

ARTICLE

Isolation of a candidate human telomerase catalytic subunit gene, which reveals complex splicing patterns in different cell types

Andrzej Kilian*, David D.L. Bowtell¹, Helen E. Abud¹, Gary R. Hime¹, Deon J. Venter¹, Paul K. Keese², Emma L. Duncan³, Roger R. Reddel³ and Richard A. Jefferson

CAMBIA, GPO Box 3200, Canberra, ACT 2601, Australia, ¹Trescowthick Research Laboratories, Peter MacCallum Cancer Research Institute, Locked Bag 1, A'Beckett St, Melbourne, VIC 3000, Australia, ²Institute of Molecular Agrobiolgy, Singapore and ³Children's Medical Research Institute, 214 Hawkesbury Rd, Westmead, Sydney, NSW 2145, Australia

Received August 25, 1997; Revised and Accepted September 7, 1997

Telomerase is a multicomponent reverse transcriptase enzyme that adds DNA repeats to the ends of chromosomes using its RNA component as a template for synthesis. Telomerase activity is detected in the germline as well as the majority of tumors and immortal cell lines, and at low levels in several types of normal cells. We have cloned a human gene homologous to a protein from *Saccharomyces cerevisiae* and *Euplotes aediculatus* that has reverse transcriptase motifs and is thought to be the catalytic subunit of telomerase in those species. This gene is present in the human genome as a single copy sequence with a dominant transcript of ~4 kb in a human colon cancer cell line, LIM1215. The cDNA sequence was determined using clones from a LIM1215 cDNA library and by RT-PCR, cRACE and 3'RACE on mRNA from the same source. We show that the gene is expressed in several normal tissues, telomerase-positive post-crisis (immortal) cell lines and various tumors but is not expressed in the majority of normal tissues analyzed, pre-crisis (non-immortal) cells and telomerase-negative immortal (ALT) cell lines. Multiple products were identified by RT-PCR using primers within the reverse transcriptase domain. Sequencing of these products suggests that they arise by alternative splicing. Strikingly, various tumors, cell lines and even normal tissues (colonic crypt and testis) showed considerable differences in the splicing patterns. Alternative splicing of the telomerase catalytic subunit transcript may be important for the regulation of telomerase activity and may give rise to proteins with different biochemical functions.

INTRODUCTION

Telomerase is a multi-subunit reverse transcriptase enzyme that adds DNA to telomeres, using its RNA component as a template for synthesis of the DNA repeat sequences (1). This activity of telomerase is capable of compensating for the loss of terminal sequence that occurs during the normal replication of linear DNA molecules. The loss is due in part to the 'end-replication problem', that is the inability to fill the terminal gap remaining after degradation of the most distal RNA primer during replication of the lagging strand of DNA. In mammalian cells loss of telomeric DNA also occurs due to a putative 5'-3' exonuclease activity that produces a long G-rich overhang at both telomeres (2). In the absence of telomerase, the combined effect of these processes is progressive loss of telomeric sequence with each round of cellular replication. It has been proposed that in normal

somatic cells this telomeric attrition acts as a cell division counting mechanism that eventually dictates that the cell must enter the state of senescence, in which cellular replication permanently ceases (3–5).

As it is necessary for telomere length to be preserved in the germ line of multicellular organisms, it is not surprising these cells possess telomerase activity. It has recently been shown that various somatic cells also possess telomerase activity, although at low levels. These include lymphocytes, endothelial cells, hair follicle cells, cells of the colonic crypt, and cells in the basal layer of the epidermis (6–11). The low levels of telomerase activity present in such cells appear to be insufficient to maintain telomere length over many cycles of cell division (6). In contrast, cells that have become immortalized usually exhibit telomerase at levels sufficient to maintain the lengths of their telomeres (12). In cell

*To whom correspondence should be addressed. Tel: +61 2 6246 5310; Fax: +61 2 6246 5303; Email: zej@cambia.org.au

culture models of immortalization, cells that escape from senescence due to the activity, for example, of DNA tumor virus oncoproteins usually arrest in crisis. A small minority of the cells may escape from crisis and become immortalized [reviewed in (13)] and this is usually associated with expression of substantial levels of telomerase activity (12,14–17). The great majority of cancer cell lines and cancer biopsies also contain telomerase activity (18–21), making telomerase an attractive target for the development of anti-cancer therapeutics (22).

Not all immortalized cell lines contain detectable telomerase activity, however (16,20,23). It has been shown that telomere lengthening occurs in telomerase-negative cell lines when they escape from crisis and become immortalized *in vitro*, and that they are characterized by telomeres of very heterogeneous length ranging from short to extremely long (16). These cells therefore exhibit an alternative mechanism for lengthening of telomeres (ALT) (16,24). Although it cannot be assumed that the mechanism of ALT will be the same in every telomerase-negative cell line, a possible mechanism involves recombination, similar to that seen in mutant yeast in the absence of telomerase (25). ALT also occurs in tumor-derived cell lines and tumors (26).

All immortalized human cell lines examined to date contain either telomerase activity or evidence of ALT (24). This represents very strong circumstantial evidence for the importance of telomere maintenance in immortalization. Elucidation of the molecular details of telomere maintenance will greatly advance our understanding of immortalization, cancer, and related processes. Considerable progress has recently been made in characterizing mammalian telomerase. Genes for the RNA component (27,28) and for a protein homologous to the p80 component of *Tetrahymena* telomerase (29,30) have been cloned.

We report here the cloning of a human gene homologous to a protein found in *Saccharomyces cerevisiae* and *Euplotes aediculatus* that has reverse transcriptase motifs and is thought to be the catalytic subunit of telomerase in those species (31). The gene, which we designate hTCS1 (human Telomerase Catalytic Subunit 1), is present in single copy in the human genome but is expressed in a complex splicing pattern that gives rise to a number of potential proteins. While this manuscript was in preparation, Nakamura *et al.* reported the cloning of the same gene, which they referred to as hTRT (32). We find that expression of hTRT/hTCS1 positively correlates with the known telomerase status of tissues and cell lines and occurs preferentially in *in vitro* cell lines after they had undergone crisis, in normal tissues with a significant stem cell component and in a range of tumors. These findings define an important component of the human telomerase complex and implicate its expression as a major determinant of the distribution of telomerase activity in mammalian cells.

RESULTS

Identification of a human homolog of yeast and *Euplotes* telomerase catalytic subunit genes

An homology search (BLAST) of sequence databases using the *E.aediculatus* telomerase catalytic subunit gene as the query sequence identified the GenBank AA281296 Expressed Sequence Tag (EST) as the top ranked match ($P = 3.2 \times 10^{-6}$). Polymerase chain reaction (PCR) primers HT1553F and HT1920R, based on the EST sequence (Table 1), were used to search for the human homolog of *E.aediculatus* telomerase. We amplified a fragment of the expected size (~350 bp) from a

colonic carcinoma cell line LIM1215 cDNA library. We cloned the PCR product into pBluescript II KS+ vector (plasmid pAKE54.8) and confirmed its identity with the EST sequence (data not shown).

