4* null mutants did not affect the axial skeleton. However, a combined null mutants of *Cdx1* and *4* or *Cdx2*(+/-) and 4 resulted in a more severe anterior transformations.[134]

Signaling Molecules and Receptors

An example of cell surface receptors regulating the anterior posterior patterning is *Fgfr1*. Fgfr1 null mutation results in early embryonic lethality. However, hypomorphic *Fgfr1* mutations were associated posterior truncations and anterior homeotic transformations (L1 has an extra pair of ribs). On the other hand, gain of function mutations in *Fgfr1* were associated with posterior homeotic transformations (such as C7 vertebrae with ribs attached, T13 vertebrae lacking their ribs and becoming like L1 vertebrae).[120,135]

1.7 Rib Anomalies (Supernumerary Ribs)

Rib anomalies in humans include supernumerary ribs (SNR), forked ribs, fused ribs and pseudoarthrosis of the first rib.[118,136] These anomalies are seen either in isolation or in association with other anomalies and as part of other syndromes.[118,136] This section will focus on SNR due to

its relevance to this project. SNR (known as accessory ribs, extra ribs) can be categorized based on their anatomic position into cervical SNR, intrathoracic SNR, lumbar SNR, pelvic ribs. Intrathoracic and pelvic ribs are rare anomalies.[118,136] Cervical ribs can be unilateral or bilateral and have a prevalence ranging from (0.04-4.5%).[118] Lumbar SNR manifest as a unilateral or bilateral extra ribs with a prevalence ranging in (0.04%-5.8%) with one study reporting a prevalence as high as 15.8%.[118,137] Cervical and lumbar ribs are commonly asymptomatic. However, their presence can cause some problems.[118] For, example cervical ribs can cause a condition known as thoracic outlet disease which has vascular (cerebral embolism) and neurologic symptoms (extreme pain, migraine, and Parkinson's like symptoms). The vascular symptoms seen are due to compression and associated reduction in the blood flow through subclavian artery and vein, and carotid arteries. The neurologic symptoms are associated with altered positioning of stellate ganglia, sympathetic ganglia and C7-T1 nerve roots.[118,136,138] Moreover, lumbar ribs can be associated with pain in the lumbar region and disc degeneration in L4-5.[118,138]

1.8 Etiology of Supernumerary Ribs

Little is known about the etiology of SNR in humans.[118,137] However, evidence from animal models and humans suggests that both genetic and environmental factors contribute to the formation of SNR.[118] For example, supernumerary cervical ribs were seen in a family across two generations. Furthermore, different mouse models targeting *Hox* genes (single and multiple) and other genes (*Cdx* and *Fgfr1*) demonstrate SNR (cervical and lumbar). The role of environmental factors in the etiology SNR is inferred mainly from animal models.[118] Exposure to certain environmental factors was shown to cause SNR in mice, such as retinoic acid, valproic acid, salicylate and maternal stress.[118,138]

1.9 SNR Mouse Models

As in CP mouse models, spontaneous SNR mouse models are available. For example, CD1 mice in which the frequency of SNR was between 14.3%- 25%. [118] SNR induced mutations in mice are also seen such as the radiation induced SNR model *Pgap1*^{oto} mice have supernumerary cervical rib at C7, [139] ENU induced SNR model such as *Rpl38*^{Rbt} mice which have an lumbar rib[140], targeted SNR models such as *Hoxa9* and *Hoxb9* single and double mutants[141] and gene trap induced SNR models such as *BC055757*$^{-/-}$ line which had an extra pair of ribs on L1.[142]

1.10 Specific Aims

OVE1328 line, is an insertional mutant mouse model generated by microinjecting the tyrosinase transgene complex into an FVB/NJ single cell embryo. The integration site of the transgene is mapped to chromosome 4 (near the centromere). The recessive homozygous mutation in OVE1328 mice is initially found to be associated with perinatal lethality and CP. This projects aims to:

- **Aim 1:** identify the phenotypic differences between OVE1328 (homozygous and hemizygous) line and the wild type FVB/NJ strain.
- **Aim 2:** determine the gene/genes disrupted by the transgene integration and characterize the resulting mutation.

CHAPTER 2: PHENOTYPIC CHARACHTERISTICS OF OVE1328 TRANSGENIC LINE

2.1 Introduction

2.1.1 Insertional Mutagenesis in Mice

In the early 1980s, successful delivery and integration of exogenous foreign DNA into the mouse genome was achieved. Mice harboring the exogenous material were known as transgenic mice.[143]Transgenic mice can be generated by either microinjecting the exogenous DNA into the pronuclei of a single cell embryo (zygote) or by lentiviral or retroviral infection of the developing embryo (at different stages of embryonic development) or by genetic alteration of embryonic stem cells (using transfection or viral infection).[111,144]

Pronuclear microinjection is mainly used to study the effect of expressing exogenous genetic material.[144]In this strategy, which is used to develop the mouse model studied in this chapter, the transgene randomly integrates into the genome, though there might be some bias to regions with open chromatin.[143]The transgene will integrate at one site or few other sites in the founder transgenic mice. One to hundreds of copies of the transgene can integrate in the genome in a tandem head to tail array. However, other less common arrangements are also seen (head to head or tail to tail). Little is known about the integration process of the transgenes.[112,143,145,146] Upon integration, the transgene will result in an injury at the insertion site: deletion, duplication mutations, rearrangements, translocations involving the insert and host genetic material.[111,112,146] These injuries induced at integration site can range from simple insertion [147] to extensive

deletions (in excess of 500kb) and complex rearrangements and chromosomal translocations.[145] This differs from viral infection strategy for gene transfer which results in a short (4-6bp) duplication in the host genome at the integration site with no other rearrangements or alterations in these sequences.[111,112,144]

During transgenic mice generation, there is a 5-10% chance that the transgene (exogenous DNA) integration will disrupt a functional gene or alter its expression resulting in an insertional mutation. The actual frequency is probably higher since the reported frequency is based on observed mutations and prenatal lethal phenotypes.[112,143,145] A major advantage of studying such mutations pertains to the presence of the transgene which can be used as a cloning tag to determine the mutation at the molecular level: a feature lacking in both spontaneous and induced mutations (chemical and radiation).[111,148] Studying mouse insertional mutants provided a great tool to identify genes important for normal development (Table 2.1). Though studying such mutations is of great benefit, insertional mutations are sometimes hard to dissect and challenging. This is more of a problem in the presence of complex rearrangements or translocations or extensive deletions that might disrupt a group of genes.[111,112]

Insertional mutation phenotype	Gene	Reference
Severe progressive glomerulonephritis	*Col4a4 and Col4a3*	149
XX sex reversal	*Ods*	150
Vestibular and cochlear defects and hearing problems	*Sfswap*	151
Fragile stratum corneum, perinatal death due to dehydration	*Spink5*	152
Cleft palate	*Ap2β1*	113
Left-right reversal of abdominal visceral organs,polycystic kidney disease, severe jaundice, and die by seven days of age	*Invs*	153
Cryptorchidism and complete sterility	*Crsp*	154
No spermatids production in male and no eggs production in females	*Stag3*	155

Table 2.1. Murine insertional mutations are helpful to understand developmental processes.

2.1.2 General Background of OVE1328 Transgenic Line

OVE1328 line, produced in Dr. Overbeek lab (Baylor College of Medicine) is one of many transgenic lines (>250 lines) generated by microinjecting tyrosinase minigene construct (alone or as a subcloned component of other plasmid constructs) into a zygote (Dr. Overbeek personal communication). The tyrosinase mini gene encodes tyrosinase enzyme, which is the first enzyme in the melanin synthesis pathway[156]. Melanin is responsible for skin, hair bulbs and eyes (retinal pigment epithelial cells) pigmentation in mammals[156]. The integration and expression of the tyrosinase transgene in the FVB/N mice rescues the albino phenotype (white fur and red eyes) resulting in an animal that has a pigmented fur and black eyes[157]. Because of these distinct physical characteristics, the use of the tyrosinase transgene represents a useful tool for distinguishing wild type mice from transgenic ones.[157]

The transgenic mice generated were bred to homozygosity and examined for recessive insertional mutations. The breeding outcome in most of the cases resulted in normal progeny (Dr.

Overbeek personal communication). However, several lines showed insertional mutations (Table 2.1). About 15 lines had no viable adult homozygotes and 8 of these lines demonstrated perinatal lethality (Dr. Overbeek personal communication). Four of the lines with perinatal lethality had a CP phenotype (OVE lines 270,427,1226b and 1328) (Dr. Overbeek personal communication). These lines were kindly provided to our lab for further interrogation as our lab is interested in understanding the genetic basis of craniofacial development.

Until now, 2 of these lines, OVE427 and OVE1226b, have been characterized while OVE270 line went extinct and was not studied. In OVE427 line[113], the insertional mutation was mapped to the *Ap2β1* gene located on mouse chromosome 11. This gene encodes β2 adaptin subunit of the heterotetrameric adaptor protein 2 (AP2) complex which is a player in the clathrin dependent endocytosis. The mutation associated with the transgene integration prevents the expression of β2 adaptin at mRNA and protein levels in homozygous mutants. Homozygous mutants did not have developmental defects other than CP.

In OVE1226b line[158], the transgene integration was mapped to the intergenic region near the 3' end of collagen, type XI, alpha 1 (*Col11a1*) gene on mouse chromosome 3. The integration results in significant reduction of *Col11a1* gene expression at both mRNA and protein levels. In addition to CP, homozygous mutants demonstrate micrognathia and chondrodysplasia.

2.1.3 Generation of OVE1328 Transgenic Line

OVE1328 founder mouse was generated by the pronuclear microinjection of tyrosinase minigene transgene complex in to an FVB/N albino single cell embryo (Dr. Overbeek personal communication). OVE1328 line was then generated by mating the OVE1328 founder mouse

(hemizygous for the transgene complex) and a Wt (FVB/N) albino mouse. The line was maintained for more than 20 generations. Homozygous mutants were generated by intercrossing OVE1328 hemizygous male and female mice. Using fluorescent in situ hybridization (FISH), the transgene integration in OVE1328 line was mapped to chromosome 4 band A2, near the centromere[159].

The transgene complex used in generating OVE1328 line is composed of the 4.1 Kb tyrosinase minigene insert (C57BL/6 tyrosinase cDNA driven by the Balb/c tyrosinase promoter (2.25Kb)) [157]subcloned into an MSCV-neo vector driven by the human phosphoglycerate kinase promoter[160]and (Dr. Overbeek personal communication) .

2.1.4 General Characteristics of OVE1328 Transgenic Line

As seen in (Fig 2.1) there are few physical features that distinguish the OVE1328 transgenic line from the Wt counterpart. The Wt albino phenotype (white fur, red eyes) is rescued in OVE1328 hemizygous mutant mice (grey fur and black eyes). Furthermore, there are no viable adult homozygous mutants as they demonstrate perinatal lethality. Some of the aforementioned differences during adulthood are also demonstrated during embryonic development. For example, transgenic OVE1328 embryos demonstrate black pigmented retina while Wt embryos lack pigmentation in their retina.

The presence of CP in OVE1328 homozygous embryos distinguished them from hemizygous mutants which had normal palates.

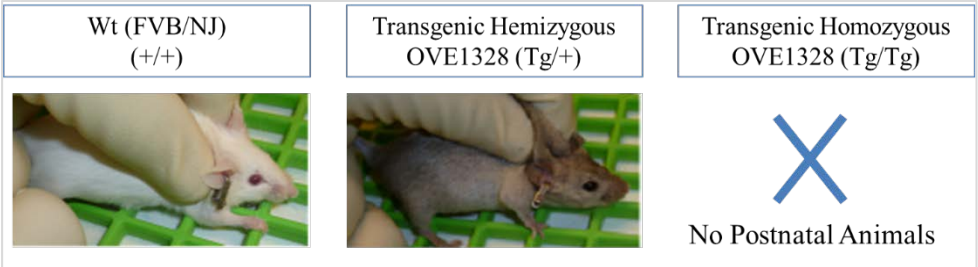


Figure 2.1 General physical differences between transgenic OVE1328 line and Wt mice.

This chapter aims to characterize the phenotypic differences between OVE1328 homozygous mutant (Tg/Tg) embryos and their hemizygous mutant (Tg/+) and Wt counterparts.

2.2. Materials and Methods

Husbandry and Timed Matings of OVE1328 Mice

Founder mice were provided as gifts from Dr. Paul Overbeek (Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas). Mice were housed in the Division of Laboratory Animal Medicine (DLAM) facility, Koury Oral Health Sciences Building at UNC at a 12 hour light/dark cycle with *ad libitum* access to water and food. OVE1328 (Tg/Tg) embryos were generated by intercrossing OVE1328 (Tg/+) females and OVE1328 (Tg/+) males which were allowed to mate for 4-8 hours. Females were checked for vaginal plugs and the day the plug was detected was designated E 0.5. At E18 pregnant dams were euthanized using CO2 inhalation and death assured by cervical dislocation. Embryos were harvested by Cesarean section and euthanized using Ketamine 90mg/kg/ Xylazine 14mg/kg. Upon harvest litter size and embryo related information (position within the uterine horn, weight, eye color and the presence of CP) were recorded. Animal studies were approved by the University of North Carolina at Chapel Hill Animal Care and Use Committees.

Genotyping OVE1328 Transgenic Line

Genomic DNA was extracted from 174 pieces of embryo tails according to manufacturer's instructions (Puregene Core Kit A, Qiagen Inc., Valencia, CA, USA). The presence of the transgene in embryos was confirmed by PCR (Promega Kit, Madison, WI, USA) using primer pairs designed across exon 1 and 2 of tyrosinase minigene as shown in (APPENDIX 2.1,2.2 and 3.9). In the presence of the transgene the primer pair amplifies a 700 bp product. However, the primer pairs are unable to amplify a product from genomic endogenous tyrosinase sequences (chromosome 7) as the amplification is hindered by the presence of intron 1

(8.495 kb) between the primer pair target sites in gDNA. This makes the primer pair very useful in distinguishing between Wt (no amplification product) and transgenic (700bp product) embryos. Furthermore, OVE1328 (Tg/+) embryos were distinguished from OVE1328 (Tg/Tg) embryos based on the CP phenotype. Embryo palates were checked for a cleft after each harvest.

Histological Analyses of OVE1328 Line

Wt, OVE1328 (Tg/+) and OVE1328 (Tg/Tg) E18 embryos were fixed in 10% neutral buffered formalin (1 week at 4^{o}C), stored in 70% v/v ethanol at 4^{o}C. Each embryo was cut into four parts (head, upper body third, middle body third, lower body third) and processed overnight to paraffin wax (containing DMSO) according to standard protocols. Paraffin infiltrated samples were embedded in DMSO free paraffin. Paraffin blocks were stored at -20^{o}C. Both coronal head (from the snout through the posterior segment of the eye) and transverse body sections of the embryos were cut (6µm) and stained with H&E according to standard protocols. An optical microscope (Nikon, Alphaphot 2) was used to examine the sections at 40X and 100X magnification.

Skeletal Clearing and Staining

Wt, OVE1328 (Tg/+) and OVE1328 (Tg/Tg) E18 embryos were cleared and stained [161] with slight modifications of incubation times as shown in (APPENDIX 2.3). Alcian blue 8GS and alizarin red were used for staining cartilage and bone, respectively. Stained skeletons were examined. Bones of right forelimbs (humerus, radius, and ulna) and hindlimbs (femur, tibia, and fibula) were dissected, mounted flat beside a calibrated ruler, and imaged using a dissection microscope with a digital image capture. Using an image and photo editing software (Paint.NET) a line was drawn through the central longitudinal axis of the bones and 3 repeated measurements

(in pixels) were recorded per bone. The average length of each bone in pixels was converted to millimeters using the imaged calibrated ruler. For rib cage examination, ribs on each side were counted 3 times on 3 isolated occasions.

Cephalometric Measurements of Embryo Heads

Wt, OVE1328 (Tg/+) and OVE1328 (Tg/Tg) E18 embryo heads were fixed in 10% neutral buffered formalin (1 week at 4ºC), stored in 70% v/v ethanol at 4ºC. Three images where taken for each head and a calibrated ruler was included in all of the captured images. Using an image and photo editing software (Paint.NET), 3 lines were drawn and were designated as the interpupillary distance (IPD, a line drawn between the centers of the pupil of both eyes), snout-occiput distance (SOD, a line drawn from the tip of the snout to the mid-occipital margin), snout width (SW) (measured at a line demarcating the end of the thickened epidermis of the muzzle). The average length of each line in pixels was converted to millimeters using the imaged calibrated ruler.

Micro-CT Scanning and Analysis

Upper body third of Wt, OVE1328 (Tg/+) and OVE1328 (Tg/Tg) embryos (stored in 70% v/v ethanol / 4ºC) were scanned in air using the Skyscan 1074HR micro-CT (Skyscan, Aartselaar, Belgium). Pixel size (µm) =20.7. Scanning conditions were as following; Source Voltage (kV) = 40, Source Current (µA) = 1000, Exposure (ms) = 420. Reconstruction was done using NRecon program (Skyscan, Aartselaar, Belgium). Following reconstruction, CTAn software® (Version 1.9.1.0, Skyscan, Aartselaar, Belgium) was used to determine the tissue mineral density (TMD) of embryo humerii. TMD measurements were performed following calibration to hydroxyapatite phantoms (250mg/cc and 750mg/cc) (CIRS Inc., Norfolk ,VA,

USA). The cortical bone 1 mm below the deltoid tuberosity (right and left humerii) was selected as the region of interest (ROI). The selected ROI contained 49 contiguous transverse slices (each 20.7μm thick). For every ROI a binary threshold was determined. An average binary threshold/genotype was calculated and named the universal threshold. Mean TMD was determined for the entire selection of slices/ sample based on the universal threshold calculated per genotype.

Statistical Analysis

For multiple comparisons one-way ANOVA followed by Bonferroni post hoc test was done. P values ≤ 0.05 were considered statistically significant. For limb measurement data, Generalized Estimating Equation (GEE) approach was used.

2.3 Results

Phenotypic Characteristics of OVE1328 Transgenic Mouse Line

Cleft of the Secondary Palate (CP)

A total of 174 embryos were examined. Validated by PCR, 112 embryos were transgene positive (Tg) and 57 embryos were transgene negative (Wt). Among transgene positive embryos, 74 embryos had normal palates (OVE1328 (Tg/+)) while 38 embryos had a V shaped CP (OVE1328 (Tg/Tg)) (Fig 2.2.A). Intercrossing OVE1328 (Tg/+) mice revealed that CP was inherited as an autosomal recessive trait. The genotype frequencies of Wt, OVE1328 (Tg/+) and OVE1328 (Tg/Tg) embryos were 32.8%, 42.5%, 21.8%, respectively (Fig 2.3).

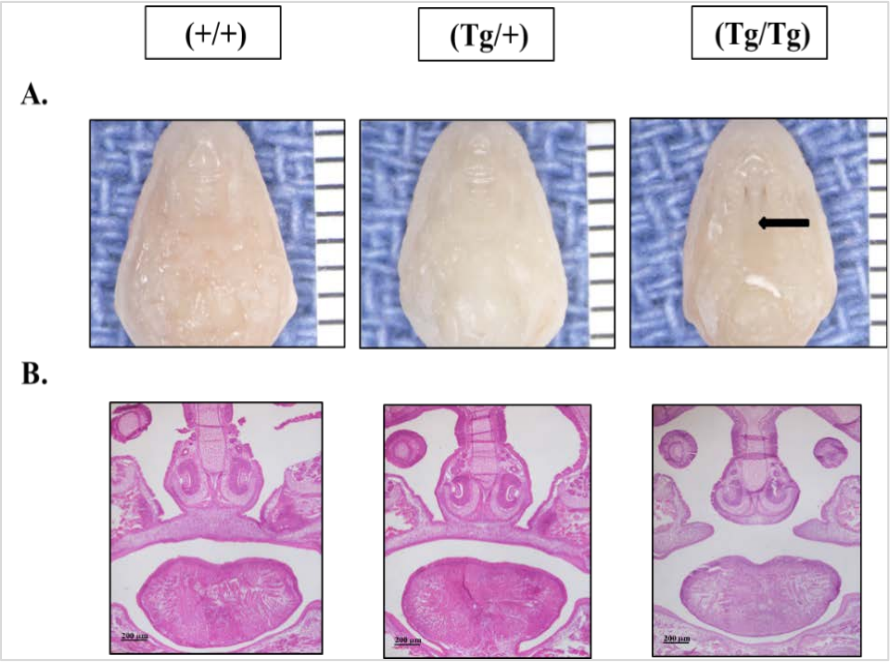


Figure 2.2. CP phenotype in OVE1328 line A. Oral view of Wt (+/+), OVE1328 (Tg/+) and OVE1328 (Tg/Tg) embryo heads sans the mandibles (E18). Black arrow pointing to nasal cavity. Scale bar (distance between 2 horizontal lines) =0.86 mm. B. H&E stained coronal sections of anterior secondary palate for Wt, OVE1328 (Tg/+) and OVE1328 (Tg/Tg) embryos (E18). Scale bar =200um.

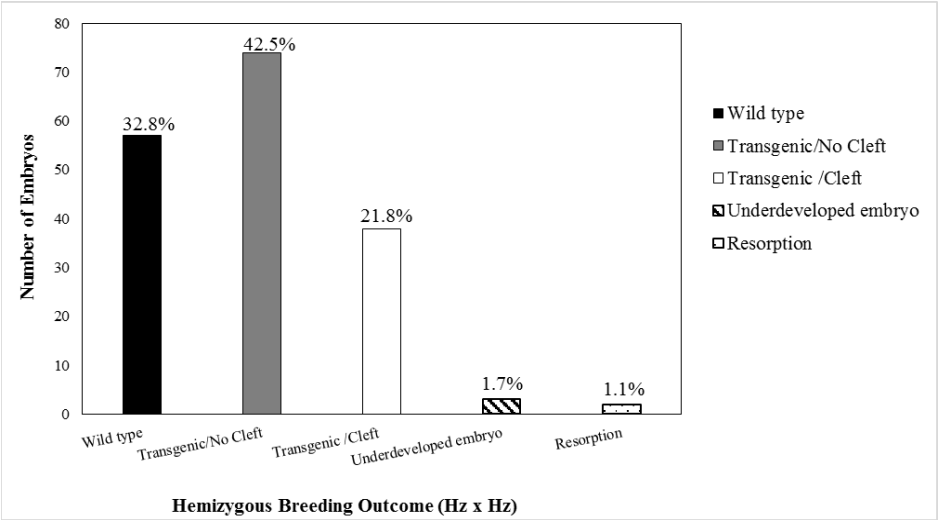


Figure 2.3. E18 Embryo genotype frequencies of OVE1328 (Tg/+) intercrosses. A total of 174 embryos were produced.

Histological examination of the palatal shelves in OVE1328 (Tg/Tg) embryos show elevated palatal shelves. However, the palatal shelves fail to make contact and fuse compared to Wt and OVE1328 (Tg/+) embryos which had intact palates (Fig 2.2.B).

Histological Analyses

Examination of histological sections of E18 embryo major organs including: heart and lungs (Fig 2.4 A); thymus (Fig 2.4 B); liver (Fig 2.5 A); kidneys and spleen (Fig 2.5 C); stomach (Fig 2.5 B) and intestines (Fig 2.6) did not reveal any significant differences in their morphologies among Wt and OVE1328 (Tg/Tg) E18 embryos.

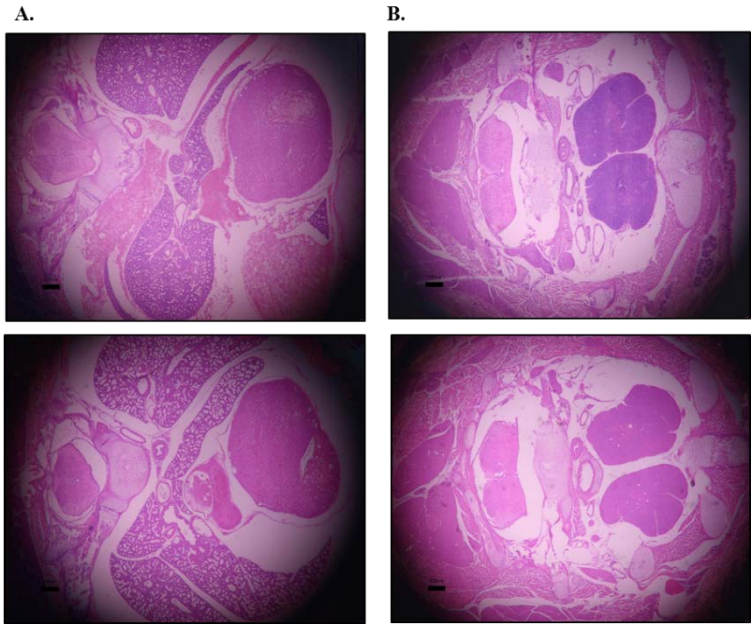


Figure 2.4. Histological analysis (H&E stained) of the heart, lungs, and thymus gland. A. Heart and lungs. B. Thymus gland. No differences were seen in general between Wt (n=3), upper panel, and OVE1328 (Tg/Tg) (n=3), lower panel, E18 embryos. Scale bar = 200 um.

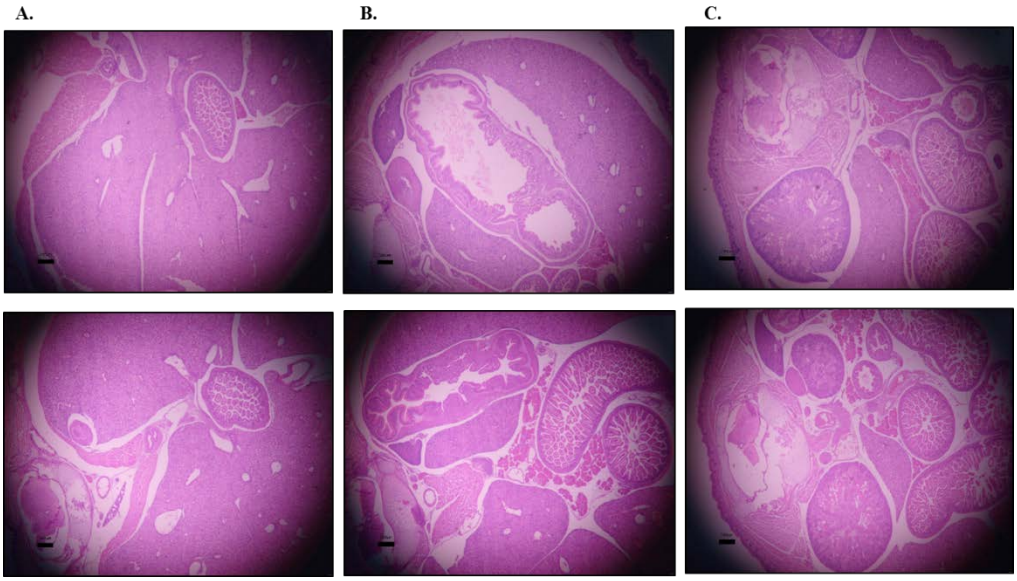


Figure 2.5. Histological analysis (H&E stained sections) of some major organs A. Liver B. Stomach C. Kidney and spleen. No significant difference was seen between Wt (n=6), upper panel, and OVE1328 (Tg/Tg) (n=6), lower panel, E 18 embryos. Scale bar=200 um.

