

Functional Assessment of Clubfoot Associated *HOXA9*, *TPM1*, and *TPM2* Variants Suggests a Potential Gene Regulation Mechanism

Katelyn S. Weymouth PhD, Susan H. Blanton PhD, Tamar Powell MS,
Chandrashekhar V. Patel PhD, Stuart A. Savill PhD, Jacqueline T. Hecht PhD

Received: 16 October 2015 / Accepted: 8 March 2016 / Published online: 28 March 2016
© The Association of Bone and Joint Surgeons® 2016

Abstract

Background Isolated nonsyndromic clubfoot is a common birth defect affecting 135,000 newborns worldwide each year. Although treatment has improved, substantial long-term morbidity persists. Genetic causes have been implicated in family-based studies but the genetic changes

This work was supported by National Institute of Child Health and Human Development Grant R01HD059862 (JTH) and Shriners Hospital for Children grants (JTH).

Each author certifies that he or she, or a member of his or her immediate family, has no funding or commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

Each author certifies that his or her institution approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.

This study was performed at the UTHealth Science Center–McGovern Medical School, Houston, TX, USA.

Electronic supplementary material The online version of this article (doi:10.1007/s11999-016-4788-1) contains supplementary material, which is available to authorized users.

K. S. Weymouth (✉), T. Powell, J. T. Hecht
Department of Pediatrics, UTHealth McGovern Medical School,
6431 Fannin St., MSB 3.306, Houston, TX 77030, USA
e-mail: katelyn.s.weymouth@uth.tmc.edu

S. H. Blanton
Department of Human Genetics, John P. Hussman Institute for
Human Genomics, University of Miami, Miller School of
Medicine, Miami, FL, USA

C. V. Patel
Department of Cell Biology and Anatomy, University of South
Carolina School of Medicine, Columbia, SC, USA

have eluded identification. Previously, using a candidate gene approach in our family-based dataset, we identified associations between clubfoot and four single nucleotide polymorphisms (SNPs) located in potential regulatory regions of genes involved in muscle development and patterning (*HOXA9*) and muscle function (*TPM1* and *TPM2*) were identified.

Questions/purposes Four SNPs, rs3801776/*HOXA9*, rs4075583/*TPM1*, rs2025126/*TPM2*, and rs2145925/*TPM2*, located in potential regulatory regions, were evaluated to determine whether they altered promoter activity.

Methods Electrophoretic mobility shift assays were performed on these four SNPs to identify allele-specific DNA-protein interactions. SNPs showing differential banding patterns were assessed for effect on promoter activity by luciferase assay. Undifferentiated (for *HOXA9*) and differentiated (for *TPM1* and *TPM2*) mouse cells were used in functional assays as a proxy for the in vivo developmental stage.

Results Functional analyses showed that the ancestral alleles of rs3801776/*HOXA9*, rs4075583/*TPM1*, and rs2025126/*TPM2* and the alternate allele of rs2145925/*TPM2* created allele-specific nuclear protein interactions and caused higher promoter activity. Interestingly, while

S. A. Savill
Diabetes and Endocrinology Research Group, Wrexham
Academic Unit, Betsi Cadwaladr University Health Board,
Wrexham, UK

J. T. Hecht
UTHealth School of Dentistry, Houston, TX, USA

rs4075583/TPM1 showed an allele-specific nuclear protein interaction, an effect on promoter activity was observed only when rs4075583/TPM1 was expressed in the 1.7kb haplotype construct.

Conclusion Our results show that associated promoter variants in *HOXA9*, *TPM1*, and *TPM2*, alter promoter expression suggesting that they have a functional role. Moreover and importantly, we show that alterations in promoter activity may be observed only in the context of the genomic architecture. Therefore, future studies focusing on proteins binding to these regulatory SNPs may provide important key insights into gene regulation in clubfoot.

Clinical Relevance Identifying the genetic risk signature for clubfoot is important to provide accurate genetic counseling for at-risk families, for development of prevention programs and new treatments.

Introduction

Clubfoot is a common congenital abnormality characterized by rigid inward turning of the contracted foot toward the midline of the body [6]. Severe calf muscle hypoplasia can be present, even after corrective treatment [24]. Approximately 20% of clubfeet occur as part of a syndrome, while the remaining 80% occur as an isolated abnormality (nonsyndromic, no other malformations) affecting one of 700 to 1000 newborns [5, 7, 9, 10, 30, 35].

Clubfoot is a complex birth defect wherein genetics and environmental factors play etiologic roles [11]. Even with overwhelming evidence for a genetic etiology, only a few genes have been implicated in clubfoot, including *TBX4*, *PITX1*, muscle contraction and apoptotic genes [2, 3, 13, 14, 20, 32, 43]. Therefore, most of the genetic variation remains to be identified.

Candidate gene approaches traditionally have been the preferred method for identifying potential clubfoot genes. Two previous studies focused on genes involved in muscle development and function because of the presence of calf muscle hypoplasia and muscle contraction in clubfoot [14, 43]. These included the *HOXA* and *D* genes that are involved in synchronized development of muscles, tendons, and cartilage of the limb; mutations in these genes have been shown to cause various limb malformations but not clubfoot [1, 8, 12, 16, 22, 27, 29, 33, 36, 37, 40, 41, 45]. Additionally, muscle contraction genes (*TPM2*, *MYH3*, *TNNT3*, and *TNNI2*) were interrogated because mutations in these genes cause distal arthrogyrosis syndromes 1, 2A, and 2B (OMIM 108120, 193700, and 601680). These autosomal dominant syndromes are characterized by multiple congenital joint contractures, clubfoot, and muscle

hypoplasia. In previous studies, a total of 94 single nucleotide polymorphisms (SNPs) spanning 35 genes involved in muscle patterning and contraction were interrogated in a large family-based clubfoot dataset [14, 43]. The strongest associations were found for SNPs in potential regulatory regions suggesting a common mechanism of gene regulation.

This led us to hypothesize that changes in gene expression contribute to clubfoot. We, therefore, asked whether four SNPs, rs3801776/*HOXA9*, rs4075583/*TPM1*, rs2025126/*TPM2*, and rs2145925/*TPM2*, located in potential regulatory regions, play a functional role in gene regulation.