Table 1. Oligonucleotide primers

Primer	Sequence
CHO17	5'-CCCTGAGGCACTCTCCAG-3'
CHO18	5'-ACTTGCCTCAGGAGGAGC-3'
EBHT18	5'-CACGAATTCGGATCCAAGCTTTTTTTTTTTTTTTTTT-3'
HT0093R	5'-GGCACGCACACCAGGCACTG-3'
HT0141F	5'-CCTGCCTGAAGGAGCTGGTG-3'
HT0142R	5'-GGACACCTGGCGGAAGGAG-3'
HT0163F	5'-CCGAGTGCTGCAGAGGCTGT-3'
HT0220R	5'-GAAGCCGAAGGCCAGCACGTTCTT-3'
HT1114R	5'-GTTCCAAGCAGCTCCAGAAACAG-3'
HT1157R	5'-GGCAGTGCCTTGTAGGAGCA-3'
HT1262F	5'-GTGCAGCTGCTCCGCCAGCACA-3'
HT1553F	5'-CACTGGCTGATGAGTGTGTAC-3'
HT1576R	5'-GACGTACACACTCATCAGCCAG-3'
HT1590F	5'-GGTCTTCTTTATGTACCGGAG-3'
HT1691F	5'-CACTTGAAGAGGGTGCAGCT-3'
HT1875F	5'-GTCTCACCTCGAGGGTGAAG-3'
HT1893R	5'-TTCACCCTCGAGGTGAGACGCT-3'
HT1920R	5'-TCGTAGTTGAGCACGCTGAAC-3'
HT2026F	5'-GCCTGAGCTGTACTTTGTCAA-3'
HTM2028F	5'-CTGAGCTGTACTTTGTCAAAGGACA-3'
HT2356R	5'-CATGAAGCGTAGGAAGACGTCGAAGA-3'
HT2482R	5'-CGCAAACAGCTGTCTCCATGTC-3'
HT2761F	5'-CTATGCCCGACCTCCATCAGA-3'
HT2781R	5'-CTGATGGAGGTCCGGGCATAG-3'
HT3114F	5'-CCTCCGAGGCCGTGCAGT-3'
Vector1	5'-CGCCAGGGTTTTCCAGTCACGA-3'
Vector2	5'-GTAATACGACTCACTATAGGGCGA-3'

To determine whether related genes exist within the human genome, Southern blot analysis was performed using normal male human DNA. The DNA was restricted separately with several different enzymes, separated on an agarose gel and transferred to a nylon membrane. The membrane was hybridized with ³²P-labeled insert from pAKE54.8 and washed under conditions of high stringency. Autoradiography of the blot revealed the presence of a single hybridizing band in most lanes (Fig. 1A). This finding indicated that hTCS1 probably exists as a single copy gene within the human genome. We cannot, however, exclude the possibility of distantly related sequences that fail to hybridize with this region of the gene. We also hybridized DNA obtained from LIM1215 cells to determine whether there were any changes in the hTCS1 gene in this cell line that could account for activation of telomerase expression. There was no evidence of gross rearrangement or amplification of hTCS1 in these cells, although this analysis does not rule out subtle mutations within the regulatory region of the gene.

We performed Northern blot analysis of RNA obtained from LIM1215 and primary fibroblast cells to provide an indication of the size and complexity of hTCS1 transcripts and to determine whether expression was regulated differently in tumor versus

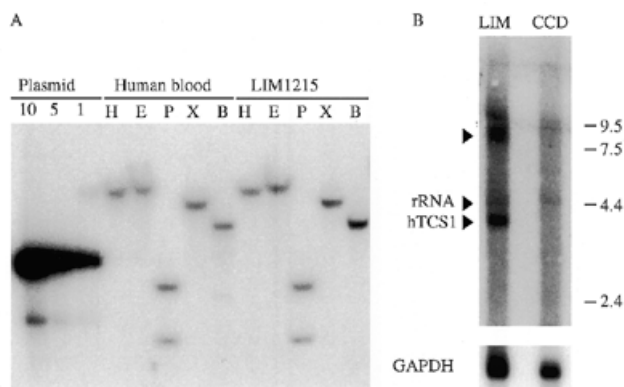


Figure 1. hTCS1 Southern and Northern analyses. (A) hTCS1 is encoded by a single gene and is not amplified in LIM1215 cells. Southern blot of genomic DNA isolated from human blood and the LIM1215 cell line probed with radiolabeled insert from plasmid pAKE54.8 (a PCR-generated hTCS1 gene fragment). The blot also contained dilutions of plasmid pAKE54.8 to control for the sensitivity of detection. The plasmid was diluted to ~10, 5 and 1 genome equivalents. H: *HindIII*; E: *EcoRI*; P: *PstI*; X: *XbaI*; B: *BamHI*. (B) hTCS1 is expressed in LIM1215 colon carcinoma cells but not in CCD primary fibroblasts. Northern blot of polyA⁺ mRNA probed as for Southern analysis, and exposed to X-ray film for 72 h. An mRNA of 3.9 kb hybridized to the hTCS1 probe. Additional cross-hybridizing mRNA of higher molecular weight is indicated by the top arrowhead. Cross-hybridization to ribosomal RNA that had contaminated the polyA⁺ RNA preparation is indicated. The blot was also hybridized to a probe from the GAPDH gene as a loading control, and exposed to film for 1 h (lower panel). Marker sizes are indicated in kb.

normal cells. Approximately 3 μ g of polyadenylated RNA from each cell type was separated on a formaldehyde agarose gel, transferred to nitrocellulose and probed with radiolabeled pAKE54.8 insert. After washing the blot at high stringency a prominent 3.9 kb band was apparent in LIM1215 RNA (Fig. 1B, upper panel). Importantly, this band was not apparent in the lane containing primary fibroblast RNA, even though this lane contained a similar amount of high quality RNA as assessed by hybridization with a glyceraldehyde 6-phosphate dehydrogenase (GAPDH) control probe (Fig. 1B, lower panel) and ethidium staining of the gel (data not shown). The presence of larger transcripts (especially an ~8 kb heterodispersed band) was also visible only in LIM1215 RNA (Fig. 1B, upper panel). These findings provided an indication of additional hTCS1-specific mRNAs and also that hTCS1 may be preferentially expressed in tumor versus primary cells.

