

RESEARCH

Open Access



Resequencing of the TMF-1 (TATA Element Modulatory Factor) regulated protein (TRNP1) gene in domestic and wild canids

James C. Sacco^{1*} , Emma Starr¹, Alyssa Weaver¹, Rachel Dietz¹ and Muhammad A. Spocter²

Abstract

Background Cortical folding is related to the functional organization of the brain. The TMF-1 regulated protein (TRNP1) regulates the expansion and folding of the mammalian cerebral cortex, a process that may have been accelerated by the domestication of dogs. The objectives of this study were to sequence the *TRNP1* gene in dogs and related canid species, provide evidence of its expression in dog brain and compare the genetic variation within dogs and across the *Canidae*. The gene was located *in silico* to dog chromosome 2. The sequence was experimentally confirmed by amplifying and sequencing the *TRNP1* exonic and promoter regions in 72 canids (36 purebred dogs, 20 Gy wolves and wolf-dog hybrids, 10 coyotes, 5 red foxes and 1 Gy fox).

Results A partial TRNP1 transcript was isolated from several regions in the dog brain. Thirty genetic polymorphisms were found in the *Canis* sp. with 17 common to both dogs and wolves, and only one unique to dogs. Seven polymorphisms were observed only in coyotes. An additional 9 variants were seen in red foxes. Dogs were the least genetically diverse. Several polymorphisms in the promoter and 3' untranslated region were predicted to alter *TRNP1* function by interfering with the binding of transcriptional repressors and miRNAs expressed in neural precursors. A c.259_264 deletion variant that encodes a polyalanine expansion was polymorphic in all species studied except for dogs. A stretch of 15 nucleotides that is found in other mammalian sequences (corresponding to 5 amino acids located between Pro58 and Ala59 in the putative dog protein) was absent from the *TRNP1* sequences of all 5 canid species sequenced. Both of these aforementioned coding sequence variations were predicted to affect the formation of alpha helices in the disordered region of the TRNP1 protein.

Conclusions Potentially functionally important polymorphisms in the *TRNP1* gene are found within and across various *Canis* species as well as the red fox, and unique differences in protein structure have evolved and been conserved in the *Canidae* compared to all other mammalian species.

Keywords TRNP1, Cortical folding, Genetic polymorphism, Dog, Canidae

*Correspondence:

James C. Sacco

james.sacco@drake.edu

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Plain English Summary

The folded shape of mammalian brains allows the cerebral cortex and other structures to attain a large surface area to fit into the skull. Over the past several thousand years, the domestication of dogs has led to increased folding as well as a great variety of skull size and shape amongst various breeds. Cortical folding, also known as gyrification, is regulated by numerous genes, including one that encodes the TMF-1 regulated protein (TRNP1). The TRNP1 protein in the brain affects the development of specialized cells in the brain that are involved in gyrification. It is not known whether a functionally distinct TRNP1 protein in dogs is responsible for the increased folding. Variations in the DNA sequence of the gene that encodes TRNP1 may be responsible for these dramatic changes in brain structure in dogs. This study sought to discover the differences in the TRNP1 DNA sequence in seventy-two canids, represented by thirty-six dogs of various breeds, twenty gray wolves and wolf-dog hybrids, ten coyotes, five red foxes and one gray fox.

After finding evidence of the expression of this gene in dog brain, we located thirty genetic changes or variants in the canids, with seventeen common to both dogs and wolves, and only one unique to dogs. Another seven of these genetic variants were observed only in coyotes. An additional nine variants were seen in red foxes. Dogs were the least genetically diverse species, an expected result of the inbreeding that characterizes domestication. Several of these changes may affect the function of the TRNP1 gene by affecting the binding of other biomolecules to regions in the DNA which regulate this gene. This study also found two other changes, one only found in dogs, and the other one only found in canids (compared to all other mammalian TRNP1 proteins) may change the length and three-dimensional structure and hence the function of the TRNP1 protein. This study concluded that numerous, potentially functionally significant dissimilarities in the TRNP1 gene exist between dogs and their wild relatives, as well as between canid and all other mammalian species.

Background

The cerebral cortex is a crucial region in the mammalian brain that regulates cognitive behavior [22]. Cortical folding (also known as gyrification), which originates during fetal and early postnatal development, is intrinsically related to the functional organization of the brain and has long been studied as a marker of both normal and pathological brain function. In humans, cortical size is crucial for normal brain function, as patients with microcephaly or macrocephaly (in other words, small or enlarged brains, respectively) show a range of cognitive deficits. As part of the evolutionary process, cortical folding has enabled the mammalian brain to grow in volume and to expand in surface area despite being restricted to a certain skull size [58].

Based on cortical folding, mammals can be divided into lissencephalic species (such as mice and manatees), which have smooth-surfaced cortices, and gyrencephalic species (such as ferrets and small primates), which exhibit convolutions in the cortex [23]. Typically, gyrencephalic brains are found in large rodents and large primates. The gyrification index (GI) is a measure of the total cortical surface area relative to the convex smooth hull that defines the outer boundaries of the cerebrum. Across mammalian species, including carnivores and canids, the GI shows a strong positive relation with brain mass [18, 37, 43].

Artificial selection has resulted in domestic dog (*Canis lupus familiaris*) breeds that have diverged significantly from the form of their closest ancestor, the grey wolf

(*Canis lupus lupus*). While brain size has decreased by about 30%, the GI in dogs has not decreased [18, 56]. A brain imaging study conducted on wild canid and domestic brains concluded that the GI in canids is positively correlated with cortical surface area, thickness and total gray matter volumes [18]. However, within domestic dogs the strength of this hypoallometric relationship is dramatically reduced. In addition, the hyperallometric relationship displayed between cortical thickness and GI for the *Canidae* indicates that global folding changes outpace changes in the underlying cortical thickness. In addition, lower local GI (i.e., regional gyrification measured within a defined functional area) in foxes compared to dogs, wolves and coyotes imply a divergent folding pattern for the anterior temporoparietal area between “fox-like” and “wolf-like” canids, [18]. The domestic dog also exhibits more morphological diversity than any other species, with the greatest variation evident in the size and shape of the skull [55], which in turn has led to distinct changes in cerebral organization between dolichocephalic (long-skulled) and brachycephalic (short-skulled) breeds [44].

The mammalian TMF-1 (TATA Element Modulatory Factor) regulated protein (*TRNP1*) is one of several genes that are known to regulate the expansion and folding of the mammalian cerebral cortex by accelerating cell cycle progression [54]. This gene is expressed in the ventricular zone and neuronal layers of the developing cortex. TRNP1 levels have contrasting effects on the expansion of the murine cerebral cortex, with high expression in the germinal layers of the precentral and parahippocampal

gyri promoting neural stem cell self-renewal and tangential expansion, whereas low expression in the germinal layers of the occipital and temporal lobes induces radial expansion and increased gyrification [52]. Recently, it has been demonstrated that TRNP1 acts as a negative transcriptional regulator of basal progenitor genesis, controlling radial glial fate by its interaction with several nuclear membrane-less organelles [11].

It is not known whether the changes observed in the dog brain due to domestication are a result of altered expression of *TRNP1* and/or sequence differences within the TRNP1 protein that alter its function. As with other proteins, the level of TRNP1 in the brain is subject to interindividual variation, which can be due to various extrinsic or intrinsic factors. An important intrinsic factor is the presence of genetic polymorphisms in the *TRNP1* gene, which may account for differences in gyrification observed within dogs and across the *Canidae*.

Therefore, we designed a study whose objectives were to sequence the gene coding for TRNP1 in both a diverse dog population and selected wild canid species, screen for genetic polymorphisms, and predict any effects these variants may have on TRNP1 function. We also aimed to provide evidence of its expression in select regions of the canid brain. Finally, we also compared the experimentally determined TRNP1 DNA and resulting inferred protein sequences in domestic dogs with published TRNP1 sequences for other canids and carnivores as well as species representative of other mammalian orders.

Methods

Animals and sample collection. A total of 72 canids were included in this study. Thirty-six pure-bred dogs (representing 33 *Canis lupus familiaris* breeds, including 2 dingoes and 3 New Guinea singing dogs) recruited for this study were either privately owned or provided through the collaboration of the Jordan Creek Animal Clinic, West Des Moines IA (Additional file 1). The American grey wolves (*Canis lupus lupus*, $n=20$), coyotes (*Canis latrans*, $n=10$), and foxes (red fox *Vulpes vulpes*, $n=5$, gray fox *Urocyon cinereoargenteus*, $n=1$) were from the Colorado Wolf and Wildlife Center (Divide, CO), Wolf Park (Battle Ground, IN), Shy Wolf Sanctuary (Naples, FL), Blanke Park Zoo (Des Moines, IA), and the JAB Canid Education and Conservation Center (Santa Ysabel, CA). Buccal cell samples were collected from the dogs using cheek swabs designed for use in canines (Performagene®, DNA Genotek Inc.). DNA collected in this manner is stable for at least one year when stored at room temperature. Brain tissue was obtained from a single dog that was euthanized due to terminal illness. All experimental procedures were approved by the Drake University Institutional Animal Care and Use Committee.

