FZD6, MATN2 AND SLC25A32, POSSIBLE CANDIDATE GENES IN NONSYNDROMIC CLEFT LIP AND PALATE

Nevena Cvjetkovic

Follow this and additional works at: http://digitalcommons.library.tmc.edu/utgsbs_dissertations

Part of the Genetics and Genomics Commons, and the Medicine and Health Sciences Commons

Recommended Citation
http://digitalcommons.library.tmc.edu/utgsbs_dissertations/281
FZD6, MATN2 AND SLC25A32, POSSIBLE CANDIDATE GENES IN NONSYNDROMIC CLEFT LIP AND PALATE

by

Nevena Cvjetkovic

APPROVED:

____________________________
Supervisory Professor
Jacqueline T. Hecht, PhD

____________________________
Eric C. Swindell, PhD

____________________________
John F. Teichgraeber, MD

____________________________
Ariadne M. Letra, DDS, MS, PhD

____________________________
Marianna Raia, MS, CGC

APPROVED:

____________________________
Dean, The University of Texas
Graduate School of Biomedical Sciences at Houston
FZD6, MATN2 AND SLC25A32, POSSIBLE CANDIDATE GENES IN NONSYNDROMIC CLEFT LIP AND PALATE

A

THESIS

Presented to the Faculty of
The University of Texas
Health Science Center at Houston
and
The University of Texas
M. D. Anderson Cancer Center
Graduate School of Biomedical Sciences
in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

by

Nevena Cvjetkovic, BS
Houston, Texas

May, 2012
ACKNOWLEDGEMENTS

First and foremost I would like to thank Dr. Hecht, my thesis advisor, for giving me the opportunity to work on a project that I have a personal connection with. Her guidance through the process allowed me to accomplish something I am extremely proud of. I would also like to thank my committee: Dr. Swindell, Dr. Teichgraeber, Dr. Letra, and Marianna Raia for all their support and advice. Additionally, I would like to thank my lab-mates: Katelyn, for helping me from day 1, showing me the ropes, and always answering my questions, no matter how trivial they were, Lori, for her devotion to this project and all the hard work she put into it, Robert for making the time in lab very entertaining, as well as the rest of Dr. Hecht’s lab for always being available and willing to help.

I would like to thank the UT Genetic Counseling program, Claire Singletary, Sarah Jane Noblin, my mentor Jennifer Czerwinski, and the numerous supervisors for their dedication to my training and education over the past two years. I would also like to thank Norma Leal and Patricia McDonald for their help behind the scenes. I would especially like to thank my wonderful classmates: Ann, Carolyn, Katie, Roya, Salma, and Sam. I could not imagine a better group of girls to experience graduate school with. When the going got tough having them right there beside me made everything more bearable. Additionally I’d like to thank my second year Emily who proved to be an awesome mentor and a true friend.

I would like to thank my boyfriend Tihomir for sticking by me through this trying time for any relationship. I would also like to thank all my family who always believed in me and encouraged me in whatever I decided to do. And most importantly, I’d like to thank my father Nenad, mother Ljubenka, and sister Jovana for their unconditional love and support. It is very simple fact of life that without them I would not be who or where I am.
Nonsyndromic cleft lip with or without cleft palate (NSCLP) is a common birth defect with a multifactorial etiology. Despite decades of research, the genetic underpinnings of NSCLP still remain largely unexplained. A genome wide association study (GWAS) of a large NSCLP African American family with seven affected individuals across three generations found evidence for linkage at 8q21.3-24.12 (LOD = 2.98). This region contained three biologically relevant candidate genes: *Frizzled-6 (FZD6) (LOD = 2.8)*, *Matrilin-2 (MATN2) (LOD = 2.3)*, and *Solute Carrier Family 25, Member 32 (SLC25A32) (LOD = 1.6)*. Sequencing of the coding regions and the 5’ and 3’ UTRs of these genes in two affected family members identified a rare intronic variant, rs138557689 (c.-153+432A>C), in *FZD6*. The rs138557689/C allele segregated with the NSCLP phenotype; *in silico* analysis predicted and EMSA analysis showed that the 138557689/C allele creates new DNA binding sites. *FZD6* is part of the WNT pathway, which is involved in craniofacial development, including midface development and upper lip fusion. Our novel findings suggest that an alteration in *FZD6* gene regulation may perturb this tightly controlled biological pathway and in turn contribute to the development of NSCLP in this family. Studies are underway to further define how the rs138557689/C variant affects expression of *FZD6*. 
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgements</td>
<td>iii</td>
</tr>
<tr>
<td>Abstract</td>
<td>iv</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>vii</td>
</tr>
<tr>
<td>Background</td>
<td>1</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>18</td>
</tr>
<tr>
<td>Results</td>
<td>21</td>
</tr>
<tr>
<td>Discussion</td>
<td>25</td>
</tr>
<tr>
<td>Appendix</td>
<td>30</td>
</tr>
<tr>
<td>References</td>
<td>33</td>
</tr>
<tr>
<td>Vita</td>
<td>51</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. Development of lip and palate .......................... 6
Figure 2. Pedigree of the African American NSCLP family ...... 10
Figure 3. Canonical WNT/β-catenin signaling pathway .......... 13
Figure 4. Segregation of rs138557689 in the African American NSCLP family .......................... 21
Figure 5. Conservation analysis of rs138557689 .................. 23
Figure 6. EMSA of rs138557689 ........................................ 24
Figure 7. Schematic representation of WNT Signaling Pathway .. 27
LIST OF TABLES

Table 1. Syndromes with CLP or CPO 3

Table 2. Frequency of rs138557689 C allele 22

Supplementary Table 1. Primers for FZD6 30

Supplementary Table 2. Primers for MATN2 31

Supplementary Table 3. Primers for SLC25A32 32
INTRODUCTION

Isolated or nonsyndromic cleft lip with or without cleft palate (NSCLP) is the fourth most common birth defect affecting more than 4,000 births each year in the United States [1,2,3]. Despite the improvements in treatment, NSCLP has considerable medical and financial implications for the affected individuals and their families, and the effects on speech, hearing, and appearance can lead to adverse psychosocial and psychiatric outcomes [4,5,6]. Individuals with NSCLP require multidisciplinary team care through adulthood and continue to demonstrate increased morbidity and mortality rates compared to the general population [7,8].

Cleft lip and palate has been recognized and documented throughout history. First evidence of clefting was noted in an Egyptian mummy dating from 2400 to 1300 BC [9]. A 2000-year-old statue of a Columbian king, a 2000-year old African mask, a Ming dynasty painting, and a 20th century Russian painting all depicted orofacial clefts [9]. What appears to be the first documented treatment of cleft lip and palate was performed in 390AD by a Chinese physician [10].

Many different explanations for causation of cleft lip and palate have evolved among the various cultures and populations in the world. Causes are often attributed to maternal impressions or supernatural events [11]. The term “hare-lip” to describe cleft lip originated from beliefs that clefting results from eating or looking at a rabbit [12,13]. A prevailing belief in many cultures is that a child will be born with a birth defect, including cleft lip and palate, if a woman pities or makes fun of an affected individual while pregnant [13,14,15]. Supernatural causes, such as the effect of the moon (lunar eclipse), are alleged causes of cleft lip and palate in Mexican folklore and among Hispanic populations [14,15,16], while populations in India and Nigeria ascribed the cleft to “God's will”, evil and ancestral spirits,
or sins committed in past lives [17,18]. Thesis work of Fogh-Andersen in 1942 was the first to widely recognize the hereditary component of CLP [19].

Despite the fact that NSCLP has been part of human life for many centuries, and the fact that many beliefs for the causation of this birth defect exist, the etiology of NSCLP is still largely unknown. Continual research in this area is needed to help elucidate the underlying etiologies of this common birth defect in order to improve genetic counseling for recurrence, diagnosis, prevention, and treatment, as well as to improve our understanding of the development of clefts.