To obtain longer clones of hTCS1 we initially surveyed a number of cDNA libraries prepared from tumor cells, by performing PCR with primers from within the EST sequence (primers HT1553F and HT1920R, followed by nested primers HT1590F and HT1893R). An amplified product of the expected size (~350 bp) was detected in only three of the 12 libraries screened (data not shown). The most abundant product was obtained from a library prepared from LIM1215 cell mRNA. In this library, and in several others, an additional fragment of ~170 bp was amplified (data not shown). We took the parallel approaches to obtaining longer clones from the LIM1215 library of screening plaques with a ³²P-labeled pAKE54.8 insert probe and also performing PCR analysis (see Materials and Methods). A single positive plaque, designated 53.2, with a 1.9 kb insert was obtained by hybridization of the library with the pAKE54.8 insert

probe. DNA sequence analysis of this clone demonstrated that it extended both 5' and 3' of the EST sequence, but did not contain a single open reading frame (ORF). A fragment obtained from PCR amplification of the library was similar in sequence to the 53.2 fragment but also contained two additional sequences of 36 bp and >300 bp. Both insertions demonstrated characteristics of splice acceptor and donor sequences at their boundaries relative to the 53.2 sequence and may represent unspliced introns. As none of these fragments defined the expected ORF when their sequences were compared with the *Euplotes* catalytic subunit we then performed PCR analysis on mRNA samples. Reverse transcriptase PCR (RT-PCR) on LIM1215 mRNA identified a number of additional PCR products including one in which a 182 bp insertion relative to the 53.2 fragment resulted in a single ORF. The sequence of this fragment was extended 5' and 3' using a combination of cRACE and 3' RACE, respectively, on LIM1215 mRNA to give a fragment of 3871 nt that we designated hTCS1. The size of this fragment accorded well with the size estimated from the Northern blot (Fig. 1B) for the most abundant RNA species in LIM1215 RNA. Interestingly, the first AUG codon in frame with the longest reading frame was >1 kb (1031 bp) from the transcription start. The sequence around this Met codon did not conform to the Kozak consensus sequence (33) and therefore is unlikely to be the translation start. There are several CUG codons at the 5' end of the mRNA matching the Kozak consensus that could support translation initiation. The sequence reported by Nakamura *et al.* (32) contains an additional 138 5' terminal nucleotides that are consistent with a Met translation start; the sequences are otherwise identical. It is not clear whether this difference reflects transcription from alternative promoters, a feature that is often associated with alternative splicing (34–36).

hTCS1 sequence and alignment with other telomerases

Multiple sequence alignment demonstrated that the predicted hTCS1 protein remained co-linear with the *Euplotes* and *Saccharomyces cerevisiae* telomerase catalytic subunits over their lengths. There are a number of CUG codons at the 5' end of the mRNA matching the Kozak consensus sequence that could support translation initiation (37). Multiple sequence alignment seems to support a CUG translation start since three of the CUG codons matching the Kozak consensus align close to the translation start for *Euplotes* (Leu 22) and *S.cerevisiae* (Leu 49 and Leu 52). Although the overall homology between the three proteins was relatively low (~40% similarity in all pairwise combinations) the overall structure of the protein seems to be well conserved. Four major domains: N-terminal, basic, reverse transcriptase (RT) and C-terminal (31) are present in all three proteins. The highest area of sequence similarity was within the RT domain. Notably, all the motifs characteristic of the *Euplotes* RT domain were present and all amino acid residues implicated in RT catalysis were conserved in the hTCS1 sequence (31).

Recently, Li *et al.* (38) demonstrated that protein phosphatase 2A treatment of human breast cancer cell extracts inhibited telomerase activity. It is not known whether this effect is direct, but it raises the possibility of regulation of telomerase activity by protein phosphorylation. Although requiring experimental verification, the predicted hTCS1 protein does contain numerous potential phosphorylation sites, including 11 SP or TP dipeptides, which are potential sites for cell cycle dependent kinases.

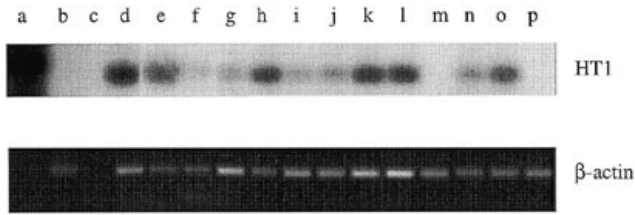


Figure 2. hTCS1 is differentially expressed in normal and tumor tissues. RT-PCR of total RNA from normal and tumor tissues is shown. Top panel (labeled HT1): PCR was performed with primers (HT1553F and HT1893R) from the hTCS1 cDNA sequence that span an intron in the hTCS1 gene, and the products were blotted and probed with a radiolabeled oligonucleotide (HT1691F) from the hTCS1 sequence. Lower panel: PCR was also performed on the same samples with a pair of primers from the β -actin gene as a loading control. Lanes: a, hTCS1 cDNA control; b, human genomic DNA control; c, no template control; d-p RNA from: d, normal colon; e, normal testis; f, normal lymphocyte; g, melanoma (cerebral metastasis); h, melanoma (subcutaneous metastasis); i, melanoma (liver metastasis); j, melanoma (lung metastasis); k, melanoma (axillary lymph node metastasis); l, melanoma (skin metastasis); m, breast carcinoma; n, breast carcinoma; o, breast carcinoma; p, breast carcinoma.

hTCS1 expression patterns

Although telomerase activity has been widely associated with tumor cells and the germline, it has recently been recognised that certain normal mammalian tissues express low levels of telomerase activity. Although we had not detected hTCS1 expression in primary fibroblast RNA we were interested to determine whether it could be detected in other normal tissues. We screened by PCR several commercially available cDNA libraries from lung, heart, liver, pancreas, hippocampus, fetal brain, and testis for the presence of hTCS1 sequences using nested primers (HT1553F and HT1920R, followed by HT1590F and HT1893R) for the EST region but none were found (data not presented). We also examined the expression of hTCS1 in normal tissues that have previously been shown to have telomerase activity (colon, testis and peripheral blood lymphocytes). To extend our analysis of tumor material, we also performed RT-PCR with PCR primers HT1553F and HT1893R on RNA from a number of melanoma and breast cancer samples. RNA was isolated from normal human colon, testis and circulating lymphocytes, and from tissue sections of tumor samples, and subjected to RT-PCR analysis. PCR products from cDNA could easily be distinguished from those due to contaminating genomic DNA, as we had determined that products of ~300 bp and 2.7 kb were obtained with this primer set from RNA and DNA, respectively (data not shown). hTCS1 transcripts were detected in both colon and testis, in the majority of tumor samples, and very weakly in the lymphocyte RNA (Fig. 2, upper panel). Interestingly, two of the breast cancer samples were negative for hTCS1 expression, despite containing comparable amounts of RNA to the other samples, as judged by PCR amplification of β -actin as a positive control (Fig. 2, lower panel).

