Mammalian Genome 12, 212–218 (2001).
DOI: 10.1007/s003350010253

Mammalian Genome
Incorporating Mouse Genome

Canine homolog of the T-box transcription factor T; failure of the protein to bind to its DNA target leads to a short-tail phenotype

Kim Haworth,[1] Wendy Putt,[1] Bruce Cattanach,[2] Matthew Breen,[3] Matthew Binns,[3] Frode Lingaas,[4] Yvonne H. Edwards[1]

[1]MRC Human Biochemical Genetics Unit, University College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, UK
[2]MRC Mammalian Genetics Unit, Harwell, Didcot, Oxfordshire, OX11 ORD, UK
[3]Animal Health Trust, Lanwades Park, Kentford, Newmarket, Suffolk CB8 7UU, UK
[4]Section of Genetics, Norwegian School of Veterinary Science PO Box 8146 Dep., N-0033 Oslo, Norway

Received: 17 August 2000 / Accepted: 13 October 2000

Abstract. Domestic dog breeds show a wide variety of morphologies and offer excellent opportunities to study the molecular genetics of phenotypic traits. We are interested in exploring this potential and have begun by investigating the genetic basis of a short-tail trait. Our focus has been on the *T* gene, which encodes a T-box transcription factor important for normal posterior mesoderm development. Haploinsufficiency of T protein underlies a short-tail phenotype in mice that is inherited in an autosomal dominant fashion. We have cloned the dog homolog of *T* and mapped the locus to canine Chromosome (Chr) 1q23. Full sequence analysis of the *T* gene from a number of different dog breeds identified several polymorphisms and a unique missense mutation in a bob-tailed dog and its bob-tailed descendants. This mutation is situated in a highly conserved region of the T-box domain and alters the ability of the T protein to bind to its consensus DNA target. Analysis of offspring from several independent bobtail × bobtail crosses indicates that the homozygous phenotype is embryonic lethal.

Introduction

Although selective breeding of the dog began between Paleolithic and Mesolithic times, 20,000 to 10,000 years ago, most contemporary breeds were established within the last 250 years (Ostrander and Giniger 1997). They show wide inter-breed variation in behavior, facial features, coat types, weight, and height, and these traits have been selected either to enhance the working performance or for aesthetic reasons. These diverse breeds with limited intra-breed genetic variation offer unique opportunities to study the molecular genetics of phenotypic traits.

One characteristic, tail length, varies naturally such that the number of caudal vertebrae ranges from 6 to 23. Inherited taillessness (anury) and short or bob-tails (brachyury) occur at low frequency in various breeds such as Beagles, Cocker Spaniels, and Pembroke Wesh Corgis and can be inherited as recessive, semidominant, or dominant characters (Pullig 1953; Burns and Fraser 1966; Cattanach 1996).

Various genes that cause short or abnormal tails in mice could be considered as candidates for these canine phenotypes; for example, Pax1 mutations of this gene are responsible for the undulated (*un*) mutant and lead to a shortened and usually kinked tail (Wilm et al. 1998), and *Wnt-3a,* which is thought to be mutated in the vestigial tail (*vt*) mutant (Greco et al. 1996). One strong candidate is the *T* gene, which encodes a member of the T-box family of transcription factors (reviewed in Smith 1999). Mice heterozygous for *T* null mutations (Beddington et al. 1992) have short tails varying in length from near absent to 50% of normal and often kinked (Dobrovolskaia-Zavadskaia 1927; Herrmann et al. 1990). In late development, the abnormal tail shows a terminal filament that is often lost before birth (Gruneberg 1958). *T* is expressed only in early development in primitive streak, prenotochordal mesoderm, and notochord and plays a central role in mesoderm determination and notochord formation (Smith 1999). Homozygosity for *T* null alleles leads to absence of a notochord, severe abnormalities of posterior structures, and morbidity in early development (Wilson et al. 1993, 1995). Several loss-of-function and dominant negative alleles have been described in mouse (brachyury, T^{2J}, T^{wis}, T^{c} etc; (Herrmann and Kispert 1994) and zebrafish (*ntl;* Schulte-Merker et al. 1994).

In this paper we describe the cloning of the dog *T* homolog and a detailed genetic analysis of this locus in a variety of dog breeds. We have identified a number of common polymorphisms in the *T* gene, some of which were used as markers in a segregation analysis of a four-generation breed cross in which bob-tail was inherited in a dominant fashion. This led to the identification of a missense mutation in the *T* gene that alters the ability of the T protein to bind to its consensus DNA target.

Materials and methods

DNA samples. DNA was prepared from small aliquots of blood, residual to samples taken for routine blood tests, from 32 dogs comprising examples of 18 breeds and one cross-bred. In addition, samples were obtained from a Pembroke Welsh Corgi and Boxer cross that included parents and three generations of descendants derived by backcrossing to Boxers. Blood samples were collected by a veterinary surgeon into Na-EDTA and stored at −70°C until required. DNA was extracted by a standard procedure (Grimberg et al.1989) involving cell lysis in cold sucrose buffer, proteinase K digestion, removal of cell debris, and ethanol precipitation.

Cloning of the canine T gene. Recombinant clones were obtained by screening a gridded canine bacterial artificial chromosome (BAC) library prepared from dog genomic DNA (Li et al. 1999; Ref. RPCI81, supplied by HGMP Resource Centre, Cambridge, UK) by using ^{32}P-labeled mouse *T* cDNA as probe. DNA from a positive clone BAC90-I8 was extracted with a Maxiprep Plasmid Kit (Qiagen, West Sussex). Exon containing fragments were identified by restriction digestion of BAC90-I8 and Southern blotting with the mouse cDNA probe, and were subcloned into pUC19 and sequenced.

Correspondence to: Y.H. Edwards; E-mail: yedwards@hgmp.mrc.ac.uk

The nucleotide sequence data in this paper have been submitted to EMBL and have been assigned the accession number AJ245513.