Identification of target region. Since the chromosomal location of *TRNP1* is unannotated in the most recent version of the canine genome (canFam6), we located the predicted transcript (XM_038462447) using the gene search function on the UCSC Genome Browser. To ensure that this transcript represents TRNP1 mRNA and also to identify exons, we then used the blastn tool (available at: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>) to search for highly similar sequences (megablast) in the mammalian nucleotide and *Canis lupus familiaris* EST collection while excluding predicted transcripts. Having established the identity of the predicted transcript as TRNP1, we used the BLAT Search Genome Tool [25] to align the sequence with the canine genome assembly. We also used the Broad Improved Canine Annotation v1 track data hub within the UCSC Genome Browser [20]) to provide evidence of tissue expression of the TRNP1 transcript.

Nucleic acid isolation, gDNA and cDNA amplification and gene resequencing

Following inactivation of nucleases and precipitations of impurities in the buccal samples, genomic DNA (gDNA) was purified via ethanol precipitation (Performagene®, DNA Genotek Inc.). We used 200 ng of genomic DNA to amplify both exons and the proximal promoter region of TRNP1. The reaction mixture consisted of a 25 μ l of reaction mixture containing 0.25 μ M of each primer (Table 1), 4–8% of GC Enhancer, and Amplitaq Gold 360 Master Mix (Applied Biosystems, Foster City, CA). After an initial incubation at 95 °C for 10 min, PCR amplification was performed for 40 cycles consisting of 95 °C for 30 s, 60 °C for 30 s and 72 °C for 45 s, followed by a final extension at 72 °C for 7 min. The specificity of each PCR was checked by electrophoresis on a 1.5% agarose gel.

Total RNA was isolated from distinct regions (cerebellum, hippocampus, occipital cortex, frontal cortex) within a single dog brain (Purelink RNA Mini Kit,) and 2 μ g was reverse transcribed into first-strand cDNA using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Two sets of primers were used in the subsequent PCR, each designed to amplify progressively longer stretches of the TRNP1 transcript (Table 1). The PCR conditions were similar to those described above, except for the exclusion of GC enhancer.

Following purification by Exo-SapIT (Affymetrix, Santa Clara, CA), the amplicons were submitted to bidirectional Sanger sequencing using the Big-Dye Terminator v3.1 (Eurofins Genomics, Louisville, KY). 5% of the samples were randomly chosen and resequenced to confirm the initial genotype result. Sequence assembly and identification of genetic polymorphisms was performed using Staden package software (<http://staden.sourceforge.net/>).

Table 1 Sequences of primers used for the amplification of the *TRNP1* gene and transcript

Primer name	Oligonucleotide sequence (5'→3')	Region/length amplified
Primers for gDNA template		
cfTRNP1_prom_F	TGGGCTGTTTCCTCTCTG	Proximal promoter/exon 1 c.-829 to c.135
cfTRNP1_prom_R	CGGAGTCAAGGTCGGAGTT	
cfTRNP1_ex1_1F	CGTCGGTCTTCCCCTCG	Exon 1/Coding sequence c.-76 to c.*94
cfTRNP1_ex1_2R	CGCTGGCCCTTCTTGAGG	
cfTRNP1_ex2_F	AAGAGTTTGGGTGGGGTTGG	Exon 2/3' UTR c.*143–169 to c.*802
cfTRNP1_ex2_R	TGCGGAATACAAGTCTGGGT	
Primers for cDNA template		
cfTRNP1_839F	CGGGAAGAGACAGTCGGG	Partial transcript (within 3' UTR) c.*97 to c.*246
cfTRNP1_968R	TGTTTAAGTTGTGGGACGTC	
cfTRNP1_868F	CAGTGAAGGAGGCTGCTC	Partial transcript (within 3' UTR) c.*126 to c.*713
cfTRNP1_1445R	CTCCCTGAGAATGAGGTGCC	

Positions are denoted with respect to the predicted start codon. 3'UTR: 3'untranslated region; underlined nucleotides indicate exon 1/2 splice junction; * indicates numbering from start of 3' UTR

Data analysis. In order to determine the degree of conservation between the canid species, DNA sequences were compared using ClustalOmega [50]. The aligning of transcription factors to the canid *TRNP1* promoter sequence was performed by using JASPAR position-specific scoring matrix models through the LASAGNA algorithm v 2.0, an integrated webtool for transcription factor binding site search and visualization [29, 30]. By inputting sequence variants, we identified the effects that promoter polymorphisms may have on DNA binding proteins that may bind to our sequence of interest. Further information on the function of these transcription factors was obtained from UniProtKB [53] and their potential relevance to the role of TRNP1 in cortical development assessed using GO annotations available at the Gene Ontology Resource [4, 17], specifically “cerebellar cortex morphogenesis” (GO:0021696), “interkinetic nuclear migration” (GO:0022027) and “neural precursor cell proliferation” (GO:0061351). The effect of 3'UTR genetic polymorphisms on microRNA (miRNA) binding was analyzed by submitting sequence variants to the MirTarget target prediction tool available at the miRDB database [9, 32]. Sequence, expression and functional information on the miRNAs of interest was obtained from the miRmine database [42] and TargetScan prediction tool [39]. The MPI bioinformatic toolkit [15] and the Phyre2 protein fold recognition server [24] were used to subject the predicted dog TRNP1 protein (XP_038318375) to sequence analysis, prediction of secondary structure, and sequence comparison with its homologs from the genus Carnivora (8 species from the *Canidae* and 32 species from other families) and with 25 other species, representative of six other mammalian genera. Since no TRNP1 experimental protein structure is available, we used the HHPred tool to select homologous templates

of known structure and forwarded the alignment to Modeller [45] in order to calculate a modeled structure for the dog TRNP1 protein.

The effect of genetic variation on the structure and function of the TRNP1 protein within and across species were predicted by the following tools: Polyphen-2 [1], Align GVDG [38] Coils [35] Deepcoil [34], and Phyre2 [24].

Results

The 1,682 nucleotides of the predicted XM038462447.1 transcript align with 100% identity to dog chromosome 2, at locus 69,744,731–69,751,120. The gene is composed of two exons (ex 1: 907 bp; ex 2: 774 bp), separated by a single intron (4827 bp). Experimental evidence for the entire sequence was inferred from partial mRNAs (all incomplete at the 5' end), recorded as canine expressed sequence tags (ESTs) isolated from kidney, ovary, heart, muscle, and cerebral cortex (<https://genome.ucsc.edu>). Only the sequence of one EST (CF411103) includes the single splice junction. The predicted full-length transcript aligned with several primate and rodent TRNP1 reference sequences. The closest match (79.23% identity) was with pig TRNP1, represented by NM_001243828. Very poor identity was obtained with non-mammalian vertebrate sequences, as reported by other studies [54].

Despite multiple attempts at optimization, we were unable to isolate the entire TRNP1 transcript (predicted length 1682 bp), probably due to the high GC content in exon 1 (80%). However, we were able to sequence a partial length (606 bp) TRNP1 transcript from all brain tissue samples, that is, cerebral cortex, hippocampus, frontal cortex and, cerebellum. This sequence (represented by GenBank accession numbers OQ191934, OQ191935) corresponds to 78% of the 3' UTR region that spans both

exons. The regions of *TRNP1* that were sequenced using gDNA as template exhibited 95% or higher similarity across the five species investigated, with a higher degree of conservation in both exons (97.8% or higher) than in the proximal promoter (95.67 or higher).

Thirty *TRNP1* polymorphisms were identified within the dog, wolf and coyote cohorts, with dogs and coyotes being the least and most genetically diverse respectively. All variants represented single base substitutions, with the exception of c.-240delA in the proximal promoter (rs852983207), a GCGGCG deletion in the coding region of exon 1 at c.259_264 (resulting in p.Ala87_Ala88del), and a variable TG dinucleotide repeat ($n=1-3$) in exon 2 at c.*370_371 (Fig. 1A). Significant sequence similarity enabled us to use the same primers to amplify the homologous *TRNP1* regions in the fox species. This enabled us to identify 9 polymorphisms in the red fox, three of which (c.-129 A>G, c.274insGCG, c.274insGCG [5]) occur at analogous loci in the *Canis* species (Fig. 1B, Additional file 2).

Comparison of the sequences of the deletion variant at c.259_264 across the five canid species indicates that this results in a trinucleotide GCG repeat polymorphism that occurs within wolf, coyote and red fox populations (Fig. 2). This locus was not polymorphic in

dogs, who have a strongly conserved consecutive series of five GCG repeats, representing 5 alanine residues (AAAAA). At least 35% of coyotes and wolves possessed the allele represented by three GCC repeats (AAA). In red foxes, three alleles were found, represented by 8 ((GCG)₂GCA(GCG)₅), 9 ((GCG)₂GCA(GCG)₆) and 13 ((GCG)₃GCA(GCG)₉) repeats that encode variable polyalanine lengths. The single grey fox sequenced had a series of 9 consecutive GCG repeats.

