Title

Full text links Actions Favorites Share Page navigation Title & authors Abstract Similar articles Cited by Publication types MeSH terms Substances Associated data Related information LinkOut - more resources Comparative Study Gene . 1998 Nov 19;2...

Journal

Gene, 222(2)

Author

Sarker, Altaf

Publication Date

1998-11-19

DOI

https://pubmed.ncbi.nlm.nih.gov/9831664/

Data Availability

The data associated with this publication are available at: https://pubmed.ncbi.nlm.nih.gov/9831664/

Peer reviewed

ELSEVIER

Gene 222 (1998) 287–295

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

Genomic structure and sequence of a human homologue (NTHL1/NTH1) of *Escherichia coli* endonuclease III with those of the adjacent parts of *TSC2* and *SLC9A3R2* genes

Kotoe Imai [a,b], Altaf H. Sarker [a], Kosuke Akiyama [a], Shogo Ikeda [c], Ming Yao [a], Ken Tsutsui [a], Toshikiyo Shohmori [b], Shuji Seki [a,*]

[a] *Department of Molecular Biology, Institute of Cellular and Molecular Biology, Okayama University Medical School, Okayama 700-8558, Japan*

[b] *Department of Neurology, Okayama University Medical School, Okayama 700-8558, Japan*

[c] *Department of Biological Chemistry, Okayama University of Science, Okayama 700-0005, Japan*

Received 26 June 1998; received in revised form 4 September 1998; accepted 4 September 1998; Received by T. Sekiya

Abstract

Genomic cloning and sequencing of a human homologue (the gene name, endonuclease III-like 1; gene symbol, *NTHL1* or *NTH1*) for *Escherichia coli* endonuclease III, that is involved in pyrimidine base excision repair, were performed. The sequence covered the entire *NTHL1* gene consisting of six exons and five introns spanning 8 kb with 5′ flanking (8 kb) and 3′ flanking (3.8 kb) regions. Southern blot analysis suggested that the *NTHL1* gene exists as a single copy in a haploid genome. The sequenced 5′ flanking region lacks typical TATA and CAAT boxes, but contains a CpG island having putative binding sites for several transcription factors such as Ets1 and Sp1. The *NTHL1* gene lies immediately adjacent to the tuberous sclerosis 2 (*TSC2*) gene on chromosome 16p13.3 in a 5′-to-5′ orientation. Transcription initiation sites of both *NTHL1* and *TSC2* genes were suggested to be multiple by 5′ RACE experiments. The northern hybridization experiment suggested that both genes are expressed in all tissues, but at different levels. Downstream of the *NTHL1* gene, the gene for the regulatory factor 2 (SLC9A3R2/E3KARP; also called OCTS2, TKA-1 and SIP-1) of the solute carrier family 9 (sodium/hydrogen exchanger), isoform A3, lies in a 3′-to-3′ orientation. This paper demonstrates for the first time the spatial relationship of these three genes (*TSC2*, *NTHL1* and *SLC9A3R2*) at the nucleotide level, and the presence of multiple transcription initiation sites of the *NTHL1* and *TSC2* genes.

Keywords: Thymine glycol DNA glycosylase; AP lyase; *NTH1*; Tuberous sclerosis 2; E3KARP; *OCTS2* gene

1. Introduction

Oxidative DNA damage produced by endogenously and exogenously generated reactive oxygen species has been implicated in cell death, mutagenesis, carcinogenesis, and aging (Lindahl, 1993; Friedberg et al., 1995; Martin et al., 1996). Thymine residue of DNA is known to be especially susceptible to oxidative DNA damage, and the major detectable product is thymine glycol (Cathcart et al., 1984; Higgins et al., 1986). Ames and co-workers (Cathcart et al., 1984) have shown that the average human cell repairs about 320 thymine glycol sites/day based on an examination of the thymine glycol content of urine. Repair of the lesion is generally initiated by the DNA glycosylase activity of such enzymes as endonuclease III (endoIII) of *Escherichia coli* (*E. coli*) which is followed by the AP lyase activity of the same enzymes (Bailly and Verly, 1987; Friedberg et al., 1995; Krokan et al., 1997). Actually, endoIII is known to have a broad specificity for DNA base excision repair in that it removes numerous forms of modified thymine and cytosine bases from DNA (Krokan et al., 1997).

* Corresponding author. Tel: +81 86 235 7385; fax: +81 86 235 7392; e-mail: spseki@cc.okayama-u.ac.jp

Abbreviations: NTHL1 (alias NTH1, OCTS3), endonuclease III-like 1; TSC2, tuberous sclerosis 2; SLC9A3R2 (alias E3KARP, OCTS2, SIP-1, TKA-1), solute carrier family 9 (sodium/hydrogen exchanger), isoform 3 regulatory factor 2; RACE, rapid amplification of cDNA ends; BLAST, basic local alignment search tool; PCR, polymerase chain reaction; cDNA, DNA complementary to RNA; dNTP, deoxyribonucleoside triphosphate; kb, kilobase(s).

0378-1119/98/$ – see front matter
PII: S0378-1119(98)00485-5

Recently, the cDNAs for the eukaryotic homologues of endoIII have been cloned from yeast (Roldan-Arjona et al., 1996; Eide et al., 1996; Augeri et al., 1997), human (Aspinwall et al., 1997; Hilbert et al., 1997; Ikeda et al., 1998, accession no. AB001575) and mouse (Sarker et al., 1998, accession no. AB006812). The human and mouse homologues (the corresponding gene name is endonuclease III-like 1 and the gene symbols are *NTHL1* and *Nth11*, respectively) have been shown to have thymine glycol DNA glycosylase, urea DNA glycosylase and AP lyase activities (Aspinwall et al., 1997; Hilbert et al., 1997; Sarker et al., 1998; Ikeda et al., 1998) similar to those of endonuclease III.

Concerning the genomic organization, the gene for the human *NTHL1* (also called *NTH1* and *OCTS3*) has been shown to lie on chromosome 16p13.3, immediately adjacent to the tuberous sclerosis 2 (*TSC2*) gene in a 5′-to-5′ orientation (Hilbert et al., 1997; Aspinwall et al., 1997), and adjacent to the *OCTS2* gene (European Chromosome 16 Tuberous Sclerosis Consortium, 1993). Several tuberous sclerosis patients with large deletions affecting both the *TSC2* and *NTHL1*/*NTH1*/*OCTS3* genes have been described (European Chromosome 16 Tuberous Sclerosis Consortium, 1993). Considering the possibility that some of the highly variable phenotypes observed in TSC patients might be due to the deletion of the *NTHL1*/*NTH1* (and/or *OCTS2*) gene contiguous to the *TSC2* gene, it is important to know the genomic structure and sequence of the *NTHL1*/*NTH1* gene and its relation to the neighboring genes (*TSC2* and *OCTS2*) for genetic analyses of TSC patients.