Materials and Methods

Electrophoretic Mobility Shift Assay

Nuclear extracts isolated from undifferentiated and differentiated C2C12 mouse muscle cells were obtained from Active Motif (Carlsbad, CA, USA). Complementary single-stranded oligonucleotides (synthesized by Integrated DNA Technologies, Coralville, IA, USA) were annealed and end-labeled with corresponding radiolabeled α -³²P nucleotides on the forward strand to be used as probes (PerkinElmer, Waltham, MA, USA). Probes were designed with approximately 10 bp on either side of the candidate SNPs (Supplemental Table 1. Supplemental material is available with the online version of *CORR*[®]). Binding reactions were performed in the absence or presence of unlabeled competitor oligonucleotides (5×, 10×, and 50×) with 4.5 μg (undifferentiated) or 5.00 μg (differentiated) C2C12 nuclear extract in 20 μL of 1× binding buffer (50 mmol/L KCl, 10% glycerol, 0.5 mmol/L EDTA, 0.5 mmol/L dithiothreitol, 0.05% NP-40, 1 mmol/L phenylmethylsulfonyl fluoride) with 1 μg deoxyguanylic/deoxycytidylic (Sigma-Aldrich, St. Louis, MO, USA) and 1 μL of radiolabeled probe for 1 hour at 4°C. Protein-DNA complexes were separated on 5% polyacrylamide gels. The gels were dried and radioactivity was observed by autoradiography on radiographic film.

Cell Culture Technique

C2C12 mouse muscle cells were obtained from the American Type Culture Collection (ATCC# CRL-1772, Manassas, VA, USA). For the undifferentiated cells, the cells were cultured in Gibco[®] DMEM High Glucose medium (Life Technologies, Grand Island, NY, USA) supplemented with 10% fetal bovine serum. To initiate

differentiation, cells first were washed with phosphate buffered saline and Gibco® DMEM High Glucose medium supplemented with 2% horse serum for 5 days.

Generation of *TPM1* and *TPM2* Promoter Constructs

Promoter constructs for *TPM1* and *TPM2* have not been characterized; therefore we targeted the first 500-bps upstream of the transcriptional start site known to contain the basal promoter [28, 34]. To obtain *TPM1* (-500 to -1 bp) and *TPM2* (-461 to -1 bp) constructs, promoter sequences were amplified using primers incorporating *KpnI* and *XhoI* sites (Supplemental Table 2, Supplemental material is available with the online version of *CORR*®) by standard PCR methods from bacterial artificial chromosome (BAC) clones (*TPM1* [RP11-244F12] and *TPM2* [RP-112J3]) and ligated into the linearized pGL4.10 basic vector (Promega, Madison, WI, USA) using the In-Fusion® HD Cloning System (Clontech, Mountain View, CA, USA). For the *TPM2* promoter, only the region spanning -34 to -494 bps could be generated because secondary structure inhibited PCR amplification. The 400 bp *HOXA9* luciferase construct was obtained from one of the authors (CVP) and construct design was described Trivedi et al. [42]. Cloning efficiency was verified by Sanger sequencing.

Generation of Promoter With Regulatory SNP Construct

Allele-specific transcription factor binding identified by electrophoretic mobility shift assays was evaluated for effect on promoter activity. Constructs were designed to incorporate the electrophoretic mobility shift assay probe sequences with ancestral and alternate forms of the SNP using *KpnI* and *XhoI* cut site overhangs for direct ligation in the corresponding linearized gene promoter luciferase construct. The ancestral (G) allele of rs3801776 located at -206 was contained in the 400 bp *HOXA9* construct. Site-directed mutagenesis was used to create the *HOXA9* alternate (A) allele construct using QuikChange® II (Agilent Technologies, Santa Clara, CA, USA) following the manufacturer’s protocol. All constructs were verified by sequencing.

Generation of *TPM1* Haplotype Frequencies

To identify common *TPM1* haplotypes, the 1774 base pair region (-1714 to +60) containing eight *TPM1* SNPs, described by Savill et al. [38], was sequenced in 64 non-

Table 1. *TPM1* haplotype frequencies in non-Hispanic Whites

| Haplotype | SNP genotype | Frequency | | | | | |
|-----------|--------------|--------------|--------------|----------------------------------|----------------------------------|------------|-------------|
| | | Controls | | Cases | | | |
| | | All (n = 85) | All (n = 64) | Positive family history (n = 28) | Negative family history (n = 36) | | |
| | rs79854225 | rs76273871 | rs4075583 | rs4075584 | rs4075047 | rs73431508 | rs111470259 |
| 1 | G | G | A | T | A | T | C |
| 2 | G | G | G | C | A | T | C |
| 3 | A | A | G | T | G | C | T |
| 4 | G | G | G | T | A | T | C |
| | | 0.641 | 0.687 | 0.642 | 0.722 | 0.14 | 0.111 |
| | | 0.194 | 0.14 | 0.178 | 0.111 | 0.055 | 0.056 |
| | | 0.082 | 0.055 | 0.054 | 0.056 | 0.055 | 0.083 |

SNP = single nucleotide polymorphism.

