European Journal of Human Genetics (2007) 15, 185–193
© 2007 Nature Publishing Group All rights reserved 1018-4813/07 $30.00
www.nature.com/ejhg

ARTICLE

Cystatin B: mutation detection, alternative splicing and expression in progressive myclonus epilepsy of Unverricht-Lundborg type (EPM1) patients

Tarja Joensuu*[,1], Mervi Kuronen[1], Kirsi Alakurtti[1], Saara Tegelberg[1], Paula Hakala[1], Antti Aalto[2], Laura Huopaniemi[3], Nina Aula[1], Roberto Michellucci[4], Kai Eriksson[5] and Anna-Elina Lehesjoki[1]

[1]*Department of Medical Genetics and Neuroscience Center, Folkhälsan Institute of Genetics, Biomedicum Helsinki, University of Helsinki, Finland;* [2]*Institute of Biotechnology, Viikki Biocenter, University of Helsinki, Finland;* [3]*Rational Drug Design Program, Biomedicum Helsinki, Helsinki, Finland;* [4]*Department of Neurosciences, Epilepsy Centre, Bellaria Hospital, Bologna, Italy;* [5]*Pediatric Neurology Unit, Department of Pediatrics, Pediatric Research Center, Medical School, University of Tampere and Tampere University Hospital, Tampere, Finland*

Progressive myoclonus epilepsy of Unverricht-Lundborg type (EPM1) is an autosomal recessive neurodegenerative disorder caused by mutations in the *cystatin B* gene (*CSTB*) that encodes an inhibitor of several lysosomal cathepsins. An unstable expansion of a dodecamer repeat in the *CSTB* promoter accounts for the majority of EPM1 disease alleles worldwide. We here describe a novel PCR protocol for detection of the dodecamer repeat expansion. We describe two novel EPM1-associated mutations, c.149G > A leading to the p.G50E missense change and an intronic 18-bp deletion (c.168 + 1_18del), which affects splicing of *CSTB*. The p.G50E mutation that affects the conserved QVVAG amino acid sequence critical for cathepsin binding fails to associate with lysosomes. This further supports the previously implicated physiological importance of the CSTB-lysosome association. Expression of *CSTB* mRNA and protein was markedly reduced in lymphoblastoid cells of the patients irrespective of the mutation type. Patients homozygous for the dodecamer expansion mutation showed 5–10% expression compared to controls. By combining database searches with RT-PCR we identified several alternatively spliced *CSTB* isoforms. One of these, *CSTB2*, was also present in mouse and was analyzed in more detail. In real-time PCR quantification, *CSTB2* expression was less than 5% of total *CSTB* expression in all human adult and fetal tissues analyzed. In patients homozygous for the minisatellite mutation, the level of *CSTB2* was reduced similarly to that of *CSTB* implicating regulation from the same promoter. The physiological significance of *CSTB2* remains to be determined.
European Journal of Human Genetics (2007) **15,** 185–193. doi:10.1038/sj.ejhg.5201723; published online 27 September 2006

Keywords: cystatin B; progressive myoclonus epilepsy; alternative splicing; real-time PCR; expansion repeat; lysosomes

*Correspondence: Dr T Joensuu, Folkhälsan Institute of Genetics, Department of Medical Genetics and Neuroscience Center, Biomedicum Helsinki, P.O. Box 63, University of Helsinki, Helsinki, Finland.
Tel: +358 9 191 25081; Fax: +358 9 191 25073;
E-mail: tarja.joensuu@helsinki.fi
Received 30 June 2006; revised 21 August 2006; accepted 23 August 2006; published online 27 September 2006

Introduction

Eight mutations in the *cystatin B* gene (*CSTB*, OMIM No. 601145) encoding a cysteine protease inhibitor have thus far been reported to associate with an autosomal recessive neurodegenerative disorder, progressive myoclonus epilepsy of Unverricht-Lundborg type (EPM1, OMIM No.

254800).[1–7] The onset of EPM1 with stimulus-sensitive myoclonus and/or tonic-clonic epileptic seizures occurs at the age of 6–15 years with progressive ataxia manifesting later.[8,9] Most of the disease alleles harbour an unstable expansion of at least 30 copies of a normally polymorphic 12-nucleotide, dodecamer repeat located in the promoter region of the *CSTB* gene.[3,5,7,10] Three reported EPM1 mutations affect splice sites (c.67-1G>C, c.168G>A, c.169-2A>G), two result in amino-acid changes (c.10G>C, p.G4R; c.212A>C, p.Q71P) and two predict truncated proteins either through creating a stop codon (c.202C>T) or producing a frameshift (c.218_219delTC).

Markedly reduced *CSTB* mRNA expression has been reported by Northern analysis in lymphoblastoid cells of EPM1 patients who are heterozygous or homozygous for the expansion mutation.[1,3,4,6,11] Consequently, decreased inhibitory activity of the CSTB protein in these cells has been demonstrated.[12] On the contrary, by RNase protection assay normal *CSTB* expression levels were detected in lymphoblastoid cells and fibroblasts of patients, while blood leukocytes showed reduced mRNA level, suggesting cell-specific regulation of *CSTB* expression.[5] Consistent with these results, the activity of the putative *CSTB* promoter, which contained ~50 copies of dodecamer repeats, was found to be cell-specific and varied from normal to two- to fourfold reduction.[13] The repression of transcription was proposed to result from the disrupted spacing of transcription factor binding sites from the transcription initiation site. Contrary to these results, but compatible with the Northern blot data on *CSTB* expression, our *in vitro* study with a promoter construct containing 19 copies of the dodecamer repeat in transiently transfected COS-1 cells showed a reduced promoter activity of 10-fold.[11] Later, the expanded dodecamer repeats have been shown to induce stable tetraplex secondary structures, which has been suggested to result in the repression of *CSTB* transcription by altered chromatin structures.[14]

Here, we have characterized the *CSTB* gene in more detail. We describe a novel direct PCR amplification protocol, which allows reliable detection of expanded *CSTB* alleles from genomic DNA and describe the identification and characterization of two novel EPM1-associated *CSTB* gene mutations. We have further quantitated the *CSTB* expression in human tissues and in patients' lymphoblastoid cells by real-time PCR. Moreover, we have experimentally verified alternative splicing of *CSTB* and describe the characterization of a ubiquitously expressed variant, *CSTB2*.