Sequence analysis. Sequence analysis was performed by using ^{33}P dideoxy terminators (Amersham Pharmacia Biotech, Buckinghamshire) and a Thermo Sequenase sequencing kit (Amersham Pharmacia Biotech) with both M13 vector-specific primers and gene-specific primers. Sequences with high GC content were sequenced by using a Big Dye terminator kit (Perkin Elmer Applied Biosystems, Warrington) and an ABI 377 automated sequencer.

PCR amplification. *T* exons and intron-exon boundaries were amplified by the polymerase chain reaction (PCR) in a 50-μl reaction mix containing 100 ng of genomic DNA; 25 pM of each primer pair (Oswel, Southhampton); 0.1mM dNTPs (BRL); 5% formamide (Sigma) and 2U Taq DNA polymerase (Advanced Biotechnologies, Surrey) or Dnazyme (Flowgen) in the case of the 5′UTR PCR. PCR conditions were 95°C, 30 s; 50–62°C, 30 s; 72°C, 45 s; for a total of 35 cycles. Exons 2–7 were amplified in 50 mM KCl; 10 mM Tris-HCl, pH 8.3; 1.5 mM $MgCl_2$. Exon 8 was amplified in the same buffer with the addition of 0.1% Triton X100, and the 5′ flanking region was amplified with the addition of 1 M betaine and 5% DMSO. Exon 1 was amplified in 1 × Hotstar buffer with 1.25 Units of Hot Star Taq DNA polymerase (Qiagen). This polymerase requires an initial 15 min denaturation at 95°C for activation and extension times of 90 s. The primers were as follows: 5' flanking Dtx1-4F AGTTCCCAAAAGGGCTCGTTCC and Dtx1-5R GTTATTCCACTTGAACTCCCGC; Exon1 5'UTR Dtx1-10F AGACCTACTCCAGCCGTGCCTC Dtx1-7R GCAGGTGGTCCACTCGGTACTG; Exon 1 Dtx1-5F AAGGTGGCCCTGCGGTAGCGAGTC and Dtx1-3R-CCAGCCGAGCAGAAAGGAGCAAG; Exon 2 Dtx2-F TGGGTGTAAGGACTCAGATCC and DTx2-1R CCCTCCGTTGAGCTTGTTGGTG; Exon 2 Dtx2-F TGGGTGTAAGGACTCAGATCC and Dtx2-2R TCTGGGACAGATGTCACCTCC; Exon 3 Dtx3-F CTGCAATAGCAGGACAAATGG and Dtx3-R CATTCTTTGTGACCTCCAAGG; Exon 4 and 5 Dtx4-F GGTTGATGGACAATGCC and Dtx5-R ACTCAGAAACGTGCTGTCAGCC; Exon 6 Dtx6-F ATGAAACCGCCCGTTACTCAGC and Dtx6-R TGGAATCCATCCCTACAGATGCC; Exon 7 Dtx7-F CTGTTATTCTTGGAAGACTCCC and Dtx7-R GATGGAAATGGCATGAGTCACC; Exon 8 Dtx8-F TGCCTTAGTAGCAGAGCAAAGG and Dtx8-R TGAGCCAGACACCCAGAAGTCC.

FISH analysis. Metaphase chromosomes were prepared from dog peripheral lymphocytes by conventional hypotonic and fixation treatments. 25 ng of BAC 90-I8 DNA labeled with digoxigenin-11-dUTP was coprecipitated with 10 μg of sonicated dog genomic DNA, resuspended in 50% deionized formamide, 10% dextran sulfate, 2 × SSC. The probe plus competitor mix were denatured for 10 minutes at 70°C and pre-annealed for 30 min at 37°C prior to being added to denatured chromosome preparations for 16 h at 37°C. Post-hybridization washes were performed as described previously (Breen et al. 1999b). The hybridization sites were detected with FITC-conjugated mouse anti-digoxigenin (Sigma) followed by signal amplification with FITC-conjugated goat anti-mouse (Sigma). Chromosomes were counterstained in 80 ng/ml 4′ 6-diamidino-2-phenylindole (DAPI) and mounted in antifade solution (Vectashield, Vector Laboratories, Peterborough). Images were processed with a fluorescence microscope (Axiophot, Zeiss) equipped with an FITC/DAPI excitation filter set, a cooled CCD camera (KAF 1400, Photometrics, Tuscon, AZ) and dedicated software (SmartCapture, Vysis Inc., Illinois). At least 20 metaphases were analyzed for the presence of hybridization signals. The digital image of each DAPI-stained metaphase spread was processed with a high-pass spatial filter to reveal enhanced DAPI bands. The physical assignment was made by reference to a DAPI-banded ideogram (Breen et al. 1999c).

In vitro transcription and translation of dog T. In order to prepare cDNA templates for in vitro transcription, dog exon 1 was joined by PCR concantenation (Tuohy and Groden 1998) to mouse exons 2 to 5. In brief, dog exon 1 was PCR amplified from genomic DNA from the Ile63Met heterozygous bob-tailed Corgi parent by using the primers DTUF CUACUACUACUAAAGGTGGCCCTGCGGTAGCGAGTC and DTUR CAUCAUCAUCAUCCAGCCGAGCAGAAAGGAGCAAG . The polyuracil ends allow subcloning into the pAMP1 vector (Gibco BRL) with uracil DNA glycosylase. Subclones containing the mutant or wildtype exons 1 were selected and PCR amplified. In each case, the forward primer contained a T7 polymerase binding site, a Kozak consensus sequence, and overlapped the natural translation start codon; DTBS GGATCCTAATACGACTCACTATAGGAACAGACCACCATGAGCTCCCCCGGCGCCGAGAGC. The reverse primer contained an extension to form an overlap with exon 2 of the mouse sequence and the 3′ end of dog exon 1; DTB1R GAACATCCTCCTGCCGTTCTTGGTCACCATCATCTCGTTGG; the base at position 14 from the 3′ end of the primer (underlined) was either a G or C depending on whether wild-type or mutant product was being amplified. Mouse *T* exons 2 to 5 were amplified from cDNA by using the primers MT3FW CAAGAACGGCAGGAGGATGTTCC and MT1RW GTTCCTCCATTACATCTTTGTGG. Equal amounts (100 ng) of the dog and mouse PCR products were mixed and concatenated by PCR in a 100-μl reaction mix containing 1× cloned-Pfu Buffer (Stratagene, Cambridge), 50 pmoles of each primer, DTBS, and MT1RW, 0.2 mM dNTPs, and 5U Pfu Taq polymerase (Stratagene). The cycling conditions were 94°C × 2 min, 5 × (94°C × 45 s, 45°C × 45 s, 72°C × 2 min), 30 × (94°C × 45 s, 48°C × 45 s, 72°C × 2 min), 72°C × 5 min. The concatenated products were subcloned into pUC19. Accuracy of sequence was checked by sequence analysis of the subcloned concatenated products.

