Mouse Brachyury the Second (T2) Is a Gene Next to Classical T and a Candidate Gene for tct

Gabriela Rennebeck,1,2 Eric Lader,1,3 Atsushi Fujimoto,4 Elissa P. Lei5 and Karen Artzt

The Institute for Cellular and Molecular Biology, Department of Microbiology, The University of Texas, Austin, Texas 78712-1064

Manuscript received December 29, 1997
Accepted for publication July 20, 1998

ABSTRACT

The mouse Brachyury the Second (T2) gene is 15 kb away from classical Brachyury (T). A mutation in T2 disrupts notochord development, pointing to the existence of a second T/t complex gene involved in axis development. T2 encodes a novel protein that is disrupted by an insertion in T2n nostalgic mice. Sequence analysis of T2 from several haplotypes shows that they all share the same changed stop codon, and, thus, T2 is a candidate gene for the t complex tail interaction factor (tct), a recessive mutation that is genetically inseparable from T (Justice and Bode 1988). T heterozygous mice have shortened tails; however, tct has no phenotype on its own but causes tailless mice when heterozygous with T (T/tct). tct thus acts as an enhancer of T. Although the close linkage of these two loci and their failure to complement suggests that they are allelic, no defect in the T gene of t haplotypes has been found (B. Herrmann, personal communication). Moreover, exogenous transgenic copies of the T gene fail to rescue the embryonic lethality caused by homoygosity for the original Brachyury deletion (Stott et al. 1993), suggesting that other essential embryonic gene(s) map to this 200-kb region.

We previously showed that Brachyury the Second (T2) interacts with tct such that T2/tct mice have tails that are short stumps (Rennebeck et al. 1995). Genetic separation between T and tct has been sought but never documented (M. Justice, personal communication), suggesting that T and tct are either the same gene or closely linked genes. As a result, the question of allelism between T and tct has remained open.

Our previous analysis of T2nostalgic showed it to be an insertional mutation mapping within the original Brachyury deletion (T). This suggested that T2nostalgic is either a novel allele of T or tct or that it acts in the same pathway as these genes because they all interact phenotypically (Rennebeck et al. 1995). The homozygous T2nostalgic syndrome is similar to but less severe than the homozygous T phenotype. Whereas T mutants do not achieve a placental connection and lack floorplate and the most posterior somites (Chesley 1935), T2nostalgic homozygous embryos achieve all these milestones. In contrast to Brachyury mutants, in which a few notochord precursors can be transiently detected but fail to persist or proliferate (Herrmann 1991; Conlon et al. 1995), T2nostalgic homoygotes contain many notochord cells distributed discontinuously along the axis. Although these cells can induce a visible floorplate, they are only distinguishable from the surrounding mesenchyme because they express the notochord markers sonic hedgehog and Brachyury (Rennebeck et al. 1995). Thus, the primary defect in T2nostalgic homoygotes appears to be the inability to organize or maintain the notochord as a structure.

Analysis of axis development by many laboratories confirms that neural tube defects in mice and, by implication, in humans are a heterogeneous group of malformations caused by numerous factors. In addition to the effects that environmental agents may have, the functions of nonallelic pairs of genes have been tied to axial development. Elevated occurrences of neural tube
defects have been found in mice that are double mutants for the transcription factors Splotch and extra toes (Copp 1995), as well as Splotch and curly tail (Estibeiro et al. 1993). Undulated-Patch double-mutant mice also have a phenotype reminiscent of an extreme form of spina bifida occulta in humans (Helwig et al. 1995).

Another valuable and well-studied model is the combination of mutant alleles of Brachyury and tct (Park et al. 1989) and, more recently, Brachyury and btm (Fujimoto et al. 1994), in which penetrance of spina bifida approaches 100%. Molecular access to multiple genetic components of the T/tct model may lead to a better understanding of this frequent birth defect.