This polyalanine region appears to be very variable across the *Carnivora*, with the number of consecutive alanine repeats ranging from 3 in wolves to 12 in tigers and leopard cats, in an otherwise highly conserved part of the protein (Fig. 3). A similar degree of variability was observed in 43 additional species representing 9 additional mammalian orders (Additional file 3).

Homology modeling predicted that this polyalanine repeat lies between two alpha helices in the N-terminal disordered region (Fig. 4). A highly conserved proline at position 80 in the dog (position 87 in the consensus sequence in Fig. 3) creates a kink in the protein chain which separates two alpha helices. This accuracy of this prediction was limited by the fact that only 40% of the sequence could be modelled with high confidence due to the disordered region.

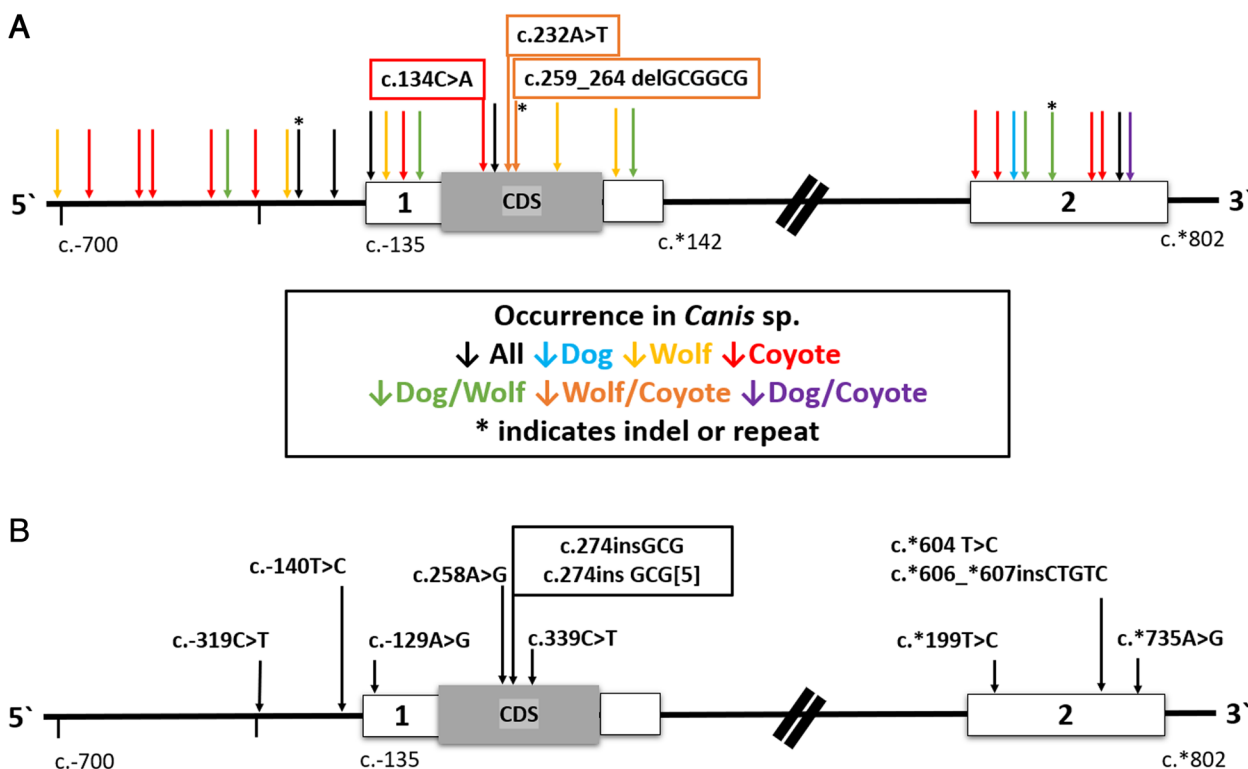


Fig. 1 *TRNP1* genetic polymorphisms detected in this study. **A** *Canis* species; **B** Red fox. Boxed polymorphisms indicate nonsynonymous variants. Intron 1 not to scale

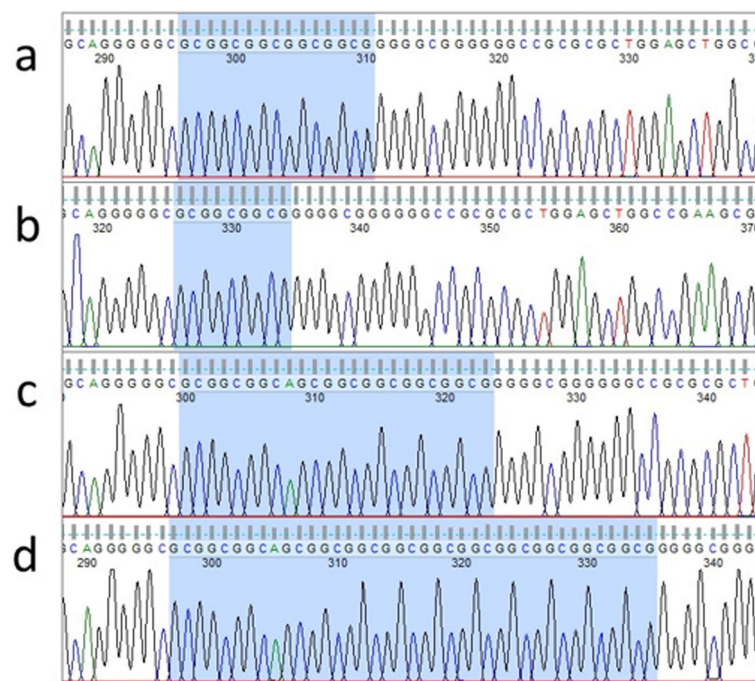


Fig. 2 DNA electropherograms representing the alignment of the GCG repeat polymorphism in (a) dog, (b) wolf, (c, d) red fox

Regardless of species, the formation of alpha helical structures is predicted to occur in this part of the protein if the number of consecutive polyalanine repeats exceeds 7, as for example, is the case with red fox (Fig. 5A). Due to the disordered nature of the region, this prediction is, however, of low confidence. Additionally, an increased number of polyalanine repeats may affect the structure and stability of the coiled coil domain, which is located just downstream of the repeat region (Fig. 5B).

There were distinct differences in the allele frequency of several polymorphisms between the wolves and dogs in the study (Fig. 1A; Table 2). Thirteen variants occurred in two or all three species, while others were only detected in one species. For example, while domestic dogs share 11 polymorphisms with one or all canid species, only one mutation was specific to dogs (c.*280 C>A). Conversely, five polymorphisms were unique to wolves and another 11 variants were only found in coyotes. The variants represented by c.-240delA, the c.-412 C>T, and the c.*370_371insTG appear to be the major allele in dogs with respect to the reference allele currently represented at the same locus in canFam6.

Several polymorphisms in the TRNP1 proximal promoter were predicted to alter the binding of transcription factors by deleting or altering the interaction with existing sites as well as creating novel sites (Table 2). A number of these transcription factors play important roles in cerebellar cortex morphogenesis, including

RE1-silencing transcription factor (REST), homeobox protein engrailed-1 (EN1), transcriptional repressor CTCF (C2CF) and zinc finger protein 423 (ZNF423). Four genetic polymorphisms in the 3'UTR region were predicted to delete or create new binding miRNA sites (Table 2). Several of the miRNAs that may be impacted, such as hsa-miR-335-3p, hsa-miR-769-3p, and hsa-miR-216a, are expressed at high levels in the human brain.

Only one polymorphism, the synonymous variant c.141 T>G, was found in the TRNP1 coding region in dogs; the variant allele G was the only or major allele in wolves and coyotes respectively. Another coding SNP, c.232 A>T (Thr>Ser78), found in wolves and coyotes, is predicted to be benign (Polyphen-2 score: 0.000; align GCVG Class C55). Conversely, another nonsynonymous polymorphism, c.134 C>A, which leads to the amino acid change Pro>Gln45, found in a single coyote, was predicted to be damaging (Polyphen-2 score: 0.988; align GCVG Class C65), as the proline at that locus is highly conserved in mammals.