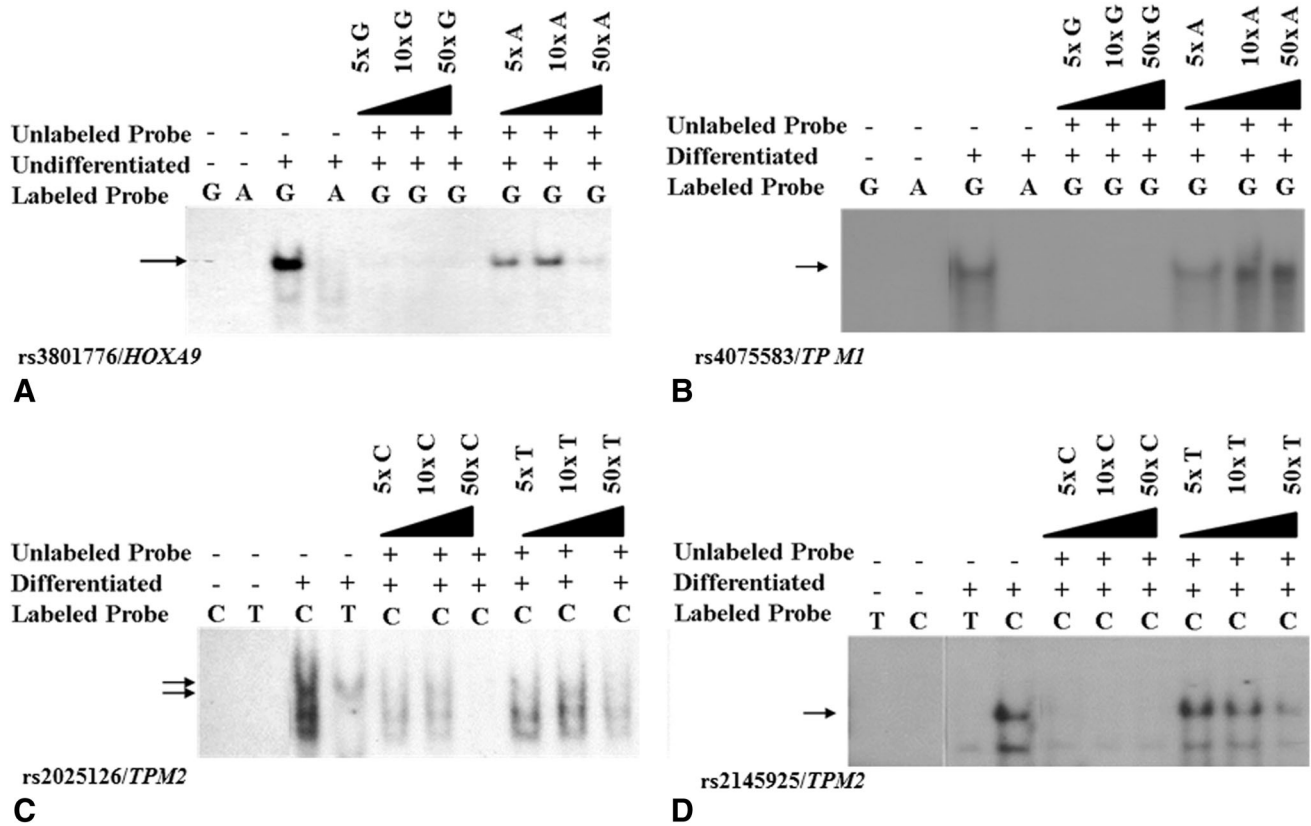


Fig. 1A–D Electrophoretic mobility shift assays were performed to evaluate whether the associated single nucleotide polymorphisms (SNPs) effected DNA-protein interactions. The ancestral G allele of (A) rs3801776/*HOXA9* and (B) rs4075583/*TPM1* created DNA-protein interactions that were eliminated when the alternate A allele was present. The (C) rs2025126/*TPM2* ancestral C allele created two DNA-protein interactions that were eliminated when the alternate T

allele was present. The (D) rs2145925/*TPM2* alternate C allele created a DNA-protein interaction that was eliminated with the ancestral T allele. The allele-specific DNA-protein interactions were confirmed through competitive assays using corresponding unlabeled probes in excess of 5x, 10x or 50x. The arrow indicates allele-specific DNA-protein complex.

Hispanic white: 28 multiplex (+FH) and 36 simplex (–FH) and 73 Hispanic: 21 multiplex (+FH) and 52 simplex (–FH) clubfoot probands. DNAs were amplified following standard PCR methods using two primers sets: set 1 (940 bp, T_m : 63°C): forward primer, ACTCACCTGAACTGACCTTCCCA; reverse primer, AAGTCACGCAGCAGGAACTAGGA; set 2: (1,281 bp, T_m : 56°C): forward primer, ATGGGCCTCAGCCTGACTCTTAAA; reverse primer, AACGGGTGGTGTGAGAAGGTTCT). Control *TPM1* haplotype frequencies were obtained from the 1000 genomes (<http://browser.1000genomes.org/index.html>) using CEU for non-Hispanic whites and MXL for the Mexican ancestry to represent the Hispanics. Haploview was used to identify the common haplotypes in cases and controls [4].

Generation of *TPM1* Haplotype Construct

Luciferase constructs were designed for the four common *TPM1* haplotypes (Table 1) for effect on *TPM1* promoter

activity. Inserts containing *TPM1* haplotypes 2 to 4 (Table 1) were obtained by *XhoI* and *SacI* double digestion of the *TPM1* constructs described by Savill et al. [38]. Haplotypes 2 to 4 inserts were ligated in the linearized *TPM1* promoter luciferase construct. *TPM1* haplotype 1 was generated using site-specific mutagenesis following the manufacturer’s protocol (QuikChange® II, Agilent Technologies). Cloning efficiency was verified by Sanger sequencing.

Luciferase Assays

C2C12 cells (100,000 cells/well) were seeded in 12-well plates for 24 hours before transfection. For transfection, 1.12 µg of luciferase reporter construct, 0.048 µg of Renilla (internal control), and Opti-MEM® (Thermo-Scientific, Grand Island, NY, USA) were incubated with FuGENE® HD (Promega) following the manufacturer’s protocol. All experiments were performed in triplicate and repeated at least three independent times. Luciferase activities for