In the present paper, the genomic organization and the sequence of a 19.8 kb genomic fragment containing the *NTHL1* gene and partial sequences of the neighboring genes (*TSC2* and *OCTS2* genes) are reported. The *OCTS2* gene is shown to be the gene for the regulatory factor 2 (SLC9A3R2/E3KARP; also called OCTS2, TKA1 and SIP1) of solute carrier family 9 (Na^+/H^+ exchanger), isoform A3, whose cDNA sequence was reported recently (Yun et al., 1997; Poulat et al., 1997; Reczek et al., 1997).

2. Materials and methods

2.1. Materials

The reagents used in these experiments were obtained from the following sources: restriction enzymes from Toyobo Biochemicals, Osaka, Japan; Moloney murine leukemia virus reverse transcriptase from Gibco BRL Life Technologies; RNasin from Promega; Megaprime™ DNA labeling system, [α-^{32}P]dCTP (3000 Ci/mmol) and positively charged nylon membrane (Hybond-N^+) from Amersham. The human *TSC2* cDNA clone used for northern blot analysis was kindly provided by Dr. Mark Nellist of the Department of Clinical Genetics, Erasmus University, Rotterdam, Netherlands. Plasmid containing cDNA for the human S26 ribosomal protein was kindly provided by Professor J. Fujita of Kyoto University, Kyoto, Japan (Kaneko et al., 1995). The cDNA for the human homologue (*NTHL1*/*NTH1*) of endonuclease III was cloned by screening a human bone marrow cDNA library with the ^{32}P-labeled mouse *Nth11*/*Nth1* cDNA (Sarker et al., 1998). A human leukocyte genomic library in EMBL-3, *E. coli* strain NM538, a 'Human Multiple Tissue Northern Blot', 'ExpressHyb Hybridization Solution' and 'Human Brain PolyA^+ RNA' were obtained from Clontech Laboratories, Inc. A human placenta genomic library in Lambda FIXII and *E. coli* strain XL1-Blue MRA(P2) were obtained from Stratagene. ABI Prism™ Dye Terminator Cycle Sequencing Ready Reaction Kit was obtained from Perkin Elmer Corp.

2.2. Southern blot analysis of human genomic DNA

High molecular weight DNA was isolated from human peripheral leukocytes as described (Sambrook et al., 1989). DNA (5 μg) digested with restriction enzymes were electrophoresed on a 0.7% agarose gel in an electrophoresis buffer containing 89 mM Tris base, 89 mM boric acid and 2 mM EDTA (pH 8.0), and transferred onto a nylon membrane (Hybond N+). The ^{32}P-labeled probe was prepared using the *NTHL1*/*NTH1* cDNA as a template and [α-^{32}P]dCTP (3000 Ci/mmol) as the labeled nucleotide by the random priming method (Sambrook et al., 1989). The blotted membrane was hybridized with the ^{32}P-labeled probe. Hybridization was carried out at 68°C for 15 h in 6 × SSC (1 × SSC: 150 mM NaCl and 15 mM sodium citrate, pH 7.0), 5 × Denhardt's reagent (1 × Denhardt's reagent: 0.02% Ficoll, 0.02% polyvinylpyrrolidone and 0.02% bovine serum albumin), 0.5% sodium dodecylsulfate (SDS) and 100 μg/ml denatured, fragmented salmon testes DNA. The membrane was washed at room temperature in 0.1% SDS-2 × SSC, and then at 68°C in 0.1% SDS-1 × SSC. After washing, the membrane was air-dried and autoradiographed at −80°C using Fuji X-ray film with an intensifying screen.

2.3. Cloning of the NTHL1 gene

Recombinant bacteriophage clones in a human leukocyte genomic library in bacteriophage vector EMBL-3 (Clontech) were plated with *E. coli* NM538. Approximately 1×10^6 bacteriophage clones were screened using the ^{32}P-labeled *NTHL1*/*NTH1* cDNA probe by the plaque hybridization technique (Sambrook et al., 1989).

Hybridization and washing were conducted as described in Section 2.2. 12 positive clones were

obtained, and the recombinant phage DNA was isolated from the positive clones by the plate lysate method. The isolated DNA was analyzed by digestion with restriction enzymes and southern blot analysis using the *NTHL1*/*NTH1* cDNA. From these clones, a 7 kb *Eco*RI fragment of the clone C was selected for further analysis. The 7 kb fragment subcloned into plasmid vector pBluescript KS(−) had a 3′ region of the *NTHL1* gene (from the intron 4 to the 3′ flanking region). The 5′ region of the *NTHL1* gene was not present in the 12 clones isolated from the human leukocyte genomic library. Then, a human placenta genomic library in bacteriophage vector λ (lambda) FIXII Stratagene was screened with a 5′ region (*Nar*I fragment) of the *NTHL1*/*NTH1* cDNA to obtain a 5′ region of the *NTHL1* gene. From eight positive clones, the clone 3 having a 17 kb insert was selected for further analysis. The 17 kb fragments were subcloned into the plasmid vector pBluescript KS(−) for sequencing.

2.4. DNA sequencing and sequence analysis

DNA sequence was determined on both strands by the dideoxy chain termination method (Sanger et al., 1977) using double-stranded templates and the modified AmpliTaq DNA polymerase (AmpliTaq DNA polymerase, FS, Perkin Elmer). Sequence primers used were the M13 universal primers and specific oligonucleotide primers synthesized according to the sequences determined.

The CpG islands (CpG-rich regions) were defined and determined according to Gardiner-Garden and Frommer (1987), as described previously (Akiyama et al., 1994, 1995).

2.5. Northern blot analysis

The 'Human Multiple Tissue Northern Blot' purchased from Clontech contained polyadenylated RNA (2 μg per lane) from the human heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas. The blotted membrane was hybridized with the ^{32}P-labeled *NTHL1*/*NTH1* cDNA probe. Hybridization was performed at 68°C for 1 h in ExpressHyb Hybridization Solution (Clontech). The membrane was washed at room temperature for 40 min in 0.05% SDS-2 × SSC, and then at 50°C in 0.1% SDS-0.1 × SSC. After washing, the membrane was processed for autoradiography. Northern blot analysis was also performed using the same membrane with the ^{32}P-labeled *TSC2* cDNA and the ^{32}P-labeled S26 ribosomal protein cDNA probes. The latter was used as an internal standard to normalize RNA levels on northern blots (Vincent et al., 1993).