All the canid TRNP1 DNA sequences determined by this study, as well as the predicted nucleotides for two others (common raccoon dog *Nyctereutes procyonoides*, and Arctic fox *Vulpes lagopus*), exhibit a distinctive 15-base pair gap in the coding region (represented by the absence of 5 amino acids located between Pro58 and Ala59 in dogs), when compared to all other mammalian TRNP1 sequences, including other species of the order *Carnivora* (Fig. 6). Species belonging to the *Canidae* lack

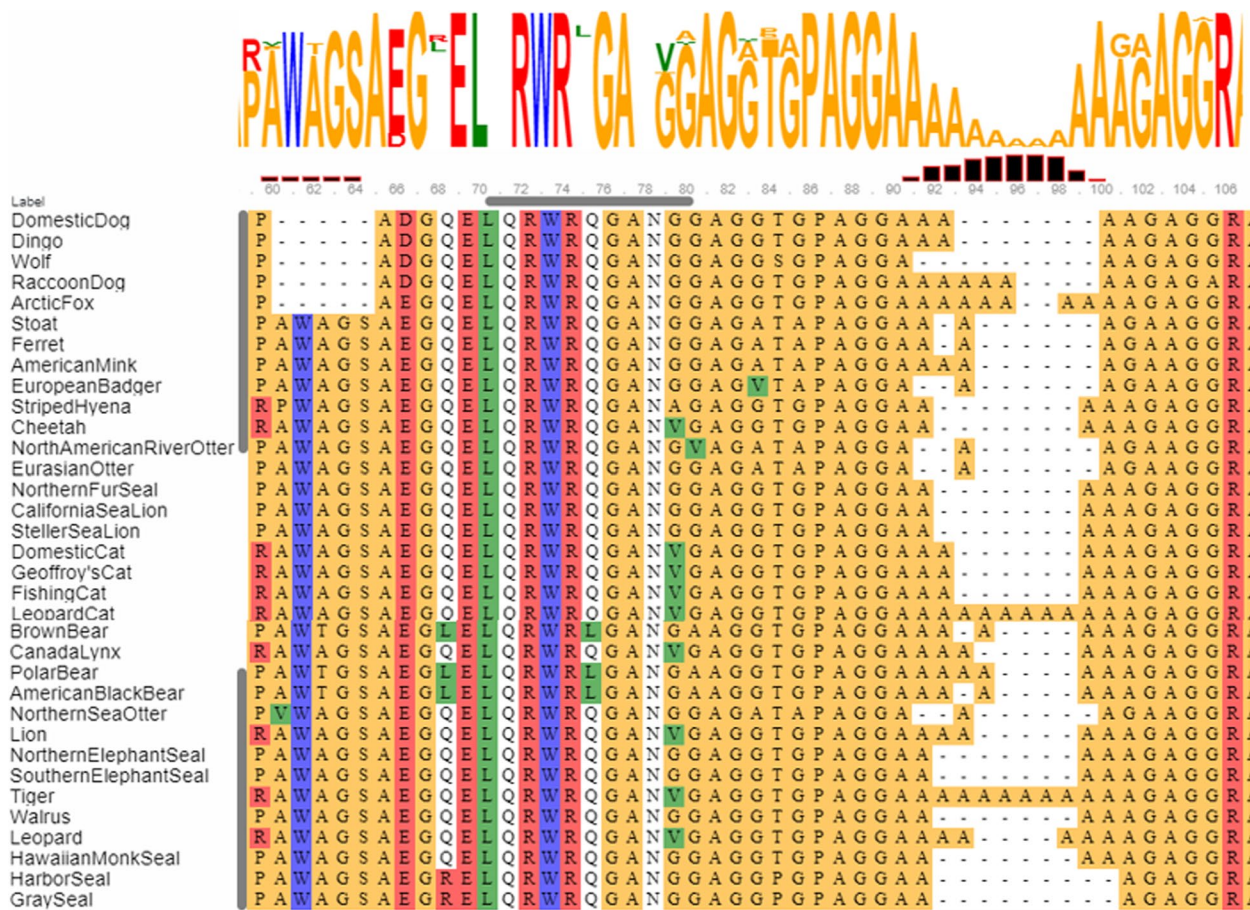


Fig. 3 Clustal Omega alignment of part of the N-terminus of the TRNP1 protein in 34 carnivore species. The variable alanine repeat is shown to the left. Color coding for individual amino acids is as follows: yellow – small/polar; red – charged; blue - aromatic; green - hydrophobic

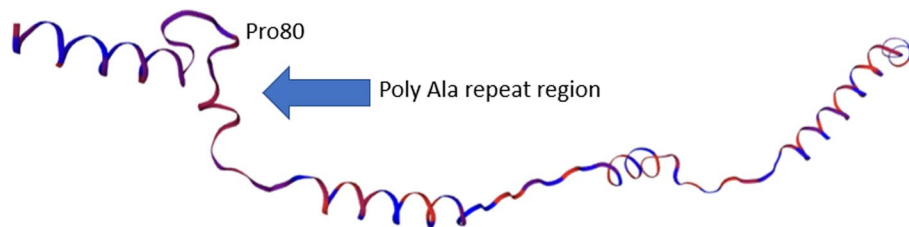


Fig. 4 Calculated model for the N-terminal region of the dog TRNP1 protein (residues Ala59-Ala166 shown), based on homology with the cryo-EM structure of the MukBEF monomer (PDB#7NYY)

an A/PWA/TGS sequence that appears to be ubiquitous to other carnivores. The end-result of this gap is a smaller stretch of small or polar amino acids, that is flanked, as with all other species by charged residues such as glutamine, aspartic and glutamic acids.

Discussion

Experimental sequencing and comparative genomic tools enabled us to confirm that, as with other mammals, the canid *TRNP1* gene, located on dog

chromosome 2, is composed of two exons. Exon 1 and the proximal promoter region are located within a 1 kb CpG island, indicating a housekeeping gene function for TRNP1. Although there is no evidence of methylation of individual CpG sites in dog brain or placental tissue [46, 47], the technique used so far is limited by its low coverage and a detailed study of the area is likely to reveal evidence of epigenetic alterations. The entire coding region (666 bp in dogs) lies within the first exon and constitutes only 40% of the transcript, while the

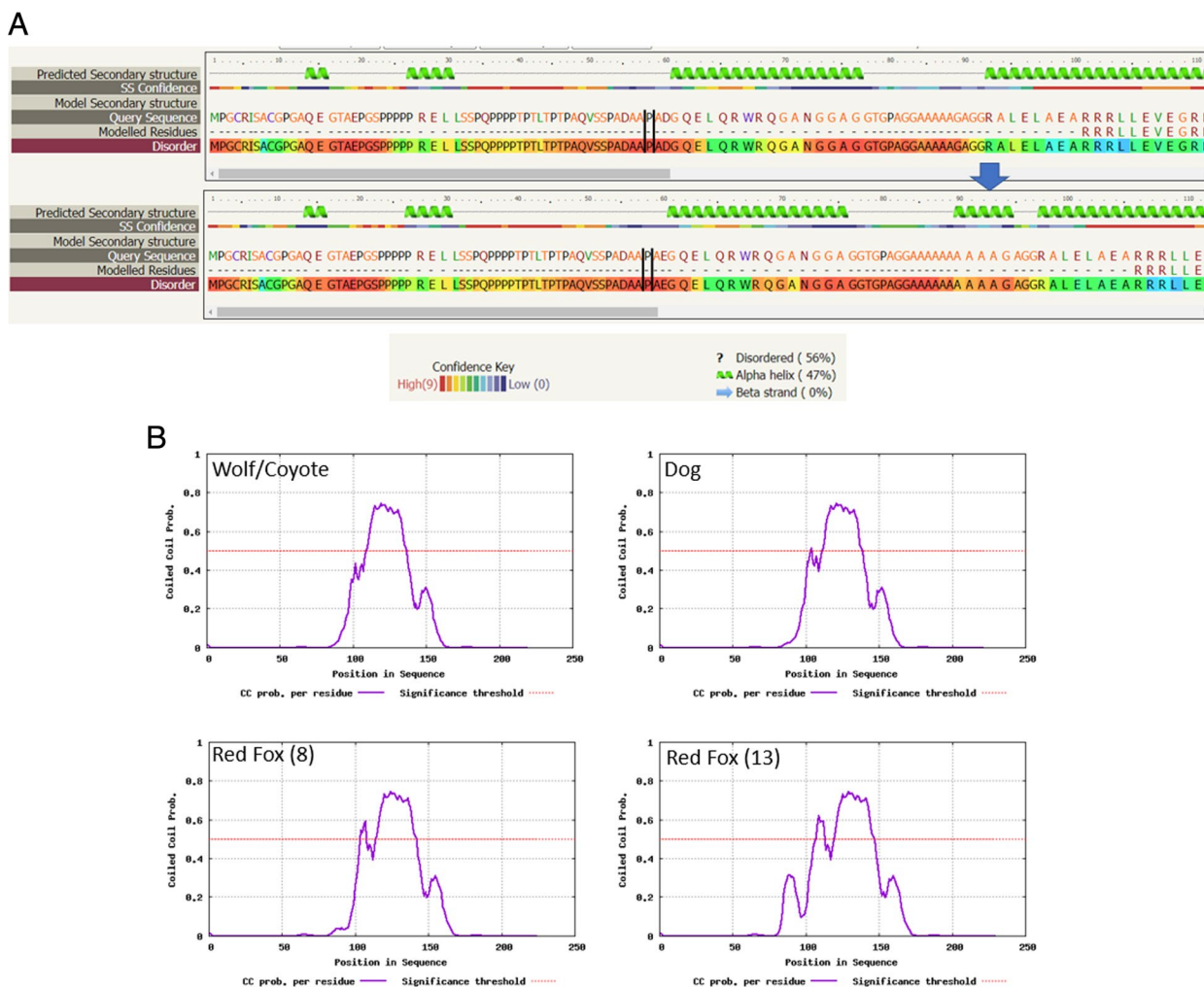


Fig. 5 Comparison of the effect of polyalanine repeats on TRNP1 protein secondary structure between dog and wild canids. **A** Alpha helix formation (top: dog; bottom: red fox (Ala₁₀)* as predicted by Phyre2 with Pro80 highlighted by parallel lines; **B** Probability of formation of coiled coil structure in wolf/coyote (Ala₃), dog (Ala₅) and red fox (Ala_{8,13}) as calculated by DeepCoil. *length representative of alleles observed in red fox

larger 3'UTR region is comprised within both exons. Disproportionately large 3'UTRs are often observed in genes that encode proteins that are involved in multiple protein interactions, and the nervous system selectively expresses isoforms with longer 3'UTRs [5].