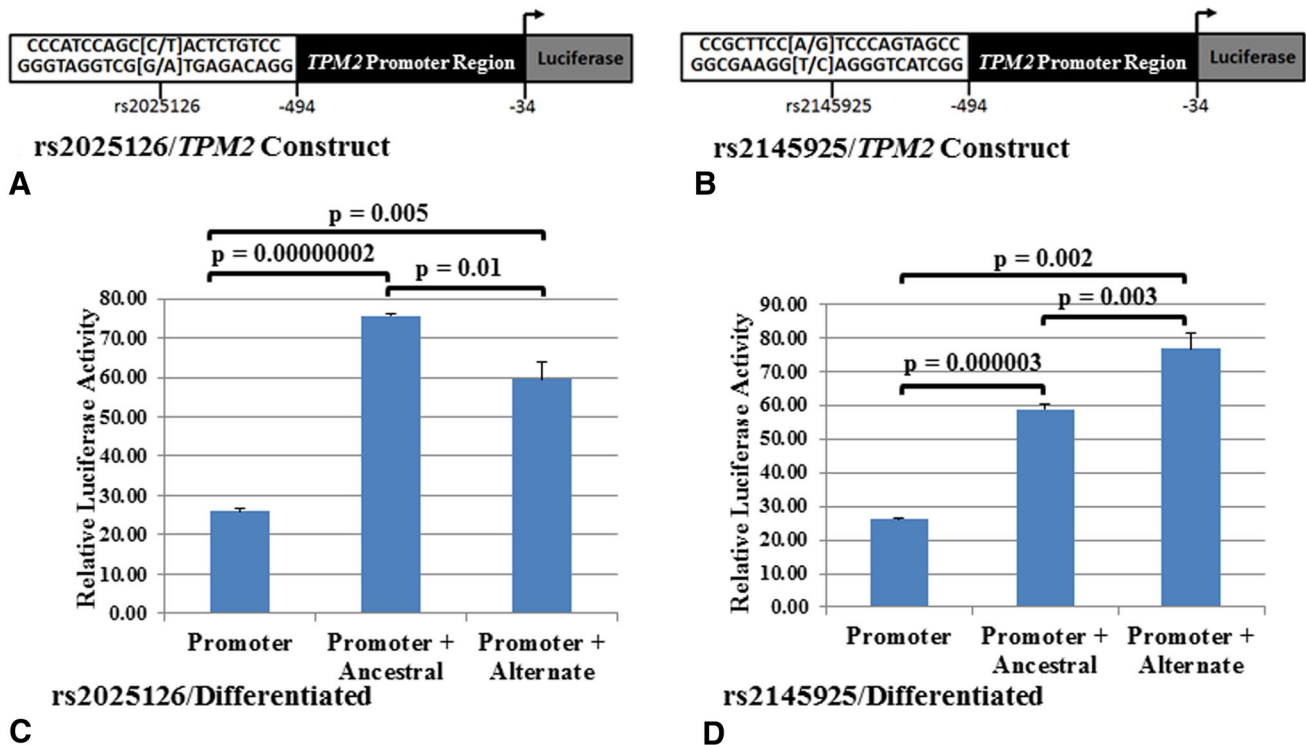


Fig. 2A–D Luciferase assays were performed to assess the effect of the DNA-protein interactions on *TPM2* promoter activity for rs2025126/*TPM2* and rs2145925/*TPM2*. The *TPM2* promoter construct incorporated 460-bps upstream of the transcriptional start site and was ligated into the pGL4.10 basic luciferase vector. The ancestral and alternate allele constructs were generated by ligating the

electrophoretic mobility shift assay double-stranded oligonucleotides that contained (A) rs2025126 or (B) rs2145925 in front of the *TPM2* promoter construct. (C) The ancestral allele of rs2025126 that creates a DNA-protein interaction caused a decrease in promoter activity. (D) The alternate allele of rs2145925 creating the DNA-protein interaction causes an increase in promoter activity.

undifferentiated C2C12 cells were determined 48 hours after transfection using the dual-luciferase system (Promega). For differentiated C2C12 cells, the media were replaced 48 hours after transfection with DMEM medium supplemented with 2% horse serum (Life Technologies) and luciferase activity was measured 5 days later. Unpaired t-tests were used to compare luciferase expression between constructs.

Results

Electrophoretic mobility shift assays identified nuclear protein interactions for the ancestral alleles of rs3801776C/*HOXA9* (Fig. 1A), rs4075583G/*TPM1* (Fig. 1B), rs2025126C/*TPM2* (Fig. 1C) and the alternate allele of rs2145925G/*TPM2* (Fig. 1D). Competitive assays for each of these nuclear protein interactions confirmed that the DNA-protein complexes were allele-specific and showed differential binding affinities.

Luciferase reporter assays were used to assess the effect of these allele-specific binding nuclear proteins on promoter activity (Figs. 2 and 3). The presence of the genomic

sequence containing either the *TPM1* or *TPM2* SNPs significantly increased promoter activity above the basal level suggesting that they have enhancer activity (Figs. 2 and 3). For rs2025126/*TPM2* (Fig. 1C), the ancestral allele nuclear protein increased promoter activity compared with the alternate form only in differentiated C2C12 mouse muscle cells (Fig. 2C), which is in contrast to rs2145925/*TPM2* alternate allele nuclear protein that increased promoter activity (Fig. 2D). Although the ancestral allele of rs4075583/*TPM1* creates a unique nuclear protein interaction, there was no difference in promoter activity in the presence of the ancestral or alternate alleles. In contrast, the ancestral allele for rs3801776/*HOXA9* increased promoter activity in undifferentiated muscle cells (Fig. 3C).

rs4075583/*TPM1* was next evaluated in the context of the genomic region containing seven additional SNPs in muscle cells (Fig. 4A). We identified the common haplotypes by genotyping probands from 64 non-Hispanic whites (Table 1) and 73 Hispanic families (Table 2). The most common in non-Hispanic whites and Hispanics was haplotype 1, while haplotypes 2 and 3 were the second most common in non-Hispanic whites and Hispanics, respectively, indicating ethnic differences. Haplotype 1

