LETTER TO JMG

Screening of *TCOF1* in patients from different populations: confirmation of mutational hot spots and identification of a novel missense mutation that suggests an important functional domain in the protein treacle

A Splendore, E W Jabs, M R Passos-Bueno

J Med Genet 2002;**39**:493–495

Treacher Collins syndrome (TCS, MIM *154500) is an autosomal dominant craniofacial disorder characterised by malar hypoplasia, micrognathia, downward slanting palpebral fissures, lower eyelid coloboma, malformed auricles, conductive deafness, and cleft palate. The estimated incidence is 1/50 000, with 60% of the cases resulting from sporadic mutations.[1] There is marked phenotypic variability among patients, ranging from perinatal death because of a compromised airway to those that go undetected by medical examination. The gene underlying this condition, *TCOF1*, was mapped in 1996[2] and since then mutation detection studies have concluded that: (1) the majority of pathogenic mutations are small deletions and insertions causing frameshifts that are predicted to result in a truncated protein; (2) mutations (both polymorphic and pathogenic) can be found throughout the 25 coding exons of the gene; (3) most mutations are family specific with the exception of a commonly occurring 5 bp deletion in exon 24 (found in approximately 16% of families); and (4) there is no correlation between type and/or localisation of the mutation and phenotypic expression.[3–6] Furthermore, the observation that more than 50% of all described pathogenic mutations known to date are clustered in five exons (10, 15, 16, 23, and 24) has led to the hypothesis that these five exons are mutational hot spots, suggesting that any effort to identify mutations in *TCOF1* would benefit from testing these five exons before extending the analysis to the rest of the gene.[6]

The protein encoded by the *TCOF1* gene, treacle, has repetitive motifs encoded by exons 7-16, which are subject to phosphorylation by casein kinase II[7] and a nucleolar localisation signal on its C-terminus.[8] It bears weak similarity to a family of nucleolar phosphoproteins and its precise function remains unknown.[9 10] In the present work, screening for *TCOF1* mutations in a large cohort of TCS patients from different populations confirms that exons 10, 15, and 16 of the repeat region and exons 23 and 24 of the nucleolar localisation signal domains are mutational hot spots, responsible for more than half of the pathogenic mutations in TCS. The identification of a pathogenic missense change in exon 2 suggests that a functionally important domain of treacle exists near the N-terminus.

MATERIALS AND METHODS

Twelve patients were referred directly to the Centro de Estudos do Genoma Humano in São Paulo, Brazil, and samples from 38 patients evaluated by a clinical geneticist or photos were received at the Johns Hopkins University School of Medicine. Of these patients, six were familial cases that showed linkage consistent with the disease locus at the chromosome 5q3 region[11] and 29 had previously been tested for mutations in exons 4-16 through heteroduplex analysis.[4]

Screening of mutations by SSCP of PCR amplified products was performed as described elsewhere.[6] Products displaying a mobility shift were sequenced in an ABIPrism Model 377 (Applied Biosystems) according to the manufacturer's instructions. Paternity was confirmed for sporadic cases where parental DNA was available by means of five highly informative tetranucleotide markers (D3S1754, D5S820, D6S477, D7S821, D12S391) following standard procedures. Segregation analysis with markers flanking the *TCOF1* gene at 5q31-34 (D5S434, D5S413, IL9, D5S820, D5S816, D5S209, D5S210, D5S211) was conducted for one family displaying a seemingly autosomal recessive mode of inheritance.

RESULTS

Of the 12 patients referred directly to CEGH, we detected alterations in SSCP pattern in 10. All alterations were confirmed to be pathogenic after sequencing, and one patient was found to have two pathogenic mutations in exon 24. The two remaining patients were subjected to direct sequencing of the 25 coding exons of the gene. One alteration was thus identified, a C→T transition at the splice acceptor region in intron 22 (c.3370-3 C→T). This alteration abolishes a restriction site for *Mbo*II, and a 100 chromosome control sample was tested for the presence of this alteration, which was found in 12% of chromosomes. Both patients without an identifiable pathogenic mutation were familial cases, with the first (TCS31) comprising an affected father and son, and the second (TCS37) two affected brothers born to normal parents. In this last case, haplotype analysis showed that both brothers shared the same markers surrounding the *TCOF1* locus on both chromosomes.