The plasmids containing the wild-type and mutant forms of the T domain were transcribed in vitro and translated by using the TNT coupled reticulocyte lysate system (Promega, Southhampton) following the manufacturer's protocol and using redivue L-[^{35}S] Met (Amersham Pharmacia Biotech UK) to label newly synthesized protein. Size and quantity of protein were checked by SDS PAGE and autoradiography.

Electrophoretic mobility shift assays (EMSAs). The DNA probe used in the electrophoretic mobility shift assays (EMSAs) was the BS.p binding site described by Kispert and Herrmann (1993), which represents a 24-bp palindromic fragment flanked by *Sma*I half-sites GGGAATTTCACACCTAGGTGTGAAATTCCC. Oligonucleotides were chemically synthesized (Sigma-Genosys, Southhampton) annealed in 100 mM NaCl and labeled with T4 polynucleotide kinase and [γ^{32}P]ATP (Amersham Pharmacia Biotech UK). Aliquots of in vitro translated T protein were incubated in 20 μl of 10% glycerol, 10 μg/ml bovine serum albumen, 75 mM NaCl, 25 mM HEPES pH 7.4, 1 mM MgCl, 0.25 mM EDTA, 0.1% Nonidet P40, 1 mM PMSF, 1 mM DTT, 1 μg/ml leupeptin, 1 μg/ml pepstatin, and 1 μg poly (dI-dC) at 25°C for 20 min prior to the addition of ^{32}P-labeled BS.p (80,000 cpm). After addition of the DNA target, the reaction mix was incubated at 25°C for a further 30 min. Binding reactions were electrophoresed in a pre-run 6% polyacrylamide gel in 44.5 mM Tris HCl, 44.5 mM borate, and 1 mM EDTA at 4°C and 10V per cm for 2.8 h. Gels were vacuum dried onto 3MM paper and autoradiographed.

Results

Cloning of the dog T homolog. Screening of 36,800 clones from a canine BAC library with a mouse T cDNA detected a single positive clone BAC90-I8. Restriction digestion of BAC90-I8 with *Bam*HI and *Hin*dIII and Southern blotting with the mouse cDNA probe identified a 4.5-kb *Hin*dIII fragment and three *Bam*HI fragments (1.4 kb, 0.5 kb and 4.5 kb) that contained all exons apart from exons 4, 5, and 6. These latter were isolated in a single 2-kb PCR product. After being subcloned into pUC19, all fragments including 1,750 bp of 5′flanking region were sequenced. Exon/intron boundaries were defined by comparison with the human (Edwards et al. 1996) and mouse (Herrmann et al. 1990) *T* genomic sequences. The dog *T* coding sequence shows 89% and 84% identity with human and mouse *T* respectively. At the amino acid (aa) level, the dog sequence (435 aa) shows 91% identity with human and mouse T and 76% with the *Xenopus* homolog Xbra (Smith et al. 1991). The identity is strongest (99%) across the T-box, where very few differences occur: aa 178 is Ala in dog and Pro in human and mouse; aa 96 is Ala in dog and human, but Thr in mouse.

Chromosome assignment of dog T. BAC clone 90-I8 DNA was used to map dog *T* to Chr 1 (CFA1 *Canis familiaris*1) at band 1q23, in the middle of a conserved segment shared with human Chr 6q (Breen et al. 1999a; Yang et al. 1999; Fig 2). The human T gene maps to the tip of Chr 6 at 6q27 (Edwards et al. 1996). In the mouse, *T* is part of an unstable region of Chr 17 designated the *t*-complex. A variant form of the *t*-complex, the so-called t-

Table 1. Polymorphic sites identified in dog *T*.

Location	Position	Variant	Enzyme[a]
5′flanking	−300	A to C	—
intron 1	+37	C to A	—
intron 1	+76	G to A	—
exon 2	511	C to T	*Bsr*I
intron 2	+22	G to A	—
intron 2	+27	G to A	—
intron 2	+71	A to G	*Msp*I
intron 3	+20	T to A	*Bsr*I
intron 6	−20	C to T	—
exon 8	1171	T to C	*Bsr*DI
exon 8	1416	T to C	*Dde*I

[a] Indicates that the base change leads to loss or gain of a restriction enzyme site. −300 indicates position is 300 bp upstream of the transcription start site. + and − numbers given for intron variants indicate whether the site is in a 3′ or 5′ position relative to the nearest exon/intron boundary.

haplotype, contains a locus *tct* that interacts with mutations of *T* to convert the short-tail into taillessness (reviewed in Schimenti 2000). Mapping in dog and human of other genes that lie within the mouse *t*-complex shows that they are dispersed among several chromosomes (Edwards et al. 1996; Langston et al. 1997) and that the *t*-complex, as such, is not a conserved feature.

Genetic analysis of T. Primers were designed from dog *T* sequence to allow the amplification of the 5′ flanking region, coding exons, and exon/intron boundaries. PCR products were searched for genetic variation by sequence analysis using DNA from dogs of five distinct breeds—a bob-tail Pembroke Welsh Corgi, long-tailed Boxer and lurcher, a short kinked-tailed French Bulldog, and a curly-tailed Pug.