The high GC content described above was likely the reason for our failure, despite extensive PCR optimization, to amplify the full, or at least the 5' end of the dog TRNP1 transcript. In fact, a number of mammalian genomes with low coverage also exhibit gaps in their homologous exon 1 regions. Nevertheless, we were able to isolate around 90% of the 3'UTR, including the region that comprises the splice junction. This transcript was present in all four dog brain regions investigated.

Amongst the *Canis* species, coyotes were observed to have the greatest number of genetic polymorphisms, 11

out of 20 (or 55%) being unique to that species. Similarly, 5 variants were observed exclusively in wolves, representing 45% of the total number of polymorphisms found in this wild canid. Conversely, dogs, even though they represented various breeds, including ancestral breeds such as dingoes and New Guinea singing dogs, had the lowest number of polymorphisms (6), with only one mutation being specific to dogs. This lower genetic diversity is a hallmark of domestication, which may also explain the wide differences in allele frequency observed for some variants (c.-412T, c.-240 delA, c.*370_371insTG) between dogs and wolves. Interestingly, some of the most striking differences were observed in the coding region: c.141G and c.232T were present as the minor allele in dogs compared to the two wild canid species, while c.259_264 delGCGGCG was only found in wolves and coyotes at allele frequencies of 0.48 and 0.35 respectively.

Table 2 Genetic polymorphisms in the canid *TRNP1* gene, variant allele frequencies (AF) and predicted effect of variants on regulatory and coding regions

Gene region	Genetic polymorphism	Variant AF D W C	Predicted effect	Known function or expression of affected transcription factor or miRNA in brain tissue
Promoter	c.-702 C>G	0.00 0.03 0.00	TCF3, NFIC binding site deleted	Both genes expressed in cerebral cortex. TCF3 is a transcriptional regulator involved in the initiation of neuronal differentiation and mesenchymal to epithelial transition
	c.-665 C>T	0.00 0.00 0.10	KLF4 binding site deleted; New NR2E3 binding site	KLF4 regulates the expression of key transcription factors during embryonic development. NR2E3 involved in cerebellar cortex morphogenesis
	c.-607 T>C	0.00 0.00 0.05	None	
	c.-584 G>A	0.00 0.00 0.10	New CTCF binding site	Negative regulation of cell population proliferation
	c.-444 C>A	0.00 0.00 0.05	MYC binding site deleted; New TBXT, ZEB1 binding sites; Increased binding for EN1	MYC expressed in the proliferating cells of the developing CNS. EN1 and TBXT involved in cerebellar cortex and neural plate morphogenesis. ZEB1 positively regulates neuronal differentiation.
	c.-412 C>T	0.60 0.15 0.00	New ZFX binding site	None reported
	c.-355 C>T	0.00 0.00 0.05	New PAX5, PAX6 binding sites; Increased binding for RUNX1	PAX important for cerebellar cortex morphogenesis, including regulation of neuroblast proliferation and neuron migration. RUNX1 involved in neuron differentiation.
	c.-256 G>C	0.00 0.03 0.00	ELK1, MIZF binding sites deleted	ELK1 is a positive regulator of neuron death while MIZF is involved in cerebellar cortex morphogenesis.
	c.-240delA	0.76 0.38 0.05	REST, ZNF423 binding sites deleted	Expressed in neural progenitor cells, REST is a transcriptional repressor that regulates neurogenesis and neuron differentiation. ZNF423 regulates cerebellar granule cell precursor cell proliferation.
	c.-161 T>G	0.04 0.10 0.10	Spl1 binding site deleted; Increased binding for NFYA	None reported
5'UTR	c.-129 A>G	0.13 0.08 0.10	New binding sites for TBP	None reported
	c.-104 T>C	0.00 0.03 0.00	None predicted	
	c.-79 C>T	0.00 0.00 0.10	New binding sites for ARNT and CREB1	CREB1 involved in axonogenesis
	c.-58 C>T	0.25 0.15 0.00	None predicted	
Coding	c.134 C>A	0.00 0.00 0.05	p.(Pro45Gln) – probably damaging	
	c.141 T>G	0.39 1.00 0.85	p.Pro47=	
	c.232 A>T	0.00 0.43 0.25	p.(Thr78Ser) – probably benign	
	c.259_264 delGCGGCG	0.00 0.48 0.35	p.(Ala87_Alal88del) – may alter spacer length	
	c.453 G>A	0.00 0.03 0.00	Gln151Gln	
3'UTR	c.*19 A>G	0.00 0.05 0.00	None	
	c.*41 G>T	0.08 0.13 0.00	None	
	c.*149 C>T	0.00 0.00 0.05	None	
	c.*257 G>A	0.00 0.00 0.15	New sites for hsa-miR-335-3p, 450b-3p, 769-3p, 5694	hsa-miR-335, 769 expressed at high levels in human brain
	c.*280 C>A	0.03 0.00 0.00	None	
	c.*286 A>G	0.13 0.23 0.00	None	
	c.*370_371 insTG	0.64 0.95 1.00	Deletion of hsa-miR-128, 216a, 3681 sites	cfa-miR-128-1 detected in canine cerebrospinal fluid; hsa-miR-128-3p, 216a-3p expressed at high levels in human brain
	c.*452 T>A	0.00 0.00 0.05	None	
	c.*454 A>G	0.00 0.00 0.05	None	
	c.*482 C>T	0.01 0.28 0.05	Deletion of hsa-miR-549a-3p site	
c.*495 C>G	0.03 0.00 0.55	Deletion of hsa-miR-6820-3p, hsa-miR-3164 sites		

D dog; W wolf; C coyote. Alleles are denoted as variant with respect to the reference allele at that locus on the (+) strand on the canFam6 assembly. Variant accession numbers are given in Additional file 2. * indicates numbering from start of 3'UTR. Software prediction programs listed in text. Function of transcription factors were based on gene ontology using the GO: biological process annotation

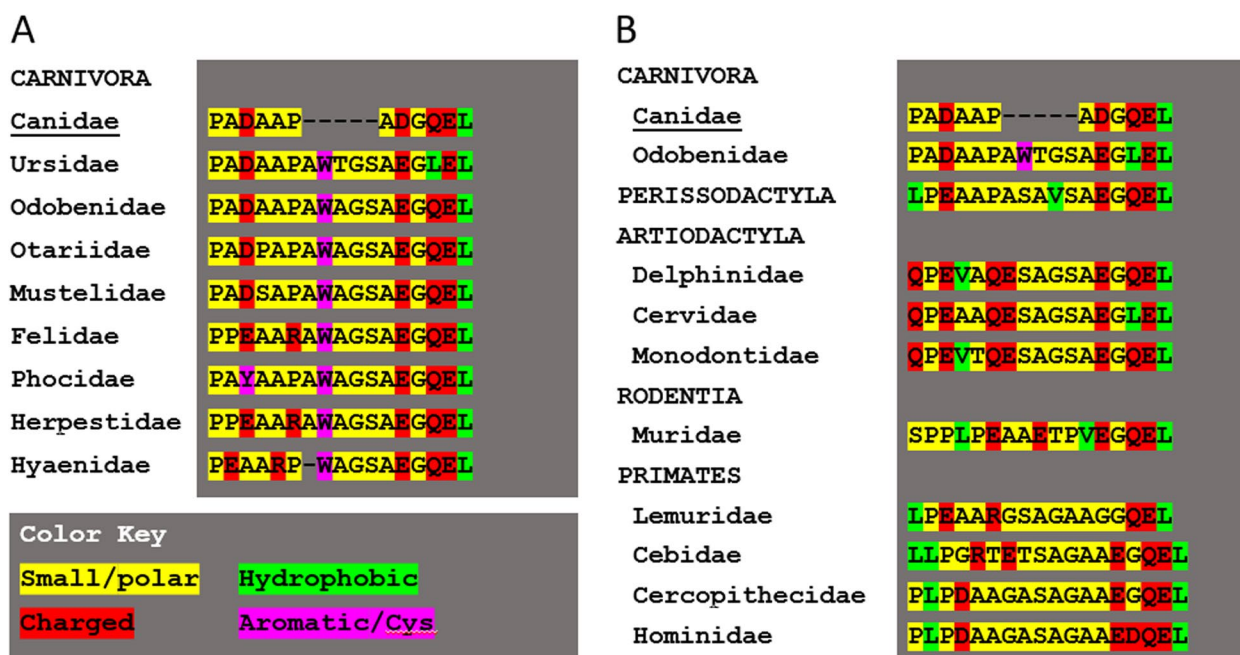


Fig. 6 Comparison of the sequence that comprises the five amino acid gap in the N-terminal region of the TRNP1 protein that is unique to the *Canidae*. Alignment with: **A**. Other families included in the order Carnivora, **B**. Other mammalian orders. Residue size, charge and water solubility are color-coded. The protein sequences used to align with the dog sequence XP_038318375 were representative species listed in Additional file 3

It has been firmly established that polymorphisms at the regulatory regions flanking the coding sequence of a gene impact the expression and function of the resulting protein. Predictive software allowed us to investigate whether *TRNP1* variants in the three *Canis* species could affect the function of transcription factors and miRNAs that interact with the 5' end of the gene and 3' end of the transcript respectively. We focused on those regulatory factors that are expressed in the cerebral cortex, especially those that may be regulating *TRNP1* and other genes involved in cerebellar cortex morphogenesis, and neural precursor cell proliferation.

