*Am. J. Hum. Genet. 64:1550–1562, 1999*

# The Spectrum of Mutations in *TBX3:* Genotype/Phenotype Relationship in Ulnar-Mammary Syndrome

M. Bamshad,[1,3] T. Le,[2] W. S. Watkins,[2] M. E. Dixon,[2] B. E. Kramer,[2] A. D. Roeder,[2] J. C. Carey,[1] S. Root,[4] A. Schinzel,[5] L. Van Maldergem,[6] R. J. M. Gardner,[7] R. C. Lin,[8] C. E. Seidman,[9] J. G. Seidman,[8] R. Wallerstein,[10] E. Moran,[10,11] R. Sutphen,[12] C. E. Campbell,[13] and L. B. Jorde[3]

Departments of [1]Pediatrics and [2]Human Genetics, University of Utah Health Sciences Center, and [3]Shriners Hospitals for Children, Intermountain Unit, Salt Lake City, Utah; [4]Division of Genetics and Dysmorphology, University of New Mexico, Albequerque; [5]Institute of Medical Genetics, University of Zurich, Zurich; [6]Center de Génétique Humaine, Institut de Pathologie et de Génétique, Loverval, Belgium; [7]Victorian Clinical Genetics Service, Royal Children's Hospital, Melbourne; [8]Department of Genetics, Howard Hughes Medical Institute and Harvard Medical School, and [9]Howard Hughes Medical Institute and Cardiovascular Division, Brigham and Women's Hospital, Boston; [10]Human Genetics Program, Department of Pediatrics, New York University Medical Center, and [11]Center for Nervous and Developmental Disorders, Hospital for Joint Diseases, Orthopedic Institute, New York; [12]Department of Pediatrics, South Florida College of Medicine, Tampa; and [13]Department of Cancer Biology, The Lerner Research Institute, Cleveland Clinic Foundation, Cleveland

## Summary

**Ulnar-mammary syndrome (UMS) is a pleiotropic disorder affecting limb, apocrine-gland, tooth, hair, and genital development. Mutations that disrupt the DNA-binding domain of the *T-box* gene, *TBX3*, have been demonstrated to cause UMS. However, the 3′ terminus of the open reading frame (ORF) of *TBX3* was not identified, and mutations were detected in only two families with UMS. Furthermore, no substantial homology outside the T-box was found among *TBX3* and its orthologues. The subsequent cloning of new *TBX3* cDNAs allowed us to complete the characterization of *TBX3* and to identify alternatively transcribed *TBX3* transcripts, including one that interrupts the T-box. The complete ORF of *TBX3* is predicted to encode a 723-residue protein, of which 255 amino acids are encoded by newly identified exons. Comparison of other *T-box* genes to *TBX3* indicates regions of substantial homology outside the DNA-binding domain. Novel mutations have been found in all of eight newly reported families with UMS, including five mutations downstream of the region encoding the T-box. This suggests that a domain(s) outside the T-box is highly conserved and important for the function of *TBX3*. We found no obvious phenotypic differences between those who have missense mutations and those who have deletions or frameshifts.**

Received November 9, 1998; accepted for publication April 19, 1999; electronically published May 7, 1999.

Address for correspondence and reprints: Dr. Mike Bamshad, Department of Pediatrics, Eccles Institute of Human Genetics, 15 North 2030 East, Room 2100, University of Utah, Salt Lake City, UT 84112-5330. E-mail: mike@genetics.utah.edu


0002-9297/99/6406-0008$02.00

## Introduction

*T-box* genes are members of a rapidly growing and highly conserved family of transcription factors that share a region of homology to the DNA-binding domain (T-box) of the mouse *Brachyury* (or *T*) gene product (Herrmann et al. 1990; Bollag et al. 1994; Papaioannou and Silver 1998). Most *T-box* genes have been found through experiments designed to identify genes that have specific activities in embryonic development or that cause developmental defects (Bollag et al. 1994). During embryogenesis the spatial and temporal expression patterns of *T-box* genes are unique although overlapping. This indicates that they are differentially regulated during development (Chapman et al. 1996). Phylogenetic analyses suggest that the origin of the *T-box* gene family predates the divergence of arthropods and chordates from a common ancestor >600 million years ago (Agulnik et al. 1996; Ruvinsky and Silver 1997; Wattler et al. 1998). Thus, *T-box* genes are likely to play critical roles in the development of all animal species.

The *T* gene and its orthologues are immediate-early-response genes (i.e., transcription is directly activated by a signal-transduction pathway) that appear to control the specification and differentiation of mesoderm subpopulations in the developing embryo (Smith et al. 1991). Consequently, many of the early studies of *T-box* genes concentrated on understanding their role in the promotion of mesoderm formation (O'Reilly et al. 1995; Schulte-Merker and Smith 1995). Recently, more effort has shifted toward explaining how *T-box* genes act to control morphogenesis and organogenesis. For example, *Tbx5* and *Tbx4* are exclusively expressed in the vertebrate forelimb and hindlimb, respectively (Gibson-Brown et al. 1996, 1998; Simon et al. 1997). This suggests that *T-box* genes may specify limb identity in tet-

