Mutation analysis of TBX1 in non-deleted patients with features of DGS/VCFS or isolated cardiovascular defects

Weilong Gong, Shoshanna Gottlieb, Joelle Collins, Adam Blescia, Harry Dietz, Elizabeth Goldmuntz, Donna M McDonald-McGinn, Elaine H Zackai, Beverly S Emanuel, Deborah A Driscoll, Marcia L Budarf

Division of Human Genetics and Molecular Biology, The Children’s Hospital of Philadelphia, 3516 Civic Center Boulevard, Philadelphia, PA 19104, USA
E H Zackai
B S Emanuel
D A Driscoll
M L Budarf

Howard Hughes Medical Institute, Institute for Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA
H Dietz

Division of Cardiology, The Children’s Hospital of Philadelphia, Philadelphia, PA, USA
E Goldmuntz

Division of Pediatrics, The Children’s Hospital of Philadelphia, Philadelphia, PA, USA
E Goldmuntz
E H Zackai
B S Emanuel
M L Budarf

Division of Obstetrics and Gynecology, University of Pennsylvania School of Medicine, Philadelphia, PA, USA
D A Driscoll

Correspondence to: Dr Budarf, budarf@mail.med.upenn.edu

Editor—Microdeletions of chromosomal region 22q11.2 (del22q11) have been associated with several genetic disorders, including Di-George syndrome (DGS), velocardiofacial syndrome (VCFS), and conotruncal anomaly face syndrome (CTAFS). The major clinical features associated with del22q11 are conotruncal heart defects, hypoplastic or aplastic thymus and parathyroid glands, facial dysmorphism, and learning difficulties. Many of the structures affected in the del22q11 syndrome are derivatives of the branchial apparatus, which is populated by the rostral neural crest cells. This has led to the hypothesis that haploinsufficiency of a gene(s) within the deleted region disrupts the development of these structures. The majority of patients carry a common ~3 Mb deletion and over 20 genes have been mapped to the deleted region. The phenotypic features seen in association with the 22q11.2 deletion are highly variable, even among family members with the same sized deletion. Monozygotic twins with discordant phenotypes have also been reported. Thus, additional factors, such as genetic background and environment, can modify the effect of haploinsufficiency of genes within 22q11.2.

Recent studies have shown that mice heterozygously deleted (Df1+/) for part of the region of mouse chromosome 16 homologous to the DGS/VCFS region of 22q11.2 have heart defects similar to those found in DGS/VCFS patients. On a different genetic background these mice have been found also to have thymic defects. Several genes are located within the Df1 deleted region, including Tbx1. Tbx1 is a member of the T box gene family of DNA binding transcription factors. T box genes have been shown to play an important role in the regulation of developmental processes. Haploinsufficiency of two T box genes, TBX3 and TBX5, are associated with the human genetic diseases ulnar-mammary syndrome and Holt-Oram syndrome, respectively.

A targeted deletion of Tbx1 has been made and heterozygous mice were found to have aortic arch defects similar to those seen in the Df1/+ mice. Mice homozygously deleted for Tbx1 die before or at birth with all the hallmark features seen in DGS/VCFS, including outflow tract defects of the heart, hypoplasia of the thymus and parathyroid glands, cleft palate, and facial abnormalities. These data strongly suggest human TBX1 as a candidate gene for many of the key features of DGS/VCFS. Mutation analysis of TBX1 in patients who have DGS/VCFS phenotypes but no detectable deletion of 22q11 will provide insight into the role of TBX1 in the aetiology of these disorders.

Methods

We screened 105 patients, who did not have deletions of 22q11.2, for mutations in TBX1. Samples were collected from two groups of patients: 40 had clinical features consistent with the diagnosis of DGS/VCFS and 65 had isolated cardiac lesions frequently seen in DGS/VCFS (11 with interrupted aortic arch (IAA) type B, one IAA unspecified, 18 with truncus arteriosus (TA), and 35 with aortic arch anomalies (AAA)). The diagnostic criteria for DGS/VCFS required that a patient have two or more of the major features associated with these disorders: congenital heart defects, immune defects (that is, absent thymus, immunodeficiency, or a history of frequent infections), palatal abnormalities and/or speech difficulties, facial dysmorphism, and a history of learning disabilities. Patients with the conotruncal cardiac defects (TA and IAA) and AAA were included in this study because 22q11 deletions have been found in 50% of patients with IAA, 35% with TA, and 22% with AAA.