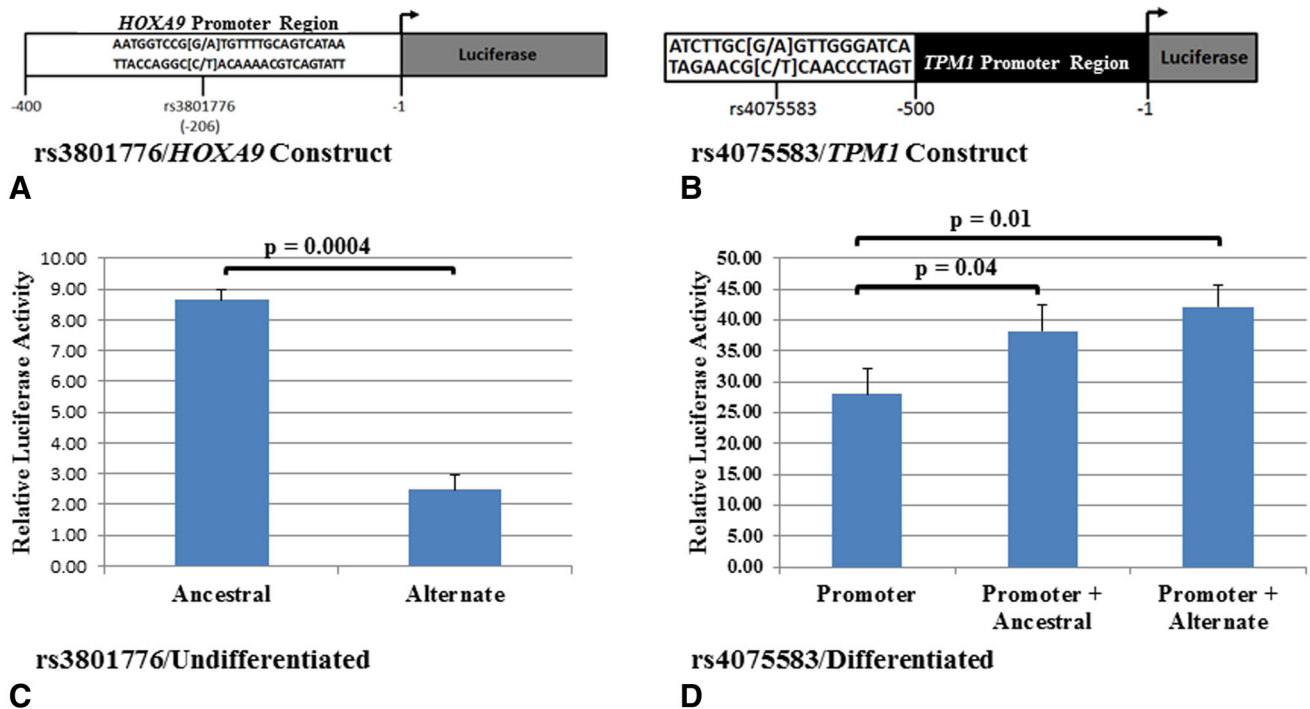


Fig. 3A–D Luciferase assays were performed to assess the effect of the DNA-protein interactions on *TPM1* promoter activity for rs4075583 and *HOXA9* promoter activity for rs3801776. **(A)** The *HOXA9* promoter construct containing the rs3801776 is shown. **(B)** The rs4075583/*TPM1* construct contained 500-bps upstream of the transcriptional start site for the *TPM1* skeletal muscle isoform ligated into the pGL4.10 basic luciferase vector. **(C)** The ancestral allele-specific DNA-protein interaction for rs3801776 increased promoter

activity in undifferentiated mouse muscle cells. **(D)** The ancestral G allele and alternate *TPM1* constructs were generated by ligating the electrophoretic mobility shift assay double-stranded oligonucleotides in front of the *TPM1* promoter construct (as shown in Illustration C). Although the ancestral allele for rs4075583/*TPM1* creates a DNA-protein interaction (as shown in Illustration B), it did not significantly affect promoter activity in differentiated mouse muscle cells.

containing the alternate allele, rs4075583A showed the highest promoter activity, while the ancestral allele in haplotypes 2 to 4 produced less promoter activity (Fig. 4 B). These findings are of interest because none of the seven other SNPs create nuclear protein interactions (data not shown). Notably, haplotype 3, containing the most allelic variation (variation at five SNP positions) in comparison to the most common haplotype (Haplotype 1) showed the lowest in promoter activity.

Discussion

This study assessed the functionality of potential regulatory variants, rs3801776/*HOXA9*, rs2025126/*TPM2*, rs2145925/*TPM2*, and rs4075583/*TPM1* previously found to be associated with clubfoot [43]. These genes initially were evaluated because of the role they play in muscle development and/or function, which is abnormal in clubfoot [15, 17, 19, 21, 23–26, 31, 39, 44]. Three SNPs, rs3801776/*HOXA9*, rs2025126/*TPM2*, and rs2145925/*TPM2*, created allele-specific nuclear protein interactions that individually

had significant functional consequences by affecting promoter activity in muscle cells [14, 43]. While rs4075583/*TPM1* creates an allele-specific nuclear protein interaction, it was not sufficient to alter promoter activity. However, it did alter activity when in the context of the surrounding approximately 1.7 kb genomic architecture. These results suggest that these variants may enhance or repress gene expression individually or in combination, which may affect muscle morphogenesis and/or contraction. This is consistent with a multifactorial inheritance model proposed for clubfoot which predicts that changes in multiple genes are necessary to cause phenotypic consequences [11].

Muscle development and function are multifaceted processes that involve multiple genes expressed at precise times and locations. Alterations during these key processes could cause a cascade of muscle anomalies such as muscle hypoplasia and abnormal muscle contraction, both phenotypic characteristics of clubfoot. To account for these differences, we tested the functionality of our SNPs in undifferentiated and differentiated mouse muscle cells as a proxy for in vivo stages of development. During the beginning stages of myogenesis (undifferentiated cells),