2.6. Analysis by 5′ RACE and RT-PCR techniques

The transcription initiation site(s) was analyzed by sequencing the 5′ ends of *NTHL1*/*NTH1*-specific or *TSC2*-specific cDNAs produced by the 5′ RACE (rapid amplification of 5′ cDNA ends) technique (Hirzmann et al., 1993). 0.5 μg of polyadenylated RNA isolated from a human whole brain (Clontech) was heated to 65°C for 3 min and cooled on ice. The reverse transcription reaction was performed at 37°C for 2 h in a 10 μl reaction mixture containing 0.5 μg of the 65°C-treated RNA, 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 1 mM each dNTP, 0.5 unit/μl RNasin (Promega Biotec), 1 μM oligonucleotide primer (the hRACE2 primer for *NTHL1*/*NTH1*, 5′-ATGGCACGGATGTTGACCAG-3′; the R10 primer for *TSC2*, 5′-TGCCCTATCATCCGGATGCG-3′), and 10 units/μl Moloney murine leukemia virus reverse transcriptase. Excess primer molecules and substrate were removed by washing six times through a Microcon-30 membrane (Amicon Corporation; washing each time with 0.5 ml of 10 mM Tris-HCl, 1 mM EDTA, pH 8.0). For dG tailing, an aliquot of the reaction products (equivalent to 140 ng mRNA) was incubated at 37°C for 20 min in a 20 μl reaction mixture containing 100 mM cacodylate buffer, pH 6.8, 1 mM $CoCl_2$, 0.1 mM DTT, 100 μg/ml bovine serum albumin, 0.2 mM dGTP, and six units terminal deoxynucleotidyl transferase (Promega Biotec). The reaction was stopped by heating at 75°C for 10 min. The dG-tailed products were used for PCR amplification with an EcoRI-BamHI-(dC)12-oligonucleotide primer (EBPC: 5′-TTCT-

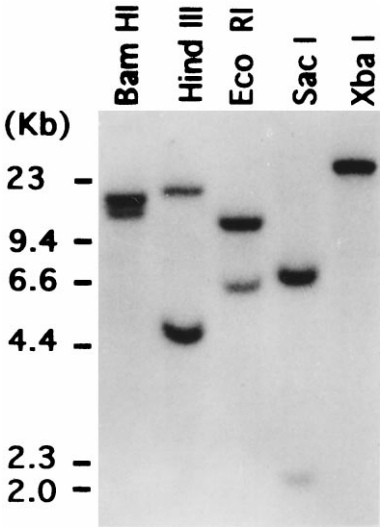

Fig. 1. Southern blot analysis of the *NTHL1* gene. Human leukocyte genomic DNA was digested with the restriction enzymes indicated on the top. 5 μg of DNA per lane was subjected to electrophoresis on 0.7% agarose gel and then transferred to a nylon membrane. The blots were hybridized with the ^{32}P-labeled *NTHL1*/*NTH1* cDNA. Numbers in the left column indicate size markers in kilobase pairs.