Further, conotruncal heart defects are the most common heart lesions seen in DGS/VCFS patients. These are also the type of heart defects seen in the mouse models of DGS/VCFS. None of the patients in this study was found to have a 22q11.2 deletion when tested by FISH using the marker N25 (D22S75). Additional FISH analysis to exclude the possibility of a smaller deletion was performed on 23 of the 40 DGS/VCFS patients and 15 of 18 patients with TA. This analysis did not detect any smaller deletions in the patients tested. Further, Southern blot
analysis using a mixture of probes from the TBX1 gene was performed on 100 of the 105 patients to exclude deletions, insertions, and rearrangements that might not be detected by PCR analysis or FISH. No rearrangements or deletions were detected. Linkage analysis to ascertain 22q11.2 involvement was not possible owing to the paucity of families with multiple affected members.

The genomic structure of TBX1 has been previously reported. Two alternatively spliced transcripts, TBXIA and TBXIB (AF012130 and AF012131, respectively), were found that differed in their terminal exons. However, when the mouse genomic sequence became available, comparison of the human and mouse genomic sequences showed that exon 9A of TBXIA and exons 9B and 10 of TBXIB are not conserved (fig 1A). Analysis of the sequence with GenScan (genes.mit.edu/GENSCAN.html) predicted this conserved region to be an excellent candidate exon. A probe, 9C, was generated from this conserved region and hybridised on a multiple tissue northern blot (16 adult and 4 fetal tissues from Clontech). The expression pattern seen with 9C (data not shown) was identical to that previously published for similar northern blots using a probe from the TBX1 T box region. A search of the human EST database using the exon 9A, exons 9B and 10, or exon 9C sequences showed over 12 EST matches for exon 9C, but none for the previously characterised terminal exons. RT-PCR with primers designed from exon 9C and RNA from human adult skeletal muscle, adult testis, and fetal lung, tissues that were previously shown to express TBX1 confirm the presence of this additional alternatively spliced TBXIC transcript (AF373867). The comparison of levels of RT-PCR products using primers from the 3'UTR of TBXIA, TBXIB, and TBXIC show that TBX1C is the major transcript.

Primers for mutation analysis were designed using the TBXIC sequence. Primers for all nine exons, including exon-intron junction regions, were selected (primer sequences available upon request). For exon 9C, the acceptor splice site, the coding region, plus an additional 30 bp 3' of the stop codon (TAA) were analysed. PCR was performed with genomic DNA from all 105 patients and normal controls. Single stranded conformational polymorphism (SSCP) and/or conformation sensitive gel electrophoresis (CSGE) were used to screen for mutations. Changes seen by SSCP or CSGE were confirmed by sequence analysis.

Results

As shown in table 1, we identified eight common polymorphisms and 10 rare variants. Only one of the common polymorphisms results in an amino acid change (1189C→A, His397Asn). In the mouse, this amino acid residue is a histidine. Of the remaining seven polymorphisms, one is a C→T transition in the 5'UTR, the second is a base pair substitution in intron 7, and the last five polymorphisms are all single base pair changes in the coding region that do not alter the amino acid sequence. These common SNPs may be useful

Figure 1. (A) Schematic representation of the gene organisation of the alternatively processed transcripts, TBXIA, TBXIB, and TBXIC. All three transcripts have the same sequence organisation from exon 1 to exon 8 and these regions are highly conserved in mouse (indicated by solid boxes). The nucleotide identity between the human and mouse sequences from exons 1–8 ranges from 80% to 91%. However, exon 9A of TBXIA and exons 9B and 10 of TBXIB are not conserved in mice (represented by unfilled boxes). For TBXIC, the first 478 nucleotides of exon 9C, representing the coding region, show 80% identity with the corresponding mouse sequence. In the 3'UTR region of exon 9C, immediately after the stop codon, the identity of a block of 250 bp drops to 19.6% compared to the mouse, shown as shaded boxes. The last 215 nucleotides of exon 9C are highly conserved between mouse and human with 88% identity. (B) Diagram showing distribution of the rare variants relative to the TBX1 gene structure. The T box domain is represented as the black boxes and the coding region outside the T box domain is shown as striped boxes. The positions of the rare variants are indicated by arrows (the numbers refer to the nucleotide sequence). (C) The conserved non-coding regions (CNSs) between the human and mouse. Comparative sequence alignment was performed using PipMaker (http://bio.cse.psu.edu). Six CNSs with 75% or higher identity over at least 100 bp were identified in the intron regions, indicated by arrows.
Table 1  Mutation screening of TBX1 in 105 patients