Among these five dogs, 11 common polymorphic sites were found (Table 1). Of these, seven are in introns but not in sequences essential to normal mRNA splicing, one occurs in 5′ flanking region, and another in 3′ UTR. Two variants are located in coding sequence, C511T and T1171C, but are silent. Further analysis of these variants in a larger panel of dogs (n = 30) from 19 breeds (Table 2) showed that none is unique to any one breed including the bob-tail Pembroke Welsh Corgi.

In addition to the dog breed panel, we genotyped members of a breed cross in which the parents were a naturally long-tailed female Boxer and a bob-tailed male Pembroke Welsh Corgi. DNA from 24 descendants was available; of these, 12 have bob-tails, indicating that this feature is inherited in an autosomal dominant manner (Cattanach 1996). In this family, as within the Corgi breed, the bob-tail character is variable in expression, ranging from 5 cm to one-quarter normal length and occasionally kinked. Typically, the tail has a fleshy pad at its distal end and terminates in a profusely-coated filament (Fig. 1).

Three polymorphisms (C511T, T1171C, and T1416C) were used as genetic markers for genotyping. Segregation analysis showed that the two chromosomes of the parental Corgi had the haplotypes 511C: 1171C: 1416C and 511T: 1171T: 1416T. The 511T: 1171T: 1416T haplotype segregated exactly, (no recombination) with the bob-tail phenotype (Lod score >+6) (Fig. 3b). This indicates that a mutation at the T locus, or another gene close by, is very likely to underlie the tail phenotype.

Full DNA sequence analysis identified a single mutation, C295G, in exon 1 that was unique to the bob-tail Corgi parent. This base change leads to an Ile to Met substitution at aa 63 (Fig. 3a) and creates a *Bst*EII restriction site. Genotyping of all dogs in the family, by PCR and *Bst*EII digestion, demonstrated that the 295G mutation segregated with the bob-tail phenotype with no exceptions (Fig. 3b). This mutation was not observed in any other dog in our extended panel of 19 breeds (30 dogs), providing further evidence that the Ile63Met change leads to the shortened tail phenotype.

Table 2. *T* Genotypes in 20 different dog breeds and 1 cross-bred dog.

Dog breed	Exon 2 C511T	Exon8 T1171C	3′UTR T1416C
Border collie	C-C	C-C	C-C
Boxer	C-C	T-C	T-T
Boxer n =5*	C-C	T-T	T-T
Cavalier King Charles	C-C	C-C	C-C
Cross breed	C-C	C-C	T-C
Dachshund	C-T	T-C	T-C
Dalmatian	T-T	C-C	C-C
French Bulldog	C-C	C-C	C-C
Giant schnauzer	C-C	C-C	T-T
Irish setter	C-C	C-C	C-C
Irish terrier	C-C	C-C	C-C
Irish wolfhound	C-C	T-T	T-T
Japanese Akita	C-C	T-C	T-C
Labrador	C-C	T-C	C-T
Labrador	C-C	C-C	C-C
Labrador	C-T	T-C	T-C
Labrador	C-C	T-T	T-T
Labrador	C-C	C-C	C-C
Lurcher	C-T	C-C	C-C
Lurcher	C-T	T-C	T-C
Newfoundland	C-C	C-C	T-C
Newfoundland	C-C	C-C	C-C
Old English sheepdog	C-T	C-C	C-C
Pembroke Welsh corgi*	C-T	T-C	T-C
Pug	C-C	T-T	T-T
Rottweiler	C-C	C-C	C-C
Shetland sheep dog	C-C	C-C	C-C
Staffordshire bull terrier	C-C	T-T	T-T

In the case of the Boxer, 6 dogs were tested; 5 showed the same genotype.
*Indicates the dogs used as parents in the Corgi × boxer cross.

In addition to the analysis described above, four independent Pembroke Welsh Corgi bob-tail × bob-tail crosses were established and the offspring genotyped. The parents used in these crosses were derived from different lines from those used to establish the cross-bred family. This experiment was designed to see whether the Met63 mutation was a characteristic of all bob-tail Corgis and also to ascertain whether homozygosity of this mutation is embryonic lethal in dogs, as might be predicted from the phenotype of *T* brachyury mice. The offspring from these crosses, which included 13 bob-tail pups, were genotyped (*Bst*EII digests shown in Fig. 4a); all the bob-tails were heterozygous for the Met63 mutation, confirming that this mutation runs throughout the Pembroke Welsh Corgi bob tail lineage. In contrast, the long-tailed offspring did not carry the mutation. Furthermore, no homozygotes were found, strongly indicating that the Met63 mutation is lethal in homozygotes ($p = 0.0045$).

Functional analysis of the Met63 mutation. Comparison of the *T* sequences from other animal species shows that Ile63 is conserved from *Drosophila* to ascidian to human (Edwards et al. 1996; Herrmann et al. 1990; Smith et al. 1991; Kispert et al. 1994, 1995; Schulte-Merker et al.1992; Yasuo and Satoh 1994; Fig. 4b). Since Ile63 lies in a highly conserved region of the T-box domain, it seemed likely that the substitution of Ile63 for Met63 might change the DNA binding properties of the mutant protein.

In order to test this idea, we used an in vitro transcription/translation system to synthesize the mutant and wild-type T DNA binding domain. Since *T* is expressed only in very early development, dog *T* mRNA is not readily available for cDNA/PCR. To overcome this, an in-frame concatenation by PCR was used to join mutant and wild-type dog exon 1 to mouse exons 2 to 5, thereby providing a template for protein synthesis. After in vitro synthesis, the protein products were analysed by SDS polyacrylamide gel electrophoresis and autoradiography (Fig. 5A). This showed that wild-type and mutant proteins were synthesized equally well and that the presence of an additional Met at position 63 in the mutant protein does not interfere with translation efficiency.

