Hum Genet (2001) 109 : 569–575
DOI 10.1007/s004390100607

ORIGINAL INVESTIGATION

**Michel Guipponi · Caroline Tapparel · Olivier Jousson
Nathalie Scamuffa · Christophe Mas · Colette Rossier
Pierre Hutter · Paolo Meda · Robert Lyle
Alexandre Reymond · Stylianos E. Antonarakis**

The murine orthologue of the Golgi-localized TPTE protein provides clues to the evolutionary history of the human TPTE gene family

Received: 5 June 2001 / Accepted: 17 August 2001 / Published online: 27 October 2001

Abstract The human *TPTE* gene encodes a testis-specific protein that contains four potential transmembrane domains and a protein tyrosine phosphatase motif, and shows homology to the tumor suppressor PTEN/MMAC1. Chromosomal mapping revealed multiple copies of the *TPTE* gene present on the acrocentric chromosomes 13, 15, 21 and 22, and the Y chromosome. Zooblot analysis suggests that mice may possess only one copy of *TPTE*. In the present study, we report the isolation and initial characterization of the full-length cDNA of the mouse homologue *Tpte*. At least three different mRNA transcripts (*Tpte.a, b, c*) are produced via alternative splicing, encoding predicted proteins that would contain four potential transmembrane domains and a protein tyrosine phosphatase motif. Transfection of a 5′EGFP-TPTE fusion protein in Hela cells revealed an intracellular localization within the Golgi apparatus. *Tpte* was mapped by radiation hybrid to a region of mouse chromosome 8 that shows conserved synteny with human 13q14.2-q21 between *NEK3* and *SGT1*. This region of the human genome was found to contain a partial, highly diverged copy of *TPTE* that is likely to represent the ancestral copy from which the other copies of *TPTE* arose through duplication events. The Y chromosome copy of *TPTE* is a pseudogene and is not therefore involved in the testis expression of this gene family.

M. Guipponi · C. Tapparel · N. Scamuffa · C. Rossier · R. Lyle
A. Reymond · S.E. Antonarakis (✉)
Division of Medical Genetics,
University of Geneva Medical School,
1 rue Michel Servet, 1211 Geneva 4, Switzerland
e-mail: stylianos.antonarakis@medecine.unige.ch,
Tel.: +41-22-7025708, Fax: +41-22-7025706

C. Tapparel · P. Hutter
Unit of Genetics, Institut Central des Hôpitaux Valaisans,
Sion, Switzerland

O. Jousson
Zoology Station, Department of Zoology and Animal Biology,
University of Geneva, 1224 Chêne-Bougeries, Switzerland

C. Mas · P. Meda
Department of Morphology,
University of Geneva Medical School,
1 rue Michel Servet, 1211 Geneva 4, Switzerland

Introduction

We have recently identified a testis-specific cDNA, *TPTE* (Transmembrane Phosphatase with TEnsin homology), that encodes a predicted protein of 551 amino acids containing four potential transmembrane domains and a tyrosine phosphatase motif (Chen H et al. 1999; Guipponi et al. 2000). *TPTE* showed significant homology to the tumor suppressor gene PTEN/MMAC1 (Simpson and Parsons 2001). Protein tyrosine phosphorylation is an essential component in intracellular signaling, with functions pertinent to cell proliferation, differentiation, migration and attachment.

TPTE is present in multiple copies in the human genome, including the acrocentric chromosomes 13, 15, 21 and 22, and the Y chromosome. In addition, several of these chromosomes contain more than one copy of *TPTE* (Chen H et al. 1999). It is not known whether all of these copies are functional, although the testis-specific expression led to the hypothesis that the active *TPTE* copy might be on the Y chromosome. The presence of two nucleotide substitutions in the 5′UTR of *TPTE* cDNAs indicated that different alleles of the *TPTE* gene are expressed (Chen H et al. 1999), although these polymorphisms could either represent variations between individuals or sequence differences between multiple expressed loci. It is therefore not clear how many *TPTE* loci are transcribed. Analysis of the human genome sequence is needed to determine the number of genes and pseudogenes in the *TPTE* family, and provide data which can be used to determine which *TPTE* locus is transcriptionally active.

Analysis of *TPTE* by Southern blot in diverse species (primate, rat, mouse, dog, cow, and rabbit) has suggested that the *TPTE* gene is conserved and that the mouse genome might contain only one copy of *TPTE* (Chen H et al. 1999). The chromosomal location of the mouse gene

may provide evidence for the location of the active *TPTE* in humans.

Here, we report: (a) the isolation of the mouse *Tpte* and its mapping to chromosome 8; (b) the sequence analysis of the homologous region of the human genome and a hypothesis for the evolutionary history of the *TPTE* gene family; (c) the analysis of the human chromosome Y genomic sequence; (d) the subcellular localization of TPTE in the Golgi apparatus.