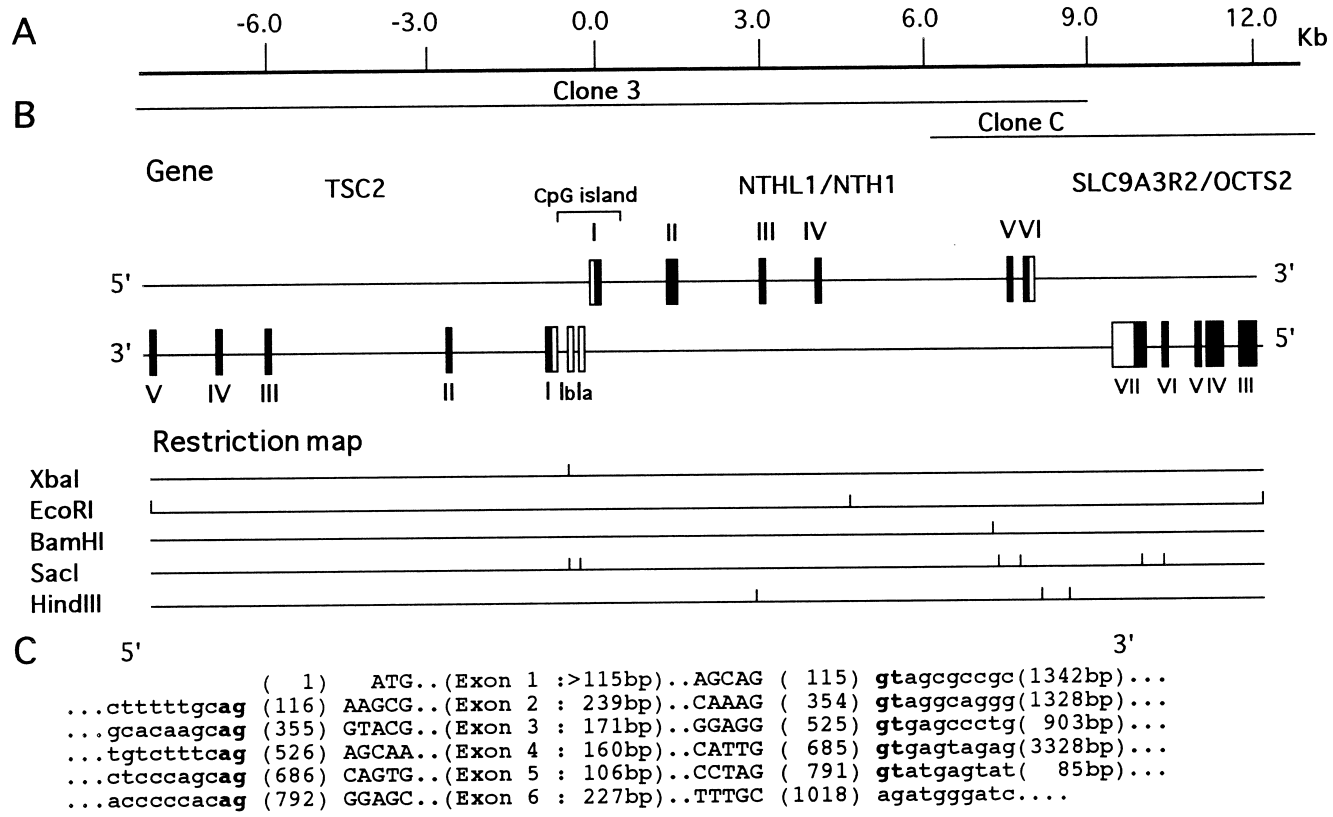

Fig. 2. Organization and restriction maps of the human *NTHL1* gene and its flanking regions. In order to assign tentatively a base position within the gene and in the flanking regions, the adenine residue in the translation initiation codon of the *NTHL1* gene was designated as position +1 and the 5′ neighboring residue cytosine as position −1. With this as a starting point, positive numbers are given to positions downstream from the starting point and negative numbers to those upstream. (A) Scale in kilobases. (B) (Top) Inserts of overlapping clones 3 and C sequenced. (Middle) A diagram of the *NTHL1* gene and portions of the neighboring *TSC2* and *SLC9A3R2/OCTS2* genes. The exons are numbered from the 5′ end of the genes. Filled and open boxes represent translated and untranslated regions, respectively. (Bottom) The restriction maps of the *NTHL1* gene and its flanking regions. (C) Nucleotide sequences of the exon–intron boundaries of the *NTHL1* gene. Sequences of introns are given in lowercase letters, and those of exons in uppercase letters. On both sides of the parentheses for the number and size of each exon, the first and last five nucleotides are shown. The nucleotide numbers of the first and last residue of each exon are shown in parentheses before and after the first and last five nucleotides, respectively. The size of each intron is shown in parentheses on the right side.

$AGAATTCGGATC_{12}$-3′; Hirzmann et al., 1993) and a specific primer for *NTHL1/NTH1* (exo-1R5 primer, 5′-CCCGGGCTCCTCCCTACACC-3′), or for *TSC2* (TSC2e1RA primer for the first round RACE, 5′-TGGCCTCGGTGTTCCCAGTC-3′; TSC2e1RB primer for the second round RACE, 5′-TGTTGGTTTGGCCATGGTGG-3′). The RACE products were cloned into a pGEM-T-plasmid vector (Promega) and sequenced.

3. Results

3.1. Southern blot analysis

Southern blot analysis of human genomic DNA was performed using the ^{32}P-labeled *NTHL1/NTH1* cDNA probe which covers almost the entire transcribed region. Restriction enzymes which do not cut the *NTHL1/NTH1* cDNA (*Xba*I, *Eco*RI, *Bam*HI and *Sac*I) or cut at a single site (*Hind*III) were used. The restriction fragment lengths determined by the southern blot experiment correspond well to the lengths deduced from the sequence analysis of the 19.8 kb genomic fragment containing the whole region of the *NTHL1* gene (Fig. 1Fig. 2). The present results suggest that the human *NTHL1* gene is present as a single copy in a haploid genome, and that a gene(s) highly homologous but not identical to the *NTHL1* gene does not exist in the human genome.

3.2. Isolation, sequencing and analysis of the human NTHL1 gene

Two human genomic libraries were screened using the *NTHL1/NTH1* cDNA as a probe. 20 positive clones were isolated from 2×10^6 independent clones. The insert DNAs from two overlapping clones (designated as clones 3 and C, respectively) were subcloned into the plasmid vector pBluescript KS(−), and sequenced. The

resultant 19.8 kb sequence contained the entire region of the *NTHL1* gene (accession no. AB014460). By comparing the sequence with the published *NTHL1/NTH1* cDNA sequence (accession no. AB001575), we assigned the exon/intron structure of the human *NTHL1* gene (Fig. 2). The gene consists of six exons and five introns, with a total length of 8 kb. All exon/intron junctions follow the GT/AG rule (Breathnach and Chambon, 1981). The nucleotide sequence of each exon is perfectly identical to that of the corresponding region of the *NTHL1/NTH1* cDNA deposited by Ikeda et al. (1998); accession no. AB001575. A base position within the *NTHL1/NTH1* gene and in the flanking regions is numbered tentatively as described in the legend to Fig. 2. The translation initiation codon (ATG) of the *NTHL1* gene is located in the exon 1 (Fig. 3). The translation termination codon (TGA) and the polyadenylation signal (AATAAA) of the *NTHL1* gene are located in the exon 6 at base positions 7899/7901 and 7983/7988, respectively.

3.3. Genes adjacent to the NTHL1 gene

A previous study on tuberous sclerosis (TSC)-associated deletion at 16p13.3 showed that a number of genes are located in regions deleted in one or more affected patients (European Chromosome 16 Tuberous Sclerosis Consortium, 1993). In addition to the *TSC2* gene and the polycystic kidney disease (*PKD1*) gene that lies immediately proximal to *TSC2*, two genes (*OCTS3* and *OCTS2*) lying immediately distal to *TSC2* were partially characterized (European Chromosome 16 Tuberous Sclerosis Consortium, 1993). The 19.8 kb genomic fragment sequenced contained the *NTHL1* gene (8 kb) with the 8 kb, 5′ flanking region and the 3.8 kb, 3′ flanking region (Fig. 2). Analysis of the 5′ flanking sequence of the *NTHL1* gene using the human *TSC2* mRNA sequence (European Chromosome 16 Tuberous Sclerosis Consortium, 1993; accession no. X75621) revealed that the *NTHL1* gene lies immediately adjacent to the *TSC2* gene in a 5′-to-5′ orientation and coincides with the *OCTS3* gene reported previously (Aspinwall et al., 1997) (Figs. 2 and 3). BLAST analysis of the 3′ flanking sequence, corresponding to the region of the *OCTS2* gene, of the *NTHL1* gene revealed that the sequence contains a partial genomic sequence for a double PDZ-containing protein (denoted in Fig. 2 as *SLC9A3R2/OCTS2*), as discussed later.

3.4. Further analysis of possible promoter regions of the NTHL1 and TSC2 genes

To identify the 5′ termini of *NTHL1* and *TSC2* transcripts, we performed 5′ RACE experiments using the polyadenylated RNA isolated from a human whole brain (Clontech). The 5′ ends of the 5′ RACE products for the *NTHL1* transcripts were mapped on multiple positions within the exon 1 of the *NTHL1* gene (Fig. 3A).

