and 0.0234, respectively) and in tumours with distant metastasis (stage IV *vs* others), but not significantly. The levels of *hTERT* mRNA expression (mean 49.9, $n=22$) and telomerase activity (mean 369.1 TPG, $n=14$) in tumours obtained after preoperative chemotherapy did not significantly differ from those in tumours obtained without any therapies (mean 54.4, $n=42$ and mean 435.4, TPG, $n=25$, respectively). There was no significant correlation between the levels of serum AFP and *hTERT* mRNA expression or telomerase activity.

Correlation between *hTERT* mRNA expression and prognosis of the patients

The median follow-up in the series of patients examined was 74 months (range, 1–288 months). Kaplan–Meier event-free survival (EFS) curves of all patients (Figure 2A) show that the 10-year EFS rate in the patients with high *hTERT* mRNA expression ($\geqslant 10$ copies 1000 copies^{-1} of the *PBDG* gene) was 38%, while that in the remaining patients was approximately 90%. The prognosis of the patients with high expression of *hTERT* was significantly worse than that of other patients ($\chi^2=23.40$, $P<0.0001$). Since the levels of *hTERT* mRNA were significantly correlated with advanced stages of tumour, the correlation between *hTERT* mRNA expression and prognosis was examined in tumours in early stages (stage I or II), and those in advanced stages (stage III or IV), separately (Figure 2B). The prognosis of the patients with high levels of *hTERT* mRNA expression was significantly poor in advanced tumours ($\chi^2=26.03$, $P<0.0001$). In 26 patients with early tumours, all 20 patients with low levels of *hTERT* mRNA expression are alive disease free and two out of six patients with high levels of *hTERT* mRNA expression showed poor prognosis ($\chi^2=3.291$, $P=0.046$).

Correlation between the levels of telomerase activity and prognosis of the patients

This study attempted to determine the effect of telomerase activity and *hTERT* mRNA expression on the prognosis of patients with hepatoblastoma. Telomerase activity was investigated in only 39 cases because frozen tumour tissue was unavailable in the remaining 24 cases. Kaplan–Meier EFS curves of these 39 patients (Figure 2C) show that the 10-year EFS rate in the patients with high telomerase activity (TPG $\geqslant 100$) was approximately 40%, while that in the remaining patients was approximately 90%. The prognosis of the patients with high telomerase activity (TPG $\geqslant 100$) was significantly worse than that of other patients ($P=0.0003$). Since the levels of telomerase activity were significantly correlated with advanced stages of tumour, the correlation between telomerase activity and prognosis was examined in the tumours in early stages (stage I or II) and those in advanced stages (stage III or IV), separately (Figure 2D). The prognosis of the patients with high telomerase activity was significantly poor in advanced tumours ($\chi^2=27.12$, $P<0.0001$). In early tumours, one of two patients with high telomerase activity showed poor prognosis, while all patients with low telomerase activity are alive disease free.