Materials and methods

Isolation and mapping of the mouse *Tpte* gene

In order to identify the mouse *Tpte* gene, two degenerate primers were designed (5′-GARMGYATHATHGCNATG-3′/5′-XARRT-CYTGXGGCCATCCA-3′) based on regions of the human *TPTE* gene sequence homologous to other phosphatases/phosphatase domains. Double-stranded cDNA was synthesized using 1 μg of mouse testis poly(A)+ RNA (Clontech) and Superscript II reverse transcriptase (Life Technologies). PCR amplification was carried out using 1 μl of the mouse testis cDNA synthesis reaction, 1 μg of each degenerate primer and Taq Start Antibody (Clontech). The expected 262-bp PCR fragment was cloned (TOPO TA cloning, Invitrogen), sequenced and used as a probe to screen 6×10^5 clones from a mouse testis cDNA library (5′STRETCH PLUS cDNA library; Clontech). Hybridization, washing and autoradiography were carried out as described (Clontech). PCR amplifications were performed to amplify the positive phage inserts using primers from the cloning vector λTRIPLEX (5′-CTCGGGAAGCGCGCCCA-TTGTGTTGGT-3′/5′-ATCGACTCACTATAGGGCGAATTGG-CC-3′) and gene-specific primers. The PCR products were purified using Qiagen purification columns and sequenced. Mouse *Tpte* was mapped by PCR using the T31 whole genome radiation hybrid (WGRH) panel (Research Genetics, Huntsville, Ala.) and a pair of selected primers (5′-TTTGGTGAGAACTAGCTGTTGAG-3′/5′-TATACATTTAATATACTTCCCAG-3′). PCR amplification data were submitted to The Jackson Laboratory for radiation hybrid mapping.

Phylogenetic analysis of the nucleotide sequence of the human *TPTEs* and orthologues

The nucleotide sequence of the different human *TPTE* copies was identified using the Human Genome BLAST option available at the NCBI (http://www.ncbi.nlm.nih.gov). The genomic structure of the different copies was determined using the EST-GENOME tool available at HGMP (http://www.hgmp.mrc.ac.uk). BLAST searches against non-redundant EST databases identified two EST clones derived from zebrafish and Xenopus *Tpte* (IMAGE no. 2602067 and GenBank no. AW643813, respectively).

The largest common region of the human, mouse, zebrafish and *Xenopus TPTE* coding sequences were aligned manually using GDE 2.2 (Larsen et al. 1993; the complete alignment is available upon request). Human genomic regions HC13 (NT_009888) and HC13 (NT_024539) were excluded from this analysis because they are incomplete working draft sequences. The alignments were analyzed using the following methods: the neighbor-joining (NJ) method (Saitou and Nei 1987), the maximum parsimony (MP) (Fitch 1971) method, and the maximum likelihood (ML) (Felsenstein 1981) method. Phylogenetic analyses were performed using PAUP* (Swofford 1998). For the MP method, analyses were conducted using unweighted characters. Heuristic search parameters included ten replicates of random addition of sequences and tree-bisection reconnection (TBR) branch swapping. For the ML and NJ analyses, likelihood scores were determined on an initial MP tree for the following models of sequence evolution: Jukes-Cantor (JC) (Jukes and Cantor 1969), Kimura 2-parameter (K2P) (Kimura 1980), Hasegawa, Kishino and Yano (HKY85) (Hasegawa et al. 1985), and general time-reversible (GTR) (Yang et al. 1994). Additionally, the incorporation into each of the models of among-site rate heterogeneity using the Γ-distribution (Yang 1993) was examined. Likelihood ratio tests were used to compare likelihood scores obtained for pairs of nested models to determine which model best fit the sequence data (Yang et al. 1994). The reliability of internal branches was assessed using the bootstrap method (Felsenstein 1988), with 1,000 replicates.

Expression pattern studies

RT-PCR with primers 5′-ATGTCATTTCCATCGTCTGGAAG-3′ and 5′-CTTGGAACTTCGAGCTGTTGC-3′ was performed on a cDNA panel containing 16 different double- stranded cDNAs from mouse tissues to obtain a semi-quantitative estimation of gene expression as described (Michaud et al. 2000). Northern blot analysis was performed on a filter containing 20 μg of total RNA extracted from mouse testis and epididymis using TRIZOL (GIBCO BRL). The radioactive probe used was generated by RT-PCR between primers 5′-ATGTCATTTCCATCGTCTGGAAG-3′ and 5′-CTT-GGAACTTCGAGCTGTTGC-3′.