Classification of Clefts

Orofacial clefts are divided into two groups based on developmental origin: anterior and posterior clefts [20]. Anterior cleft anomalies include cleft lip (CL) only, or cleft lip and primary palate (CLP) [20]. Anterior clefts can extend through the lip and alveolar part of the maxilla (primary or hard palate) to the incisive fossa [20]. Posterior cleft anomalies, which include clefts of the secondary or soft palate, present as cleft palate only (CPO) [9,20]. Posterior clefts extend through the soft and hard sections of the palate to the incisive fossa [20].

Clefts can further be divided into syndromic or nonsyndromic. Syndromic clefts are distinguished by the presence of other congenital malformations and may be associated with specific genes, chromosomal abnormalities, or maternal teratogenic exposures [21]. Approximately 30% of CLP and 50% of CPO are associated with other syndromes or anomalies, and there are currently over 400 different conditions listed on Online Mendelian Inheritance of Man in which clefts occur (http://www.ncbi.nlm.nih.gov/omim) [21]. Table 1
is an abbreviated list of syndromes which have CLP or CPO as a phenotypic finding and the
genes associated with them. The remaining 70% of cases of CLP and 50% of cases of CPO
are isolated, and therefore are referred to as nonsyndromic [21]. Nonsyndromic CLP
(NSCLP) is the focus of this study.

Table 1.

<table>
<thead>
<tr>
<th>Syndromes with CLP</th>
<th>Syndromes with CPO</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Syndrome</strong></td>
<td><strong>Gene</strong></td>
</tr>
<tr>
<td>Autosomal dominant developmental malformations, deafness, and dystonia</td>
<td>ACTB</td>
</tr>
<tr>
<td>Craniofrontonasal</td>
<td>EFN1B</td>
</tr>
<tr>
<td>Roberts</td>
<td>ESCO2</td>
</tr>
<tr>
<td>Holoprosencephaly</td>
<td>GLI2, SHH, SIX3, TGIF</td>
</tr>
<tr>
<td>Oro-facial-digital</td>
<td>GLI3</td>
</tr>
<tr>
<td>Van der Woude</td>
<td>IRF6</td>
</tr>
<tr>
<td>X-linked mental retardation and CL/P</td>
<td>PHF8</td>
</tr>
<tr>
<td>Gorlin</td>
<td>PTCH1</td>
</tr>
<tr>
<td>CLP – ectodermal dysplasia</td>
<td>PVRL1</td>
</tr>
<tr>
<td>Branchio-oculo-facial</td>
<td>TFAP2A</td>
</tr>
<tr>
<td>Ectrodactyly-ectodermal dysplasia-clefting-</td>
<td>TP63</td>
</tr>
<tr>
<td>Tetra-amelia with CLP</td>
<td>WNT3</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Syndrome</strong></td>
<td><strong>Gene</strong></td>
</tr>
<tr>
<td>Oculofaciocardiodental</td>
<td>BCOR</td>
</tr>
<tr>
<td>CHARGE</td>
<td>CHD7</td>
</tr>
<tr>
<td>Escobar multiple pterygium</td>
<td>CHRNA1</td>
</tr>
<tr>
<td>Stickler</td>
<td>COL2A1, COL11A1, COL11A2</td>
</tr>
<tr>
<td>Smith-Lenli-Opitz</td>
<td>DHCR7</td>
</tr>
<tr>
<td>Crouzon/Apert</td>
<td>FGFR2</td>
</tr>
<tr>
<td>Kabuki</td>
<td>MLL2</td>
</tr>
<tr>
<td>Cornelia de Lange</td>
<td>NIPBL</td>
</tr>
<tr>
<td>Diastrophic dysplasia</td>
<td>SLC26A2</td>
</tr>
<tr>
<td>Campomelic dysplasia</td>
<td>SOX9</td>
</tr>
<tr>
<td>Treacher Collins</td>
<td>TCOF1</td>
</tr>
<tr>
<td>Loeys-Dietz</td>
<td>TGFR1, TGFR2</td>
</tr>
</tbody>
</table>

Adapted from M.J. Dixon, M.L. Marazita, T.H. Beaty, J.C. Murray, 2011

A cleft can be unilateral, occurring on one side of the face, or bilateral, occurring on
both sides of the face. Unilateral clefts account for 90% of CLP (~60% of which occur on
the left side) while bilateral clefts account for the remaining 10% [21].

Recurrence of syndromic CLP depends on the underlying genetic condition, while
recurrence risk for NSCLP depends on several factors including number of affected
individuals in family, relation of affected family member, and laterality of the cleft [22].
The empirical recurrence risk for siblings is 2-3% in a case of unilateral CLP and 3-5% in the case of bilateral CLP [22].

**Birth Prevalence**

Cleft lip with or without palate is estimated to occur in approximately 1/700 to 1/1000 births but the prevalence of clefts varies by ethnicity [21,23]. The frequency of NSCLP appears to be highest among Native Americans (~3.6/1000 births) followed by Japanese (~2.1/1000 births) and Chinese (~1.7/1000 births) populations, and appears to be the lowest among African derived populations (~0.3/1000 births) [9]. In Texas, the prevalence of NSCLP is 0.67/1000 for Non-Hispanic Whites, 0.63/1000 for US-born Hispanics, 0.65/1000 for non-US born Hispanics, and 0.40/1000 for African Americans [1].

Birth prevalence of NSCLP also varies by gender with a 2:1 male to female ratio [24]. An association between NSCLP and socioeconomic status is implicated by studies which show groups from rural, lower socioeconomic areas having a higher birth prevalence of NSCLP compared to ethnically similar groups within higher socioeconomic status [21].

The birth prevalence of isolated CPO is approximately 1/2000 births and does not appear to differ much between various ethnicities [21]. In contrast to NSCLP, there is a female predominance of isolated CPO with a 3:2 female to male ratio [25].

**Normal Development of the lip and palate**

Understanding of normal development of the upper lip, primary, and secondary palates is important to help explain the reason for the altered development seen in CLP. Normal craniofacial development is a complex process directed by intricate pathways
important for cell induction, differentiation, proliferation, migration, patterning, and apoptosis [8,26]. Molecular pathways, including the Bmp, Fgf, Shh, and Wnt pathways, coordinate craniofacial development through both synergistic and antagonistic signaling [26]. Numerous genes, which code for growth factors, transcription factors, cell adhesion molecules, and signaling molecules, strictly regulate the development of the face [19,27,28,29,30]. Theoretically, perturbations within any of these genes and pathways could alter craniofacial morphogenesis and result in clefting.

The human face begins developing in the 4th week of embryogenesis [28]. Neural crest cells (NCCs), which are the major source of connective tissue components, migrate from the neural folds and combine with the core mesoderm and the epithelia to establish five facial primordia [20,26]. As depicted in Fig. 1a, the primodia consist of a single frontonasal prominence, two mandibular and two maxillary prominences, all of which surround the primitive oral cavity [8,28]. Ectodermal thickenings form nasal placodes and results in the formation of nasal pits by the end of week 4. The nasal pits further divide the frontonasal prominence into two horseshoe-shaped medial and lateral nasal processes (Fig. 1b) [8,26].

Rapid growth of the underlying mesenchyme of the facial primordia causes the medial nasal prominences to merge with each other and with the maxillary and lateral nasal prominences between the 7th to 10th week of embryogenesis [20]. The contact epithelia between the prominences are broken up by cell apoptosis to give rise to the intermaxillary segment from which the philtrum of the upper lip, the premaxillary part of the maxilla and the primary palate are formed (Fig. 1c) [20,26].