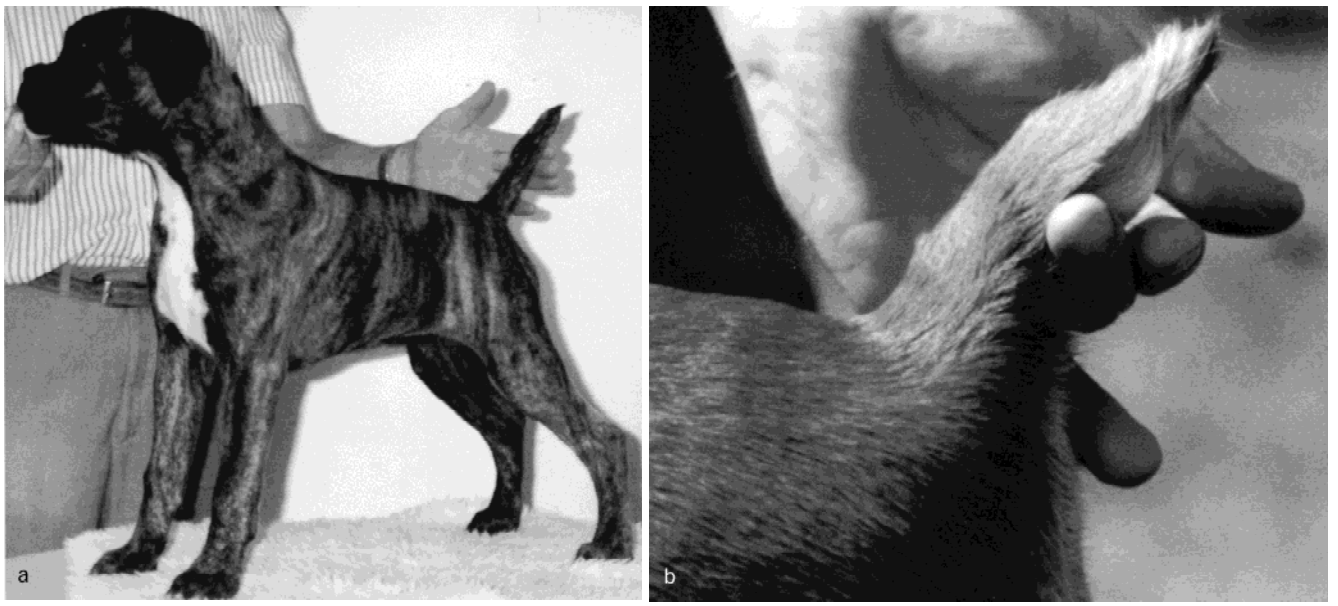

Fig. 1. a. A female puppy showing the short-tail phenotype. This dog is the result of a cross between a fourth generation bob-tail female and an unrelated Boxer male. **b.** bob-tail of a third-generation female dog. Both tails show the terminal filament.

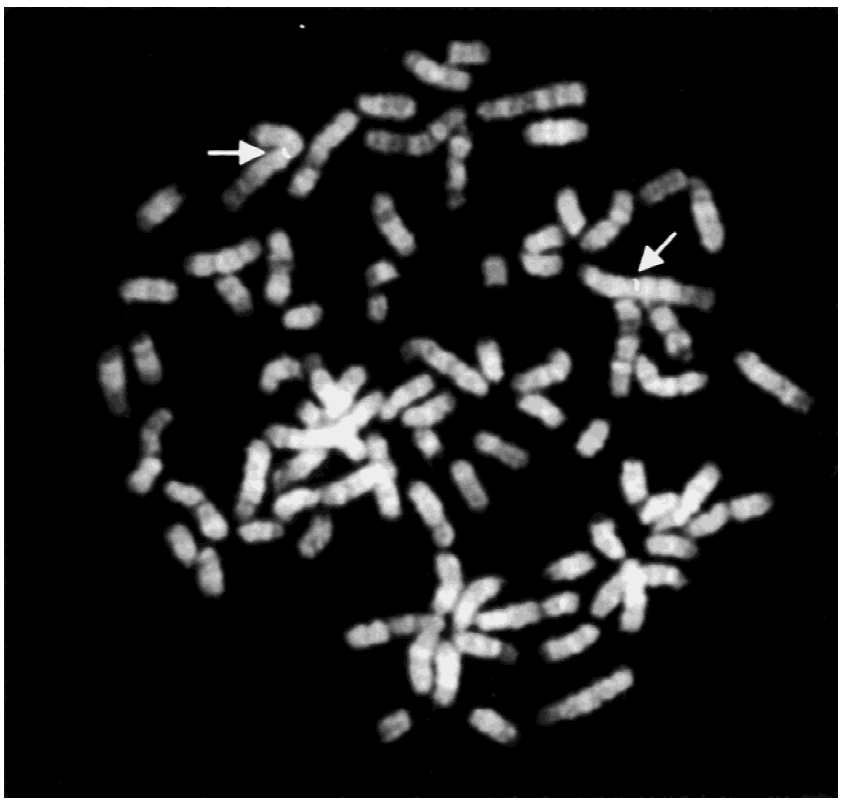

Fig. 2. FISH analysis using the dog *T* BAC clone 90-I8 DNA as probe and showing localization of the *T* gene to CFA 1q23.