Subcellular localization of *TPTE*

Hela cells were transfected with a 5′EGFP-TPTE expression vector using FuGENETM (Life Technologies) as described by Reymond et al. (2001). Twenty-four hours after transfection, EGFP fluorescence was detected on 5% PFA-fixed cells. Images were analyzed with Adobe Photoshop software. The anti-giantin (324450, Calbiochem) compartment-specific marker was used in the co-localization experiments. A pEGFP vector without insert was expressed in cells as a control.

Results and discussion

Isolation and mapping of the mouse *Tpte* gene

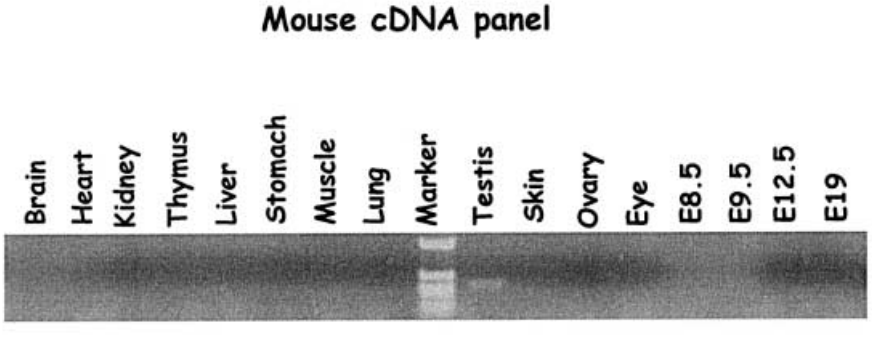
To isolate the full-length cDNA for the mouse homologue of *TPTE*, we first designed degenerate primers in regions where the protein sequences of several tyrosine protein phosphatases, including *TPTE*, were highly homologous. A 262-bp RT-PCR fragment was amplified and cloned. Five clones were sequenced and were found identical. TBLASTN analysis against non-redundant protein databases revealed that this mouse sequence showed high homology to the human *TPTE*, higher than to other mouse tyrosine phosphatase proteins. Screening of a mouse testis

Fig. 1 **A** Sequences alignment of the predicted human *TPTE* ▶ (Genbank No. NM_0133159), the mouse Tpte.a protein, the zebrafish ESTs (GenBank Nos. fi24b06.x1 and fi24b06.y1) and the *Xenopus* EST (GenBank no. AW643813). BLAST search against EST databases revealed the existence of a zebrafish and *Xenopus* orthologues. The zebrafish gene was found to be unique and mapped to linkage group 10 (Woods et al. 2001). Identical or conserved amino acids in all four sequences are shaded in *black*. Identical or conserved amino acids in three of the four proteins are shaded in *grey*. The four potential transmembrane domains and the protein tyrosine phosphatase domain are underlined and named underneath. The tyrosine phosphatase motif signature within the protein tyrosine phosphatase domain is *boxed*. **B**, **C** Expression pattern of mouse *Tpte*. RT-PCR (**B**) and Northern blot (**C**) analysis showed that the mouse *Tpte* is only expressed in testis