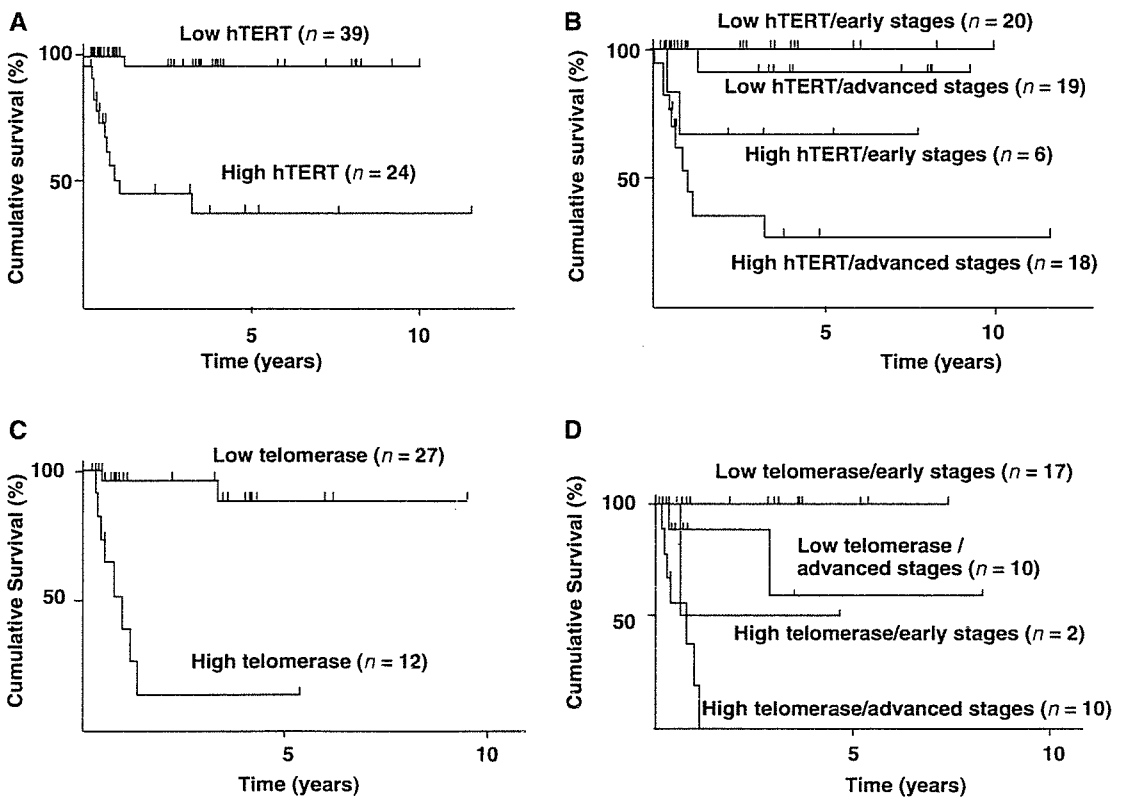

Figure 2 Kaplan–Meier cumulative survival spots for patients with hepatoblastoma. (**A**) Survival according to the levels of *hTERT* mRNA expression. High *hTERT*: *hTERT* mRNA $\geqslant 10$ copies 1000 copies^{-1} of the *PBDG* genes, low *hTERT*: *hTERT* mRNA <10 copies 1000 copies^{-1} of the *PBDG* genes. The patients with tumours with high *hTERT* mRNA expression showed significantly worse survival ($P<0.0001$). (**B**) Survival according to the levels of *hTERT* mRNA expression and stages. Early stages: I and II, advanced stages: III and IV in the stage classification of the Japanese Society of Pediatric Surgeons. Patients having advanced tumours with high *hTERT* mRNA expression showed significantly worse survival ($P<0.0001$). (**C**) Survival according to the levels of telomerase activity. High telomerase: TPG value $\geqslant 100$, low telomerase: TPG value <100. The patients having tumours with high telomerase activity showed significantly worse survival ($P<0.0001$). (**D**) Survival according to the levels of telomerase activity and stages. The patients having advanced tumours with high telomerase activity showed significantly worse survival ($P<0.0001$).

Prognostic factors

By univariate analysis, we analysed clinical parameters such as PRETEXT classification, distant metastasis, stage classification, serum levels of AFP, histological classification, preoperative chemotherapy, and total surgical resection, for the correlation with prognosis of the patients. PRETEXT 2, 3, and 4 tumours and the tumours with distant metastasis showed poorer prognosis, but not significantly. On the other hand, advanced stages (stages 3 and 4) were significantly correlated with poor prognosis of the patients ($P=0.022$). Thus, both telomerase activation and advanced stages were correlated with poor prognosis of patients.

To identify which independent factors had a significant influence on survival, multivariate survival analysis using the Cox proportional hazard regression model was performed. In this multivariate analysis, we assessed the prognostic value for event-free survival of the parameters that were significant in univariate analysis: stage and *hTERT* mRNA expression. For this multivariate analysis, variables with *P*-value lower than 0.30 in the univariate analysis were also selected: gender, age at diagnosis, curative surgery, stage/PRETEXT classification, histology, and hTERT/telomerase activity. As stage classification was significantly associated with PRETEXT and distant metastasis, we used stage classification in this multivariate analysis. As the levels of *hTERT* mRNA expression were significantly correlated with the levels of telomerase activity, and telomerase activity could not be analysed in 24 cases, we analysed these two factors separately in different multivariate analysis sheets. In the multivariate analysis including *hTERT* mRNA expression and the other five factors for 63 cases, *hTERT* mRNA and poorly differentiated histology were independent predictors of EFS. The hazard ratios were 50.0 (95% confidence interval of 5.07–492.9, $P=0.0008$) and 5.11 (95% confidence interval of 1.16–22.5, $P=0.031$). In the multivariate analysis including telomerase activity and the other five factors for 39 cases, the level of telomerase activity was also an independent predictor of EFS. The hazard ratio of telomerase activity levels was 17.7 (95% confidence interval of 2.16–120.1, $P=0.0032$). In advanced stages, the hazard ratios for *hTERT* mRNA and telomerase activity levels were 9.221 ($P=0.043$) and 5.248 ($P=0.188$), respectively.

DISCUSSION

Clinical investigation revealed that the prognosis of children with hepatoblastoma correlates with multifocal growth in the liver, invasion of blood vessels, distant metastasis, and either very low or high levels of serum AFP (von Schweinitz *et al*, 1997; Brown *et al*, 2000). Survival rates of children with more than two of these factors were less than 10%. Thus, these findings discriminate a subgroup of hepatoblastoma with more aggressive biological properties, which correlate with a poor prognosis. However, these factors are not sufficient to predict the prognosis of children with hepatoblastoma. Recently, high-dose chemotherapy with stem cell transplantation has become effective in some patients with aggressive hepatoblastoma with metastasis (Nishimura *et al*, 2002). Thus, to identify such high-risk patients with hepatoblastoma, we need additional useful prognostic markers for evaluation of aggressive biological properties.