Materials and methods

EPM1 patient and control samples

The novel PCR-based method to amplify the dodecamer repeat expansion from genomic DNA was validated in 16 patients with homozygous mutation for the expansion, five heterozygous expansion mutation carriers and three control individuals with no expansion, as determined by Southern blot hybridizations. The method was applied in molecular diagnostics of an Italian patient with a clinical suspicion of EPM1.

Genomic DNA from a Finnish EPM1 patient previously identified to be a heterozygous carrier of the dodecamer expansion mutation, as well as from an Italian patient (see above) with no previously known *CSTB* mutations were scanned for novel mutations by sequencing. The identified novel mutations were screened in a panel of 73 Finnish and 93 CEPH control individuals.

Total RNA and/or protein extracted from lymphoblastoid cells of eight EPM1 patients of which two were homozygous for the minisatellite expansion and six compound heterozygous for the expansion and either the splice site c.67-1G>C ($N=4$), the p.G50E missense c.149G>A ($N=1$) or the nonsense c.202C>T ($N=1$) mutation were analyzed for the *CSTB* gene and protein expression. One expansion mutation heterozygous carrier and two non-carrier individuals were examined as controls. *CSTB* and *CSTB2* expression was also examined from fibroblast RNA of two expansion mutation homozygous patients and two controls. For the first-strand cDNA synthesis, total RNA was isolated using the RNAqueous™–4PCR kit (Ambion, Austin, TX, USA). The c.168+1_18del mutation was characterized by RT-PCR from total RNA extracted from a fresh blood using the PAXgene Blood RNA Validation kit (Qiagen, Hilden, Germany).

The study has been approved by an institutional ethical review board at the University of Helsinki.

Mutation analysis

The amplification conditions across the dodecamer repeat expansion were modified from the Expand Long Template PCR system (Roche Diagnostics, Mannheim, Germany) and performed in a total of 20 μl reaction mix using primers 2F (5′-CCC GGA AAG ACG ATA CCA G-3′) and 1R (5′-GAG GAG GCA CTT TGG CTT C-3′) (Figure 1) with 15 ng genomic DNA, 0.5 mM dNTP, 0.5 μM of each primer, 0.12 U enzyme mix, 1 X BSA, 5% DMSO, and 1.0 M GC-melt mix (Clontech, Palo Alto, CA, USA). Following denaturation at 94°C for 2 min, PCR was carried out in 40 cycles of 94°C for 10 s, 57°C for 45 s, and 68°C for 8 min. A final elongation was at 68°C for 7 min.

For mutation scanning, 2949 bp of the *CSTB* gene (GenBank AF208234), including the exons, the introns, and the 5′ and 3′ untranslated regions, was amplified in overlapping fragments (Figure 1) using the following primers: 1F (5′-AAA CGC AAA TTC CAC CAG AG-3′) and 1R (5′-GAG GAG GCA CTT TGG CTT C-3′), 2F (5′-CCC GGA AAG ACG ATA CCA G-3′) and 2R (5′-CGG CTT CTT TCG CTC CAG-3′), 3F (5′-GCC GAG ACC CAG CAC ATC-3′) and 3R (5′-CCT GTG GAC CTT TTA TGC AG-3′), 4F (5′-GCA AGA GGT CCC CAG TGA TA-3′) and 4R (5′-TGA CAC

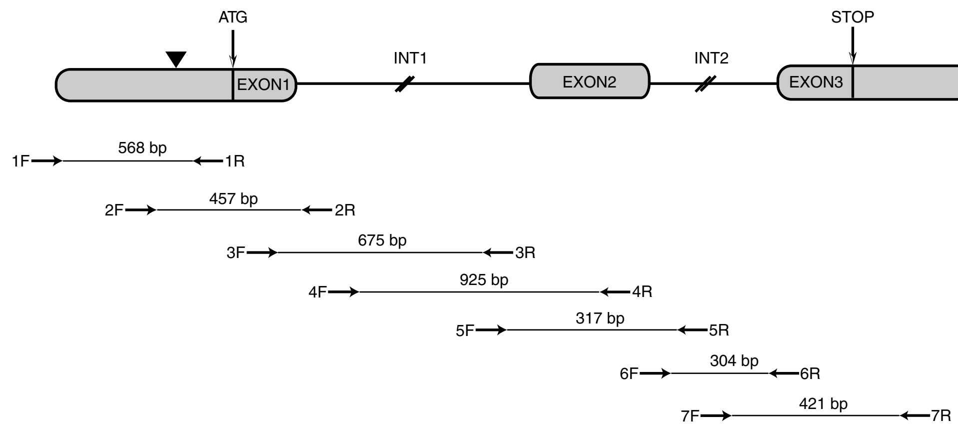

Figure 1 Schematic structure of the *CSTB* gene and the localization of primers used in mutation analysis with lengths of the respective PCR products. The primer sequences are given in the Materials and methods. The exons of *CSTB* are depicted as boxes. The translation initiation site and the stop codon are indicated by arrows. The position of the dodecamer repeat is shown by an arrowhead.

GGC CTT AAA CAC AG-3′), 5F (5′-ACC GTA CCC AGC TGG AAC TGT-3′) and 5R (5′-GCT TAT CTC AGG GGG CAG CCA CAG-3′), 6F (5′-GTA GAG TGT GGG CCT CAG GA-3′) and 6R (5′-AAG CCT CTG ATC CCA AGT CA-3′), 7F (5′-ATT GTC TTC AGC TGG CTG CTA AT-3′) and 7R (5′-AAG ATC ACC TAT TGG GAA GGA AAG A-3′). The PCR fragments were sequenced by an ABI 3730 DNA analyzer (Applied Biosystems, Foster City, CA, USA) and the sequences analyzed for mutations with Sequencher 4.0 program (Gene Codes Corporation, Ann Arbor, MI, USA).