Patients from JHUSM were sorted into two groups: those who had not been previously screened for mutations (group I) and those who had had exons 4-16 previously analysed by heteroduplex analysis (group II).[4] Of the 19 patients in group I, 13 shifts were detected through SSCP. Through sequencing of abnormally migrating fragments, we determined that 11 patients had mutations causing frameshifts and two patients had missense changes. Patient G1568 had a c.149A→G transition in exon 2 (Y50C) and patient G1428 a c.3389 T→A transversion in exon 22 (M1130K). Neither alteration was detected in 100 control chromosomes in the Brazilian population. Patient G1568 is a sporadic case and analysis of parental DNA indicated that the mutation arose de novo, thus suggesting its pathogenic nature. Parental samples were not available for patient G1428. Of the 29 patients in group II, pathogenic mutations could be identified in 11; two of these mutations were located in exons that had been previously screened by heteroduplex (one in exon 4 and another in exon 12). All mutations detected are detailed in table 1.

DISCUSSION

In a previous study using the same methodology as the one used here, we were able to identify pathogenic mutations in

Table 1 Thirty-three *TCOF1* changes detected by SSCP in 60 patients

Patient	Location	DNA mutation	Consequence	Centre
G1568	Exon 2	c.149A>G (Y50C)	Missense	JHUSM
TC286	Intron 3	c.305-1 G>A	Splice	JHUSM
G1775	Exon 7	c.720-727delAGCACCCC	Frameshift	JHUSM
G1481	Exon 7	c.786delA	Frameshift	JHUSM
TCS30	Exon 8	c.1015 G>T (E339X)	Nonsense	CEGH
G1589	Exon 9	c.1098delC	Frameshift	JHUSM
TCS33	Exon 9	c.1215-1216insA	Frameshift	CEGH
G1608	Exon 10	c.1408-1409delAG[4]	Frameshift	JHUSM
TC410	Exon 12	c.1872-1875delTGAG	Frameshift	JHUSM
TCS40	Exon 13	c.2026C>T (Q676X)	Nonsense	CEGH
TCS42	Exon 13	c.1926-1927insG	Frameshift	CEGH
G1448	Exon 15	c.2375-2376delGG	Frameshift	JHUSM
TC546	Intron 16	c.2629-1G>C	Splice	JHUSM
TCS32	Intron 17	c.2815+1delG	Splice	CEGH
TC603	Exon 18	c.2846delC	Frameshift	JHUSM
G1942	Exon 18	c.2881delG	Frameshift	JHUSM
TCS41	Exon 20	c.3084delA	Frameshift	CEGH
G1282	Exon 23	c.3639delG	Frameshift	JHUSM
TC213	Exon 23	c.3779-3780insC	Frameshift	JHUSM
TCS35	Exon 23	c.3830delC	Frameshift	CEGH
TC450	Exon 23	c.3899delT	Frameshift	JHUSM
TC512	Exon 23	c.3933-3934insG	Frameshift	JHUSM
TC521	Exon 23	c.3975-3979delGAAAG	Frameshift	JHUSM
TC601	Exon 23	c.3987-3988insG[6]	Frameshift	JHUSM
G1455	Exon 23	c.3987-3988insG[6]	Frameshift	JHUSM
TC376	Exon 23	c.4019-4032delCAGAAGAGGAGCTT	Frameshift	JHUSM
G1685	Exon 24	c.4124-4125delAA	Frameshift	JHUSM
G1482	Exon 24	c.4134delA	Frameshift	CEGH
TCS36	Exon 24	c.4130-4134delAAAAA[5]	Frameshift	CEGH
TCS39	Exon 24	c.4135-4136insG; c.4156-4157delAA	Frameshift	CEGH
TCS34	Exon 24	c.4135-4139delGAAAA[5 6]	Frameshift	CEGH
TC389	Exon 24	c.4135-4139delGAAAA[5 6]	Frameshift	JHUSM
G1352	Exon 24	c.4135-4139delGAAAA[5 6]	Frameshift	JHUSM
G1428	Exon 22	c.3389T>A (M1130K)	Missense ?Pathogenic	JHUSM

References are for previously reported mutations.

26/28 (93%) patients through SSCP.[6] These findings indicate that the SSCP conditions we are currently using are a good method for mutational screening of *TCOF1*. Indeed, we detected in the group II sample from JHUSM two pathogenic changes that were not detected by heteroduplex analysis in a previous work.[4]

In the present study, we were able to identify a pathogenic mutation in 33 out of 60 patients, 10/12 (83%) among CEGH patients and 23/48 (48%) among JHUSM patients. The overall mutation detection rate was 33/60 (55%). Among patients from JHUSM, a much lower mutation detection rate was obtained when compared to that achieved in the group from CEGH. This could be because of less stringent diagnosis among the patients, as there is wide phenotypic variability observed in TCS. Also, since there are some unmistakable cases of TCS which do not display a mutation in the coding region of *TCOF1*,[6] it is possible that mutations in the promoter region or 3' UTR, gross rearrangements, or other pathogenic mechanisms could have different frequencies in these two populations.