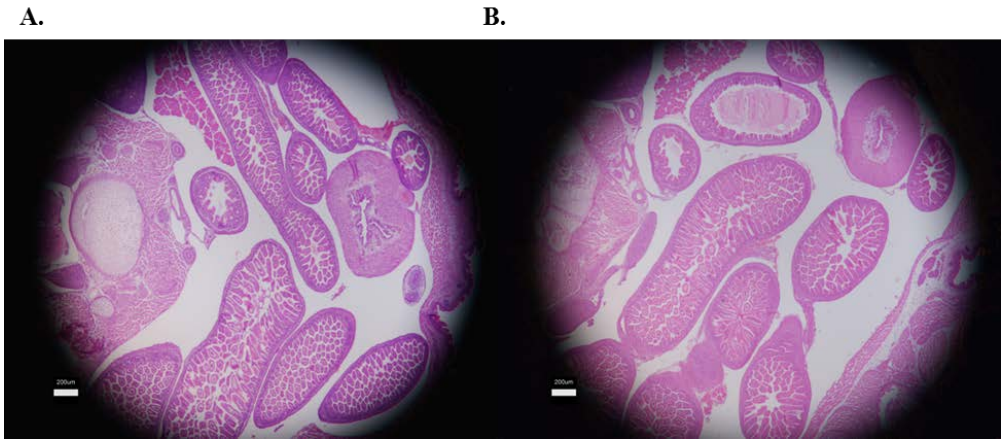


Figure 2.6. Histological analysis (H&E stained sections) of the intestines in Wt and OVE1328 (Tg/Tg) E18 embryos. A. Wt (n=6) a B. OVE1328 (Tg/Tg) (n=6). No significant difference was seen in the intestinal structure. Scale bar=200 um.

Growth delay

Embryos with CP, OVE1328 (Tg/Tg), appear smaller than OVE1328 (Tg/+) and Wt littermates. As part of phenotyping and characterizing the OVE1328 line, OVE1328 (Tg/Tg) embryos (n=25) weighed on average ~13% less (P<0.001) than Wt (n=36) and OVE1328 (Tg/+) embryos (n=60) (Fig 2.7).

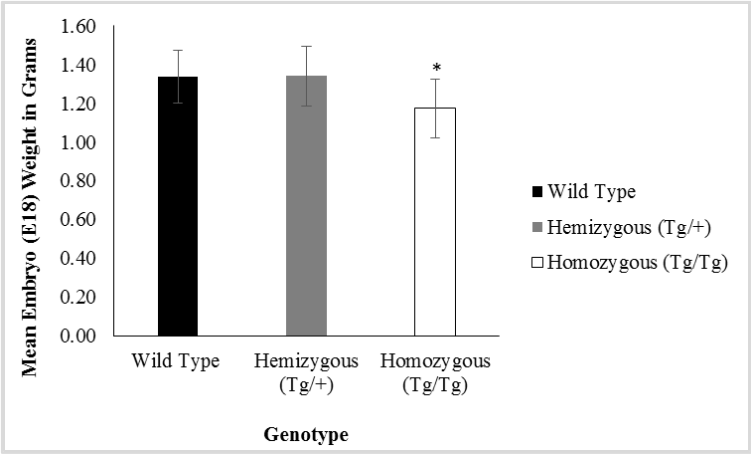


Figure 2.7. Average E18 embryo weights. OVE1328 (Tg/Tg) embryos weighed ~ 13% less (P<0.001) than Wt and OVE1328 (Tg/+) embryos of same gestational age. Data presented as Mean± SD. (*): statistically significant difference.

Furthermore, examination of the cleared skeletons demonstrated that OVE1328 (Tg/Tg) embryos (E18) were generally smaller compared to Wt and OVE1328 (Tg/+) embryos (Fig 2.8). Reduced weight and size in OVE1328 (Tg/Tg) was an indication of probable growth delay.

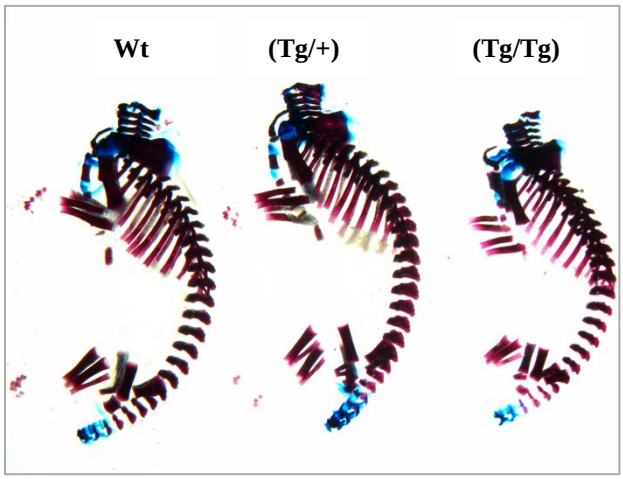


Figure 2.8. Cleared skeletons of E18 embryos. OVE1328 (Tg/Tg) skeletons were generally smaller than Wt and OVE1328 (Tg/+) counterparts.

Further characterization of the cleared skeletons of E18 OVE1328 (Tg/Tg) embryos revealed shorter forelimb bones (humerus, radius, ulna), when controlling for the side, compared to Wt (P=0.0011, P=0.0008, P=0.0014) and OVE1328 (Tg/+) (P=0.0006, P=0.0022, P=0.0033) embryos. On the other hand, the mean lengths of hindlimb bones were not statistically significantly different across all genotypes (femur (P=0.0775), tibia (P=0.0767) and fibula (P= 0.0972)) (Fig 2.9) and (APPENDIX 2.5, 2.6).

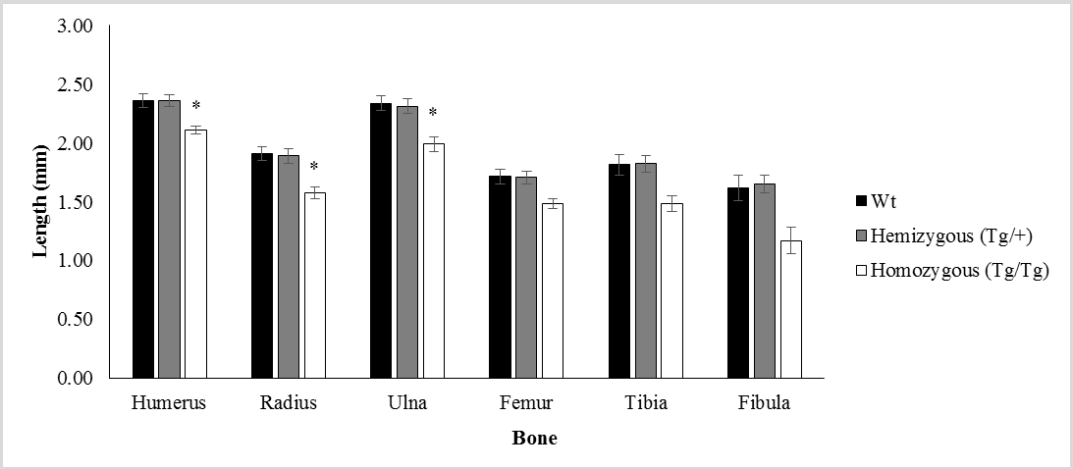


Figure 2.9.Mean lengths of forelimb and hindlimb bones of E18 Wt (n=10), OVE1328 (Tg/+) (n=12) and OVE1328 (Tg/Tg) (n=12) embryos. Controlling for the side, OVE1328 (Tg/Tg) embryos had shorter forelimb bones compared to Wt and OVE1328 (Tg/+) embryos. Data presented as Mean± SE. (*): statistically significant difference.

Limb Mineralization

Occasionally the cleared skeletons of OVE1328 (Tg/Tg) E18 embryos showed diminished mineralization (reduced alizarin staining) of metatarsals, metacarpals and phalanges compared to the Wt and OVE1328 (Tg/+) groups. To assess the degree of mineralization micro-CT analysis of right (n=6) and left humerii (n=6) of OVE1328 (Tg/Tg) E18 embryos was performed and showed no significant differences in the mean TMD compared to Wt and OVE1328 (Tg/+) groups (n=6/genotype/side) (one way ANOVA, right humerus P=0.807, left humerus P=0.154) (Fig.2.10) and (APPENDIX 2.7).

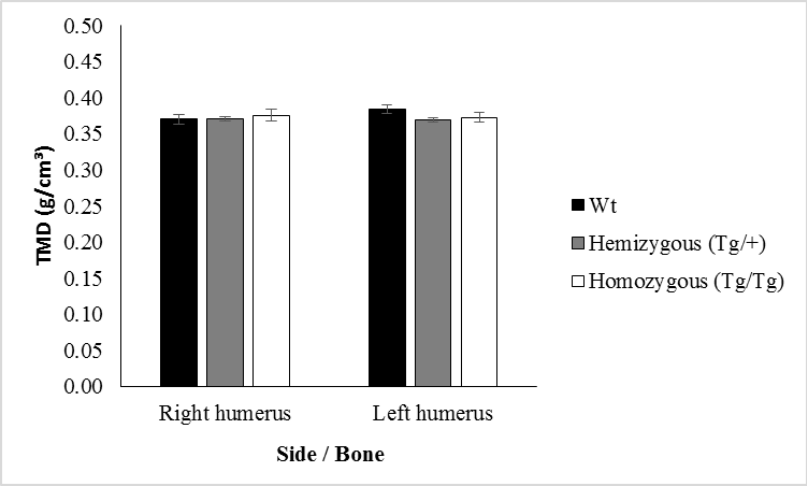


Figure 2.10. Average TMD of right and left humerii from E18 Wt, OVE1328 (Tg/+) and OVE1328 (Tg/Tg) embryos (n=6/genotype/side). There were no significant differences in TMD between all genotypes. Data presented as Mean ± SE.

Cephalometric Measurements

As part of interrogating size differences between OVE1328 line and Wt mice, E18 fixed embryo heads were examined. Mean head measurements of OVE1328 (Tg/Tg) embryo heads were similar to Wt and OVE1328 (Tg/+) groups as shown in (Fig 2.11) and (APPENDIX 2.8). There was no statistically significant difference between all genotypes with respect to mean length of inter-pupillary distance (IPD) (one way ANOVA, P= 0.062), snout width (SW) (one way ANOVA, P=0.811) and snout-occiput distance (SOD) (one way ANOVA, P= 0.146).

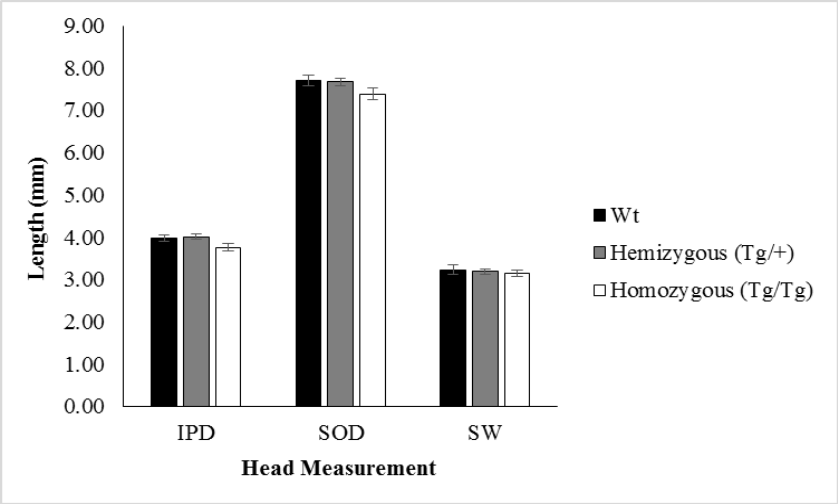


Figure 2.11. Comparison of E18 embryo head measurements between Wt, OVE1328 (Tg/+) and OVE1328 (Tg/Tg) embryos. Data presented as Mean± SE. IPD: inter-pupillary distance. SOD: snout occiput distance. SW: snout width.

Supernumerary Ribs

Thoracic cage of mice normally contains 13 pairs of ribs that originate from the vertebral bodies of 13 thoracic vertebrae. General examination of E18 embryo cleared skeletons showed that 85.71% of OVE1328 (Tg/Tg) embryos (n=7) had rudimentary supernumerary ribs (extra ribs, accessory ribs) (Fig2.12 and 2.13) and (APPENDIX 2.4). The unilateral (50%) or bilateral

(50%) accessory ribs arose from the vertebral bodies of the first lumbar vertebrae (Fig 2.14). All of Wt (n=5) and OVE1328 (Tg/+) (n=9) embryos had normal rib counts (13 pairs).

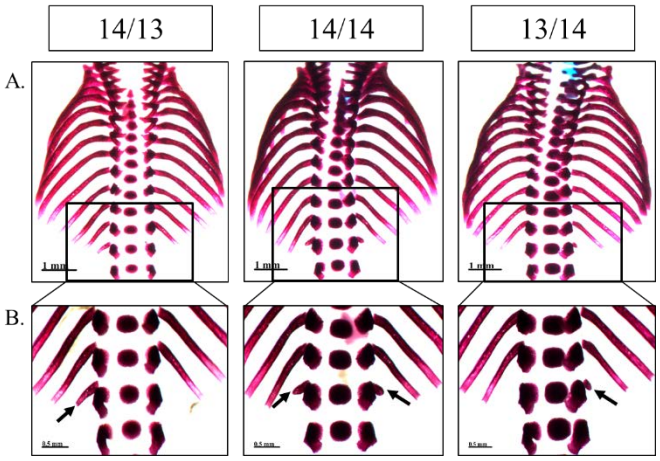


Figure 2.12. A. Supernumerary ribs in OVE1328 (Tg/Tg) E18 embryos. A. General view. Scale bar = 1 mm. B. Enlarged image of the rectangles. The supernumerary ribs are seen either unilaterally or bilaterally (black arrowhead). Scale bar= 0.5mm.

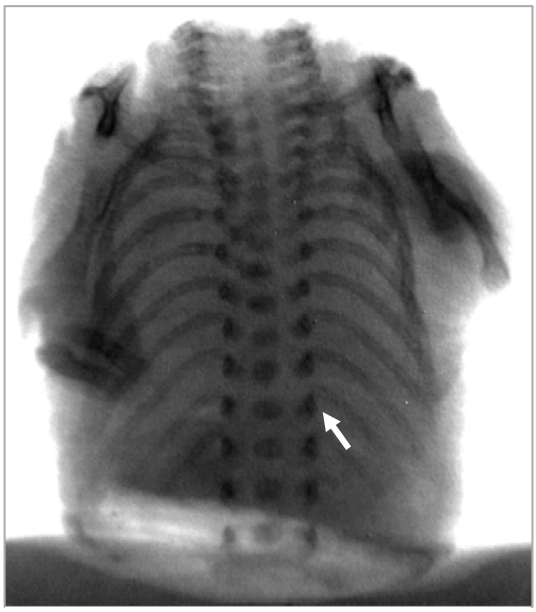


Figure 2.13. Micro-CT X-ray projection image of OVE1328 (Tg/Tg) E18 embryo rib cage. Note the rudimentary accessory rib on the right side (white arrowhead).

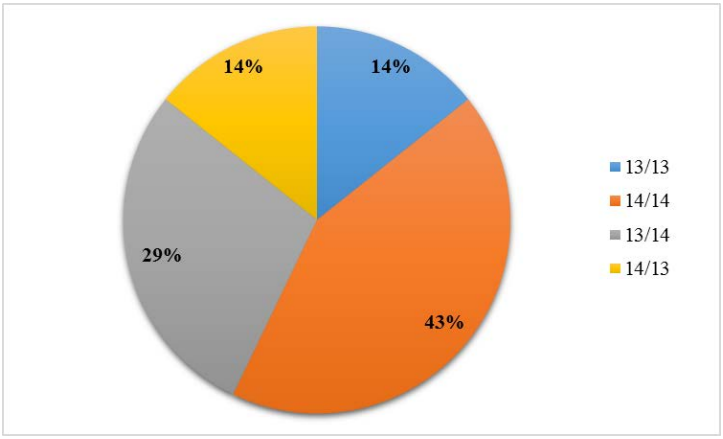


Figure 2.14. Supernumerary ribs in OVE1328 (Tg/Tg) embryos**.** The pie chart shows the percentage of OVE1328 (Tg/Tg) embryos (n=7) having supernumerary ribs and depicts the variation in the supernumerary rib phenotype; unilateral, bilateral or absent. Pie chart key: 13/13 rib count (right side/left side).

2.4 Discussion

Several phenotypic differences were observed between OVE1328 (Tg/Tg) embryos and Wt embryos. While OVE1328 (Tg/+) embryos were phenotypically similar to Wt embryos, OVE1328 (Tg/Tg) demonstrated growth delay, cleft of the secondary palate and supernumerary lumbar ribs (bilateral or unilateral).

Examination of harvested embryo heads and H&E stained head sections revealed that palatal shelves successfully elevated but failed to fuse. A cleft of the secondary palate can be due to factors intrinsic to palatal shelves, such as lack of growth of palatal shelves, delayed or failed elevation, or in ability of the shelves to adhere and fuse properly. Additionally, cleft of the secondary palate can be due to factors extrinsic to the developing shelves such as small mandible, abnormally large tongue, and aberrant craniofacial growth.[11,20,40] It would be useful to histologically examine the development of palatal shelves in OVE1328 (Tg/Tg) at different time points starting at E11.5-E16. By comparing palatogenesis (growth, elevation, horizontal growth, adhesion and fusion) in OVE1328 (Tg/Tg) to that of wild type, it would be possible to determine at which stage during secondary palate formation did the shelves deviate from the norm. Gross examination of the OVE1328 (Tg/Tg) embryo heads with dissected mandibles did not reveal significant differences in head dimensions compared to Wt group. However, it would be useful to examine the intermaxillary relations and the mandibular size in OVE1328 (Tg/Tg) embryos compared to Wt counterparts by examining μCT sagittal sections and skeletally cleared embryo heads.[162]

Other than CP, OVE1328 (Tg/Tg) embryos had lumbar SNRs seen on L1 vertebrae. The lumbar vertebrae were seen either unilateral or bilateral and was not fully penetrant (85%). The presence of a lumbar rib on L1 indicates that L1 vertebrae are gaining a different identity that is

similar to T13 vertebrae, such transformation is known as anterior homeotic transformation. *Hox* genes are major regulators of anterior-posterior patterning.[129] The lumbar rib phenotype is seen in several single, double and paralogous murine mutants of *Hox* genes (Table 2.2).[163-167]

Hox gene null mutations	L1 rudimentary or supernumerary rib	Reference
c8	Yes	163
c8 & d8	Yes	163
b8 & c8 & d8	Yes	163
a9	Yes	164
c9	Yes, transitional L1 vertebrae	165
a10	Yes	166
a10 & c10 & d10	Yes, involving all lubar vertebrae	167

Table 2.2. Single, double and paralogous null mutations in *Hox* genes that contribute to L1 supernumerary ribs.

It is possible that the lumbar supernumerary rib phenotype seen in OVE1328 (Tg/Tg) embryos is associated with a defective *Hox* gene (code) or expression pattern. However, the phenotype in OVE1328 (Tg/Tg) does not fully capture the phenotype in any of the *Hox* mutants. This could be due to the effect of examining the phenotype in mice with different genetic backgrounds[168] or the reduced expression of the *Hox* gene in OVE1328 (hypomorphic variant).However, the mutation in OVE1328 (Tg/Tg) might have affected other transcription factors (possibly unidentified yet) or other proteins known to be important for determining segmental identity. For example, hypomorphic *Fgfr1* mutant mice (targeted mutation) demonstrate supernumerary rib phenotype at the lumbar vertebra L1, in addition to other developmental defects.[135] It would be useful to examine the expression pattern of relevant *Hox* genes (in which an alteration is associated with lumbar SNR phenotype, (Table 2.2)) and the expression pattern of *Fgfr1* and *Fgf1* signaling pathway in OVE1328 (Tg/Tg). Using RNA extracted from dissected palatal shelves

CHAPTER 3: MOLECULAR CHARACHTERIZATION OF THE INSERTION MUTATION IN OVE1328 TRANSGENIC LINE

3.1. Introduction

During the process of generating transgenic mice there is a 5-10% chance that integration of the transgene complex results in a secondary mutation not related to the transgene phenotype[112]. During integration of the transgene complex in the mouse genome there can be genomic rearrangements that include deletions and inversions[112]. When bred to homozygosity transgenic mice can display abnormal phenotypes most likely due to disruption a functional gene preventing or altering its expression[112]. In OVE1328 line, expression of the tyrosinase enzyme from the transgene complex rescued the albino phenotype in both hemizygous and homozygous mice. In addition to the rescued albino phenotype OVE1328 (Tg/Tg) mice had other phenotypic characteristics not related to the ectopic tyrosinase expression. This was confirmed by the fact that OVE1328 (Tg/+) mice did not show those phenotypic characteristics. Therefore it was of great interest to determine accurately the transgene integration site in the OVE1328 line and to identify potential gene/genes in which disruption is associated with the phenotype seen. To better understand the molecular basis of the mutation in OVE1328 line, a brief and general introduction to the transgene integration process and its sequelae at insertion site will be presented in the following section.

3.1.1 Transgene Integration Process

During the generation of transgenic mice, the plasmid DNA injected in the pronucleus of mouse zygotes can undergo concatenation followed by illegitimate recombinaton.[146,169] Concatenation is the formation of an array (concatemers) of exogenous DNA molecules. The associations of the DNA molecules can take different forms (head to tail, head to head and tail to tail).[169] Head to tail association is seen more than the other forms at a ratio of 2:1:1, respectively.[169] Furthermore, there is evidence that linear exogenous DNA undergoes circularization and random linearization.[170] Extrachromosomal homologous recombination is believed to be involved in the process of joining exogenous DNA molecules in head to tail association. For the less frequent random associations (head to head, tail to tail) non-homologous end joining (NHEJ) is believed to be responsible for the formation of such random concatemers.[146,171] In NHEJ, short sequence homology or deletions or insertions of up to 25 nucleotides is involved in the process of joining the ends of two recombining DNA molecules.[146]

The injected transgene or transgene array randomly integrates into the host (mouse) genome. The integration is mostly at one site or might be at few sites.[146] In general, little is known about transgene integration mechanisms.

3.1.2 Transgene Integration: Effects on Transgene and Genomic Loci at Insertion Site

There is evidence that mutations such as point mutations, deletions and complex rearrangements (insertion of gDNA) modify the transgene sequence.[111,112] It is however unknown whether such modification take place before or during the integration process.[146] Furthermore, genomic loci at insertions sites can also be mutated and rearranged. Such mutations include deletions, duplications and translocations.[146] The disruption at a genomic locus could involve genes (coding sequence or the non-coding sequence) or predicted genes.[146]

3.1.3 Transgene Integration in OVE1328 Transgenic Line

OVE1328 mice homozygous for the transgene complex die during the perinatal period suggesting that the transgene complex integrated is a critical region of the genome. Fluorescent *in situ* hybridization (FISH) was used to localize the transgene complex to chromosome 4 band A2 (chr4:14,882,674-17,763,191 bp ± 2,000,000 bp).[159] This significantly reduced the number of candidate genes to ~52 (NCBI Map Viewer as of 8.2014) (Table 3.1).

Gene Symbol	Genomic locus/strat	Genomic locus/stop	Gene Symbol	Genomic locus/strat	Genomic locus/stop
Triqk	12906837	12981485	**Decr1**	15917240	15945507
LOC102640073	13091351	13108752	**Mir6400**	15942900	15942991
LOC102640185	13283930	13307483	**Nbn**	15957967	15992589
LOC102639470	13444763	13447593	**LOC102638365**	15969473	15970509
Gm11826	13482690	13483193	**Osgin2**	15997121	16014725
LOC102640424	13515142	13584228	**LOC102631758**	16010943	16013487
LOC102640354	13585545	13592621	**Ripk2**	16122741	16163676
Runx1t1	13743302	13895056	**A530072M11Rik**	16164110	16266225
D930021N14	13919954	13995193	**Gm11865**	16246060	16250256
Gm11824	14170063	14170911	**Gm11864**	16390462	16391664
LOC102640675	14240289	14298100	**Gm11862**	16714956	16715620
Gm11836	14433877	14434481	**Gm11848**	17032691	17033656
Slc26a7	14502430	14621990	**LOC102631866**	17547344	17559733
Lrrc69	14623618	14796463	**Mmp16**	17853482	18118926
Otud6b	14809505	14826587	**Gm11860**	18009236	18010245
Tmem55a	14864219	14915260	**Gm11853**	18472198	18481668
Gm11837	14929908	14953031	**Cnbd1**	18767985	19122566
Necab1	14952245	15149804	**Gm11868**	18844205	18845679
LOC102640902	15021890	15031776	**Gm12393**	19259090	19259943
LOC102640824	15092189	15126287	**Cngb3**	19280850	19510623
Tmem64	15265820	15286753	**Gm11877**	19298257	19299183
LOC102631581	15287985	15328393	**Cpne3**	19519252	19570104
LOC102631507	15309164	15317408	**Rmdn1**	19575066	19606932
Gm11857	15636169	15636485	**LOC102632171**	19590105	19593590
Gm11855	15819062	15821373	**Wwp1**	19608296	19709004
Calb1	15881264	15906709	**LOC102632246**	19716576	19720179

Table 3.1. List of candidate genes present on chromosome 4 band A2 ± 2,000,000bp.

To further narrow down the number of candidate genes genotyping microarrays were used. The rationale for performing genome-wide profiling of single nucleotide polymorphisms (SNPs) was based upon the hypothesis that transgene complex integration in the OVE1328 line resulted in deletion of genomic DNA, including single nucleotide polymorphisms (SNPs). SNPs are normal variations in the nucleotide sequence of the genome.[172] Microarrays are designed to identify such variations utilizing the principle of hybridization between different DNA molecules. If no hybridization takes place a single spot in the microarray, a no call is retrieved from that spot. Therefore, based on the hypothesis a pattern of lost SNPs (seen as a no call on a microarray) would be only seen with gDNA of OVE1328 (Tg/Tg) samples but not Wt samples.

To test the hypothesis a JAX Mouse Diversity Genotyping Array (The Jackson Laboratory, Bar Harbor, ME) was used. This custom Affymetrix array interrogated 623,124 single nucleotide polymorphisms (SNPs) and 916,269 invariant genomic probes (IGPs) in a mouse DNA sample. The average no call rate across 623,124 SNPs was 0.695% for Wt gDNA samples (n=3) and 0.635% for OVE1328 (Tg/Tg) gDNA (n=3) samples. Counting the no call rate across all chromosomes (excluding Y chromosome), chromosome 4 had the highest no call rate in OVE1328 (Tg/Tg) gDNA (52/36,748 SNPs) compared to Wt gDNA (26/36,748 SNPs). One group of contiguous SNPs (10 SNPs) showed a pattern of no call on chromosome 4 in OVE1328 (Tg/Tg) samples only. This group included a region of 34,443 bp (13,881,698-13,916,141bp/ GRCm Build 38) of genomic DNA. Two of the SNPs in this group (rs28297363 and rs28297360) were located in intron 13 of *Runx1t1*gene. The remaining SNPs were located in the intergenic region 3' to *Runx1t1*gene.[159] This preliminary data lead to the interrogation of *Runx1t1* gene in OVE1328 (Tg/Tg) embryos (Fig 3.1).

3.1.4 *Runx1t1* a Member of ETO (Eight Twenty One) Gene Family

Runx1t1 is the murine ortholog of human *RUNX1T1* (Table 3.2). *RUNX1T1* a member of ETO gene family of transcriptional co-factors. The family is composed of closely related members; *RUNX1T1* (also known as *CDR; ETO; MTG8; AML1T1; ZMYND2; CBFA2T1*), *CBFA2T2* (also known as *EHT; p85; MTGR1; ZMYND3*) and *CBFA2T3* (also known as *ETO2; MTG16; MTGR2; ZMYND4*). A great degree of homology is shared between ETO family members (Table 3.3). Furthermore, ETO family members structurally share 4 conserved domains/regions known as nervy homology region (NHR) due to their homology to *Drosophila*

A.

OVE1328 (+,+)	OVE1328 (+,+)	OVE1328 (+,+)	OVE1328 Tg/Tg	OVE1328 Tg/Tg	OVE1328 Tg/Tg	rs Number	Chromosome	SNP Position (Build 38) bp	Locus Symbol	Class Function
G/G	G/G	G/G	No Call	No Call	No Call	rs28297363	Ch4	13881698	Runx1t1	Intron 13
A/A	A/A	A/A	No Call	No Call	No Call	rs28297360	Ch4	13881892	Runx1t1	Intron 13
C/C	C/C	C/C	No Call	No Call	No Call	rs28297303	Ch4	13897399	N/A	Intergenic
C/C	C/C	C/C	No Call	No Call	No Call	rs28297268	Ch4	13904977	N/A	Intergenic
C/C	C/C	C/C	No Call	No Call	No Call	rs28297266	Ch4	13905077	N/A	Intergenic
A/A	A/A	A/A	No Call	No Call	No Call	rs28297254	Ch4	13906924	N/A	Intergenic
T/T	T/T	T/T	No Call	No Call	No Call	rs28297239	Ch4	13908743	N/A	Intergenic
A/A	A/A	A/A	No Call	No Call	No Call	rs28297235	Ch4	13909408	N/A	Intergenic
C/C	C/C	C/C	No Call	No Call	No Call	rs28297207	Ch4	13913138	N/A	Intergenic
G/G	G/G	G/G	No Call	No Call	No Call	rs6332151	Ch4	13916141	N/A	Intergenic

B.