The secondary palate, which comprises majority of the soft palate, except for the anterior portion that holds the incisor teeth, has a different developmental origin [20].
Palatal shelves, which originate from the maxillary processes, appear in the 6th week of development and grow downward, vertical to the sides of the developing tongue (Fig. 1d) [8,9]. The palatal shelves then elevate to a horizontal position above the tongue (Fig. 1e), and fuse together (Fig. 1f) [8,20]. This is thought to occur earlier in females than in males [9]. Additionally, the two palatal shelves merge with the primary palate and the nasal septum dividing the oral from the nasal cavity [8]. This process is complete by the end of the 10th week [9,20].


**Development of Cleft Lip and Palate**

Cleft lip and palate can result when any of the multifaceted factors involved in normal facial morphogenesis are disrupted or altered. These may include failure of fusion of
the processes that form the face, distortions in epithelial movement, deficiency in epithelial and mesenchymal transformation (EMT), or failure in apoptosis [26].

Deficiency of mesenchyme in the maxillary prominence(s) and the median palatal process leads to anterior clefts [20]. Failure of one maxillary prominence to unite with the two merged medial nasal prominences results in unilateral cleft of the upper lip, while failure of both maxillary prominences to unite with the merged medial nasal prominences results in bilateral cleft of the upper lip [9,20].

Defective development of the secondary palate leads to posterior palatal clefts [20]. This is mainly due to the failure of mesenchymal masses in the lateral palatal processes to meet and fuse with each other and the nasal septum [9,20].

**Treatment of NSCLP**

Treatment of a patient with CLP requires a multidisciplinary approach which includes craniofacial and maxillofacial surgery, dentistry and orthodontic, audiology, otolaryngology, speech and language, nursing, pediatrics, genetics, and social services [31]. Many different surgical protocols exist to repair CLP and treat ensuing complications [21]. Currently, however, there is not an internationally accepted protocol for CLP repair [32].

Surgical repair of the CLP usually occurs within the first six months of life, the general rule being to perform surgical repair when the child is approximately 10 weeks of age, weighs at least 10 lb, and has achieved serum hemoglobin of 10 mg/ml [33]. There are various techniques of CLP repair which have been refined throughout time and some of these include LeMesurier (1955), Tennison (1952), Millard (1950), Skoog (1969), Randall
The choice of technique, however, most often depends on surgeon preference and severity of cleft [21].

**Etiology of NSCLP**

NSCLP does not follow a traditional Mendelian pattern of inheritance but is believed to result from a combination of several genetic variants which act in an additive fashion and interact with the environment to produce the phenotype [21]. NSCLP is therefore considered to be a multifactorial disorder that results from the interaction of both genetic and environmental factors [21]. Genetic evidence for NSCLP comes from multiple studies. These studies show that NSCLP aggregates in families, that there is a family history for clefting in 24-33% of NSCLP patients with a 10 to 32-fold increase in recurrence risk to first degree relatives, that the heritability of NSCLP in the NHW population is approximately 76%, and that the rate of concordance is higher in monozygotic (25-40%) than dizygotic (3-6%) twins [9,21,34,35,36,37,38,39]. Various environmental factors have also been associated with an increased risk for NSCLP including smoking, alcohol, diet, infections, fever, drugs, and teratogenic agents during early pregnancy [21,40,41,42,43].

Only a handful of genes associated with NSCLP have been identified to date and these comprise approximately 20% of the genetic causes of NSCLP [44]. Interferon regulator factor 6 (IRF6) is one of the genes that has shown an association with NSCLP in multiple populations [45,46,47,48,49,50]. Mutations in IRF6 are also known to cause van der Woude syndrome [51]. Various other genes, which include growth factors (TGFα, TGFβ), transcription factors (MSX1, TBX22), genes involved in xenobiotics (CYP1A1, GSTM1, NAT2), and other genes (RARA, MTHFR, CRISPLD2), have also been associated
with NSCLP [52,53,54,55,56,57,58,59,60,61]. The constant identification of novel genes for the pathogenesis of NSCLP across various populations highlights the genetic heterogeneity of this complex birth defect.

**Gene identification in NSCLP**

Multiple genetic approaches have been used to identify the genes and pathways contributing to NSCLP. These approaches include assessment of mouse models, linkage analysis using large multiplex families and affected relative pairs, association studies using family-based or case-control samples, identification of chromosomal anomalies or microdeletions in affected cases, and candidate gene studies [62].

Recently, genome-wide association studies (GWAS) have expanded the research in NSCLP and markedly helped in the identification of new candidate genes. GWAS studies use known single nucleotide polymorphisms (SNPs) that cover the genome to identify regions that are linked to and/or associated with NSCLP [62]. Identification of candidate genes by GWAS permits further analysis of individual genes through direct sequencing [62]. Sequencing of the candidate genes allows for identification of sequence variants within the gene which could potentially contribute to the development of NSCLP [62]. The sequence variants can be located within protein coding regions as well as noncoding regions, which include promoter/enhancer regions, or introns of the gene [63].

**8q21.3-24.12 Chromosomal Region**

A GWAS study performed in 2009 identified a major locus for NSCLP on chromosome 8q24 but the region with highest linkage did not contain any protein coding
genes [64]. Subsequent GWAS studies confirmed the association with the 8q24 locus and identified additional loci including 10q25, 7q22, 20q12, and 1p22 [62]. A smaller GWAS study reported by Chiquet et al. evaluated 10 multiplex NSCLP families including one large African-American family with 11 affecteds across 3 generations (Fig. 2). It was performed to confirm existing and identify new NSCLP chromosomal regions [65]. The study found evidence for linkage on 8q21.3-24.12 (LOD=2.98) [65]. Further evaluation of this region using linkage analysis generated a maximum multipoint LOD score of 2.8 for SNPs in *Frizzled 6 (FZD6)*, 2.3 in *Matrilin-2 (MATN2)*, and 1.6 in *Solute carrier family 25, member 32 (SLC26A32)*, prompting further analysis of these three genes and their connection to NSCLP.

**Figure 2. Pedigree of the African American NSCLP family**

*DNA included in the genome scan performed by Chiquet et al., 2009. Filled symbols denote affected individuals.*
**Frizzled-6 and WNT pathway**

Human *Frizzled-6* (*FZD6*) gene maps to chromosome 8q22.3-q23.1 and encodes a 706 amino-acid seven-pass transmembrane protein with a cystine-rich domain in the N-terminal extracellular region, two-N linked glycosylation sites, and two cystine residues in the second and third extracellular loops [66]. A total of 7 exons, 6 of which are coding, comprise the *FZD6* gene [66]. FZD6 amino acid sequence is well conserved between mouse and human genes with an amino acid identity of 83.3% [66].

*FZD6* is part of the *Frizzled* gene family which code for a group of receptors critical for initiation of wingless-type (WNT) signaling pathway [67]. The Frizzled receptors share conserved features which include a cystine-rich domain in the N-terminal extracellular region implicated in the binding of WNT ligands, N-linked glycosylation sites, two cystine residues in the second and third extracellular loops, and a Ser/Thr-x-Val motif in the C-terminus which acts as a binding site for the cytoplasmic protein containing the PDZ domain [66,68,69].

*FZD6* codes for three mRNA isoforms that are detected in both adult and fetal tissues [70]. *FZD6* was shown to be important for hair patterning in both *Drosophila* and mice as well as mice claw morphogenesis and appears to act within the WNT/PCP (planar cell polarity) pathway [71,72,73]. Mutations in *FZD6* have been associated with isolated autosomal recessive nail dysplasia and open neural tube defects in humans [71,74,75].