Characterization of the alternatively spliced forms of *CSTB*

The alternatively spliced *CSTB* transcripts were verified and isolated by PCR from the Marathon Ready Human Brain cDNA (Clontech) and/or Human Multiple Tissue cDNA panels I and II (Clontech) and sequenced. Two variant transcripts were identified from EST sequences present in ASAP or AceView databases in NCBI, while two were identified by RT-PCR. The complete coding region of mouse *Cstb2* was amplified from Marathon Ready Mouse Brain cDNA library (Clontech) and testis cDNA.

Quantitative real-time PCR analysis of human *CSTB* and *CSTB2* expression

The primers and probes for the real-time expression analysis of *CSTB* and *CSTB2* were designed to overlap exonic splice junctions with Primer Express® program (version 2.0, Applied Biosystems). Specificity for the two variants was achieved with different reverse primers. Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*),[15] and TATA-box binding protein (*TBP*; Assay no. 4333769F) were used as endogenous controls.

The mRNA quantifications were carried out using the ABI Prism® 7000 Sequence Detection System (Applied Biosystems). The 25-μl reaction volume contained 30–50 ng of lymphoblastoid cell or fibroblast cDNA or 2 ng cDNA from Human adult and fetal MTC™ Panels I and II (Clontech), 300 nM forward primer (except for *GAPDH*; 900 nM), 300 nM reverse primer, and 200 nM probe specific to each amplicon in TaqMan Universal PCR Master Mix (Applied Biosystems). All reactions were prepared in triplicate, and a minimum of three separate runs was performed for each sample. The mRNA expression levels in different human adult and fetal tissues were calculated using the comparative C_T method (user bulletin 2, ABI PRISM 7000 Sequence Detection System). The proportion of *CSTB2* expression was calculated from total *CSTB* and *CSTB2* expression. In lymphoblastoid cells and fibroblasts, the expression levels were compared to the mean percentage ($\pm$SD) of the expression in controls. For mutations with more than one patient, the mean value ($\pm$SD) of the expression level of all patients studied was calculated.

Expression analysis of the mutant CSTB proteins

Lymphoblastoid cell pellets were resuspended in lysis buffer (10 mM Tris-HCl, 100 mM NaCl, 2% Triton X-100, 1% SDS, 1 mM EDTA), 25 μg of proteins were electrophoresed on 8–16% gradient SDS-polyacrylamide gels (Bio-Rad, Hercules, CA, USA) under denaturing conditions, and transferred onto a Hybond™-C Extra filter (Amersham Biosciences, Uppsala, Sweden). After incubation in 1% nonfat milk in TBST overnight at 4°C with a 1:1000 dilution of the CSTB polyclonal antibody (Biogenesis, Kingston, NH, USA) the protein-antibody complexes were detected with horse-radish-peroxidase-conjugated goat anti-mouse IgG (Dako, Copenhagen, Denmark) using enhanced chemiluminescence (Amersham Pharmacia Biotech, Piscataway, NJ, USA).

The *CSTB* coding region PCR-amplified from lymphoblastoid cDNA of the missense c.149G>A mutation patient was subcloned into pcDNA3.1 (+) (Invitrogen, Carlsbad, CA, USA; Life Technologies). *CSTB* in pcDNA3.1 (+)[16] was used as a control. 1.5 μg of expression plasmids were

transfected into BHK-21 cells (American Type Culture Collection; ATCC, Rockville, MD, USA) using FuGENE 6 Transfection Reagent (Roche Diagnostics Corporation, Indianapolis, IN, USA). The cells were fixed with 4% paraformaldehyde and incubated with ice-cold methanol for 3 min before incubation with antibodies. The monoclonal 2E7 antibody[16] was used to detect CSTB and the polyclonal lysosome-associated membrane protein 1 (LAMP1) antibody (lgp120; a kind gift from Dr van der Sluijs, the Netherlands) was used as a lysosome specific marker. As secondary antibodies Cy2- or Cy3-conjugated anti-mouse and anti-rabbit antibodies (Jackson Immunochemicals, West Grove, PA, USA) were used.

Databases

BLAST program (http://www.ncbi.nlm.nih.gov/BLAST/) was used for similarity and homology searches against human and mouse EST and genome sequence databases. Repeat Masker (http://www.repeatmasker.org/) was used to screen for the repetitive elements before any similarity searches. The primer sequences were designed using the program Primer 3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Comparative analysis of alternatively spliced cDNAs and ESTs were performed using ASAP (http://www.bioinformatics.ucla.edu/HASDB/) and AceView in NCBI.

Results

Identification of EPM1 mutations

We designed a novel PCR protocol (see Materials and methods) to replace the Southern blot method used widely in the detection of the dodecamer repeat expansion in the *CSTB* promoter (Figure 2). The method was validated in 21 known homozygous and heterozygous expansion mutation carriers in whom the repeat copy number ranged from approximately 55 to 85. The method was applied to molecular diagnosis of an Italian EPM1 patient (patient J) with a clinical suspicion of EPM1, who was heterozygous for the expansion (Figure 2).

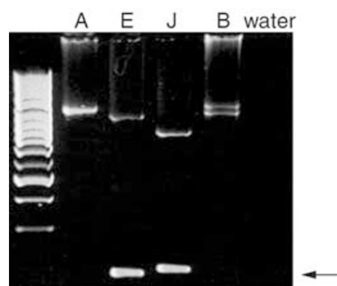

Figure 2 Detection of the expanded *CSTB* dodecamer repeat mutation by PCR. (A, B) Expansion homozygous patients previously determined by Southern analysis. Individual E is compound heterozygous for the expansion and the missense c.149G>A mutations (see Figure 3); Individual J is compound heterozygous for the expansion and the deletion c.168 + 1_18del mutations (see Figure 3). As a size marker, the 100-bp DNA ladder was used. The arrow indicates the fragment of an approximately 193-bp normal allele with three copies of the repeat in heterozygous expansion carriers. A and B patients homozygous for the expansion mutation.