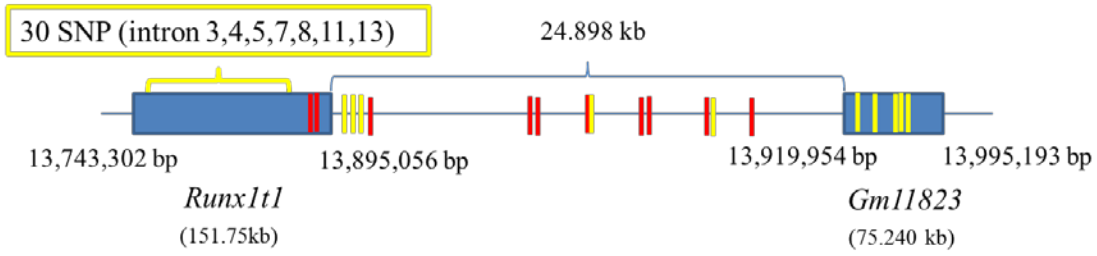


Figure 3.1 Diversity genotyping microarray data**.** A. SNP genotyping microarray data showing a group of no calls from contiguous SNPs across chromosome 4. The pattern of no calls is only seen with gDNA of OVE1328 (Tg/Tg) samples but not Wt samples. B. A schematic representation of the position of the contiguous SNPs with no calls seen among OVE1328 (Tg/Tg) gDNA samples. Key: red rods are SNPs with no calls and yellow rods are SNPs with calls.

Nervy protein and are discussed in the following section.[173,174] Murine orthologs of ETO family members are presented in Table (3.2). Protein alignments of murine (C57BL/6J strain) ETO family members are 99%, 96% and 88% identical to human ETO family members (Table 3.3), respectively (Homologene, NCBI database 8.2014*).*

ETO family member	Chromosome	Murine Ortholog	Chromosome
RUNX1T1	8q22	**Runx1t1** (aliases RP23-134H12.1, Cbfa2t1h, Eto Mtg8)	4 band A
CBFA2T2	20q11	**Cbf2t2** (aliases RP23-109C16.1, A430091M07, C330013D05Rikh, Mtgr1)	2 band C
CBFA2T3	16q24	**Cbfa2t3** (aliases A630044F12Rik, AI465270, AW229127h, Eto-2, Eto2, Mtgr2)	8 band E

Table 3.2. Murine orthologs of ETO family members and their aliases.

Similarity to RUNX1t1 Isoform 1	Query Coverage	Percent Identity
RUNX1t1 Isoform 1	100%	100%
RUNX1t1 Isoform 2	95%	99%
RUNX1t1 Isofrom 3	95%	99%
CBFA2t2 Isoform 1	89%	66%
CBFA2t2 Isoform 2	89%	66%
CBFA2t2 Isoform 3	57%	59%
CBFA2T3 Isoform 1	94%	71%
CBFA2T3 Isoform 2	93%	71%
CBFA2T3 Isoform 3	93%	74%

Table 3.3 ETO family members are closely related proteins. The table demonstrates the high homology between murine family members when compared to RUNX1T1 isoform 1. Refer to (APPENDIX 3.13) for defining terms used in this table.

3.1.5 *ETO* Gene Family and Acute Myelogenous Leukemia (AML)

The discovery of *RUNX1T1* gene (Runt related transcription factor 1 translocated to 1; cyclin D related) goes back to early nineties where it was first identified in humans as part of the translocation (8; 21) (q22:q22); a commonly seen chromosomal abnormality in AML.[175,176]. AML is a clonal malignant expansion of hematopoietic stem cells in the bone marrow.[177,178] Chromosomal aberrations are commonly seen in AML with frequencies reaching up to 55% in adulthood AML and 76% in childhood AML.[179] One of the most common translocations in AML involve *RUNX1T1* gene in t (8;21) (q22;q22) where *RUNX1-RUNX1T1* (also known as *AML1-MTG8, AML1-ETO*) fusion gene is formed.[180] *CBFA2T3* gene is involved in the translocation t (16; 21) (q24; q22) where *AML1-MTG16* fusion gene is formed.[181] Recently, CBFA2T2 was found to be involved the insertional translocation ins (21; 20)(q22.12;q11.22q13.33).[182]

3.1.6 Murine *Runx1t1* Gene, mRNA and Protein Structure

Murine *Runx1t1* (NCBI Reference Sequence: NC_000070.6) gene was then cloned by Erickson et al in 1994.[183] *Runx1t1*, formerly known as eight twenty one (Eto) gene or myeloid translocation chromosome 8 gene (Mtg8), is located on mouse chromosome 4 band A1 near the centromere (chr4:13,743,302-13,895,056/ + strand, GRCm Build 38) encompassing 151.75 kb of DNA. The gene is composed of 14 exons and 13 introns. Alternative splicing of the first 4 exons results in 3 distinct mRNA variants[184] encoding 3 protein isoforms (Fig3.2) and (APPENDIX 3.11,3.12).

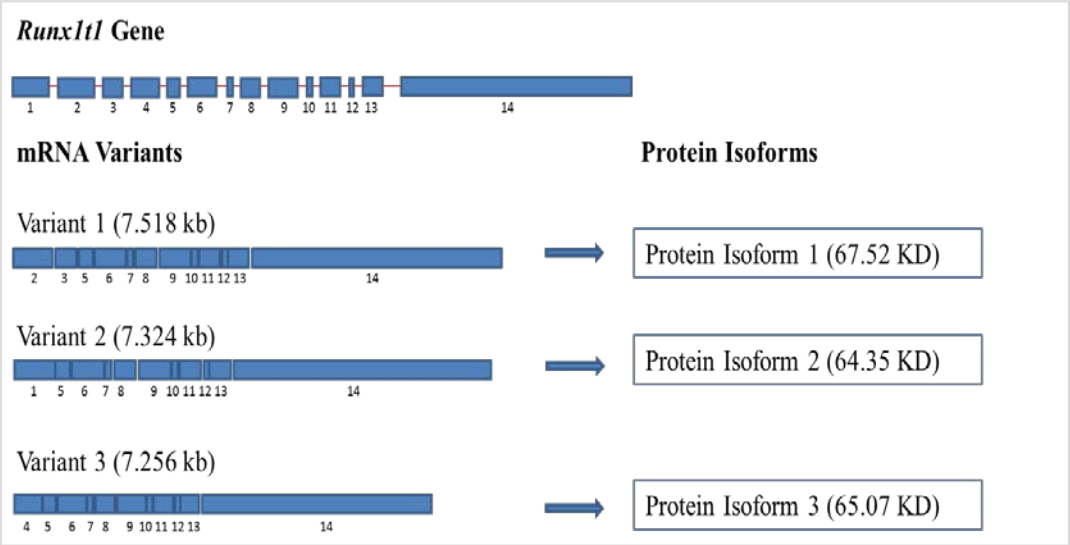


Figure 3.2. Runx1t1 gene, mRNA structure. Alternative splicing of the first 4 exons results in 3 distinct mRNA variants. All variants share exons 5-14. However, exon 2 and 3 are included in variant (1) only, the largest variant. While exons 1 and 4 are included in variants 2 and 3, respectively.

As previously mentioned, RUNX1T1 protein (Fig 3.3) APPENDIX (3.12) is composed of 4 conserved domains known as Nervy Homology Region (NHR) since homology to *Drosophila* protein Nervy.[173,174] NHR1 (also known as TAFH) has homology to a few TATA binding protein associated factors ,such as human hTAF 130 and *Drosophila* TAF110, and is required for subnuclear localization of RUNX1T1.[185] NHR2 domain (also known as hydrophobic heptad repeat region) is required for homodimerization of RUNX1T1, heterodimerization of RUNX1T1 with other ETO family members and binding to the co-repressor protein Sin3a.[186] NHR3 domain is unremarkable but its presence is required for binding N-CoR.[186] NHR4 domain has two zinc-finger motifs. However, there is no evidence of DNA binding activity of RUNX1T1.[187-189] On the other hand, there is evidence that NHR4 is required for RUNX1T1 interaction with other co-repressors such as N-CoR, SMRT and for histone deactylases.[174,186,190]In addition to NHR domains there are several Proline-Serine-Threonine (PST) rich regions especially at the N and C termini of RUNX1T1. RUNX1T1is a phosphorylated protein and kinases phosphorylating

RUNX1T1 were discovered. However, little is known about the kinases and the physiological significance of RUNX1T1 phosphorylation.[180,188,191,192]

Figure 3.3. Domain structure of RUNX1T1 protein.

3.1.7 Orthologs of *Runx1t1*

Runx1t1 has orthologs in human, rat, chimpanzee, monkey, dog, chicken, and zebra fish species (Table 3.4). (Homologene, NCBI database 8.2014)

Species	Percent Identity	Query Coverage M.Musculus (Runx1t1 Isoform 1)
D.rario (Zebra fish)	91%	92%
G.gallus (Chicken)	97%	93%
C. lupus (Dog)	98%	96%
M.Musculus (Mouse)	100%	100%
R.norvegicus (Rat)	99%	95%
H.sapiens (Human)	99%	95%
P.troglodytes (Chimpanzee)	99%	95%
M.mulatta (Monkey)	96%	95%

Table 3.4. Orthologs of *Runx1t1*. *(Homologene, NCBI database 8.2014)*

3.1.8 *Runx1t1* Expression

In mouse species, *Runx1t1* message is detected in whole mouse embryo extracted RNA.[193] RT-PCR demonstrates an expression of *Runx1t1* as early as Theiler stage 1 (E0-0.5) (Gene Expression Database, Aug 2014)[194]. *Runx1t1* expression showed a pattern where the gene expression greatly increased in developing embryo at E7 to E11. The expression was maintained

through E17.[193]Furthermore, *Runx1t1* is expressed in the brain, lung, ovary, testis, heart, spleen based on northern blot[195] and RNA-seq data (EMBL-EBI Expression Atlas, Aug 2014 http://www.ebi.ac.uk/gxa/genes/ENSMUSG00000006586)[196]. The highest expression for *Runx1t1* is found in mouse brain as was demonstrated by northern blot data [183] and RNA-seq data (EMBL-EBI Expression Atlas, Aug 2014 http://www.ebi.ac.uk/gxa/genes/ENSMUSG00000006586)[196]. However, the expression is higher in the developing mouse brain when compared to the adult one as demonstrated by northern blot data.[183] *Runx1t1* message is also moderately expressed in the axial skeleton and in the developing limb buds at around E14 as was shown by in situ hybridization and whole mount in situ hybridization (Gene Expression Database, Aug 2014)[194] and EMAGE gene expression database (http://www.emouseatlas.org/emage/)[197]

3.1.9 Subcellular Location of *Runx1t1*

RUNX1T1 is mostly seen in the nucleus.[173,191,198] Evidence supports the presence of RUNX1T1 in specific nuclear bodies known as ETO-nuclear bodies (ENB).[185] It was demonstrated that a RUNX1T1 (241-280 a.a), located between NHRI and NHRII domains, contain the nuclear localization signal (NLS) and is required for the nuclear entry of RUNX1T1 through an importin dependent pathway.[185] In addition, RUNX1T1 (114-216 a.a), most of which is represented in NHRI domain), was shown to be required but not sufficient for the formation of ENB.[185]

Immunohistochemical staining of RUNX1T1 demonstrated its presence in the cytoplasmic and nuclear compartments of neural cells.[199] For example, the presence of RUNX1T1 in the cytoplasm was demonstrated in Purkinje cells and the nuclei of the neurons in the molecular layer of adult mouse and human cerebellum.[199] Furthermore, RUNX1T1 was

localized in the nucleus of the developing brain of E12.5 mouse embryo. However, primary culture of hippocampal neurons of newborn mice showed both nuclear and cytoplasmic localization of RUNX1T1.[199]Based on the available evidence nuclear and cytoplasmic localization of RUNX1T1 seems to be cell type dependent and developmental stage dependent.[200]

3.1.10 Biological Functions of *Runx1t1*

Cancer

Most of the studies looking into the biological functions of RUNX1T1 has examined its role as part of the fusion protein RUNX1-RUNX1T1 in AML with t (8; 21). The fusion protein seems to regulate hematopoietic stem cell differentiation, proliferation and apoptosis.[201]For example, expression of RUNX1-RUNX1T1 inhibited the differentiation of several hematopoietic stem cell lines such as 32D , L-G, MEL, U937 and K562.[201] Furthermore, the fusion protein was shown to enhance G-CSF dependent cellular proliferation in L-G and 32D myeloid progenitor cell lines[202,203] and induce apoptosis as seen in U937 cell line through JNK signaling pathway.[204] in addition to the pro-apoptotic roles there is evidence that the fusion protein induces the expression of BCL2 (anti-apoptotic factor) as seen in U937 cell line expressing RUNX1-RUNX1T1.[205]

In vivo studies, on the contrary to the aforementioned in vitro studies, demonstrated that the expression of RUNX1-RUNX1T1 is unable to promote leukemogenesis on its own and that additional secondary mutations are needed for AML formation.[201,206]

Other than AML, *RUNX1T1* may have potential roles in certain solid cancers as such as ovarian[207], breast, lung, colorectal and head[208,209] and neck squamous cell carcinoma.[210] In those

studies, *RUNX1T1* was significantly down-regulated[207], up-regulated[210] or had somatic point mutations.[208,209]

Development

N**eurogenesis:** though RUNX1T1 has been implicated in few developmental processes little is known about how RUNX1T1 regulates those processes. For example, RUNX1T1 and other ETO family members have distinct expression patterns in developing nervous system as was demonstrated in mouse and chick developing spinal cords.[211,212]Specifically the expression pattern of RUNX1T1 changes from being expressed ventrally (in motor neurons) to dorsal expression (involving post-mitotic neurons) between E4-E7 embryonic days as was demonstrated in chick embryo spinal cord. In the same study, transcription assay revealed that RUNX1T1 inhibited the transcription of NEUROG2 and ASCL1, neuronal differentiation bHLH transcription factors, however to a lesser extent compared to other ETO family members.[211]

Adipogenesis: RUNX1T1 has been shown to regulate early adipogenesis. RUNX1T1 was shown to inhibit the transcription from C/CAAT-enhancer binding protein /alpha (C/EBPα), a master regulator of adipogenesis, by preventing the binding of C/CAAT-enhancer binding protein /beta (C/EBPβ), to C/EBPα promoter and therefore preventing the activation of the adipogenic program required for the formation of the differentiated adipocytes.[198]

Pancreatic development: recently, RUNX1T1 has been implicated in the normal and ectopic development of pancreatic beta cells of *Xenopus* species.[213]

Gut development: the role of RUNX1T1 in normal gut development was demonstrated in *Runx1t1* knockout murine model.[214] The midgut was absent in 25% of the knockout animals. The rest of the knockouts had an abnormal gut structure where they had thin intestinal walls and

widened lumens. The villi were reduced in length, disorganized, thicker and fewer in number. A closer examination of the knockout model is discussed in the following section.

3.1.11 OVE1328 Transgenic Line and *Runx1t1* Knockout Model

Calabi et al[214] generated a *Runx1t1* knockout model by targeted disruption of exon 2. The disruption of exon 2 was shown to knockout RUNX1T1. Homozygous mutants had other problems in addition to the gut problems just discussed. Homozygous mutant (knockout) mice had reduced survival While 100% of the homozygous mutants (knockouts) viability. While 100% of homozygous mutants survived at birth only ~19% survived till postnatal day 15 (P15). Furthermore, growth impairment was seen in pups surviving beyond postnatal day (P2) where the weights of homozygous mutants were 30-50% less than wild type counterparts. Furthermore, sterility was seen in male homozygous mutants only albeit having normal testicular size and morphology. There were no abnormalities detected in other organs where RUNX1T1 is known to be expressed.

When comparing OVE1328 model to *Runx1t1* knockout model several major differences are seen. These are summarized in (Table 3.5).

Mouse Model	**OVE1328 Model**	**Runx1t1 Knockout Model**
Genetic bacground	FVB/NJ strain	Mixed background (129/Sv and C57BL/6)
Viability	no viable adults (perinatal lethal)	Reduced viability, few mice make it till adulthood
Growth	OVE(Tg/Tg) have growth delay and weigh ~13% less than Wt counterparts	Growth retardation in > 2 days old pups. Homozygous mutants weigh 30-50% less than Wt conterparts.
Cleft palate	Cleft of secondary palate	No evidence
Gut phenotype	Normal gut	Missing midgut (25%) and abnormal structure (75%)

Table 3.5. Comparison of OVE1328 and *Runx1t1* knockout models.

3.1.12 *Runx1t1* Neighborhood: *Gm11823* Gene

Downstream (26.305 kb) of *Runx1t1* resides a long intergenic non coding RNA predicted gene (lincRNA gene) known as *Gm11823*.The gene (75.24 kb) is located on the minus strand at 13,919,954-13,995,193 bp , GRCm build 38. *Gm11823* (also known as *D930021N14,RP23-134H12.2*) has 3 exons and 2 introns and is predicted to encode a single transcript (2.278 kb). Little is known about the function and expression of *Gm11823*. Till present no knockout mouse model is available for *Gm11823*. (NCBI data base, Gene ID: 329797 and Ensembl data base, release 76, Aug 2014)[215]

3.1.13 *Gm11823*: Potential Roles as a Long Intergenic Non-Coding RNA Gene

Long non coding RNAs (LncRNA) are RNA species that are more than 200 nucleotides (Nt) in size with no protein coding ability.[216-218] This size limit distinguishes lncRNA from small RNAs (such as miRNA, siRNA, snRNA, piwiRNA) based on RNA preparation methods.[216-218] There are some caveats with this definition as some lncRNA have both coding and non-coding abilities.[219-221] Furthermore, many of lncRNAs are transcribed by RNA polymerase II (pol II) and might be polyadenylated.[222]With regard to their cellular location, lncRNAs are detected in both the nuclear and cellular compartments.[216]There are around ~14,000 and ~5,000 lncRNA genes in humans and mice species (GENCODE data base version 20/2013, M3/2013), respectively.

LncRNA can be categorized into 5 main groups[223] based on the genomic locus from which these RNAs are transcribed with respect to the surrounding genomic markers: (1)Stand-alone RNA (also known as long intergenic or intervening non coding RNA (lincRNA)). This group of RNA is transcribed from genomic loci that are outside the boundaries of protein coding

genes. They are transcribed by pol II and polyadenylated and can be spliced. The average size for such RNAs is 1 kb. *Gm11823* gene belongs to this group. (2) Natural antisense transcripts (NAT) are RNA molecules transcribed from the opposite DNA strand of coding and non-coding genes. [224] The NAT are usually clustered around the 5' end or the 3' end of the sense transcript. Splicing and polyadenylation is seen few NAT. (3) Transcripts from pseudogenes: pseudogenes are genes with structural similarity to other coding genes with mutations that render them unable to produce fully functional proteins. Those genes can be produced by gene duplication or retrotransposition.[225]About 2-20% of those genes have transcripts. (4). Long intronic non coding RNA: those are RNAs transcribed from introns of annotated genes. (5) Divergent, promoter associated and enhancer RNA: group is sense or antisense RNA species (20-2500 nt in size) that are transcribed from regions close to the transcription start sites such as promoter and enhancer regions.[223]

Little is known about the biological functions of lncRNA.[219] LncRNA have been implicated in X chromosome inactivation such as (*XIST*), genomic imprinting such as (*AIRN*), neuronal development such as (*EVF2*) and in many types of cancer including breast, liver, pancreatic, colorectal, prostate such as (*MALAT1*, *HOTAIR).*[216,223] The mechanism of action of lncRNA is not fully understood but there is evidence that lncRNA can function in cis or trans.[219] Epigenetic (specifically chromatin modulation), transcriptional and posttranscriptional regulation of gene expression is seen with different lncRNAs (Fig 3.4).[226]

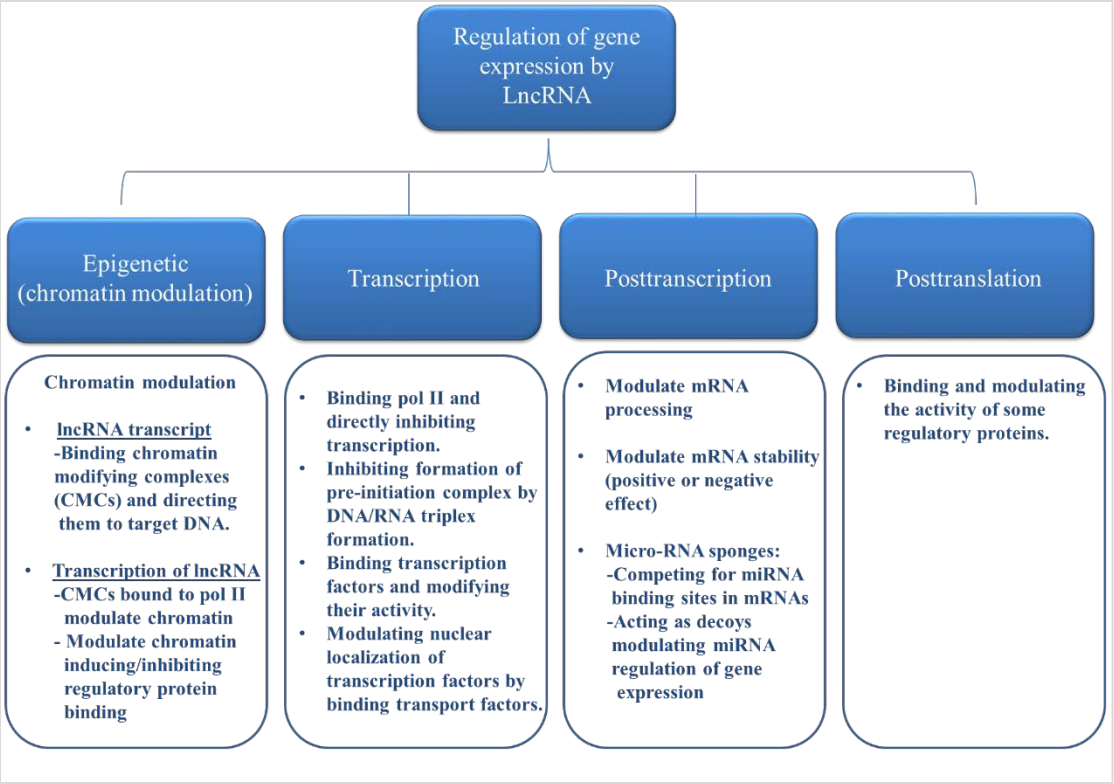


Figure 3.4. Gene regulation by lncRNAs. Gene regulation by lncRNAs can take place at different levels; epigenetics, transcriptional, posttranscriptional and posttranslational levels. Figure content adopted from Geisler et al.[226]

3.1.14 *Runx1t1*, *Gm11823* and Their Alignment to Human Genome

Runx1t1 human ortholog (*RUNX1T1)* is present on chromosome 8q22 (chr8: 91,954,967-92,103,226/ minus strand). Furthermore, Gm11823 gene aligns as well to part of another lincRNA gene present on chromosome 8 (chr8: 91,542,924-91,907,619/ minus strand) and is known as RP11-122C21.1, ENSG00000253901 (Ensembl Database, release 76, Aug 2014)[215] (Fig 3.5). Closer examination of these orthologs reveals that the order of these genes is conserved between human and mouse chromosomes (synteny) (Ensembl Database, release 76, Aug 2014)[215] as shown in (Fig 3.6).

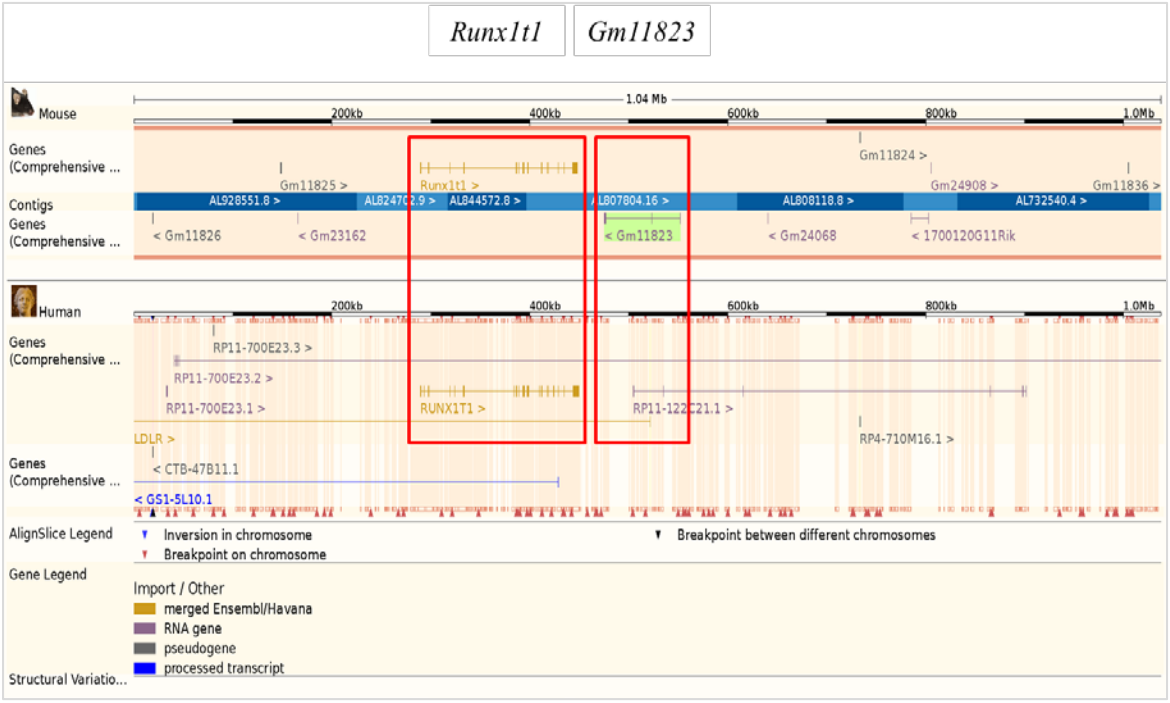


Figure 3.5 Graphical alignment of Gm11823 (lincRNA) to RP11-122C21.1 human (lincRNA) on chromosome 8. RP11-122C21.1 is downstream of RUNX1T1 gene. *Screenshot adapted with permission from Ensembl database (release 76, August 2014).* *http://Oct2014.archive.ensembl.org/info/website/archives/index.html*

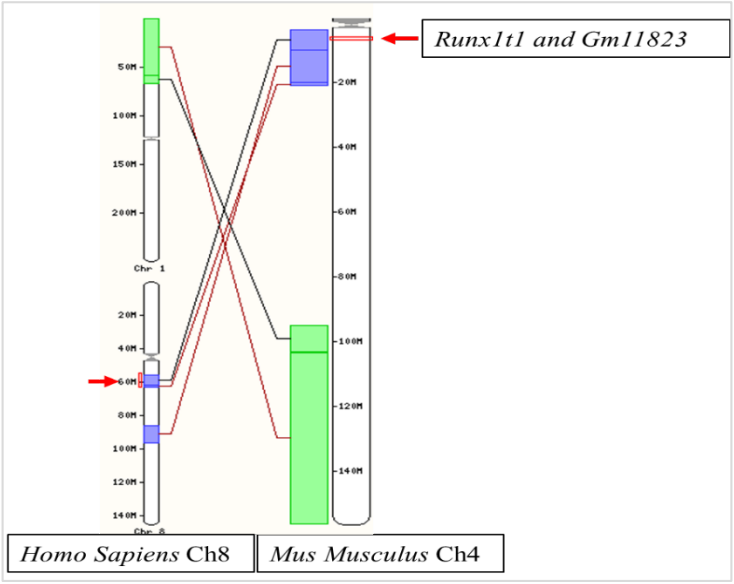


Figure 3.6. Syntenic regions between mouse chromosome 4 and human chromosome 8. *Screenshot adapted with permission from Ensembl database (release 76, August 2014).* *http://Oct2014.archive.ensembl.org/info/website/archives/index.html*

3.2 Materials and Methods

RNA Extraction

Total RNA was extracted from Wt, OVE1328 (Tg/+) and OVE1328 (Tg/Tg) E18 embryo heads using RNeasy Midi Kit (Qiagen Inc., Valencia, CA, USA) according to manufacturer recommendations. Extracted RNA was dissolved in RNase free water. RNA quality (RIN cutoff value = 9) and quantity was assessed using Agilent Bioanalyzer 2100 (Lineberger Comprehensive Cancer Center Genomics Core Facility) and NanoDrop 1000 ® spectrophotometer NanoDrop Technologies Inc., Wilmington, DE, USA), respectively. RNA samples were stored at -80 °C.