WNT signaling regulates differentiation and proliferation of a variety of cell types during development and is an important regulator of various functions within the cells [76]. Signaling by the WNT pathway directs cell proliferation, cell polarity, and cell fate determination during embryogenesis and homeostasis, regulates multiple developmental
processes, and plays a critical role in embryogenesis through both canonical (β-catenin dependent) and noncanonical (WNT/PCP and Ca2+/CAMKII) signaling [76, 77].

The canonical/β-catenin pathway is the best understood. As depicted in Fig. 3a, when a WNT ligand is absent, cytoplasmic β-catenin forms a complex with the scaffolding protein Axin, the tumor suppressor adenomatous polyposis coli gene product (APC), glycogen syntahse kinase 3B (GSK3B) and casein kinase 1 (CK1), and is phosphorylated by CK1 and subsequently by GSK3B [67]. Phosphorylated β-catenin is then recognized by the E3 ubiquitin ligase (UB), and is targeted for degradation within the proteosome [78]. This degradation, in turn, prevents β-catenin from reaching the nucleus, where the WNT target genes are repressed by the DNA-bound T cell factor/lymphoid enhancer factor (TCF-TLE)/Groucho and histone deacetylases (HDAC) [67].

Transcription of WNT target genes is activated in the presence of WNT ligand. As shown in Fig. 3b, the WNT ligand binds to a receptor complex formed between Frizzled and low-density lipoprotein receptor related protein 5 or 6 (LRP5/6). This complex then recruits a scaffolding protein Dishevelled (Dvl) and leads to LRP5/6 phosphorylation and Axin recruitment [67, 77]. This interrupts Axin-mediated phosphorylation and subsequent degradation of β-catenin, and allows β-catenin to build up in the nucleus where it can activate TCF/LEF and initiate transcription of WNT responsive genes [67].
In contrast to other Frizzled receptors, FZD6 protein lacks the Ser/Thr-X-Val motif at the C-terminal common to other frizzled receptors and has been shown to repress canonical WNT signaling through the noncanonical Ca2+/CaMKII pathway by inhibiting the TCF/LEF binding activity and down-regulating β-catenin targeted transcription of WNT genes [66,70]. Removal of FZD6’s N or C terminal sequences, however, abolishes the repressive activity of the protein [70].

Mutations in the WNT pathway have been linked to a variety of birth defects, cancers, and other diseases [67]. Importantly, WNT signaling has been shown to play a role in craniofacial development, including regional specification in the vertebrate face, neural crest induction and differentiation, mid-face development and upper lip fusion in mice, and
facial morphogenesis in mice [26,79,80,81,82,83,84,85]. Craniofacial abnormalities, including orofacial clefts, are found in WNT knockout mice and zebrafish [86,87]. An association with individual WNT genes and NSCLP has been found in humans (WNT3, WNT3A, WNT5A, WNT7A, WNT8A, WNT9B and WNT11) [87,88,89,90]. Based on this information, FZD6 is a strong candidate gene for NSCLP. Coding mutations or variants in the regulatory regions of FZD6 may alter the tightly regulated WNT pathway, and in turn contribute to NSCLP.

Matrilin-2 and the extracellular matrix assembly

Matrilin-2 (MATN2) gene maps to chromosome 8q22.1-q22.3 and encodes a 956 amino-acid protein whose structure consists of a putative signal peptide, two von Willerbrand factor A (vWFA)-like domains parted by ten epidermal growth factor (EGF)-like domains, a unique segment not seen in other matrilins, and an α-helical coiled-coil (CC) domain [91,92,93]. MATN2 has 19 exons and is transcribed from two alternative promoters, one of which is an upstream housekeeping-type promoter functional in fibroblasts and other cell types, and a downstream TATA-like promoter restricted to only embryo fibroblast and certain cell lines [94]. Matrilin-2 protein functions as a part of a tightly regulated pathway in the extracellular matrix (ECM) assembly and it is found in a wide variety of connective tissue cells, smooth muscle cells, and both the dermis and the epidermis in humans [93,95]. Human and mice Matrilin-2 proteins show 86.5% identity [93].

Matrilin-2 belongs to a family of four ECM proteins (matrilin-1, -2, -3, and -4), all of which share a structure made of vWFA and EGF domains, and a α-helical CC segment [92].
The matrilin genes also share a phase I, U-12-type AT-AC intron that is located in a strictly conserved position that separates the two exons for the CC domain [94]. While matrilin-1 and matrilin-3 are expressed mainly in cartilage and skeletal tissue, matrilin-2 and -4 have wider tissue distribution, including some non-skeletal tissues [91]. Matrilins -1, -2, and -3 can form filamentous networks that can connect to collagen fibers [93]. While Matrilin-2 is expressed in the mesoderm and epithelium in mice it is expressed in the dermal side of the basement membrane at the dermal-epidermal junction, as well as in keratinocytes and fibroblasts in humans [95].

Mutations in MATN3 gene were found to be associated with autosomal dominant forms of multiple epiphyseal dysplasia (MED) and mutations in MATN1 have been associated with osteoarthritis and relapsing polychondritis. MATN2 and MATN4, to date, have not been associated with any disease [92,96]. Knock-out mice lacking Matn2 have been shown to develop without any obvious abnormalities [97]. Few disorders, including Klippel-Fiel syndrome with laryngeal malformations and Cohen syndrome (which consists of facial, oral, ocular, and limb deformities, and well as intellectual disability), have been linked to the 8q22 position [93]. MATN2, however, has not been identified as a candidate gene in these syndromes [93].

Despite the lack of obvious craniofacial phenotype in Matn2 knock-out mice, interactions of ECM components, growth factors, and embryonic tissues are known to be involved in the normal development of cranial priomordia and orofacial structures [98,99] and alterations in the ECM components could potentially lead to NSCLP.
**Solute carrier family 25, member 32 and Folate metabolism**

*Solute carrier family 25, member 32* (*SLC25A32*) gene, also called *Mitochondrial Folate Transporter (MTF)*, is located on chromosome 8q21.2. *SLC25A32* consists of 7 exons and encodes a 315 amino-acid protein transporter which shuttles folates from the cytoplasm into the mitochondria [100]. The protein consists of six transmembrane domains: three loops facing the mitochondrial matrix and three repeats of an energy transfer sequence [100]. The protein does not contain an ATP-binding motif and probably functions as an ATP-independent transporter [100].

Solute carriers are comprised of 43 identified families which control the movement of various substances across the cell membrane [101]. Members of the solute carrier family 25 (*SLC25*) are known to transfer a variety of substances across the mitochondrial membrane, and are therefore referred to as mitochondrial solute carriers [102]. The *SLC25A32* protein is shown to transport folate across the mitochondrial membrane [100].

Folate metabolism is a complex process consisting of multiple genes and pathways. Folate is absorbed from the environment, transported across the cell membrane, and compartmentalized between the mitochondria and the cell cytoplasm before it is transported into the mitochondria [103]. The transport of folate into the mitochondria involves various transporters and carriers, including the *SLC25A32* transporter [100,103].

The importance of folate in prevention of birth defects has been investigated since folate fortification reduced the occurrence of spina bifida, an open neural tube defect (ONTD), by approximately 70% [104]. Because ONTDs and NSCLP both develop from migrating neural crest cells, studies have been performed to assess the role of folic acid in NSCLP in hopes to validate the hypothesis that folic acid deficiency may also contribute to
NSCLP [105,106]. Observational studies of folic acid supplementation and the recurrence of NSCLP have not reliably shown the same decrease as with ONTD. A few meta-analysis studies, however, showed decreases of 18-23% in the risk of NSCLP with supplementation of folic acid and/or multivitamins during pregnancy [107,108]. Studies also suggest that variations in the folate pathway genes may contribute to NSCLP [105]. Disturbances in the folate transport, therefore, may also contribute to NSCLP.