We screened genomic DNA of the Italian patient and one Finnish patient previously identified by Southern blot to be a heterozygous expansion mutation carrier for *CSTB* mutations by PCR and sequencing (Figure 1). A novel c.149G>A change was identified in the Finnish patient E (Figure 3a). It results in substitution of glutamic acid for glycine at codon 50 (p.G50E) affecting the highly conserved QVVAG papain-binding region on the CSTB protein. A second novel mutation, observed in the Italian patient J (Figure 3b), is a deletion of the first 18 nucleotides downstream from the 5′ splice donor site of intron 2. On cDNA level, this results in aberrant splicing of *CSTB* with two different transcripts (Figure 3c). One has an in-frame deletion of exon 2 and predicts a deletion of 34 amino acids (p.delV23_K56), while the other contains an insertion of 25 nucleotides from intron 2 resulting in a frameshift and predicting a premature stop codon (p.V57EfsX28). None of the controls carried these mutations.

Characterization of alternative splice forms of *CSTB*

ASAP and AceView databases revealed several overlapping ESTs that suggested alternative splicing of *CSTB*. By combining *in silico* analyses with direct PCR amplification from a brain cDNA library, we isolated a variant with retention of the entire intron 2. This novel variant, *CSTB2*, is transcribed from two exons with an ORF of 249 bp. It encodes a putative 9.0-kDa protein of 83 amino acids, including 57 identical to CSTB followed by 26 amino acids encoded by the intron 2 sequences. The mouse *Cstb2* orthologue was PCR-amplified from mouse brain and testis cDNAs producing a fragment with an ORF of 192 bp. The predicted Cstb2 protein of 64 amino acids is 81% identical with human CSTB2 in the first 57 amino acid overlap. A nucleotide BLAST search using intron 2 sequences followed by PCR amplification of human tissues cDNAs with different intronic and exonic primer combinations resulted in the identification of at least three additional partial *CSTB* splice variants with limited expression in colon, intestine, and prostate (data now shown). We were not able to fully assemble these splice forms by PCR, since the amplification from the 5′ end of the cDNA favoured the major isoform.

Real-time quantitative PCR of *CSTB* isoforms

CSTB and *CSTB2* mRNA expression in different human adult and fetal tissues was quantitated using real-time PCR. In peripheral tissues, the *CSTB* and *CSTB2* mRNAs are ubiquitously expressed, most abundantly in lung and least in heart and skeletal muscle (Figure 4a, b). The level of

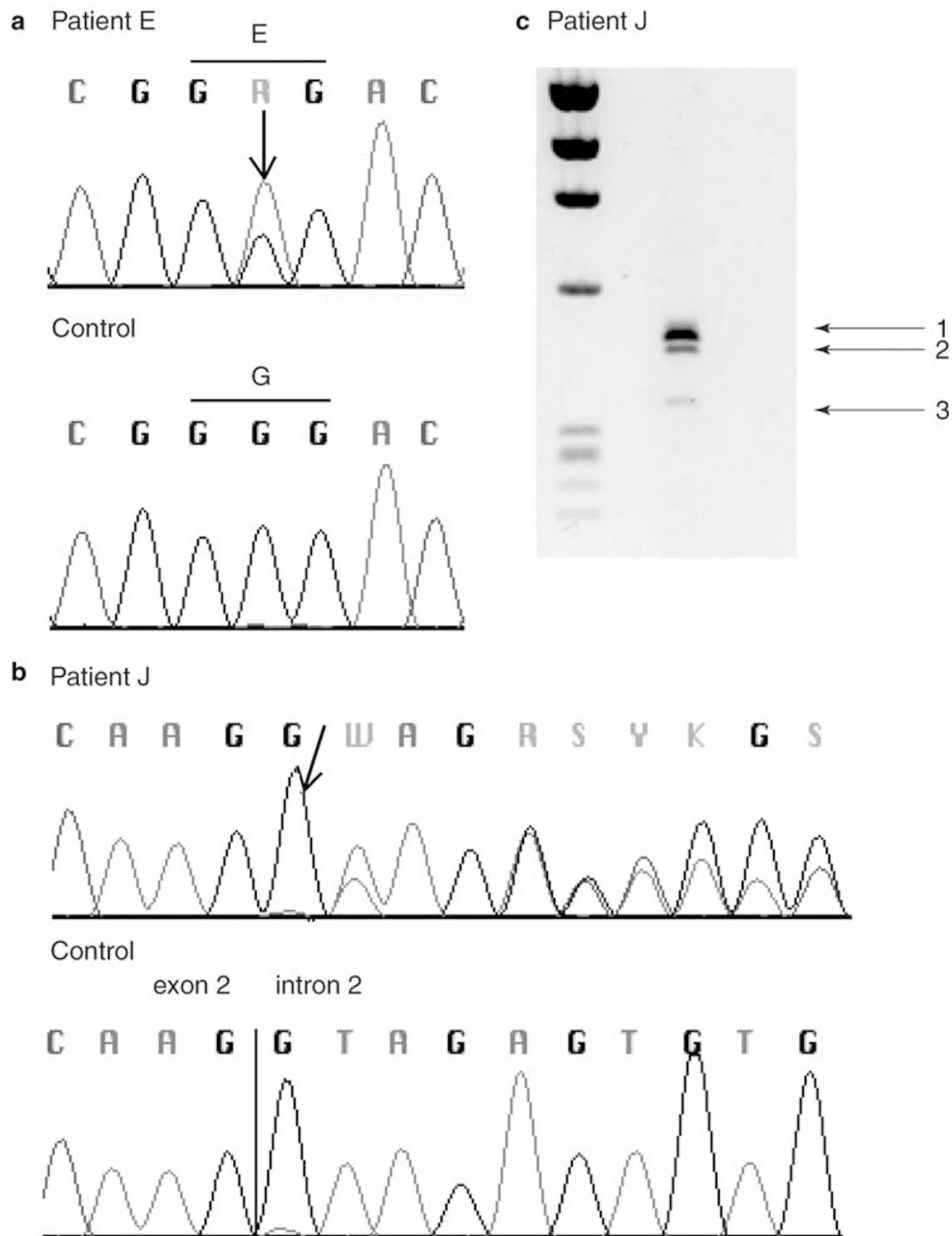

Figure 3 Novel EPM1-associated *CSTB* mutations. Sequence chromatograms showing (**a**) the heterozygous c.149G>A mutation (↓) in exon 2 in the Finnish patient (E in Figure 2) and (**b**) the heterozygous 18-bp deletion (c.168+1_18del) in intron 2 in the Italian patient (J in Figure 2). The deletion start site is marked with an ↓. (**c**) RT-PCR analysis of the c.168+1_18del mutation. In addition to the expected 438-bp product (2), two splice variants are identified. In the longer variant (1), 25 bp of intron 2 sequence is retained and in the shorter variant (3) exon 2 is skipped.

endogenous *CSTB2* relative to *CSTB* was low in all tissues examined, the ratio varying from 0.4% to approximately 4% (Figure 4c).