While the promoter region of a gene can extend thousands of base pairs upstream of the transcription start site, we decided to focus on the 800 nucleotides immediately upstream of the start codon. This proximal promoter region typically contains binding sites for transcription factors that are essential for the proper function for the gene. While a number of transcription binding sites were predicted to be affected by genetic polymorphisms, we focused on those that are known to be expressed in the mammalian embryonic cerebral cortex. Deletion of adenine at -240, which represented the major allele in dogs, but was present at much lower frequencies in wolves and virtually absent in coyotes, was predicted to remove binding sites for REST and ZNF423, both of which regulate neuron differentiation [7]. REST

is a transcriptional repressor that is expressed in neural progenitor cells, neurons of the prefrontal cortex, in hippocampal pyramidal neurons, dentate gyrus granule neurons and cerebellar Purkinje and granule neurons [21, 33]. As with TRNP1, the level of expression of REST regulates the migration of radial glia during neocortical development [36]. Most significantly, ChIP-seq experiments indicate that REST binds to the human *TRNP1* promoter in neural progenitor cells [19].

The 3'UTR of transcripts regulate mRNA localization, mRNA stability, and translation into protein. This study showed a high degree of sequence homology between the *TRNP1* 3'UTR across canid species, indicating the importance of conserved sites which likely bind miRNAs and RNA-binding proteins. MicroRNAs play an important role in post-transcriptional gene regulation. For example, miR-128, which is highly expressed in neurons [51], has been shown to regulate the proliferation and neurogenesis of neural precursors, as well as neuronal migration and outgrowth [8, 13, 57]. Specifically, as neurogenesis progresses, neural stem cells reduce miR-128 expression in the developing neocortex [49]. Significantly, the canine homolog cfa-miR-128-1 has been detected in canine cerebrospinal fluid [14]. The sole binding site for miR-128 at the 3'UTR for TRNP1 is predicted to be located at *367-*374 which contains a polymorphic TG repeat (c.*370_*371insTG). Complementary base

pairing is only possible if the number of TG repeats is 2 (Additional file 4). All the coyotes and foxes in this study had 3 repeats, as did all the wolves with the exception of an individual originating from the Alaskan interior with a unique $(TG)_1/(TG)_2$ genotype. Interestingly, only 35% of the dogs had the $(TG)_2$ allele which would allow miR-128 binding, with, however, 64% of this subgroup being homozygous for this variant.

The length of the TRNP1 protein varies widely across mammalian species, from 219 residues in wolves and coyotes to 233 amino acids in the tiger (*Panthera tigris*) and leopard cat (*Prionailurus bengalensis*) (Additional file 2). This variability is primarily due to a polyalanine repeat (p.Ala84-Ala88) in exon 1, encoded for by GCG repeats located in the c.250-c.264 TRNP1 sequence in the domestic dog. It is probable that these expansions and deletions arise from replication slippage and/or unequal recombination. This polyalanine repeat is located at the N-terminal region of the protein, adjacent to the predicted coiled-coil SNARE domain (located between residues 110 and 171 in the dog). SNARE (soluble N-ethylmaleimide-sensitive factor attachment protein receptor) proteins typically consist of a central SNARE motif that is linked to a disordered or low structural complexity N-terminal domain which may serve as a binding interface with other proteins [26]. Disordered regions are dynamically flexible and often initiate low-affinity transient interactions with several different proteins. The polyalanine repeat lies within a glycine/alanine-rich domain (that extends from Gly69 to Gly92 in the dog) similar to the one needed for example, for the trans-activation function of the Oct-6 transcription factor which is expressed in glial progenitor and other cells involved in neural development [40]. In general, homopolymeric tracts such as the one observed in TRNP1 are abundant in transcription factors and DNA-binding proteins and play a role in transcriptional repression and protein-protein interactions [2, 3, 28, 40]. The variable polyalanine repeat is predicted to form an alpha-helix in a hydrophobic region at the interface between a disordered region and the coiled-coil domain that comprises the SNARE motif. Comparison of dog ($n=5$) with red fox ($n=8,9,13$), as well as in the *Delphininae* (Pacific white-sided dolphin, $n=5$, vs. common bottle-nosed dolphin, $n=9$) and *Felidae* (domestic cat, $n=6$, vs. leopard cat, $n=12$) demonstrates that increasing polyalanine lengths extend the coiled-coil structure. In addition to differences in polyalanine length and composition between species, within the *Canidae* we observed that this repeat is also polymorphic within wolves, coyotes, and red foxes, but not in dogs. All dogs in the study, irrespective of breed, were observed to have a sequence of 5 alanine repeats. Wolves and coyotes have either a sequence of 3 or 5 repeats, supporting previous observations that even sequences coding for short polyamine domains can be

polymorphic [28]. Red foxes may also be subject to significant variability at this locus, as even in the small number of individuals studied, we observed 8, 9 and 13 polyalanine repeats. Conservation of this repeat region across various dog breeds may indicate that expansion or contraction in this region may significantly affect protein function. In genes involved in development, polymorphisms in these polyalanine expansions are associated with various hereditary disorders in humans [27, 41]. Even the addition of a single alanine residue can cause disease [6]. In dogs and other species in the order Carnivora, polyalanine repeats in *FOXI3* have been shown to impact ectodermal development [10], while repeats in *RUNX-2*, *TWIST* and *DIX-2* impact limb and skull morphology [12] and facial length [48]. Thus, changes in the length of these tandem repeats in canine genes regulating growth and development often have the potential to drive rapid interspecific phenotypic evolution [12, 16]. The fact that the TRNP1 five alanine repeat motif was highly conserved across the dog breeds in this study may underlie a functionally important role for this repetitive sequence.

When aligned with other species representing different families in the order Carnivora, the coding region of TRNP1 in the *Canidae* exhibits a 15-base pair gap at c.174_175, which, in the predicted protein corresponds to the position between Pro58 and Ala59 (Fig. 6), just downstream from the conserved proline-rich region in the N-terminal disordered region. In all of the 32 other carnivores for whom transcript or genomic sequences are available, this position is occupied by five amino acids (AWAGS in *Felidae*, *Mustelidae*, *Otariidae*, and *Phocidae* or AWTGS in *Ursidae* or, uniquely, four amino acids (WAGS) in the striped hyena (*Hyaena hyaena*)). Mammalian species belonging to other orders, including primates, also have amino acid sequences that bridge this 'gap'. A potentially important consequence is that canids, and wolf-like canids in particular have the shortest TRNP1 protein known to exist amongst mammalian species. This potentially significant change in the TRNP1 protein must have occurred early in the evolution of canids, as it is also present in the grey fox, whose lineage is the most primitive amongst the *Canidae* [31] (Fig. 3, Additional file 2). One has to wonder whether the evolutionary loss of five amino acids been retained throughout the *Canidae* because it has no impact on TRNP1 function, or whether it is crucial to a specific property of the said protein in this family of canid species.

Conclusions

The TRNP1 gene, expressed in several regions of the brain in domestic dogs, displays a number of potentially functionally significant genetic polymorphisms in the coding

and flanking regulatory regions. While some of these variants are shared with other canid species, others are species-specific. A polyalanine repeat, which is predicted to affect the formation of alpha helices and coiled-coils, was found to be polymorphic in wolves, coyotes and red foxes, but not in the dog breeds investigated. Compared to other carnivores and mammalian species in general, the *Canidae* species all have a shorter TRNP1 protein due to a five-amino acid 'gap' in the N-terminal domain.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40575-023-00133-0>.

Additional file 1. Details of individual pure-bred dogs, wolves, coyotes and foxes used in the study.

Additional file 2. Accession numbers for *Canis lupus familiaris* and *Vulpes vulpes* polymorphisms found in the study.

Additional file 3.