Telomerase activity has been reported in many kinds of malignant tumours, including gastric cancer (Hiyama *et al*, 1995b), hepatocellular carcinoma (Nouso *et al*, 1996; Nakashio *et al*, 1997), pancreatic cancer (Hiyama *et al*, 1997b; Iwao *et al*, 1997), and colorectal cancer (Chadeneau *et al*, 1995; Naito *et al*, 2001). Approximately 80–90% of these malignant tumours showed telomerase activity (Shay and Bacchetti, 1997). In some kinds of tumours, high telomerase activity has been reported as a marker of tumour aggressiveness and poor prognosis (Hiyama *et al*, 1995a, b; Shay *et al*, 1997; Marchetti *et al*, 1999). In childhood tumours, telomerase activity and *hTERT* mRNA expression were also detected in a majority of cases of neuroblastoma, retinoblastoma, and nephroblastoma. In neuroblastoma, we have already reported a significant correlation between high telomerase activity and poor outcomes of patients (Hiyama *et al*, 1995a, 1997a). In retinoblastoma, telomerase activity was detected in about 50% of the tumours and such tumours showed a high recurrence rate (Gupta *et al*, 1996). In the present study, *hTERT* mRNA expression and telomerase activity were correlated with poor prognosis of patients, indicating these factors are useful prognosis-predicting factors in hepatoblastoma. Thus, activation of telomerase may correlate with malignant potential in most childhood malignant tumours including hepatoblastoma, neuroblastoma, and retinoblastoma.

This is the first report to show an association between the levels of *hTERT* mRNA expression or telomerase activity and patient prognosis in hepatoblastoma. In the multivariate analysis, activation of telomerase, stage of disease, and histological type were significantly correlated with the outcome of patients. In these three independent parameters, the risk of *hTERT* mRNA or telomerase activation was highest, indicating that telomerase reactivation is the most useful prognosis-associating factor in hepatoblastoma. In the present study, four (15.4%) of the 26 cases with early stage hepatoblastoma showed recurrence of tumours, and all four cases showed high telomerase activity (TPG $\geqslant 100$ or *hTERT* mRNA $\geqslant 100$ copies). In contrast, three (16.7%) of the 18 advanced cases with high telomerase activity remain disease-free. Since all stage 4 cases underwent different chemotherapeutic regimen in the JPLT study (Sasaki *et al*, 2002), one explanation for this result is that the high-dose chemotherapy might have been effective in preventing recurrence in these four early cases. Thus, in early stage tumours, selection of patients for high-dose chemotherapy based on high telomerase activity (TPG $\geqslant 100$) might be an effective method to improve the prognosis of this category of patient. Moreover, the exclusion of low-risk patients from postoperative chemotherapy could spare some of its serious side effects. In advanced hepatoblastoma with low malignancy, complete resection and chemotherapy should be performed, but in such tumours with high malignancy, complete resection and chemotherapy might be insufficient and new aggressive strategies should be implemented. The observations in our study suggest that telomerase inhibition is an effective strategy for the reversal of tumour growth. Since most somatic cells do not have detectable telomerase activity and telomerase shows a tumour-specific expression in general, telomerase is an important target for new anticancer therapy. A number of different approaches have been developed for telomerase inhibition in human cancer. Different components and type of inhibitors targeting various regulatory levels have been regarded as useful telomerase inhibitors and seem to be most efficient when combined with conventional chemotherapy (Saretzki, 2003). Telomerase inhibition, which may be involved in triggering apoptosis, may be a new strategy for curing hepatoblastoma in the future.

In the present study, we analysed the clinical variables of hepatoblastoma cases, but did not find significant correlation between the levels of *hTERT* mRNA or telomerase activity and these variables except for PRETEXT system and disease-stage. It is well-known that prognosis of the cases with pure-foetal histology is good (Finegold, 2002). In the present study, we had only two pure-foetal subtypes in 33 well-differentiated tumours. Although the levels of *hTERT* mRNA or telomerase activity in them were relatively low, further study with large number of this subtype is necessary to analyse statistically.

Some noncancerous childhood liver tissues showed low levels of telomerase activity and *hTERT* mRNA expression. In childhood liver tissue infiltrating lymphocytes, multipotential stem cells, and their daughter cells might have telomerase activity, resulting in positive results by the contamination of lymphocytes and stem

cells with telomerase activity. Recently, it was reported that telomere maintenance by the existence of telomerase activity is necessary for the proliferation of normal human cells (Masutomi *et al*, 2003). Thus, low levels of telomerase activity may reflect the proliferation of normal hepatocytes in children. To solve this false-positive problem, *in situ* evaluation is necessary to analyse the origin of telomerase expression in clinical samples using *hTERT* mRNA ISH (Chou *et al*, 2001; Kumaki *et al*, 2001; Kotoula *et al*, 2002) or hTERT immunohistochemistry (Yasui *et al*, 1999; Hiyama *et al*, 2001).

In summary, we show that an increased level of *hTERT* mRNA expression or telomerase activity is a prognostic indicator of poor outcome in patients with hepatoblastoma, independent of disease stage and histological classification. Although it would need large series to clarify the correlation between clinical variables and the levels of *hTERT* mRNA or telomerase activity, high telomerase activity may stratify patients that are likely to have cancer recurrence requiring postoperative aggressive chemoadjuvant therapy, or, in the future, telomerase-targeting therapy.