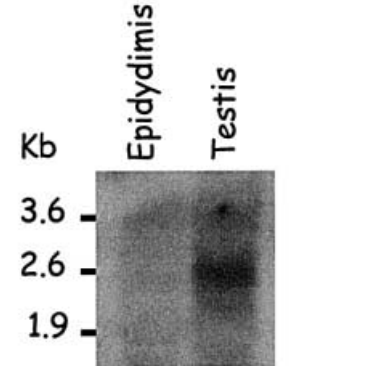
A.

```
mTPTE : MYGEKKSHLYLWMEHYGYDMPANIYKMYSQPSRKTDDANKKVSVSASRTIKLNGSTGYDTNEQITLITNGSSLSYPDEIKSASYA
hTPTE : ----------------------------------------------------------------------------------------
zTPTE : ----------------------------------------------------------------------------------------
xTPTE : ----------------------------------------------------------------------------------------

mTPTE : DPISTKAYTNDSSVYDPGGASSSTTLYELNSLSEVSKEIITQGESALLRDKEATSELKIPSTLQTQTSMSTNTLSLSDLSSDYQE
hTPTE : ------------------------MNESPDPTDLAGVIIELGPN----DSPQTSEFKG----ATEEAPAKESPHTSEFKGAARV
zTPTE : ----------------------------------------------------------------------------------------
xTPTE : ----------------------------------------------------------------------------------------

mTPTE : EQMNKCKLNQMSKLYDDDERTDIQKSYWNVVKKFVGILVSSVAFRIFGIFLVILDVFLVVVDLNVSEKKIYIPLDYRSISLAIAL
hTPTE : SPISESVLARLSKFEVEDAEN--VASYDSKIKKIVHSIVSSFAFGLFGVFLVLLDVTLILADLIFTDSKLYIPLEYRSISLAIAL
zTPTE : ----------------------------------------------------------------------------------------
xTPTE : ----------------------------------------------------------TM1-------------------TM2---

mTPTE : FFLVDILLRVSVEGRRRYFSDVLNTLDAVVIGVTVVVAVIYALYDKHFLRDIPRLAVLLRPLRLLILIRILQLAHQKRQLERLTR
hTPTE : FFLMDVLLRVFVERRQQYFSDLFNILDTAIIVLLLVDVVYIFFDIKLLRNIPRWTHLLRLLRLIILLRIFHLFHQKRQLEKLIR
zTPTE : ------------------YFSSKLNIVDACIVVITLVVTMIYAFSDFSGASLIPRVVTFLRSLRILILVRIFRLASQKRELEKVTR
xTPTE : ----------TM3-------------------------------------TM4-------------------------------

mTPTE : KLVSGNKRRYKKDGFDLDLTYVTERIIAMSFPSSGRESFYRNPIKEVVRFLDTKHPNHYQVYNLCSERAYDPKHFHYRVRRIMID
hTPTE : RRVSENKRRYTRDGFDLDLTYVTERIIAMSFPSSGRQSFYRNPIKEVVRFLDKKHRNHYRVYNLCSERAYDPKHFHNRVVRIMID
zTPTE : RMVSENKRRYQKDGFDLDLTYVTERVIAMSFPSSGKQALYRNPIREVVRFLDTKHMDHYKVFNLCSEKGYDPKFFHYRVERVMID
xTPTE : ---------SARGGFDLDLTYITDRIIAMSFPSSGKQSFYRNPIEDVVRFLDSKHNDHYKIYNLCSEKGYDPKYFHYRVERVFID

mTPTE : DHNVPTLEEMLLFSKEVNNWMAQDPENVVAIHCKGGKGRTGTMVCACLIASEIVLNAKESLYFFGERRTDKSNSSKFQQGIETPSQ
hTPTE : DHNVPTLHQMVVFTKEVNEWMAQDLENIVAIHCKGGTDRTGTMVCAFLIASEICSTAKESLYYFGERRTDKTHSEKFQGVETPSQ
zTPTE : DHNVPSLDDMLRYTACVRDWMAADSRNVIAIHCKGGKGRTGTMVCTYLIDSDQFESAQESLDYFGERXTDKSMSSKFQGVETPSQ
xTPTE : DHNVFVLADMLTFTASVRAWMAYDPQNVIAIHCKGGKGRTGTMVCTYLVDSDQFESAAESLDYFGERRTDKS------------
                                                                  Protein Tyrosine Phosphatase domain

mTPTE : NRYVKYFEKLKINYQLTLPPKKVLVIKRLVVYSIHGVGKGDGSDLEVQIIMWQETVFSFCNSRNCMIFHDPETDRAIINVFHCPA
hTPTE : KRYVAYFAQVKHLYNWNLPPRRILFIKHFIIYSIPRYVR----DLKIQIEMEKKVVFSTISLGKCSVLDNITTDKILIDVFDGPP
zTPTE : SRYVGYYEIMKDQYN-----------------------------------------------------------------------
xTPTE : ----------------------------------------------------------------------------------------

mTPTE : LYDDVKVKFLSPNLPKYYDDCPFFFWFHTSFIKNNRLYLPRNELDNTHKPKTWKIYGEKFAVEVDFGEN-----------
hTPTE : LYDDVKVQFFYSNLPTYYDNCSFYFWLHTSFIENNRLYLPKNELDNLHKQKARRIYPSDFAVEILFGEKMTSSDVVAGSD
zTPTE : ----------------------------------------------------------------------------------
xTPTE : ----------------------------------------------------------------------------------
```

B.



C.



cDNA library with one of these five clones resulted in the detection of 15 positive phage clones. Sequence analysis of the longest phage clones revealed an mRNA transcript (*Tpte.a*) of 2,552-nucleotide cDNA with a 1,995-nucleotide open reading frame encoding a predicted 664-amino-acid protein (Fig. 1A). The nucleotide sequence of two further independent clones revealed the existence of two additional transcripts. Thus, *Tpte* locus expresses at least three different transcripts that are likely to be produced by alternatively spliced exons: *Tpte.a* encoding a 664-residue protein; *Tpte.b* encoding a 645-amino-acid protein and *Tpte.c* encoding a 510-residue protein (GenBank nos. AJ311311, AJ311312, AJ311313, respectively). RT-PCR experiments and sequencing of additional positive phage clones and RT-PCR products revealed that the *Tpte.a* transcript is the most frequent. A protein sequence comparison between human *TPTE* and mouse *Tpte.a* showed 61% identity (Fig. 1A). Compared to the human *TPTE* cDNA, the mouse *Tpte* cDNA was found to contain additional 5′ coding sequences. BLAST searches against the human genome did not identify any homologous human sequences, and pattern and profile searches using the ExPASy Proteomics tools (http://www.expasy.org/tools/) did not reveal the presence of additional known protein domains (Fig. 1A).