Additional file 4. Mirtarget prediction of binding of cfa-miR-128 To dog TRNP1 3'UTR with c*370_371insAG ALLELE.

Acknowledgements

The authors would like to express their thanks to all the private dog owners that consented to have their dogs participate in this study, as well as the Des Moines Training & Obedience Club, and the Des Moines Kennel Club. Special thanks also to the following individuals and wild canid education centers for providing DNA samples from wolves, coyotes and foxes: Jami Hammer, Wolf Park, Colorado Wolf and Wildlife Center, JAB Canid Education and Conservation Center, Shy Wolf Sanctuary, and Blanke Park Zoo. Anastasia Yablochkin, Claire Steinbronn, and Sarah Mann contributed to some of the experiments in this study.

Authors' contributions

JS designed the study and wrote the majority of the manuscript. ES, RD and AW performed the experiments. JS and MS carried out the data analysis. All authors read and approved the final manuscript.

Funding

Drake College of Pharmacy and Health Science Harris Endowment Grant (J.C.S). Des Moines University IOER Grant (M.A.S).

Availability of data and materials

All polymorphism locus and population data for canids with GenBank reference genomes, that is, dog (GCA_011100685.1) and red fox (GCA_003160815.1), were deposited in the European Variation archive <https://www.ebi.ac.uk/eva/>, project numbers PRJEB55868 (dog), PRJEB55814 (red fox). The TRNP1 partial transcript sequences were submitted to GenBank <https://www.ncbi.nlm.nih.gov/genbank/> and designated by accession numbers OQ191934 and OQ191935.

Declarations

Ethics approval and consent to participate

These experiments did not require IACUC approval since no animal handling or use was needed since saliva samples had been collected for an earlier study.

Consent for publication

All contributors have provided written consent to be listed as authors.

Competing interests

None.

Author details

¹Ellis Pharmacogenomics Laboratory, College of Pharmacy and Health Sciences, Drake University, 50311 Des Moines, IA, USA. ²Department of Anatomy, Des Moines University, 50266 Des Moines, IA, USA.

Received: 6 July 2023 Accepted: 27 October 2023

Published online: 15 November 2023

References

- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248–9. <https://doi.org/10.1038/nmeth0410-248>.
- Alba MM, Guigo R. Comparative analysis of amino acid repeats in rodents and humans. *Genome Res*. 2004;14(4):549–54. <https://doi.org/10.1101/gr.1925704>.
- Amiel J, Trochet D, Clement-Ziza M, Munnich A, Lyonnet S. Polyalanine expansions in human. *Hum Mol Genet*. 2004;13 Spec No 2:R235–243. <https://doi.org/10.1093/hmg/ddh251>.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25(1):25–9. <https://doi.org/10.1038/75556>.
- Bae B, Miura P. Emerging roles for 3'UTRs in neurons. *Int J Mol Sci*. 2020;21(10):3413. <https://doi.org/10.3390/ijms21103413>.
- Brais B, Bouchard JP, Xie YG, Rochefort DL, Chretien N, Tome FM, Lafreniere RG, Rommens JM, Uyama E, Nohira O, Blumen S, Korczyn AD, Heutink P, Mathieu J, Duranceau A, Codere F, Fardeau M, Rouleau GA. Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat Genet*. 1998;18(2):164–7. <https://doi.org/10.1038/ng0298-164>.
- Casoni F, Croci L, Bosone C, D'Ambrosio R, Badaloni A, Gaudesi D, Barili V, Sarna JR, Tessarollo L, Cremona O, Hawkes R, Warming S, Consalez GG. Zfp423/ZNF423 regulates cell cycle progression, the mode of cell division and the DNA-damage response in Purkinje neuron progenitors. *Development*. 2017;144(20):3686–97. <https://doi.org/10.1242/dev.155077>.
- Cernilogar FM, Di Giaimo R, Rehfeld F, Cappello S, Lie DC. RNA interference machinery-mediated gene regulation in mouse adult neural stem cells. *BMC Neurosci*. 2015;16:60. <https://doi.org/10.1186/s12868-015-0198-7>.
- Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. *Nucleic Acids Res*. 2020;48(D1):D127–31. <https://doi.org/10.1093/nar/gkz757>.
- Drogemuller C, Karlsson EK, Hytonen MK, Perloski M, Dolf G, Sainio K, Lohi H, Lindblad-Toh K, Leeb T. A mutation in hairless dogs implicates FOXI3 in ectodermal development. *Science*. 2008;321(5895):1462. <https://doi.org/10.1126/science.1162525>.
- Esglesas M, Falk S, Forne I, Thiry M, Najas S, Zhang S, Mas-Sanchez A, Geerlof A, Niessing D, Wang Z, Imhof A, Gotz M. Trnp1 organizes diverse nuclear membrane-less compartments in neural stem cells. *EMBO J*. 2020;39(16):e103373. <https://doi.org/10.15252/embj.2019103373>.
- Fondon JW 3, Garner HR. Molecular origins of rapid and continuous morphological evolution. *Proc Natl Acad Sci U S A*. 2004;101(52):18058–63. <https://doi.org/10.1073/pnas.0408118101>.
- Franzoni E, Booker SA, Parthasarathy S, Rehfeld F, Grosser S, Srivatsa S, Fuchs HR, Tarabykin V, Vida I, Wulczyn FG. miR-128 regulates neuronal migration, outgrowth and intrinsic excitability via the intellectual disability gene Phf6. *Elife*. 2015;4:e04263. <https://doi.org/10.7554/eLife.04263>.
- Friedlander MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, Rajewsky N. Discovering microRNAs from deep sequencing data using miRDeep. *Nat Biotechnol*. 2008;26(4):407–15. <https://doi.org/10.1038/nbt1394>.
- Gabler F, Nam SZ, Till S, Mirdita M, Steinegger M, Soding J, Lupas AN, Alva V. Protein sequence analysis using the MPI bioinformatics toolkit. *Curr Protoc Bioinform*. 2020;72(1):e108. <https://doi.org/10.1002/cpbi.108>.
- Gemayel R, Vincens MD, Legendre M, Verstrepen KJ. Variable tandem repeats accelerate evolution of coding and regulatory sequences.