ACKNOWLEDGEMENTS

We acknowledge I Fukuba from Department of Surgery, Graduate School of Biomedical Sciences (Hiroshima University) for technical assistance and E Isogai, A Morohasi, and N Sugimitsu (Chiba Center Cancer Institute) for preparing RNA. We also thank Drs H Kenmotsu (Division of Surgery, Ibaraki Children's Hospital), O Ijichi (Department of Pediatrics, Kagoshima University School of Medicine), H Ikawa (Department of Pediatric Surgery Kanazawa Medical University), H Nakadate (Department of Pediatrics, Kitasato University School of Medicine), H Hosoi (Department of Pediatrics, Kyoto Prefectural University of Medicine), T Noda (Department of Pediatrics, Kochi Municipal Central Hospital), H Fujita (Department of Pediatrics, Juntendo University School of Medicine), S Hasegawa (Division of Pediatrics, Nagoya Memorial Hospital), M Iwafuchi (Department of Pediatric Surgery, Niigata University School of Medicine), E Ito (Department of Pediatrics, Hirosaki University School of Medicine), H Ayukawa (Department of Pediatrics, Yamaguchi University School of Medicine), Y Tsuchida (Gunma Children's Hospital Medical Center), J Yokoyama (Department of Surgery, Keio University School of Medicine), A Hayashi (Division of Surgery, Tokyo Metropolitan Kiyose Children's Hospital), M Miyake (Department of Pediatrics, Osaka Medical College), T Matsuyama, T Sugito (Department of Pediatrics, Nagoya First Red Cross Hospital), H Kurosawa (Department of Pediatrics, Dokkyo University School of Medicine) and K Ohtsu (Hiroshima Prefectural Hospital) for providing the HB tissue samples to JPLT or Hiroshima University. This work was supported by a Grant-in-Aid for Scientific Research (A) (No. 13307050) and (B) (No. 12470372) from the Ministry of Education, Culture, Sports, and Technology Science in Japan.

REFERENCES

Albrecht S, von Schweinitz D, Waha A, Kraus JA, von Deimling A, Pietsch T (1994) Loss of maternal alleles on chromosome arm 11p in hepatoblastoma. *Cancer Res* **54:** 5041–5044

Blackburn EH (1991) Structure and function of telomeres. *Nature (London)* **350:** 569–573

Brown J, Perilongo G, Shafford E, Keeling J, Pritchard J, Brock P, Dicks-Mireaux C, Phillips A, Vos A, Plaschkes J (2000) Pretreatment prognostic factors for children with hepatoblastoma – results from the International Society of Paediatric Oncology (SIOP) study SIOPEL 1. *Eur J Cancer* **36:** 1418–1425

Chadeneau C, Hay K, Hirte HW, Gallinger S, Bacchetti S (1995) Telomerase activity associated with acquisition of malignancy in human colorectal cancer. *Cancer Res* **55:** 2533–2536

Chou SJ, Chen CM, Harn HJ, Chen CJ, Liu YC (2001) *In situ* detection of *hTERT* mRNA relates to Ki-67 labeling index in papillary thyroid carcinoma. *J Surg Res* **99:** 75–83

Chromzynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidium thiocyanate phenol-chloroform extraction. *Anal Biochem* **162:** 156–159

Dome JS, Chung S, Bergemann T, Umbricht CB, Saji M, Carey LA, Grundy PE, Perlman EJ, Breslow NE, Sukumar S (1999) High telomerase reverse transcriptase (*hTERT*) messenger RNA level correlates with tumor recurrence in patients with favorable histology Wilms' tumor. *Cancer Res* **59:** 4301–4307

Finegold MJ (2002) Chemotherapy for suspected hepatoblastoma without efforts at surgical resection is a bad practice. *Med Pediatr Oncol* **39:** 484–486

Fuchs J, Rydzynski J, Von Schweinitz D, Bode U, Hecker H, Weinel P, Burger D, Harms D, Erttmann R, Oldhafer K, Mildenberger H (2002) Pretreatment prognostic factors and treatment results in children with hepatoblastoma: a report from the German Cooperative Pediatric Liver Tumor Study HB 94. *Cancer* **95:** 172–182

Gupta J, Han LP, Wang P, Gallie BL, Bacchetti S (1996) Development of retinoblastoma in the absence of telomerase activity. *J Natl Cancer Inst* **88:** 1152–1157

Haas JE, Muczynski KA, Krailo M, Ablin A, Land V, Vietti TJ, Hammond GD (1989) Histopathology and prognosis in childhood hepatoblastoma and hepatocarcinoma. *Cancer* **64:** 1082–1095

Hata Y (1990) The clinical features and prognosis of hepatoblastoma: follow-up studies done on pediatric tumors enrolled in the Japanese pediatric tumor registry between 1971 and 1980. Part I. committee of malignant tumors, Japanese Society of Pediatric Surgeons. *Jpn J Surg* **20:** 498–502

Harley CB, Kim NW, Prowse KL, Weinrich SL, Hirsh KS, West MD, Bacchetti S, Hirre HW, Counter CM, Greider CW, Piatyszek MA, Wright WE, Shay JW (1994) Telomerase, cell immortality, and cancer. *Cold Spring Harbor Symp Quant Biol* **59:** 307–315

Hiyama E, Hiyama K (2002) Clinical utility of telomerase in cancer. *Oncogene* **21:** 643–649

Hiyama E, Hiyama K (2003) Telomerase as tumor marker. *Cancer Lett* **194:** 221–233

Hiyama E, Hiyama K, Ohtsu K, Yamaoka H, Ichikawa T, Shay JW, Yokoyama T (1997a) Telomerase activity in neuroblastoma: is it a prognostic indicator of clinical behavior? *Eur J Cancer* **33:** 1932–1936

Hiyama E, Hiyama K, Shay JW, Yokoyama T (2001) Immunohistochemical detection of telomerase (hTERT) protein in human cancer tissues and a subset of cells in normal tissues. *Neoplasia* **3:** 17–26