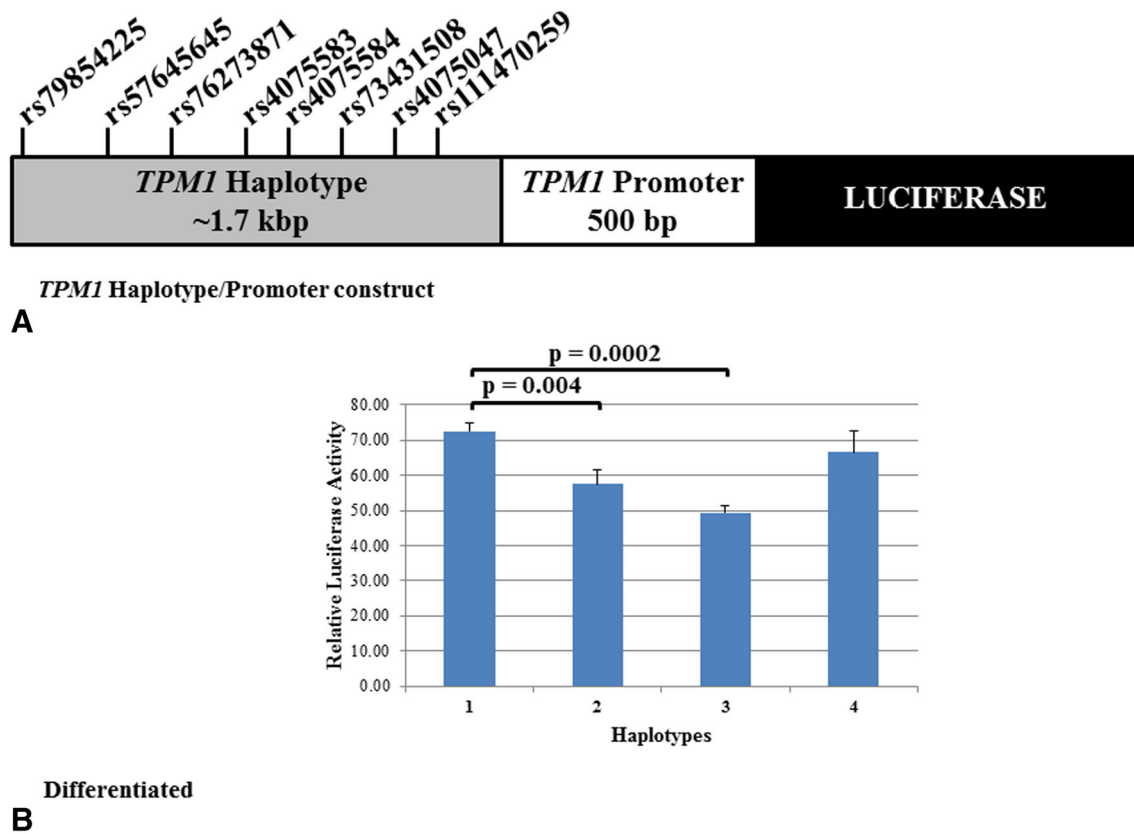


Fig. 4A–B Luciferase assays were performed to evaluate whether specific *TPM1* haplotypes affected *TPM1* promoter activity. The 1774 bp *TPM1* region containing rs4075583/*TPM1* has been shown to influence expression depending on cell type and haplotype. Based in that finding, (A) constructs were designed to incorporate the four most common *TPM1* haplotypes and the *TPM1* promoter region. (B) The

four common haplotypes produced varying degrees of promoter activity in differentiated mouse muscle cells with the least promoter activity found with haplotype 3, which contains the alternate A allele of rs40755863 which eliminates the DNA-protein interaction observed with the ancestral G allele.

muscle cell migration and patterning are highly regulated with specific genes playing important roles in limb and foot development. *HOXA9*, a member of the homeobox A gene transcription factor cluster plays a role in synchronized patterning and differentiation of muscles, tendons, and cartilages in embryonic fore- and hindlimbs. *HOXA9* regulates other transcription factors such as *LBX1*, which are important in the migration of muscle precursor cells into the developing hindlimb bud. The observed decrease in promoter activity with the alternate allele of rs3801776/*HOXA9* (Fig. 3C) could lead to alterations in muscle structure, calf muscle size, and slow and fast-twitch muscle composition, all of which have been reported in clubfoot [17, 24, 25].

The maturing of myoblasts (differentiated muscle cells) into long, multinucleated muscle fibers facilitates the functional properties of muscle such as contraction, which is a highly regulated process involving many proteins including *TPM1* in the quick, easily fatigued fast-twitch muscle fibers and *TPM2* in the slow, long contracting,

slow-twitch fibers [18]. Based on this information, we hypothesized that variation in expression of *TPM1* and *TPM2* could alter muscle contraction related to changes in binding of regulatory proteins. We found that two SNPs in *TPM2* had functional activity that significantly differed depending on which variant was present. For example, rs2025126, located upstream of *TPM2*, had decreased promoter activity when the alternate allele was present because it eliminated nuclear protein interactions (Fig. 2 C). Decreased *TPM2* could allow for continuous actin-myosin interaction causing perpetual muscle contraction resulting in a contracted foot. In contrast, rs2145925, located in the first intron of *TPM2*, created a nuclear protein interaction in the presence of the alternate allele that increased promoter activity (Fig. 2D). Increased expression of *TPM2* could lead to the actin inhibition limiting muscle contraction and foot movement thereby potentially causing muscle wasting as manifested by calf muscle hypoplasia. Imbalance of these important muscle functions could contribute to clubfoot.

Table 2. TPM1 haplotype frequencies in Hispanics

| SNP genotype | | Frequency | | | |
|--------------|---|-------------------|--------------|----------------------------------|----------------------------------|
| | | Controls (n = 85) | | Cases (n = 64) | |
| Haplotype | rs79854225 rs57645645 rs76273871 rs4075583 rs4075584 rs4075047 rs73431508 rs111470259 | All (n = 85) | All (n = 64) | Positive family history (n = 28) | Negative family history (n = 36) |
| 1 | G G G A A T T C | 0.621 | 0.699 | 0.667 | 0.712 |
| 2 | G G G G C C T C | 0.106 | 0.048 | 0.048 | 0.048 |
| 3 | A A G G T G C T | 0.220 | 0.171 | 0.190 | 0.163 |
| 4 | G G G A G T C C | 0.030 | 0.014 | 0.000 | 0.019 |

SNP = single nucleotide polymorphism.