Conventional PCR

RNA samples were subjected to cDNA synthesis using (High Capacity cDNA reverse transcription kit, Applied Biosystems, Grand Island, NY, USA). Custom oligonucleotide primers (APPENDIX 3.1-3.10) were designed using Primer 3® software[227] and generated by (Invitrogen, Grand Island, NY, USA). PCR products were generated (APPENDIX 3.9) using (Go Taq ® PCR Core System I, Promega, Madison, WI, USA) and electrophoresed in 2% agarose gel at 100 Volts (V) for 2-3 hours. Fidelity of PCR products were confirmed by DNA sequencing (Eton Bioscience Inc., Durham, NC, USA).

RNA-Seq

Total RNA was extracted from E18 Wt (n=3) and OVE1328 (Tg/Tg) (n=3) embryo heads. RNA was checked for quality and quantity as previously mentioned. cDNA libraries were constructed for RNA samples after rRNA depletion (RiboZero Magnetic kit, Epicenter, Madison,

WI, USA) at the High Throughput Sequencing Facility (HTSF)/UNC. The libraries (SMARTer Universal Low Input RNA Library Prep Kit, Mountain View, CA, USA) were subjected to 100 bp paired end sequencing using Illumina platform (Illumina HiSeq 2500). Samples were multiplexed in a single flow cell. Reads were aligned to mm10 genome assembly using TopHat2 (2.0.11).[228] To determine the transgene insertion site, the following reads from the three OVE1328 (Tg/Tg) mice were pooled: unmapped reads, reads aligning to a region containing *Runx1t1* and *GM11823* (chr4:13,743,102-13,995,393), and reads aligning to a region containing *Tyr* (chr7:87,427,205-87,493,611) (aligned reads extracted using SAMtools 0.1.19)[229]. De novo assembly of the selected reads was conducted using Velvet 1.2.10[230] and Oases 0.2.08[231], merging transcripts built using 27- and 29-mers. The resulting contigs were aligned (BLAST+ 2.2.29)[232] to a custom BLAST database consisting of mm10 genome assembly and sequences known to be contained in the transgene construct. The BLAST results were filtered to find contigs that align to part of the transgene construct and at least one other sequence (from mm10 genome assembly or transgene). Only 2 contigs had part of the transgene and *Runx1t1* (contig 15360) or the transgene and *Gm11823* (contig 4450). PCR was used to validate the contigs using custom oligonucleotide primers (APPENDIX 3.7 & 3.8).

Northern Blot

Total RNA (30ug) was electrophoresed in formaldehyde-agarose gel (1.8% -1.2%, respectively) pre-stained with 0.4% Gel Red (TM) Nucleic Acid Gel Stain (Biotium Inc., Hayward, CA, USA). The samples were allowed to run for 4-4.5 hours at 45-50 V. Electrophoresed RNA was subjected to alkaline treatment (0.05 N sodium hydroxide (NaOH)) for 20 min,[233] transferred to uncharged nylon membrane (MagnaGraph, M.S.I., Westboro, MA, USA) in 20X saline-sodium citrate (SSC) buffer for 4 hours. Pre-hybridization/ hybridization (at

62 °C) were done using DIG Easy Hyb (Roche Diagnostics Corporation, Indianapolis, IN, USA) according to manufacturer instructions. Digoxigenin (DIG)-labelled cDNA probes (1-2ul/ml hybridization buffer) were used for hybridization. The DIG-labelled probes were generated using Go Taq ® PCR Core System I (Promega, Madison, WI, USA) and Digoxigenin-11-dUTP (Roche Diagnostics Corporation, Indianapolis, IN, USA). *Runx1t1* cDNA custom primers (5-16) (APPENDIX 3.1& 3.2) were used for probe generation. Membranes were washed at high stringency conditions (0.1X SSC, 1% w/v sodium dodecyl sulfate (SDS) at 62 °C). Chemiluminescent detection (CSPD, Roche Diagnostics Corporation, Indianapolis, IN, USA) was used according to manufacturer instructions. Blots were exposed to film (Amersham Hyperfilm ^{TM} ECL, GE Healthcare, Pittsburgh, PA, USA) for 4 hours. The membranes were then stripped and reprobed with DIG- labelled β-actin RNA probe (Roche, Roche Diagnostics Corporation, Indianapolis, IN, USA) according to manufacturer instructions.

Western Blot

Wt, OVE1328 (Tg/+) and OVE1328 (Tg/Tg) embryo heads (with cut snouts and mandibles) were homogenized with SDS buffer (0.1% w/v SDS, 2.42% tris-base, 3.5% sodium chloride (NaCl), 0.117% ethylenediaminetetraacetic acid (EDTA)) and using a tissue homogenizer (Tissue Tearor, Biospec Products Inc., Bartlesville, OK, USA). Protein concentration was determined using BCA Protein Assay Kit (Pierce/Thermo Scientific, Rockford, IL, USA) according to manufacturer instructions. Protein samples (20-25µg) were resolved on NuPAGE 4-12% Bis-Tris gel (Novex®, Life Technologies, Grand Island, NY, USA) and transferred on to a (0.45 µm) nitrocellulose membrane (Protran BA 85 Nitrocellulose, GE Healthcare, Pittsburgh, PA, USA) at 250V, for 2 hours. The membranes were blocked with 5% skim milk + 2% goat serum (Santa Cruz Biotechnology Inc., Dallas, TX, USA) in phosphate

buffered saline with tween (PBST) (0.1% Tween 20 ®) at room temperature for 1 hour. The blots were then incubated with *Runx1t1* primary antibody (1:1000, Cell Signaling cat# 4498, Beverly, MA, USA) in 5% skim milk/PBST for 1 hour/room temperature. The blots were washed 3-4 times each 5 minutes (min) in PBST and incubated with a secondary antibody (1:10,000, Promega cat# W401B, Madison, WI, USA) in PBST(0.1% Tween 20 ®) for 1 hour/ room temperature. Chemiluminescent detection (SuperSignal West Pico Chemiluminescent Substrate, Pierce/ Thermo Scientific, Rockford, IL, USA) was used according to manufacturer instructions. Blots were exposed to film (Amersham Hyperfilm ^{TM} ECL, GE Healthcare, Pittsburgh, PA, USA) for 5min. Membranes were stripped and reprobed with β-Actin antibody (1:3000, Abcam cat# ab8227, Cambridge, MA, USA) according to standard protocols.

3.3 Results

Disruption of the *Runx1t1* Gene by Transgene Integration in the OVE1328 Mouse Line

Deletion Mutation of the 3' End of *Runx1t1* Gene

Prior studies of the OVE1328 mouse line have mapped the transgene integration site to chromosome 4 band A2.[159] Genomic SNP microarrays of Wt and OVE1328 (Tg/Tg) embryos identified a series of no calls (absence SNP hybridization signals) involving OVE1328 (Tg/Tg) embryos but not Wt embryos indicating that the 3' end of *Runx1t1* gene was disrupted (Fig 3.1).[159]

To validate this, a set of primers were designed to amplify exon-intron junctions and exons of the 3'end of *Runx1t1* starting at intron 11-exon12-intron 12 junction of *Runx1t1*. Since the no calls in the microarray were detected from SNPs in intron 11, the primers were designed starting at intron11-exon12-intron12 junction (APPENDIX 3.3 and 3.4). Using OVE1328 (Tg/Tg) gDNA, no amplification products were seen with primers amplifying part of *Runx1t1* exon 13 (primer pair 132), exon 14 (primer pairs 143-147), or primers amplifying intron 12-exon13-intron13 (primer pair 131) or intron 13-exon14 junctions (primer pairs 141,142) compared to Wt gDNA (Fig 3.7), (Fig 3.8).

Figure 3.7. Primer map for *Runx1t1* exon-intron junction primers. Key: (*) primer pairs amplifying intron 11-exon12-intron 12 junction. (**) primer pairs amplifying intron 12-exon13-intron 13 junction. (***) primer pairs amplifying intron 13-exon14 junction. (****) primer pairs amplifying exon14 up to Chr4: 13,891,269 (GRCm38). Red shaded region represent parts of *Runx1t1* gene that were not successfully amplified using the respective primer pairs with OVE1328 (Tg/Tg) gDNA.

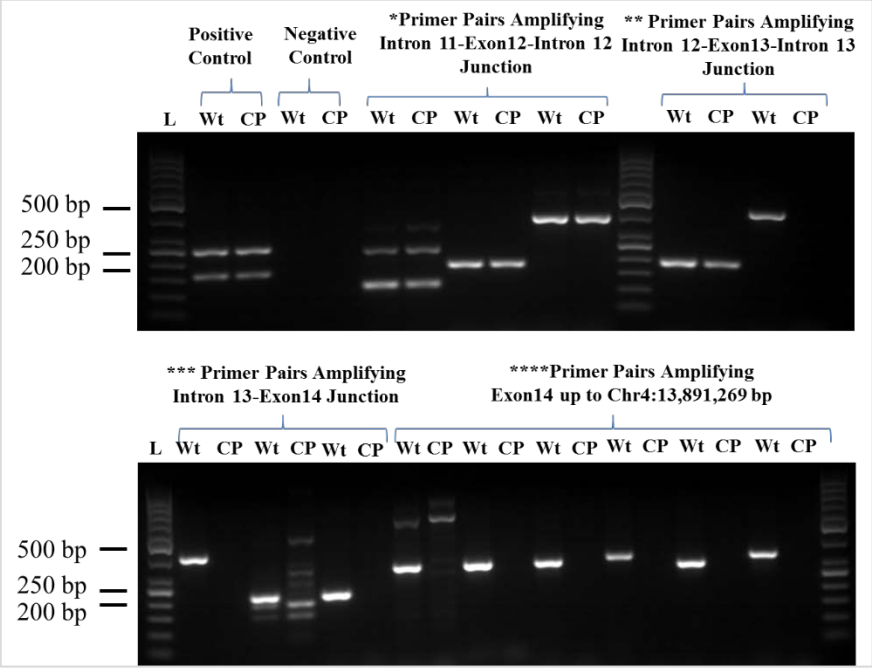


Figure 3.8. Conventional PCR data of exon-intron junction amplifications. PCR data demonstrate failure of amplification of exon13, exon14 (up to chr4:13,891,269), intron 12-exon13 and intron 13-exon14 junctions with OVE1328 (Tg/Tg) gDNA. Key: (*) primer pairs amplifying intron 11-exon12-intron 12 junction. (**) primer pairs amplifying intron 12-exon13-intron 13 junction. (***) primer pairs amplifying intron 13-exon14 junction. (****) primer pairs amplifying exon14 up to Chr4: 13,891,269 (GRCm38). L: molecular weight marker (50bp DNA ladder).

Altered *Runx1t1* mRNA in the OVE1328 Mouse Line

To examine the effect of the deletion mutation on the mRNA, custom primers were designed to amplify overlapping regions of the longest mRNA variant of *Runx1t1* (variant 1) (Fig 3.9) and APPENDIX 3.1, 3.2. Interrogation of *Runx1t1* transcript (variant 1) revealed that (Tg/Tg) and Wt transcripts share the 5' end of the message but not the 3' end (Figure 3.10). Primer pairs amplifying exons 2, 3,5,6,7,8,9,10,11 and12 successfully produced PCR products with both OVE1328 (Tg/Tg) and Wt cDNA. However, primers amplifying exons 13 and 14 failed to amplify products with OVE1328 (Tg/Tg) samples indicating that the intragenic deletion

of *Runx1t1* gene resulted in the disruption of the 3'end of *Runx1t1* transcript; a transcript lacking exons 13 and 14 respectively.

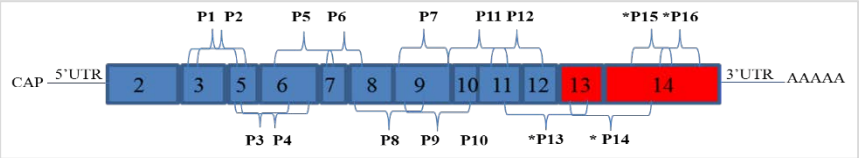


Figure 3.9. Primer map for *Runx1t1* message (variant1). The primer pairs amplify overlapping regions of the message. (P1-16): primer pairs 1-16. (*): primer pairs that failed to produce PCR products with OVE1328 (Tg/Tg) cDNA only. Red shaded region represents the region being amplified by the asterisk labeled primer pairs.

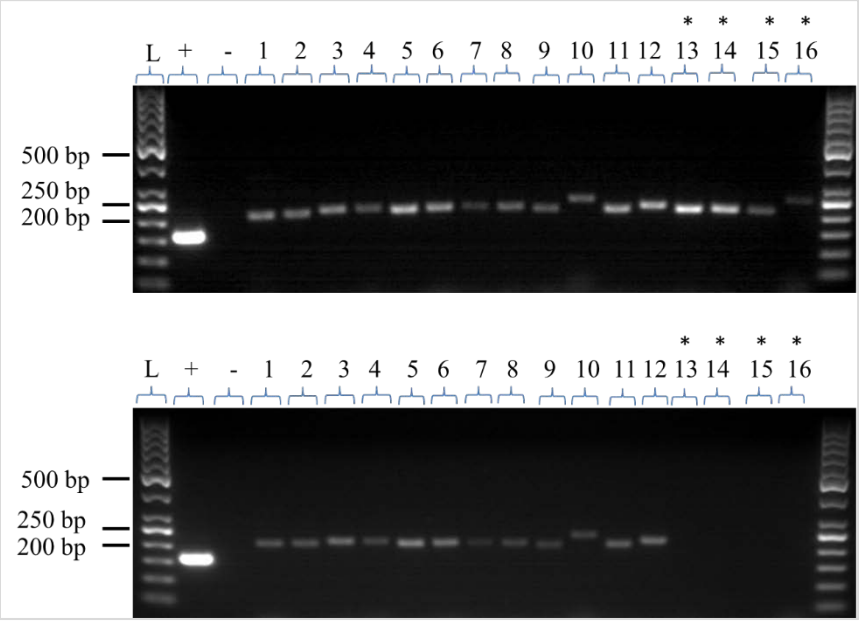


Figure 3.10. Conventional PCR data of *Runx1t1* message. The data shows failure of amplification with OVE1328 (Tg/Tg) cDNA upon using primer pairs amplifying part of exon 13 and exon14. (*): primer pairs that failed to produce PCR products with OVE1328 (Tg/Tg) cDNA. (+): positive control. (-): negative control. (L): molecular weight marker (50bp DNA ladder).

Determination of Transgene Flanking Sequences: *Runx1t1*-Transgene Junction

RNA-seq data confirmed previous PCR findings. Close examination of the RNA-seq data using UCSC genome browser (UCSC)[234] showed an abrupt loss of reads, in general, retrieved from the 3'end of *Runx1t1* message of OVE1328 (Tg/Tg) RNA (n=3). Specifically, no reads were retrieved beyond chr4: 13,876,840 bp (intron 12) of OVE1328 (Tg/Tg) RNA compared to Wt group (n=3) (Fig 3.11).

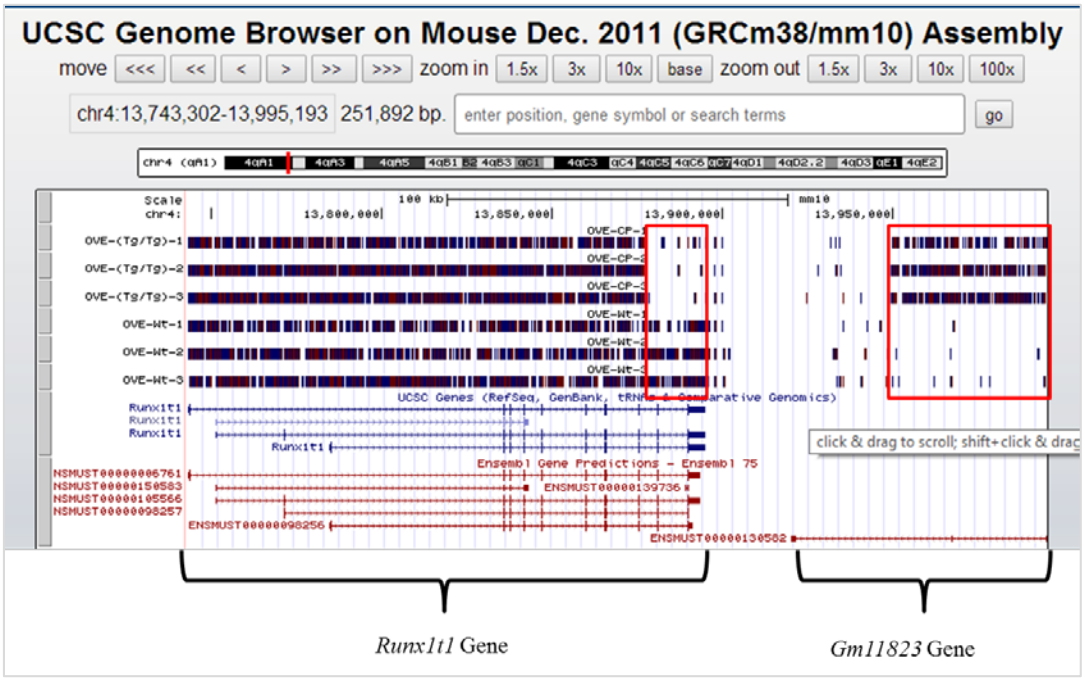


Figure 3.11. RNA-Seq data. Reads from OVE1328 (Tg/Tg) and Wt RNA (n=3/genotype) mapped to GRCm38/mm10 mouse assembly showing *Runx1t1* and a downstream neighborhood gene *Gm11823*. Red rectangles represent regions where a difference in read mapping is detected between(Tg/Tg) and Wt groups. *Screenshot modified from UCSC genome browser (http://genome.ucsc.edu), assembly Dec. 2011(GRCm38/mm10).*

In addition, the transgene insertion site at *Runx1t1* was determined by bridging paired end reads that mapped to *Runx1t1*and those reads that mapped to the transgene sequence. One contig (contig-15360) 437 bp mapped to *Runx1t1* chr4:13876578-13876840bp and *Tyrosinase (Tyr)*

complement 2108-2284bp (Fig 3.12). To validate the contig, custom primers were designed to flank *Runx1t1*-transgene complex junction (Fig 3.12) and (APPENDIX 3.7, 3.8). The primer pair successfully produced a PCR product with OVE1328 (Tg/Tg) gDNA (n=3) but not with Wt gDNA (n=1). Furthermore, the primer pair successfully amplified a product with OVE1328 (Tg/Tg) cDNA (gDNA free) (n=3) but not with Wt cDNA (n=3) (Fig3.13). This indicated that *Runx1t1* message contained part of the transgene in OVE1328 (Tg/Tg) embryos.

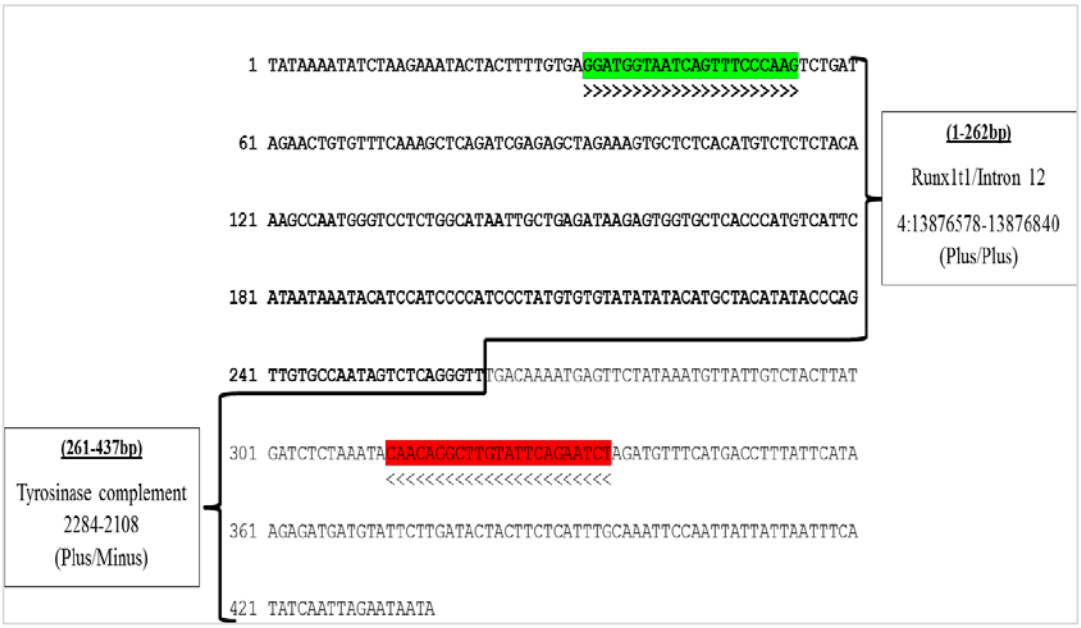


Figure 3.12. Contig 15360 primer map. The contig is 437 bp in size where 1-262 bp correspond to intron 12 of *Runx1t1* chr4: 13,876,578-13,876,840 bp, respectively. The next 261-437 bp correspond to *Tyr* complement (part of the TYBS transgene complex) 2284-2108 bp, respectively. Bold sequence: mouse *Runx1t1* sequence (intron12). Grey sequence: *Tyr* complement of the transgene complex. Green shaded sequence: forward primer. Red shaded sequence: reverse primer.

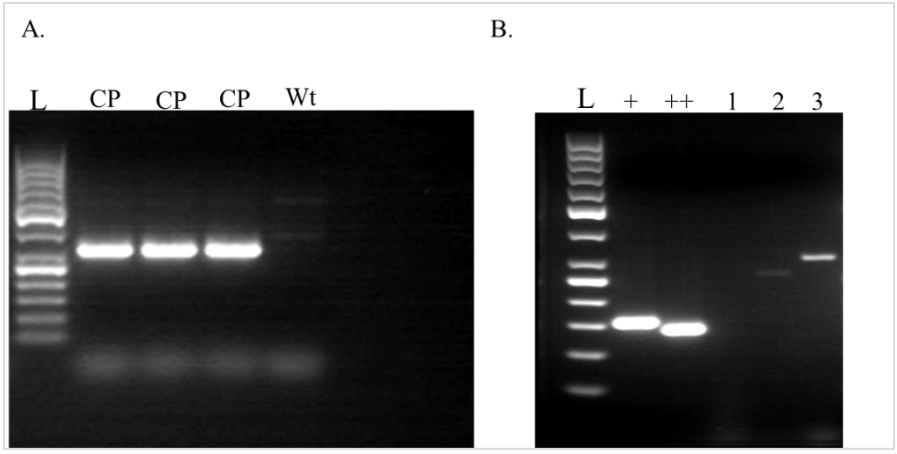


Figure 3.13. Conventional PCR data of 15360 contig primer pair**.** A. Successful amplification of contig (15360) when using 15360 contig primer pair and OVE1328 (Tg/Tg) gDNA (n=3). No amplification products were detected with Wt gDNA (n=1). B. Amplification products of (4550 and 15360) contig primer pairs (lanes 2 and 3, respectively) with OVE1328 (Tg/Tg) cDNA (gDNA free) (n=3). (+): positive control (β-actin). (++): positive control Gapdh. Lane1: OVE1328 (Tg/Tg) RNA with 4550 contig primer pair. (L): molecular weight marker (50bp

The data presented above showed the integration site of the transgene complex in OVE1328 line. The integration site was mapped to *Runx1t1* gene (intron 12, chr4:13,876,840bp). The integration resulted in a deletion mutation of *Runx1t1* gene starting at intron 12 (chr4:13,876,841bp). Furthermore, there is compelling evidence that the transgene integration affected the splicing of intron 12. The PCR data with OVE1328 (Tg/Tg) cDNA (Fig3.13.B /lane3) indicated the presence of part of intron 12 in *Runx1t1* message. For further validation, previous exon-intron primer pairs amplifying exon12 junctions (Fig 3.14) were used this time with OVE1328 (Tg/Tg) cDNA. The data demonstrated that *Runx1t1* message not only contained part of intron 12 but also part of the 3' end of intron 11 (chr4 13,875,423-13,875,476) as seen with primer pairs 121 and 122 (Fig 3.15).

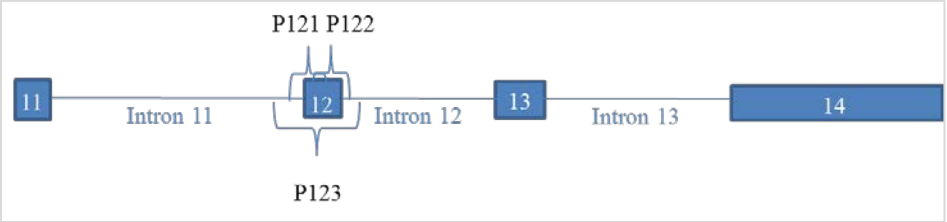


Figure 3.14. Map for exon–intron junction primer pairs 121,122 and 123.

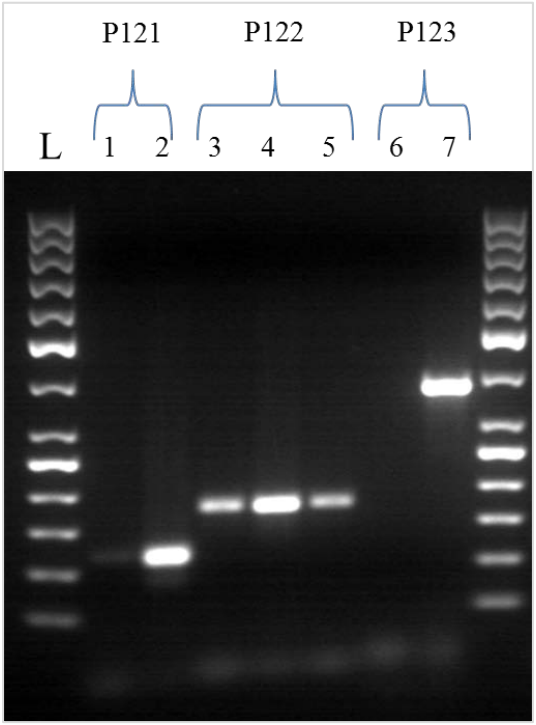


Figure 3.15. Interrogation of exon 12- intron 12 junction of *Runx1t1*. PCR data demonstrates amplification products with OVE1328 (Tg/Tg) cDNA (gDNA free) (lanes 1, 3, 6) and gDNA (lanes 2, 4, 7) when using 121 and 122, 123 primer pairs. Amplification products with OVE1328 (Tg/Tg) cDNA are only seen 121and 122 primer pairs. Lane5: Wt cDNA (non-gDNA

Closer examination of the RNA-seq data in (Fig 3.11) revealed that the transgene integration not only disrupted *Runx1t1* gene but also a downstream gene, *Gm11823* which normally resides 26.305 kb from *Runx1t1*. *Gm11823* is located on chr4:13919954-13995193 bp and predicted to be transcribed on the opposite strand from *Runx1t1*. *Gm11823* is predicted to encode a large intergenic non-coding RNAs (lincRNAs). There are a number of reads mapping to *Gm11823* gene from OVE1328 (Tg/Tg) RNA (n=3) but only a few sporadic reads mapping to Wt RNA. The interrogation of the effect of the transgene complex on *Gm11823* gene will be discussed in the *GM11823* section.