**Significance of this study**

NSCLP is a common birth defect caused by both genetic and environmental factors. Despite decades of research, genetic contributions to NSCLP have yet to be explained, and there remains the need to identify candidate genes. The goal of this project is to further elucidate the genetic etiology of NSCLP by sequencing the coding and 5’ and 3’ UTR regions of three biologically relevant candidate genes, *FZD6, MATN2,* and *SLC25A32,* previously identified in a genome wide scan in a large multiplex African-American family. This family is unique because NSCLP has a low prevalence in the African American population yet there are numerous affected individuals in multiple generations. Sequence variants in any of these three genes may disrupt the highly regulated craniofacial development and give rise to NSCLP. Identification of new genes for NSCLP would not only benefit in diagnosis, prevention, and counseling for this birth defect, but may help in understanding the developmental pathways involved in craniofacial morphogenesis.
MATERIALS AND METHODS

IRB Approval

This study was approved by the Committee for the Protection of Human Subjects at the University of Texas Health Science Center at Houston (HSC-MS-03-090 and HSC-MS-11-0336).

Family

The focus of this study is an African American NSCLP family with 11 affected individuals, 7 of whom were available for evaluation (Fig. 2). All family members were evaluated by one author (JTH). No other anomalies or lip pits were present in any family members. DNA samples from these 7 individuals and 13 connecting or related unaffected family members were subjected to a previously described 6K Illumina IVb genome scan and linkage analysis [65]. A maximum multipoint LOD score of 3.0 can be obtained for this family.

Sequencing

The genomic structure of FZD6, MATN2, and SLC25A32 was determined using the NCBI GenBank (www.ncbi.nlm.nih.gov). Forward and reverse primers were designed to capture the sequence of each exon and approximately 50-100 bps upstream and downstream of the intron/exon junction, as well as the complete 5’ and 3’ untranslated regions (UTRs) for all three isoforms of FZD6, two isoforms of MATN2, and SLC25A32 (Supplemental Tables 1-3). DNA samples from two affected family members (Fig 2: III-15 and IV-2) were sequenced for each gene. These individuals were chosen from the opposite sides of the family since they are less likely to share common familial variants. After initial data analysis, 24 additional family members (5 affected and 19 unaffected) were sequenced for
rs138557689 using \textit{FZD6} primer set E1C (Supplementary Table 1). Standard PCR amplification conditions were used and the annealing temperatures for each primer set are shown in the Supplemental Tables 1-3. Amplified PCR product was purified according to manufacturer’s protocol (Qiagen, Valencia, CA). Sequencing results were compared to consensus sequences obtained from NCBI public database and analyzed using Sequencer v4.9 (Gene Codes, Ann Arbor, MI).

\textbf{Variant Analysis}

Sequence variants were identified using dbSNP (\url{www.ncbi.nlm.nih.gov/projects/SNP}). Only sequence changes shared by both affected individuals (Fig 2: III-15 and IV-2) were considered. SNPs identified in the potential regulatory regions, 5’ UTR and the first two introns of the gene, were assessed for their effect on DNA binding using three online \textit{in silico} analyses: Alibaba2, Patch, and Transcription Element Search Software (TESS) [109,110,111]. SNPs identified in the 3’UTR region were assessed for their effect on microRNA binding sites using microRNAMap and miRBAse databases [112,113]. SNPs identified in the coding region were analyzed using PolyPhen and SIFT [114,115].

\textbf{Species Conservation Analysis}

SNPs were assessed for evolutionary conservation using the UCSC Genome Browser Multiz Alignments of 46 Vertebrates track (\url{www.genome.ucsc.edu/cgi-bin/hgGateway}) and the ECR Browser tool (\url{http://ecrbrowser.dcode.org}). SNPs were further assessed using the genomic evolutionary rate profiling (GERP) track within the UCSC browser to estimate the evolutionary constraint rates for individual nucleotide positions [116]. GERP scores are positive in constrained regions and negative in neutral DNA [116].
Electrophoretic mobility shift assay (EMSA)

EMSA was used to evaluate whether an ancestral or alternate allele in the SNP of interest altered DNA binding. 20-mer oligonucleotide sequences were hybridized incorporating either the ancestral or alternate allele (Integrated DNA Technologies, Coralville, IA) and labeled with P\(^{32}\) dCTP (40,000cpm). The binding reaction was carried out in 20 uL binding buffer mix (1M Tris (pH 7.5), 1M KCl, 80% glycerol, 10% NP-40, 0.5M EDTA, 100 mM PMSF, 1M DTT and H2O), containing 1 uL poly dGdC and 1 uL Cos7 cell nuclear extract, and was incubated on ice for 20 mins. The sample was incubated with the radiolabeled probes for 20 mins at room temperature. The complexes were resolved on a 5% polyacrylamide gel at 150V for 3hrs in TBE buffer. Gels were dried and exposed at -80° C for 48 hrs. Negative controls were run with cold oligonucleotides using the labeled probes and binding buffer without the nuclear extract.

Genotyping of NSCLP Probands and Unaffected Controls

A custom TaqMan Genotyping Assay (Applied Biosystems, Foster City, CA) was designed to genotype rs138557689 using our standard protocol and was detected on ABI Viia7 RUO Machine (Applied Biosystems, Foster City, CA). A total of 579 controls and 836 NSCLP probands were genotyped. Allele calls were determined using the Viia7 Software (ABI) and the allele call rate was greater than 98%.
RESULTS

Sequencing of the coding and 5’ and 3’ UTR regions for FZD6, MATN2, and SLC25A32 identified three sequence variants in III-15 and IV-2 (Fig. 4). One sequence change, a homozygous deletion of nucleotide T (c.*164delT), was identified in the 3’UTR of SLC25A32. This sequence variation is listed in dbSNP as rs11345830, but the frequency in the general population was not available. However, this homozygous deletion is highly conserved among primates, is not predicted to affect microRNA binding sites, and was therefore excluded from further analysis. One variant, rs113199627 was identified in the 5’UTR of MATN2. This is a common variant in the population (MAF=14.5%) and was also excluded from further analysis.

A single base pair c.-153+432A>C change in FZD6 was present in both individuals (Fig. 4). Rs138557689, a SNP in intron 1, is located 681bps upstream of the start site.

Figure 4. Segregation of rs138557689 in the African American NSCLP family. A large African American family with 11 affected individuals across 3 generations is shown. Filled symbols denote affected individuals. Asterisks (*) denote individuals included in the genome scan reported in Chiquet et al., 2009. The C allele of rs138557689 segregates with the affected individuals and is transmitted by 4 unaffected individuals originating from individual I-3. The C allele is present in 5 additional unaffected individuals.
Based on the availability of DNA samples, 24 additional relatives were sequenced for the same variant. The C allele was found to segregate with the cleft phenotype where 100% of the affected individuals had the CA genotype (Fig 4: II-9, III-2, III-15, III-17, IV-2, IV-6, IV-8). The C allele was transmitted by 4 unaffected individuals (Fig. 4: I-3, II-13, III-4, III-11) and there were 5 other unaffected family members with the CA genotype (Fig. 4: II-12, III-13, IV-5, IV-10, IV-12). The CA genotype was present in 60% of the unaffected family members. There were no affected individuals, however, with the AA genotype.