## Table 1

***TBX3* Mutations and Clinical Findings in Families with Ulnar-Mammary Syndrome**

| *TBX3* Mutation and Individual[a] | Clinical Finding(s)[b] | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Apocrine | | Breast | | | |
| | Limb[c] | Axillary Hair | Perspiration | Lactation[d] | Breasts | Endocrine (Puberty) | Other |
| S343ter: | | | | | | | |
| 1 | Stiff digit (B 5) | Absent | Reduced | Reduced | Hypoplasia (B) | Normal | Anteposition of the anus |
| 2 | Stiff digit (B 5) | Prepubertal | Prepubertal | N/A | Normal | Normal | Cryptorchidism, imperforate anus |
| 1301G…G1302: | | | | | | | |
| 1 | Extra digit (B 5) | Reduced | Reduced | N/A | Deceased | Deceased | |
| 2 | Extra digit (B 5) | Absent | Reduced | N/A | Normal | Delayed | Extra digit (L foot) |
| 3 | Absent digit (L 5) | Absent | Reduced | Normal | Hypoplasia (U) | Normal | Small uterus |
| 4 | Absent ulna (B), absent digits (L 3–5, R 5), fused digits (R 3–4) | Prepubertal | Prepubertal | Nulliparous | Normal | Prepubertal | |
| IVS2+1G→C: | | | | | | | |
| 1 | Extra digit (B 5) | Absent | Absent | Normal | Hypoplasia (U) | Normal | |
| 2 | Extra digit (U 5) | Prepubertal | Prepubertal | Nulliparous | Normal | Prepubertal | Kyphosis, scoliosis, recto-perineal fistula |
| Y149S: | | | | | | | |
| 1 | Duplicated nail (B 5), tapered digit (B 5) | Reduced | Reduced | N/A | Hypoplasia (B) | Delayed | Short stature |
| 1857delC: | | | | | | | |
| 1 | Extra digit (B 5), stiff digit (R 5) | Normal | Normal | N/A | Normal | Normal | |
| 2 | Extra digit (B 5), stiff digit (5) | Absent | Reduced | None | Hypoplasia (B), absent areola (B) | Normal | |
| 3 | Stiff digit (R 5) | Absent | Absent | Normal | Absent areola (B) | Normal | Single kidney, septate uterus |
| 4 | Normal | Absent | Reduced | None | Absent areola (B) | Normal | |
| 5 | Extra digit (L 5) | Normal | Normal | Normal | Normal | Normal | |
| 6 | Normal | Absent | Absent | Nulliparous | Absent areola (B) | Normal | |
| 7 | Absent digits (L 4–5), stiff digit (R 5) | Absent | Reduced | N/A | Normal | Normal | Cryptorchidism |
| 8 | Extra digit (L 5) | Absent | Absent | None | Hypoplasia (B), absent areola (B) | Normal | |
| 9 | Normal | Absent | Absent | None | Absent areola (B) | Normal | |
| 10 | Normal | Absent | Absent | None | Absent areola (B) | Normal | |
| 11 | Extra digit (L 5) | Prepubertal | Prepubertal | Nulliparous | Absent areola (B) | Prepubertal | Absent teeth (incisors) |
| 12 | Absent ulna (L), absent digits (L 3–5), fused digit (R 5) | Prepubertal | Prepubertal | N/A | Absent areola (B) | Prepubertal | Anteposition of the anus |
| 227delT: | | | | | | | |
| 1 | Deceased | Absent | Absent | N/A | Normal | Normal | Ectopic canine (R) |
| 2 | Stiff digit (L 5) | Absent | Absent | None | Hypoplastic areola (B) | Normal | Obesity, ectopic canine (R) |
| 3 | Stiff digit (B 5) | Absent | Absent | Nulliparous | Normal | Normal | Ectopic canine (B) |
| 4 | Absent digit (R 4), fused digits (R 2–3), stiff digit (R 5) | Absent | Absent | N/A | Hypoplastic areola (B) | Delayed | Ectopic canine (B), obesity, inguinal hernia, micro-penis, hypoplastic testis (R), shawl scrotum |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| IVS6+2T→A: | | | | | | | |
| 1 | Reduced digit (B 5) | Reduced | Absent | N/A | Hypoplastic areola (B) | Delayed | Anal stenosis, subglottic cartilaginous webs, cryptorchidism (L), micropenis, hypoplastic testis (B) |
| 2 | Absent digits (L 5, R 4–5) | Absent | Absent | N/A | Hypoplastic areola (B) | Delayed | Pyloric stenosis, anal stenosis, obesity, cryptorchidism, micropenis |
| 3 | Reduced digit (B 5) | Reduced | Reduced | N/A | Hypoplastic areola (B) | Delayed | Cryptorchidism (B), micropenis |
| 4 | Reduced digit (B 5) | Absent | Absent | N/A | Hypoplastic areola (B) | Delayed | Inguinal hernia (B), scoliosis, obesity, absent testis (R), micropenis |
| L143P: | | | | | | | |
| 1 | Deceased | Absent | Absent | N/A | Deceased | Delayed | Delayed growth |
| 2 | Reduced humerus (B), absent ulna (B), absent digits (L 4–5, R 3–5) | Absent | Absent | N/A | Hypoplastic areola (B) | Delayed | Delayed growth, cryptorchidism (B) |
| 3 | Stiff digit (B 5) | Absent | Absent | N/A | Hypoplastic areola (B) | Delayed | Scoliosis, delayed growth, cryptorchidism (B) |
| 4 | Reduced humerus (B), absent ulna (B), absent digits (L 4–5, R 3–5) | Absent | Absent | N/A | Hypoplastic areola (B) | Delayed | Cryptorchidism (B) |
| 5 | Reduced humerus (L), absent ulna (L), absent digits (L 3–5, R 5) | Prepubertal | Prepubertal | Nulliparous | Hypoplastic areola (B) | Prepubertal | |
| 465T…T466: | | | | | | | |
| 1 | Stiff digit (R 5), reduced digit (L 5) | Absent | Absent | N/A | Hypoplastic areola (B) | Delayed | Pectus carinatum |
| 2 | Normal | Absent | Absent | N/A | Hypoplastic areola (B) | Delayed | Pectus carinatum |
| 3 | Reduced digit (B 5) | Absent | Absent | N/A | Hypoplastic areola (B) | Delayed | Pectus carinatum, duplicated teeth, hypospadias |
| 4 | Stiff digit (B 5) | Prepubertal | Prepubertal | N/A | Hypoplastic areola (B) | Prepubertal | |
| 5 | Stiff digit (U 5) | Prepubertal | Prepubertal | N/A | Hypoplastic areola (B) | Prepubertal | |
| 6 | Stiff digit (R 5) | Prepubertal | Prepubertal | Nulliparous | Hypoplastic areola (B) | Prepubertal | Anteposition of the anus |

[a] The clinical findings in families with Q360ter, 227delC, and IVS6+2T→A have been published elsewhere (Schinzel et al. 1987; Franceschini et al. 1992; Bamshad et al. 1996).
[b] U = unilateral; B = bilateral; L = left; R = right.
[c] Digits are numbered "1"–"5" from anterior (thumb) to posterior (little finger).
[d] N/A = not applicable.

rapods (Logan et al. 1998; Ohuchi et al. 1998; Logan and Tabin 1999).

Mutations in *T-box* genes have been associated with defects of development in *Drosophila melanogaster* (Kispert et al. 1994), *D. rerio* (Schulte-Merker et al. 1994), and *Mus musculus* (Herrmann et al. 1990). In humans, mutations in *TBX3* cause limb, apocrine-gland, hair, genital, and dental defects in ulnar-mammary syndrome (UMS [MIM 181450]; Bamshad et al. 1997), and mutations in *TBX5* cause limb and heart defects in Holt-Oram syndrome (HOS [MIM 142900]; Basson et al. 1997, 1999; Li et al. 1997; Bamshad et al. 1999). Thus, the medical importance of *T-box* gene mutations as causes of human birth defects is well established (Bamshad et al. 1999).

Direct evidence of a role for *TBX3* and *TBX5* in control of the anterior/posterior axis of the tetrapod forelimb comes from the analysis of human birth defects of the limb. Defects of the anterior elements of the upper limb are the most common abnormalities observed in individuals with *TBX5* mutations (Basson et al. 1994, 1997; Li et al. 1997), whereas mutations in *TBX3* cause deficiencies or duplications of the posterior elements of the upper limb (Bamshad et al. 1996, 1997; see fig. 1*A*). The pattern of limb defects observed in UMS is concordant with the finding that *Tbx3* is expressed by cells of the posterior region of the chick wing that are fated to become posterior skeletal elements (Gibson-Brown et al. 1998). Thus, analogous to the role of *T* in the specification of axial mesodermal subpopulations, *TBX3* may specify posterior mesodermal subpopulations in the tetrapod forelimb.

*TBX3* may also play a role in the specification of the dorsal/ventral axis of the forelimb. This inference is based on the observation that the ventral surface of the posterior digits of the upper limb are dorsalized in some individuals with UMS, as manifested by both the presence of a nail and the lack of normal dermatoglyphic patterning (fig. 1*B*). Thus, mutations in *TBX3* can disturb the development of all three axes (i.e., proximal/distal, anterior/posterior, and dorsal/ventral) of the limb.

Here we report the characterization of *TBX3* mutations in 75 individuals with UMS who are members of eight newly identified families and two families reported previously (Bamshad et al. 1997). We emphasize that half the mutations are found downstream of exons encoding the T-box domain, suggesting the existence of an additional functional domain(s) in *TBX3*. Relationships between the types of mutations in *TBX3* and the distribution and severity of developmental defects observed in patients with UMS are examined. Defining these relationships is necessary to allow us to further understand both the role of *TBX3* in human development and the phenotypic consequences of specific *TBX3* mutations. Furthermore, a clear understanding of the mutation spectrum and the genotype/phenotype relationship will permit molecular testing and will facilitate genetic counseling of families with limb defects.

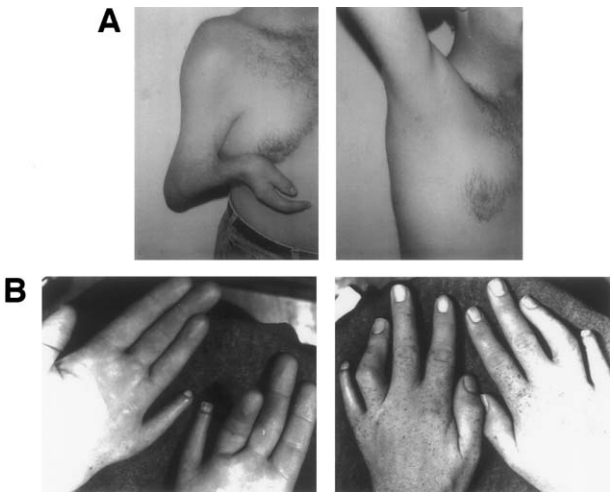


**Figure 1** Limb defects observed in individuals with a *TBX3* (L143P) mutation causing UMS. *A*, Photograph of right upper limb and chest (*left*) and axilla (*right*) of an adult male with UMS. Note the shortened forearm and absence of the 3d–5th digits. Axillary hair is absent, and the amount of hair on the anterior-lateral chest is reduced. Additionally, the right areola is hypoplastic. *B*, View of ventral surface of right and left hands of an adult male with UMS. Note the dorsalized ventral surface of the right and left 5th digits, including a duplication of the nailbed of these digits.