Acquisition of telomerase activity appears to be an important aspect of the immortalization process. We analyzed the expression of hTCS1 in a number of matched pairs of pre- and post-crisis cell cultures using RT-PCR followed by amplification with nested primers (Fig. 3, upper panel). These cultures were

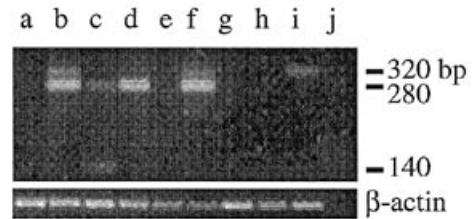


Figure 3. hTCS1 expression in pre-crisis cells and post-crisis cell lines. Upper panel: Nested amplification using nested primers (HT1553F and HT1920R, followed by HT1590F and HT1893R). Lower panel: Control RT-PCR using β -actin primers. Lanes: a, BET-3K passage (p) 7 (pre-crisis); b, BET-3K p32 (post-crisis); c, BFT-3K p14 (pre-crisis); d, BFT-3K p 22 (post-crisis); e, BFT-3B p15 (pre-crisis); f, BFT-3B p29 (post-crisis); g, GM847 (ALT); h, IICF/c (ALT); i, IICF-T/B1 (ALT); j, No template control

found to be telomerase negative and positive, respectively, using the TRAP assay (16). In two matched pairs (BFT-3B and BET-3K) hTCS1 was detected only in the post-crisis cell lines. Whilst the post-crisis line in the BFT-3K set showed an abundant hTCS1 band, a fragment of the same size was also weakly present in the pre-crisis culture sample. Interestingly, two of the three post-crisis cell lines demonstrated the presence of an additional unexpected fragment of 320 bp and this product was also found when colon and testis mRNA was analyzed on high resolution gels (data not shown).

Three immortalized telomerase-negative (ALT) cell lines were also analyzed for hTCS1 expression. We were unable to detect expression in two of the lines, but in one line (IICF-T/B1) a product of ~320 bp was again amplified (Fig. 3), as in the post-crisis lines, and the colon and testis samples. Direct sequencing of the 320 bp PCR product from the line IICF-T/B1 (ALT) revealed the presence of a 38 bp insertion (that we refer to as Insertion 1), relative to the expected product. The possibility that this is an amplification product from genomic DNA rather than mRNA was ruled out by performing PCR with the same primers but using genomic DNA as the template. Under these conditions a 2.7 kb fragment was amplified and its authenticity was confirmed by partial sequencing (data not presented).

hTCS1 shows multiple RNA variants suggesting alternative splicing

Sequencing of the clones from the LIM1215 cDNA library, and the RT-PCR data presented above for the pre-crisis and post-crisis cultures, identified a number of different sequence variants of the hTCS1 transcript. To systematically survey for variants we performed RT-PCR with a number of primer pairs covering the whole sequence. We did not find any variants in the N-terminal and the basic domains (data not presented), but the RT domain (and, to a lesser extent, the C-terminal domain) contained several variants. Most notably, there were several RNA variants between RT Motif A and RT Motif B (Fig. 4A). RT-PCR using primers that span Motifs A and B (HT1875F and HT2781R, followed by amplification with nested primers HT2026F and HT2482R), on RNA samples from numerous tumors showed the presence of four different PCR products: 220 bp (band 1), 250 bp (band 2), 400 bp (band 3) and 430 bp (band 4). Strikingly, we noticed considerable variation among the tumor samples tested both in

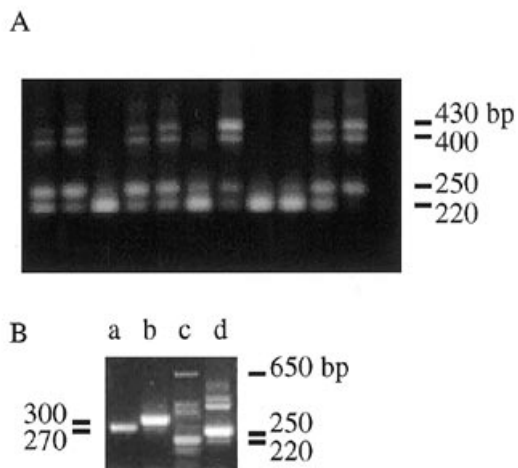


Figure 4. hTCS1 splicing patterns. (A) hTCS1 shows various splicing patterns in different cells and tumor samples. Nested amplification (14 cycles) using HT2026F and HT2482R primers on the primary RT-PCR products generated with HT1875F and HT2781R primers. Lanes: a, lung carcinoma; b, lymphoma; c, lung carcinoma; d, medulloblastoma; e, lymphoma; f, lymphoma; g, T47D breast carcinoma cell line; h, pheochromocytoma; i, lymphoma; j, glioma; k, lymphoma; l, no template control. (B) RT-PCR evidence that some hTCS1 transcripts have a 36 bp deletion (α). RT-PCR was carried out on LIM1215 RNA with the following primer combinations: a, HTM2028F + HT2356R; b, HT2026F + HT2356R; c, HTM2028F + HT2482R; d, HT2026F + HT2482R. Primer HTM2028F was designed to generate a PCR product only if the mRNA lacked the 36 bp sequence (see Results).

the total number of PCR products and in the quantitative distribution among the products. We sequenced three of these products from a number of tumor tissues and showed that one of them, a 220 bp fragment, is equivalent to the 53.2 cDNA from the LIM1215 library. The fragment of ~250 bp (band 2) was found to contain a 36 bp in-frame insertion, the same insertion that was identified in a PCR product from LIM1215 cDNA library. As the RT-PCR product had the same sequence as the PCR product from the cDNA library, it is apparent that the 36 bp of additional sequence was not a library construction artifact. The largest product (band 4) contained an additional block of 182 bp (the same as the larger product amplified earlier from LIM1215 RNA) compared with the 250 bp amplicon. We could not obtain an unambiguous sequence for the 400 bp band (band 3). Based on its size, we hypothesized that it may contain the 182 bp sequence but be missing the 36 bp sequence present in bands 2 and 4 but absent from band 1. To test the hypothesis that such a transcript exists, we designed the primer (HTM2028F) that allowed amplification only when the 36 bp fragment was missing. Amplification using HTM2028F and HT2026F primers in combination with HT2356R demonstrated that transcripts containing the 182 bp fragment but missing the 36 bp fragment are present in LIM1215 RNA (Fig. 4B, lanes a and b). The same top strand primers (HTM2028F and HT2026F) in combination with HT2482R primer amplified a number of products from LIM1215 RNA (Fig. 4B, lanes c and d), most of which corresponded to bands 1–4 as determined by direct sequencing of PCR products (data not presented). A band of 650 bp amplified using HTM2028F and HT2482R represents another, not yet fully

characterized, alternatively spliced hTCS1 variant in the RT-MotifA/RT Motif B region.

Thus we detected a number of hTCS1 RNA variants in normal tissues, immortalized cell lines and tumor samples (Fig. 5). For clarity of presentation we used the protein sequence giving the best match with the *Euplotes* and *S.cerevisiae* proteins as the reference sequence. This reference sequence commences 28 amino acids after the first methionine encoded by the sequence deposited in GenBank by Nakamura *et al.* (32), and is otherwise identical. Transcripts that are missing sequences that are present in the reference sequence are described as having deletions of α or β (Fig. 5A). For example, band 1 in Figure 4A contains deletions α and β , band 2 contains deletion β , band 3 contains deletion α , and band 4 has no deletions. In addition to the 38 bp Insertion 1 referred to above, similar analyses have defined two other blocks of sequence that are not present in the reference sequence; these are referred to as Insertions 2 and 3 (Fig. 5). Interestingly, all of the variations but one (Insertion 3) are within the RT domain and none was detected within the N-terminal or basic domains.