DNA protein binding efficiency was explored by electrophoretic mobility shift assays (EMSAs), by using as DNA target the preferred T recognition sequence GGGAATTTCACACCTAGGTGTGAAATTCCC and equal concentrations of each protein (Kispert and Herrmann 1993). The wild-type protein bound to the DNA to give two complexes, representing T as a monomer and as a dimer, as we have demonstrated previously for human and mouse T (Papapetrou et al. 1997). In contrast, the mutant protein failed to bind to the DNA target. We do not know whether the wild type and mutant proteins can form heterodimers, or whether the Ile63 to Met63 amino acid substitution affects dimer formation. The Ile 63 residue was not identified as important for dimerization from the analysis of the crystal structure (Müller and Herrmann 1997). In the EMSAs, an equal mixture of mutant and wild-type protein showed approximately 50% binding (Fig. 5B). In addition, the proportions bound as monomer and dimer were very similar for the wild type and the 50:50 mixture. This might be interpreted as suggesting that heterodimers do not form or, if they do form, they

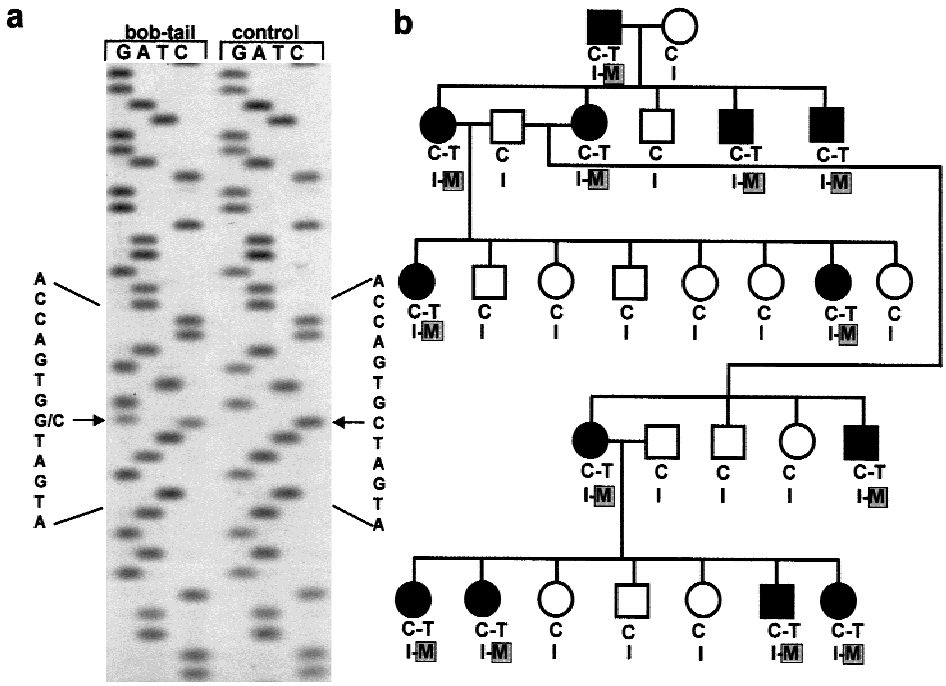

Fig. 3. a. Sequence of *T* exon 1 in DNA from the bob-tail male parent (Pembroke Welsh Corgi) and the female Boxer parent (control), across the region where the bob-tail is heterozygous for a C295G change. This sequence change gives rise to an Ile63Met substitution. **b.** Genotypes of dogs in the four generation pedigree showing segregation of the bob-tail phenotype (black symbols) and the C511T polymorphism in exon 2. Heterozygosity for the rare mutation which gives rise to the Ile to Met substitution in exon 1 is also indicated. This mutation and the 511T allele of the exon 2 polymorphism segregate exactly with the bob-tail phenotype with no exceptions.

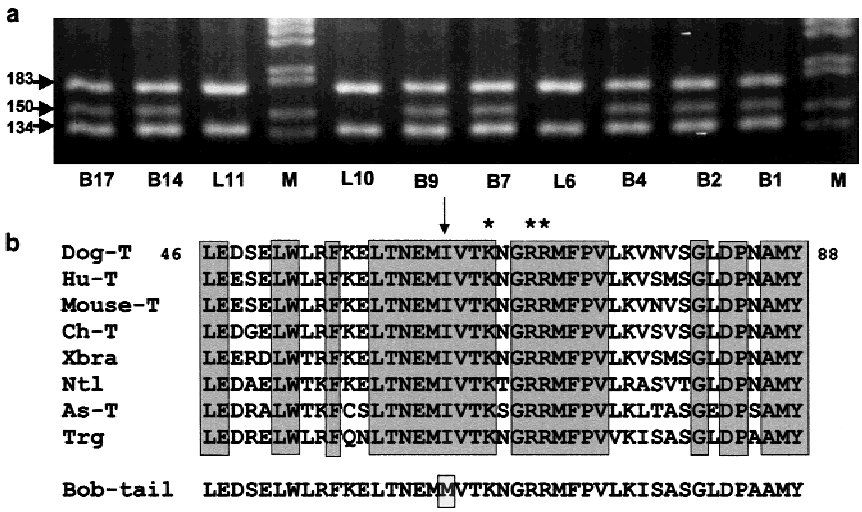

Fig. 4. a. Genotyping by *Bst*EII digestion of the exon 1 PCR product. Homozygotes show two bands of 183 bp and 134 bp while heterozygotes are three banded with an extra band of 150 bp. The samples shown are of DNA from dogs derived from bob-tail × bob-tail crosses. Bob-tail = B; long-tail = L. M = DNA size marker. **b.** Comparison of the amino acid sequence around the conserved Ile63 residue. Sequences shown are dog T (this report); Hu-T human (Edwards et al. 1996); mouse T (Herrmann et al. 1990); Ch-T chicken (Kispert et al. 1995), Xbra *Xenopus* (Smith et al. 1991); Ntl zebrafish (Schulte-Merker 1992); As-T ascidian (Yasuo and Satoh 1994); and Trg *Drosophila* (Kispert et al 1994). Conserved and identical amino acid sequences are indicated as grey blocks. The amino acids marked with asterisks are in contact with DNA during protein/DNA binding (Müller and Herrmann 1997). The rare mutation leading to the Ile63Met substitution in the bob-tail Corgi protein is indicated by an arrow and is boxed.

bind to DNA less efficiently. If heterodimers bound with the same efficiency as the wild-type dimers, then there would be a relative increase in dimer binding compared with monomers in the 50:50 mixture.