Hiyama E, Hiyama K, Yokoyama T, Matsuura Y, Piatyszek MA, Shay JW (1995a) Correlating telomerase activity levels with human neuroblastoma outcomes. *Nat Med* **1:** 249–255

Hiyama E, Kodama T, Sinbara K, Iwao T, Itoh M, Hiyama K, Shay JW, Matsuura Y, Yokoyama T (1997b) Telomerase activity is detected in pancreatic cancer but not in benign tumors. *Cancer Res* **57:** 326–331

Hiyama E, Yokoyama T, Tatsumoto N, Hiyama K, Imamura Y, Murakami Y, Kodama T, Piatyszek MA, Shay JW, Matsuura Y (1995b) Telomerase activity in gastric cancer. *Cancer Res* **55:** 3258–3262

Iwao T, Hiyama E, Yokoyama T, Tsuchida A, Hiyama K, Murakami Y, Shimamoto F, Shay JW, Kajiyama G (1997) Telomerase activity for the preoperative diagnosis of pancreatic cancer. *J Natl Cancer Inst* **89:** 1621–1623

Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PLC, Coviello GM, Wright WE, Weinrich SL, Shay JW (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science* **266:** 2011–2015

Koch A, Denkhaus D, Albrecht S, Leuschner I, von Schweinitz D, Pietsch T (1999) Childhood hepatoblastomas frequently carry a mutated degradation targeting box of the beta-catenin gene. *Cancer Res* **59:** 269–273

Kotoula V, Hytiroglou P, Pyrpasopoulou A, Saxena R, Thung SN, Papadimitriou CS (2002) Expression of human telomerase reverse transcriptase in regenerative and precancerous lesions of cirrhotic livers. *Liver* **22:** 57–69

Kumaki F, Kawai T, Hiroi S, Shinomiya N, Ozeki Y, Ferrans VJ, Torikata C (2001) Telomerase activity and expression of human telomerase RNA component and human telomerase reverse transcriptase in lung carcinomas. *Hum Pathol* **32:** 188–195

Marchetti A, Bertacca G, Buttitta F, Chella A, Quattrocolo G, Angeletti CA, Bevilacqua G (1999) Telomerase activity as a prognostic indicator in stage I non-small cell lung cancer. *Clin Cancer Res* **5:** 2077–2081

Masutomi K, Yu EY, Khurts S, Ben-Porath I, Currier JL, Metz GB, Brooks MW, Kaneko S, Murakami S, DeCaprio JA, Weinberg RA, Stewart SA, Hahn WC (2003) Telomerase maintains telomere structure in normal human cells. *Cell* **114:** 241–253

Naito Y, Takagi T, Handa O, Ishikawa T, Matsumoto N, Yoshida N, Kato H, Ando T, Takemura T, Itani K, Hisatomi H, Tsuchihashi Y, Yoshikawa T (2001) Telomerase activity and expression of telomerase RNA component and catalytic subunits in precancerous and cancerous colorectal lesions. *Tumour Biol* **22:** 374–382

Nakashio R, Kitamoto M, Tahara H, Nakanishi T, Ide T, Kajiyama G (1997) Significance of telomerase activity in the diagnosis of small differentiated hepatocellular carcinoma. *Int J Cancer* **74:** 141–147

Nishimura S, Sato T, Fujita N, Yamaoka H, Hiyama E, Yokoyama T, Ueda K (2002) High-dose chemotherapy in children with metastatic hepatoblastoma. *Pediatr Int* **44:** 300–305

Nouso K, Urabe Y, Higashi T, Nakatsukasa H, Hino N, Ashida K, Kinugasa N, Yoshida K, Uematsu S, Tsuji T (1996) Telomerase as a tool for the differential diagnosis of human hepatocellular carcinoma. *Cancer* **78:** 232–236

Piatyszek MA, Kim NW, Weinrich SL, Hiyama K, Hiyama E, Wright WE, Shay JW (1995) Detection of telomerase activity in human cells and tumors by a telomeric repeat amplification protocol (TRAP). *Methods Cell Sci* **17:** 1–15

Samuel DP, Tsokos M, DeBaun MR (1999) Hemihypertrophy and a poorly differentiated embryonal rhabdomyosarcoma of the pelvis. *Med Pediatr Oncol* **32:** 38–43

Saretzki G (2003) Telomerase inhibition as cancer therapy. *Cancer Lett* **194:** 209–219

Sasaki F, Matsunaga T, Iwafuchi M, Hayashi Y, Ohkawa H, Ohira M, Okamatsu T, Sugito T, Tsuchida Y, Toyosaka A, Nagahara N, Nishihira H, Hata Y, Uchino J, Misugi K, Ohnuma N (2002) Outcome of hepatoblastoma treated with the JPLT-1 (Japanese study group for pediatric liver tumor) protocol-1: a report from the Japanese study group for pediatric liver tumor. *J Pediatr Surg* **37:** 851–856

Shay JW (1995) Aging and cancer: are telomeres and telomerase the connection? *Molec Med Today* **1:** 378–384

Shay JW, Bacchetti S (1997) A survey of telomerase activity in human cancer. *Eur J Cancer* **33:** 787–791

Shay JW, Werbin H, Wright WE (1997) Telomerase assay in the diagnosis and prognosis of cancer. *Ciba Found Symp* **211:** 148–155