DISCUSSION

We have identified a candidate human telomerase catalytic subunit gene. The gene we called hTCS1 is represented in the genome as a single copy sequence. This simple genomic context contrasts with a complex expression pattern. We identified a number of RNA variants suggesting that this single genetic locus may encode a number of proteins, quite possibly with different biochemical properties. The transcript that best matches putative telomerase catalytic subunits from lower eukaryotes is nearly 4 kb long. It displays relatively low overall sequence homology with the *S.cerevisiae* and *Euplotes* proteins yet critical motifs within the RT domain are conserved and there are extensive areas of homology outside this domain. During the preparation of this manuscript Nakamura *et al.* (32) also reported the cloning of the same gene and share the conclusion that it represents a human telomerase catalytic subunit.

Consistent with this conclusion we find good correspondence between the expression of hTCS1 and telomerase activity. This was shown by RT-PCR amplification of hTCS1 transcripts from normal tissues, tumors and immortalized cell lines that are telomerase-positive and general lack of this RT-PCR product in normal fibroblasts and pre-crisis cells and also in telomerase-negative immortalized (ALT) cells. The expression of hTCS1 in testis, colonic epithelium, and lymphocytes correlates with previous observations that these normal cells have telomerase activity. Surveys of large numbers of tumors have shown that ~85% of all tumors contain detectable telomerase activity (21), and similarly we found that most, but not all, of the tumors we surveyed had hTCS1 expression. Two of the pairs of matched pre-crisis (mortal) and post-crisis (telomerase-positive) fibroblast and epithelial cell cultures showed a clear change in hTCS1 expression from undetectable to detectable, in accord with their telomerase status (16). In the single pre-crisis culture in which hTCS1 expression was detectable, it is possible that the subclone of immortalized cells that eventually escaped from crisis had already arisen at the stage when the hTCS1 analysis was performed. In another apparent exception to the correlation between TRAP activity and hTCS1 expression, one of the ALT

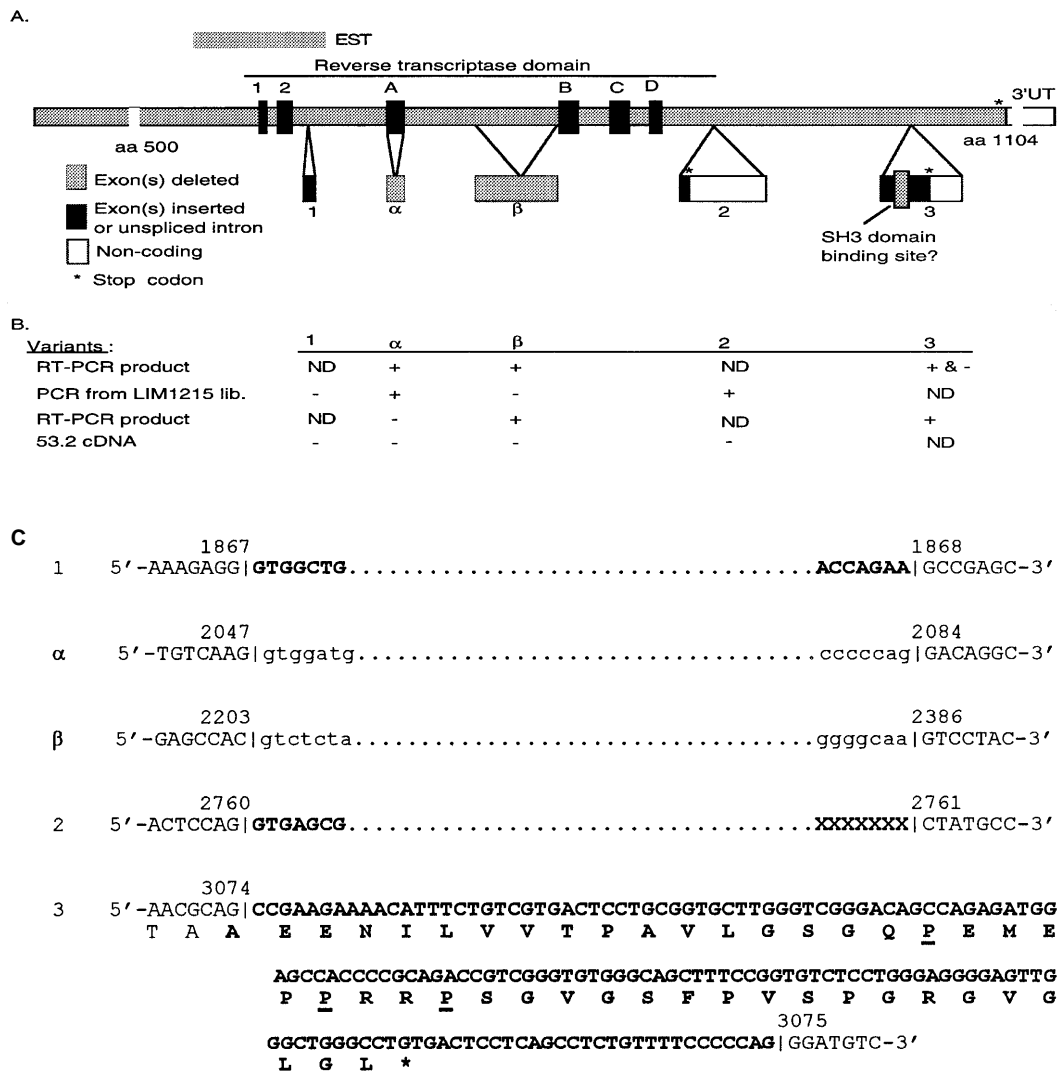


Figure 5. Alternative splicing of the hTCS1 transcript may lead to many isoforms of the protein. (A) Schematic representation of the splicing variants sequenced. Amino acid 500 of the reference sequence corresponds to 528 in the sequence of Nakamura *et al.* (32). Domains 1 and 2 are telomerase-specific and domains A–D are common to reverse transcriptases (31). The positions of Deletions α and β and Insertions 1, 2, and 3 are shown. (B) Combinations of insertions and deletions in RNA variants identified. '+' and '-' indicate that the sequence is present (+) or absent (-). (C) Sequences of putative exon/intron junctions of RNA variants. Variants are designated as in part A, and putative exon/intron junctions are marked with |. Putative spliced exons are in lower case and putative unspliced introns are in bold type. Sequence coordinates: nucleotide 1 corresponds to nucleotide 139 of the sequence in GenBank Accession number AF015950 (32). A complete DNA sequence (with protein translation) of Insertion 3 is presented. Amino acids corresponding to the putative c-Abl/SH3 binding site are underlined.

cells yielded a variant hTCS1 PCR, which raises the subject of the role of the variant mRNAs.