Despite being submitted by the 1000 Genomes project and listed in NCBI, frequency data on rs138557689 variant was not available. To assess the frequency of this allele, we genotyped 579 controls of Non-Hispanic White (NHW), Hispanic, and African American ethnicity, and found the C allele in 3 individuals (1 NHW and 2 African Americans) who were heterozygous. No homozygous individuals for the C allele were found. The frequency of the C allele was 0.8% in the African American, 0.3% in NHW, and 0% in Hispanic controls (Table 2A). We then genotyped 836 probands from our NSCLP dataset and found the C allele in two African American individuals, who were both heterozygous. This includes one African American proband from the family described here. The C allele was not found in the NHW and Hispanic

<table>
<thead>
<tr>
<th>Table 2. Frequency of the rs138557689/C allele</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Controls</strong></td>
</tr>
<tr>
<td><strong>Alleles</strong></td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td><strong>C allele frequency</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>B. NSCLP</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alleles</strong></td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td><strong>C allele frequency</strong></td>
</tr>
</tbody>
</table>

NHW= Non-Hispanic White, African A= African American, Hisp= Hispanic
NSCLP probands. The frequency of the C allele in the NSCLP dataset was 1.2% for the African Americans and 0% for the NHW and Hispanics (Table 2B).

To assess whether this SNP was in a conserved region, we compared the sequence in different species. The ancestral A allele of rs138557680 showed conservation in chimps, rhesus monkeys, gorillas, marmoset, mouse lemur, and opossum while the alternate C allele was only conserved in elephants and armadillos (Fig. 5). These results suggest that C is a rare variant. The position of the rs138557689 variant showed moderate evolutionary constraint (GERP score=-2.98). This suggests that the A allele is moderately conserved at this location and that nucleotide changes at this position are tolerated and not selected against through evolution.

Figure 5. Conservation analysis of rs138557689. The UCSC Genome Browser Multiz Alignments of 46 Vertebrates track depicting the nucleotide alignment of the human FZD6 gene and homologues from other species. The position of the rs138557689 variant is indicated by the arrow. The rs138557689 A allele is conserved. Mouse, rat, X. tropicalis, and zebrafish lack the region containing this variant.
*In silico* analyses predicted that the alternate C allele alters DNA binding. The ancestral allele was predicted to create a C/EB alpha transcription factor binding site by Patch. The alternate allele was predicted to create a Sp1 site by AliBaba2, and NF-1/L binding site by Patch and NIP and NF-1/L binding sites by TESS. EMSA analysis was used to determine whether a DNA binding site was present with either allele. As shown in Fig. 6, there were two bands present for the probe containing the C allele but not for the A allele. These results indicate that the C allele creates new DNA binding sites.

![EMSAs image](image_url)

**Figure 6. EMSA of rs138557689.** EMSA was performed using nuclear extract from Cos7 cells. Samples were incubated with ancestral A or alternate C allele alpha P32 labeled DNA oligonucleotide probes. Poly-dCdG was used as a nonspecific competitor. Negative controls were run with cold oligonucleotides using the labeled probes without the nuclear extract. Bands were observed with the C allele only indicating that the C allele creates a DNA binding site. No binding was observed in the presence of negative control A and C probes.
DISCUSSION

We identified a potentially functional variant, rs138557689, in intron 1 of FZD6 gene segregating with NSCLP phenotype in a large African American family. *In silico* analysis suggested and EMSA results showed that the alternate C allele creates transcription binding sites. Additionally, the ancestral A allele is highly conserved in primates and moderately conserved in other vertebrates, suggesting that it may have important functional significance in the regulation of FZD6. While the CA genotype for the rs138557689 variant appears to contribute to the development of NSCLP in this family, analyses are underway to assess how the variant affects gene expression.

The rs138557689/C allele is rare in the general population and was found in less than 1% of NHW and African American controls. The C allele was also rare in the NSCLP probands being identified in only two of 1672 African American chromosomes. Even though there appears to be an enrichment for the C allele in the African American population in our data set, there was no significant difference in the C allele frequency between the African American controls and NSLCP probands (p-value=1.0, Fisher exact test). Based on this information the variant does not appear to be relevant at the population level. However, the large number of individuals with the C allele in the African American family suggests that the rare rs138557689/C genotype may act as a predisposing risk factor for NSCLP when superimposed on the genetic background present in this family. The presence of the C allele in the unaffected family members as well as controls suggests that a person must have a specific combination of variants in one or multiple genes to exhibit the phenotype. This is consistent with the multifactorial model and is supported by the observed reduced penetrance in this African American family. Our results also suggest that different
ethnic-specific genetic risk profiles may contribute to NSCLP, and therefore the underlying disease mechanisms may also be different between ethnic groups. This is consistent with the observed differences in prevalence of NSCLP among different populations and the observed genetic heterogeneity that characterizes this complex disorder. Ethnic-specific differences warrant further studies.

Variants in noncoding regions that modify transcription and expression have increasingly been implicated in complex diseases because they perturb transcription by creating or removing binding sites for specific transcription factors [21,117,118]. We found that the presence of the rs138557689/C allele creates DNA binding sites that could potentially repress or enhance the expression of FZD6. Because the development of facial structures is a highly regulated process that relies on crosstalk and convergence of different developmental pathways, alterations in the expression of FZD6 could affect the cascade of molecular events that the gene is involved in.

While the exact function of FZD6 in human craniofacial development is not well defined, FZD6 is diffusely expressed in the craniofacial mesenchyme of zebrafish and chick suggesting that it plays a role during craniofacial development [119,120]. FZD6 is part of the Frizzled family of genes which encode a group of G-coupled receptors critical for initiation of WNT signaling [66,67]. WNT signaling is a highly controlled cellular pathway that regulates multiple functions during development through both canonical/β-catenin and noncanonical signal transduction [76,77]. Importantly, WNT signaling has been shown to play a role in craniofacial development, including regional specification in the vertebrate face, neural crest induction and differentiation, and mid-face development and upper lip fusion in mice [26,80,81,82,83,84,85,121]. Craniofacial abnormalities, including orofacial
clefts, are found in WNT knock-out mice and zebrafish, while mutations in WNT3 cause tetra-amelia with CLP in humans [86,87,122]. Additionally, associations between WNTs (WNT3, WNT3A, WNT5A, WNT7A, WNT8A, WNT9B and WNT11) and NSCLP have been found, with strongest associations seen in WNT3 and WNT3A [87,88,89,90,123]. Unlike the other Frizzleds, FZD6 represses canonical WNT signaling through the noncanonical Ca2+/CaMKII pathway by down-regulating TCF/LEF binding activity and subsequent transcription of WNT target genes (Fig. 7) [70].

**Figure 7. Schematic representation of WNT signaling pathway**

![Diagram of WNT signaling pathway](image)

WNT= WNT ligand; FZD=Frizzled receptor; LRP=Low-density lipoprotein receptor related protein; Dvl=Dishkevelled; APC=adenomatous polyposis coli; GSK3B=glycogen synthase kinase 3B; CK1=casein kinase 1; TCF/LEF=T cell factor/lymphoid enhancer factor; PLC=Phospholipase C; Ca2+=calcium; CAMKII=calcium/calmodulin-dependent protein kinase; NLK=Nemo-like Kinase; Rac/Rho=small GTP-binding proteins; JNK=c-Jun N-terminal kinase
An increase in the expression of FZD6 may further repress canonical WNT signaling crucial for normal craniofacial development while a decrease in FZD6 expression may cause a loss of negative regulation of the WNT pathway. Negative regulation of the WNT pathway is equally important in craniofacial development since mice that are deficient in the Dkk1 gene (another negative regulator of the WNT pathway) lack craniofacial structures [124].

FZD6 has also been shown to mediate the non-canonical planar cell polarity (WNT/PCP) pathway in mice (Fig. 7) [72,73,125]. The WNT/PCP pathway controls the polarity and orientation of the migrating neural crest cells [81,126]. During normal craniofacial development NCCs migrate from the neural folds to fill the facial prominences with mesenchyme and contribute to the formation of the nose and upper lip [20,26,28,127,128]. Defects in NCC formation, induction, differentiation, or migration can result in craniofacial abnormalities [129,130,131,132]. Therefore, a decrease in FZD6 expression may perturb the WNT/PCP pathway and alter neural cell migration or mesenchymal planar cell polarity in craniofacial structures and may potentially lead to NSCLP.