The 5′ ends of the *TSC2* transcripts were distributed into three exons (exons 1, 1a and 1b) (Fig. 3A). The exon 1 of the *TSC2* gene was the exon which was previously assigned by the European Chromosome 16 Tuberous Sclerosis Consortium (1993). Significant numbers of the 5′ RACE products were also mapped in two novel exons (designated exons 1a and 1b). The presence of the exon 1a in the *TSC2* gene was recently suggested by Kobayashi et al. (1997) in the rat *Tsc2* gene and human *TSC2* gene and Seki et al. (1997) in the mouse *Tsc2* gene. Our RACE experiment confirmed the presence of the exon 1a in the human *TSC2* gene. The 5′ ends of the RACE products belonging to exons 1 and 1a were mapped on multiple positions as shown in Fig. 3A. Some of the RACE products carried an extra G residue in their 5′ end sequences that is not present in the sequence of the genomic clone. This extra G residue is thought to be derived from the terminal cap G nucleotide, as reported by Hirzmann et al. (1993) (Akiyama et al., 1995). Another novel leader sequence, exon 1b, was also detected in the RACE products (Figs. 2 and 3A). The RACE products belonging to the exon 1b carried the extra G residue. The exon 1a-containing transcripts showed that the exon 1a joins not via the exon 1b, but directly with the exon 1 (Fig. 3B).

The present results indicate that the transcription initiation sites of the *TSC2* are multiple and distributed into three exons (the exons 1, 1a and 1b). The closest exons for the *NTHL1* and *TSC2* genes, the exon 1 (125 bp) of the *NTHL1* gene and the exon 1a (83 bp) of the *TSC2* gene, lie within the 357 bp-sequence in a head-to-head orientation (Fig. 3A).

Multiple transcription initiation sites suggest the presence of multiple (alternative) promoters (Ayoubi and Van de Ven, 1996). Referring to their patterns of structural organization of genes containing alternative promoters, the case of the *NTHL1* gene is thought to be classified in the pattern C-like, where the leader exons are overlapping, and the case of the *TSC2* gene in the pattern B-like, where alternative promoter usage will result in alternative first exons that will be spliced to a common second exon (Ayoubi and Van de Ven, 1996). The promoters of *NTHL1* and *TSC2* genes may be overlapped or bidirectional. A CpG island is found between the sequences of the intron 1b of the *TSC2* gene and the intron 1 of the *NTHL1* gene. The 5′ flanking region of the *NTHL1* gene contains consensus sequences for Sp1, Ets1, LBP-1 and others. The 5′ flanking region of the *TSC2* gene also contains consensus sequences for Sp1, Ets1, NF-IL6, AP2 and others. The putative promoter regions of the *NTHL1* and

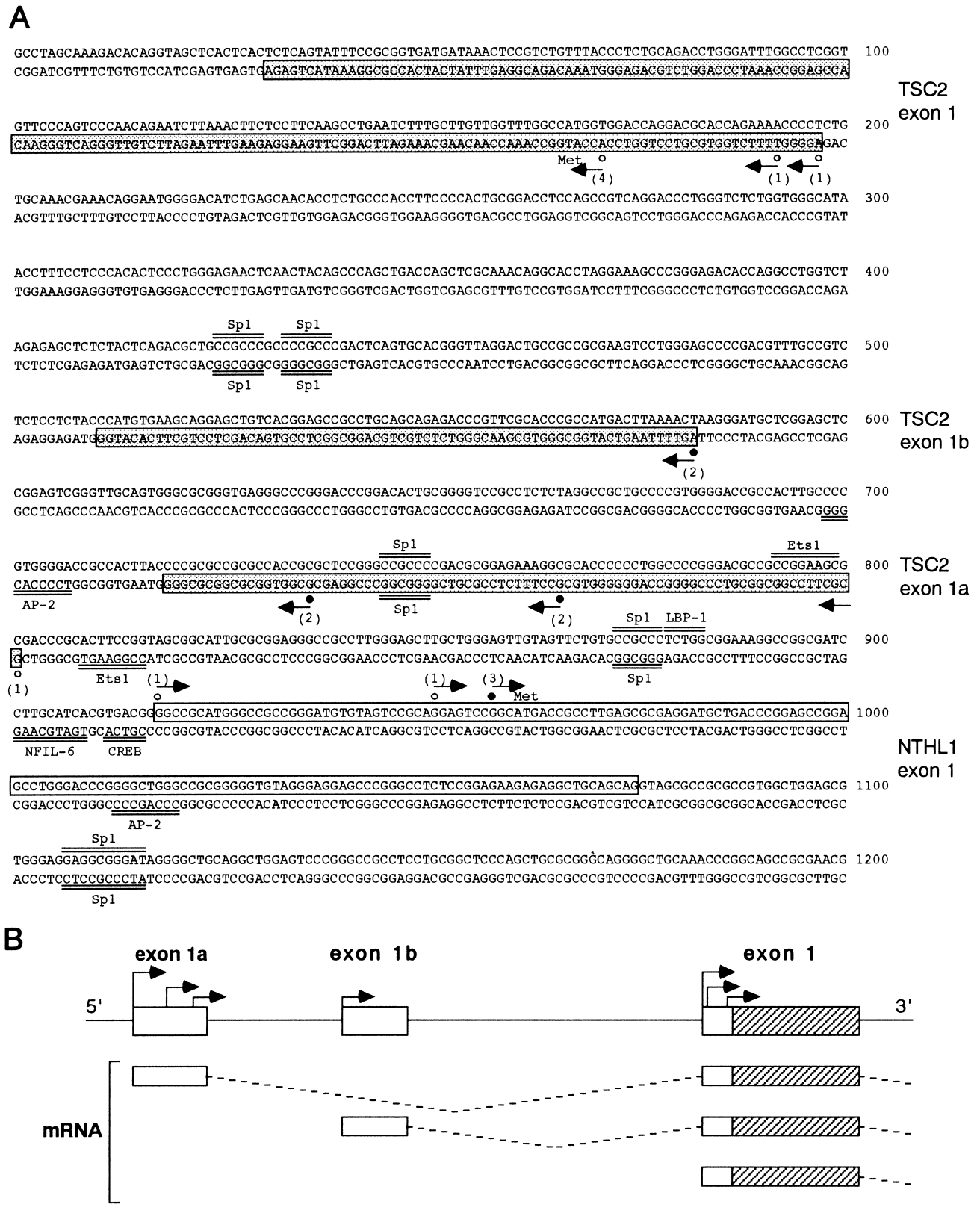

Fig. 3. Putative promoter regions and multiple transcription initiation sites of the *NTHL1* and *TSC2* genes. (A) Nucleotide sequence of the 5′ flanking and 5′ end region of the human *NTHL1* gene. Two line sequences complementary to each other are shown. The upper, downward sequence is the sense strand for *NTHL1*, and the lower, upward sequence is the sense strand for *TSC2*. Open- and shaded-boxed sequences are the *NTHL1* and *TSC2* exons, respectively, as indicated in the right column. The 5′ ends were determined by the 5′ RACE technique (see Section 2.6). The RACE products were cloned into a pGEM-T-plasmid vector (Promega), and the positive clones were picked up randomly and sequenced. The 5′

TSC2 genes both have no typical TATA and CAAT boxes.