Using the mouse T31 radiation hybrid panel, the mouse *Tpte* gene was found to be linked to marker D8Mit257 with a LOD score of 15.5. The best-fit localization was between markers D8Mit257 and D8Mit143 close to *Nek3* (Chen A et al. 1999) and *Atp7b* (Reed et al. 1995). This high LOD score value and the Southern blot analysis of *TPTE* sequence in the mouse genome, which showed a single band (Chen H et al. 1999), support the hypothesis that the mouse genome might contain only one copy of *TPTE*. This region of mouse chromosome 8 shows conserved synteny to human chromosome 13q14.2-q21 (http://www.ncbi.nlm.nih.gov/Homology/view.cgi?chr=13&tax_id=9606), where *NEK3* and *ATP7B* have been mapped. Analysis of the nucleotide sequence of this region of human chromosome 13q allowed the identification of a presumably inactive copy of *TPTE* characterized by exon deletions and multiple nucleotide substitutions. This *TPTE* copy was named *TPTEps1* for *TPTE* pseudogene 1 (Fig. 2). These data indicate that the human chromosome 13q14.2-q21 *TPTE* copy is likely to be the ancestral human gene from which the different *TPTE* copies arose through duplication events that occurred after the separation of primates and rodents.

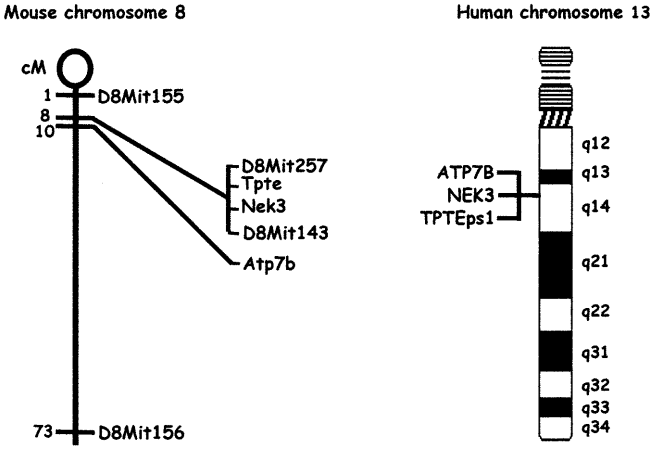


Fig. 2 Mapping of mouse *Tpte*. The human chromosome 13q14.2-q21 and its homologous region on mouse chromosome 8 are shown. Gene symbols are *NEK3*/*Nek3* (never in mitosis gene a-related kinase 3), *ATP7B*/*Atp7b* (ATPase, Cu(2+)-transporting, β polypeptide), *TPTEps1* (*TPTE* pseudogene1)/*Tpte* (transmembrane phosphatase with tensin homology). The map position is based on the NCBI human genome map (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/hum-srch) and the mouse genome database, Jackson Laboratory Bar Harbor (http://www.informatics.jax.org/menus/map_menu.shtml)

The human *TPTE* gene family

In order to investigate the evolutionary history of the *TPTE* gene family, the *TPTE* mRNA sequence (Genbank no. NM_013315) was compared to the human genomic sequence to identify all *TPTE*-related sequences. Significant homologies were found on chromosomes 13 (GenBank nos. NT_009888; NT_024514; NT_009967; NT_024539; NT_024506), 15q (GenBank no. NT_024122), 21p (GenBank no. NT_011511), 22 (GenBank no. NT_025936), and Y (GenBank no. NT_025973). A complete and non-mutated copy of *TPTE* was found on three acrocentric chromosomes: 13 (GenBank NT_024506, not precisely localized on the chromosome), 15q (GenBank NT_024122) and 21p (GenBank NT_011511). This would suggest that the active *TPTE* locus is on one of these chromosomes. In agreement with this conclusion, all *TPTE* cDNA clones (identified either by cDNA library screening or by BLAST against EST databases) showed identity only to these three loci. The other partial and mutated copies of *TPTE* identified on chromosomes 13, 22, and Y are therefore pseudogenes. However, the existence of additional complete and non-mutated copies could not be ruled out because of the incomplete sequencing of some chromosomal regions that contain *TPTE* sequences.

Due to the testis-specific expression, it has been proposed that only the copy on the Y chromosome is expressed and that the copies on the acrocentric chromosomes are inactive. A copy of *TPTE* was mapped to the long arm of HCY close to *PRY* (testis-specific PTP-BL-related protein) and *RBMY2* (RNA binding motif protein, Y chromosome, family 2, member B). Similar to the HC21p *TPTE* copy (Guipponi et al. 2000) the HCY copy is divided into 23 exons over 85 kb of genomic sequence. The sizes of the exons and introns are almost identical to those of the HC21p copy, however introns 6 and 14 are longer and intron 17 is shorter, and the untranslated exon 1 of *TPTE* was not found in the HCY genomic sequence. The translation of the reconstructed HCY *TPTE* cDNA revealed stop codons in all reading frames indicating that the copy on the Y chromosome is not transcribed.