<table>
<thead>
<tr>
<th>Position</th>
<th>Nucleotide change*</th>
<th>Amino acid change</th>
<th>No of patients†</th>
<th>No of controls‡</th>
<th>Patient</th>
<th>Diagnosis</th>
<th>Parents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon 2</td>
<td>39nt 5’ to ATG C→T</td>
<td>No</td>
<td>1/105</td>
<td>0/101</td>
<td>p23</td>
<td>DGS</td>
<td>Mother is carrier</td>
</tr>
<tr>
<td>Exon 5</td>
<td>576 C→T</td>
<td>192(Arg/Arg)</td>
<td>1/105</td>
<td>0/101</td>
<td>p60</td>
<td>TA</td>
<td>Not available</td>
</tr>
<tr>
<td>Exon 7</td>
<td>886 C→A</td>
<td>296(Arg/Arg)</td>
<td>1/105</td>
<td>1/101</td>
<td>p89</td>
<td>IAA typeB</td>
<td>Not available</td>
</tr>
<tr>
<td>Exon 9</td>
<td>1049 G→A</td>
<td>350(Gly/Gly)</td>
<td>1/105</td>
<td>0/101</td>
<td>Ap22</td>
<td>AAA</td>
<td>Not available</td>
</tr>
<tr>
<td>1187 C→T</td>
<td>1049 G→A</td>
<td>350(Gly/Gly)</td>
<td>1/105</td>
<td>0/101</td>
<td>Ap18</td>
<td>AAA</td>
<td>Not available</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p13</td>
<td>DGS</td>
<td>Mother is carrier</td>
</tr>
<tr>
<td>Exon 10</td>
<td>1334-1348del15bp‡</td>
<td>445-449delGYHPH</td>
<td>1/105</td>
<td>0/101</td>
<td>p23</td>
<td>DGS</td>
<td>Father is carrier</td>
</tr>
<tr>
<td>Exon 9</td>
<td>1371-1372insCAC</td>
<td>457-458insHis</td>
<td>1/105</td>
<td>0/101</td>
<td>p22</td>
<td>VCFS</td>
<td>Mother is carrier</td>
</tr>
<tr>
<td>Exon 8</td>
<td>1399-1427dup30bp§</td>
<td>466-476dup10Ala</td>
<td>1/105</td>
<td>0/101</td>
<td>p46</td>
<td>IAA</td>
<td>Not available</td>
</tr>
<tr>
<td>Exon 9</td>
<td>1404-1412delGCTGCCGCC</td>
<td>468-471del3Ala</td>
<td>1/105</td>
<td>0/101</td>
<td>p65</td>
<td>DGS/VCFS</td>
<td>Mother is carrier</td>
</tr>
</tbody>
</table>

*B1 refers to the first nucleotide of the initiation codon ATG.
†The numbers of patients or controls with polymorphisms or variants are reported for heterozygotes. Complete genotyping was not performed in this study.
‡Deletion of GCTACCACCACGCAG.
§Duplication of GCCGCCGCCGCCGCCGCCGCCGCC.
¶Two normal subjects were found, each with a different 9 bp deletion in the same region.

for association studies. Among the rare variants, we identified three deletions, one insertion, one duplication, and five single nucleotide substitutions (table 1, fig 1B). To determine if these variants were low frequency polymorphisms, we screened over 100 normal controls (>200 chromosomes). Only one of the infrequent changes, the C→A transition at nucleotide 886, was identified in one of the control samples, giving a frequency of ~0.5%. Interestingly, two control samples had deletions similar to the del(9 bp) at nucleotides 1404-1412. All three 9 bp deletions occurred in a tract of 15 alanines. However, when the sequences were aligned, it was possible to show that each 9 bp deletion occurred in a slightly different region of the alanine repeat. Thus, although the nucleotide sequence in these variants is not identical, the resulting change in the protein function is unlikely to be associated with alteration in the protein function. None of the remaining eight rare variants were found in the control samples.