## Patients and Methods

### *Clinical Evaluation*

All studies were performed with the approval of the Institutional Review Board of the University of Utah and the General Counsel of the Shriners Hospitals for Children. After informed consent was obtained, each participant was evaluated by history and physical examination and/or review of medical records. Radiography was performed on selected individuals. The diagnosis of UMS was based on criteria described elsewhere (Bamshad et al. 1996). At least one affected individual in each family had limb, apocrine-gland, dental, and/or axillary-hair defects (table 1). Affected individuals in 10 unrelated families with UMS were selected for mutation analysis of *TBX3*. Vertical transmission of UMS, with affected individuals in two or more generations, was documented in 9 of 10 families.

### *Analysis of* TBX3

Isolation and analysis of a P1 clone containing *TBX3* has been reported previously (Bamshad et al. 1997). *TBX3* cDNAs were isolated from a 15-wk-old human fetal kidney cDNA library, a human placental cDNA

library, and an adult pancreatic cDNA library. The *TBX3* exon 1/7 and exon 2a alternative transcripts were obtained from a Soares human senescent fibroblast NbHSF DNA library and a Soares human infant brain 1NIB cDNA library, respectively. Exon/intron boundaries corresponding to the *TBX3* cDNA sequence were defined by comparison of cDNA and genomic sequences (table 2 and fig. 2). The BLAST algorithm (Altschul et al. 1990) was used to search the expressed sequence tag (EST) database of the National Center for Biotechnology Information, for sequences homologous to *TBX3*.To extend the *TBX3* contig, EST sequences were grouped into overlapping sets.

To isolate *Tbx3,* murine cDNA was synthesized by means of a Superscript Preamplification System (GIBCOBRL) and ~2 μg of total RNA from adult mouse kidney. Since we wished to extend the partial sequence for *Tbx3* available from GenBank to include the entire T-box domain, a number of primer combinations were tested. The primers that gave us the desired murine *Tbx3* product were derived from the 5′ end of the human *TBX3* DNA-binding domain (5′-CCTGACGCTGCCTCCCAACGG-3′) and from the 3′ end of murine *Tbx2* (5′-AGGCAGTGACAGCGATGAA-3′). The PCR product was cloned into pCRII-topo (Invitrogen) and was sequenced on an ABI 377 automated sequencer (Applied Biosystems). The sequence agreed with the partial *Tbx3* sequence found in GenBank and extended this sequence as shown in figure 2.

### *Northern Blot Analysis*

A human Multi-Tissue northern blot (Clontech) was screened with [$^{32}$P]-labeled PCR products that included either exon 1 or exon 7. The filter was hybridized and washed under high-stringency conditions, in accordance with the manufacturer's recommendations. The quantity of RNA from each tissue was normalized against the expression of eight housekeeping genes. Volume-based pixel densities for each dot were quantified by means of a BioRad imager and MULTIANALYST software (BioRad).

### *RNAse Protection Assay*

A PCR product containing part of exon 1 through to exon 5, including the 60-nt alternative exon (exon 2a), was cloned into pCRII-topo (Invitrogen). The template DNA was linearized at a *Bsr*GI site within exon 2, and an ~460-nt riboprobe was synthesized by means of T7 polymerase. The riboprobe includes 428 nt of *TBX3* sequence and ~40 nt of sequence derived from the vector. Six micrograms of total RNA from normal mammary epithelial cells, fetal kidney, or three breast carcinoma–derived cell lines were hybridized with 500,000 cpm of *TBX3* plus 5,000 cpm of β-actin probe synthesized at 1/100 the specific activity of the *TBX3* probe.

**Table 2**

**DNA Sequence of *TBX3* Exon Boundaries**

| Exon/ Intron | Sequence[a] |
|---|---|
| 1 | ATGAGCCTCT....AGTCGGGAAGgtaagcagtg |
| 2 | cgttttatagGCGAATGTTT....ACATGGATTTgtaagtttca |
| 3 | gcccctgcagACTATATTGA....GAATGATAAGgtcanctcaa |
| 4 | cttccaatagATAACCCAGT....GAGAAAAAAGgtgagttgaa |
| 5 | tctgccacagAAAACAGCTC....AACCTCAAAGgtaaaccatg |
| 6 | cctgtggtagATTTATGTCC....GGCCTCTCAGgtatggatcc |
| 7 | ctcaccccagGGCCTGGCCA....GTCCCCGTAG |

[a] Exonic regions are shown as uppercase, and intronic regions are shown as lowercase.

After hybridization, the reactions were processed according to the instructions provided by Ambion's RNase protection kit and were subjected to electrophoresis on 5% denaturing polyacrylamide gels.

### *Mutation Analysis of* TBX3

Genomic DNA was prepared from peripheral lymphocytes and/or Epstein-Barr virus–transformed lymphoblastoid cell lines derived as previously described. Exons 1–5 of *TBX3* were PCR-amplified by means of previously published exon-flanking primers (Bamshad et al. 1997). Exons 6 and 7 were PCR-amplified by means of 5′-GGAGATAACGCCCTTCTGCCTTGG-3′ (forward)/5′-TTACAGCTACTAGGCCAAAGGGATC-3′ (reverse) and 5′-AGAGGAGAGGGATGAGATAAGCTC-3′ (forward) and 5′-CAACTGCAAAAGGAAGGGCTAACG-3′ (reverse), respectively. Genomic DNA sequences were amplified in 1 × buffer (10 mM Tris pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 10% dimethylsulfoxide) by 20 ng of template genomic DNA, 50 μM of each dNTP, 20 pmol of each primer, and 1 unit of *Taq* polymerase, in a total reaction volume of 25 μl. Samples were cycled 30 times in a Perkin-Elmer 9600 PCR machine using a standard three-step PCR profile with an annealing temperature of 60°C. PCR products were purified on a 2% NuSieve gel and were sequenced by means of either ABI Dye-primer or dRhodamine sequencing reagents. Sequenced products were loaded on an ABI 377 automated sequencer and were analyzed by SEQUENCHER software (Genecodes). The forward and reverse strand of every exon, including flanking splice-recognition sequences, was PCR-amplified and sequenced twice in each direction. For all missense and nonsense mutations, the presence of a mutation was confirmed by restriction digestion, and genomic DNA samples from 100 individuals representing 200 control chromosomes were screened.

To confirm the presence of mutations that did not alter a restriction site, PCR-amplified products from individuals with insertion mutations were ligated into a pAMP1 plasmid (BRL), according to the manufacturer's rec-