3.5. Northern blot analysis

Northern blot analyses of expression of the *NTHL1* and *TSC2* genes using the same 'Northern Blot Membrane' showed that all of the examined tissues express *NTHL1/NTH1* mRNA as well as *TSC2* mRNA, although the variations of expression levels among different tissues are high (Fig. 4A,B). The results also showed the tendency of some tissues (heart and brain), which show relatively high expression of *NTHL1/NTH1* mRNA, to show high expression of *TSC2* mRNA. Further analysis, in the form of well-controlled experiments, is necessary to better understand these tissue-to-tissue variations in expression.

4. Discussion

A 19.8 kb genomic fragment containing the *NTHL1* gene, a gene for a human homologue (*NTHL1/NTH1*) of *E. coli* endonuclease III, and portions of *TSC2* and *SLC9A3R2* genes was cloned and sequenced.

The present study confirmed at the sequence level that the human *NTHL1/NTH1* gene lies immediately adjacent to the *TSC2* gene in a 5′-to-5′ orientation as reported previously (Aspinwall et al., 1997; Hilbert et al., 1997; Seki et al., 1997; Kobayashi et al., 1997; Sarker et al., 1998). Genomic deletions at the TSC2 locus on chromosome 16p13.3 are known to cause tuberous sclerosis (TSC) (European Chromosome 16 Tuberous Sclerosis Consortium, 1993). TSC is a dominantly inherited hamartomatous disorder characterized by seizures, mental retardation, skin lesions, eye lesions and widespread development of hamartomatous and tumorous growths. The *TSC2* gene has also been suggested to function as a tumor suppressor gene, in accordance with Knudson's hypothesis (Green et al., 1994; Yeung et al., 1994). The TSC-associated genomic deletions at the TSC2 locus are fairly large, frequently involving not only the *TSC2* gene but also the neighboring genes termed *OCTS2* and *OCTS3* genes (European Chromosome 16 Tuberous Sclerosis Consortium, 1993). Recent and present studies on the *NTHL1/NTH1* gene revealed that the *NTHL1/NTH1* gene coincides with the *OCTS3* gene (Aspinwall et al., 1997; Hilbert et al., 1997; Seki et al., 1997).

Both the *NTHL1* and *TSC2* genes possess such properties common to housekeeping genes as ubiquitous expression, putative promoters in a CpG island, multiple transcription initiation sites and no typical TATA box. These genes may mutually influence each other's expression, because of the fact that putative promoters of both genes are overlapping (or bidirectional). The results of northern blot analysis showed that both *NTHL1* and *TSC2* genes are expressed ubiquitously, and that there is a tendency for some tissues showing a relatively high expression of *NTHL1/NTH1* mRNA also to show a high expression of *TSC2* mRNA, although the correlation between their expression levels may not be so tight.

The 3′ flanking sequence (3.8 kb) of the *NTHL1* gene contained a 3′ region of the *OCTS2* gene, which lies adjacent to the *NTHL1* (*OCTS3*) gene in a 3′-to-3′ orientation (European Chromosome 16 Tuberous Sclerosis Consortium, 1993). BLAST analyses of the 3′ flanking sequence of the *NTHL1* gene revealed that the *OCTS2* gene coincides with the gene for a double PDZ domain-containing protein variably designated as follows: Based on the protein properties characterized, it has been called Na^+/H^+ exchanger 3 kinase A regulatory protein (E3KARP; Yun et al., 1997), SRY-interacting protein 1 (SIP-1; Poulat et al., 1997), tyrosine kinase activator 1 (TKA-1; Reczek et al., 1997), and Na^+/H^+ exchanger regulatory factor 2 (NHERF-2 in Data Bases, accession no. AF035771). The HUGO/GDB-approved human gene symbol for Na^+/H^+ exchanger 3 (NHE3) is the SLC9A3 [solute carrier family 9 (sodium/hydrogen exchanger), isoform A3]. Two regulatory factors (NHERF and E3KARP/NHERF-2) for the SLC9A3 protein have been reported (Yun et al., 1997). Therefore, the *SLC9A3R2* (regulatory factor 2 for the SLC9A3) is adopted here as the gene symbol for the *OCTS2* gene encoding E3KARP/NHERF-2. The human *SLC9A3R2/OCTS2* gene consists of seven exons spanning 11.5 kb (Seki et al., unpublished results). The 3′ flanking sequence (3.8 kb) of the *NTHL1* gene contained exons 3 through 7 of the SLC9A3R2 gene.

Combining the present results with the previous information (European Chromosome 16 Tuberous Sclerosis Consortium, 1993) on genomic organization of polycystic kidney disease 1 (*PKD1*) and *TSC2* genes, these

ends determined are indicated by open and dark circles. The number in brackets over or under each circle indicates the number of independent clones in which the 5′ end of the insert was located at this position. Within these positions, dark circles indicate the positions assigned by the 5′ RACE products having an extra G residue which is thought to be derived from the terminal cap G nucleotides. Arrows indicate the direction of transcription. The codons indicated by Met over the *NTHL1* exon 1 sequence and under the *TSC2* exon 1 sequence indicate the translation initiation sites for the *NTHL1* and *TSC2* genes, respectively. Consensus sequences of potential binding sites for some transcription factors are indicated with double lines. (B) Structural organization of 5′ end regions of TSC2 transcripts. Exons are depicted as boxes. Coding regions are shaded. Intervening sequences are depicted by solid lines. Dotted lines connecting exons indicate splicing patterns.