For the rare variants, only two of the five single base pair substitutions are predicted to cause a change in the encoded amino acid. The 1049 G→A changes the Gly at 350 to Asp (G350D) in a patient with AAA. However, the sequence of the mouse at this position is not conserved (the mouse has a single Gly whereas the human has two). The second missense mutation is a C→T transition at nucleotide 1187, resulting in Pro→Leu at amino acid residue 396 (P396L) in a patient with AAA. A proline is found at this position in the mouse. Unfortunately, DNA from the parents of these two patients is not available to clarify whether the changes are de novo. Of the three remaining single base pair substitutions, one is the C→A transversion at position 886 discussed above, one is a silent change (576C→T; R192R), and the last a C→T transition at 39 bp 5’ of the presumed start codon. The patient with the −39C→T (p23 in table 1) also has a 15 bp deletion (1334-1348del) that results in an in frame deletion of five amino acids (GYHPH) starting at codon 445. The five amino acids are highly conserved with 100% identity to mouse and 40% to Drosophila TBX1 sequence (CAA76529). Analysis of samples from this patient’s parents and his unaffected brother showed that he had inherited the 15 bp deletion from his father. The −39C→T change was found in his unaffected mother. Sequence analysis shows that his normal brother has the same two changes, the −39C→T and the del(15 bp). Patient 23 had clinical findings that are consistent with DGS, including absent thymus with partial T cell insufficiency, low parathyroid levels, minor ear anomalies and complex conotruncal heart defect (double outlet right ventricle with pulmonary atresia, VSD, and ASD).

One of the remaining three variants was the result of a small insertion within exon 9C in a patient (p22 in table 1) with a phenotype suggestive of VCFS (cleft palate, dysmorphic facial features, autism with moderate-severe mental retardation, and developmental delay). The patient has a 3 bp insertion (1371-1372insCAC), resulting in an additional histidine within a run of five. The mouse has five histidines in this region but the sequence immediately flanking the histidines is somewhat variable. Family analysis indicated that the 1371-1372insCAC was transmitted to the patient from her unaffected mother. This variant was not found in 139 control samples. In a second patient (p46), we identified a 30 bp in frame duplication (1399-1427dup30), resulting in a tract of 25 alanines instead of the 15 normally seen in this position. The comparative region of the mouse sequence has 10 alanines. DNA from the patient’s family is not available and so we are not able to determine whether this duplication is de novo. Analysis of 139 controls did not detect this change in any samples. Finally, we found an 8 bp deletion in a patient with the diagnosis of DGS. The deletion occurs between nucleotides 1274-1281 and is predicted to cause a frameshift at amino acid 425, resulting in the replacement of the
Molecular studies have shown that the majority of patients with DiGeorge syndrome/velocardiofacial syndrome (DGS/VCFS) have a common pathogenesis, deletion of chromosomal region 22q11.2. Recent studies using mouse models suggest that TBX1, a gene localized in the common deleted region, is a strong candidate for many of the key features of 22q11 deletion syndromes.

To investigate the role of TBX1 in DGS/VCFS further, we performed mutation analysis of this gene in 105 patients, who did not have deletions or rearrangements of 22q11.2.12–16 By FISH and Southern blot analysis, but had clinical features consistent with DGS/VCFS or had cardiovascular defects associated with deletion of 22q11.

We identified eight common polymorphisms and 10 rare variants. Among the rare variants, three are deletions, one is an insertion, one is a duplication, and five are single nucleotide substitutions. One of the rare variants results in a frameshift. In several cases, samples from the parents were available for study and analysis suggests that most rare variants are transmitted to the patient from an unaffected parent.

Although the experiments in the mouse suggest that TBX1 may be the key gene for the majority of features of DGS/VCFS, our results indicate that isolated loss of TBX1 function is unlikely to be a common cause of DGS/VCFS or the associated cardiovascular defects.
In this study, only the exons and splice junctions of \( \text{TBX1} \) were screened. Thus, mutations in the regulatory regions would be missed. In addition to the promoter, there are at least six blocks of sequence greater than 130 bp in the intronic regions of \( \text{TBX1} \) that share 75-93\% of similarity to the corresponding mouse sequence (fig 1C). These highly conserved regions may harbour regulatory elements. Mutation screening of these control regions will provide more complete analysis. Finally, although \( \text{TBX1} \) may be a critical gene for the aetiology of DGS/VCFS, the possibility that haploinsufficiency of more than one gene in 22q11.2 is necessary for the expression of the DGS/VCFS phenotype cannot be excluded at present.

We thank Joanne Campanile, Erica Tompkins, Stacy Harris, and Dan Gray for technical assistance. We would like to thank Dr V. Papasodros for critical reading of the manuscript and helpful discussions. We would also like to thank the numerous clinicians who provided samples for these studies. This research was supported by NIH grants DC02027 and HL62174.