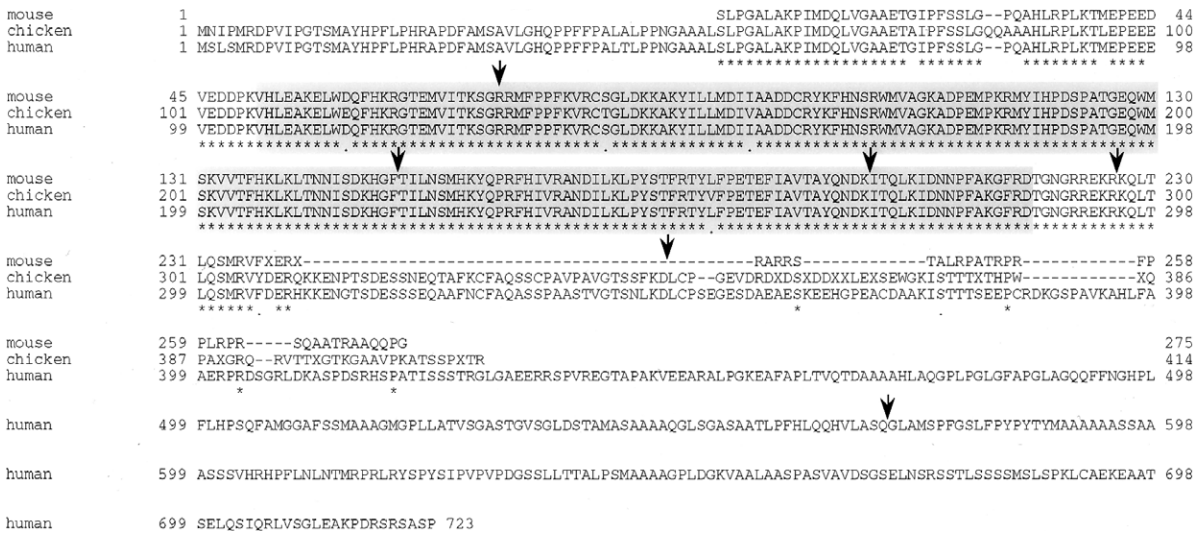


**Figure 2** Deduced amino acid sequence of partial Tbx3 protein, in mouse and chicken, and of complete TBX3 protein, in human. The conserved T-box region is shaded. Asterisks (*) denote identical amino acids, and dots (.) denote conserved amino acid changes. Locations of intron/exon boundaries in TBX3 protein are indicated by arrows (↓). The deduced amino acid residues encoded by exon 2a (TLAFPSDHA-TWQGNYSFGTQ) are not shown.

ommendations. Ligation mixes were used to transform DH5$\alpha$ cells (Stratagene), and transformants were selected on ampicillin plates. From each individual, 10 transformed colonies were picked, and DNA was prepared by means of an alkaline/SDS miniprep protocol. DNA from each clone was bidirectionally sequenced by previously described methods using ABI dRhodamine sequencing reagents.

## Results

### *Construction of the Full-Length Open Reading Frame (ORF) of* TBX3

We previously have reported the partial cDNA sequence of *TBX3* (Bamshad et al. 1997). However, the 3′ boundary of the ORF of *TBX3* was not determined, and, as a result, mutations were identified in only two of three families studied. We characterized both the 3′ boundary of the ORF of *TBX3* and the 3′ UTR, to help us to identify additional disease-causing mutations in *TBX3*. The 3′ UTR of *TBX3* extends 1,587 bp beyond the termination codon in exon 7. The transcription-termination site is demarcated by a consensus polyadenylation sequence (AATAAA) and a poly-A tail (data not shown).

*TBX3* is composed of at least seven exons (fig. 3) with an ORF of 2,172 bp that is predicted to encode a protein containing 723 amino acid residues. The predicted DNA-binding domain of TBX3 protein is 98% identical to that of murine and chick Tbx3 protein (fig. 2). The regions 5′ and 3′ of the T-box in human TBX3 protein also exhibit a high degree of homology to the human TBX2 protein (fig. 4). Compared with TBX2 protein, there are two domains of ~75–100 amino acid residues, encoded by portions of exons 6 and 7, that exhibit nearly 70% identity to TBX3 protein.

### *Expression Studies and Identification of Alternatively Spliced* TBX3 *Transcripts*

Northern blot analyses reveal a major *TBX3* transcript of ~5.2 kb in adult human tissues (Bamshad et al. 1997). At least three additional transcripts can be observed (fig. 5*A*). Alternative transcription-start sites, polyadenylation sites, or splicing could generate these *TBX3* transcripts of different lengths. Quantification of *TBX3* RNA from human fetal and adult tissues (fig. 5*B*) reveals abundant signal in placenta, adrenal gland, thyroid, prostate, breast, bladder, uterus, and liver. A weaker signal is detected in RNA isolated from ovary, lung, salivary gland, kidney, small bowel, and skeletal muscle. In fetal tissue, a signal is detected in lung, kidney, heart, liver, and spleen. In the adult brain, the pituitary gland is the only tissue in which a signal is observed. *TBX3* expression was not detected in mRNA isolated from lymphoblastoid cell lines (data not shown).

An alternative transcript of *TBX3* contains an additional 60 bp inserted between exons 2 and 3 (fig. 5*C*). Examination of the intronic region between exons 2 and 3 reveals a 60-bp exon (exon 2a in fig. 3) bracketed by consensus splice-acceptor/-donor (ag/gt) sequences. The region of *TBX3* into which exon 2a is spliced encodes the highly conserved DNA-binding domain of TBX3 protein. Insertion of a novel 20 amino acid residues into the middle of the T-box is likely to alter its DNA-binding

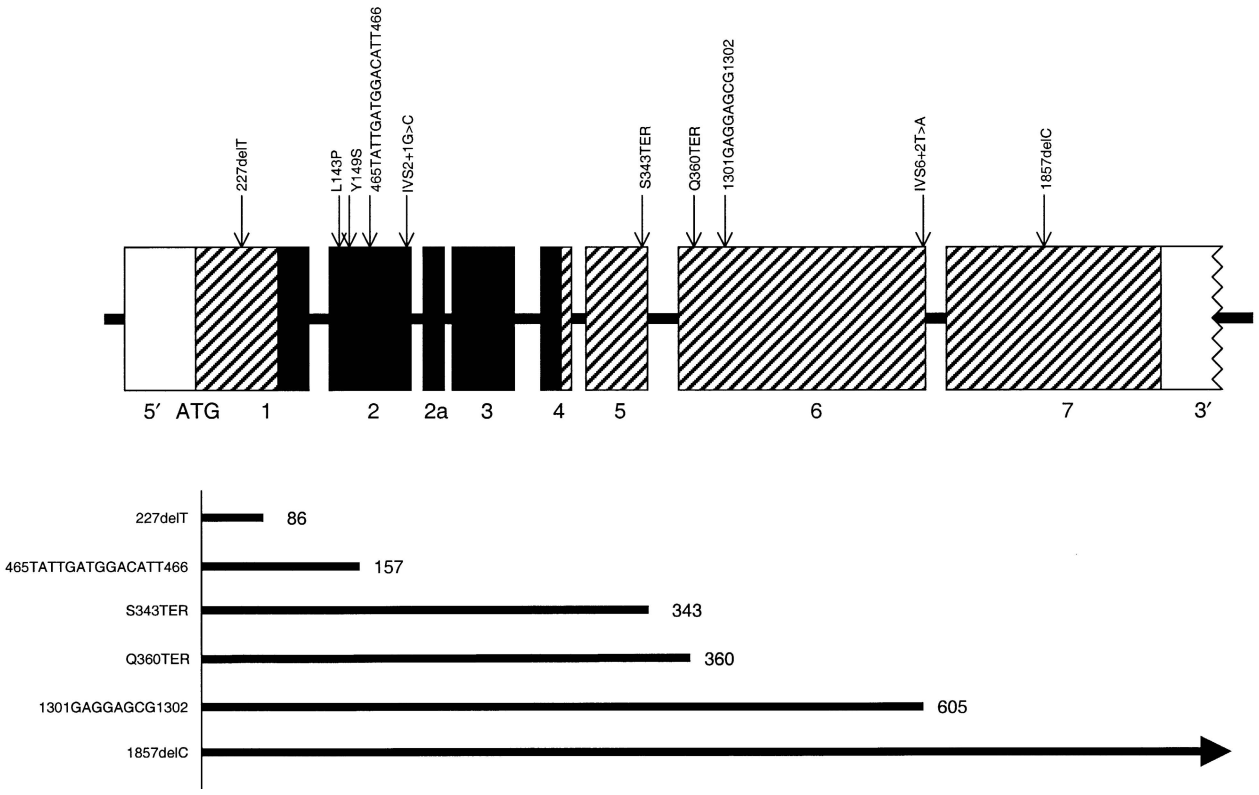


**Figure 3** Genomic structure of exons encoding the ORF of *TBX3*. *TBX3* is composed of at least seven exons (Bamshad et al. 1997) that contain untranslated sequence (*unblackened boxes*), highly conserved T-box sequences (*blackened boxes*), and protein-encoding sequence (*striped boxes*). Exon 2a is alternatively transcribed. Locations of *TBX3* mutations found in 10 families with UMS are indicated by arrows (↓). Predicted polypeptides encoded by mutant *TBX3* alleles are shown below the genomic structure of *TBX3*.

properties. Another alternative transcript of *TBX3* splices the 5′ sequence of exon 1 to exon 7 in frame and eliminates the T-box. This is analogous to the few transcripts lacking a homeobox that have been reported for several *Hox* genes (Chariot et al. 1995). Furthermore, if this *TBX3* transcript is translated, the predicted isoform would include only the highly homologous regions conserved with *TBX2* protein.

