

Mutation watch: Mouse brachyury (*T*), the T-box gene family, and human disease

Miriam H. Meisler

Department of Human Genetics, University of Michigan, Ann Arbor, Michigan 48109-0618, USA

Received: 27 May 1997 / Accepted: 11 July 1997

The classical mouse mutation brachyury (“short tail”), first described in 1927, is located in the T/t region of proximal mouse Chromosome (Chr) 17. Heterozygotes for null alleles of brachyury have short tails and mild skeletal defects due to haploinsufficiency. Homozygous null mice have a severe developmental disorder that includes defective mesoderm formation and regression of the notochord, and they do not survive beyond E10. Seventy years of work on the short tail mouse has resulted in isolation of an important family of transcriptional regulators of early development, as well as the recent identification of the genes responsible for two complex human developmental disorders: the Holt-Oram syndrome of cardiac and limb defects, and the ulnar-mammary syndrome affecting limb, tooth, and genital development.

Cloning of brachyury and characterization of transcriptional activation

The isolation of the brachyury gene was a landmark accomplishment, the first successful identification of a mouse mutant gene by pure positional cloning (Herrmann et al. 1990). The strategy included microdissection of metaphase chromosomes to obtain probes for screening chromosome jumping libraries. Hybridization screening of an E8.5 embryonic cDNA library with a cosmid clone identified a cDNA that was located within the interval missing in a deletion allele of *T*. The *T* transcript was exclusively expressed in early mesoderm and the notochord, with an open reading frame of 436 codons encoding a protein of predicted molecular weight of approximately 44 kDa. The sequence of brachyury did not match anything in the databases until the *Drosophila* optomotor-blind gene was cloned. The two genes share a conserved domain of roughly 200 amino acids that has DNA binding activity (Pflugfelder et al. 1992). The *T* protein binds a 24 bp palindromic element (AATTC ACACCT AGGTGT GAAATT) and regulates transcription through two pairs of activation and repression domains (Kispert et al. 1995). Downstream regulated genes remain to be identified.

An evolutionarily conserved family of transcription factors, the T-box genes

The human ortholog of *T* was mapped to the conserved linkage group on Chr 6q27 that corresponds to proximal mouse Chr 17 (Edwards et al. 1996). What was more surprising was the discovery of a family of related genes with conserved DNA binding domains but divergence of the remaining protein sequence. The first three family members were identified in mouse genomic DNA by degenerate PCR, using primers corresponding to conserved sequences in brachyury and optomotor-blind (Bollag et al. 1994). The conserved domain present in all of the genes was designated the T-box, and the three new genes, *Tbx1*, *Tbx2*, and *Tbx3*, were

mapped to three different mouse chromosomes. Each gene had a unique temporal and spatial pattern of expression in midgestation embryos, suggesting divergent roles in development. At least five more T-box genes are also present in the mouse genome (Bollag et al. 1994; Bulfone et al. 1995). In lower vertebrates, homologs of brachyury, such as the *Xenopus* gene *Xbra* and the zebrafish gene no tail (*ntl*) are widely used as markers for cells of the development notochord. How ancient is the T-box gene family? Agulnik and associates (1995) identified a cluster of four T-box genes in the genome of *C. elegans* by cDNA library screening and analysis of genomic DNA sequence. Phylogenetic analysis predicted the existence of at least three T-box genes in the early metazoan ancestor of nematodes and vertebrates.

Mutations in human TBX3 and TBX5 are associated with developmental defects

The medical pay-off from basic research on the T-box genes came earlier this year with the association of two inherited developmental disorders with closely linked human T-box genes on Chr 12q24 (Li et al. 1997; Basson et al. 1997; Bamshad et al. 1997). The autosomal dominant Holt-Oram syndrome (MIM142900) with cardiac and distal limb defects had been mapped to Chr 12q2 by linkage analysis. Exon amplification of cosmid clones from the target region generated three exons that proved to be derived from two T-box genes separated by less than 300 kb of genomic DNA (Li et al. 1997). The genes were identified as TBX3 and TBX5 by their sequence similarity to the corresponding mouse genes. Mutations in TBX5 were identified in Holt-Oram patients by SSCP screening and sequencing of the PCR products. Six distinct truncating mutations were discovered, three in families and three in sporadic cases. Expression of TBX5 in developing heart was detected in human embryos between 26 and 52 days of gestation, consistent with the cardiac defects in Holt-Oram patients. Variable phenotypes were observed in patients with identical mutations. Independently, Basson and colleagues (1997) carried out exon amplification from a P1 clone contig of the Holt-Oram gene region and obtained exons that matched T-box genes in the sequence databases. A contig of the TBX5 cDNA was generated, and patients were screened for mutations by sequencing exons amplified from genomic DNA. Two mutations that co-segregated with the syndrome were detected, a nonsense mutation and an Arg to Gln substitution in a conserved Arg residue. Haploinsufficiency of TBX5 seems to account for many of the abnormalities in human patients.