Takayasu H, Horie H, Hiyama E, Matsunaga T, Hayashi Y, Watanabe Y, Suita S, Kaneko M, Sasaki F, Hashizume K, Ozaki T, Furuuchi K, Tada M, Ohnuma N, Nakagawara A (2001) Frequent deletions and mutations of the beta-catenin gene are associated with overexpression of cyclin D1 and fibronectin and poorly differentiated histology in childhood hepatoblastoma. *Clin Cancer Res* **7:** 901–908

Tatsumoto N, Hiyama E, Murakami Y, Imamura Y, Shay JW, Matsuura Y, Yokoyama T (2000) High telomerase activity is an independent prognostic indicator of poor outcome in colorectal cancer. *Clin Cancer Res* **6:** 2696–2701

von Schweinitz D, Hecker H, Schmidt-von-Arndt G, Harms D (1997) Prognostic factors and staging systems in childhood hepatoblastoma. *Int J Cancer* **74:** 593–599

von Schweinitz D, Kraus JA, Albrecht S, Koch A, Fuchs J, Pietsch T (2002) Prognostic impact of molecular genetic alterations in hepatoblastoma. *Med Pediatr Oncol* **38:** 104–108

Watson JD (1972) Origin of concatemeric T7 DNA. *Nat New Biol (London)* **239:** 197–201

Weinberg AG, Finegold MJ (1983) Primary hepatic tumors of childhood. *Hum Pathol* **14:** 512–537

Wick M, Zubov D, Hagen G (1999) Genomic organization and promoter characterization of the gene encoding the human telomerase reverse transcriptase (hTERT). *Gene* **232:** 97–106

Yasui W, Tahara E, Tahara H, Fujimoto J, Naka K, Nakayama J, Ishikawa F, Ide T (1999) Immunohistochemical detection of human telomerase reverse transcriptase in normal mucosa and precancerous lesions of the stomach. *Jpn J Cancer Res* **90:** 589–595

Differential gene expressions during immortalization of normal human fibroblasts and endothelial cells transfected with human telomerase reverse transcriptase gene

TSUTOMU KUMAZAKI[1,5], KEIKO HIYAMA[1], TOMOKO TAKAHASHI[3,6], HIDEAKI OMATSU[1,4], KEIJI TANIMOTO[1], TAKUYA NOGUCHI[1], EISO HIYAMA[2], YOUJI MITSUI[3,6] and MASAHIKO NISHIYAMA[1]

[1]Department of Translational Cancer Research, Research Institute for Radiation Biology and Medicine (RIRBM), [2]Natural Science Center for Basic Research and Development, Hiroshima University, Hiroshima; [3]Center for Tsukuba Advanced Research Alliance, University of Tsukuba, Tsukuba; [4]Department of Hospital Pharmacy, School of Medicine, Kobe University, Kobe, Japan

Received December 29, 2003; Accepted February 12, 2004

Abstract. It is widely accepted that telomerase, which compensates for telomere shortening, is finally activated in almost all kinds of human malignant neoplasms, and ectopic expression of telomerase may endow some kinds of human somatic cells with indefinite proliferation capacity, i.e., immortality. To clarify the intrinsic responses required in acquiring immortality, we investigated the chronological changes in the expression levels of the cell cycle and apoptosis-related genes by real-time RT-PCR in human normal fibroblasts and endothelial cells after *hTERT* transfection. We found that fibroblast MJ90 required intrinsic responses including reversible upregulation of cell-cycle promoting genes and down-regulation of apoptosis-inducing genes in early phase after transfection, whereas the endothelial cell HUE142-2 did not. In addition, the microarray analysis of the fibroblast strains revealed that the dysregulated genes during cellular immortalization were different from those reported in fibroblasts probably having acquired telomere maintenance mechanism concomitant with *hTERT* induction. These findings indicate that cell-type specific differential gene expression after telomerase activation may be important to acquire telomere-maintenance capacity and immortality in some non-cancerous human cells. Investigation of these molecules may elucidate the differences in the capacity of acquiring immortality in cancer and normal somatic cells in future.

Correspondence to: Dr Keiko Hiyama, Department of Translational Cancer Research, Research Institute for Radiation Biology and Medicine, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8553, Japan
E-mail: khiyama@hiroshima-u.ac.jp

Present address: [5]Suzugamine Women's College, Hiroshima; and [6]Department of Pharmaceutical Technology, Faculty of Pharmaceutical Science at Kagawa Campus, Tokushima Bunri University, Kagawa, Japan

Key words: cellular immortalization, telomerase, telomere, hTERT, fibroblast, endothelial cell, apoptosis

Introduction

Most cancer cells in advanced malignancies and established cancer cell lines have the capacity to undergo indefinite cell divisions. However, most normal somatic cells can divide no more than several dozens of times. Thus, cellular immortality is one of the hallmarks of cancer cells *in vivo* and *in vitro*.

Recent research has shown that telomerase, a highly conserved reverse transcriptase that adds G-rich nucleotide repeats onto the ends of chromosomal DNAs (i.e., telomeres), can endow some kinds of mortal cells with the immortal capacity (1,2). Telomerase expression has been observed in almost all immortal cell lines, proliferating germ-line cells, and about 85% of human tumors (3). Furthermore, normal somatic cells with self-renewal capacity over lifespan, such as lymphocytes and hematopoietic progenitor cells, can upregulate telomerase activity upon proliferation (4). Except for rare cases like alternative lengthening of telomere (ALT) cells (5), activation of telomerase is a prerequisite for human cells to acquire immortal capacity.