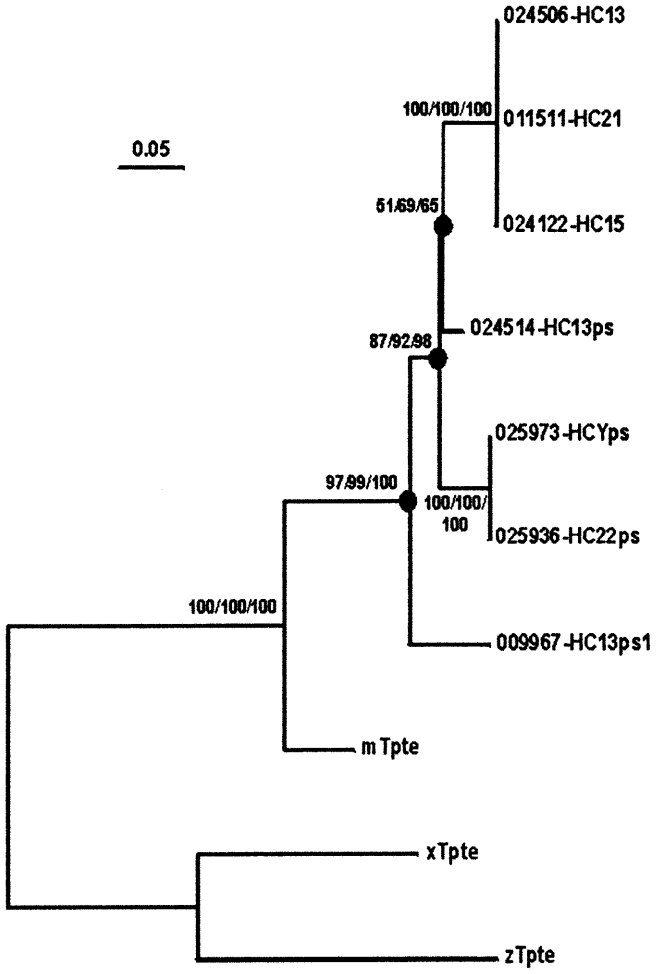


Fig. 3 Phylogenetic analysis of human and vertebrate *Tpte* sequences, inferred by a maximum likelihood (ML) analysis using the JC+Γ model. *Number* at nodes represent bootstrap values obtained with ML/MP/NJ methods (1,000 replicates). *Black circles* correspond to the putative *Tpte* gene duplication events in human. *Scale bar* Number of substitutions/site. The different human *TPTE* copies are named as follows: 011511-HC21 stands for NT_011511 *TPTE* copy on chromosome 21p; 024506-HC13 for NT_024506 *TPTE* copy on chromosome 13; 024122-HC15 for NT_024122 *TPTE* copy on chromosome 15; 024514-HC13 for NT_024514 *TPTE* pseudogene; 025936-HCY for NT_025936 *TPTE* pseudogene; 025973-HC22 for NT_025973 *TPTE* pseudogene; 009967-HC13 for NT_009967 *TPTE* pseudogene1

Phylogenetic analysis of the *TPTE* gene family

A total of nine human *TPTE* sequences were identified, in addition to single copies in mouse, *Xenopus* and zebrafish. Since not all of these sequences were complete, the longest sequence which was present in all the copies of *TPTE* was used for the analysis. The analyzed data set comprises 294 sites, of which 151 are variable (51.7%), and 89 informative (30.1%). For ML and NJ phylogenetic analyses, the comparison of likelihood scores obtained for each model of sequence evolution and the likelihood ratio test led us to choose a JC (Jukes and Cantor) +Γ model (–lnL=1336.45; α-parameter=1.0074), which imposes a single substitution rate as well as equal base frequencies. Phylogenetic analyses using the MP method produced a single most-parsimonious tree with a length of 227 steps, a consistency index (CI)=0.899, a retention index (RI)= 0.827, and a homoplasy index (HI)=0.101. The topology was identical to that obtained with NJ and ML methods (Fig. 3). *TPTE* human sequences appear to be monophyletic (100% bootstrap support) using *TPTE* homologues from *Xenopus* and *Danio rerio* (zebrafish) as an outgroup. The sequence divergence among human *TPTE* copies reaches 16.8%. The divergence between human sequences and their vertebrates homologues reaches up to 33.5%. The sequences HC13, HC15 and HC21 are identical, as are HC22ps and HCYps.