- Annu Rev Genet. 2010;44:445–77. <https://doi.org/10.1146/annurev-genet-072610-155046>.
17. Gene Ontology, C. The gene ontology resource: enriching a GOLD mine. *Nucleic Acids Res.* 2021;49(D1):D325–34. <https://doi.org/10.1093/nar/gkaa1113>.
 18. Grewal JS, Gloe T, Hegedus J, Bitterman K, Billings BK, Chengetanai S, Bentil S, Wang VX, Ng JC, Tang CY, Geletta S, Wicinski B, Bertelson M, Tendler BC, Mars RB, Aguirre GK, Rusbridge C, Hof PR, Sherwood CC, Spocter MA. Brain gyrfication in wild and domestic canids: has domestication changed the gyrfication index in domestic dogs? *J Comp Neurol.* 2020;528(18):3209–28. <https://doi.org/10.1002/cne.24972>.
 19. Hammal F, de Langen P, Bergon A, Lopez F, Ballester B. ReMap 2022: a database of human, mouse, drosophila and Arabidopsis regulatory regions from an integrative analysis of DNA-binding sequencing experiments. *Nucleic Acids Res.* 2022;50(D1):D316–325. <https://doi.org/10.1093/nar/gkab996>.
 20. Hoepfner MP, Lundquist A, Pirun M, Meadows JR, Zamani N, Johnson J, Sundstrom G, Cook A, FitzGerald MG, Swofford R, Mauceli E, Moghadam BT, Greka A, Alfoldi J, Abouelleil A, Aftuck L, Bessette D, Berlin A, Brown A, et al. An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. *PLoS One.* 2014;9(3):e91172. <https://doi.org/10.1371/journal.pone.0091172>.
 21. Huang Z, Wu Q, Guryanova OA, Cheng L, Shou W, Rich JN, Bao S. Deubiquitylase HAUSP stabilizes REST and promotes maintenance of neural progenitor cells. *Nat Cell Biol.* 2011;13(2):142–52. <https://doi.org/10.1038/ncb2153>.
 22. Kaas JH. The evolution of brains from early mammals to humans. *Wiley Interdiscip Rev Cogn Sci.* 2013;4(1):33–45. <https://doi.org/10.1002/wcs.1206>.
 23. Kelava I, Lewitus E, Huttner WB. The secondary loss of gyrencephaly as an example of evolutionary phenotypical reversal. *Front Neuroanat.* 2013;7:16. <https://doi.org/10.3389/fnana.2013.00016>.
 24. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* 2015;10(6):845–58. <https://doi.org/10.1038/nprot.2015.053>.
 25. Kent WJ. BLAT—the BLAST-like alignment tool. *Genome Res.* 2002;12(4):656–64. <https://doi.org/10.1101/gr.229202>.
 26. Khvotchev M, Soloviev M. SNARE modulators and SNARE mimetic peptides. *Biomolecules.* 2022;12(12):1779. <https://doi.org/10.3390/biom12121779>.
 27. Laumonier F, Ronce N, Hamel BC, Thomas P, Lespinasse J, Raynaud M, Paringaux C, Van Bokhoven H, Kalscheuer V, Frysns JP, Chelly J, Moraine C, Briault S. Transcription factor SOX3 is involved in X-linked mental retardation with growth hormone deficiency. *Am J Hum Genet.* 2002;71(6):1450–5. <https://doi.org/10.1086/344661>.
 28. Lavoie H, Debeane F, Trinh QD, Turcotte JF, Corbeil-Girard LP, Dicaire MJ, Saint-Denis A, Page M, Rouleau GA, Brais B. Polymorphism, shared functions and convergent evolution of genes with sequences coding for polyaniline domains. *Hum Mol Genet.* 2003;12(22):2967–79. <https://doi.org/10.1093/hmg/ddg329>.
 29. Lee C, Huang CH. LASAGNA: a novel algorithm for transcription factor binding site alignment. *BMC Bioinformatics.* 2013;14:108. <https://doi.org/10.1186/1471-2105-14-108>.
 30. Lee C, Huang CH. LASAGNA-Search 2.0: integrated transcription factor binding site search and visualization in a browser. *Bioinformatics.* 2014;30(13):1923–5. <https://doi.org/10.1093/bioinformatics/btu115>.
 31. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, Clamp M, Chang JL, Kulbokas EJ 3, Zody MC, Mauceli E, Xie X, Breen M, Wayne RK, Ostrander EA, Ponting CP, Galibert F, Smith DR, DeJong PJ, Lander ES. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature.* 2005;438(7069):803–19. <https://doi.org/10.1038/nature04338>.
 32. Liu W, Wang X. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. *Genome Biol.* 2019;20(1):18. <https://doi.org/10.1186/s13059-019-1629-z>.
 33. Lu T, Aron L, Zullo J, Pan Y, Kim H, Chen Y, Yang TH, Kim HM, Drake D, Liu XS, Bennett DA, Colaiacovo MP, Yankner BA. REST and stress resistance in ageing and Alzheimer’s disease. *Nature.* 2014;507(7493):448–54. <https://doi.org/10.1038/nature13163>.
 34. Ludwiczak J, Winski A, Szczepaniak K, Alva V, Dunin-Horkawicz S. DeepCoil—a fast and accurate prediction of coiled-coil domains in protein sequences. *Bioinformatics.* 2019;35(16):2790–5. <https://doi.org/10.1093/bioinformatics/bty1062>.
 35. Lupas A, Van Dyke M, Stock J. Predicting coiled coils from protein sequences. *Science.* 1991;252(5009):1162–4. <https://doi.org/10.1126/science.252.5009.1162>.
 36. Mandel G, Fiondella CG, Covey MV, Lu DD, Loturco JJ, Ballas N. Repressor element 1 silencing transcription factor (REST) controls radial migration and temporal neuronal specification during neocortical development. *Proc Natl Acad Sci U S A.* 2011;108(40):16789–94. <https://doi.org/10.1073/pnas.1113486108>.
 37. Manger PR, Prowse M, Haagensen M, Hemingway J. Quantitative analysis of neocortical gyrencephaly in African elephants (*Loxodonta africana*) and six species of cetaceans: comparison with other mammals. *J Comp Neurol.* 2012;520(11):2430–9. <https://doi.org/10.1002/cne.23046>.
 38. Mathe E, Olivier M, Kato S, Ishioka C, Hainaut P, Tavtigian SV. Computational approaches for predicting the biological effect of p53 missense mutations: a comparison of three sequence analysis based methods. *Nucleic Acids Res.* 2006;34(5):1317–25. <https://doi.org/10.1093/nar/gkj518>.
 39. McGeary SE, Lin KS, Shi CY, Pham TM, Bisaria N, Kelley GM, Bartel DP. The biochemical basis of microRNA targeting efficacy. *Science.* 2019;366(6472):eaav1741. <https://doi.org/10.1126/science.aav1741>.
 40. Meijer D, Graus A, Grosveld G. Mapping the transactivation domain of the Oct-6 POU transcription factor. *Nucleic Acids Res.* 1992;20(9):2241–7. <https://doi.org/10.1093/nar/20.9.2241>.
 41. Muragaki Y, Mundlos S, Upton J, Olsen BR. Altered growth and branching patterns in synpolydactyly caused by mutations in HOXD13. *Science.* 1996;272(5261):548–51. <https://doi.org/10.1126/science.272.5261.548>.
 42. Panwar B, Omenn GS, Guan Y. miRmine: a database of human miRNA expression profiles. *Bioinformatics.* 2017;33(10):1554–60. <https://doi.org/10.1093/bioinformatics/btx019>.
 43. Pillay P, Manger PR. Order-specific quantitative patterns of cortical gyrfication. *Eur J Neurosci.* 2007;25(9):2705–12. <https://doi.org/10.1111/j.1460-9568.2007.05524.x>.
 44. Roberts T, McGreevy P, Valenzuela M. Human induced rotation and reorganization of the brain of domestic dogs. *PLoS One.* 2010;5(7):e11946. <https://doi.org/10.1371/journal.pone.0011946>.
 45. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol.* 1993;234(3):779–815. <https://doi.org/10.1006/jmbi.1993.1626>.
 46. Schroeder DJ, Blair JD, Lott P, Yu HO, Hong D, Cray F, Ashwood P, Walker C, Korf I, Robinson WP, LaSalle JM. The human placenta methylome. *Proc Natl Acad Sci U S A.* 2013;110(15):6037–42. <https://doi.org/10.1073/pnas.1215145110>.
 47. Schroeder DJ, Jayashankar K, Douglas KC, Thirkill TL, York D, Dickinson PJ, Williams LE, Samollow PB, Ross PJ, Bannasch DL, Douglas GC, LaSalle JM. Early developmental and evolutionary origins of gene body DNA methylation patterns in mammalian placentas. *PLoS Genet.* 2015;11(8):e1005442. <https://doi.org/10.1371/journal.pgen.1005442>.
 48. Sears KE, Goswami A, Flynn JJ, Niswander LA. The correlated evolution of Runx2 tandem repeats, transcriptional activity, and facial length in carnivora. *Evol Dev.* 2007;9(6):555–65. <https://doi.org/10.1111/j.1525-142X.2007.00196.x>.
 49. Shu P, Wu C, Ruan X, Liu W, Hou L, Fu H, Wang M, Liu C, Zeng Y, Chen P, Yin B, Yuan J, Qiang B, Peng X, Zhong W. Opposing gradients of MicroRNA expression temporally pattern layer formation in the developing neocortex. *Dev Cell.* 2019;49(5):764–785e764. <https://doi.org/10.1016/j.devcel.2019.04.017>.
 50. Sievers F, Higgins DG. The clustal omega multiple alignment package. *Methods Mol Biol.* 2021;2231:3–16. https://doi.org/10.1007/978-1-0716-1036-7_1.
 51. Smirnova L, Grafe A, Seiler A, Schumacher S, Nitsch R, Wulczyn FG. Regulation of miRNA expression during neural cell specification. *Eur J Neurosci.* 2005;21(6):1469–77. <https://doi.org/10.1111/j.1460-9568.2005.03978.x>.
 52. Stahl R, Walcher T, De Juan Romero C, Pilz GA, Cappello S, Irmiler M, Sanz-Aquila JM, Beckers J, Blum R, Borrell V, Gotz M. Trnp1 regulates expansion and folding of the mammalian cerebral cortex by control of radial glial fate. *Cell.* 2013;153(3):535–49. <https://doi.org/10.1016/j.cell.2013.03.027>.
 53. UniProt, C. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* 2021;49(D1):D480–9. <https://doi.org/10.1093/nar/gkaa1100>.
 54. Volpe M, Shpungin S, Barbi C, Abraham G, Malovani H, Wides R, Nir U. Trnp: a conserved mammalian gene encoding a nuclear protein that accelerates cell-cycle progression. *DNA Cell Biol.* 2006;25(6):331–9. <https://doi.org/10.1089/dna.2006.25.331>.

55. Wayne RK. Cranial morphology of domestic and wild canids: the influence of development on morphological change. *Evolution*. 1986;40(2):243–61. <https://doi.org/10.1111/j.1558-5646.1986.tb00467.x>.
56. Wosinski M, Schleicher A, Zilles K. Quantitative analysis of gyrification of cerebral cortex in dogs. *Neurobiology (Bp)*. 1996;4(4):441–68.
57. Zhang W, Kim PJ, Chen Z, Lokman H, Qiu L, Zhang K, Rozen SG, Tan EK, Je HS, Zeng L. MiRNA-128 regulates the proliferation and neurogenesis of neural precursors by targeting PCM1 in the developing cortex. *Elife*. 2016;5:e11324. <https://doi.org/10.7554/eLife.11324>.
58. Zilles K, Palomero-Gallagher N, Amunts K. Development of cortical folding during evolution and ontogeny. *Trends Neurosci*. 2013;36(5):275–84. <https://doi.org/10.1016/j.tins.2013.01.006>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