### *Identified Mutations*

Mutation analysis was conducted on at least two affected individuals in each multiplex family. DNA sequence changes were found in the affected individuals of all 10 families (fig. 6). Two individuals were found to have single-base deletions (227delT and 1857delC) that produced frameshift mutations. Three individuals had a 1-bp change that produced either a missense (L143P and Y149S) or nonsense (Q360TER) mutation. One individual had a 2-bp change (S343TER) resulting in a nonsense mutation. Presumably, this was produced by an initial silent substitution, G→A, followed by a pathogenic mutation, C→A. Two individuals had splice-site mutations (IVS2+1G→C and IVS6+2T→A) that are predicted to produce a truncated protein product. Frameshift mutations caused by insertions of small duplications (465TATTGATGGACATT466 and 1301GAGGAGCG1302) were identified in two individuals. Each specific mutation was identified in only one kindred. Sequence and/or restriction analysis confirmed that each mutation segregated only with affected individuals. None of the changes predicted to cause missense or nonsense mutations were found in 200 control chromosomes. Additional sequence variants were identified within the intronic regions of affected and unaffected individuals and were classified as nonpathogenic polymorphisms (data not shown).

## Discussion

### *Structure and Expression Pattern of* TBX3

Here we have reported the characterization of the complete ORF of *TBX3,* its 3′ UTR with a transcription termination site, and a *TBX3* cDNA containing a partial ORF that, we suggest, encodes a novel isoform of TBX3 protein. Because the Northern blots indicate that the predominant *TBX3* transcript is ~5.2 kb in length (fig.

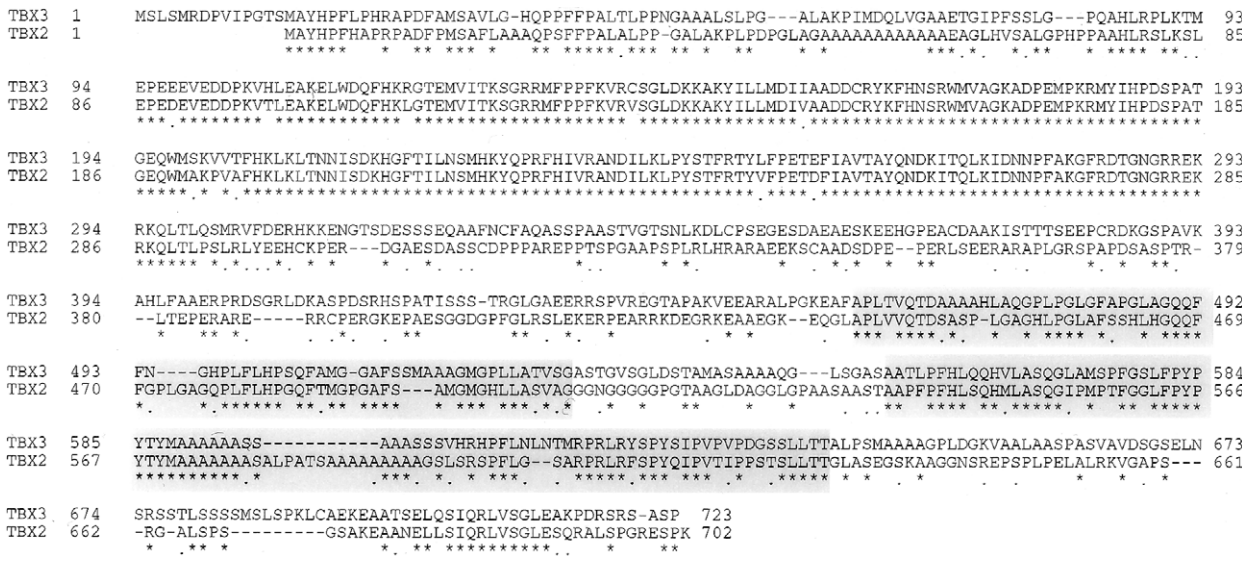


**Figure 4** Amino acid alignment of human TBX3 and TBX2 proteins. Asterisks (*) denote identical amino acids, and dots (.) denote conserved amino acid changes. The conserved T-box region is shaded.

*5A*), the full-length transcript of *TBX3* is likely to include a long 5′ UTR that remains to be characterized. Comparisons with chick and mouse *Tbx3* reveal that the nucleotide and amino acid identity of *TBX3* and TBX3 protein, respectively, are highly conserved (figs. 2 and 4).

Among different *T-box* gene–family members, the size range of the T-box domain is ~180–190 amino acids, and the amino acid identity is highly conserved. The T-box may be located at various positions within each protein, and sequence similarities outside the T-box are observed only among closely related members of the *T-box* gene family (Papaioannou and Silver 1998). Thus, it appears that the DNA-binding activity between *T-box*–gene–family members has been retained while the domains that may be involved in protein-protein interactions have diversified. A comparison of TBX2 and TBX3 proteins indicates that, in addition to their T-boxes, which are ~95% identical, the amino-terminal domains of the two proteins are also similar, sharing ~70% identity over two regions composed of ~70 and ~100 amino acids, respectively. The significance of these regions of shared homology is not known, but the similarity suggests that TBX2 and TBX3 may share functional domains in addition to their DNA-binding domain.

Analysis of the crystal structure of the DNA-binding domain of *Xenopus* T protein reveals that T is bound as a homodimer, interacting with the major and minor grooves of DNA (Müller and Herrmann 1997). At least two mutations in *TBX3* disrupt the DNA-binding domain and produce a protein fragment that presumably has reduced, if any, DNA-binding activity. Two missense mutations in exon 2 produce nonconservative amino acid substitutions near a critical dimerization site in the DNA-binding domain of *Xenopus* T protein. These two sites are highly conserved across *T-box* genes. However, five mutations in *TBX3* are located 3′ of the region encoding the T-box. This suggests that, although these mutant TXB3 proteins may bind DNA, their function is substantially impaired. The observation that many of the mutations in *TBX3* do not disrupt the T-box underscores the existence of an additional functional domain(s) in TBX3 protein. However, it remains to be demonstrated that mutant *TBX3* transcripts encode a truncated protein product.

*TBX3* is widely expressed in a variety of tissues and organs that are not affected in individuals with UMS; in other words, development of most tissues in which *TBX3* is expressed is normal in individuals with UMS. This suggests that different tissues may require quantitatively different levels of normal TBX3 protein. Alternatively, other genes, including *T-box* genes, with expression domains overlapping *TBX3* (e.g., *TBX2*), may be able to compensate for reduced levels of normal TBX3 protein in certain tissues and organs. Such functional redundancy is also observed for other families of transcription factors that play prominent roles in limb development and that cause human birth-defect syndromes (e.g., *HOX* genes).