Absence of *Runxt1t1* Message in OVE1328 (Tg/Tg) Embryos; Northern Blot Data

To further investigate the findings of PCR and RNA-seq data, northern blot was done to interrogate the effects of the transgene integration on *Runx1t1* message in OVE1328 (Tg/Tg) embryos. *Runx1t1* message was detected with Wt and OVE1328 (Tg/+) RNA at the expected molecular weight. No message was detected in OVE1328 (Tg/Tg) embryos (Fig 3.16).

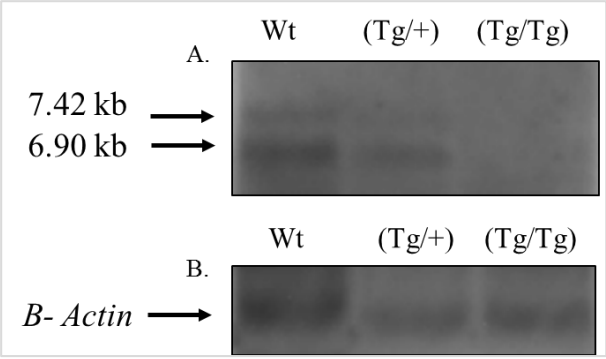


Figure 3.16. Northern blot data. A. Blot probed with *Runx1t1*-DIG labeled PCR probe. No message is detected in OVE1328 (Tg/Tg) embryo heads compared to Wt and OVE1328 (Tg/+) groups. B. Blot probed with *β-actin*-DIG labeled RNA probe.

No Differences in RUNX1T1 Protein among OVE1328 Transgenic Line and Wt Embryos

Interrogation of RUNX1T1 in OVE1328 (Tg/Tg) revealed no differences from Wt RUNX1T1 which showed two isoforms at the expected molecular weights (Fig 3.17). The size and the number of isoforms detected in Western blot were consistent with previous data presented by Calabi et al [214].

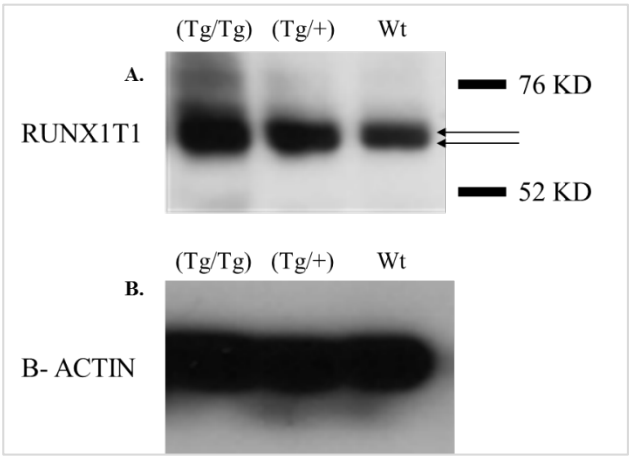


Figure 3.17. Western blot data. A. Blot probed with anti-RUNX1T1 (Cell Signaling 4498) antibody. No difference detected between all genotypes in RUNX1T1 immunoreactive band. B. Blot probed with anti β-ACTIN antibody.

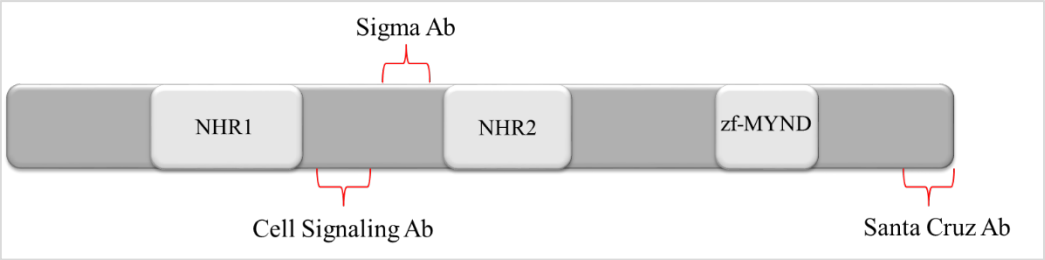


Figure 3.18. Epitope map for RUNX1T1 antibodies used for Western blotting. Cell Signaling and Sigma antibodies bind to an epitope close to the N-terminal end. Santa Cruz antibody recognizes an epitope close to the C-terminal end of RUNX1T1.

The antibody (Cell Signaling 4498) used for probing RUNX1T1 in (Fig 3.17) recognizes an epitope close to the N-terminal end of RUNX1T1 (Fig 3.18) and (APPENDIX 3.10). To further validate the previous findings with Cell Signaling Ab, other blots were probed with 2 different anti-RUNX1T1 antibodies those were: (SIGMA C5616) and (Santa Cruz sc-9737). As seen in (Fig 3.18) and the antibodies recognize an epitope in proximity to the N-terminal and C-terminal ends of RUNX1T1 (APPENDIX 3.10), respectively. Interestingly, blots with those antibodies did not reveal any differences in RUNX1T1 between all genotypes. Furthermore, blots probed with each of the aforementioned antibodies did not show any difference in the banding pattern between Wt, OVE1328 (Tg/+) and OVE1328 (Tg/Tg) samples.

Disruption of the *Gm11823* Gene by Transgene Integration in the OVE1328 Mouse Line

Transgene Integration Induces *Gm11823* gene expression

RNA-seq data provided evidence that *Gm11823* was another gene to be affected by the TYBS transgene integration. A number of reads mapped to *Gm11823* gene from OVE1328 (Tg/Tg) RNA but only a few sporadic reads mapped to Wt RNA. The reads retrieved from OVE1328 (Tg/Tg) RNA covered the most 5’ end of *Gm11823* gene up to intron 2 (ch4:13,949,713 bp). Analysis of RNA seq data yielded one contig (contig 4550) that had part of *GM11823* (chr 4:13,949,452-13,949,713) and part of the transgene complex corresponding to MSCV-neo vector sequence 2785-2862 bp and the 3’enhancer sequence (3E:3563-3655) bp (Fig.3.19).

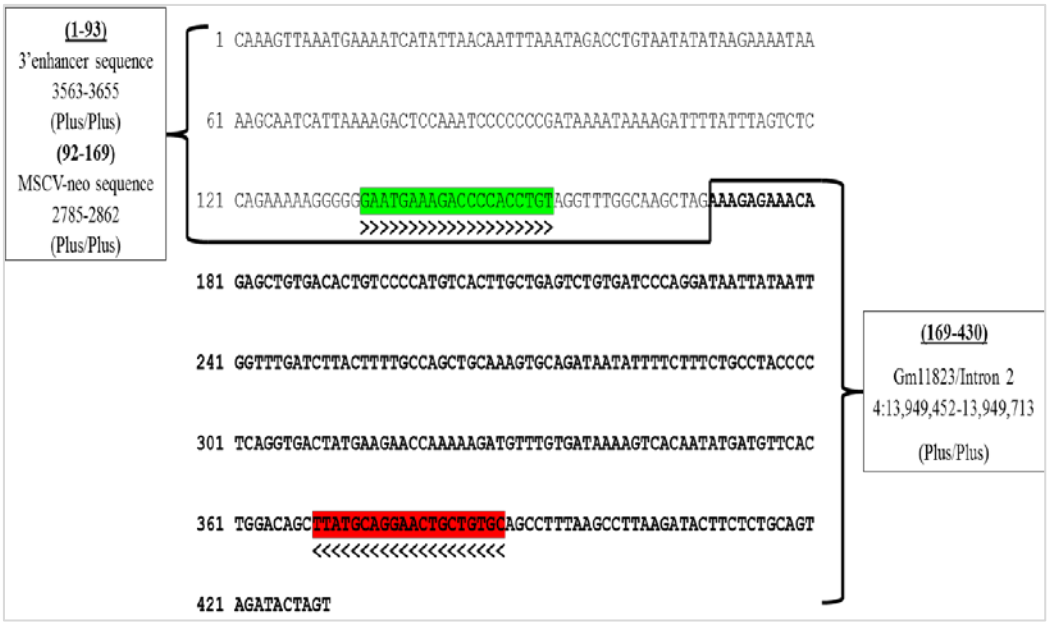


Figure 3.19. Contig 4550 primer map. The contig is 430 bp where 1-93 bp correspond to the 3' enhancer sequence 3563-3655 bp of the TYBS transgene complex, respectively. The next 92-169 bp correspond to the MSCV-neo vector sequence 2785-2862 bp of the TYBS transgene complex, respectively. The last 169-430 bp correspond to intron 2 chr 4:13,949,452-13,949,713 bp, respectively. Bold sequence: mouse *Gm11823* sequence (intron 2). Grey sequence: TYBS transgene complex related sequences. Green shaded sequence: forward primer. Red shaded sequence: reverse primer.

RNA-seq data from OVE1328 (Tg/Tg) RNA indicated that *Gm11823* gene was disrupted by the TYBS transgene complex. The disruption induced the expression of *Gm11823* in OVE1328 (Tg/Tg) RNA. This is in contrast to the RNA-seq data from Wt RNA where *Gm11823* was not expressed. To further interrogate these observations, custom primers were designed to amplify *Gm11823* RNA (Fig 3.20) and (APPENDIX 3.7, 3.8).

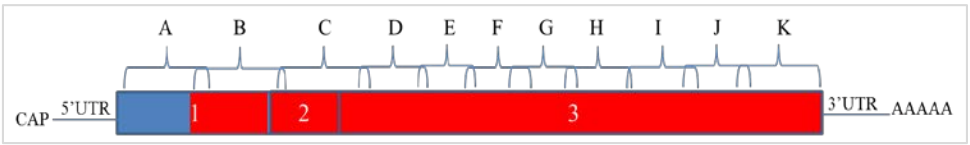


Figure 3.20. *Gm11823* primer map. Custom primer pairs were designed to amplify overlapping regions of *Gm11823* message. Red shaded areas represent primer pairs that did not amplify a product with OVE1328 (Tg/Tg) cDNA and gDNA.

The primers were tested with gDNA of Wt and OVE1328 (Tg/Tg) samples. All primer pairs (except primers B and C which amplify exon-exon junctions) amplified PCR products with Wt gDNA (Fig 3.21.A). On the other hand, only one primer pair (A), which amplifies part of exon1, produced a PCR product with OVE1328 (Tg/Tg) gDNA. No PCR products were detected when primer pairs amplifying exon 3 of *Gm11823* (E-K) were used with OVE1328 (Tg/Tg) gDNA (Fig 3.21.B). This was consistent with the previous findings of RNA-seq data where no reads mapped beyond intron 2 (ch4:13,949,714 bp) of *Gm11823* (Fig3.11).

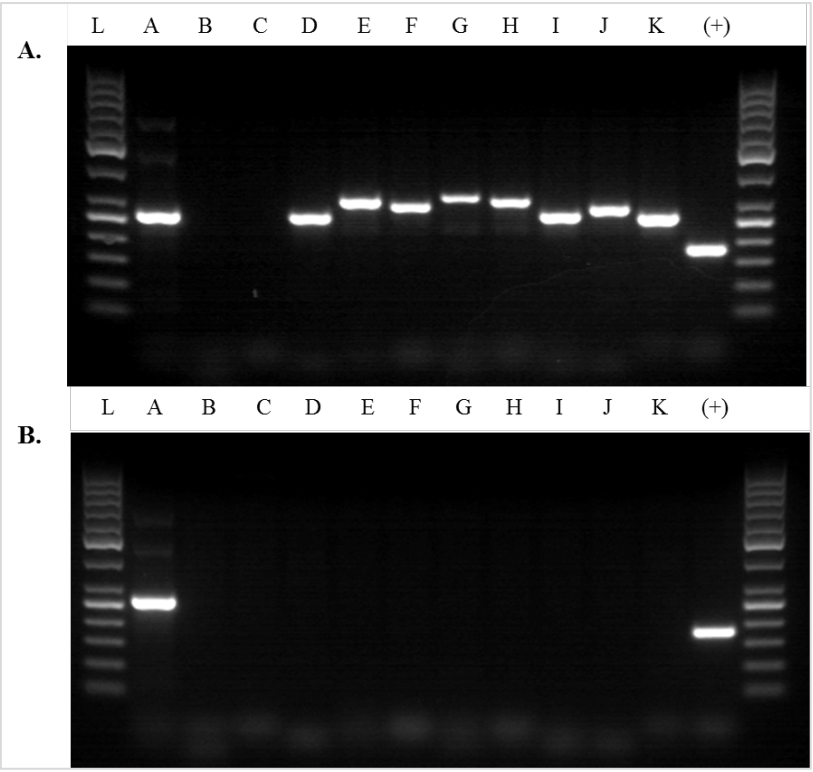


Figure 3.21. Conventional PCR data of Gm11823 genomic locus. Amplification products of *GM11823* custom primers tested with Wt and OVE1328 (Tg/Tg) gDNA. A. Amplification with Wt gDNA (n=3).B. Amplification with OVE1328 (Tg/Tg) gDNA (n=3). (A-K): GM11823 custom primer pairs. (+): positive control mouse *β-Actin* primer (SA Biosciences, Valencia, CA, USA). (L): molecular weight marker (50bp DNA ladder).

Moreover, PCR data further confirmed the RNA-seq findings with respect to *Gm11823* expression (Fig 3.22). No amplification products were detected when using *GM11823* primer pairs with Wt cDNA (Fig 3.22.A). On the other hand, although in general no PCR products were detected when using *Gm11823* custom primers with OVE1328 (Tg/Tg) cDNA, one primer pair (A) consistently yielded a PCR product at the expected size with OVE1328 (Tg/Tg) cDNA (Fig 3.22.B). The product corresponds to part of exon 1 of *Gm11823* gene and was verified by direct sequencing.

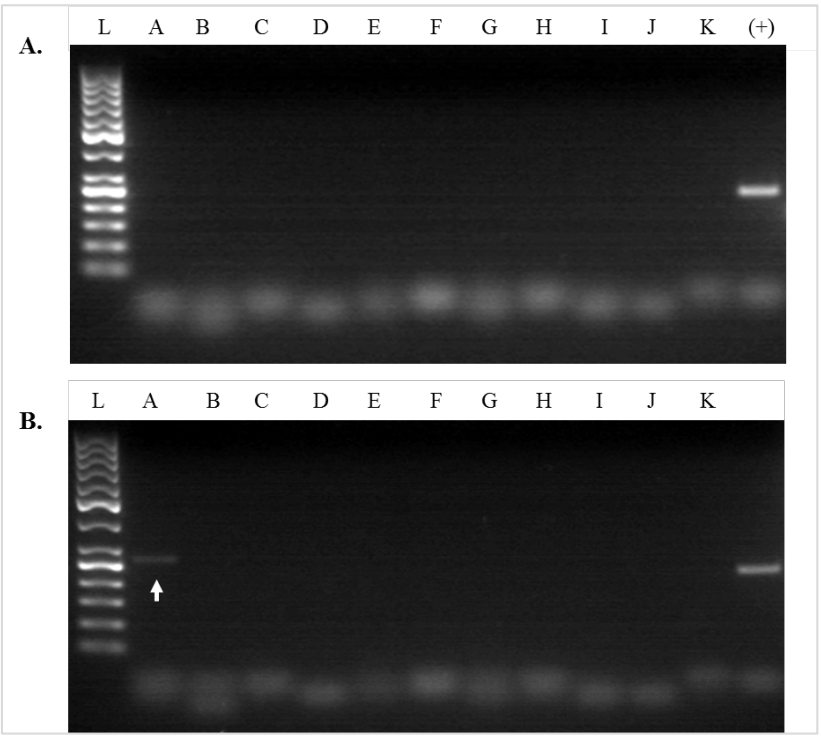


Figure 3.22. Conventional PCR data of Gm11823 message**.** Amplification products of *GM11823* custom primers with Wt and OVE1328 (Tg/Tg) cDNA. A. Amplification with Wt cDNA (n=3). B. Amplification with OVE1328 (Tg/Tg) cDNA (n=3). White arrow: amplification product of primer (A) which corresponds to part of exon 1 of *Gm11823.* (A-K): *GM11823* custom primer pairs. (+): positive control (Runx1t1 cDNA custom primer 3). (L): molecular weight marker (50bp ladder).

Validation of Transgene Flanking Sequences: *Gm11823*-Transgene Junction

Validation of contig sequence (contig 4550) was done by testing the custom primer pair (contig 4550 primer pair) with OVE1328 (Tg/Tg) cDNA and gDNA. The primer pair was designed to flank the junction of *Gm11823* and the transgene complex. In agreement with RNA-seq data, PCR products were produced with both (Tg/Tg) gDNA (Fig 3.23) and cDNA (gDNA free) (Fig 3.13 lane2). This confirmed the integration site of the transgene at intron 2 (ch4:13,949,714 bp) of *Gm11823* gene. Furthermore, PCR data indicated the expression of an altered *Gm11823* message in OVE1328 (Tg/Tg) embryos where part of the transgene sequence is included in the message (Fig 3.19).

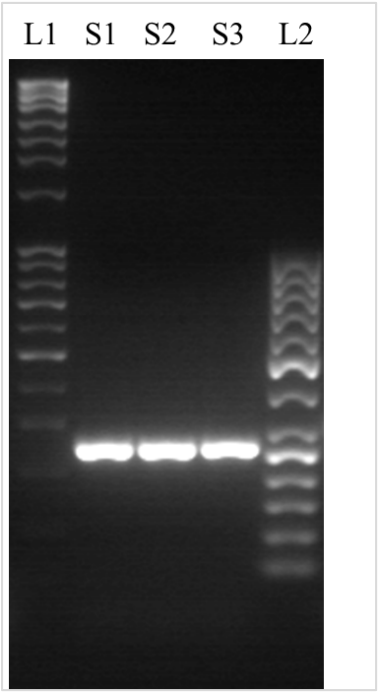


Figure 3.23. Conventional PCR data using 4550 contig primer pair. Amplification products of contig (4550) using 4550 contig primer pair with OVE1328 (Tg/Tg) gDNA (n=3). (L1): molecular weight marker (1kb DNA ladder). (L2): molecular weight marker (50bp DNA

3.4 Discussion

In the OVE1328 model the transgene integration site was accurately determined. The data presented confirmed previous FISH analysis and microarray data. The transgene integrated at intron 12 of *Runx1t1* gene /chr4: 13,876,840 bp. The data also indicated the disruption of a downstream gene known as *Gm11823,* lincRNA gene, where the integration took place at intron 2 of Gm11823/chr4:13,949,713.

An Altered Expression of *Runx1t1* in OVE1328 model

In OVE1328 model, *Runx1t1* message is being formed and expressed at similar levels to wild type counterpart (data not shown). However, structurally the *Runx1t1* message is different than wild type counterpart (deletion of exons 13and14, inclusion of part of intron 11-exon12 junction inclusion of part of intron 12, and inclusion of part of the transgene complex) which was validated by PCR data. The deletion of exons 13 and 14 would result in the loss of two functional domains of *Runx1t1* (NHR3 and NHR4). Both of those domains were shown to be important for protein-protein interactions such as its interaction with histone deacetylases and transcriptional corepressors N-CoR and SMRT.[174] The deletion of NHR3 and NHR4 domains in OVE1328 would probably affect RUNX1T1 interaction with these proteins and therefore its function as a transcriptional cofactor. It has been shown/ predicted that RUNX1T1 can have various binding partners. RUNX1T1 can homodimerize with another RUNX1T1 molecules, it can heterodimerize with other ETO family members, and it can bind other transcription factors such as C/EBPbeta and Neurod1[198,235]

In addition, PCR data indicated that the *Runx1t1* message contains part of the transgene complex as shown in Fig (3. 14). It is possible that such insertion, in addition to the deletion

mutation, would alter RUNX1T1 secondary and tertiary structures. Alterations of a protein folding can significantly alter its function and its binding with other interaction partners.[236]

Though PCR data indicated the presence of such message, northern blot data did not demonstrate the expression of *Runx1t1* in homozygous mutants. A possible explanation is that the *Runx1t1* message is very large (more than 10Kb). A large RNA species would require longer running time on formaldehyde agarose gels in northern blots. The running time used in the shown northern blot (Fig 3.16) was sufficient to detect RNA species less than 10Kb (4hours running time) as was demonstrated by detecting wild type *Runx1t1* RNA and the 10Kb band of the RNA ladder in the FA gel (data not shown). The ladder successfully transferred to the nylon membrane but the 10kb band position was close to the well's margin. Therefore, it would be useful to do another northern blotting where the RNA is allowed to run for a longer time (~7 hours) and then probe for *Runx1t1* message.

Furthermore, another controversial piece of evidence with qRT-PCR and conventional PCR data is the absence of any significant difference in the size of RUNX1T1 between all genotypes. Moreover, no differences were detected in the banding pattern seen among all genotypes. We tested 3 different commercially available antibodies (2 binding to an N-terminal epitope and 1 binding to a C-terminal epitope of RUNX1T1). If those antibodies specifically bound RUNX1T1, then it would be anticipated that there will be a difference in the banding pattern between those binding an N-terminal epitope and C-terminal epitope. The deletion mutation affecting the 3'end (exon13, exon 14) of *Runx1t1* would result in the loss of the epitope recognized by the latter antibody in the OVE1328 (Tg/Tg) protein samples and so no immunoreactive band will be detected. However, this was not the case. Furthermore, closer examination of the sequence similarity of the immunogen used for generating the Cell Signaling

and Sigma anti-RUNX1T1 antibodies showed high sequence identity to CBFA2T2 and CBFA2T3 (other ETO family members). These proteins have isoforms with molecular weights similar to RUNX1T1. The previous findings indicate that there could be a problem with the specificity of the antibodies used to RUNX1T1. Mass spectrometry can be used to verify and identify the RUNX1T1 immunoreactive band detected in OVE1328 (Tg/Tg) lysates.

The finding that the transgene disrupted not only the *Runx1t1* gene but also another downstream predicted gene known as *Gm11823* (lincRNA gene) adds another layer of complexity to the OVE1328 model. The transgene integration at *Gm11823* not only disrupted its structural integrity (deletion of exon3 and part of intron 2) but also induced an ectopic expression of *Gm11823* gene in OVE1328 line. At this stage, it is not clear how would such expression contribute to the phenotype seen in OVE1328 (Tg/Tg) embryos. It is unclear if the phenotype in this model is related to *Runx1t1* alteration or *Gm11823* alteration or the alteration of both genes.

OVE1328: from Genotype to Phenotype

To date, none of the phenotypes discussed in chapter 2 have been linked to mutations in either *Runx1t1* or *Gm11823* in mouse models or humans. However, several observations and findings are worth mentioning here. *Runx1t1* message is moderately expressed (whole embryo-in situ hybridization) in the palatal shelf of E14.5 (TS22) mouse embryo.[194] Furthermore, the tongue and the anterior and lower borders of the developing mandible showed a mild to moderate expression of *Runx1t1*.[194] At the same stage, *Runx1t1* was shown to be moderately expressed in the developing axial skeleton (which includes the developing rib cage).[194] Furthermore, *Runx1t1* was shown to be expressed in the developing murine fore and hindlimbs at E10.5, E11.5 and E14.5.[194] This being said, little is known about *Runx1t1* expression pattern during palatogenesis, rib and limb development in FVB/NJ strain. Furthermore, as far as our

knowledge no studies looked into the expression pattern of *Runx1t1* at protein level in those developing structures.

It is not known how *Runx1t1* or *Gm11823* or both would play a role in the pathogenesis of the affected structures in OVE1328 (Tg/Tg) embryos. For the time being, there is no knockout model of *Gm11823* gene whereas for *Runx1t1*, Calabi et al [214] generated a knockout model for *Runx1t1*. There is no evidence for the disruption of palate development or the rib development in this model. However, there is a concern about the Calabi mouse model. The paper used the S1 nuclease protection assay to demonstrate the disruption of the targeted exon (exon2). However, no other experiment was done to examine *Runx1t1* message using northern blot which will be helpful to exclude the possibility of the formation of other alternative isoforms that lack exons 2, 3 and 4 or the formation of novel *Runx1t1* variant. It is predicted that a message lacking those exons would translate into a protein that is ~ 53 KD. The Calabi paper indeed shows Western blot data in which a band of similar size to the predicted protein (translated from a message lacking exons2, 3, and 4) is detected. The authors pointed that this band could be another ETO family member detected by the antibodies. Though that could hold true, the band could also be a novel variant of RUNX1T1where exons 2, 3, 4 are spliced out. Such isoform has a molecular weight similar to the band shown in their western blot (~53Kd).

Future Studies

Our understanding to how *Runx1t1* and *Gm11823* are altered in OVE1328 (Tg/Tg) embryos is still limited. Further characterization of *Runx1t1* and *Gm11823* at the genomic DNA, mRNA and protein level is needed. At the mRNA level, having another RNA-seq done using gDNA free RNA samples will be helpful to get more information about the global gene expression profile in OVE1328 (Tg/Tg) mutants and to better dissect the mutation at the

integration site. At the protein level, mass-spectrometry can be very helpful to verify the identity of the band detected in the western blot and better characterize the mutation at the gene level.

Moreover, little is known about the spatio-temporal expression pattern of *Runx1t1* and *Gm11823* in the developing palatal shelves and vertebral column. This can be characterized in both the Wt (FVB/NJ) and OVE1328 (Tg/Tg) mutants using *in situ hybridization* of frozen head (coronal) and body sections (sagittal) of embryos harvested at different embryonic days .[237] qRt-PCR (using RNA extracted from dissected palatal shelves or developing vertebral column) can be an alternative approach to determine the expression pattern of both genes at different time points of these developing structures.

In addition to characterizing the expression profile of both genes, it would be interesting to examine the role of each gene or both on the development of the secondary palate. This can be done using the palatal organ culture system.[238] To understand the role of *Runx1t1* on palatogenesis, it will be useful to examine the effect of down regulating *Runx1t1* expression (using siRNA) in a suspension based palatal organ culture system as previously done.[239] This can be done by dissecting the developing midfacial region from a Wt embryo. Using the same culture system (midfacial region of Wt embryo) can be helpful to dissect the role of *Gm11823*. Ectopic expression of *Gm11823* in the developing Wt palatal shelves (using suspension based palatal organ culture system) can be achieved by lentiviral transduction.[240]

In addition to investigating the role of each gene in palatogenesis, it would be of great interest to examine if *Gm11823 gene* regulates *Runx1t1* expression. There is evidence that lincRNA genes can regulate the expression of other genes in cis or trans.[226] The regulation can take place at transcriptional and posttranscriptional levels.[226] This can be tested *in vitro* in a cell line that expresses *Runx1t1* but not *Gm11823*. It was previously shown that the preadipocyte 3T3-

L1 cell line expresses *Runx1t1* but there is no evidence if this cell line expresses *Gm11823.*[198] Therefore, *Gm11823* expression in 3T3-L1 cell line can be determined first using qRT-PCR. If the cell line does not express *Gm11823,* the cells can be transfected with a plasmid expressing *Gm11823.* RNA and protein will then be harvested from the cells to determine the expression level of *Runx1t1* (using qRT-PCR and Western blot, respectively) after validating the expression of *Gm11823* gene. If the cell line expresses *Gm11823* another alternative approach would be using primary cell line (such as mouse embryonic fibroblasts) from the Wt embryos (FVB/NJ) which normally does not express *Gm11823.*

Moreover, it is very useful to determine RUNX1T1 binding partners during palatogenesis in OVE 1328 (Tg/Tg). This can be examined using co-immunoprecipitation assay[241]. Such assay will enable predicting what pathways RUNX1T1 is involved in during palatogenesis[241].

Significance and Potential Impact

The study of OVE1328 model identifies new genes; *Runx1t1* and *Gm11823* which seem to be potentially important for normal secondary palate development and vertebral column development. This study demonstrates the significance and the justification for examining the role of *Runx1t1* and *Gm11823* in palatogenesis and the vertebral column development. It also ascertains the importance of understanding how both genes regulate each other and what molecular pathways they are involved in during normal palatal and vertebral column development in a murine model.