The *CSTB* mRNA expression in lymphoblastoid cells of expansion mutation homozygous EPM1 patients was markedly reduced, the mean expression (±SD) being 8.8±1.0% (Figure 5a) of controls. A heterozygous expansion mutation carrier had 60.6±12.7% of the expression in controls. In patients compound heterozygous for the expansion and the c.67-1G>C splice site mutations, the level of *CSTB* mRNA was 17.9±3.8% compared to controls. In a patient compound heterozygous for the expansion and the c.202C>T nonsense mutations, and in a patient compound heterozygous for the expansion and the novel

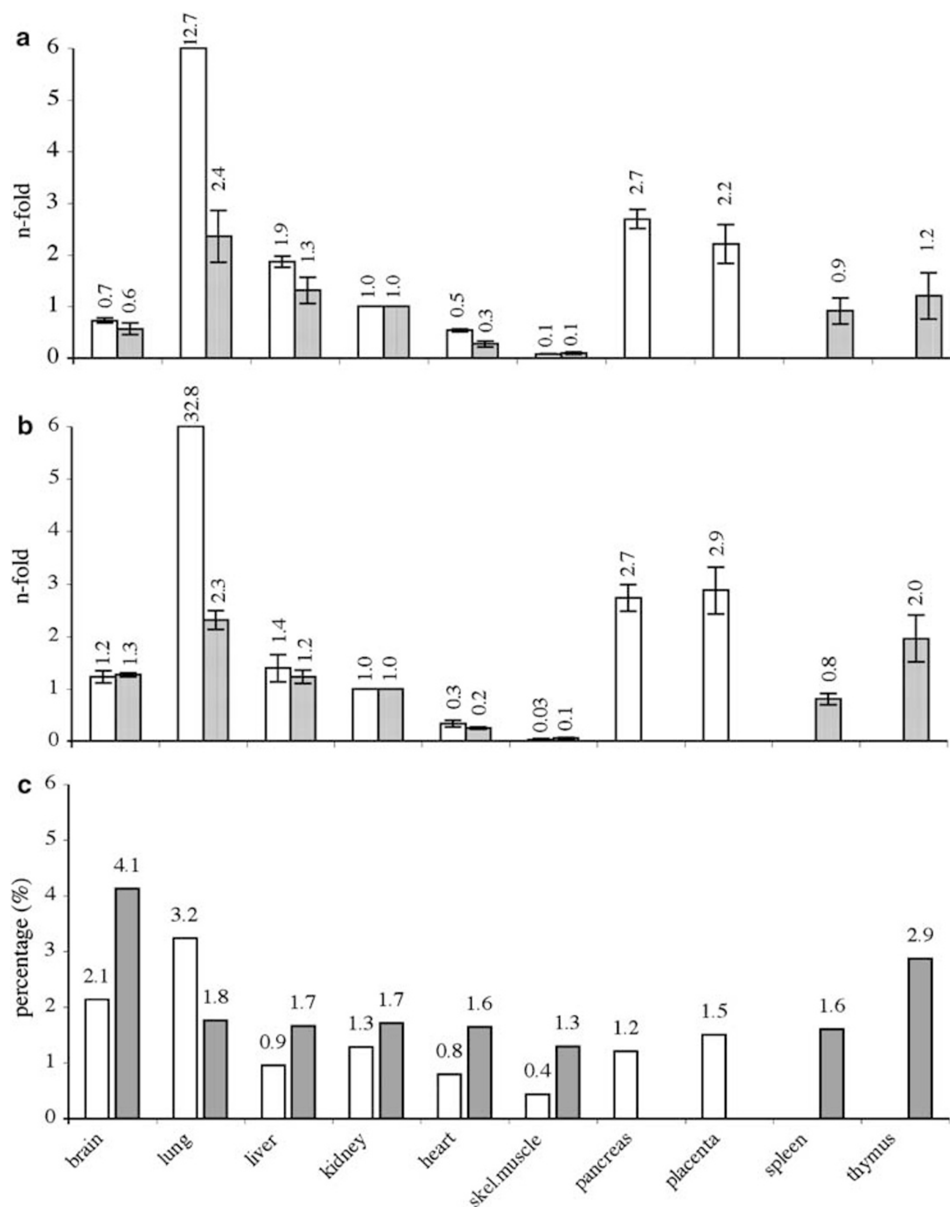

Figure 4 Quantification of human *CSTB* and *CSTB2* tissue expression by real-time PCR. The expression levels of (**a**) *CSTB* and (**b**) *CSTB2* mRNAs in various human adult (white columns) and fetal (grey columns) tissues. The expression of *CSTB* and *CSTB2* in the kidney, given a value of 1, was used as a reference. The expression in the other tissues is shown as a fold of the reference. The *n*-fold value of expression is marked above to columns. The error bars indicate $\pm$SD. (**c**) Proportion of *CSTB2* expression from total *CSTB* and *CSTB2* expression in various human adult (white columns) and fetal (grey columns) tissues. The percentage of *CSTB2* expression is marked above the columns.

c.149G > A missense mutations, expression was 35.4 $\pm$ 4.5% and 15.6 $\pm$ 3.8%, respectively.

We were not able to quantitate reliably the amount of the *CSTB2* mRNA in lymphoblastoid cells of EPM1 patients or controls due to inter-assay variation. Instead, using mRNA extracted from fibroblasts of two patients homozygous for the minisatellite expansion mutation, similar reduction in the levels of *CSTB* and *CSTB*2 compared to controls was detected (data not shown).