The phylogenetic analysis shows that the mouse *Tpte* sequence branches as a sister group of all human sequences, suggesting a common origin of mouse and human *TPTE* genes, compared to the outgroup taxa used (Fig. 3). Several points can be made about the evolution of the human *TPTE* family from the topology we obtained. First, the monophyletic group formed by all human sequences indicates that the human *TPTE* gene family originated from a single ancestral copy, and at least three gene radiation events occurred. Second, after the earliest radiation event, the original active copy of *TPTE* on HC13 mutated to become the pseudogene HC13ps1. Third, subsequent radiation events produced further pseudogenes and at least three possibly active copies on chromosomes 13, 15 and 21.

It should be noted, however, that the group including HC13, HC15, HC21 and HC13ps is not well supported (69/65/51 bootstrap values). This means that the branching order of HC13ps, the HC13-HC15-HC21 group, and the HCYps-HC22ps group is uncertain.

Expression profile of mouse *Tpte*

PCR of a panel of normalized mouse cDNAs was performed to study the expression profile of the *Tpte.a* transcript in a semi-quantitative manner (Michaud 2000). Expression of the mouse *Tpte* was found only in testis, revealing a similar pattern to the human gene, and not in brain, heart, kidney, thymus, liver, stomach, muscle, lung, skin, ovary, eye and embryos at different stages of development (E.8.5, 9.5, 12.5, 19) (Fig. 1B). Northern blot analysis showed a single band of approximately 2.6 kb that corresponds to the three *Tpte* transcripts identified, *Tpte.a*, *b*, c of 2,552, 2,637, and 2,675 bp, respectively (Fig. 1C). The *Tpte* transcript was found in testis but not in epididymis suggesting that *TPTE* is probably not involved in spermatozoa maturation and/or storage. In order to determine which cell types of the testis express *Tpte*, adult mouse testis sections were hybridized using different cDNA probes, including the probe used for the Northern blot analysis. Despite several different attempts, we were unable to determine the cells that express *Tpte* in the testis.

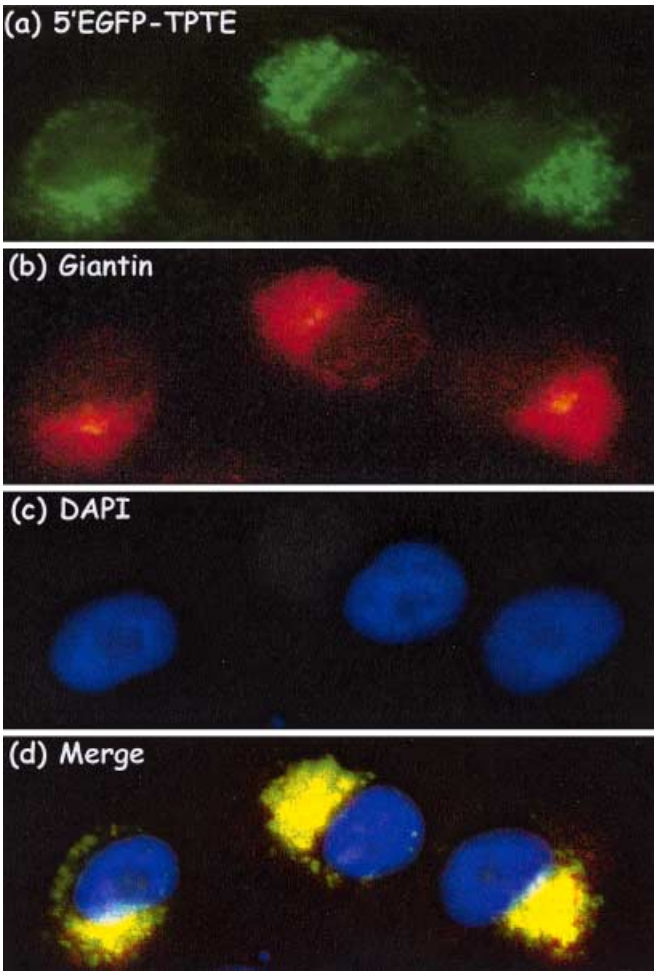


Fig. 4a–d Intracellular localization of 5′EGFP-TPTE fusion protein in the Golgi apparatus of Hela cells. Fluorescence microscopy analysis of fixed Hela cells transiently transfected with 5′EGFP-TPTE expression vector. **a** EGFP-TPTE fluorescence **b** stained using an anti-giantin antibody, **c** stained with DAPI, and **d** superimposition of panels **a**, **b** and **c**

Protein motifs and subcellular localization

Multiple forms of protein tyrosine phosphatase (PTPase) have been characterized and can be classified into two categories: soluble PTPases and transmembrane receptor proteins that contain PTPase domain(s). Structurally, all known receptor PTPases, are composed of a variable length extracellular domain, followed by a transmembrane region and a C-terminal catalytic cytoplasmic domain. The cytoplasmic region generally contains two copies of the PTPase domain. The first seems to have enzymatic activity, while the second is inactive but seems to affect substrate specificity of the first. TPTE is the only PTPase described to date with the combination of four N-terminal transmembrane domains and a C-terminal protein tyrosine phosphate domain and may represent the first member of a new subclass of PTPase.