We have identified a new gene, T2, that maps 15 kb centromeric to T (hereafter referred to as T1). The two genes are transcribed in the same direction. The T2^Bob allele is a 200-kb transgenic insertion with a concomitant deletion of 3 kb containing two of its exons. Despite this disruption, the genomic locus of T1 and its early phase of expression are apparently not affected in the T2^Bob mutant. Although this suggests the presence of a second axis-determining gene in this region, the possibility remained that the insertion disrupted some previously undefined 5' regulatory element of notochord expression of T1. The cloning of a novel gene, Brachyury the Second (T2), bridging the insertion site but not sharing any sequence with T1 may resolve this issue and explain the Brother of Brachyury (Bob) phenotype. Moreover, the sequence of T2 in three different t haplotypes reveals a common change that leads to 12 extra hydrophobic amino acids (aa) at the C terminus of the conceptual protein. Taken together, the map position of the T2 gene, its mutant phenotype, the interaction of T2 and tct, and the presence of a t haplotype-specific change strongly suggest that the t haplotype copy of T2 is a candidate gene for tct.

**MATERIALS AND METHODS**

**Nomenclature:** Because the phenotypes of T1 and T2 are so similar and genetic background is known to affect the penetrance of axis defects, it is possible that some alleles originally classified as T may really be mutations in T2. To avoid confusion, it becomes necessary to determine whether in fact T1 or T2 (or both) are mutated in the more than a dozen known "T alleles." In consultation with the mouse nomenclature committee, we propose that the original Brachyury deletion and any deletion known to affect both genes retain the designation T, and the previously characterized Brachyury gene should be called T1 with a superscript for the allele, (e.g., as in T1^111). Likewise, the T2 allele Brother of Brachyury, described by Rennebeck et al. (1995), is T2^112. Finally, if tct proves to be an allele of T2, it should be designated T2^80.

**Molecular techniques:** The Exon Trapping II System from Bethesda Research Laboratories (BRL, Gaithersburg, MD) was used according to the manufacturer's instructions. Random access retrieval of genetic information by PCR (rargip) screening was done according to Abe (1992). Biotinylation of probe was done with the Bionick kit (BRL). Streptavdin-coated magnetic beads were purchased from Promega Biotech (Madison, WI). Exon trapping and rargip clones were authenticated by analysis of the genomic sequence flanking the two trapped exons. This revealed that the splice donor/acceptor sites of the exons in the rargip clones were identical to those used by the exon trapping system. Subcloning of PCR fragments was done using the pGEM-TA vector (Promega Biotech). Sequencing was done using Sequenase version 2.0 (United States Biochemical, Cleveland, OH) or an ABI automated sequencer (Applied Biosystems, Foster City, CA). Total RNA was isolated by the LiCl/urea method (Gellibert et al. 1987). mRNA selection was done with the poly(A) Tract system (Promega Biotech).

Reverse transcription (rt) reactions were done using Superscript II (BRL) according to the manufacturer's recommendations using either oligo(dT) or gene-specific primers and 1-μg aliquots of RNA; 5' RACE was done with the BRL 5' RACE kit. PCR reactions used the Expand PCR system (Boehringer Mannheim, Indianapolis, IN) and an MJ thermal cycler. rtPCR from embryos used total RNA isolated from embryonic day (E)8 and E8.5 embryos using Trizol (BRL). First-strand cDNA was synthesized using oligo(dT) and the Superscript II RNAse H^-Reverse Transcriptase kit from BRL. PCR was performed using primers 107962 (5' ACT ATG TGT AAG ACA AGG ACG 3') and 5681C A 5' CAC ATT GTT CAC CCA GTA TCG 5'). PCR conditions were 30 cycles of denaturation for 1 min at 94°C, annealing for 30 sec at 55°C, extension for 1 min at 72°C, and final extension at 72°C for 10 min. PCR reactions were analyzed by agarose gel using a 100-bp ladder molecular marker from BRL. All radiolabeling, DNA preparation, and library screenings were performed using standard techniques. Sequence analysis was performed using the MacVector software package (IBI/Kodak). Database searches were performed using the BLAST, TBLASTN, and TBLASTX programs on the NIH/NLM World Wide Web server (http://www.ncbi.nlm.nih.gov/blast/).

**RESULTS**

Analysis of 30 kb around the site of the T2^80 transgenic insertion using the conventional techniques for gene searching failed to turn up definitive evidence of a gene. Therefore, we extended the search using two complimentary approaches: exon trapping and a cDNA direct selection technique (rargip, Abe 1992). Whereas the candidate exons from trapping yield genomic fragments with functional splice donor/acceptor sites, rargip allows the retrieval of very rare cDNA sequences from complex pools of uncloned cDNAs.