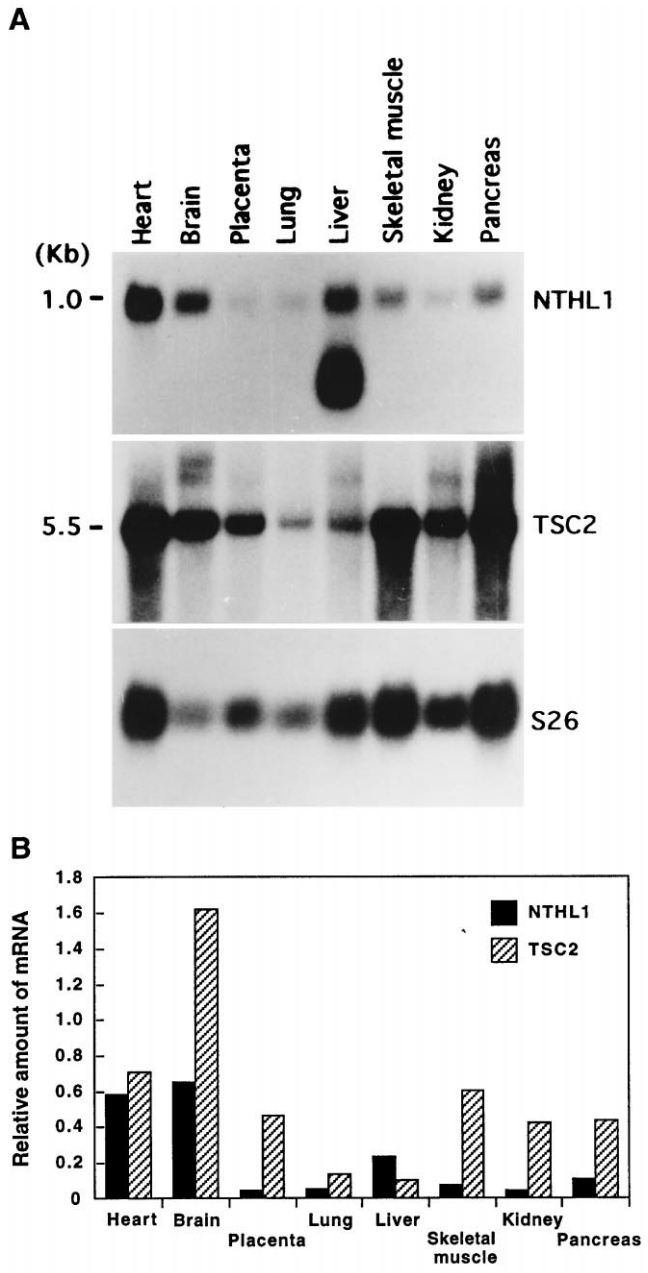

Fig. 4. Expression of *NTHL1* and *TSC2* genes in human tissues. (A) Northern blot analyses. Polyadenylated RNA obtained from the indicated human tissues was electrophoresed at 2 μg in each lane and blotted on a nylon membrane ('Human Multiple Tissue Northern Blot', Clontech). The membrane was hybridized with the ^{32}P-labeled *NTHL1*/*NTH1* cDNA probe and later with the ^{32}P-labeled *TSC2* cDNA probe. The same membrane was rehybridized with the cDNA probe for the human S26 ribosomal protein to correct for the amount of mRNA loaded. The smaller massage (~0.5 kb) present only in liver may be due to alternate splicing or specific degradation. (B) Relative levels of *NTHL1* and *TSC2* mRNAs in various tissues. Autoradiograms of the northern blots shown above were scanned and hybridization signals were quantified using a densitometric software (Intelligent Quantifier, Bio Image). Data were normalized against the S26 ribosomal standard.

genes on chromosome 16p13.3 are arranged in order of *PKD1*, *TSC2*, *NTHL1* and *SLC9A3R2*/*OCTS2* genes from centromere to telomere, and the gene orientations are 3′-to-3′ between the *PKD1* and *TSC2* genes, 5′-to-5′ between *TSC2* and *NTHL1*, and 3′-to-3′ between *NTHL1* and *SLC9A3R2*/*OCTS2* (Fig. 2). Considering the fact that tuberous sclerosis-associated deletion at 16p13.3 is fairly large and frequently involves deletion of the *NTHL1*/*NTH1*/*OCTS3* and *SLC9A3R2*/*OCTS2* genes in addition to the deletion of the *TSC2* gene (European Chromosome 16 Tuberous Sclerosis Consortium, 1993), some of a wide variety of signs, symptoms and complications observed in TSC patients may be caused by abnormal *NTHL1*/*NTH1* and/or *SLC9A3R2*/*OCTS2* gene products.

The present study will further facilitate our understanding of the genetic bases of tuberous sclerosis as well as the products of the *NTHL1*/*NTH1* and *SLC9A3R2*/*OCTS2* genes.

5. Note

The nucleotide sequence data reported in this paper have been deposited in the GSDB, DDBJ, EMBL, and NCBI nucleotide databases under the accession no. AB014460.

Acknowledgement

We thank Dr. Mark Nellist of the Department of Clinical Genetics, Erasmus University, Rotterdam, Netherlands for providing a human *TSC2* cDNA clone, Dr. J. Fujita of Kyoto University for providing a plasmid containing cDNA of the human S26 ribosomal protein, and Mr. T. Nakamura and Ms. K. Tabuchi for their technical assistance. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture and in part by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports and Culture of Japan (Repair, Recombination and Mutagenesis, 08280101).

References

Akiyama, K., Seki, S., Oshida, T., Yoshida, M.C., 1994. Structure, promoter analysis and chromosomal assignment of the human *APEX* gene. Biochim. Biophys. Acta 1219, 15–25.

Akiyama, K., Nagao, K., Oshida, T., Tsutsui, K., Yoshida, M.C., Seki, S., 1995. Cloning, sequence analysis, and chromosomal assignment of the mouse *Apex* gene. Genomics 26, 63–69.

Aspinwall, R., Rothwell, D.G., Roldan-Arjona, T., Anselmino, C., Ward, C.J., Cheadle, J.P., Sampson, J.R., Lindahl, T., Harris, P.C., Hickson, I.D., 1997. Cloning and characterization of a functional

human homolog of *Escherichia coli* endonuclease III. Proc. Nat. Acad. Sci. USA 94, 109–114.

Augeri, L., Lee, Y.M., Barton, A.B., Doetsche, P.W., 1997. Purification, characterization, gene cloning, and expression of saccharomyces cerevisiae redoxyendonuclease, a homologue of *Escherichia coli* endonuclease III. Biochemistry 36, 721–729.

Ayoubi, T.A.Y., Van de Ven, W.J.M., 1996. Regulation of gene expression by alternative promoters. FASEB J. 10, 453–460.