Previously, we found an association with rs4075583, a variant located in intron 1 of the *TPM1* skeletal muscle isoform, in a potential enhancer or suppressor regulatory region [43]. While the ancestral allele of rs4075583/*TPM1* created a nuclear protein interaction, this complex alone did not alter promoter activity (Fig. 4B). However, different levels of expression were found when rs4075583/*TPM1* was evaluated in the context of the approximate 1.7 kb genomic region (Fig. 4B). Haplotype 3, the only haplotype containing the alternate allele, had the lowest activity consistent with elimination of the nuclear protein interaction (Fig. 4B and C). These findings show that the binding affinity of the ancestral allele nuclear protein is influenced by the allelic composition of the surrounding genomic region and also suggest the importance of evaluating each SNP in its genomic context to understand its functional role.

The lack of an animal model and/or human muscle cell lines for use in functional studies is a limitation of this study. To circumvent this problem, C2C12 mouse muscle cells were used for the functional analyses because the phenotype of these cells can be manipulated to mimic different stages of muscle development, thereby allowing us to assess the effects that may occur at different times in development. For example, in the undifferentiated state, myoblast cells are mononucleated like those cells involved in early muscle patterning where *HOXA9* plays an important role. In contrast, differentiated muscle cells are long multinucleated myotubes that contain the contractile proteins including *TPM1* and *TPM2* that have specific functions. While this system may not perfectly mimic the in vivo state, the observed changes provide evidence that these SNPs play a functional role and this expression may be observed only at different stages of development. Therefore, this study lays the groundwork for future studies.

In this study, we show that variants associated with clubfoot have functional consequences individually and in the context of the surrounding genomic architecture. This latter finding is important because lack of a positive functional outcome for an individual SNP may lead to the inappropriate exclusion of some variants or genes as playing a role in clubfoot. Therefore, all variants need to be analyzed individually and then in combination with surrounding SNPs. The future challenge resides in identifying all the common and rare variants, along with the haplotypes that confer increased risk for clubfoot. Next-generation sequencing approaches applied to family-based datasets should begin to yield all potential rare and common variants that will need to be functionally assessed. Knowledge of genetic variation underlying clubfoot will provide important insights into the etiologic pathways, provide a means to assess genetic risk and accurate genetic

counseling based on individual risks, and develop prevention programs and potentially new treatments.

Acknowledgments We thank the families who participated in this study. We also are indebted to Elena Serna, Rosa Martinez (both from the Department of Pediatrics, UT Health McGovern Medical School, Houston, TX), and S. Shahrukh Hashmi MD, PhD (Department of Pediatrics, UT Health, Houston, TX) for collecting the families, tissue samples, and maintaining the database.