Exon trapping using Sad genomic fragments from around the insertion site yielded two potential exons of 107 and 228 bp. rargip selection was done using a 15-kb genomic clone containing the insertion site as a probe against E9.5 and E10.5 cDNA pools. rargip clones were screened with the two trapped exons as probes. Clones positive with both exons were isolated and analyzed. Although the two trapped exons map 5 kb apart in the genome, they were 80 bp apart in the cDNA clones. The largest clone (1119 bp) was used to screen six embryonic cDNA libraries from stages E7.5–E11.5. The fact that we retrieved only two positive clones from a total of 12 × 10^6 independent recombinants attests to the rarity of the T2 message. The two positive clones, isolated from E10.5 and E11.5 libraries, were contiguous...
with our largest rargip clones, but they did not extend as far 5' as those clones. Sequence analysis reveals an open reading frame whose beginning is coincident with the 5'-end of the sequence (Figure 1). Because there are no stop codons before the first AUG at position 223, it is possible that an ATG might lie 5' to the start of the known sequence. Two lines of evidence argue against this. First, the only clones we retrieved from 5'-RACE terminate downstream of the start of the known cDNA sequence (data not shown). Second, there is an in-frame stop codon in the genomic sequence 20 bp upstream of the 5'-most nucleotide in exon 1 with no obvious splice consensus sequence between it and the start of the cDNA sequence. There are Kozak consensus translational start sequences (boxed) centered around the fixed two alternatively processed forms of the transcript. One of these uses an alternative polyadenylation site ATGs at nucleotides (nt) 304-306 and 352-354. The stop codon is at nt 772, with a polyadenylation consensus and results in a 200-bp truncation of the 3' UTR. The other isoform has exons 6 and 7 spliced out and codes sequence at nt 1051. Thus, translation may produce a protein with a molecular weight of 17.8 kD and another for a 55-aa protein that lacks the 13-aa overlapping repeat. This is in fact an isoform that is similar to the protein with a molecular weight of 14.3 kD. A curious feature of either protein is a 13-aa overlapping repeat genomic structure we have found in Mus molossinus, where the repeat is missing (see below).

Northern blots and both whole-mount in situ from E8.5-10.5 and 35S in situ to sections at E10.5 were negative. However, the message is readily detectable by rtPCR to position the T2 gene relative to Brachyury, two contigs were constructed consisting of one cosmid and four lambda clones and covering a total of 80 kb. There is a 5-kb gap between the T1-containing cosmid and the closest lambda clone, as determined by pulsed-field gel electrophoresis (PFGE; Figure 2 and data not shown). The insertion in the T2 gene occurred 15 kb centromeric to the T1 gene (Figure 2). Thus, the classic T mutation, which is a 160- to 200-kb deletion with the T1 gene located in the approximate center (Herrmann et al. 1990), includes both T1 and T2. T2 contains nine exons spanning 30 kb and is in the same transcriptional orientation as T1. As diagrammed in Figure 2, the T2^{20} insertion has exons 5 and 6 deleted and exons 7-9 displaced, leaving coding potential intact for only the first 14 aa. Because the presumed 5' promoter region is intact, some aberrant transcripts may be made in the mutant.

Among the T2 rtPCR products analyzed, we identified two alternatively processed forms of the transcript. One of these uses an alternative polyadenylation site and results in a 200-bp truncation of the 3' UTR. The other isoform has exons 6 and 7 spliced out and codes for a 55-aa protein that lacks the 13-aa overlapping repeat. This is in fact an isoform that is similar to the genomic structure we have found in Mus molossinus, where the repeat is missing (see below).

The T2 message is extremely low in abundance. Northern blots and both whole-mount in situ from E8.5-10.5 and 35S in situ to sections at E10.5 were negative. However, the message is readily detectable by rtPCR from E7.5 to E11.5, the period in which notochord differentiation and development occurs. Thus, the expression of T2 is notably lower than that of T1, which is easily detected by Northern hybridization and local-
Figure 2.—Structure of the T2 gene and orientation relative to T1. Black boxes indicate size and position of T2 exons. Exons 4 and 6 were found by exon trapping. Exons 5 and 6 are deleted by the transgenic insertion. CpG island containing two BssHII restriction sites is located between exons 2 and 3. Arrows indicate transcriptional orientation. Represented restriction sites: SacI (S), BamHI (B), EcoRI (E), XbaI (X), PmeI (P), and KpnI (K).

ized by whole-mount in situ hybridization to the primitive streak and notochord (Herrmann 1991).