To determine where TPTE might exert its function, a 5′EGFP-TPTE fusion protein was expressed in Hela cells and was shown to localize in a discrete vesicular and perinuclear structure (Fig. 4). Control experiments with GFP empty vectors showed a predominantly nuclear staining pattern (data not shown). To better define the compartment identified by TPTE proteins, we performed co-localization studies using compartment-specific antibodies. Signal comparison revealed that 5′EGFP-TPTE co-localized with the Golgi marker giantin, as shown in Fig. 4.

Due to its expression pattern in testis, its Golgi localization and its homology to PTEN/MMAC1, which is mutated in different sporadic cancers and in two hamartoma syndromes (Simpson and Parsons 2001), *TPTE* may be involved in tumors or male reproductive deficiencies.

Acknowledgements This work was supported by grants from the Swiss FNRS 31.57149.99, the European Union/OFES and funds from the Jérôme Lejeune fondation, the “Child Care” foundation and the University and Cantonal Hospital of Geneva to SEA and Ligue Suisse Contre le Cancer grant to AR. CT is supported by a Marie Heim-Vögtlin grant from the Swiss FNRS.

References

Chen A, Yanai A, Arama E, Kilfin G, Motro B (1999) NIMA-related kinases: isolation and characterization of murine nek3 and nek4 cDNAs, and chromosomal localization of nek1, nek2 and nek3. Gene 234:127–137

Chen H, Rossier C, Morris MA, Scott HS, Gos A, Bairoch A, Antonarakis SE (1999) A testis-specific gene, *TPTE*, encodes a putative transmembrane tyrosine phosphatase and maps to the pericentromeric region of human chromosomes 21 and 13, and to chromosomes 15, 22, and Y. Hum Genet 105:399–409

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17:268–376

Felsenstein J (1988) Phylogenies from molecular sequences: inference and reliability. Ann Rev Gen 22:521–565

Fitch WM (1971) Toward defining the course of evolution: minimum change for a specific tree topology. Syst Zool 20:406–416

Guipponi M, Yaspo ML, Riesselman L, Chen H, De Sario A, Roizes G, Antonarakis SE (2000) Genomic structure of a copy of the human TPTE gene which encompasses 87 kb on the short arm of chromosome 21. Hum Genet 107:127–131

Hasegawa M, Kishino H, Yano TA (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) Mammalian protein metabolism III. Academic, New York pp 21–132

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Larsen N, Osen GJ, Maidak BL, Mc Caughey MJ, Overbeek R, Macke TJ, Marsh TL, Woese CR (1993) The ribosomal database project. Nucleic Acids Res 21:3021–3023

Michaud J, Kudoh J, Berry A, Bonne-Tamir B, Lalioti MD, Rossier C, Shibuya K, Kawasaki K, Asakawa S, Minoshima S, Shimizu N, Antonarakis SE, Scott HS (2000) Isolation and characterization of a human chromosome 21q22.3 gene (WDR4) and its mouse homologue that code for a WD-repeat protein. Genomics 15:71–79

Reed V, Williamson P, Bull PC, Cox DW, Boyd Y (1995) Mapping of the mouse homologue of the Wilson disease gene to mouse chromosome 8. Genomics 10:573–575

Reymond A, Meroni G, Fantozzi A, Merla G, Cairo S, Luzi L, Riganelli D, Zanaria E, Messali S, Cainarca S, Guffanti A, Minucci S, Pelicci PG, Ballabio A (2001) The tripartite motif family identifies cell compartments. EMBO J 20:2140–2151

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Simpson L, Parsons R (2001) PTEN: life as a tumor suppressor. Exp Cell Res 264:29–41

Swofford DL (1998) PAUP*: phylogenetic analysis using parsimony, version 4.0b2a. Sinauer Associates, Sunderland, Massachusetts

Woods IG, Kelly PD, Chu F, Ngo-Hazelett P, Yan YL, Huang H, Postlethwait JH, Talbot WS (2000) A comparative map of the zebrafish genome. Genome Res 10:1903–1914

Yang Z (1993) Maximum likelihood estimation of phylogeny from DNA sequences when substitutions rates differ over sites. Mol Biol Evol 10:1396–1401

Yang Z, Goldman N, Friday A (1994) Comparison of models for nucleotide substitution used in maximum likelihood phylogenetic estimation. Mol Biol Evol 11:316–324