A prominent feature of our analysis was the detection of a number of variant hTCS1 transcripts. A full understanding of the significance of these variants awaits characterization of the functional properties of all the protein derivatives and direct assessment of proteins produced in various cell types. Although some of the variants may reflect incompletely processed mRNA, it is noteworthy that the variants were abundant in an RNA sample (LIM1215) preselected for polyadenylated mRNA. These findings, together with their clustering in the RT domain, suggest that the insertion variants more likely reflect regulation of hTCS1 protein expression. We can, however, be more confident that variants in which exons are deleted (see α and β in Fig. 5) reflect alternative mature mRNAs coding for variant proteins.

Additional evidence in support of this assertion comes from the sequencing of both cDNA clones identified in the LIM1215 cDNA library (one by plaque screening, and the other by PCR amplification), as these clones contained both deletions and insertions compared with the reference sequence. Again, it is noteworthy that these variants are located within the RT domain. Deleted α -exon results in a small in-frame 12 amino acid deletion, encompassing a conserved sequence motif (motif A) that is critical for RT function. Intriguingly, a single amino acid mutation within this domain in the yeast EST2 protein resulted in a protein that acted as a dominant negative and resulted in cellular senescence and telomere shortening (32).

Some of the variant sequences, including the β -exon deletion, encode truncated proteins. Insertion 2 is particularly interesting as it truncates the protein precisely at the end of the RT domain so that

it lacks a C-terminal domain. We also identified a variant with an alternative C-terminal domain (Insertion 3, Fig. 5). This alternative C-terminus includes the peptide SGQPEMEPPRRPSPGCVG which matches the consensus c-Abl SH3 binding peptide (PXXXXPXXP), found in proteins such as ataxia telangiectasia mutated (ATM) (39). A second example of this motif is found within the N-terminal end of the hTCS1 protein in the peptide HAGPPSTSRPPRPWDTP. Interaction of SH3 domain containing proteins with hTCS1 warrants further scrutiny.

Another sequence motif (AVRIRGKS) identified in hTCS1 matches a P-loop motif consensus AXXXXGK(S) (40). This motif is found in a large number of protein families including a number of kinases, bacterial dnaA, recA, recF, mutS and ATP-binding helicases (41). While it is difficult to speculate at this point about the functional importance of this motif for hTCS1 function, it is interesting that the P-loop is present within the 182 bp fragment (spliced exon β) that is present only in a subpopulation of the mRNA in most RNA samples analyzed and completely absent from several tumor samples (Fig. 4A).

The variant detected in one ALT cell line (Fig. 3, lane i) opens up the possibility that the basic domain of hTCS1 may contribute to the ALT mechanism in at least some ALT cell lines. Interestingly, this ALT cell line expresses the hTR gene (42). We have previously hypothesized that one possible mechanism of ALT could involve dysregulated telomerase components that are inactive in the TRAP assay (16).

Alternative mRNA splicing is a common mechanism for regulating gene expression in higher eukaryotes and there are many examples of tissue-specific, development-specific and sex-specific alterations in splicing events (43). Although we cannot rule out the involvement of novel minor alternatively spliced hTCS1 variants in immortalization and tumorigenesis, the altered relative expression levels of the major transcripts found in various tumors compared with normal cells, and in post-crisis cell lines compared with limited life span pre-crisis cells, may be more important. In addition, the existence of the alternatively spliced variants of hTCS1 that were seen in both testis and colonic crypt, suggests complex regulation of this gene in normal development.

We have shown that, although some hTCS1 splice variants may be expressed in normal, pre-crisis, and ALT cells that lack detectable telomerase activity, expression of the major hTCS1 products was found in most tumors and in all telomerase-positive immortalized cell lines. Transcriptional control of hTCS1 may therefore be a major aspect of the regulation of telomerase activity, but it also seems likely that the function and regulation of telomerase is complex. There is good evidence that telomerase is involved in the healing of chromosome breaks (44) in addition to its role in maintaining telomere length in the germline, and it is possible that it has other functions. It is possible that the subunit composition of telomerase, and the regulatory mechanisms, may vary according to these functional roles and that, as a result, there may be several telomerase species with their own control mechanisms. It seems unlikely that transcriptional control of the hTR RNA subunit gene is a dominant controlling mechanism for telomerase activity (45). Major questions about subunit composition and regulation include whether there is a mammalian homolog of the *Tetrahymena* p95 protein, the role post-translational processing (29) of the TLP1/TP1 p80 homolog plays in activation of telomerase, the nature and role of the products encoded by the splice variants of the hTCS1 putative catalytic

subunit identified in this study, and whether post-translational changes such as phosphorylation of hTCS1 products is an important control mechanism.

MATERIALS AND METHODS

Cell lines

GM847, HICF/c and HICF-T/B1 are immortalized, telomerase-negative (ALT) fibroblast cell lines (16). BET-3K, BFT-3B and BFT-3K are telomerase-positive cell lines (16), derived from bronchial epithelial cells (BET-3K) or bronchial fibroblasts (BFT-3B and -3K) (46). The LIM1215 colon carcinoma cell line was from the Ludwig Institute for Cancer Research, Melbourne branch, and the CCD fibroblasts and the T47D breast cancer cell line were obtained from the American Type Culture Collection. The LIM1215 and T47D cell lines are both telomerase-positive (26).

Southern and Northern analyses

Southern and Northern analyses and probe preparation were carried out according to standard methods (47). For Northern analysis, $\sim 3 \mu\text{g}$ of polyA+ RNA per lane was fractionated on a 0.85% formaldehyde agarose gel and transferred to a Zetaprobe membrane. The blot was probed with ^{32}P -labeled insert from pAKE54.8 and a ^{32}P -labeled GAPDH cDNA clone as a loading control.

For Southern analysis, $\sim 10 \mu\text{g}$ of endonuclease restricted genomic DNA was fractionated per lane on a 1% agarose gel and transferred to Zetaprobe membrane. The blot was probed with the same labeled plasmid insert used for Northern analysis. One hundred, 50 and 10 μg of this plasmid was also electrophoresed on the gel as a sensitivity control to ~ 10 , 5 and 1 single gene copy equivalents. Peripheral blood DNA was obtained from a normal human male volunteer.

RT-PCR analysis

RNA from normal and tumor tissues was kindly provided by Drs G. Somers, D. Germain and A. Hutchins. RT-PCR followed by nested PCR was performed using the Titan RT-PCR system (Boehringer Mannheim) according to the manufacturer's recommendation. PCR cycling after the RT step was with primers HT1553F and HT1920R as follows: 95°C for 2 min, two cycles of 94°C for 30 s, 65°C for 30 s, and 68°C for 3 min; two cycles of 94°C for 30 s, 63°C for 30 s, and 68°C for 3 min; and 34 cycles of 94°C for 30 s, 60°C for 30 s, and 68°C for 3 min. RT-PCR products were diluted 100-fold and 1 μl was used for nested PCR using *Taq* polymerase with buffer Q (Qiagen) and primers HT1590F and HT1893R. PCR conditions were as in RT-PCR, but the final cycling was for 14 cycles only. For normal tissues and tumors, RT-PCR products were resolved by electrophoresis in a 1.5% agarose gel, transferred to Zetaprobe membrane and probed with radiolabeled oligonucleotide HT1691F (Table 1). As a control, a 213 bp fragment of β -actin was PCR-amplified from reverse transcribed RNA using primers CHO17 and CHO18 under the same conditions but without a nested PCR step.