When the second gene in the T deletion affecting mesoderm development was identified, an obvious question was whether it was an allele of tct or possibly a third gene. To examine the T2 gene in t haplotypes, rTPCR was used to amplify the T2 transcript from total RNA isolated from individual T/t<sup>E</sup>10.5 embryos. Because T is a deletion, these embryos are hemizygous, containing only the t copy of T2. Analysis of six clones from t<sup>E</sup> revealed a change in the sequence of T2 at the third position of the putative stop codon. As expected from their common ancestry (Silver et al. 1987), two other t haplotypes (t<sup>23</sup> and t<sup>121</sup>) contain the identical change, even though these particular t haplotypes come from three very separate wild populations: New York, Montana, and France (Guenet et al. 1980), respectively. This A → C transition changes the stop codon to a tyrosine. These extra bases, combined with the mutated stop codon, cause the conceptual T2 protein of t haplotypes to contain an additional 12 aa at the C terminus (Figure 3B). Eight of these residues are phenylalanine, creating a very hydrophobic tail. This molecular lesion seems reasonable because tct homozygotes have no visible phenotype. It should be kept in mind, however, that there is currently no way to distinguish whether this change is functionally significant or merely a t-to-wild-type polymorphism.

The allelism of T2 and tct was not confirmed using the only known independent tct mutant, tct<sup>4</sup>, induced by ethylnitrosourea mutagenesis in a wild-type chromosome (Bode 1984). Sequence analysis of coding regions from genomic PCR products amplified from tct<sup>4</sup> and wild-type controls revealed no tct<sup>4</sup>-specific modification. It remains possible that the mutation in tct<sup>4</sup> has a difficult-to-detect point mutation in a noncoding region. Expression analysis of T2 is not possible because these mice are extinct.

In an attempt to find other alleles of T2, we sequenced several candidates. Among them was btm, a mutation derived from M. molossinus that is recessive and causes a short tail, but in compound heterozygotes with T, causes spina bifida (Fujimoto et al. 1995). Interestingly both btm and its molossinus parent were polymorphic for position of the putative stop codon. As expected from their common ancestry (Silver et al. 1987), two other t haplotypes (T1<sup>LAF</sup> and T1<sup>3B</sup>) contain the identical change, similar to the alternate spliced form. No significant changes were found in btm or other haplotypes sequenced. A list of the various polymorphisms is presented in Table 1.

The most relevant genetic test of "the two T's" would be to look for complementation of T2<sup>Bob</sup> with a null allele of only the T1 gene; however, none were known to exist. The T mutations described are large deletions, an ENU mutation in a t haplotype already carrying the tct mutation, or gain-of-function alleles such as T1<sup>LAF</sup> or T1<sup>3B</sup>. T1<sup>LAF</sup> is a new spontaneous allele of Brachyury. Its phenotype resembles the standard Brachyury null. Southern analysis of T1<sup>LAF</sup> homozygotes shows an insertion in the 5′ end of the T1 gene that disrupts the T box (G. Rennebeck, L. Flaherty and K. Artzt, unpublished results). Nine of 31 embryos dissected from crosses of T1<sup>LAF</sup>/+ × T2<sup>Bob</sup>/+ at E9.5-E10.5 had grossly abnormal posterior ends similar to the T1<sup>LAF</sup> phenotype. Thus, given the genomic structure of T2 and the available mutations, T1 and T2 appear to be noncomplementing, nonallelic genes.
<table>
<thead>
<tr>
<th>Exon no.</th>
<th>Base pair changes</th>
<th>Nucleotide</th>
<th>Amino acid</th>
<th>Common haplotype specific</th>
<th>Ethylnitrosourea mutation specific</th>
<th>Randomly bred strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No change</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>41</td>
<td>GTT→GCT</td>
<td>Precoding</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>61</td>
<td>AAG→GAG</td>
<td>Precoding</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>71</td>
<td>ATA→AAA</td>
<td>Precoding</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>116</td>
<td>CGA→CGG</td>
<td>Precoding</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>126</td>
<td>ACA→ACG</td>
<td>Precoding</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>No change</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>No change</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>320</td>
<td>AAA→AGA</td>
<td>Lys→Arg</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>1094–1123</td>
<td>Delete 1 repeat</td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>436</td>
<td>CGT→TGT</td>
<td>Arg→Cys</td>
<td>+</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>487</td>
<td>ACC→GCC</td>
<td>Thr→Ala</td>
<td>+</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>494</td>
<td>GAG→GGG</td>
<td>Glu→Gly</td>
<td>+</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>585</td>
<td>AGA→AGG</td>
<td>Arg→Arg</td>
<td>+</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>590</td>
<td>CAA→CGA</td>
<td>Gin→Arg</td>
<td>+</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>7</td>
<td>No change</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 has TAA</td>
<td>789–800</td>
<td>Oneless T</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>No change</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
DISCUSSION