*TBX3* is widely expressed in adult tissues, although *TBX3*'s role beyond organogenesis is largely unexplored. Because delayed puberty is a consistent feature of individuals with UMS, it has been suggested that

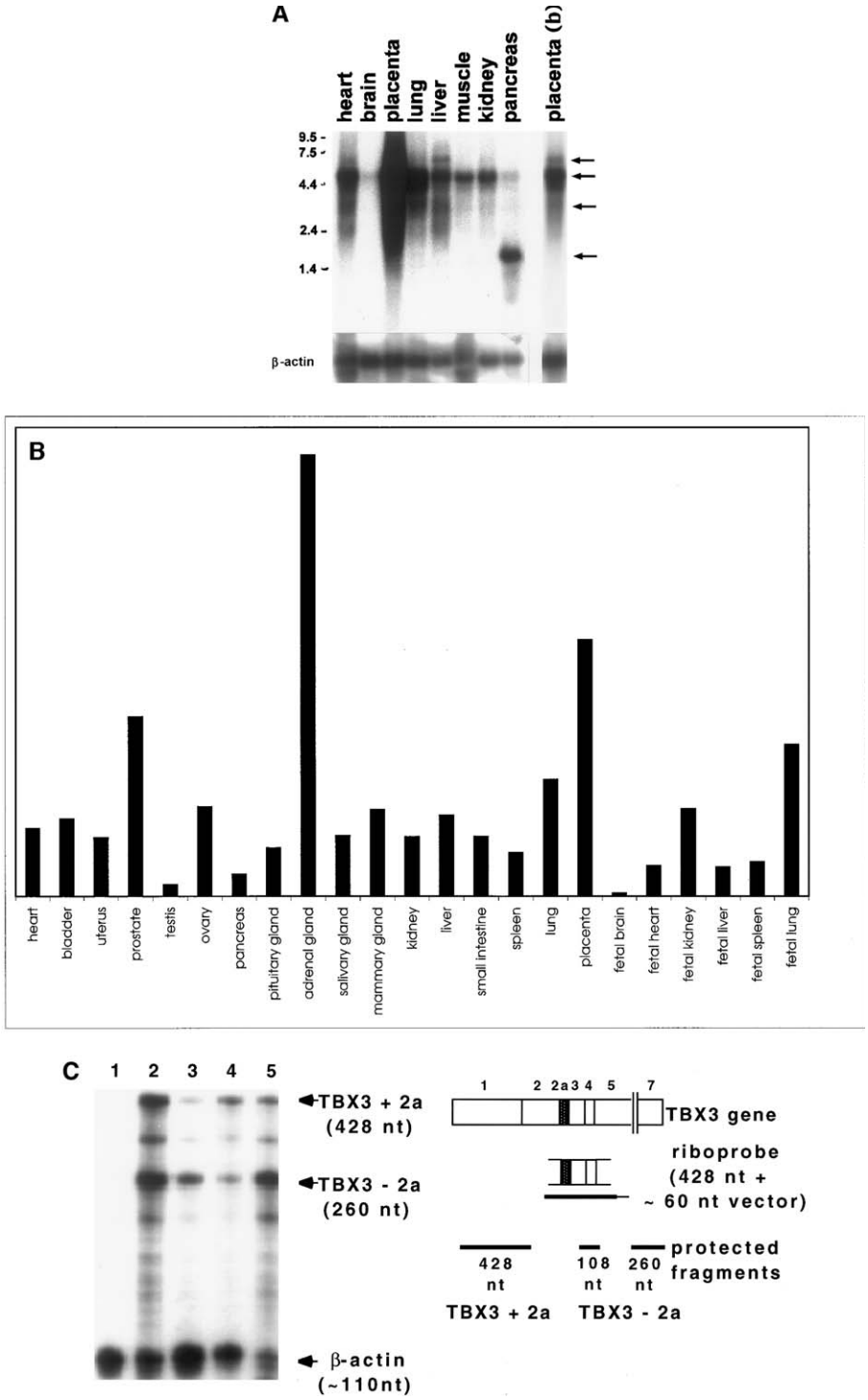


**Figure 5** *A,* Northern blot of *TBX3* mRNAs (*arrows* [←]) isolated from heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas. The predominant *TBX3* transcript is 5.2 kb in length. Alternatively spliced/transcribed forms of *TBX3* are found in the placenta, liver, and pancreas. *B,* Histogram of relative mRNA expression patterns of *TBX3* in various adult and fetal tissues. *C,* RNAse protection assay illustrating the alternative transcript of *TBX3,* which has a 60-bp insertion within the region encoding its DNA-binding domain. Lane 1, Breast-carcinoma cell line not expressing *TBX3*. Lanes 2 and 3, Breast-carcinoma cell lines expressing different amounts of *TBX3*. Lane 4, Mammary epithelial cells. Lane 5, Human fetal kidney. The diagram indicates the derivation of the riboprobe, which encompasses part of exon 2 through part of exon 5 and includes all but the first ~250 nt of the DNA-binding domain. The alternative *TBX3* transcript containing exon 2a is indicated by a stippled box. The predicted sizes of the protected fragments are 428 nt, for transcripts containing exon 2a (i.e., *TBX3* + 2a), and 260 nt plus 108 nt, for transcripts without exon 2a. Note that the signal for the 108-nt protected fragment is obscured by the signal produced by *β*-actin.

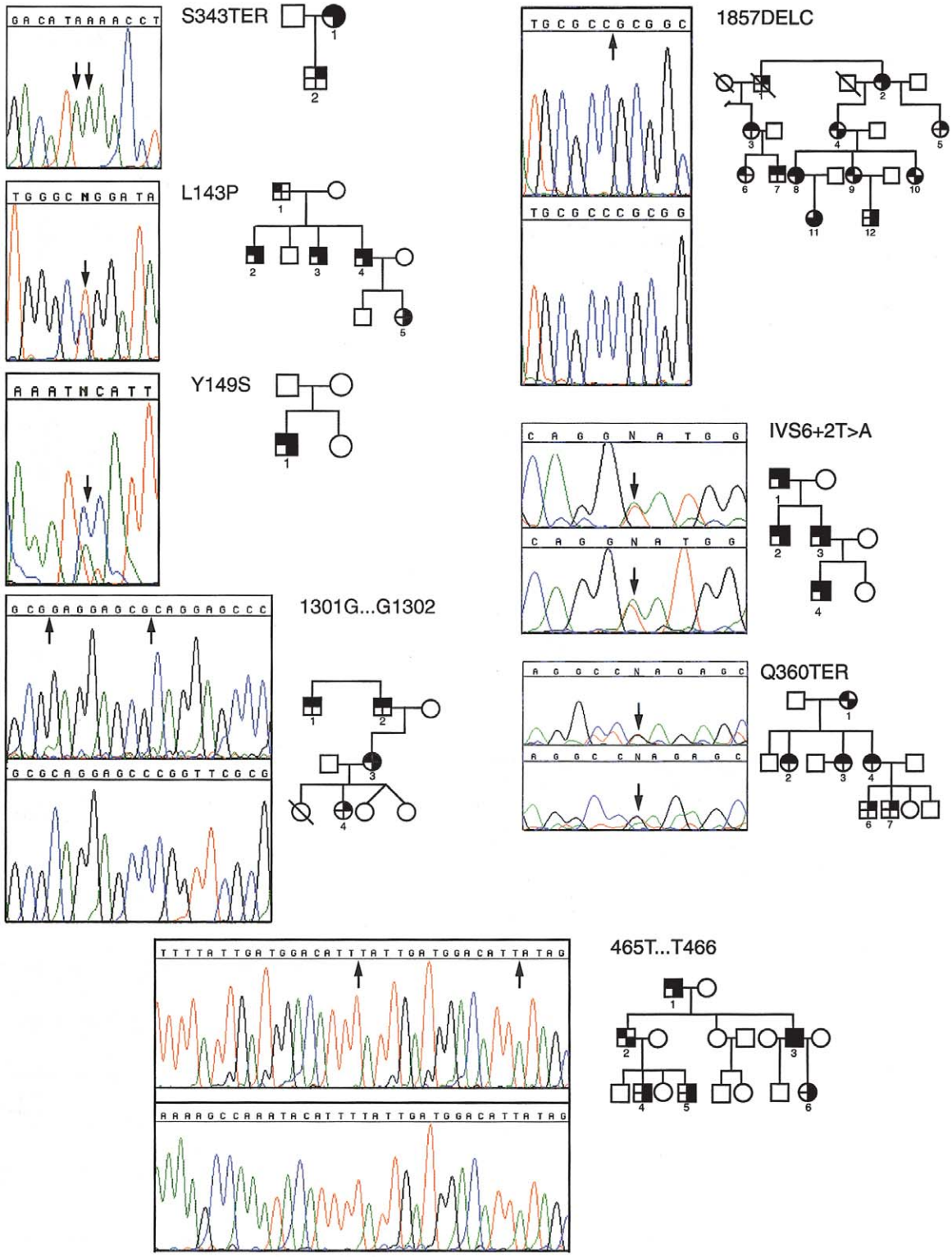


**Figure 6** Pedigrees of eight newly identified families with UMS, with electropherograms of the mutations identified in each kindred. Blackened quadrants within each symbol represent the different organ systems affected in each individual: Upper-right quadrant = limb deficiencies and/or duplications; upper-left quadrant = apocrine-gland dysfunction and/or abnormal axillary hair; lower-right quadrant = abnormalities of the breast ; lower-left quadrant = dental defects. Numbers below each symbol in the pedigree correspond to the identification numbers listed in table 1.Table 1 provides a more detailed summary of clinical findings. Panels for S343TER, L143P, and Y149S represent the forward strand of each mutant sequence. For 1301G…G1302, 465T…T466, and 1857delC, the upper panel represents the forward strand of the mutant sequence, and the lower panel represents the forward strand of the wild-type sequence. For IVS+2T→A and Q360TER, the upper panel represents the forward strand of the mutant sequence, and the lower panel represents the reverse strand of the mutant sequence.