The combination of CP and supernumerary rib phenotype in OVE1328 model is of great interest. Using OMIM database, 3 human syndromes were identified in which CP and supernumerary ribs are part of their clinical manifestations (Table 3.6).

Syndrome	Gene	SMCP	CP	CL	CL&P	Rib Abnormality	Others
SIMPSON-GOLABI-BEHMEL SYNDROME, TYPE 1; SGBS1	GPC3	X	X	X	.	Extra ribs (13 ribs)	Short hands and feet,short fingers
MICROPHTHALMIA, SYNDROMIC 3; MCOPS3	SOX2	.	X	.	.	Abnormal number of ribs (13 ribs, 11 ribs)	Short stature
FRYNS SYNDROME; FRNS	N/A	.	X	X	X	Extra pair of ribs	Hypoplasia distal phalanges

Table 3.6 Human syndromes with clinical manifestations of CP and supernumerary ribs. Key: SMCP: sub-mucous cleft palate. CP: Cleft palate. CL: Cleft lip. CL&CP: Cleft lip and Cleft palate. OMIM database.

In humans, RUNX1T1 is potentially important for brain and heart development. A patient with a translocation t (5; 8) (q32; q21.3) results in a deletion mutation in *RUNX1T1* near its 5'end. The patient has congenital heart disease (ventricular septal defect) and mild to moderate mental retardation[242]. As far as CP there is no evidence that links *Runx1t1* to CP in humans, however, this gene might have been overlooked in the genome wide association studies as till now there is no evidence supporting a possible role for *Runx1t1* gene in human or mouse palatogenesis.

Although still in its experimental phase, gene therapy seems to be a promising approach for treating many types of human genetic diseases [243]. Till now, no human clinical trial in which gene therapy is/was employed to correct cleft palate in utero. However, successful correction of such defect in utero was demonstrated in a *Tgf-β3* knockout mouse model[244]. Therefore identifying the genes essential for normal palate development would potentially be employed to prevent human CP in utero and help prevent the suffering associated with such anomaly.

APPENDIX 2.1: TYROSINASE PRIMER PAIRS (SEQUENCE)

FW primer	Sequence	RV Primer	Sequence	Product Size (bp)
TYPS-W	CTGTCCAGTGCACCATCTGGACCTC	**TYBS-RV**	GATTACGTAATAGTGGTCCCTCAGG	700

APPENDIX 2.2: TYROSINASE PRIMER PAIRS (GENOMIC COORDINATES)

FW primer	Genomic coordinates	RV Primer	Genomic coordinates	Description Exon
TYPS-FW	87,493,149-87,493,173 or TYBS complement (1661-1685)	**TYBS-RV**	87,483,956-87,483,980 or TYBS complement (962-986)	Part of exons 1 and 2

APPENDIX 2.3: PROTOCOL MODIFICATIONS FOR SKELETAL CLEARING AND STAINING

Material	Incubation Time (Days)	Temperature (°C)
95% (v/v) Ethanol	10	22-25
Acetone	2-3	22-25
Alcian Blue (0.3% in 70% ethanol) Alizarin Red (0.1% in 95% ethanol)	3-4	37
1% KOH	14	22-25
20% Glycerol in 1% KOH	19	22-25
50% Glycerol in 1% KOH	21	22-25
80% Glycerol in 1% KOH	21-40	22-25
100% Glycerol	∞	22-25

APPENDIX 2.4: RIB CAGE DATA

Embryo ID	Genotype	Rib Count 1		Rib Count 2		Rib Count 3	
		Rt side	Lt side	Rt side	Lt side	Rt side	Lt side
2838-1	Wt	13	13	13	13	13	13
2838-2	Hz	13	13	13	13	13	13
2838-3	CP	13	13	13	13	13	13
2838-4	Hz	13	13	13	13	13	13
2838-5	CP	13	14	13	14	13	14
2838-6	Hz	13	13	13	13	13	13
2838-7	Hz	13	13	13	13	13	13
2838-8	Hz	13	13	13	13	13	13
2838-9	Wt	13	13	13	13	13	13
2830-1	CP	14	14	14	14	14	14
2830-2	CP	13	14	13	14	13	14
2830-3	CP	14	14	14	14	14	14
2830-4	Wt	13	13	13	13	13	13
2830-5	Hz	13	13	13	13	13	13
2830-6	Wt	13	13	13	13	13	13
2830-7	CP	14	14	14	14	14	14
2830-8	Hz	13	13	13	13	13	13
2830-9	Hz	13	13	13	13	13	13
2830-10	Hz	13	13	13	13	13	13
2830-11	Wt	13	13	13	13	13	13
2843-1	Wt	13	13	13	13	13	13
2843-2	Hz	13	13	13	13	13	13
2843-3	Wt	13	13	13	13	13	13
2843-4	Hz	13	13	13	13	13	13
2843-5	Hz	13	13	13	13	13	13
2843-6	Hz	13	13	13	13	13	13
2843-7	Wt	13	13	13	13	13	13
2843-8	Wt	13	13	13	13	13	13
2843-9	Hz	13	13	13	13	13	13
2843-10	CP	14	13	14	13	14	13

APPENDIX 2.5: MEAN LENGTH OF RIGHT FORELIMB AND HINDLIMB BONES DATA

ID	Genotype	Rt Humerus (mm)	Rt Radius (mm)	Rt Ulna (mm)	Rt Femur (mm)	Rt Tibia (mm)	Rt Fibula (mm)
2838-1	Wt	2.17	1.77	2.19	1.46	1.50	1.31
2838-9	Wt	2.00	1.64	1.99	1.51	1.58	1.28
2830-4	Wt	2.37	1.91	2.35	1.82	1.98	1.85
2830-6	Wt	2.43	2.11	2.53	1.87	2.17	1.94
2830-11	Wt	2.44	2.05	2.52	1.92	2.10	1.92
2838-2	(Tg/+)	2.20	1.62	1.98	1.62	1.60	1.32
2838-4	(Tg/+)	2.00	1.62	1.97	1.39	1.45	1.33
2830-5	(Tg/+)	2.26	1.81	2.28	1.75	2.01	1.77
2830-8	(Tg/+)	2.48	2.08	2.54	1.97	2.17	1.98
2830-9	(Tg/+)	2.47	2.20	2.64	1.90	2.12	1.92
2830-10	(Tg/+)	2.35	1.76	2.28	1.76	1.88	1.63
2838-3	(Tg/Tg)	2.03	1.50	1.95	1.44	1.56	1.29
2838-5	(Tg/Tg)	2.00	1.40	1.84	1.38	1.17	0.64
2830-1	(Tg/Tg)	2.16	1.47	1.72	1.43	1.45	1.27
2830-2	(Tg/Tg)	1.99	1.30	1.79	1.39	1.38	0.85
2830-3	(Tg/Tg)	2.33	1.75	2.23	1.66	1.75	1.59
2830-7	(Tg/Tg)	2.14	1.84	2.27	1.60	1.80	1.60

APPENDIX 2.6: MEAN LENGTH OF LEFT FORELIMB AND HINDLIMB BONES DATA

ID	Genotype	Lt Humerus (mm)	Lt Radius (mm)	Lt Ulna (mm)	Lt Femur (mm)	Lt Tibia (mm)	Lt Fibula (mm)
2838-1	Wt	2.28	1.64	2.08	1.54	1.46	1.03
2838-9	Wt	2.30	1.81	2.25	1.48	1.53	1.35
2830-4	Wt	2.61	1.99	2.44	1.86	1.86	1.73
2830-6	Wt	2.45	2.15	2.50	1.86	1.97	1.87
2830-11	Wt	2.54	2.01	2.50	1.83	1.99	1.89
2838-2	(Tg/+)	2.41	1.78	2.20	1.55	1.57	1.43
2838-4	(Tg/+)	2.19	1.71	2.15	1.45	1.51	1.30
2830-5	(Tg/+)	2.42	1.92	2.36	1.70	1.83	1.68
2830-8	(Tg/+)	2.56	2.09	2.51	1.88	2.07	1.97
2830-9	(Tg/+)	2.61	2.20	2.51	1.90	1.94	1.84
2830-10	(Tg/+)	2.39	1.89	2.32	1.60	1.75	1.63
2838-3	(Tg/Tg)	2.01	1.68	2.20	1.49	1.52	1.32
2838-5	(Tg/Tg)	2.06	1.50	1.92	1.27	1.07	0.47
2830-1	(Tg/Tg)	2.12	1.52	1.92	1.47	1.47	1.22
2830-2	(Tg/Tg)	1.97	1.43	1.72	1.28	1.25	0.77
2830-3	(Tg/Tg)	2.18	1.70	2.07	1.64	1.68	1.51
2830-7	(Tg/Tg)	2.31	1.81	2.26	1.74	1.68	1.51

APPENDIX 2.7: MEAN TMD DATA (RT AND LT HUMERUS) / UNIVERSAL THRESHOLD

ID	Genotype	RT TMD	Threshold	Lt TMD	Threshold
2616-1	Wt	0.37382	47/127	0.39672	47/127
2619-5	Wt	0.38606	47/127	0.39855	47/127
2619-7	Wt	0.35408	47/127	0.39773	47/127
2616-3	Wt	0.38624	47/127	0.37717	47/127
2875-6	Wt	0.37041	47/127	0.36471	47/127
2875-7	Wt	0.3508	47/127	0.37434	47/127
2606-1	(Tg/+)	0.37614	47/127	0.36748	47/127
2616-7	(Tg/+)	0.36203	47/127	0.38328	47/127
2616-8	(Tg/+)	0.3731	47/127	0.37672	47/127
2619-6	(Tg/+)	0.38458	47/127	0.36362	47/127
2875-1	(Tg/+)	0.36672	47/127	0.3636	47/127
2875-3	(Tg/+)	0.36356	47/127	0.36323	47/127
2606-8	(Tg/Tg)	0.3985	47/127	0.38148	47/127
2616-9	(Tg/Tg)	0.37215	47/127	0.3797	47/127
2619-8	(Tg/Tg)	0.37964	47/127	0.3565	47/127
2606-7	(Tg/Tg)	0.38265	47/127	0.39765	47/127
2879-2	(Tg/Tg)	0.38069	47/127	0.35847	47/127
2879-4	(Tg/Tg)	0.33983	47/127	0.36108	47/127

APPENDIX 2.8: MEAN HEAD MEASUREMENTS DATA

Sample ID	Genotype	Muzzle width (mm)	Snout Occiput length (mm)	IPD (mm)
2518-4	Wt	2.82	7.26	3.68
2518-5	Wt	3.08	7.54	3.85
2830-4	Wt	3.65	8.18	4.13
2830-6	Wt	3.31	7.82	4.08
2830-11	Wt	3.35	7.80	4.05
2838-9	Wt	3.14	7.59	4.08
2838-6	(Tg/+)	3.08	7.64	4.07
2518-2	(Tg/+)	3.05	7.77	3.85
2838-7	(Tg/+)	3.38	7.41	4.00
2838-8	(Tg/+)	3.12	7.50	4.02
2838-4	(Tg/+)	3.12	8.01	3.83
2830-9	(Tg/+)	3.38	7.71	4.27
2518-3	(Tg/Tg)	2.80	6.91	3.58
2838-5	(Tg/Tg)	3.27	7.30	3.81
2830-2	(Tg/Tg)	3.17	7.40	3.60
2830-1	(Tg/Tg)	3.00	7.49	3.62
2830-3	(Tg/Tg)	3.30	7.35	4.07
2830-7	(Tg/Tg)	3.32	7.89	3.89

PPENDIX 3.1:RUNX1T1 CDNA PRIMER PAIRS (SEQUENCE)

FW primer	Sequence	RV Primer	Sequence	Product Size (bp)
1L	GATCTGTGGGCTGGTGAACT	**1R**	CAGGTGAGTCTGGCATTGTG	201
2L	GAGCAGCAGAGGAGATTAGCA	**2R**	CTGGGGGAGTCAGCCTAGA	200
3L	TCCTCCAACCACTCAAGGAG	**3R**	AGCTTGCTGAGTTGCCTAGC	204
4L	AGCTCATTTACGCCAACGAC	**4R**	CCAAACTGCTGCAGGGTAGT	206
5L	TGTGGTGCTAGGCAACTCAG	**5R**	AAACGGGATGACAAAAGGTC	201
6L	TGTCATCCCGTTTTTGAAGG	**6R**	TCTGTCTGGAGTTCGCCTCT	202
7L	GTGAACGAAAACGGGAAGAG	**7R**	CATCCAATCGGTAATGCTGA	205
8L	CCTGCTCCAGCGTGAACT	**8R**	GGCTCTCTGTCAAAGCCATT	201
9L	TTGACAGAGAGCCTTTGCAC	**9R**	GCTGGGGTGTCGATAGGAG	200
10L	CCACCTCAGCATTACCGATT	**10R**	TGTCTTCTCCACCATGTCCA	228
11L	TTACATGGCACACGTCAAGAA	**11R**	ATCACTGTACCGCCGTATCC	201
12L	TGGACATGGTGGAGAAGACA	**12R**	GTGAAGGAATTCCCGATGTG	209
13L	CACTCCAGGCAGCAGAGTC	**13R**	CATGTCATGGGCTTTCCTCT	201
14L	AATGACGGAGCTGCAAAAG	**14R**	TGCACGTCTCACTTGCTTTC	200
15L	TCAACCAGCAGGAAGACTCC	**15R**	GGGTGTGACTGAGGAGCTGA	203
16L	TGGGAGAAGCATCACCACAT	**16R**	TTGCTTTCGTGTTGGTTGTG	237

APPENDIX 3.2: RUNX1T1 cDNA PRIMER PAIRS (GENOMIC COORDINATES)

FW primer	Genomic coordinates	RV primer	Genomic coordinates	Description/ Exon
1L	13,771,384 -13,771,403	**1R**	13,835,652-13,835,671	Part of 3&5
2L	13,771,421-13,771,441	**2R**	13,835,689-13,835,707	Part of 3&5
3L	13,835,718-13,835,737	**3R**	13,837,891-13,837,910	Part of 5&6
4L	13,835,749- 13,835,768	**4R**	13,837,924-13,837,943	Part of 5&6
5L	13,837,885- 13,837,904	**5R**	13,841,862-13,841,881	Part of 6, Most of 7
6L	13,841,869-13,841,888	**6R**	13,846,939- 13,846,957	Part of 7 and 8
7L	13,846,923-13,846,942	**7R**	13,859,961-13,859,980	Part of 9
8L	13,846,787-13,846,804	**8R**	13,859,821-13,859,840	Most of 8, Part of 9
9L	13,859,828- 13,859,847	**9R**	13,860,009-13,860,027	Most of 9, Part of 10
10L	13,859,956-13,859,975	**10R**	13,865,847-13,865,866	Part of 9, All 10, Most of 11
11L	13,865,203-13,865,223	**11R**	13,865,928-13,865,947	All10, Part of 11
12L	13,865,847-13,865,866	**12R**	13,875,480-13,875,499	Most of 11, Small Part of 12
13L	13,865,987-13,866,005	**13R**	13,881,116-13,881,135	Part of 11, All12, Part of 13
14L	13,881,081-13,881,098	**14R**	13,889,632-13,889,651	Part of 13, Small part 14
15L	13,881,215-13,881,234	**15R**	13,889,769-13,889,788	Part of 14
16L	13,889,702-13,889,721	**16R**	13,889,919-13,889,938	Part of 14

APPENDIX 3.3: RUNX1T1 EXON-INTRON JUNCTION PRIMER PAIRS PRIMER PAIRS (SEQUENCE)

FW primer	Sequence	RV primer	Sequence	Product Size (bp)
121 L	GCAATGCTGATGGTGCTCT	**121 R**	CAGATCTCCTCTGGCACGTA	112
122 L	ATCGGGAATTCCTTCACAGG	**122 R**	ATTGTTGCTGTTCGCAGTGA	172
123 L	AAGTGGCAGAGTTGCTTGCT	**123 R**	ATCAGGCCTGGTTTGAGAAT	365
131 L	TGAATGAGGTGAAGCGACAG	**131 R**	CTGGAGTCTTCCTGCTGGTT	178
132 L	TTTGGATGAGAGCTGGGTTT	**132 R**	AAGTTTTTCCCAGCGATCTG	400
141 L	CAGCCAGGGTCTTCTGAATC	**141 R**	TCTATGGTAGAGGGGGTTCC	390
142 L	CTTGGAGATGTGGGAGCAGT	**142 R**	GTCCACAGATGTGGTGATGC	207
143 L	CTGTAACACGGCCCGATACT	**143 R**	GTTCCCGGGGTGGTAGAC	201
144 L	CTCAGTCACACCCAGCAGTG	**144 R**	CCGTTACTGGCCTCTGTGTT	302
145 L	GCCAGTAACGGGTCGTAATG	**145 R**	CTGTATCCACCGCTCATGC	305
146 L	GCTGTTGAAAGCAAAAATGC	**146 R**	ACCGCACTGGAAATCATCTT	302
147 L	GCAGCAGATTGGAAGGAGAC	**147 R**	TGGAACCTTGCAAAAGTGAA	319
148 L	CAGCCAGCTGCCCTAAATAA	**148 R**	CGCATCTTCACATTTGTCCA	282
149 L	TGAGGTTTTCCCAATGGTGT	**149 R**	ATTCCATGAAGCATGCTGGT	324

APPENDIX 3.4: RUNX1T1 EXON-INTRON JUNCTION PRIMER PAIRS (GENOMIC COORDINATES)

FW primer	Genomic coordinates	RV primer	Genomic coordinates	Description Exon/Intron
121 L	13,875,423-13,875,441	**121 R**	13,875,515-13,875,534	Part of Intron 11 & Most of Exon12
122 L	13,875,483-13,875,502	**122 R**	13,875,635-13,875,654	Most of Exon12 and Part of Intron 12
123 L	13,875,384-13,875,403	**123 R**	13,875,729-13,875,748	Part of Introns 11,12,and All Exon 12
131 L	13,881,059-13,881,078	**131 R**	13,881,217-13,881,236	Most of Exon13
132 L	13,880,949-13,880,968	**132 R**	13,881,329-13,881,348	Part of Introns 12,13 and All Exon 13
141 L	13,889,482-13,889,501	**141 R**	13,889,852-13,889,871	Part of Intron 13 and Part of Exon 14
142 L	13,889,523-13,889,542	**142 R**	13,889,710-13,889,729	Part of Intron 13 and Part of Exon 14
143 L	13,889,656-13,889,675	**143 R**	13,889,839-13,889,856	Part of Exon 14
144 L	13,889,776-13,889,795	**144 R**	13,890,058-13,890,077	Part of Exon 14
145 L	13,890,067-13,890,086	**145 R**	13,890,353-13,890,371	Part of Exon 14
146 L	13,890,264-13,890,283	**146 R**	13,890,546-13,890,565	Part of Exon 14
147 L	13,890,456-13,890,475	**147 R**	13,890,755-13,890,774	Part of Exon 14
148 L	13,890,721-13,890,740	**148 R**	13,890,983-13,891,002	Part of Exon 14
149 L	13,890,946-13,890,965	**149 R**	13,891,250-13,891,269	Part of Exon 14

APPENDIX 3.5: GM11823 PRIMER PAIRS (SEQUENCE SECTION)

FW primer	Sequence	RV primer	Sequence	Product Size (bp)
A Left	TGCAAAGAAGACAGCAGTGG	**A Right**	GCTTGGGAGCTGTTCTCTGA	257
B Left	CCAAAAGGAATGTTCGTCAGA	**B Right**	CCCAGAACTCATGTCAGGTG	304
C Left	TGTCGGATCTCCTAGAGCTTG	**C Right**	TCCTCACCAACAGAGATCACC	276
D Left	TCATCTCCCTGAAATCAGCA	**D Right**	TGCAAAGATGGTAAGCTGGTC	256
E Left	CCAACAAATCATGCCAAAAA	**E Right**	TGCCAGACATCTGTGAACCT	304
F Left	GCATCTGTGTGAAGCTGACG	**F Right**	TTTCCAACCATTGTGACAGC	287
G Left	CCACTGCACCTAGCAGAAGA	**G Right**	TTTGTGATTGGTTGCCTGAG	310
H Left	TGAAAAATGAGGAATTGAGGAA	**H Right**	TCCCATCTGATTGTGTGGAA	300
I Left	GTTGACTCTTGGTCGCCTTG	**I Right**	TGTTTCCATGAGCCTCCAAT	360
J Left	ATTGGAGGCTCATGGAAACA	**J Right**	CCCCAGCCAAAAGGTTAAA	278
K Left	CGAAAAGCACCAGAAGCAAC	**K Right**	GAAACGTTTGGCAGTCAATG	252

APPENDIX 3.6: GM11823 PRIMER PAIRS (GENOMIC COORDINATES)

FW primer	Genomic coordinates	RV primer	Genomic coordinates	Description/ Exon
A Left	13,995,159-13,995,178	**A Right**	13,994,922-13,994,941	Part of exon 1
B Left	13,994,991-13,995,011	**B Right**	13,967,140-13,967,159	Part of exons 1 & 2
C Left	13,967,193-13,967,213	**C Right**	13,921,544-13,921,564	Part of exons 2 & 3
D Left	13,921,577-13,921,596	**D Right**	13,921,341-13,921,361	Part of exon 4
E Left	13,921,485-13,921,504	**E Right**	13,921,201-13,921,220	Part of exon 4
F Left	13,921,324-13,921,343	**F Right**	13,921,057-13,921,076	Part of exon 4
G Left	13,921,057-13,921,076	**G Right**	13,920,834-13,920,853	Part of exon 4
H Left	13,920,893-13,920,914	**H Right**	13,920,616-13,920,635	Part of exon 4
I Left	13,920,657-13,920,676	**I Right**	13,920,417-13,920,436	Part of exon 4
J Left	13,920,417-13,920,436	**J Right**	13,920,159-13,920,177	Part of exon 4
K Left	13,920,196-13,920,215	**K Right**	13,919,964-13,919,983	Part of exon 4

APPENDIX 3.7: CONTIG PRIMER PAIRS (SEQUENCE)

FW primer	Sequence	RV Primer	Sequence	Product Size (bp)
Contig 15360L	GGATGGTAATCAGTTTCCCAAG	**Contig 15360R**	AGATTCTGAATACAAGCCTGTTG	303

APPENDIX 3.8: CONTIG PRIMER PAIRS (GENOMIC COORDINATES)

FW primer	Genomic coordinates	RV Primer	Genomic coordinates	Description Exon/Intron
Contig 15360L	13,876,611-13,876,632	**Contig 15360R**	87,493,698-87,493,720 or TYBS complement 2210-2232	(Intron 12) Runx1t1-Transgene
Contig 4550L	MSCV-neo vector 2827-2846	**Contig 4550R**	13,949,652-13,949,671	Transgene- (Intron 2) Gm11823

APPENDIX 3.9: THERMAL CYCLER PROGRAMMING CONDITIONS AND GENERAL PRIMER INFORMATION

*Step	Temperature (°C)	Time (minutes)	Hold/ Cycles	Comments
Initial Denaturation	95	3	1 hold	3 min including preheating for (1 min)
Denaturation	95	1	30-35 cycles	
Annealing	54-55 or 60	1	30-35 cycles	contig primers & tyrosinase primers worked best at 60°C all other primers worked at 54-55°C
Extension	72	1	30-35 cycles	
Final Extension	72	5	1 hold	
Soak	4	∞	1 hold	

* PCR was done using Gene Amp ® PCR System 9700 (Applied Biosystems, Grand Island, NY, USA)

General primer information

Custom primers (Invitrogen, Grand Island, NY, USA) were received in a lyophilized form. Upon arrival primers were reconstituted in PCR-Water to get at 1mM concentration. To get a 1mM primer stock, X µl of PCR water were added to X nmoles of the primer. The nmole value was provided by the manufacturer. Primer working concentration was 10µM. The optimum $MgCl_2$ concentration was determined for each primer pair separately by doing $MgCl_2$ curve (1, 1.5, 2, 2.5, 3.3.5 µM). Generally, most primer pairs worked well at 1.5-2 µM $MgCl_2$. An exception is contig primers that worked best at high $MgCl_2$ concentrations 3.5 µM.

APPENDIX 3.10: RUNX1T1 ANTIBODIES: CLONALITY I, MMUNOGEN SEQUENCE AND EPITOPE BINDING SITE

Antibody ID	Clonality	Immunogen Sequence
Cell Signaling 4498	polyclonal	NGFDREPLHSEHPSKRPCTISPGQRYSPNNGLSYQPNGLPH
Sigma C5616	polyclonal	RDSYRHPSHRDLRDRNR
Santa Cruz sc-9737	polyclonal	LQAQQQGDTPAVSSSVTPNSGAGSPMDTPPAATPRSTTPGTPSTIETTPR

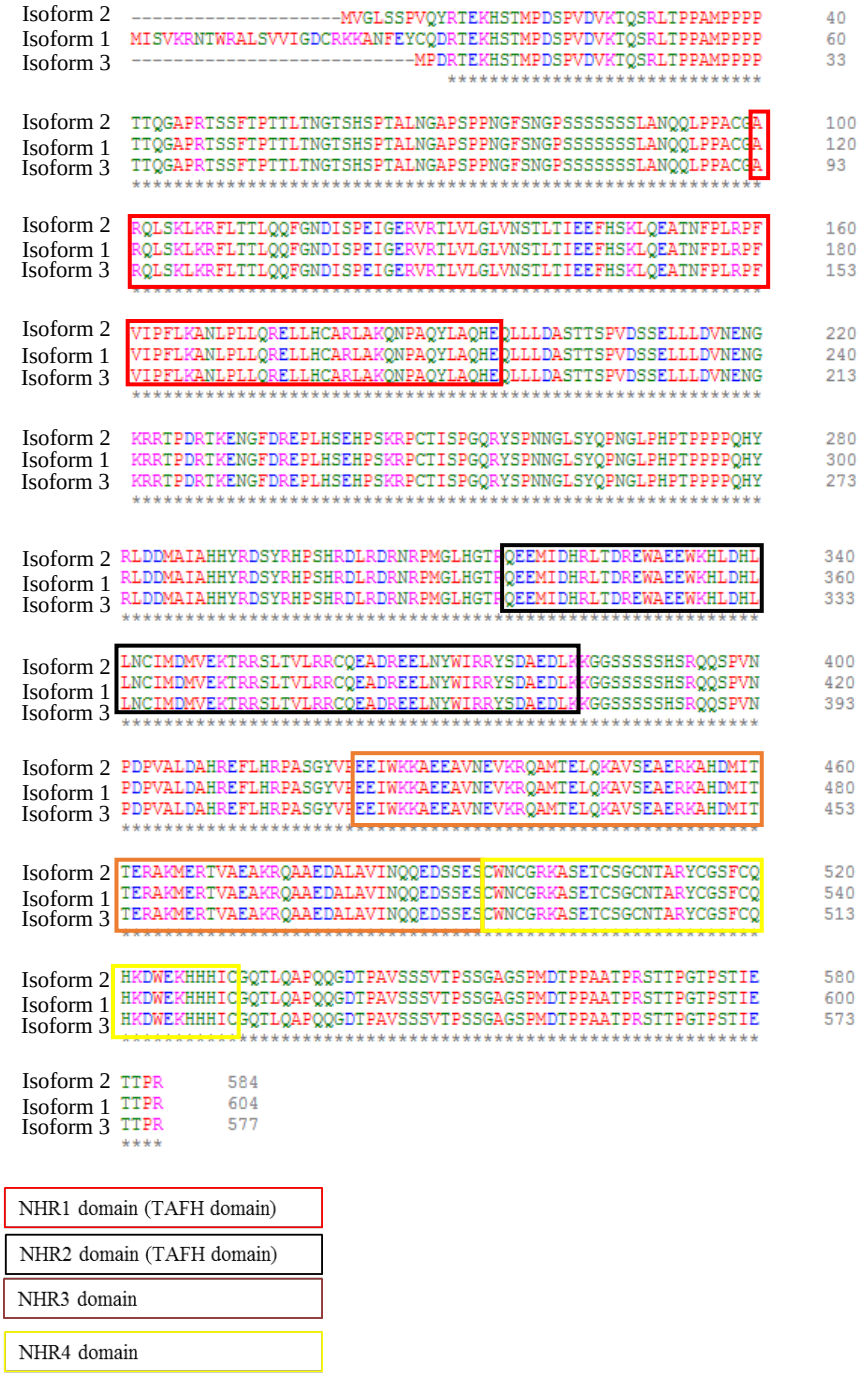
Antibody ID	Epitope Binding site isoform 1	Epitope Binding site isoform 2	Epitope Binding site isoform 3
Cell Signaling 4498	251-291 a.a	231-271 a.a	224-264 a.a
Sigma C5616	312-328 a.a	292-308 a.a	285-301 a.a
Santa Cruz sc-9737	555-604 a.a	535-584 a.a	528-577 a.a

APPENDIX 3.11: PART OF 5' END (A) AND 3'END (B) MESSENGER RNA SEQUENCE ALIGNMENTS OF RUNX1T1

A.