However, induction of telomerase activity by transfection of *hTERT* gene, which encodes telomerase protein component, does not always provide immortality (6). Although inactivation of the Rb/p16 pathway but not p53 was proposed to be necessary for human cultured cells to be immortalized (7), recent reports searching the genes responsible for the immortalization of *hTERT*-transfected somatic cells demonstrated different genes as candidates (8,9). As the dominancy of cellular senescence over the immortal phenotype was demonstrated (10), these findings indicate that induction of telomerase activity is not sufficient to acquire immortality but requires some intrinsic responses in each cell. Until now, upregulation of a subset of growth inhibiting genes or pro-apoptotic genes at the onset of the senescence and frequent

inactivation of cell cycle-related genes during immortalization have been proposed (11-13).

The above information led us to the hypothesis that differential expression of some apoptosis- or cell cycle-related genes in addition to the activation of telomerase are required for cellular immortalization. We therefore investigated the chronological changes in the expression levels of the genes involved in cellular proliferation and apoptosis, during cellular immortalization of human fibroblasts and endothelial cells.

Materials and methods

Cell culture. A normal neonatal human diploid fibroblast strain MJ90 was kindly provided by Dr J.R. Smith (University of Texas, Health Science Center, San Antonio, TX) and a normal human umbilical vein endothelial cell strain HUE142-2 was isolated by Dr Y. Mitsui. The fibroblasts were cultured in Dulbecco's modified Eagle's minimal essential medium (DMEM; Sigma-Aldrich Japan, Tokyo, Japan) supplemented with 2 mM L-glutamine (Life Technologies, Rockville, MD, USA) and 10% fetal calf serum (FCS; Life Technologies) as previously described (12). Endothelial cells were grown in MCDB151 medium (Sigma-Aldrich Japan) supplemented with 15% fetal bovine serum, 5 μg/ml of heparin (Sigma-Aldrich Japan), and 5 ng/ml of recombinant acidic fibroblast growth factor in a plastic flask (Corning, NY, USA) precoated with bovine fibronectin solution (1 μg/cm^2, Wako, Osaka, Japan) (14).

Transfection of hTERT. The human telomerase reverse transcriptase (*hTERT*) expression plasmids hTRT/FLAG and hTERTn2 were kindly provided by Dr F. Ishikawa (Kyoto University, Kyoto, Japan) (15). The 40 μg of plasmid hTRT/FLAG or hTERTn2 with single cut by a restriction enzyme *Bgl*II or *Nru*I, respectively, was transfected into 10^7 MJ90 cells at PDL (population doubling level) 41 and the HUE142-2 cells at PDL 35 by electroporation (Invitrogen, San Diego, CA, USA) at 330 V. Clones harboring the plasmids were obtained by 400 μg/ml of G418 selection (Wako) and named as TF1 from the MJ90 and nTE4-5 from the HUE142-2 cells. The TF1 cells at PDL 104, 148, and 216 and nTE4-5 cells at PDL 105, 145, and 216 were collected and stored at -80°C until used. Expression of *hTERT* and telomerase activity in TF1 and nTE4-5 cells and absence of them in MJ90 and HUE142-2 cells were confirmed by real-time RT-PCR and telomeric repeat amplification protocol (TRAP) assay, respectively.

Preparation of RNA. Total RNA was extracted from the cultured cells at each PDL using an RNAeasy™ Mini kit (Qiagen, Hilden, Germany) according to the protocol of the manufacturer. The extracted RNA was stored at -80°C until used.

TRAP assay (16). Using a TRAPeze™ telomerase detection kit (Serological Co., Gaithersburg, MD, USA), frozen cell pellet was lysed in ice-cold CHAPS lysis buffer™. After 30 min of incubation on ice, the lysate was centrifuged at 16,000 g for 20 min at 4°C, and the supernatant was stored at -80°C. An aliquot of the extract equivalent to ~$2x10^4$ cells was used for each TRAP assay using the TRAPeze kit: after 30 min of incubation at 30°C for telomerase mediated extension of the TS™ primer, the reaction mixture was subjected to PCR for 30 cycles of 94°C x 30 sec, 59°C x 30 sec, and 72°C x 30 sec. The PCR product was electrophoresed on a 12.5% acrylamide gel, and the telomerase signals were detected as 6-bp ladders after staining with SYBR™ Green (Molecular Probes, Eugene, OR, USA).

Southern blot analysis. Terminal restriction fragment (TRF) length was determined by Southern blot analysis. For MJ90 and TF1 strains, genomic DNA was extracted from the cell pellets using DNA Extractor WB Kit™ (Wako). DNA (5 μg) was digested to completion with *Hin*fI, subjected to electrophoresis on a 0.8% agarose gel, blotted onto a nitrocellulose filter, and then hybridized to a [γ-^{32}P]ATP labeled $(TTAGGG)_4$ probe at 50°C. The filter was washed in 4X SSC (1X = 0.15 M NaCl, 0.015 M sodium citrate) and 0.1% SDS at 55°C four times and exposed to X-ray film. For HUE142-2 and nTE4-5 strains, TRF lengths were measured as previously described (11). The mean length of TRFs was visually determined as the peak size of the smear signals and confirmed by BAS™ 2000 (Fuji, Tokyo, Japan).