*TBX3* might participate in the transcriptional control system of the hypothalamic-pituitary-adrenal axis (Bamshad et al. 1996). This is consistent with the observation that *TBX3* is expressed in the adult pituitary and adrenal glands (fig. *5B*).

### *Mechanism of Disease*

The mouse mutant, *Brachyury,* is caused by a 200-kb deletion that removes the entire *T* gene and effectively produces a null mutation. Additionally, correlation between *T* gene dosage and severity of the phenotype has been suggested (McMurray and Shin 1988). Consequently, the observed defects in *Brachyury* are thought to result from haploinsufficiency of T protein. In contrast, other *T* alleles (e.g., $T^{wis}$, $T^{c}$, and $T^{c\text{-}2H}$) are due to insertions and deletions that cause frameshifts and truncated T proteins (Herrmann 1991). These alleles retain their DNA-binding activity in vitro and produce developmental defects more severe than those observed in *T* null mutants. Thus, it has been suggested that they act as dominant-negative mutations (Herrmann 1991).

Mutations in *TBX3* could alter normal development by means of at least three different mechanisms—that is, by (1) disturbing the normal proportions of different *TBX3* transcripts produced by a mutant allele; (2) generating templates for the synthesis of abnormal proteins; or (3) reducing the production of normal TBX3 protein, leading to functional haploinsufficiency. There are at least four different transcripts produced by *TBX3,* including one in which exon 1 is directly spliced to exon 7. This transcript could still be produced by ⩾80% of mutant *TBX3* alleles that are not able to make a full-length *TBX3* transcript. This could alter the proportion of alternative transcript (i.e., exon 1/7) to full-length *TBX3* transcript. However, each mutation in *TBX3,* including those in exons 1 and 7, results in a similar UMS phenotype. Thus, it is unlikely that UMS results from alteration of the ratios among different transcripts of *TBX3.* Half the observed *TBX3* mutations could produce a mutant TBX3 protein that retains its DNA-binding domain. These truncated proteins could exert a functional effect through their remaining amino-terminal domains, resulting in dominant-negative interactions with normal TBX3 protein. However, analysis of the clinical features of 75 individuals with UMS shows no obvious phenotypic differences between those who have missense mutations and those who have deletions or frameshifts. Thus, it is unlikely that mutant *TBX3* alleles behave as dominant-negative mutations.

Most of the mutations in *TBX3* encode a truncated protein with a long carboxy-terminal tract of nonsense sequence, and five of six transcripts end with a premature termination codon. In general, transcription from such mutant alleles may be reduced as a result of nonsense-mediated mRNA decay (Maquat 1995), producing a functional null allele (Willing et al. 1996). Even if mutant *TBX3* alleles that encode a truncated protein were accurately translated, functional haploinsufficiency of TBX3 protein could be produced if mutations 5′ and 3′ disrupt important functional domains. Loss of function is easier to reconcile with the observation that the various *TBX3* mutations appear to produce the same functional effects. This is consistent with the wide range of types and of severity of birth defects among individuals in the same kindred. These data support the conclusions that UMS is caused by haploinsufficiency of TBX3 protein and that the wide variety of defects observed within each kindred with UMS is caused by epistatic modifiers.

There are substantial data to suggest that different mutations in the same *T-box* gene may disrupt development of different subpopulations of mesodermal cells. For example, a deletion in the promoter region of *optomotor blind* (*omb*), a *Drosophila T* homologue, causes both a dramatic reduction in *omb* expression and abnormal development of the optic neural pathways. Yet other *omb* alleles, such as *bifid* and *Quadroon,* have normal optic lobes but exhibit defects of wings and tergites, respectively (Kopp and Duncan 1997). However, no *TBX3* mutations are associated with defects of only a specific organ.

### *The Role of* Tbx3 *in Development*

At least five *T-box* genes are expressed in the developing vertebrate limb. These include *Tbx2, Tbx3, Tbx4, Tbx5,* and *Tbx15* (Agulnik et al. 1998; Gibson-Brown et al. 1998). In chick and mouse, *Tbx5* and *Tbx3* are expressed in the lateral-plate mesoderm prior to limb-bud initiation. In the forelimb, *Tbx3* is expressed along the entire length of the posterior margin of the limb-bud mesoderm. *Tbx3* expression does not extend as far distally along the anterior margin of forelimb mesoderm (Gibson-Brown et al. 1998). This is consistent with the finding that the defects observed in individuals with UMS are typically deficiencies and/or duplications of the posterior skeletal elements of the upper limb.

Defects of the apocrine glands, teeth, and external genitalia are common findings in individuals with UMS (Bamshad et al. 1996). This is consistent with the observation that *Tbx3* is expressed in mammary buds, jaw mesenchyme, and the genital tubercle (Chapman et al. 1996). Mammary buds are epithelial structures derived from the surface ectoderm. The mammary buds interact with a thickened ridge of underlying mesenchyme (i.e., the milk ridge) to form the branching-duct system of the breast (Hennighausen and Robinson 1998). Similarly, tooth development is initiated by the inductive effect of the oral epithelium on the underlying mesenchyme

(Thesleff and Sharpe 1997), and the epithelium of the genital tubercle has an inductive effect on the underlying mesenchyme that will eventually form the external genitalia (Dolle et al. 1991). Although the apical ectodermal ridge of the limb bud is organized differently, the mammary buds, the oral epithelium, and the epithelium of the genital tubercle each exhibit a similar inductive effect of epithelium on mesenchyme, or vice versa.

It has been suggested that the regulatory system of limb development was coapted for more evolutionarily derived features, such as scales, feathers, hair, and teeth (Thesleff et al. 1995). This is based on anatomical and molecular studies indicating that the early developmental anatomy of limbs, teeth, hair, and apocrine glands, including the breasts, is similar (Schwabe et al. 1998). As might be predicted, it appears that the transcriptional control systems involved in the development of the limb are also used to build other appendages of the body wall. Thus, understanding the mechanisms by which *TBX3* mutations result in malformations in developing nonlimb tissues should reveal general insights about the developmental programs controlling and coordinating organogenesis.

## Acknowledgments

We would like to thank the families for their participation, generosity, and patience. We would like to thank J. Opitz for discussion and review and would like to acknowledge the comments of an anonymous reviewer. This project was supported by Shriners Hospitals for Children, grant SHC 9510 (to. M.B.); by the Primary Children's Foundation (support to M.B.); by the General Clinical Research Center at the University of Utah, grant PHS MO1-00064; by National Institutes of Health grant DK48796 (to C.E.C.), and by the Howard Hughes Medical Institute (support to J.G.S. and C.E.S.).