Discussion

This study shows that a C295G mutation in exon 1 of the T-box gene *T*, leading to the substitution of Met for Ile at residue 63, interferes with the binding of the T protein to its DNA target. There is very strong evidence that this functional failure results in a bob-tail phenotype in dogs heterozygous for the mutation; thus, the bob-tail seems to be due to haploinsufficiency of T protein. Animals homozygous for the mutation do not occur in the progeny from bob-tail × bob-tail crosses and presumably die in utero. The Met63 *T* mutation in the dog is very similar in its outcome to loss-of-function mutations described in the mouse. Heterozygous mice show short tails, while homozygous embryos die in utero at 10.5 days post coitum, displaying severe defects in posterior mesodermal tissues, with the notochord, allantois, and somitic mesoderm most strongly affected (reviewed in Beddington et al. 1992 and Wilson et al. 1993). However, the mutations that lead to abnormal development of the posterior tissues in mouse and zebrafish are either large deletions, insertions, or protein truncation mutations (Herrmann and Kispert 1994; Schulte-Merker et al.

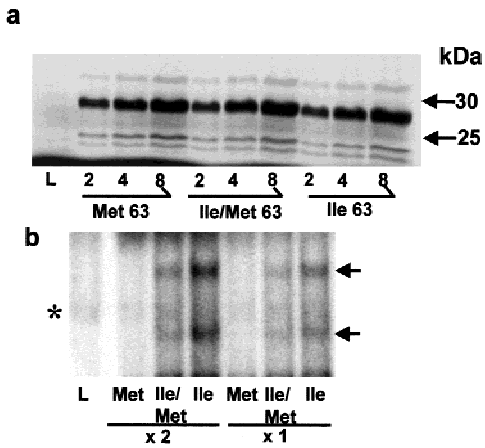

Fig. 5. a. in vitro synthesis of ^{35}S-labeled wild-type (Ile 63) and mutant (Met 63) T protein. Analysis of synthesized protein by SDS PAGE showing both proteins and an equal mixture of the wildtype and mutant form (Ile/Met 63), loaded in different amounts (arbitrary units). L = lysate only. **b.** Electrophoretic mobility shift assay using mutant (Met) and wild-type protein (Ile), and a mixture of equal amounts of both proteins (Ile/Met) at two different concentrations of the ^{32}P-labeled target DNA sequence (×1 and ×2). L = lysate only. The shifted bands that represent binding of monomeric (high mobility) and dimeric (lower mobility) T are indicated with arrows. A band due to non-specific binding of probe to a component of the reticulocyte lysate is marked with an asterisk.

1994), and as far as we are aware, this the first functional missense mutation described in the *T* gene.

A study by Woollard and Hodgkin (2000) of the *C. elegans mab-9* mutants has a bearing on the results reported here. The *mab-9* mutants are defective in hindgut and form grossly disorganized male tails. *MAB-9* turns out to be a T-box gene and is most closely related to the human *TBX20* and *Drosophila H15* genes. Two of the *C. elegans mab-9* mutant lines could be attributed to amber stop mutations and another to a frame-shift insertion. A fourth mutant was found to carry a Met to Ile change at Met97, a position which corresponds to residue 62 in the mammalian T-box domain. Remarkably, this *C. elegans* mutation in the *MAB-9* gene affects the amino acid residue adjacent to that described here in the dog *T* gene. This juxtaposition of mutations indicates that this particular region of the DNA binding domain is highly sensitive to structural changes which might subtly alter the protein conformation. Examination of the crystal structure of the T DNA binding domain shows that these residues lie very close to those identified as being in contact with the DNA during protein DNA binding (Müller and Herrmann 1997). There is no obvious *T* homolog in the *C.elegans* genome, but a search among the *C. elegans* T-box genes shows that the *mab-9* sequence is the closest to *T* (Woollard and Hodgkin 2000), and there is a clear functional relationship between these genes.

The adult *T* brachyury mouse which carries a large, 200-kb, deletion that encompasses the entire *T* gene, shows in addition to a short tail some pathological changes in the lumbo-sacral region and ankylosis of the posterior vertebrae (Dobrovolskaia-Zavadaskaia et al. 1934; Gruneberg 1958). In humans, a rare variant of T in a patient with sacral agenesis, and a genetic association between T alleles and susceptibility to spina bifida, have been reported (Papapetrou et al.1999; Morrison et al. 1998). However clinical investigations, including radiographic studies, of 19 randomly selected, naturally occurring bob-tail Corgis have not discovered any associated pathology (Indrebø and Langeland 1997). This may reflect species-to-species variation in response to gene product dosage. There is a growing number of examples of this type. For example, a striking difference is shown by mice and humans to sonic hedgehog (SHH) levels; in humans loss of product from one allele is sufficient to cause holoprosencephaly, whereas, in the mouse, the products of both alleles need to be lost to produce a similar phenotype (Chiang et al. 1996). Such differences might be explained in part by species-specific requirements for different concentrations of the gene product to produce the necessary biological responses and/or the presence of species-specific modifier genes.

Acknowledgments. We thank Caroline Burton, Greg Firth, and Sue Povey for supplying dog samples for the breed panel. We also thank the Norwegian Welsh Corgi club for providing blood samples from the bob-tail × bob-tail crosses as well as Dr Anita Sørensen for technical assistance. Canine genome mapping at the AHT (MB,MMB) is generously supported by the Guide Dogs for the Blind Association. We also thank the makers of the dog BAC library, Robin Li and Pieter de Jong, and the Wellcome Trust (grant no 052908/2/97) and the UK HGMP resource centre for supplying the library filters and clones.