B.

APPENDIX 3.12: PROTEIN SEQUENCE ALIGNEMNTS AND DOMAIN STRUCTURE OF RUNX1T1 ISOFORMS



APPENDIX 3.13: DEFINITIONS

Term	Definition
Query	is the input sequence to which other data base sequences are compared to.
Query coverage	is the percentage of the query sequence that is shared with the subject sequence.
Percent identity	is the extent of the relation between compared sequences expressed as a percentage.

The content adopted from Fassler et al[245] and NCBI data base 10.2014.

REFERENCES

1 Soukup, V., Horácek, I. & Cerny, R. Development and evolution of the vertebrate primary mouth. *Journal of Anatomy* **222**, 79-99, doi:10.1111/j.1469-7580.2012.01540.x (2013).

2 Jiang, R., Bush, J. O. & Lidral, A. C. Development of the upper lip: Morphogenetic and molecular mechanisms. *Developmental Dynamics* **235**, 1152-1166, doi:10.1002/dvdy.20646 (2006).

3 Sperber, G. & Sperber, S. in *Cleft Lip and Palate* (ed Samuel Berkowitz) Ch. 1, 3-33 (Springer Berlin Heidelberg, 2013).

4 Yoon, H., Chung, I. S., Seol, E. Y., Park, B. Y. & Park, H. W. Development of the lip and palate in staged human embryos and early fetuses. *Yonsei Med J* **41**, 477-484 (2000).

5 Cox, T. C. Taking it to the max: The genetic and developmental mechanisms coordinating midfacial morphogenesis and dysmorphology. *Clinical Genetics* **65**, 163-176, doi:10.1111/j.0009-9163.2004.00225.x (2004).

6 Bush, J. O. & Jiang, R. Palatogenesis: morphogenetic and molecular mechanisms of secondary palate development. *Development* **139**, 231-243, doi:10.1242/dev.067082 (2012).

7 Gritli-Linde, A. Molecular control of secondary palate development. *Dev Biol* **301**, 309-326, doi:10.1016/j.ydbio.2006.07.042 (2007).

8 Levi, B. *et al.* Palatogenesis: engineering, pathways and pathologies. *Organogenesis* **7**, 242-254, doi:10.4161/org.7.4.17926 (2011).

9 Jin, J. Z., Li, Q., Higashi, Y., Darling, D. S. & Ding, J. Analysis of Zfhx1a mutant mice reveals palatal shelf contact-independent medial edge epithelial differentiation during palate fusion. *Cell Tissue Res* **333**, 29-38, doi:10.1007/s00441-008-0612-x (2008).

10 Yu, K. & Ornitz, D. M. Histomorphological study of palatal shelf elevation during murine secondary palate formation. *Developmental Dynamics* **240**, 1737-1744, doi:10.1002/dvdy.22670 (2011).

11 Diewert, V. M. Craniofacial growth during human secondary palate formation and potential relevance of experimental cleft palate observations. *J Craniofac Genet Dev Biol Suppl* **2**, 267-276 (1986).

12 Dudas, M., Li, W.-Y., Kim, J., Yang, A. & Kaartinen, V. Palatal fusion – Where do the midline cells go?: A review on cleft palate, a major human birth defect. *Acta Histochemica* **109**, 1-14, doi:http://dx.doi.org/10.1016/j.acthis.2006.05.009 (2007).

13 Fitchett, J. E. Medial edge epithelium transforms to mesenchyme after embryonic palatal shelves fuse. *Developmental Biology* **131**, 455-474 (1989).

14 Cuervo, R. & Covarrubias, L. Death is the major fate of medial edge epithelial cells and the cause of basal lamina degradation during palatogenesis. *Development* **131**, 15-24, doi:10.1242/dev.00907 (2004).

15 Charoenchaikorn, K. *et al.* Runx1 is involved in the fusion of the primary and the secondary palatal shelves. *Developmental Biology* **326**, 392-402, doi:http://dx.doi.org/10.1016/j.ydbio.2008.10.018 (2009).

16 Jin, J.-Z. & Ding, J. Analysis of cell migration, transdifferentiation and apoptosis during mouse secondary palate fusion. *Development* **133**, 3341-3347, doi:10.1242/dev.02520 (2006).

17 Kumar, H. P. *Orthodontics*. (Elsevier India Pvt. Limited, 2008).

18 Baek, J.-A. *et al.* Bmpr1a signaling plays critical roles in palatal shelf growth and palatal bone formation. *Developmental Biology* **350**, 520-531, doi:http://dx.doi.org/10.1016/j.ydbio.2010.12.028 (2011).

19 Stanier, P. & Moore, G. E. Genetics of cleft lip and palate: syndromic genes contribute to the incidence of non-syndromic clefts. *Human Molecular Genetics* **13**, R73-R81, doi:10.1093/hmg/ddh052 (2004).

20 Gritli-Linde, A. The etiopathogenesis of cleft lip and cleft palate: usefulness and caveats of mouse models. *Curr Top Dev Biol* **84**, 37-138, doi:10.1016/s0070-2153(08)00602-9 (2008).

21 Burdi, A. R. Morphogenesis of the palate in normal human embryos with special emphasis on the mechanisms involved. *American Journal of Anatomy* **120**, 149-159, doi:10.1002/aja.1001200112 (1967).

22 Bloch-Zupan, A., Hunter, N., Manthey, A. & Gibbins, J. R-twist gene expression during rat palatogenesis. *Int J Dev Biol* **45**, 397-404 (2001).

23 Alappat, S. R. *et al.* The cellular and molecular etiology of the cleft secondary palate in Fgf10 mutant mice. *Developmental Biology* **277**, 102-113, doi:http://dx.doi.org/10.1016/j.ydbio.2004.09.010 (2005).

24 Wu, C. *et al.* Intra-amniotic Transient Transduction of the Periderm With a Viral Vector Encoding TGF[beta]3 Prevents Cleft Palate in Tgf[beta]3-/- Mouse Embryos. *Mol Ther* **21**, 8-17 (2013).

25 Zhang, Z. *et al.* 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**, 4135-4146 (2002).

26 Yu, L. *et al.* Shox2-deficient mice exhibit a rare type of incomplete clefting of the secondary palate. *Development* **132**, 4397-4406, doi:10.1242/dev.02013 (2005).

27 *Comparative Anatomy and Histology : A Mouse and Human Atlas*. (Academic Press, 2011).

28 Katebi, N., Kolpakova-Hart, E., Lin, C. Y. & Olsen, B. R. The mouse palate and its cellular responses to midpalatal suture expansion forces. *Orthodontics & Craniofacial Research* **15**, 148-158, doi:10.1111/j.1601-6343.2012.01547.x (2012).

29 Harville, E. W., Wilcox, A. J., Lie, R. T., Vindenes, H. & Åbyholm, F. Cleft Lip and Palate versus Cleft Lip Only: Are They Distinct Defects? *American Journal of Epidemiology* **162**, 448-453, doi:10.1093/aje/kwi214 (2005).

30 Dixon, M. J., Marazita, M. L., Beaty, T. H. & Murray, J. C. Cleft lip and palate: understanding genetic and environmental influences. *Nat Rev Genet* **12**, 167-178, doi:10.1038/nrg2933 (2011).

31 McMillan, J. A., Feigin, R. D., DeAngelis, C. & Jones, D. *Oski's Pediatrics: Principles & Practice*. (Lippincott Williams & Wilkins, 2006).

32 Cobourne, M. T. The complex genetics of cleft lip and palate. *The European Journal of Orthodontics* **26**, 7-16, doi:10.1093/ejo/26.1.7 (2004).

33 Gong, S.-G., White, N. J. & Sakasegawa, A. Y. The Twirler mouse, a model for the study of cleft lip and palate. *Archives of Oral Biology* **45**, 87-94, doi:http://dx.doi.org/10.1016/S0003-9969(99)00101-6 (2000).

34 Juriloff, D. M. & Fraser, F. C. Genetic maternal effects on cleft lip frequency in A/J and CL/Fr mice. *Teratology* **21**, 167-175, doi:10.1002/tera.1420210206 (1980).

35 Trasler, D. ASPIRIN-INDUCED CLEFT LIP AND OTHER MALFORMATIONS IN MICE. *The Lancet* **285**, 606-607, doi:http://dx.doi.org/10.1016/S0140-6736(65)91192-X (1965).

36 Engelking, L. J. *et al.* Severe facial clefting in Insig-deficient mouse embryos caused by sterol accumulation and reversed by lovastatin. *J Clin Invest* **116**, 2356-2365, doi:10.1172/jci28988 (2006).

37 Colmenares, C. *et al.* Loss of the SKI proto-oncogene in individuals affected with 1p36 deletion syndrome is predicted by strain-dependent defects in Ski-/- mice. *Nat Genet* **30**, 106-109 (2002).

38 Källén, B. in *Epidemiology of Human Congenital Malformations* Ch. 13, 73-77 (Springer International Publishing, 2014).

39 Apesos, J. & Anigian, G. M. Median cleft of the lip: its significance and surgical repair. *Cleft Palate Craniofac J* **30**, 94-96, doi:10.1597/1545-1569(1993)030<0094:mcotli>2.3.co;2 (1993).

40 Ferguson, M. W. J. Palate development. *Development* **103**, 41-60 (1988).

41 Mangold, E., Ludwig, K. U. & Nöthen, M. M. Breakthroughs in the genetics of orofacial clefting. *Trends in Molecular Medicine* **17**, 725-733, doi:http://dx.doi.org/10.1016/j.molmed.2011.07.007 (2011).

42 Parker, S. E. *et al.* Updated national birth prevalence estimates for selected birth defects in the United States, 2004–2006. *Birth Defects Research Part A: Clinical and Molecular Teratology* **88**, 1008-1016, doi:10.1002/bdra.20735 (2010).

43 Jugessur, A., Farlie, P. G. & Kilpatrick, N. The genetics of isolated orofacial clefts: from genotypes to subphenotypes. *Oral Diseases* **15**, 437-453, doi:10.1111/j.1601-0825.2009.01577.x (2009).

44 Mossey, P. A. Epidemiology of oral clefts 2012: an international perspective. *Frontiers of oral biology* **16**, 1-18, doi:10.1159/000337464 (2012).

45 Murray, J. C. Gene/environment causes of cleft lip and/or palate. *Clinical Genetics* **61**, 248-256, doi:10.1034/j.1399-0004.2002.610402.x (2002).

46 *Pediatric Otolaryngology*.

47 Kulkarni, A. P. *Objective Anaesthesia: A Comprehensive Textbook for the Examinee*. (Jaypee Brothers, Medical Publishers, 2013).

48 Kummer, A. *Cleft Palate & Craniofacial Anomalies: Effects on Speech and Resonance*. (Cengage Learning, 2007).

49 Mossey, P. A., Little, J., Munger, R. G., Dixon, M. J. & Shaw, W. C. Cleft lip and palate. *The Lancet* **374**, 1773-1785, doi:http://dx.doi.org/10.1016/S0140-6736(09)60695-4.

50 Program, D. M. D. E. B. U. S. M. S. P. H. D. F. W. G. *Cleft Lip and Palate : From Origin to Treatment: From Origin to Treatment*. (Oxford University Press, USA, 2002).

51 Devi, E. S., Sai Sankar, A. J., Manoj Kumar, M. G. & Sujatha, B. Maiden morsel - feeding in cleft lip and palate infants. *Journal of International Society of Preventive & Community Dentistry* **2**, 31-37, doi:10.4103/2231-0762.109350 (2012).

52 Coran, A. G. *et al. Pediatric Surgery: Expert Consult - Online and Print*. (Elsevier Health Sciences, 2012).

53 Vieira, N. A., Borgo, H. C., da Silva Dalben, G., Bachega, M. I. & Pereira, P. C. Evaluation of fecal microorganisms of children with cleft palate before and after palatoplasty. *Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology]* **44**, 835-838 (2013).

54 Desalu, I., Adeyemo, W., Akintimoye, M. & Adepoju, A. Airway and respiratory complications in children undergoing cleft lip and palate repair. *Ghana medical journal* **44**, 16-20 (2010).

55 Kuo, C.-L., Lien, C.-F., Chu, C.-H. & Shiao, A.-S. Otitis media with effusion in children with cleft lip and palate: A narrative review. *International Journal of Pediatric Otorhinolaryngology* **77**, 1403-1409, doi:http://dx.doi.org/10.1016/j.ijporl.2013.07.015 (2013).

56 Takasaki, K., Sando, I., Balaban, C. D. & Ishijima, K. Postnatal development of eustachian tube cartilage. A study of normal and cleft palate cases. *International Journal of Pediatric Otorhinolaryngology* **52**, 31-36, doi:http://dx.doi.org/10.1016/S0165-5876(99)00292-X (2000).

57 Akcam, M. O., Evirgen, S., Uslu, O. & Memikoğlu, U. T. Dental anomalies in individuals with cleft lip and/or palate. *The European Journal of Orthodontics* **32**, 207-213, doi:10.1093/ejo/cjp156 (2010).

58 Camporesi, M., Baccetti, T., Marinelli, A., Defraia, E. & Franchi, L. Maxillary dental anomalies in children with cleft lip and palate: a controlled study. *International Journal of Paediatric Dentistry* **20**, 442-450, doi:10.1111/j.1365-263X.2010.01063.x (2010).

59 Schroeder, D. C. & Green, L. J. Frequency of Dental Trait Anomalies in Cleft, Sibling, and Noncleft Groups. *Journal of Dental Research* **54**, 802-807, doi:10.1177/00220345750540041801 (1975).

60 Letra, A., Menezes, R., Granjeiro, J. M. & Vieira, A. R. Defining Subphenotypes for Oral Clefts Based on Dental Development. *Journal of Dental Research* **86**, 986-991, doi:10.1177/154405910708601013 (2007).

61 Al Jamal, G. A., Hazza'a, A. M. & Rawashdeh, M. a. A. Prevalence of Dental Anomalies in a Population of Cleft Lip and Palate Patients. *The Cleft Palate-Craniofacial Journal* **47**, 413-420, doi:10.1597/08-275.1 (2010).

62 Shetye, P. R. Facial growth of adults with unoperated clefts. *Clinics in Plastic Surgery* **31**, 361-371, doi:http://dx.doi.org/10.1016/S0094-1298(03)00137-8 (2004).

63 Vettore, M. V. & Sousa Campos, A. E. Malocclusion characteristics of patients with cleft lip and/or palate. *The European Journal of Orthodontics* **33**, 311-317, doi:10.1093/ejo/cjq078 (2011).

64 Baek, S. H., Moon, H. S. & Yang, W. S. Cleft type and Angle's classification of malocclusion in Korean cleft patients. *The European Journal of Orthodontics* **24**, 647-653, doi:10.1093/ejo/24.6.647 (2002).

65 Kuehn, D. P. & Moller, K. T. Speech and Language Issues in the Cleft Palate Population: The State of the Art. *The Cleft Palate-Craniofacial Journal* **37**, 348-348, doi:10.1597/1545-1569(2000)037<0348:SALIIT>2.3.CO;2 (2000).

66 Wyatt, R. *et al.* Cleft palate speech dissected: a review of current knowledge and analysis. *British Journal of Plastic Surgery* **49**, 143-149, doi:http://dx.doi.org/10.1016/S0007-1226(96)90216-7 (1996).

67 Leslie, E. J. & Marazita, M. L. Genetics of cleft lip and cleft palate. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics* **163**, 246-258, doi:10.1002/ajmg.c.31381 (2013).

68 Rahimov, F., Jugessur, A. & Murray, J. C. Genetics of Nonsyndromic Orofacial Clefts. *The Cleft Palate-Craniofacial Journal* **49**, 73-91, doi:10.1597/10-178 (2011).

69 Grosen, D. *et al.* Risk of oral clefts in twins. *Epidemiology (Cambridge, Mass.)* **22**, 313-319, doi:10.1097/EDE.0b013e3182125f9c (2011).

70 Setó-Salvia, N. & Stanier, P. Genetics of cleft lip and/or cleft palate: Association with other common anomalies. *European Journal of Medical Genetics* **57**, 381-393, doi:http://dx.doi.org/10.1016/j.ejmg.2014.04.003 (2014).

71 Meng, L., Bian, Z., Torensma, R. & Von den Hoff, J. W. Biological mechanisms in palatogenesis and cleft palate. *J Dent Res* **88**, 22-33, doi:10.1177/0022034508327868 (2009).

72 Riley, B. M. *et al.* A genome-wide linkage scan for cleft lip and cleft palate identifies a novel locus on 8p11-23. *American Journal of Medical Genetics Part A* **143A**, 846-852, doi:10.1002/ajmg.a.31673 (2007).

73 Riley, B. M. *et al.* Impaired FGF signaling contributes to cleft lip and palate. *Proceedings of the National Academy of Sciences* **104**, 4512-4517, doi:10.1073/pnas.0607956104 (2007).

74 Snyder-Warwick, A. K. *et al.* Analysis of a gain-of-function FGFR2 Crouzon mutation provides evidence of loss of function activity in the etiology of cleft palate. *Proceedings of the National Academy of Sciences* **107**, 2515-2520, doi:10.1073/pnas.0913985107 (2010).

75 De Moerlooze, L. *et al.* An important role for the IIIb isoform of fibroblast growth factor receptor 2 (FGFR2) in mesenchymal-epithelial signalling during mouse organogenesis. *Development* **127**, 483-492 (2000).

76 Rice, R. *et al.* Disruption of Fgf10/Fgfr2b-coordinated epithelial-mesenchymal interactions causes cleft palate. *The Journal of Clinical Investigation* **113**, 1692-1700, doi:10.1172/JCI20384 (2004).

77 Gritli-Linde, A. p63 and IRF6: brothers in arms against cleft palate. *J Clin Invest* **120**, 1386-1389, doi:10.1172/jci42821 (2010).

78 Ingraham, C. R. *et al.* Abnormal skin, limb and craniofacial morphogenesis in mice deficient for interferon regulatory factor 6 (Irf6). *Nat Genet* **38**, 1335-1340, doi:10.1038/ng1903 (2006).

79 Moretti, F. *et al.* A regulatory feedback loop involving p63 and IRF6 links the pathogenesis of 2 genetically different human ectodermal dysplasias. *The Journal of Clinical Investigation* **120**, 1570-1577, doi:10.1172/JCI40267 (2010).

80 Geiss-Friedlander, R. & Melchior, F. Concepts in sumoylation: a decade on. *Nat Rev Mol Cell Biol* **8**, 947-956, doi:http://www.nature.com/nrm/journal/v8/n12/suppinfo/nrm2293_S1.html (2007).

81 Carter, T. C. *et al.* Testing reported associations of genetic risk factors for oral clefts in a large Irish study population. *Birth Defects Research Part A: Clinical and Molecular Teratology* **88**, 84-93, doi:10.1002/bdra.20639 (2010).

82 Jia, Z. L., Shi, B., Xu, X. & Kong, X. L. Interactions between small ubiquitin-like modifier 1 and nonsyndromic orofacial clefts. *DNA and cell biology* **30**, 235-240, doi:10.1089/dna.2010.1110 (2011).

83 Alkuraya, F. S. *et al.* SUMO1 Haploinsufficiency Leads to Cleft Lip and Palate. *Science* **313**, 1751, doi:10.1126/science.1128406 (2006).

84 Melkoniemi, M. *et al.* Collagen XI sequence variations in nonsyndromic cleft palate, Robin sequence and micrognathia. *Eur J Hum Genet* **11**, 265-270, doi:10.1038/sj.ejhg.5200950 (2003).

85 Lavrin, I. O., McLean, W., Seegmiller, R. E., Olsen, B. R. & Hay, E. D. The mechanism of palatal clefting in the Col11a1 mutant mouse. *Archives of Oral Biology* **46**, 865-869, doi:http://dx.doi.org/10.1016/S0003-9969(01)00044-9 (2001).

86 van Aalst, J. A., Kolappa, K. K. & Sadove, M. MOC-PSSM CME article: Nonsyndromic cleft palate. *Plast Reconstr Surg* **121**, 1-14, doi:10.1097/01.prs.0000294706.05898.f3 (2008).

87 Honein, M. A. *et al.* Maternal smoking and environmental tobacco smoke exposure and the risk of orofacial clefts. *Epidemiology (Cambridge, Mass.)* **18**, 226-233, doi:10.1097/01.ede.0000254430.61294.c0 (2007).

88 Rothman, K. J. *et al.* Teratogenicity of High Vitamin A Intake. *New England Journal of Medicine* **333**, 1369-1373, doi:doi:10.1056/NEJM199511233332101 (1995).

89 Shaw, G. M. *et al.* Orofacial clefts, parental cigarette smoking, and transforming growth factor-alpha gene variants. *Am J Hum Genet* **58**, 551-561 (1996).

90 Hwang, S.-J. *et al.* Association Study of Transforming Growth Factor Alpha (TGFα) TaqI Polymorphismand Oral Clefts: Indication of Gene-Environment Interaction in a Population-based Sample of Infants with Birth Defects. *American Journal of Epidemiology* **141**, 629-636 (1995).

91 Fraser, F. C., Kalter, H., Walker, B. E. & Fainstat, T. D. The experimental production of cleft palate with cortisone and other hormones. *Journal of cellular physiology. Supplement* **43**, 237-259 (1954).

92 Juriloff, D. M. & Harris, M. J. Mouse genetic models of cleft lip with or without cleft palate. *Birth Defects Research Part A: Clinical and Molecular Teratology* **82**, 63-77, doi:10.1002/bdra.20430 (2008).

93 Gritli-Linde, A. The mouse as a developmental model for cleft lip and palate research. *Front Oral Biol* **16**, 32-51, doi:10.1159/000337523 (2012).

94 Juriloff, D. M., Harris, M. J. & Brown, C. J. Unravelling the complex genetics of cleft lip in the mouse model. *Mammalian Genome* **12**, 426-435, doi:10.1007/s003350010284 (2001).

95 Vekemans, M. & Fraser, F. C. Stage of palate closure as one indication of "liability" to cleft palate. *American Journal of Medical Genetics* **4**, 95-102, doi:10.1002/ajmg.1320040111 (1979).

96 Schwartz, J. L., Jordan, R., Sun, J., Ma, H. & Hsieb, A. W. Dose-dependent changes in the spectrum of mutations induced by ionizing radiation. *Radiation research* **153**, 312-317 (2000).

97 Grosovsky, A. J., de Boer, J. G., de Jong, P. J., Drobetsky, E. A. & Glickman, B. W. Base substitutions, frameshifts, and small deletions constitute ionizing radiation-induced point mutations in mammalian cells. *Proceedings of the National Academy of Sciences* **85**, 185-188 (1988).

98 Breimer, L. H. Ionizing radiation-induced mutagenesis. *Br J Cancer* **57**, 6-18 (1988).

99 Culiat, C. T. *et al.* Concordance between isolated cleft palate in mice and alterations within a region including the gene encoding the beta 3 subunit of the type A gamma-aminobutyric acid receptor. *Proceedings of the National Academy of Sciences* **90**, 5105-5109 (1993).

100 Acevedo-Arozena, A. *et al.* ENU Mutagenesis, a Way Forward to Understand Gene Function. *Annual Review of Genomics and Human Genetics* **9**, 49-69, doi:doi:10.1146/annurev.genom.9.081307.164224 (2008).

101 Probst, F. J. & Justice, M. J. in *Methods in Enzymology* Vol. Volume 477 (eds M. Wassarman Paul & M. Soriano Philippe) 297-312 (Academic Press, 2010).

102 Bjork, B. C., Turbe-Doan, A., Prysak, M., Herron, B. J. & Beier, D. R. Prdm16 is required for normal palatogenesis in mice. *Hum Mol Genet* **19**, 774-789, doi:10.1093/hmg/ddp543 (2010).

103 Wu, S., Ying, G., Wu, Q. & Capecchi, M. R. A protocol for constructing gene targeting vectors: generating knockout mice for the cadherin family and beyond. *Nat. Protocols* **3**, 1056-1076, doi:http://www.nature.com/nprot/journal/v3/n6/suppinfo/nprot.2008.70_S1.html (2008).

104 Doyle, A., McGarry, M., Lee, N. & Lee, J. The construction of transgenic and gene knockout/knockin mouse models of human disease. *Transgenic Research* **21**, 327-349, doi:10.1007/s11248-011-9537-3 (2012).

105 Manis, J. P. Knock Out, Knock In, Knock Down — Genetically Manipulated Mice and the Nobel Prize. *New England Journal of Medicine* **357**, 2426-2429, doi:doi:10.1056/NEJMp0707712 (2007).

106 Proetzel, G. *et al.* Transforming growth factor-[beta]3 is required for secondary palate fusion. *Nat Genet* **11**, 409-414 (1995).

107 Satokata, I. & Maas, R. Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development. *Nat Genet* **6**, 348-356 (1994).

108 Richardson, R. J. *et al.* Irf6 is a key determinant of the keratinocyte proliferation-differentiation switch. *Nat Genet* **38**, 1329-1334, doi:10.1038/ng1894 (2006).

109 Xu, X. *et al.* Cell autonomous requirement for Tgfbr2 in the disappearance of medial edge epithelium during palatal fusion. *Developmental Biology* **297**, 238-248, doi:http://dx.doi.org/10.1016/j.ydbio.2006.05.014 (2006).

110 Gu, S. *et al.* Mice with an anterior cleft of the palate survive neonatal lethality. *Developmental Dynamics* **237**, 1509-1516, doi:10.1002/dvdy.21534 (2008).

111 Gridley, T., Soriano, P. & Jaenisch, R. Insertional mutagenesis in mice. *Trends in Genetics* **3**, 162-166, doi:http://dx.doi.org/10.1016/0168-9525(87)90218-6 (1987).

112 Meisler, M. H. Insertional mutation of 'classical' and novel genes in transgenic mice. *Trends in Genetics* **8**, 341-344, doi:http://dx.doi.org/10.1016/0168-9525(92)90278-C (1992).

113 Li, W., Puertollano, R., Bonifacino, J. S., Overbeek, P. A. & Everett, E. T. Disruption of the murine Ap2beta1 gene causes nonsyndromic cleft palate. *Cleft Palate Craniofac J* **47**, 566-573, doi:10.1597/09-145 (2010).

114 Carlson, C. M. & Largaespada, D. A. Insertional mutagenesis in mice: new perspectives and tools. *Nat Rev Genet* **6**, 568-580 (2005).

115 Stanford, W. L., Cohn, J. B. & Cordes, S. P. Gene-trap mutagenesis: past, present and beyond. *Nat Rev Genet* **2**, 756-768 (2001).

116 Song, G. *et al.* Effective Gene Trapping Mediated by <italic>Sleeping Beauty</italic> Transposon. *PLoS ONE* **7**, e44123, doi:10.1371/journal.pone.0044123 (2012).

117 Okada, I. *et al.* SMOC1 Is Essential for Ocular and Limb Development in Humans and Mice. *The American Journal of Human Genetics* **88**, 30-41, doi:http://dx.doi.org/10.1016/j.ajhg.2010.11.012 (2011).