Expression analysis of the mutant CSTB proteins

While reduced levels of CSTB in lymphoblastoid cells of EPM1 patients either homozygous for the expansion mutation or compound heterozygous for the expansion mutation and the c.202C > T nonsense mutation have previously been described,[16] we here investigated the expression of the c.149G > A (p.G50E) and the c.67-1G > C (p.delV23_K56) mutants by Western analysis (Figure 5b). In the patient with the p.G50E missense

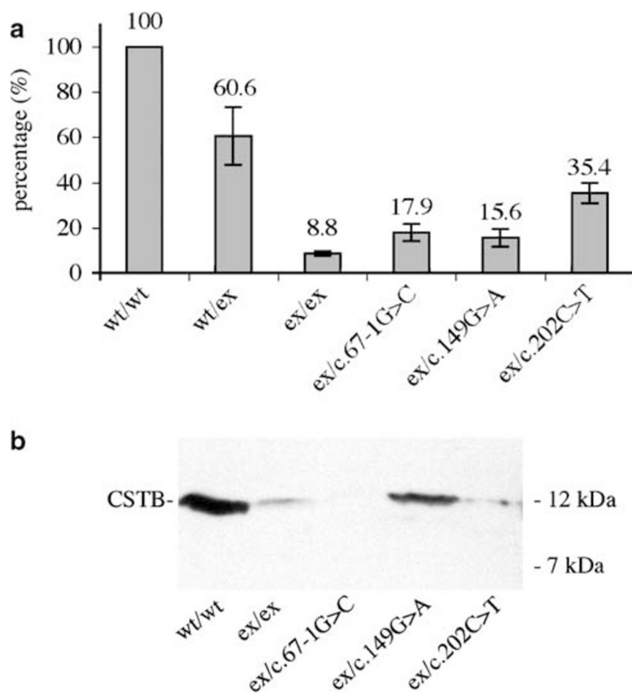

Figure 5 *CSTB* mRNA and protein expression in EPM1 patients. (**a**) Real-time PCR quantification in total lymphoblastoid RNA of EPM1 patients with various mutations. The expression levels were compared to the mean expression level of the control individuals, which was set at 100%. The level of *CSTB* expression is marked above the columns. Error bars indicate ±SD. (**b**) Western blot analysis of CSTB protein expression in lymphoblastoid cell extracts of EPM1 patients with various mutations. Full-length CSTB migrates at about 12 kDa.

mutation, a slight decrease in the amount of the CSTB protein was observed compatible with the other allele harbouring the dodecamer expansion mutation. In the patient compound heterozygous for the c.67-1 > G splice and the expansion mutations, no detectable CSTB was present.

Next, we studied the subcellular localization of the overexpressed mutant p.G50E protein in BHK-21 cells. By immunofluorescent staining, this mutant protein was localized in the nucleus and diffusely in the cytoplasm (Figure 6b). In contrast to the wild-type CSTB (Figure 6a), neither granular cytoplasmic staining nor colocalization with the lysosomal LAMP1 marker protein were detected.

Discussion

The visualization of the dodecamer minisatellite expansion mutation in the *CSTB* promoter by PCR amplification has been challenging due to a high GC content.[3,5,7] The expansion has been visualized with Southern hybridization, PCR amplification followed by hybridization or PCR amplification under special conditions with deaminated DNAs as templates.[10,17–19] Using a new PCR-based protocol to detect the expansion mutation, we were able to correctly determine the genotype previously determined by Southern blotting.[7] Expansions with repeat copy numbers of up to approximately 85, that is in the size range of most of the reported pathogenic alleles,[5,18–20] were readily amplified. Therefore, we considered this protocol to be applicable also in clinical setting and applied it correctly in molecular diagnostics of one newly identified Italian patient.

While the minisatellite expansion mutation accounts for more than 90% of EPM1 alleles worldwide, rare EPM1-associated mutations are occasionally reported.[1–4,6,21] We here report two novel mutations that affect the *CSTB* transcript. The mutation in the Italian patient, an 18-bp intronic deletion, was shown to affect splicing of *CSTB*, with two predicted abnormal protein products. However, it remains unknown whether these are translated *in vivo* or are degraded after transcription. The novel p.G50E mutation in the Finnish patient affects the highly conserved QVVAG-motif in the first beta-hairpin loop important for the complex formation with the target proteases, cathepsins.[22] Amino acid changes in the QVVAG pentapeptide are likely to affect the stability or life span of the inhibitory effect of the protein, and the interactions with their target proteins.[23,24] As in Western blot analysis of lymphoblastoid cells expression from the p.G50E mutant allele is evident, the deleterious effect of this mutation is likely to be due to impaired cathepsin interaction. In line with previous results on two EPM1-associated missense mutations (p.G4R and p.Q71P)[16], also the mutant p.G50E protein showed loss of lysosomal localization. This further supports the physiological importance of CSTB-lysosome association, but whether this occurs through cathepsin interaction or through some other mechanism remains to be studied.

Using Northern analysis and RNase protection assay the results of *CSTB* expression in different cell types have been somewhat controversial. Both reduced and normal expression in lymphoblastoid cells has been reported.[1,3–6,11] We here applied TaqMan quantitative real-time PCR to determine the level of *CSTB* expression in lymphoblastoid cells of EPM1 patients with different mutations. Our results unequivocally show that all mutations analyzed result in down-regulation of *CSTB* expression. In patients homozygous for the expansion mutation, the mRNA expression level was <10%, compatible with the *in vitro* promoter studies that have indicated significantly reduced *CSTB* promoter activity.[11,13] Concordant with the gene expression results the amount of the CSTB protein was significantly reduced both in lymphoblastoid cells and fibroblasts of expansion mutation homozygous patients. The contradictory findings of either normal or only slightly reduced *CSTB* mRNA by RNase protection assay[5] may be due to a more complicated experimental set-up in the assay.