Bailly, V., Verly, W.G., 1987. *Escherichia coli* endonuclease III is not an endonuclease but a β-elimination catalyst. Biochem. J. 242, 565–572.

Breathnach, R., Chambon, P., 1981. Organization and expression of eukaryotic split genes coding for proteins. Ann. Rev. Biochem. 50, 349–383.

Cathcart, R., Schwiers, E., Saul, R.L., Ames, B.N., 1984. Thymine glycol and thymidine glycol in human and rat urine: A possible assay for oxidative DNA damage. Proc. Nat. Acad. Sci. USA 81, 5633–5637.

Eide, L., Bjoras, M., Pirovano, M., Alseth, I., Berdal, K.G., Seeberg, E., 1996. Base excision of oxidative purine and pyrimidine DNA damage in *Saccharomyces cerevisiae* by a DNA glycosylase with sequence similarity to endonuclease III from *Escherichia coli*. Proc. Nat. Acad. Sci. USA 93, 10735–10740.

European Chromosome 16 Tuberous Sclerosis Consortium, 1993. Identification and characterization of the tuberous sclerosis gene on chromosome 16. Cell 75, 1305–1315.

Friedberg, E.E., Walker, G.C., Siede, W., 1995. DNA Repair and Mutagenesis. ASM Press, Washington, DC.

Gardiner-Garden, M., Frommer, M., 1987. CpG islands in vertebrate genomes. J. Mol. Biol. 196, 261–282.

Green, A.J., Smith, M., Yates, J.R., 1994. Loss of heterozygosity on chromosome 16p13.3 in hamartomas from tuberous sclerosis patients. Nature Genet. 6, 193–196.

Higgins, S.A., Frenkel, K., Cummings, A., Teebor, G.W., 1986. Definitive characterization of human thymine glycol N-glycosylase activity. Biochemistry 26, 1683–1688.

Hilbert, T.P., Chuang, W., Boorstein, R.J., Cunningham, R.P., Teebor, G.W., 1997. Cloning and expression of the cDNA encoding the human homologue of the DNA repair enzyme, *Escherichia coli* endonuclease III. J. Biol. Chem. 272, 6733–6740.

Hirzmann, J., Luo, D., Hahnen, J., Hobom, G., 1993. Determination of messenger RNA 5′-ends by reverse transcription of the cap structure. Nucleic Acids Res. 21, 3597–3598.

Ikeda, S., Biswas, T., Roy, R., Izumi, T., Boldogh, I., Kurosky, A., Sarker, A.H., Seki, S., Mitra, S., 1998. Purification and characterization of a human homolog (hNTH1) of *Escherichia coli* endonuclease III: Direct identification of lys-212 as the nucleophilic residue. J. Biol. Chem. 273, 21585–21593.

Kaneko, Y., Nonoguchi, K., Fukuyama, H., Takano, S., Higashitsuji, H., Nishiyama, H., Takenawa, J., Nakayama, H., Fujita, J., 1995. Presence of alternative 5′ untranslated sequences and identification of cells expressing ctk transcripts in the brain and testis. Oncogene 10, 945–952.

Kobayashi, T., Urakami, S., Cheadle, J.P., Aspinwall, R., Harris, P., Sampson, J.R., Hino, O., 1997. Identification o a leader exon and a core promoter for the rat tuberous sclerosis 2 (*TSC2*) gene and structural comparison with human homolog. Mammalian Genome 8, 554–558.

Krokan, H.E., Standal, R., Slupphaug, G., 1997. DNA glycosylases in the base excision repair of DNA. Biochem. J. 325, 1–16.

Lindahl, T., 1993. Instability and decay of the primary structure of DNA. Nature (London) 362, 709–715.

Martin, G.M., Austad, S.N., Johnson, T.E., 1996. Genetic analysis of ageing: role of oxidative damage and environmental stresses. Nature Genet. 13, 25–34.

Poulat, F., Barbara, P.S., Desclozeaux, M., Soullier, S., Moniot, B., Bonneaud, N., Boizet, B., Berta, P., 1997. The human testis determining factor SRY binds a nuclear factor containing PDZ protein interaction domains. J. Biol. Chem. 272, 7167–7172.

Reczek, D., Berryman, M., Bretscher, A., 1997. Identification of EBP50: a PDZ-containing phosphoprotein that associates with members of the ezrin–radixin–moesin family. J. Cell Biol. 139, 169–179.

Roldan-Arjona, T., Anselmino, C., Lindahl, T., 1996. Molecular cloning and functional analysis of a *Schizosaccharomyces pombe* homologue of *Escherichia coli* endonuclease III. Nucleic Acids Res. 24, 3307–3312.

Sambrook, J., Eritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Plainview, NY.

Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Nat. Acad. Sci. USA 74, 5463–5467.

Sarker, A.H., Ikeda, S., Nakano, H., Terato, H., Ide, H., Akiyama, K., Tsutsui, K., Bo, Z., Kubo, K., Yamamoto, K., Yasui, A., Yoshida, M.C., Seki, S., 1998. Cloning and characterization of a mouse homologue (mNthl1) of *Escherichia coli* endonuclease III. J. Mol. Biol. 282, 761–774.

Seki, S., Sarker, A.H., Ikeda, S., Akiyama, K., Imai, K., Tsutsui, K., Bo, Z., Yamamoto, K., Yasui, A., Kubo, K., Yoshida, M.C., 1997. cDNA and gene cloning of mammalian homologues of *Escherichia coli* endonuclease III. FASEB J. 11, A1365

Vincent, S., Marty, L., Fort, P., 1993. S26 ribosomal protein RNA: an invariant control for gene regulation experiments in eukaryotic cells and tissues. Nucleic Acids Res. 21, 1498

Yeung, R.S., Xiao, G.-H., Jin, F., Lee, W.-C., Testa, J.R., Knudson, A.G., 1994. Predispositon to renal carcinoma in the Eker rat is determined by germ-line mutation of the tuberous sclerosis 2 (*TSC2*) gene. Proc. Nat. Acad. Sci. USA 91, 11413–11416.

Yun, C.H.C., Oh, S., Zizak, M., Steplock, D., Tsao, S., Tse, C.-M., Weinman, E.J., Donowitz, M., 1997. cAMP-mediated inhibition of the epithelial brush border Na^+/H^+ exchanger, NHE3, requires an associated regulatory protein. Proc. Nat. Acad. Sci. USA 94, 3010–3015.