References

- Akarsu AN, Stoilov I, Yilmaz E, Sayli BS, Sarfarazi M. Genomic structure of HOXD13 gene: a nine polyaniline duplication causes synpolydactyly in two unrelated families. *Hum Mol Genet.* 1996; 5:945-952.
- Alvarado DM, Aferol H, McCall K, Huang JB, Tychy M, Buchan J, Cady J, Gonzales PR, Dobbs MB, Gurnett CA. Familial isolated clubfoot is associated with recurrent chromosome 17q23.1q23.2 microduplications containing TBX4. *Am J Hum Genet.* 2010;87:154-160.
- Alvarado DM, McCall K, Aferol H, Silva MJ, Garbow JR, Spees WM, Patel T, Siegel M, Dobbs MB, Gurnett CA. Pitx1 haploinsufficiency causes clubfoot in humans and a clubfoot-like phenotype in mice. *Hum Mol Genet.* 2011;20:3943-3952.
- Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* 2005;21:263-265.
- Beals RK. Club foot in the Maori: a genetic study of 50 kindreds. *N Z Med J.* 1978;88:144-146.
- Bohm M. Embryologic origin of clubfoot. *J Bone Joint Surg.* 1929;11:229-259.
- Carey M, Bower C, Mylvaganam A, Rouse I. Talipes equinovarus in Western Australia. *Paediatr Perinat Epidemiol.* 2003;17: 187-194.
- Carpenter EM, Goddard JM, Davis AP, Nguyen TP, Capocchi MR. Targeted disruption of Hoxd-10 affects mouse hindlimb development. *Development.* 1997;124:4505-4514.
- Ching GH, Chung CS, Nemechek RW. Genetic and epidemiological studies of clubfoot in Hawaii: ascertainment and incidence. *Am J Hum Genet.* 1969;21:566-580.
- Chung CS, Nemechek RW, Larsen IJ, Ching GH. Genetic and epidemiological studies of clubfoot in Hawaii: general and medical considerations. *Hum Hered.* 1969;19:321-342.
- Dobbs MB, Gurnett CA. Genetics of clubfoot. *J Pediatr Orthop B.* 2012;21:7-9.
- Dobbs MB, Gurnett CA, Pierce B, Exner GU, Robarge J, Morcuende JA, Cole WG, Templeton PA, Foster B, Bowcock AM. HOXD10 M319K mutation in a family with isolated congenital vertical talus. *J Orthop Res.* 2006;24:448-453.
- Ester AR, Tyerman G, Wise CA, Blanton SH, Hecht JT. Apoptotic gene analysis in idiopathic talipes equinovarus (clubfoot). *Clin Orthop Relat Res.* 2007;462:32-37.
- Ester AR, Weymouth KS, Burt A, Wise CA, Scott A, Gurnett CA, Dobbs MB, Blanton SH, Hecht JT. Altered transmission of HOX and apoptotic SNPs identify a potential common pathway for clubfoot. *Am J Med Genet A.* 2009;149A:2745-2752.
- Feldbrin Z, Gilai AN, Ezra E, Khermosh O, Kramer U, Wientroub S. Muscle imbalance in the aetiology of idiopathic club foot: an electromyographic study. *J Bone Joint Surg Br.* 1995;77:596-601.
- Fromental-Ramain C, Warot X, Messadecq N, LeMeur M, Dolle P, Chambon P. Hoxa-13 and Hoxd-13 play a crucial role in the patterning of the limb autopod. *Development.* 1996;122:2997-3011.
- Fukuhara K, Schollmeier G, Uthoff HK. The pathogenesis of club foot: a histomorphometric and immunohistochemical study of fetuses. *J Bone Joint Surg Br.* 1994;76:450-457.
- Gordon AM, Homsher E, Regnier M. Regulation of contraction in striated muscle. *Physiol Rev.* 2000;80:853-924.
- Gupta P, Singla R, Gupta R, Jindal R, Bahadur R. Accessory soleus muscle in clubfoot deformity: a report in four feet. *J Pediatr Orthop B.* 2007;16:106-109.
- Gurnett CA, Alaei F, Kruse LM, Desruisseau DM, Hecht JT, Wise CA, Bowcock AM, Dobbs MB. Asymmetric lower-limb malformations in individuals with homeobox PITX1 gene mutation. *Am J Hum Genet.* 2008;83:616-622.
- Herceg MB, Weiner DS, Agamanolis DP, Hawk D. Histologic and histochemical analysis of muscle specimens in idiopathic talipes equinovarus. *J Pediatr Orthop.* 2006;26:91-93.
- Houghton L, Rosenthal N. Regulation of a muscle-specific transgene by persistent expression of Hox genes in postnatal murine limb muscle. *Dev Dyn.* 1999;216:385-397.
- Ippolito E, De Maio F, Mancini F, Bellini D, Orefice A. Leg muscle atrophy in idiopathic congenital clubfoot: is it primitive or acquired? *J Child Orthop.* 2009;3:171-178.
- Irani RN, Sherman MS. The pathological anatomy of idiopathic clubfoot. *Clin Orthop Relat Res.* 1972;84:14-20.
- Isaacs H, Handelsman JE, Badenhorst M, Pickering A. The muscles in club foot: a histological histochemical and electron microscopic study. *J Bone Joint Surg Br.* 1977;59:465-472.
- Karapinar L, Kaya A, Altay T, Ozturk H, Surenkok F. Congenital clubfoot associated with an accessory soleus muscle. *J Am Podiatr Med Assoc.* 2008;98:408-413.
- Kardon G. Muscle and tendon morphogenesis in the avian hind limb. *Development.* 1998;125:4019-4032.
- Khambata-Ford S, Liu Y, Gleason C, Dickson M, Altman RB, Batzoglou S, Myers RM. Identification of promoter regions in the human genome by using a retroviral plasmid library-based functional reporter gene assay. *Genome Res.* 2003;13:1765-1774.
- Kmita M, Tarchini B, Zakany J, Logan M, Tabin CJ, Duboule D. Early developmental arrest of mammalian limbs lacking HoxA/HoxD gene function. *Nature.* 2005;435:1113-1116.
- Lochmiller C, Johnston D, Scott A, Risman M, Hecht JT. Genetic epidemiology study of idiopathic talipes equinovarus. *Am J Med Genet.* 1998;79:90-96.
- Loren GJ, Karpinski NC, Mubarak SJ. Clinical implications of clubfoot histopathology. *J Pediatr Orthop.* 1998;18:765-769.
- Lu W, Bacino CA, Richards BS, Alvarez C, Vandermeer JE, Vella M, Ahituv N, Sikka N, Dietz FR, Blanton SH, Hecht JT. Studies of TBX4 and chromosome 17q23.1q23.2: an uncommon cause of nonsyndromic clubfoot. *Am J Med Genet A.* 2012; 158A:1620-1627.
- Mark M, Rijli FM, Chambon P. Homeobox genes in embryogenesis and pathogenesis. *Pediatr Res.* 1997;42:421-429.
- Maston GA, Evans SK, Green MR. Transcriptional regulatory elements in the human genome. *Annu Rev Genomics Hum Genet.* 2006;7:29-59.
- Moorthi RN, Hashmi SS, Langois P, Canfield M, Waller DK, Hecht JT. Idiopathic talipes equinovarus (ITEV) (clubfeet) in Texas. *Am J Med Genet A.* 2005;132A:376-380.
- Muragaki Y, Mundlos S, Upton J, Olsen BR. Altered growth and branching patterns in synpolydactyly caused by mutations in HOXD13. *Science.* 1996;272:548-551.
- Rodriguez-Guzman M, Montero JA, Santesteban E, Ganay Y, Macias D, Hurler JM. Tendon-muscle crosstalk controls muscle bellies morphogenesis, which is mediated by cell death and retinoic acid signaling. *Dev Biol.* 2007;302:267-280.

38. Savill SA, Leitch HF, Daly AK, Harvey JN, Thomas TH. Polymorphisms in the tropomyosin TPM1 short isoform promoter alter gene expression and are associated with increased risk of metabolic syndrome. *Am J Hypertens.* 2010;23:399-404.
39. Shimode K, Miyagi N, Majima T, Yasuda K, Minami A. Limb length and girth discrepancy of unilateral congenital clubfeet. *J Pediatr Orthop B.* 2005;14:280-284.
40. Shrimpton AE, Levinsohn EM, Yozawitz JM, Packard DS Jr, Cady RB, Middleton FA, Persico AM, Hootnick DR. A HOX gene mutation in a family with isolated congenital vertical talus and Charcot-Marie-Tooth disease. *Am J Hum Genet.* 2004;75:92-96.
41. Tarchini B, Huynh TH, Cox GA, Duboule D. HoxD cluster scanning deletions identify multiple defects leading to paralysis in the mouse mutant Ironside. *Genes Dev.* 2005;19:2862-2876.
42. Trivedi CM, Patel RC, Patel CV. Differential regulation of HOXA9 expression by nuclear factor kappa B (NF-kappaB) and HOXA9. *Gene.* 2008;408:187-195.
43. Weymouth KS, Blanton SH, Bamshad MJ, Beck AE, Alvarez C, Richards S, Gurnett CA, Dobbs MB, Barnes D, Mitchell LE, Hecht JT. Variants in genes that encode muscle contractile proteins influence risk for isolated clubfoot. *Am J Med Genet A.* 2011;155A:2170-2179.
44. Windisch G, Anderhuber F, Haldi-Brandl V, Exner GU. Additional muscle in idiopathic club foot. *Eur J Pediatr Surg.* 2006; 16:294-296.
45. Zakany J, Duboule D. The role of Hox genes during vertebrate limb development. *Curr Opin Genet Dev.* 2007;17:359-366.