The c.67-1G>C change located at an invariant splice junction has been shown to result in skipping of exon 2

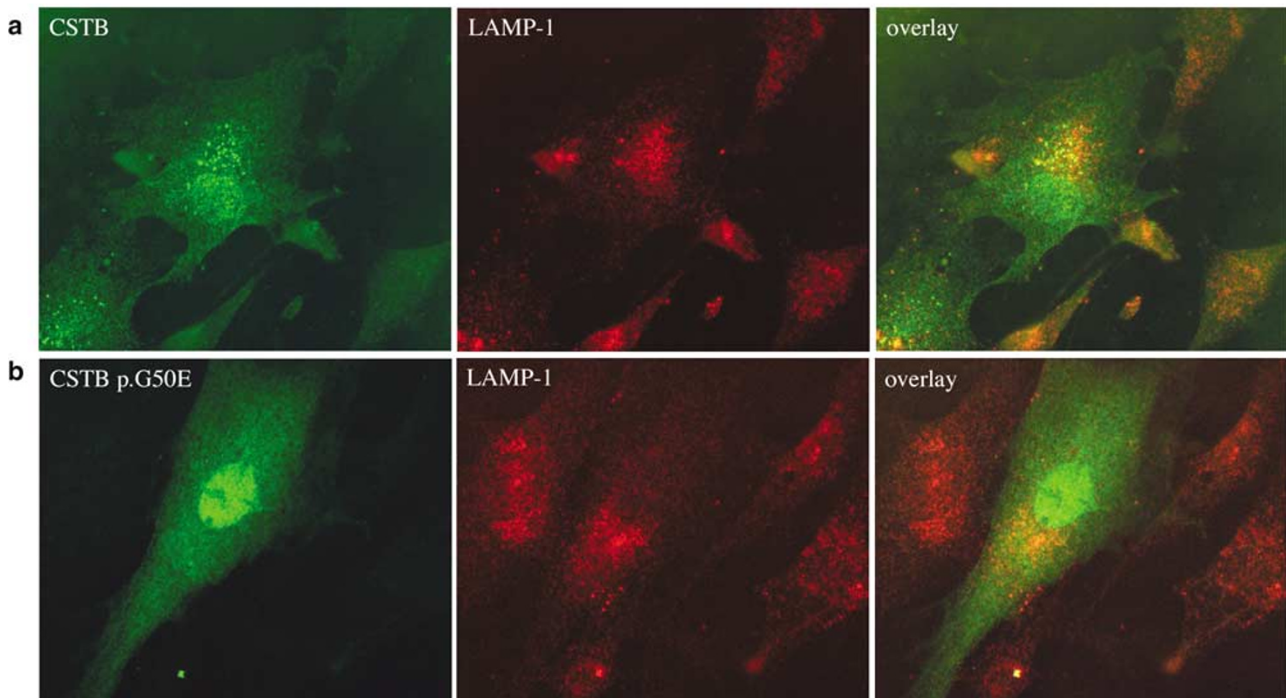

Figure 6 Subcellular localization of the p.G50E mutant CSTB protein. Double immunofluorescence staining in transiently transfected BHK-21 cells using the monoclonal 2E7 CSTB antibody (red) and polyclonal lgp tail 120 antibody (green) to visualize lysosome-associated membrane protein 1 (LAMP1). (**a**) Colocalization of the punctuate cytoplasmic wild-type CSTB with the LAMP1 protein is indicated with yellow colour in the overlay image. (**b**) The p.G50E mutant protein is detected in the nucleus and diffusely the cytoplasm, with less marked punctuate cytoplasmic staining and absence of colocalization with the lysosomal LAMP1 marker protein.

thus predicting an in-frame-deleted CSTB protein.[1] As our TaqMan assay is designed to detect transcripts with both exons 2 and 3, we expected no mRNA from the c.67-1G > C allele. However, the amount of mRNA was higher than expected implicating that the splicing defect leaks. Such partial penetrance mutations have been reported[25] and the possibility has also been discussed in relation to this mutation.[1] The c.67-1G > C mutant mRNA or the encoded polypeptide seem to be unstable, as no mutant protein was detected in either Western analysis of patient cells or in cellular transfection studies even in the presence of a proteosomal inhibitor (data not shown). In the patient compound heterozygous for the expansion mutation and the p.R68X nonsense mutation, reduced amount of the *CSTB* mRNA was detected by real-time PCR, which indicates the instability of the mutant transcripts. This is in line with our earlier findings that the p.R68X protein is rapidly degraded in transiently expressed BHK-21 cells.[16] In the patient compound heterozygous for the expansion mutation and the novel p.G50E missense mutation, the results between mRNA and protein expression remain controversial, as significantly reduced mRNA expression was detected in repeated TaqMan assays, whereas Western analysis suggested expression from the missense mutant allele.

Up to 60% of human genes and a minimum of 41% of mouse genes are estimated to be alternatively spliced resulting in a large number of proteins with a variety of functions or tissue-specificities.[26,27] The amount of tissue-specific splice variants has been shown to be highest in brain.[28,29] By combining EST database searches with RT-PCR, we were able to identify at least four novel *CSTB* splice forms. While three of the variants showed limited tissue expression and were not expressed in the brain, one, *CSTB2*, is ubiquitously expressed and conserved among human and mouse. Owing to intronic sequence retention, some of the *CSTB* variants are truncated on their C-terminal ends and are probable targets for nonsense-mediated decay. On the other hand, they might function in regulation of *CSTB* expression in post-transcriptional level by inhibiting the splicing of the pre-mRNA or by representing a storage form of the immature transcripts, whose introns are removed in certain cells or tissues. The *CSTB2* mRNA is also raised by intronic retention and is transcribed at very low levels. Its presence as multiple ESTs from different cDNA libraries would indicate its functional relevance. Whether it is degraded, serves a regulatory function for *CSTB* expression or is translated into a protein that has distinct functions from that of CSTB remains to be investigated.

Acknowledgements

This work was supported by the Folkhälsan Research Foundation, the Sigrid Juselius Foundation and the Academy of Finland (Projects 44870 and 213506, Finnish Centre of Excellence programmes 2000-2005 and 2006-2011).