In summary, we report that a rare genetic variant, rs138557689, which alters DNA binding in FZD6, may act as a predisposing risk factor for NSCLP in a large African American family with 11 cases of NSCLP segregating in 3 generations. The alteration in FZD6 gene regulation may perturb highly controlled biological pathways, in particular the WNT pathway, therefore focusing on the genes within the WNT pathway and their interaction with FZD6 may provide further insight into the genetic etiology of NSCLP. Assessing the affected and unaffected family members who have the CA genotype for the presence of other variants in NSCLP susceptibility genes may delineate the underlying
genetic risk profile contributing to development of NSCLP in this family. This outcome would not only improve genetic counseling for this family but may improve the understanding of the biological mechanisms linking \textit{FZD6} to NSCLP.
APPENDIX

Supplementary Table 1. *FZD6* Primers

<table>
<thead>
<tr>
<th>Name</th>
<th>Forward</th>
<th>Reverse</th>
<th>Length (bp)</th>
<th>Tm(°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1A</td>
<td>CACTGCTACCTGAGCATCCA</td>
<td>GGA AGC CAC CTC CAC CTT</td>
<td>318</td>
<td>65.1</td>
</tr>
<tr>
<td>E1B</td>
<td>AGA GCC AGC GCC AAG AGC TTC A</td>
<td>AGG GTG TAG TGC TGC CBT CGA AA</td>
<td>874</td>
<td>65.1</td>
</tr>
<tr>
<td>E1C</td>
<td>GGG AAC CGG CTC TGA AAG GCG</td>
<td>CCT TAA TAA GCT TTC CAA CAG GGG CCC</td>
<td>466</td>
<td>69.4</td>
</tr>
<tr>
<td>E1D</td>
<td>CTA GTT GCC CTT AGG AAA ATC GAG</td>
<td>GAC CCA GGA CTC ATT TTC AGG</td>
<td>514</td>
<td>60.2</td>
</tr>
<tr>
<td>E2</td>
<td>CAAATGTTGCTGATACACCCCTC</td>
<td>GCAAATCTGAAGAAATCGGCTC</td>
<td>473</td>
<td>54.3</td>
</tr>
<tr>
<td>E3</td>
<td>GATTTCAATGCTGATAGGCCAG</td>
<td>CTGTAAGTTCCCTGTAGACAGAAG</td>
<td>544</td>
<td>54.3</td>
</tr>
<tr>
<td>E4A</td>
<td>CCCATTTAACAGCCGACGGTTTT</td>
<td>CGCCTAGAAAAATCCCAATGAAG</td>
<td>593</td>
<td>54.3</td>
</tr>
<tr>
<td>E4B</td>
<td>GCAACTCTGTACATCATTACCTAC</td>
<td>GTCGAAATGCGAGAGAATCCGATCC</td>
<td>481</td>
<td>54.3</td>
</tr>
<tr>
<td>E4C</td>
<td>GCAGTGTGTTTCTCATGTGTTG</td>
<td>GCTCTTTGATTTTCTGACTCTG</td>
<td>538</td>
<td>54.3</td>
</tr>
<tr>
<td>E5</td>
<td>GATAAAAAATGTGGTTGCACGTAGACG</td>
<td>GATAAAAAATGTGGTTGCACGTAGACG</td>
<td>486</td>
<td>54.3</td>
</tr>
<tr>
<td>E6A</td>
<td>CTGCAAAAAGCAACTTAGAGTG</td>
<td>CCTCTCTCATGTGTTGCTGCTG</td>
<td>472</td>
<td>54.3</td>
</tr>
<tr>
<td>E6C</td>
<td>GTACATCTCTGATCCATGGGG</td>
<td>CAAGACATTCTCTGCTGCTGCTG</td>
<td>427</td>
<td>54.3</td>
</tr>
<tr>
<td>E7A</td>
<td>GATAAAAAGGCTGACCTAGAGT</td>
<td>GAAGATACGCTCTTTCTGTG</td>
<td>530</td>
<td>54.3</td>
</tr>
<tr>
<td>E7B</td>
<td>GATATGGCCTACTGTTATACTG</td>
<td>CCACCTCTTTTCTGGTGTAGCACC</td>
<td>592</td>
<td>54.3</td>
</tr>
<tr>
<td>E7C</td>
<td>CCCACTTATGTTATACCTTACCAC</td>
<td>CAAAGATACGAAAATCTTGGCCC</td>
<td>576</td>
<td>54.3</td>
</tr>
<tr>
<td>E7D</td>
<td>GCCAATCAATGGAAAAAAGGCTAG</td>
<td>GCATTCTCATTGAAAGGTTAAAAAG</td>
<td>450</td>
<td>54.3</td>
</tr>
</tbody>
</table>

*E=Exon

*Betaine was added to primer E1B during the PCR reaction.

*DMSO and magnesium chloride were added to primer E1C during the PCR reaction.
Supplementary Table 2. *MATN2* Primers