118 Tyl, R. W., Chernoff, N. & Rogers, J. M. Altered axial skeletal development. *Birth Defects Research Part B: Developmental and Reproductive Toxicology* **80**, 451-472, doi:10.1002/bdrb.20134 (2007).

119 Marcelle, C., Lesbros, C. & Linker, C. in *Vertebrate Myogenesis* Vol. 38 *Results and Problems in Cell Differentiation* (ed Beate Brand-Saberi) Ch. 4, 81-108 (Springer Berlin Heidelberg, 2002).

120 Mallo, M., Vinagre, T. & Carapuco, M. The road to the vertebral formula. *Int J Dev Biol* **53**, 1469-1481, doi:10.1387/ijdb.072276mm (2009).

121 Saga, Y. The mechanism of somite formation in mice. *Current Opinion in Genetics & Development* **22**, 331-338, doi:http://dx.doi.org/10.1016/j.gde.2012.05.004 (2012).

122 Dequeant, M.-L. & Pourquie, O. Segmental patterning of the vertebrate embryonic axis. *Nat Rev Genet* **9**, 370-382 (2008).

123 Saga, Y. & Takeda, H. The making of the somite: molecular events in vertebrate segmentation. *Nat Rev Genet* **2**, 835-845 (2001).

124 Resende, T. P., Andrade, R. P. & Palmeirim, I. Timing embryo segmentation: dynamics and regulatory mechanisms of the vertebrate segmentation clock. *BioMed research international* **2014**, 718683, doi:10.1155/2014/718683 (2014).

125 Iulianella, A., Melton, K. R. & Trainor, P. A. Somitogenesis: Breaking New Boundaries. *Neuron* **40**, 11-14, doi:http://dx.doi.org/10.1016/S0896-6273(03)00604-4 (2003).

126 Mead, T. J. & Yutzey, K. E. Notch pathway regulation of chondrocyte differentiation and proliferation during appendicular and axial skeleton development. *Proceedings of the National Academy of Sciences* **106**, 14420-14425, doi:10.1073/pnas.0902306106 (2009).

127 Saladin, K. *Human Anatomy' 2007 Ed.2007 Edition*. (Rex Bookstore, Inc.).

128 Twietmeyer, A. & McCracken, T. *Coloring Guide to Human Anatomy*. 18 (Lippincott Williams & Wilkins, 2001).

129 Carapuço, M., Nóvoa, A., Bobola, N. & Mallo, M. Hox genes specify vertebral types in the presomitic mesoderm. *Genes & Development* **19**, 2116-2121, doi:10.1101/gad.338705 (2005).

130 Casaca, A., Santos, A. C. & Mallo, M. Controlling Hox gene expression and activity to build the vertebrate axial skeleton. *Developmental Dynamics* **243**, 24-36, doi:10.1002/dvdy.24007 (2014).

131 Pang, D. & Thompson, D. P. Embryology and bony malformations of the craniovertebral junction. *Child's Nervous System* **27**, 523-564, doi:10.1007/s00381-010-1358-9 (2011).

132 Alexander, T., Nolte, C. & Krumlauf, R. Hox Genes and Segmentation of the Hindbrain and Axial Skeleton. *Annual Review of Cell and Developmental Biology* **25**, 431-456, doi:doi:10.1146/annurev.cellbio.042308.113423 (2009).

133 Wellik, D. M. Hox patterning of the vertebrate axial skeleton. *Developmental Dynamics* **236**, 2454-2463, doi:10.1002/dvdy.21286 (2007).

134 van Nes, J. *et al.* The Cdx4 mutation affects axial development and reveals an essential role of Cdx genes in the ontogenesis of the placental labyrinth in mice. *Development* **133**, 419-428, doi:10.1242/dev.02216 (2006).

135 Partanen, J., Schwartz, L. & Rossant, J. Opposite phenotypes of hypomorphic and Y766 phosphorylation site mutations reveal a function for Fgfr1 in anteroposterior patterning of mouse embryos. *Genes & Development* **12**, 2332-2344, doi:10.1101/gad.12.15.2332 (1998).

136 Guttentag, A. R. & Salwen, J. K. Keep Your Eyes on the Ribs: The Spectrum of Normal Variants and Diseases That Involve the Ribs. *RadioGraphics* **19**, 1125-1142, doi:doi:10.1148/radiographics.19.5.g99se011125 (1999).

137 Nakajima, A. *et al.* The prevalence of morphological changes in the thoracolumbar spine on whole-spine computed tomographic images. *Insights into imaging* **5**, 77-83, doi:10.1007/s13244-013-0286-0 (2014).

138 Rengasamy, P. & Padmanabhan, R. R. Experimental studies on cervical and lumbar ribs in mouse embryos. *Congenital Anomalies* **44**, 156-171, doi:10.1111/j.1741-4520.2004.00029.x (2004).

139 Zoltewicz, J. S., Plummer, N. W., Lin, M. I. & Peterson, A. S. oto is a homeotic locus with a role in anteroposterior development that is partially redundant with Lim1. *Development* **126**, 5085-5095 (1999).

140 Hustert, E., Scherer, G., Olowson, M., Guénet, J. L. & Balling, R. Rbt (rabo torcido), a new mouse skeletal mutation involved in anteroposterior patterning of the axial skeleton, maps close to the ts (tail-short) locus and distal to the sox9 locus on chromosome 11. *Mammalian Genome* **7**, 881-885, doi:10.1007/s003359900261 (1996).

141 Chen, F. & Capecchi, M. R. Targeted Mutations inHoxa-9andHoxb-9Reveal Synergistic Interactions. *Developmental Biology* **181**, 186-196, doi:http://dx.doi.org/10.1006/dbio.1996.8440 (1997).

142 Schmahl, J., Raymond, C. S. & Soriano, P. PDGF signaling specificity is mediated through multiple immediate early genes. *Nat Genet* **39**, 52-60, doi:http://www.nature.com/ng/journal/v39/n1/suppinfo/ng1922_S1.html (2007).

143 Palmiter, R. D. & Brinster, R. L. Germ-Line Transformation of Mice. *Annual Review of Genetics* **20**, 465-499, doi:doi:10.1146/annurev.ge.20.120186.002341 (1986).

144 Overbeek, P. A. in *Transgenic Animal Technology (Third Edition)* (ed Carl A. Pinkert) 71-107 (Elsevier, 2014).

145 Woychik, R. P. & Alagramam, K. Insertional mutagenesis in transgenic mice generated by the pronuclear microinjection procedure. *Int J Dev Biol* **42**, 1009-1017 (1998).

146 Wurtele, H., Little, K. C. & Chartrand, P. Illegitimate DNA integration in mammalian cells. *Gene Ther* **10**, 1791-1799, doi:10.1038/sj.gt.3302074 (2003).

147 Schrick, J. J., Dickinson, M. E., Hogan, B. L., Selby, P. B. & Woychik, R. P. Molecular and phenotypic characterization of a new mouse insertional mutation that causes a defect in the distal vertebrae of the spine. *Genetics* **140**, 1061-1067 (1995).

148 Reith, A. D. & Bernstein, A. Molecular basis of mouse developmental mutants. *Genes & Development* **5**, 1115-1123, doi:10.1101/gad.5.7.1115 (1991).

149 Lu, W. *et al.* Insertional Mutation of the Collagen Genes Col4a3 and Col4a4 in a Mouse Model of Alport Syndrome. *Genomics* **61**, 113-124, doi:http://dx.doi.org/10.1006/geno.1999.5943 (1999).

150 Bishop, C. E. *et al.* A transgenic insertion upstream of Sox9 is associated with dominant XX sex reversal in the mouse. *Nat Genet* **26**, 490-494 (2000).

151 Moayedi, Y. *et al.* The Candidate Splicing Factor Sfswap Regulates Growth and Patterning of Inner Ear Sensory Organs. *PLoS Genet* **10**, e1004055, doi:10.1371/journal.pgen.1004055 (2014).

152 Yang, T. *et al.* Epidermal detachment, desmosomal dissociation, and destabilization of corneodesmosin in Spink5-/- mice. *Genes & Development* **18**, 2354-2358, doi:10.1101/gad.1232104 (2004).

153 Morgan, D. *et al.* Inversin, a novel gene in the vertebrate left-right axis pathway, is partially deleted in the inv mouse. *Nat Genet* **20**, 149-156 (1998).

154 Overbeek, P. A. *et al.* A transgenic insertion causing cryptorchidism in mice. *genesis* **30**, 26-35, doi:10.1002/gene.1029 (2001).

155 Caburet, S. *et al.* Mutant Cohesin in Premature Ovarian Failure. *New England Journal of Medicine* **370**, 943-949, doi:doi:10.1056/NEJMoa1309635 (2014).

156 Hearing, V. J. & Tsukamoto, K. Enzymatic control of pigmentation in mammals. *The FASEB Journal* **5**, 2902-2909 (1991).

157 Yokoyama, T. *et al.* Conserved cysteine to serine mutation in tyrosinase is responsible for the classical albino mutation in laboratory mice. *Nucleic Acids Research* **18**, 7293-7298, doi:10.1093/nar/18.24.7293 (1990).

158 Katz L, R. K., Overbeek PA, Hoglund V, Everett ET Reduced expression of COL11A1 associated with cleft palate and chondrodysplasia. *J Dent Res 93(Spec Iss A):1382, 2* (2014).

159 Alazzam M, B. E., Ryan K, Overbeek PA, Everett ET. Mutation of Runx1t1: Growth Delay, Cleft Palate and Rib Anomalies. *J Dent Res 93(Spec Iss A): 1390*, doi:www.dentalresearch.org (2014).

160 Hawley, R. G. Versatile retroviral vectors for potential use in gene therapy. *Gene therapy* **1**, 136-138 (1994).

161 McLeod, M. J. Differential staining of cartilage and bone in whole mouse fetuses by alcian blue and alizarin red S. *Teratology* **22**, 299-301, doi:10.1002/tera.1420220306 (1980).

162 Perlyn, C. A. *et al.* The Craniofacial Phenotype of the Crouzon Mouse: Analysis of a Model for Syndromic Craniosynostosis Using Three-Dimensional MicroCT. *The Cleft Palate-Craniofacial Journal* **43**, 740-748, doi:10.1597/05-212 (2006).

163 van Den Akker, E. *et al.* Axial skeletal patterning in mice lacking all paralogous group 8 Hox genes. *Development* **128**, 1911-1921 (2001).

164 Fromental-Ramain, C. *et al.* Specific and redundant functions of the paralogous Hoxa-9 and Hoxd-9 genes in forelimb and axial skeleton patterning. *Development* **122**, 461-472 (1996).

165 Suemori, H., Takahashi, N. & Noguchi, S. Hoxc-9 mutant mice show anterior transformation of the vertebrae and malformation of the sternum and ribs. *Mechanisms of development* **51**, 265-273 (1995).

166 Rijli, F. M. *et al.* Cryptorchidism and homeotic transformations of spinal nerves and vertebrae in Hoxa-10 mutant mice. *Proceedings of the National Academy of Sciences* **92**, 8185-8189 (1995).

167 Wellik, D. M. & Capecchi, M. R. Hox10 and Hox11 Genes Are Required to Globally Pattern the Mammalian Skeleton. *Science* **301**, 363-367, doi:10.1126/science.1085672 (2003).

168 Doetschman, T. in *Gene Knockout Protocols* Vol. 530 *Methods in Molecular Biology* (eds Wolfgang Wurst & Ralf Kühn) Ch. 23, 423-433 (Humana Press, 2009).

169 Smith, K. Theoretical mechanisms in targeted and random integration of transgene DNA. *Reprod Nutr Dev* **41**, 465-485 (2001).

170 Yan, B.-W., Zhao, Y.-F., Cao, W.-G., Li, N. & Gou, K.-M. Mechanism of random integration of foreign DNA in transgenic mice. *Transgenic Research* **22**, 983-992, doi:10.1007/s11248-013-9701-z (2013).

171 Folger, K. R., Wong, E. A., Wahl, G. & Capecchi, M. R. Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules. *Molecular and Cellular Biology* **2**, 1372-1387, doi:10.1128/mcb.2.11.1372 (1982).

172 Roboz, J. *Mass Spectrometry in Cancer Research*. (Taylor & Francis, 2002).

173 Davis, J. N., McGhee, L. & Meyers, S. The ETO (MTG8) gene family. *Gene* **303**, 1-10, doi:http://dx.doi.org/10.1016/S0378-1119(02)01172-1 (2003).

174 Hug, B. A. & Lazar, M. A. ETO interacting proteins. *Oncogene* **23**, 4270-4274, doi:10.1038/sj.onc.1207674 (2004).

175 Miyoshi, H. *et al.* t(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, AML1. *Proceedings of the National Academy of Sciences* **88**, 10431-10434 (1991).

176 Erickson, P. G. J. C. K. S. L. T. W. E. R. S. L. R. T. J. R. J. D. H. Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, AML1/ETO, with similarity to Drosophila segmentation gene, runt. *Blood* **80**, 1825-1831 (1992).

177 Scandura, J. M., Boccuni, P., Cammenga, J. & Nimer, S. D. Transcription factor fusions in acute leukemia: variations on a theme. *Oncogene* **21**, 3422-3444, doi:10.1038/sj.onc.1205315 (2002).

178 Ferrara, F. & Schiffer, C. A. Acute myeloid leukaemia in adults. *The Lancet* **381**, 484-495, doi:http://dx.doi.org/10.1016/S0140-6736(12)61727-9.

179 Mrózek, K., Heerema, N. A. & Bloomfield, C. D. Cytogenetics in acute leukemia. *Blood Reviews* **18**, 115-136, doi:http://dx.doi.org/10.1016/S0268-960X(03)00040-7 (2004).

180 Peterson, L. F. B. A. A. E.-Y. B. J. R. O. A. J. L. M.-C. Y. M. Z. D.-E. Acute myeloid leukemia with the 8q22;21q22 translocation: secondary mutational events and alternative t(8;21) transcripts. *Blood* **110**, 799-805 (2007).

181 Ajore, R., Kumar, P., Dhanda, R., Gullberg, U. & Olsson, I. The leukemia associated nuclear corepressor ETO homologue genes MTG16 and MTGR1 are regulated differently in hematopoietic cells. *BMC Molecular Biol* **13**, 1-16, doi:10.1186/1471-2199-13-11 (2012).

182 Guastadisegni, M. C. *et al.* CBFA2T2 and C20orf112: two novel fusion partners of RUNX1 in acute myeloid leukemia. *Leukemia* **24**, 1516-1519, doi:http://www.nature.com/leu/journal/v24/n8/suppinfo/leu2010106s1.html (2010).

183 Erickson, P. F., Robinson, M., Owens, G. & Drabkin, H. A. The ETO portion of acute myeloid leukemia t(8;21) fusion transcript encodes a highly evolutionarily conserved, putative transcription factor. *Cancer Res* **54**, 1782-1786 (1994).

184 Eppig, J. T. *et al.* The Mouse Genome Database (MGD): comprehensive resource for genetics and genomics of the laboratory mouse. *Nucleic Acids Research* **40**, D881-D886, doi:10.1093/nar/gkr974 (2012).

185 Odaka, Y., Mally, A., Elliott, L. T. & Meyers, S. Nuclear import and subnuclear localization of the proto-oncoprotein ETO (MTG8). *Oncogene* **19**, 3584-3597, doi:10.1038/sj.onc.1203689 (2000).

186 Hildebrand, D., Tiefenbach, J., Heinzel, T., Grez, M. & Maurer, A. B. Multiple Regions of ETO Cooperate in Transcriptional Repression. *Journal of Biological Chemistry* **276**, 9889-9895, doi:10.1074/jbc.M010582200 (2001).

187 Wang, J., Hoshino, T., Redner, R. L., Kajigaya, S. & Liu, J. M. ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex. *Proceedings of the National Academy of Sciences* **95**, 10860-10865 (1998).

188 Licht, J. D. AML1 and the AML1-ETO fusion protein in the pathogenesis of t(8;21) AML. *Oncogene* **20**, 5660-5679, doi:10.1038/sj.onc.1204593 (2001).

189 Kozu, T. *et al.* Significance of MTG8 in leukemogenesis. *Leukemia* **11 Suppl 3**, 297-298 (1997).

190 Lutterbach, B. *et al.* ETO, a Target of t(8;21) in Acute Leukemia, Interacts with the N-CoR and mSin3 Corepressors. *Molecular and Cellular Biology* **18**, 7176-7184 (1998).

191 Erickson, P. F. D. G. L. R. S. P. G. R. M. D. H. A. ETO and AML1 phosphoproteins are expressed in CD34+ hematopoietic progenitors: implications for t(8;21) leukemogenesis and monitoring residual disease. *Blood* **88**, 1813-1823 (1996).

192 Komori, A., Sueoka, E., Fujiki, H., Ishii, M. & Kozu, T. Association of MTG8 (ETO/CDR), a leukemia-related protein, with serine/threonine protein kinases and heat shock protein HSP90 in human hematopoietic cell lines. *Jpn J Cancer Res* **90**, 60-68 (1999).

193 Davis, J. N., Williams, B. J., Herron, J. T., Galiano, F. J. & Meyers, S. ETO-2, a new member of the ETO-family of nuclear proteins. *Oncogene* **18**, 1375-1383, doi:10.1038/sj.onc.1202412 (1999).

194 *Gene Expression Database (GXD), Mouse Genome Informatics Web Site. World Wide Web*, <URL: http://www.informatics.jax.org> (10, 2014).

195 Miyoshi, H. *et al.* The t(8;21) translocation in acute myeloid leukemia results in production of an AML1-MTG8 fusion transcript. *EMBO J* **12**, 2715-2721 (1993).

196 Petryszak, R. *et al.* Expression Atlas update—a database of gene and transcript expression from microarray- and sequencing-based functional genomics experiments. *Nucleic Acids Research* **42**, D926-D932, doi:10.1093/nar/gkt1270 (2014).

197 Richardson, L. *et al.* EMAGE mouse embryo spatial gene expression database: 2014 update. *Nucleic Acids Research* **42**, D835-D844, doi:10.1093/nar/gkt1155 (2014).

198 Rochford, J. J. *et al.* ETO/MTG8 Is an Inhibitor of C/EBPβ Activity and a Regulator of Early Adipogenesis. *Molecular and Cellular Biology* **24**, 9863-9872, doi:10.1128/mcb.24.22.9863-9872.2004 (2004).

199 Sacchi, N. *et al.* Subcellular localization of the oncoprotein MTG8 (CDR/ETO) in neural cells. *Oncogene* **16**, 2609-2615, doi:10.1038/sj.onc.1201824 (1998).

200 Terman, J. R. & Kolodkin, A. L. Response to Comment on "Nervy Links Protein Kinase A to Plexin-Mediated Semaphorin Repulsion". *Science* **309**, 558, doi:10.1126/science.1109259 (2005).

201 Peterson, L. F. & Zhang, D. E. The 8;21 translocation in leukemogenesis. *Oncogene* **23**, 4255-4262, doi:10.1038/sj.onc.1207727 (2004).

202 Ahn, M.-Y. *et al.* Negative regulation of granulocytic differentiation in the myeloid precursor cell line 32Dcl3 by ear-2, a mammalian homolog of Drosophila seven-up, and a chimeric leukemogenic gene, AML1/ETO(MTG8). *Proceedings of the National Academy of Sciences* **95**, 1812-1817 (1998).

203 Kitabayashi, I. *et al.* The AML1-MTG8 Leukemic Fusion Protein Forms a Complex with a Novel Member of the MTG8(ETO/CDR) Family, MTGR1. *Molecular and Cellular Biology* **18**, 846-858 (1998).

204 Elsasser, A. *et al.* The fusion protein AML1-ETO in acute myeloid leukemia with translocation t(8;21) induces c-jun protein expression via the proximal AP-1 site of the c-jun promoter in an indirect, JNK-dependent manner. *Oncogene* **22**, 5646-5657, doi:10.1038/sj.onc.1206673 (2003).

205 Klampfer, L., Zhang, J., Zelenetz, A. O., Uchida, H. & Nimer, S. D. The AML1/ETO fusion protein activates transcription of BCL-2. *Proceedings of the National Academy of Sciences* **93**, 14059-14064 (1996).

206 Reikvam, H., Hatfield, K. J., Kittang, A. O., Hovland, R. & Bruserud, O. Acute myeloid leukemia with the t(8;21) translocation: clinical consequences and biological implications. *J Biomed Biotechnol* **2011**, 104631, doi:10.1155/2011/104631 (2011).

207 Yeh, K.-T. *et al.* Aberrant TGFβ/SMAD4 signaling contributes to epigenetic silencing of a putative tumor suppressor, RunX1T1 in ovarian cancer. *Epigenetics* **6**, 727-739 (2011).

208 Kan, Z. *et al.* Diverse somatic mutation patterns and pathway alterations in human cancers. *Nature* **466**, 869-873, doi:http://www.nature.com/nature/journal/v466/n7308/abs/nature09208.html#supplementary-information (2010).

209 Kim, Y. R., Kim, M. S., Lee, S. H. & Yoo, N. J. Mutational analysis of RUNX1T1 gene in acute leukemias, breast and lung carcinomas. *Leukemia Research* **35**, e157-e158, doi:http://dx.doi.org/10.1016/j.leukres.2011.04.024 (2011).

210 Smith, I. M., Mithani, S. K., Liu, C. & et al. NOvel integrative methods for gene discovery associated with head and neck squamous cell carcinoma development. *Archives of Otolaryngology–Head & Neck Surgery* **135**, 487-495, doi:10.1001/archoto.2009.43 (2009).

211 Aaker, J. D. *et al.* Interaction of MTG family proteins with NEUROG2 and ASCL1 in the developing nervous system. *Neuroscience Letters* **474**, 46-51, doi:http://dx.doi.org/10.1016/j.neulet.2010.03.004 (2010).

212 Alishahi, A., Koyano-Nakagawa, N. & Nakagawa, Y. Regional expression of MTG genes in the developing mouse central nervous system. *Developmental Dynamics* **238**, 2095-2102, doi:10.1002/dvdy.22021 (2009).

213 Oropez, D. & Horb, M. Transient expression of Ngn3 in Xenopus endoderm promotes early and ectopic development of pancreatic beta and delta cells. *genesis* **50**, 271-285, doi:10.1002/dvg.20828 (2012).

214 Calabi, F., Pannell, R. & Pavloska, G. Gene targeting reveals a crucial role for MTG8 in the gut. *Mol Cell Biol* **21**, 5658-5666, doi:10.1128/mcb.21.16.5658-5666.2001 (2001).

215 Flicek, P. *et al.* Ensembl 2014. *Nucleic Acids Res* **42**, D749-755, doi:10.1093/nar/gkt1196 (2014).

216 Shi, X., Sun, M., Liu, H., Yao, Y. & Song, Y. Long non-coding RNAs: A new frontier in the study of human diseases. *Cancer Letters* **339**, 159-166, doi:http://dx.doi.org/10.1016/j.canlet.2013.06.013 (2013).

217 Gibb, E. A., Brown, C. J. & Lam, W. L. The functional role of long non-coding RNA in human carcinomas. *Mol Cancer* **10**, 38, doi:10.1186/1476-4598-10-38 (2011).

218 Kornienko, A. E., Guenzl, P. M., Barlow, D. P. & Pauler, F. M. Gene regulation by the act of long non-coding RNA transcription. *BMC Biol* **11**, 59, doi:10.1186/1741-7007-11-59 (2013).

219 Ulitsky, I. & Bartel, David P. lincRNAs: Genomics, Evolution, and Mechanisms. *Cell* **154**, 26-46, doi:http://dx.doi.org/10.1016/j.cell.2013.06.020 (2013).

220 Kondo, T. *et al.* Small Peptides Switch the Transcriptional Activity of Shavenbaby During Drosophila Embryogenesis. *Science* **329**, 336-339, doi:10.1126/science.1188158 (2010).

221 Ulveling, D., Francastel, C. & Hubé, F. When one is better than two: RNA with dual functions. *Biochimie* **93**, 633-644, doi:http://dx.doi.org/10.1016/j.biochi.2010.11.004 (2011).

222 Mercer, T. R. & Mattick, J. S. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat Struct Mol Biol* **20**, 300-307 (2013).

223 Kung, J. T. Y., Colognori, D. & Lee, J. T. Long Noncoding RNAs: Past, Present, and Future. *Genetics* **193**, 651-669, doi:10.1534/genetics.112.146704 (2013).

224 Pelechano, V. & Steinmetz, L. M. Gene regulation by antisense transcription. *Nat Rev Genet* **14**, 880-893, doi:10.1038/nrg3594 (2013).

225 Poliseno, L. *Pseudogenes: Newly Discovered Players in Human Cancer*. Vol. 5 (2012).

226 Geisler, S. & Coller, J. RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat Rev Mol Cell Biol* **14**, 699-712, doi:10.1038/nrm3679 (2013).

227 Rozen, S. & Skaletsky, H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**, 365-386 (2000).

228 Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105-1111, doi:10.1093/bioinformatics/btp120 (2009).

229 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).

230 Zerbino, D. R. & Birney, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* **18**, 821-829, doi:10.1101/gr.074492.107 (2008).

231 Schulz, M. H., Zerbino, D. R., Vingron, M. & Birney, E. Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* **28**, 1086-1092, doi:10.1093/bioinformatics/bts094 (2012).

232 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol* **215**, 403-410, doi:10.1016/s0022-2836(05)80360-2 (1990).

233 Northern blotting: transfer of denatured RNA to membranes. *Nat Meth* **2**, 997-998 (2005).

234 Kent, W. J. *et al.* The Human Genome Browser at UCSC. *Genome Research* **12**, 996-1006, doi:10.1101/gr.229102 (2002).

235 Ravasi, T. *et al.* An Atlas of Combinatorial Transcriptional Regulation in Mouse and Man. *Cell* **140**, 744-752, doi:http://dx.doi.org/10.1016/j.cell.2010.01.044 (2010).

236 Reva, B., Antipin, Y. & Sander, C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Research* **39**, e118, doi:10.1093/nar/gkr407 (2011).

237 Asp, J., Abramsson, A. & Betsholtz, C. in *In Situ Hybridization Protocols* Vol. 326 *Methods in Molecular Biology*™ (eds IanA Darby & TimD Hewitson) Ch. 6, 89-102 (Humana Press, 2006).

238 Abbott, B. in *Developmental Biology Protocols: Volume II* Vol. 136 *Methods in Molecular Biology*™ (eds RockyS Tuan & CeciliaW Lo) Ch. 21, 195-201 (Humana Press, 2000).

239 Lee, J. M. *et al.* Wnt11/Fgfr1b cross-talk modulates the fate of cells in palate development. *Dev Biol* **314**, 341-350, doi:10.1016/j.ydbio.2007.11.033 (2008).

240 Shin, J.-O. *et al.* MiR-200b is involved in Tgf-β signaling to regulate mammalian palate development. *Histochem Cell Biol* **137**, 67-78, doi:10.1007/s00418-011-0876-1 (2012).

241 Berggård, T., Linse, S. & James, P. Methods for the detection and analysis of protein–protein interactions. *PROTEOMICS* **7**, 2833-2842, doi:10.1002/pmic.200700131 (2007).

242 Zhang, L. *et al.* Characterization of a t(5;8)(q31;q21) translocation in a patient with mental retardation and congenital heart disease: implications for involvement of RUNX1T1 in human brain and heart development. *Eur J Hum Genet* **17**, 1010-1018, doi:http://www.nature.com/ejhg/journal/v17/n8/suppinfo/ejhg2008269s1.html (2009).

243 Ginn, S. L., Alexander, I. E., Edelstein, M. L., Abedi, M. R. & Wixon, J. Gene therapy clinical trials worldwide to 2012 – an update. *The Journal of Gene Medicine* **15**, 65-77, doi:10.1002/jgm.2698 (2013).

244 Wu, C. *et al.* Intra-amniotic transient transduction of the periderm with a viral vector encoding TGFbeta3 prevents cleft palate in Tgfbeta3(-/-) mouse embryos. *Mol Ther* **21**, 8-17, doi:10.1038/mt.2012.135 (2013).

245 Fassler J, C. P. BLAST Glossary. 2011 Jul 14. In: BLAST® Help [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2008.