In this study, we confirmed our previous observations that five exons accounted for over 50% of pathogenic mutations, with exons 23 and 24 being the main hot spots. As in our previous publication,[6] the most frequent kind of alterations were deletions (21), followed by insertions (7), splice mutations (3), nonsense mutations (2), and one missense change. Of all 31 different pathogenic mutations found in 33 patients, 27 are novel and four have already been described. Two of the mutations already described were observed in more than one patient independently: the frequent 5 bp deletion in exon 24 (c.4135-4139delGAAAA) was observed in three patients, and the insertion of a guanine residue in exon 23 (c.3987-3988insG), already described in two unrelated cases,[6] was observed in two index patients. One patient (TCS39) had two different pathogenic changes in exon 24 (c.4135-4136insG and c.4156-4157delAA), being the second reported case with two mutations in the same exon. The only other TCS patient with two mutations in the same exon, in this case both in exon 16 (c.2552delA and c.2561delA), was described by Wise *et al*.[4]

One of the Brazilian families that did not have an identifiable pathogenic mutation after direct sequencing of the whole coding region of *TCOF1* displays a seemingly autosomal recessive mode of inheritance. This is the second instance of a seemingly autosomal recessive form of TCS where the affected sibs share the same haplotypes on chromosome 5,[6] even though there was not a haplotype in common shared by the two families. Therefore, *TCOF1* could not be exempted from having a causative role in either an autosomal dominant or autosomal recessive mode of inheritance.

One of the novel missense mutations (M1130K) lies in exon 22, in a region of the protein that bears 68.8% identity with the mouse and canine homologues, and causes the substitution of a hydrophilic amino acid (lysine) for a hydrophobic one (methionine in humans and mice, valine in dogs) (fig 1, above). Even though this mutation was not observed in 100 control chromosomes in the Brazilian population, we cannot definitely classify this change as either pathogenic or polymorphic since the control population tested may not reflect the variation found in the American population. The M1130K mutation may be a rare, non-pathogenic change, like others that have already been described.[6]

The missense mutation in exon 2 deserves special attention. This is only the second pathogenic missense change described associated with TCS among the 75 known to date. The first description of a pathogenic missense change was made by Edwards *et al*,[5] also in exon 2 (W53R). This missense change

```
Mutations        K
H sapiens    SLLSGYMTPG  LTPANSQASK  ATPKLDSSPS  VSSTLAAKDD  PDGKQEAKPQ  QAAGMLSPKT G
M musculus   ..........  ..V.......  ...RP..NSL  A..AP.T..N  .....KS.S.  H..DTAL... .
C familiaris ......V...  P.L.......  ...RP.LN..  A...SIS..A  ......VE..  HV..TV.... S

Mutations                                                              C    R
H sapiens    MAEARKRREL  LPLIYHHLLR  AGYVRAAREV  KEQSGQKCFL  AQPVTLLDIY  THWQQ
M musculus   ..........  .........Q  ..........  .......S..  T.........  .....
C familiaris ..........  .....Q...Q  ..........  .......S..  T.....M...  .....
```

Figure 1 (Above) Alignment of the 61 amino acids encoded by human exon 22 and the corresponding region of murine and canine treacle. The missense mutation identified in TCS patient G1428 is indicated above the human sequence. The original amino acid at this position is conserved in mice but not in dogs. (Below) A comparison of the amino acid sequence encoded by exons 1 and 2 in human, murine, and canine treacle. Pathogenic mutations in the human sequence (indicated above the human sequence) are located at conserved sites.

Key points

- We describe 27 novel pathogenic mutations in *TCOF1* and confirm that exons 10, 16, 15, 23, and 24 are hot spots, accounting for over 50% of all known pathogenic mutations.
- The identification of a second missense pathogenic mutation in *TCOF1*, also located in exon 2, suggests that the N-terminus of the protein is functionally important and less tolerant of changes.

was regarded as pathogenic because it was not detected among 200 control chromosomes and the proband's parents were not carriers. The authors considered the possibility that, because this change falls close to the end of exon 2, it could interfere with normal splicing.

As previously noted,[6] the treacle protein tolerates a fair amount of amino acid substitutions in its coding region, especially in the repetitive domains. However, the 55 amino acid sequence encoded by human exons 1 and 2 has 94.5% and 90.9% identity with mouse and canine sequences respectively, being the most conserved region between species[9 10 12] (fig 1, below). The two amino acids encoded by exon 2 involved in human pathogenic mutations are conserved in both mice and dogs; furthermore, both pathogenic changes result in the substitution of an amino acid with a short side chain for another with an aromatic side chain. These observations support an important functional domain present at the N-terminus of the protein, and future work attempting to unravel treacle's function or determine interacting molecules should especially focus on this region.