Brachyury the Second (T2) is a novel gene that partially fulfills the criteria of being the second gene in the T/t complex involved in axis development. T2 maps within the T deletion, ~15 kb upstream of the classical Brachyury gene. T2 consists of nine exons. None of its sequence shares homology with the available databases. rtPCR experiments show that the T2 mRNA is present as early as E8 (Figure 4), consistent with the prediction that the mutation in t-int is a similarly minor modification to a protein that may have a major role in axial development. However, inability to detect T2 expression in the embryo suggests that its product is extremely low in abundance, or that T2, like nodal, has a short window of expression that was missed in the in situ hybridization experiments (Zhou et al. 1993a; Collignon et al. 1996). Both our rtPCR data and phenotypic analysis (Rennebeck et al. 1995) suggest that T2 expression is probably restricted to the latter phase of T1 expression, in the notochord.

The close linkage to the classic Brachyury gene positions T2 as a candidate gene for tct. Sequence analysis of the T2 gene in t haplotypes shows that the stop codon is altered in all t haplotypes, and this change generates a longer T2 protein with a hydrophobic carboxy tail. It is not possible to know if this change is significant because the function of T2 is unknown. However, this relatively minor modification may explain why tct has no phenotype by itself.

There is a third component of nonallelic, noncomplementing genes in the T system. It is the specific interaction of an unlinked recessive mutation t-int with T1, T2, and tct. The t-int mutation, like tct, has no phenotype on its own, but is an enhancer of tail phenotype in all mice heterozygous for T1, T2Bob or tct (Artzt et al. 1987; Rennebeck et al. 1995). Because the interaction of t-int with T1 and T2 closely resembles that of tct, it is likely that the mutation in t-int is a similarly minor modification to a protein that may have a major role in axial development.

T2Bob is probably a dominant-negative mutation. Homozygotes of the T1c mutant, which makes a protein with a modified C terminus, have a more severe phenotype than the T1 null alleles, suggesting that the mutant protein interferes with the function of other protein(s) (Searle 1966; Herrmann et al. 1990). There is compelling evidence that T2Bob is also a gain-of-function mutation. First, although + T1LAF/T2Bob embryos have the same complement of T1 and T2 mutations as a T/+ deletion heterozygote, namely one good copy of each, the latter has a short tail and the former is an embryonic lethal. Thus, the two genes have a worse phenotype in trans than the nulls do in cis. Second, the molecular lesion in T2Bob predicts a truncated message and possibly a mutant protein missing the C-terminal 150 aa.

Figure 4.—T2 mRNA is detected as early as E8. Primers in exon 2 at position 210 bp (107962) and in exon 7 at position 750 bp (5681CAC) of the T2 cDNA (2-1) were used to amplify a 600-bp fragment from a pool of cDNAs from E8-E8.5 (lanes 2 and 3). BRL 100-bp ladder molecular markers (lane 1) and dH2O control (lane 4).
The sequence of T2 suggests that the close proximity of T1 and T2 is not the result of a gene duplication event, nor do they appear to be distantly related. While their proximity might be simply fortuitous, an alternative is that the coordinate regulation of T1 and T2 during development is controlled by shared or intertwined regulatory elements. There is precedent for this with RAG-1 and RAG-2, which both function in V(DJ) recombination and are only 8 kb apart (Oettinger et al. 1990). The same is true for the Tap1:Tap2:Lmp2:Lmp7 gene complex in the major histocompatibility complex (Zanelli et al. 1993; Zhou et al. 1993b). Whether the tight linkage of T1 and T2 is by accident or necessity can be resolved by their linkage analysis in other chordates.

This work was supported by National Institutes of Health grants HD10668 and HD30658 (K.A.).

LITERATURE CITED


Geliebter, J., 1987 Dideoxynucleotide sequencing of RNA and uncloned cDNA. Focus 9: 5-6.