References

Beddington RSP, Rashbass P, Wilson V (1992) *Brachyury*—a gene affecting mouse gastrulation and early organogenesis. Development Suppl. 157–165

Breen M, Thomas R, Binns MM, Carter NP, Langford CF (1999a) Reciprocal chromosome painting reveals detailed regions of synteny between the karyotypes of the domestic dog (*Canis familiaris*) and human. Genomics 61, 145–155

Breen M, Langford CF, Carter NP, Holmes NG, Dickens HF, et al. (1999b) FISH mapping and identification of canine chromosomes. J Hered 90, 27–30

Breen M, Bullerdiek J, Langford CF (1999c) The DAPI banded karyotype of the domestic dog (*Canis familiaris*) generated using chromosome specific paint probes. Chromosome Res 7, 401–406

Burns M, Fraser MN (1966) *The Genetics of the Dog*, 2nd edn. (Edinburgh: Oliver and Boyd)

Cattanach B (1996) Genetics can be fun. Dog World, August to September issues

Chiang C, Litingtung-Y, Lee E, Young-KE, Corden-JL, et al. (1996) Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. Nature 383, 407–413

Dobrovolskaia-Zavadskaia N (1927) Sur la mortification spontanée de la queue chez la souris nouveau-née et sur l'existence d'un caractère heriditaire 'non-viable.' C R Soc Biol 97, 114–116

Dobrovolskaia-Zavadskaia N, Kobozieff N, Verentennikoff (1934) Etude morphologique et génétique de la brachyourie chez les descendants de souris à testicules irradiés. Arch Zool Exp Gén 76, 249–358

Edwards YH, Putt W, Lekoape K, Stott D, Fox M, et al. (1996) The human homolog of the mouse *T Brachyury* gene: gene structure, cDNA sequence, and chromosomal localisation. Genome Res 6, 226–233

Greco TL, Takada S, Newhouse MM, McMahon JA, McMahon AP et al. (1996) Analysis of the vestigial tail mutation demonstrates that Wnt-3a gene dosage regulates mouse axial development. Genes Dev. 10, 313–324.

Grimberg J, Nawoschik L, Belluscio L, McKee R, Turck A et al. (1989) A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. Nucleic Acids Res 17, 8390

Gruneberg H (1958) Development of brachyury and anury. J Embryol Exp Morphol 6, 424–443.

Herrmann BG, Kispert A (1994) The *T* genes in embryogenesis. Trends Genet 10, 280–286

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

Indrebø A, Langeland M (1997) Is there a connection between natural occurring bobtails and defects in the spinal vertebrae in Welch Corgi Pembroke? "*CorgiPost*," the Annual of the Norwegian Welsh Corgi Club, No. 4 pp 28–29

Kispert A, Herrmann BG (1993) The *Brachyury* gene encodes a novel DNA binding protein. EMBO J 12, 3211–3222

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

Kispert A, Ortner-H, Cooke J, Herrmann BG (1995) The chick *Brachyury*

gene: developmental expression pattern and response to axial induction by localized activin. Dev Biol 168, 406–15
Langston A, Mellersh C, Neal C, Ray K, Acland G et al. (1997) Construction of a panel of canine rodent hybrid cell lines for use in partitioning of the canine genome. Genomics 46, 317–325
Li-R, Mignot-E, Faraco-J, Kadotani H, Cantanese J et al. (1999) Construction and characterization of an eightfold redundant dog genomic bacterial artificial chromosome library. Genomics 58, 9–17
Morrison K, Papapetrou C, Hol F, Mariman E, Lynch S et al. (1998) Susceptibility to spina bifida; an association study of five candidate genes. Ann Hum Genet 62, 379–396
Müller CW, Herrmann BG (1997) Cystallographic structure of the T domain–DNA complex of the *Brachyury* transcription factor. Nature 389, 884–888
Ostrander EA, Giniger E (1997) Semper fidelis. What man's best friend can teach us about human biology and disease. Am J Hum Genet 61, 475–480
Papapetrou C, Edwards YH, Sowden J (1997) The T transcription factor functions as a dimer and exhibits a common human polymorphisms Gly177Asp in the conserved DNA binding domain. FEBS Lett 409, 201–206
Papapetrou C, Drummond F, Reardon W, Winter R, Spitz L et al. (1999) A genetic study of human *T* gene and its exclusion as a major candidate gene for sacral agenesis with anorectal atresia. J Med Genet 36, 208–213
Pullig T (1953) Anury in cocker spaniels. J Heredity 44, 105–107
Schimenti J (2000) Segregation distortion of mouse *t*-haplotypes: the molecular basis emerges. Trends Genet 16, 240–243
Schulte-Merker S, Ho RK, Herrmann BG, Nusslein-Volhard C (1992) The protein product of the zebrafish homolog of the mouse *T* gene is expressed in nuclei of the germ ring and the notochord of the early embryo. Development 116, 1021–1032
Schulte-Merker S, van Eeden F, Halpern M, Kimmel C, Nusslein-Volhard C (1994) No tail (ntl) is the zebrafish homolog of the mouse T (Brachyury) gene. Development 120, 1009–1015
Smith J (1999) T-box genes; what they do and how they do it. Trends Genet 15, 154–158
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
Tuohy T, Groden J (1998) Exons-introns—Lexons: In-frame concatenation of exons by PCR. Hum Mutat 12, 122–127
Wilm B, Dahl E, Peters H, Balling R, Imai K (1998) Targeted disruption of Pax1 defines its null phenotype and proves haploinsufficiency. Proc Natl Acad Sci USA 95, 8692–8697.
Wilson V, Rashbass P, Beddington RSP (1993) Chimeric analysis of *T* (*Brachyury*) gene function. Development 117, 1321–1331
Wilson V, Manson L, Skarnes WC, Beddington RSP (1995) The *T* gene is necessary for normal mesodermal morphogenetic cell movements during gastrulation. Development 121, 877–886
Woollard A, Hodgkin J (2000) The *Caenorhabditis elegans* fate-determining gene *mab-9* encodes a T-box protein require to pattern the posterior hindgut. Genes Dev 14, 596–603
Yang F, O'Brien P, Milne B, Graphodatsky A, Slanky V et al. (1999) A complete comparative chromosome map for the dog, red fox and human and its integration with canine genetic maps. Genomics 62, 189–202
Yasuo H, Satoh N (1994) An ascidian homolog of the mouse *Brachyury* (*T*) gene is expressed exclusively in notochord cells at the fate restricted stage. Dev Growth Differ 36, 9–18