Genomic PCR

The same genomic DNA samples that were subjected to Southern analysis were analyzed by PCR. Genomic DNA (50–100 ng) per

reaction was amplified with primers HT1553F and HT1893R using the XL-PCR system (Perkin-Elmer) and the PCR conditions recommended by the manufacturer.

cDNA library screening

1.2 million plaques of a cDNA library constructed from LIM1215 cell mRNA in λ ZAPII were screened by hybridization to radiolabeled insert from pAKE54.8. After re-screening, a single positive clone was purified and designated 53.2.

PCR screening of cDNA libraries

In all PCR amplifications 2×10^6 p.f.u. was used as a template. A 350 bp fragment corresponding to the EST was amplified using HT1553F and HT1920R primers. The PCR program was as follows: 95°C for 4 min followed by 30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min. This was followed by PCR using nested primers HT1590F and HT1893R under the same conditions. To extend the hTCS1 sequence we performed PCR on the LIM1215 library using the XL-PCR system in two rounds of amplification. In the first round we used HT1553F and Vector1 primer. After 30 cycles of amplification PCR products were column purified and 1% used as a template in 30 cycle PCR using HT1590F and Vector 2 primer combination. All PCR was done with 'hot start' by adding polymerase at 80°C (after the initial denaturation step).

cRACE

Two rounds of cRACE were carried out to extend the sequence of hTCS1 and map the transcription initiation site, essentially as described (48). LIM1215 polyA+ RNA (500 ng) was used as the template. The first strand cDNA was primed using the HT1576R primer. The first round of PCR on the ligation product (using the XL-PCR system) employed the HT1157R and HT1262F primers. PCR products were purified using Qiagen columns and amplified using primers HT1114R and HT1553F. A 1.4 kb band generated from this amplification was sequenced and a new set of primers was designed based on the new sequence. For the second round of cRACE the first strand cDNA was primed with the HT0220R primer. The first round of PCR on the ligation product utilized the HT0142R and HT0141F primers. PCR products were purified as before and amplified using HT0093R and HT0163F primers. A product of 100 bp was amplified and sequenced in two independent experiments defining the 5' end of the hTCS1 transcript.

3' RACE

The most 3' sequences of the transcript were obtained by two rounds of PCR (XL-PCR system) using EBHT18 in both rounds as the reverse primer, and HT2761F and HT3114F as the forward primers in the first and second rounds, respectively (Table 1).

Sequencing and sequence analysis

All sequencing reactions were done using the Perkin-Elmer dye-terminator cycle sequencing kit. Reactions were analyzed on an ABI373A sequencer. Most of the sequence analysis was done using the GCG sequence analysis package (41), and the sequences were aligned with ClustalW (49).

ACKNOWLEDGEMENTS

We thank Doug Hilton, Anne Marie Hutchins, Doris Germain, Gino Somers, Donna Dorow, Bob Whitehead, and Kate Loveland for materials; and Gosia Aschenbrenner-Kilian, Colin House, Ross Dickens, Hannah Robertson, Ralf Schnall, Dennis Wang, Tanya Vaughan, and Nadia Traficante for intellectual and technical contributions. This research was supported by Cambia Biosystems LLC, a Wellcome Trust Senior Research Fellowship and a Howard Hughes Medical Institute International Research Scholarship (to D.D.L.B.), and the Carcinogenesis Fellowship of the New South Wales Cancer Council (to R.R.R.).

REFERENCES

- Greider, C.W. and Blackburn, E.H. (1985) Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell*, **43**, 405–413.
- Makarov, V.L., Hirose, Y. and Langmore, J.P. (1997) Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell*, **88**, 657–666.
- Olovnikov, A.M. (1971) Principle of marginotomy in template synthesis of polynucleotides. *Doklady Akademii Nauk SSR*, **201**, 1496–1499.
- Olovnikov, A.M. (1973) A theory of marginotomy. *J. Theor. Biol.*, **41**, 181–190.
- Harley, C.B., Vaziri, H., Counter, C.M. and Allsopp, R.C. (1992) The telomere hypothesis of cellular aging. *Exp. Gerontol.*, **27**, 375–382.
- Broccoli, D., Young, J.W. and de Lange, T. (1995) Telomerase activity in normal and malignant hematopoietic cells. *Proc. Natl. Acad. Sci. USA*, **92**, 9082–9086.
- Counter, C.M., Gupta, J., Harley, C.B., Leber, B. and Bacchetti, S. (1995) Telomerase activity in normal leukocytes and in hematologic malignancies. *Blood*, **85**, 2315–2320.
- Hiyama, K., Hirai, Y., Kyoizumi, S., Akiyama, M., Hiyama, E., Piatyszek, M.A., Shay, J.W., Ishioka, S. and Yamakido, M. (1995) Activation of telomerase in human lymphocytes and hematopoietic progenitor cells. *J. Immunol.*, **155**, 3711–3715.
- Härle-Bachor, C. and Boukamp, P. (1996) Telomerase activity in the regenerative basal layer of the epidermis in human skin and in immortal and carcinoma-derived skin keratinocytes. *Proc. Natl. Acad. Sci. USA*, **93**, 6476–6481.
- Yasumoto, S., Kunimura, C., Kikuchi, K., Tahara, H., Ohji, H., Yamamoto, H., Ide, T. and Utakoji, T. (1996) Telomerase activity in normal human epithelial cells. *Oncogene*, **13**, 433–439.
- Hsiao, R., Sharma, H.W., Ramakrishnan, S., Keith, E. and Narayanan, R. (1997) Telomerase activity in normal human endothelial cells. *Anticancer Res.*, **17**, 827–832.
- Counter, C.M., Avilion, A.A., LeFeuvre, C.E., Stewart, N.G., Greider, C.W., Harley, C.B. and Bacchetti, S. (1992) Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J.*, **11**, 1921–1929.
- Bryan, T.M. and Reddel, R.R. (1994) SV40-induced immortalization of human cells. *Crit. Rev. Oncogenesis*, **5**, 331–357.
- Klingelutz, A.J., Barber, S.A., Smith, P.P., Dyer, K. and McDougall, J.K. (1994) Restoration of telomeres in human papillomavirus-immortalized human anogenital epithelial cells. *Mol. Cell. Biol.*, **14**, 961–969.
- Counter, C.M., Botelho, F.M., Wang, P., Harley, C.B. and Bacchetti, S. (1994) Stabilization of short telomeres and telomerase activity accompany immortalization of Epstein-Barr virus-transformed human B lymphocytes. *J. Virol.*, **68**, 3410–3414.
- Bryan, T.M., Englezou, A., Gupta, J., Bacchetti, S. and Reddel, R.R. (1995) Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO J.*, **14**, 4240–4248.
- Montalto, M.C. and Ray, F.A. (1996) Telomerase activation during the linear evolution of human fibroblasts to tumorigenicity in nude mice. *Carcinogenesis*, **17**, 2631–2634.
- Morin, G.B. (1989) The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell*, **59**, 521–529.
- Counter, C.M., Hirte, H.W., Bacchetti, S. and Harley, C.B. (1994) Telomerase activity in human ovarian carcinoma. *Proc. Natl. Acad. Sci. USA*, **91**, 2900–2904.

20. Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C.B., West, M.D., Ho, P.L.C., Coviello, G.M., Wright, W.E., Weinrich, S.L. and Shay, J.W. (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science*, **266**, 2011–2015.
21. Shay, J.W. and Bacchetti, S. (1997) A survey of telomerase activity in human cancer. *Eur. J. Cancer*, **33A**, 787–791.
22. Morin, G.B. (1995) Is telomerase a universal cancer target? *J. Natl. Cancer Inst.*, **87**, 859–861.
23. Murnane, J.P., Sabatier, L., Marder, B.A. and Morgan, W.F. (1994) Telomere dynamics in an immortal human cell line. *EMBO J.*, **13**, 4953–4962.
24. Bryan, T.M. and Reddel, R.R. (1997) Telomere dynamics and telomerase activity in *in vitro* immortalised human cells. *Eur. J. Cancer*, **33A**, 767–773.
25. McEachern, M.J. and Blackburn, E.H. (1996) Cap-prevented recombination between terminal telomeric repeat arrays (telomere CPR) maintains telomeres in *Kluyveromyces lactis* lacking telomerase. *Genes Dev.*, **10**, 1822–1834.
26. Bryan, T.M., Englezou, A., Dalla-Pozza, L., Dunham, M.A. and Reddel, R.R. (1997) Evidence for an alternative mechanism for maintaining telomere length in human tumours and tumour-derived cell lines. *Nature Med.*, (in press).
27. Feng, J., Funk, W.D., Wang, S.-S., Weinrich, S.L., Avilion, A.A., Chiu, C.-P., Adams, R.R., Chang, E., Allsopp, R.C., Yu, J.H., Le, S.Y., West, M.D., Harley, C.B., Andrews, W.H., Greider, C.W. and Villeponteau, B. (1995) The RNA component of human telomerase. *Science*, **269**, 1236–1241.
28. Blasco, M.A., Funk, W., Villeponteau, B. and Greider, C.W. (1995) Functional characterization and developmental regulation of mouse telomerase RNA. *Science*, **269**, 1267–1270.
29. Nakayama, J.-I., Saito, M., Nakamura, H., Matsuura, A. and Ishikawa, F. (1997) *TLPI*: A gene encoding a protein component of mammalian telomerase is a novel member of WD repeats family. *Cell*, **88**, 1–20.
30. Harrington, L., McPhail, T., Mar, V., Zhou, W., Oulton, R., Amgen EST Program, Bass, M.B., Arruda, I. and Robinson, M.O. (1997) A mammalian telomerase-associated protein. *Science*, **275**, 973–977.
31. Lingner, J., Hughes, T.R., Shevchenko, A., Mann, M., Lundblad, V. and Cech, T.R. (1997) Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science*, **276**, 561–567.
32. Nakamura, T.M., Morin, G.B., Chapman, K.B., Weinrich, S.L., Andrews, W.H., Lingner, J., Harley, C.B. and Cech, T.R. (1997) Telomerase catalytic subunit homologs from fission yeast and human. *Science*, **277**, 955–959.
33. Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, **44**, 283–292.
34. Giguere, V., Tini, M., Flock, G., Ong, E., Evans, R.M. and Otulakowski, G. (1994) Isoform-specific amino-terminal domains dictate DNA-binding properties of ROR alpha, a novel family of orphan hormone nuclear receptors. *Genes Dev.*, **8**, 538–553.
35. Ferrigno, O., Virolle, T., Galliano, M.F., Chauvin, N., Ortonne, J.P., Meneguzzi, G. and Aberdam, D. (1997) Murine laminin alpha3A and alpha3B isoform chains are generated by usage of two promoters and alternative splicing. *J. Biol. Chem.*, **272**, 20502–20507.
36. Jitrapakdee, S., Booker, G.W., Cassady, A.I. and Wallace, J.C. (1997) The rat pyruvate carboxylase gene structure. Alternate promoters generate multiple transcripts with the 5'-end heterogeneity. *J. Biol. Chem.*, **272**, 20522–20530.
37. Florkiewich, R.V. and Sommer, A. (1989) Human basic fibroblast growth factor gene encodes four polypeptides: three initiate translation from non-AUG codons. *Proc. Natl. Acad. Sci. USA*, **86**, 3978–3981.
38. Li, H., Zhao, L.-L., Funder, J.W. and Liu, J.-P. (1997) Protein phosphatase 2A inhibits nuclear telomerase activity in human breast cancer cells. *J. Biol. Chem.*, **272**, 16729–16732.
39. Shafman, T., Khanna, K.K., Kedar, P., Spring, K., Kozlov, S., Yen, T., Hobson, K., Gatei, M., Zhang, N., Watters, D., Egerton, M., Shiloh, Y., Kharbanda, S., Kufe, D. and Lavin, M.F. (1997) Interaction between ATM protein and c-Abl in response to DNA damage. *Nature*, **387**, 520–523.
40. Saraste, M., Sibbald, P.R. and Wittinghoffer, A. (1990) The P-loop—a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.*, **15**, 430–434.
41. Devereaux, J., Haeblerli, P. and Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.*, **12**, 387–395.
42. Bryan, T.M., Marusic, L., Bacchetti, S., Namba, M. and Reddel, R.R. (1997) The telomere lengthening mechanism in telomerase-negative immortal human cells does not involve the telomerase RNA subunit. *Hum. Mol. Genet.*, **6**, 921–926.
43. Adams, M.D., Rudner, D.Z. and Rio, D.C. (1996) Biochemistry and regulation of pre-mRNA splicing. *Curr. Opin. Cell Biol.*, **8**, 331–339.
44. Kramer, K.M. and Haber, J.E. (1993) New telomeres in yeast are initiated with a highly selected subset of TG_{1–3} repeats. *Genes Dev.*, **7**, 2345–2356.
45. Avilion, A.A., Piatyszek, M.A., Gupta, J., Shay, J.W., Bacchetti, S. and Greider, C.W. (1996) Human telomerase RNA and telomerase activity in immortal cell lines and tumor tissues. *Cancer Res.*, **56**, 645–650.
46. De Silva, R. and Reddel, R.R. (1993) Similar simian virus 40-induced immortalization frequency of fibroblasts and epithelial cells from human large airways. *Cell. Mol. Biol. Res.*, **39**, 101–110.
47. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York.
48. Maruyama, I.N., Rakow, T.L. and Maruyama, H.I. (1995) cRACE: a simple method for identification of the 5' end of mRNAs. *Nucleic Acids Res.*, **23**, 3796–3797.
49. Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.

NOTE ADDED IN PROOF

hTCS has also been cloned by Meyerson *et al.* [*Cell* (1997) **90**, 785–795] who refer to it as hEST2.