References

1 Bespalova IN, Pranzatelli M, Burmeister M: G to C transversion at a splice acceptor site causes exon skipping in the cystatin B gene. *Mutat Res* 1997; **382**: 67–74.
2 Kagitani-Shimono K, Imai K, Okamoto N, Ono J, Okada S: Unverricht-Lundborg disease with cystatin B gene abnormalities. *Pediatr Neurol* 2002; **26**: 55–60.
3 Lafreniere RG, Rochefort DL, Chretien N *et al*: Unstable insertion in the 5′flanking region of the cystatin B gene is the most common mutation in progressive myoclonus epilepsy type 1, EPM1. *Nat Genet* 1997; **15**: 298–302.
4 Lalioti MD, Mirotsou M, Buresi C *et al*: Identification of mutations in cystatin B, the gene responsible for the Unverricht-Lundborg type of progressive myoclonus epilepsy (EPM1). *Am J Hum Genet* 1997; **60**: 342–351.
5 Lalioti MD, Scott HS, Buresi C *et al*: Dodecamer repeat expansion in cystatin B gene in progressive myoclonus epilepsy. *Nature* 1997; **386**: 847–851.
6 Pennacchio LA, Lehesjoki AE, Stone NE *et al*: Mutations in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1). *Science* 1996; **271**: 1731–1734.
7 Virtaneva K, D'Amato E, Miao J *et al*: Unstable minisatellite expansion causing recessively inherited myoclonus epilepsy, EPM1. *Nat Genet* 1997; **15**: 393–396.
8 Koskiniemi M, Donner M, Majuri H, Haltia M, Norio R: Progressive myoclonus epilepsy. A clinical and histopathological study. *Acta Neurol Scand* 1974; **50**: 307–332.
9 Norio R, Koskiniemi M: Progressive myoclonus epilepsy: genetic and nosological aspects with special reference to 107 Finnish patients. *Clin Genet* 1979; **15**: 382–398.
10 Lalioti MD, Scott HS, Genton P *et al*: A PCR amplification method reveals instability of the dodecamer repeat in progressive myoclonus epilepsy (EPM1) and no correlation between the size of the repeat and age at onset. *Am J Hum Genet* 1998; **62**: 842–847.
11 Alakurtti K, Virtaneva K, Joensuu T, Palvimo JJ, Lehesjoki AE: Characterization of the cystatin B gene promoter harboring the dodecamer repeat expanded in progressive myoclonus epilepsy, EPM1. *Gene* 2000; **242**: 65–73.
12 Rinne R, Saukko P, Jarvinen M, Lehesjoki AE: Reduced cystatin B activity correlates with enhanced cathepsin activity in progressive myoclonus epilepsy. *Ann Med* 2002; **34**: 380–385.
13 Lalioti MD, Scott HS, Antonarakis SE: Altered spacing of promoter elements due to the dodecamer repeat expansion contributes to reduced expression of the cystatin B gene in EPM1. *Hum Mol Genet* 1999; **8**: 1791–1798.
14 Saha T, Usdin K: Tetraplex formation by the progressive myoclonus epilepsy type-1 repeat: implications for instability in the repeat expansion diseases. *FEBS Lett* 2001; **491**: 184–187.
15 Medhurst AD, Harrison DC, Read SJ, Campbell CA, Robbins MJ, Pangalos MN: The use of TaqMan RT-PCR assays for semiquantitative analysis of gene expression in CNS tissues and disease models. *J Neurosci Methods* 2000; **98**: 9–20.
16 Alakurtti K, Weber E, Rinne R *et al*: Loss of lysosomal association of cystatin B proteins representing progressive myoclonus epilepsy, EPM1, mutations. *Eur J Hum Genet* 2005; **13**: 208–215.
17 Horiuchi H, Osawa M, Furutani R *et al*: Polymerase Chain Reaction-Based Analysis Using Deaminated DNA of Dodecamer Expansions in CSTB, Associated with Unverricht-Lundborg Myoclonus Epilepsy. *Genet Test* 2005; **9**: 328–333.
18 Larson GP, Ding S, Lafreniere RG, Rouleau GA, Krontiris TG: Instability of the EPM1 minisatellite. *Hum Mol Genet* 1999; **8**: 1985–1988.
19 Weinhaeusel A, Morris MA, Antonarakis SE, Haas OA: DNA deamination enables direct PCR amplification of the cystatin B (CSTB) gene-associated dodecamer repeat expansion in myoclonus epilepsy type Unverricht-Lundborg. *Hum Mutat* 2003; **22**: 404–408.
20 Virtaneva K, Paulin L, Krahe R, de la Chapelle A, Lehesjoki AE: The minisatellite expansion mutation in EPM1: resolution of an initial discrepancy. Mutatations in brief no. 186. online. *Hum Mutat* 1998; **12**: 218.
21 de Haan GJ, Halley DJ, Doelman JC *et al*: Univerricht-Lundborg disease: underdiagnosed in the Netherlands. *Epilepsia* 2004; **45**: 1061–1063.
22 Stubbs MT, Laber B, Bode W *et al*: The refined 2.4 A X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *EMBO J* 1990; **9**: 1939–1947.
23 Auerswald EA, Nagler DK, Assfalg-Machleidt I, Stubbs MT, Machleidt W, Fritz H: Hairpin loop mutations of chicken cystatin have different effects on the inhibition of cathepsin B, cathepsin L and papain. *FEBS Lett* 1995; **361**: 179–184.
24 Jerala R, Trstenjak-Prebanda M, Kroon-Zitko L, Lenarcic B, Turk V: Mutations in the QVVAG region of the cysteine proteinase inhibitor stefin B. *Biol Chem Hoppe Seyler* 1990; **371** (Suppl): 157–160.
25 Svenson IK, Ashley-Koch AE, Gaskell PC *et al*: Identification and expression analysis of spastin gene mutations in hereditary spastic paraplegia. *Am J Hum Genet* 2001; **68**: 1077–1085.
26 Lander ES, Linton LM, Birren B *et al*: Initial sequencing and analysis of the human genome. *Nature* 2001; **409**: 860–921.
27 Okazaki Y, Furuno M, Kasukawa T *et al*: Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* 2002; **420**: 563–573.
28 Grabowski PJ, Black DL: Alternative RNA splicing in the nervous system. *Prog Neurobiol* 2001; **65**: 289–308.
29 Xu Q, Modrek B, Lee C: Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic Acids Res* 2002; **30**: 3754–3766.