<table>
<thead>
<tr>
<th>Name</th>
<th>Forward</th>
<th>Reverse</th>
<th>Length (bp)</th>
<th>Tm(°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1A</td>
<td>GCT ACT CTG AGG CAG GAG AAT GG</td>
<td>GGC TAA TTG ATG ACA AGC GCC AAG</td>
<td>572</td>
<td>63.5</td>
</tr>
<tr>
<td>E1B</td>
<td>GAC TTT CCC TGC TCC TTC GGG GTA GG</td>
<td>C CGG CCG GAA GGA GGG GGG GTG TCC</td>
<td>538</td>
<td>65.1</td>
</tr>
<tr>
<td>E1C</td>
<td>GGGAGCGCTTGGAGATGGGAC</td>
<td>CTTCTG TCTG ACA ATG GAA GCC CAT TGT C</td>
<td>374</td>
<td>60.2</td>
</tr>
<tr>
<td>E1D</td>
<td>CAT CCCGC GTC AGT GGG TGC</td>
<td>GAG GAC GGA GCC CCA CAT TCT TGT A</td>
<td>499</td>
<td>63.1</td>
</tr>
<tr>
<td>E2</td>
<td>CGGCATTTTGGAGTCAAGTG</td>
<td>CCAGCACATCCAACAAAGGAC</td>
<td>417</td>
<td>63.5</td>
</tr>
<tr>
<td>E3A</td>
<td>GTAGAGACAGGGCAAACTTTACAGTG</td>
<td>CTTCTG TCTG ACA ATG GAA GCC CAT TGT C</td>
<td>465</td>
<td>63.5</td>
</tr>
<tr>
<td>E3B</td>
<td>CTTGCCTGTCAGAGGATGCG</td>
<td>CCTATTACAGCAGATCATCTACCTG</td>
<td>459</td>
<td>63.5</td>
</tr>
<tr>
<td>E4</td>
<td>CTGCCCAAGAGGAGAGAGTAG</td>
<td>GGGATGACA ACTGGGTTGGGCG</td>
<td>342</td>
<td>60.2</td>
</tr>
<tr>
<td>E5</td>
<td>GTACTTTTGGTGAGGCTCTG</td>
<td>CTGAGCACATTGGTGAGGAAACCC</td>
<td>352</td>
<td>63.5</td>
</tr>
<tr>
<td>E6</td>
<td>GCCCTCATCTACCATTCCC</td>
<td>GTTGAGATACAGTGCTATGGT C</td>
<td>383</td>
<td>63.5</td>
</tr>
<tr>
<td>E7</td>
<td>CCCCTCATGTCCTGTGACTC</td>
<td>GCATTAAAGTTTGCGAGGGAAC</td>
<td>338</td>
<td>60.2</td>
</tr>
<tr>
<td>E8</td>
<td>GAAATGACCGCTCCCGGCTTG</td>
<td>GACCATCGAAACCCCGGCTG</td>
<td>375</td>
<td>63.5</td>
</tr>
<tr>
<td>E9</td>
<td>GACACCTTCTCGTGTGGCTTG</td>
<td>GACACCTTCTCGTGTGGCTTG</td>
<td>305</td>
<td>63.5</td>
</tr>
<tr>
<td>E10</td>
<td>GCATGCTCTGAGGGAGGCG</td>
<td>GCCGCCAGCAAAAGGCTGGAGAC</td>
<td>315</td>
<td>63.5</td>
</tr>
<tr>
<td>E11</td>
<td>GCCGCCAGAAGGACCTGGAGAC</td>
<td>CCATTTTCTTCTTCAGGGTACC</td>
<td>400</td>
<td>60.2</td>
</tr>
<tr>
<td>E12</td>
<td>CCATGGACCTGAGCTGACAGG</td>
<td>GAGCAACCAAGGCAAGGAAACTG</td>
<td>379</td>
<td>63.5</td>
</tr>
<tr>
<td>E13</td>
<td>CCCCTCATGACGATGACG</td>
<td>GCATTAAAGGCTTCCTTCTCC</td>
<td>440</td>
<td>63.5</td>
</tr>
<tr>
<td>E14</td>
<td>GCTTCCCCAAAAAGTGGTTATGCC</td>
<td>CTGATACACTCGAAGAAGGCTTACCC</td>
<td>625</td>
<td>63.5</td>
</tr>
<tr>
<td>E15</td>
<td>GCTGGACATGAGCTCTTCAATC</td>
<td>GCTGAAAGCATGAGAAACTCTTTG</td>
<td>446</td>
<td>63.5</td>
</tr>
<tr>
<td>E16</td>
<td>GCTTTTGGATACTCTTGGTGTTACC</td>
<td>GAAAAATGATACTATCTTGGGGTTGG</td>
<td>372</td>
<td>63.5</td>
</tr>
<tr>
<td>E17</td>
<td>CCAAGTGGCTAGGAAATCCAAG</td>
<td>CTATTTCCTACCTCCTTCTCTCTTCC</td>
<td>338</td>
<td>56.2</td>
</tr>
<tr>
<td>E18</td>
<td>GTAATCTGTGGATCTGGG</td>
<td>CTGGCCCTGCCTCCTCCCTCC</td>
<td>425</td>
<td>63.5</td>
</tr>
<tr>
<td>E19A</td>
<td>GAG TTC TAC AAA TTT ACA AGT CAG G</td>
<td>CA AAT GAG ATT GCA CAC TAA GGC</td>
<td>542</td>
<td>56.2</td>
</tr>
<tr>
<td>E19B</td>
<td>GTGAGAATGAAATGAAGCTATGGAAG</td>
<td>CCACATCCTACCAACGACCTTG</td>
<td>562</td>
<td>62</td>
</tr>
<tr>
<td>E19C</td>
<td>GAG AAA TGG CAR ACA TGC CTA TGA AAA</td>
<td>CTA CAG GCA CCC ACT ACC ACA CCC G</td>
<td>595</td>
<td>68.2</td>
</tr>
</tbody>
</table>

E=Exon

DMSO and magnesium chloride were added to primer E2 during the PCR reaction.
### Supplementary Table 3. SLC25A32 Primers

<table>
<thead>
<tr>
<th>Name</th>
<th>Forward</th>
<th>Reverse</th>
<th>Length (bp)</th>
<th>Tm(°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1A</td>
<td>CAT CTC GGT TGC TCT TCC GGC</td>
<td>CGC CCC TTG TGA GCG CAA C</td>
<td>431</td>
<td>67</td>
</tr>
<tr>
<td>E1B</td>
<td>CCC CTC CAT CGC GCT TTC CG</td>
<td>CAG GTT AGC CAA CGC GGA CAG</td>
<td>513</td>
<td>67</td>
</tr>
<tr>
<td>E2</td>
<td>CGG GAC TCT GAC ACA AAA ATG</td>
<td>GCC ATT TTG TTC TTC CTG ATC C</td>
<td>519</td>
<td>62</td>
</tr>
<tr>
<td>E3</td>
<td>GCA AAC CAC TTC CAG CAA ATT C</td>
<td>CTG AGG CAG GAT TAT CGC TTG</td>
<td>530</td>
<td>58.3</td>
</tr>
<tr>
<td>E4</td>
<td>CCT TTT GAC CCT AAG ACT GTG C</td>
<td>GAT ACA TCC TAT GTT AAT GGG GAC</td>
<td>537</td>
<td>58.3</td>
</tr>
<tr>
<td>E5</td>
<td>GTG CTG TAA TGA GAG AGA AGA G</td>
<td>GAC ATT TGT GTG GCT TCA</td>
<td>374</td>
<td>58.3</td>
</tr>
<tr>
<td>E6</td>
<td>GAT GAG TAC TGG CTC TGC CA</td>
<td>CGT AAT AAC TGG GAA AGC AGG</td>
<td>473</td>
<td>58.3</td>
</tr>
<tr>
<td>E7A</td>
<td>GAC TGC TGC TTG CTC CAT GCC</td>
<td>GGC AGC CAT TTC AGG CAG AGG</td>
<td>454</td>
<td>58.3</td>
</tr>
<tr>
<td>E7B</td>
<td>GAA GCC AGA GAA CTG CTA AGT C</td>
<td>CGA CAA AGC AAG ACT CCA TCT C</td>
<td>404</td>
<td>62</td>
</tr>
<tr>
<td>E7C</td>
<td>GTG AGC TTA CTT GCC TGG ATT GC</td>
<td>CGC TAG GTA GTG CAT CCC AAC TG</td>
<td>543</td>
<td>62</td>
</tr>
<tr>
<td>E7D</td>
<td>GGATGGTCTCAATCTCCTGAC</td>
<td>CAA CCT GAA TTT GAG AAA CCA ATG AAG</td>
<td>564</td>
<td>58.3</td>
</tr>
<tr>
<td>E7E</td>
<td>GTTGACAAGGAATTGAAAATGAG</td>
<td>GTT CAC TTT TTC CGT GTG GGG</td>
<td>517</td>
<td>58.3</td>
</tr>
<tr>
<td>E7F</td>
<td>GCT GTT GCA CTA CCA TCT ATT TG</td>
<td>CTC AGT GCT TGG TGA CGT AC</td>
<td>556</td>
<td>52</td>
</tr>
</tbody>
</table>

E=Exon
REFERENCES


Ellington, C. Du, J. Dixon, M.J. Dixon, P.A. Trainor, Prevention of the
neurocristopathy Treacher Collins syndrome through inhibition of p53 function, Nat

[132] F. Vitelli, M. Morishima, I. Taddei, E.A. Lindsay, A. Baldini, Tbx1 mutation causes
multiple cardiovascular defects and disrupts neural crest and cranial nerve migratory
VITA

Nevena Cvjetkovic was born in Bosnia and Herzegovina on January 29th, 1988, the first daughter of Nenad and Ljubenka Cvjetkovic. She immigrated to Florida with her parents and sister Jovana in 1998 at the age of 10. After graduation from Lakewood High School in St. Petersburg, FL she entered the University of Florida in Gainesville. She received her Bachelor of Science degree with a major in Microbiology and Cell Science and minor in Classical Studies in May of 2010. In August of 2010 she entered the Genetic Counseling program at The University of Health Science Center Graduate School of Biomedical Sciences at Houston.

Permanent Address:
5939 32nd St. North
St. Petersburg, FL, 33714