The ulnar-mammary syndrome (MIM 181450) is an autosomal dominant disorder characterized by malformation of upper limb structures and teeth, apocrine/mammary hypoplasia, and genital anomalies, with highly variable clinical expression. The pleiotropic features suggested that the responsible gene product medi-

ated epithelial-mesodermal interactions. This syndrome had previously been mapped to 12q2 in a large Utah kindred, a position close to the Holt-Oram syndrome. Bamshad and coworkers (1997) determined the intron/exon structure of human *TBX3* and amplified each exon by PCR of genomic DNA from three unrelated individuals. Sequencing of the exons identified two mutations, one resulting in a truncated protein with only 86 amino acids, and the other causing aberrant splicing of the T-box domain.

Three interesting features of these newly recognized T-box disorders are haploinsufficiency, variability among affected individuals with the same mutation, and pleiotropic effects on multiple organ systems. Like the Hox and Pax transcription factors, the T-box genes have duplicated, diverged, and acquired multiple developmental roles. The other 10–20 human T-box genes will certainly be scrutinized in the future as candidates for developmental disorders. The stage is also set for analysis of interacting upstream and downstream partners of the T-box proteins, and further progress towards elucidation of the molecular circuits controlling mammalian development. All from a short-tailed mouse!

References

- Agulnik SI, Bollag RJ, Silver LM (1995) Conservation of the T-box gene from *Mus musculus* to *Caenorhabditis elegans*. *Genomics* 25, 214–219
- Bamshad M, Lin RC, Law DJ, Watkins WS, Krakowiak PA, Moore ME, Franceschini P, Lala R, Holmes LB, Gebuhr TC, Bruneau BG, Seidman JG, Seidman CE, Jorde LB (1997) Mutations in human *TBX3* alter limb, apocrine and genital development in ulnar-mammary syndrome. *Nature Genet*, in press
- Basson CT, Bachinsky DR, Lin RC, Levi T, Elkins JA, Soultis J, Grayzel D, Kroumpouzou E, Traill TA, Leblanc-Straceski J, Renault B, Kucheralapati R, Seidman JG, Seidman CE (1997) Mutations in human *TBX5* cause limb and cardiac malformation in Holt-Oram syndrome. *Nature Genet* 15, 30–35
- Bollag RJ, Siegfried Z, 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. *Nature Genet* 7, 383–389
- Bullfone A, Smiga SM, Shimamura K, Peterson A, Puellas L, Rubenstein JL (1995) T-brain-1: a homolog of brachyury whose expression defines molecularly distinct domains within the cerebral cortex. *Neuron* 15, 63–78
- Edwards YH, Putt W, Lekoape KM, Stott D, Fox M, Hopkinson DA, Sowden J (1996) The human homolog T of the mouse T (brachyury) gene; gene structure, cDNA sequence, and assignment to chromosome 6q27. *Genome Res* 6, 226–233
- Herrmann BG, Labeit S, Poustka A, King TR, Lehrach H (1990) Cloning of the T gene required in mesoderm formation in the mouse. *Nature* 343, 617–622
- Kispert A, Koschorz B, Herrmann BG (1995) The T protein encoded by brachyury is a tissue-specific transcription factor. *EMBO J* 14, 4763–4772
- Li QY, Newbury-Ecob RA, Terrett JA, Wilson DI, Curtis ARJ, Yi CH, Gebuhr T, Bullen PJ, Robson SC, Strachan T, Bonnet D, Lyonnet S, Young ID, Raeburn JA, Buckler AJ, Law DJ, Brook JD (1997) Holt-Oram Syndrome is caused by mutations in *TBX5*, a member of *Brachyury* (T) family. *Nature Genet* 15, 21–29
- Pflugfelder GO, Roth H, Poeck B (1992) A homology domain shared between *Drosophila* optomotor-blind and mouse brachyury is involved in DNA binding. *Biochem Biophys Res Commun* 186, 918–925