## Electronic-Database Information

Accession numbers and URLs for the data in this article are as follows:

GenBank, http://www.ncbi.nlm.nih.gov (for partial*TBX3* cDNA [AF002228], *TBX3* cDNA containing alternatively-spliced T-box [U69556], alternatively spliced transcript of *TBX3* composed of exons 1 and 7 [N94306], exon 7 and the 3′ UTR of *TBX3* [AF140240], murine *Tbx3* [U57328]), and chick *Tbx3* [AF033669])

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim (for HOS [MIM 142900] and UMS [MIM 181450])

## References

Agulnik SI, Garvey N, Hancock S, Ruvinsky I, Chapman DL, Agulnik I, Bollag R, et al (1996) Evolution of mouse T-box genes by tandem duplication and cluster dispersion. Genetics 144:249–254

Agulnik SI, Papaioannou VE, Silver LM (1998) Cloning, mapping, and expression analysis of TBX15, a new member of the T-box gene family. Genomics 51:68–75

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Bamshad M, Lin RC, Law DJ, Watkins WS, Krakowiak PA, Moore ME, Franceschini P, et al (1997) Mutations in human TBX3 alter limb apocrine and genital development in ulnar-mammary syndrome. Nat Genet 16:311–315

Bamshad M, Root S, Carey JC (1996) Clinical analysis of a large kindred with the Pallister ulnar-mammary syndrome. Am J Med Genet 65:325–331

Bamshad M, Watkins WS, Dixon ME, Le T, Roeder AD, Kramer BE, Carey JC, et al (1999) Reconstructing the history of human limb development: lessons from birth defects. Pediatr Res 45:291–299

Basson CT, Bachinsky DR, Lin RC, Levi T, Elkins JA, Soults J, Grayzel D, et al (1997) Mutations in human TBX5 cause limb and cardiac malformation in Holt-Oram syndrome. Nat Genet 15:30–35

Basson CT, Cowley GS, Solomon SD, Weissman B, Poznanski AK, Traill TA, Seidman JG, et al (1994) The clinical and genetic spectrum of the Holt-Oram syndrome (heart-hand syndrome). N Engl J Med 330:885–891 (erratum: N Engl J Med 330:1627)

Basson CT, Huangt T, Lin RC, Bachinsky DR, Weremowicz S, Vaglio A, Bruzzone R, et al (1999) Different TBX5 interactions in heart and limb defined by Holt-Oram syndrome mutations. Proc Natl Acad Sci USA 96:2919–2924

Bollag RJ, Siegfried RJ, Cebra-Thomas JA, Garvey N, Davison EM, Silver LM (1994) An ancient family of embryonically expressed mouse genes sharing a conserved protein motif with the T locus. Nat Genet 7:383–389

Chapman DL, Garvey N, Hancock S, Alexiou M, Agulnik SI, Gibson-Brown JJ, Cebra-Thomas J, et al (1996) Expression of the T-box family genes, Tbx1-Tbx5, during early mouse development. Dev Dyn 206:379–390

Chariot A, Moreau L, Senterre G, Sobel ME, Castronovo V (1995) Retinoic acid induces three newly cloned HOXA1 transcripts in MCF7 breast cancer cells. Biochem Biophys Res Commun 215:713–720

Dolle P, Izpisua-Belmonte JC, Brown JM, Tickle C, Duboule D (1991) Hox-4 genes and the morphogenesis of mammalian genitalia. Genes Dev 5:1767–1776

Franceschini P, Vardeu MP, Dalforno L, Signorile F, Franceschini D, Lala R, Matarazzo P (1992) Possible relationship between ulnar-mammary syndrome and split hand with aplasia of the ulna syndrome. Am J Med Genet 44:807–812

Gibson-Brown JJ, Agulnik SI, Chapman DL, Alexious M, Garvey N, Silver LM, Papaioannou VE (1996) Evidence of a role for T-box genes in the evolution of limb morphogenesis and the specification of forelimb/hindlimb identity. Mech Dev 56:93–101

Gibson-Brown JJ, Agulnik SI, Silver LM, Niswander L, Papaioannou VE (1998) Involvement of T-box genes Tbx2-Tbx5 in vertebrate limb specification and development. Development 125:2499–2509

Hennighausen L, Robinson GW (1998) Think globally, act

locally: the making of a mouse mammary gland. Genes Dev 12:449–455

Herrmann BG (1991) Expression pattern of the Brachyury gene in whole-mount Twis/Twis mutant embryos. Development 113:913–917

Herrmann BG, Labeit S, Poustka A, King TR, Lehrach H (1990) Cloning of the T gene required in mesoderm formation in mouse. Nature 343:617–622

Kispert A, Herrmann BG, Leptin M, Reuter R (1994) Homologs of the mouse Brachyury gene are involved in the specification of posterior terminal structures in Drosophila, Tribolium, and Locusta. Genes Dev 8:2137–2150

Kopp A, Duncan I (1997) Control of cell fate and polarity in the adult abdominal segments of Drosophila by optomotor-blind. Development 124:3715–3726

Li QY, Newbury-Ecob RA, Terrett JA, Wilson DI, Curtis ARJ, Yi CH, Gebuhr T, et al (1997) Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family. Nat Genet 15:21–29

Logan M, Simon HG, Tabin C (1998) Differential regulation of T-box and homeobox transcription factors suggests roles in controlling chick limb-type identity. Development 125: 2825–2835

Logan M, Tabin C (1999) Role of Pitx1 upstream of Tbx4 in specification of hindlimb identity. Science 283:1736–1739

MacMurray A, Shin H-S (1988) The antimorphic nature of the $T^c$ allele at the mouse T locus. Genetics 120:545–550

Maquat LE (1995) When cells stop making sense: effects of nonsense codons on RNA metabolism in vertebrate cells. RNA 1:453–465

Müller CW, Herrmann BG (1997) Crystallographic structure of the T domain-DNA complex of the Brachyury transcription factor. Nature 389:884–888

Ohuchi H, Takeuchi J, Yoshioka H, Ishimaru Y, Ogura K, Takahashi N, Ogura T, et al (1998) Correlation of wing-leg identity in ectopic FGF-induced chimeric limbs with the differential expression of chick Tbx5 and Tbx4. Development 125:51–60

O'Reilly M-AJ, Smith JC, Cunliffe V (1995) Patterning of the mesoderm in Xenopus: dose-dependent and synergistic effects of Brachyury and Pintallavis. Development 121: 1351–1359

Papaioannou VE, Silver LM (1998) The T-box gene family. BioEssays 20:9–19

Ruvinsky I, Silver LM (1997) Newly identified paralogous groups on mouse chromosome 5 and 11 reveal the age of a T-box cluster duplication. Genomics 40:262–266

Schinzel A, Illig R, Prader A (1987) The ulnar-mammary syndrome: an autosomal dominant pleiotropic gene. Clin Genet 32:160–168 (erratum: Clin Genet 32:425)

Schulte-Merker S, Smith JC (1995) Mesoderm formation in response to Brachyury requires FGF signalling. Curr Biol 5: 62–67

Schulte-Merker S, van Eeden FM, Halpern ME, Kimmel CB, Nusslein-Volhard (1994) no tail (ntl) is the zebrafish homologue of the mouse T (Brachyury) gene. Development 120:1009–1015

Schwabe JWR, Rodriguez-Esteban C, Izpisúa-Belmonte JC (1998) Limbs are moving: where are they going. Trends Genet 14:229–235

Simon HG, Kittappa R, Khan PA, Tsilfidis C, Liversage RA, Oppenheimer S (1997) A novel family of T-box genes in urodele amphibian limb development and regeneration: candidate genes involved in vertebrate forelimb/hindlimb patterning. Development 124:1355–1366

Smith JC, Price BMJ, Green JBA, Weigel D, Herrmann BG (1991) Expression of a Xenopus homolog of Brachyury (T) is an immediate-early response to mesoderm induction. Cell 67:79–87

Thesleff I, Sharpe P (1997) Signalling networks regulating dental development. Mech Dev 67:111–123

Thesleff I, Vaahtokari A, Partanen AM (1995) Regulation of organogenesis: common molecular mechanisms regulating the development of teeth and other organs. Int J Dev Biol 39:35–50

Wattler S, Russ A, Evans M, Nehls M (1998) A combined analysis of genomic and primary protein structure defines the phylogenetic relationship of new members of the T-box family. Genomics 48:24–33

Willing MC, Deschenes SP, Slayton RL, Roberts EJ (1996) Premature chain termination is a unifying mechanism for COL1A1 null alleles in osteogenesis imperfecta type 1 cell strains. Am J Hum Genet 59:799–809