ACKNOWLEDGEMENTS

The authors would like to thank the TCS families for their collaboration in this study; Drs Martin Hergersberg, Jan Miertus, Elias Olveira da Silva, Têmis Maria Félix, and Nivaldo Alonso; and Constância Gotto Urbani and Elisângela Quedas for secretarial and technical help. The financial support of FAPESP, CNPq, and Pronex is also gratefully acknowledged. MRPB is supported in part by an International Research Scholars grant from the Howard Hughes Medical Institute. This work was supported in part by NIH P60 DE13078, NIH HD24061, and NIH RR00052 (EWJ).

Authors' affiliations

A Splendore, M R Passos-Bueno, Centro de Estudos do Genoma Humano, Departamento de Genética, Instituto de Biociências Universidade de São Paulo, São Paulo, SP, Brazil

E W Jabs, Department of Pediatrics, Medicine, and Surgery, Center for Medical Genetics, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Correspondence to: Dr M R Passos-Bueno, Departamento de Genética, Instituto de Biociências, Universidade de São Paulo, Rua do Matão 277, CEP 05508-900, Cidade Universitária, São Paulo, SP, Brazil; passos@ib.usp.br

REFERENCES

1 **Gorlin RJ**, Cohen MM, Levin LS. *Syndromes of the head and neck*. 3rd ed. New York: Oxford University Press, 1990.
2 **The Treacher Collins Collaborative Group**. Positional cloning of a gene involved in the pathogenesis of Treacher Collins syndrome. *Nat Genet* 1996;**12**:130-6
3 **Gladwin AJ**, Dixon J, Loftus SK, Edwards SJ, Wasmuth JJ, Hennekam RCM, Dixon MJ. Treacher Collins syndrome may result from insertions, deletions or splicing mutations, which introduce a termination codon into the gene. *Hum Mol Genet* 1996;**5**:1533-8.
4 **Wise CA**, Chiang LC, Paznekas WA, Sharma M, Musy MM, Ashley JA, Lovett M, Jabs EW. TCOF1 gene encodes a putative nucleolar phosphoprotein that exhibits mutations in Treacher Collins syndrome throughout its coding region. *Proc Natl Acad Sci USA* 1997;**94**:3110-15.
5 **Edwards SJ**, Gladwin, AJ, Dixon MJ. The mutational spectrum in Treacher Collins syndrome reveals a predominance of mutations that create a premature-termination codon. *Am J Hum Genet* 1997;**60**:515-24.
6 **Splendore A**, Silva EO, Alonso LG, Richieri-Costa A, Alonso N, Rosa A, Carakushansky G, Cavalcanti DP, Brunoni D, Passos-Bueno MR. High mutation detection rate in TCOF1 among Treacher Collins syndrome patients reveals clustering of mutations and 16 novel pathogenic changes. *Hum Mutat* 2000;**16**:315-22.
7 **Jones NC**, Farlie PG, Minichiello J, Newgreen DF. Detection of an appropriate kinase activity in branchial arches I and II that coincides with peak expression of the Treacher Collins syndrome gene product, treacle. *Hum Mol Genet* 1999;**8**:2239-45.
8 **Winokur ST**, Shiang R. The Treacher Collins syndrome (TCOF1) gene product, treacle, is targeted to the nucleolus by signals in its C-terminus. *Hum Mol Genet* 1998;**7**:1947-52.
9 **Dixon J**, Hovanes K, Shiang R, Dixon MJ. Sequence analysis, identification of evolutionary conserved motifs and expression analysis of murine tcof1 provide further evidence for a potential function for the gene and its human homologue, TCOF1. *Hum Mol Genet* 1997;**6**:727-37.
10 **Paznekas WA**, Zhang N, Gridley T, Jabs EW. Mouse TCOF1 is expressed widely, has motifs conserved in nucleolar phosphoproteins, and maps to chromosome 18. *Biochem Biophys Res Commun* 1997;**238**:1-6.
11 **Jabs EW**, Li X, Lovett M, Yamaoka LH, Taylor E, Speer MC, Coss C, Cadle R, Hall B, Brown K, Kidd KK, Dolganov G, Polymeropoulos MH, Meyers DA. Genetic and physical mapping of the Treacher Collins syndrome locus with respect to loci in the chromosome 5q3 region. *Genomics* 1993;**18**:7-13.
12 **Haworth KE**, Islam I, Breen M, Putt W, Makrinou E, Binns M, Hopkinson D, Edwards Y. Canine TCOF1; cloning, chromosome assignment and genetic analysis in dogs with different head types. *Mam Genome* 2001;**12**:622-9.