Sequence analysis. For *TP53*, exons 1-11 were amplified from the cDNA of each strain using primers located in exons 1 and 11: 5'-TTCCACGACGGTGACACG-3' and 5'-GTGGG AGGCTGTCAGTGGGGAACAA-3'. After confirming the existence of 1.3-kb normal-size band, 1 μl aliquot of the first-step PCR product was subjected to the nested PCR for exons 5-8, the hot spots of mutations, to carry out sequence analysis using primers located in exons 4 and 9: 5'-TCTGTCCCTTC CCAGAAAACC-3' and 5'-AGAGGAGCTGGTGTTGTT GG-3'.

In addition, for MJ90 and TF1 cells, exons 3-4, 5-6, 7, 8-9 were amplified from the genomic DNA for confirmation using primers located in introns as follows: exons 3-4, 5'-GG ACTGACTTTCTGCTCTTG-3' and 5'-TGAAGTCTCATGG AAGCCAG-3'; exons 5-6, 5'-TGTTCACTTGTGCCCTGA CT-3' and 5'-GAGGTCAAATAAGCAGCAGG-3'; exon 7, 5'-CTTGCCACAGGTCTCCCCAA-3' and 5'-AGGGGTCAGC GGCAAGCAGA-3'; exons 8-9, 5'-TGGGACAGGTAGGAC CTGAT-3' and 5'-ACTTGATAAGAGGTCCCAAG-3'. The PCR products purified by QIAquick™ Gel Extraction Kit (Qiagen) were amplified using BigDye™ Terminator Cycle Sequencing Kit (Applied Biosystems, CA, USA), purified, and subjected to automated sequencer ABI PRISM™ 310 (Applied Biosystems), as recommended by the manufacturer.

For *RB1*, exons 10-27 were amplified from the cDNA of each cell strain using primers as follows: for exons 10-19, 5'-CTAATGGACTTCCAGAGGTT-3' and 5'-CGGAGATAGG CTAGCCGATA-3'; for exon 20-5'-site of exon 27, 5'-TACT GCAAATGCAGAGACAC-3' and 5'-GAAGAGGAAACAA TCTGCTA-3', to cover the tumor-suppressor pocket domain where most of the Rb-binding cellular and viral proteins make their primary contacts (17).

Real-time RT-PCR. Total RNA (2 μg) extracted from each cell strain was reverse-transcribed using High-Capacity cDNA Archive™ Kit (Applied Biosystems). One hundred

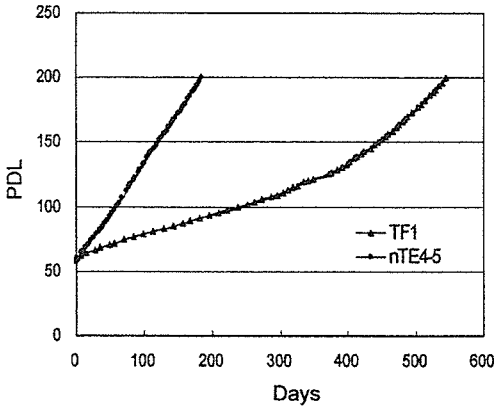

Figure 1. Growth curve of the TF1 and the nTE4-5 strains. The growth of the TF1 strain (closed triangles) was slow until ~120 PDL, +80 PDL after the *hTERT* transfection, and turned rapid thereafter. The growth of the nTE4-5 strain (closed circles) was comparably rapid at any PDL.

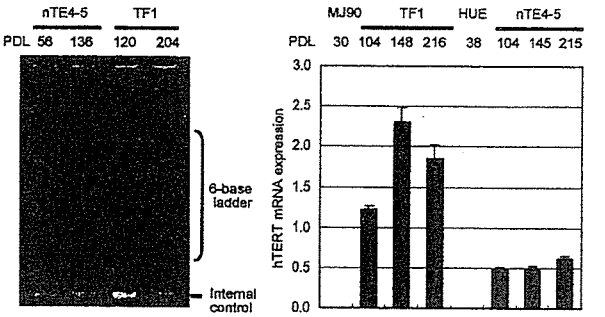

Figure 2. Telomerase activity and *hTERT* mRNA expression in *hTERT* transfected cells. The *hTERT* transfected endothelial cells (nTE4-5) and fibroblasts (TF1) showed 6-base ladder signals by TRAP assay (left) and *hTERT* mRNA expression by real-time RT-PCR using full-length mRNA specific primer set (right), while no *hTERT* mRNA was detected in either strain before transfection, MJ90 and HUE142-2.

aliquot of the cDNA (equivalent to 20 ng total RNA) was subjected to real-time RT-PCR using Assays-on-Demand™ Gene Expression products (Applied Biosystems) for *PCNA*, *RB1*, *TP53*, *APAF1*, *CDKN1A*, *CCND1*, and internal controls *ACTB*, *GAPD*, and *TBP*, Pre-Developed TaqMan™ Assay Reagents (Applied Biosystems) for *BAX*, *FAS*, *MYC*, *BCL2*, and *BCL2L1*, and originally designed TaqMan probe and primer set for *hTERT*. The feasibility of using *ACTB* as the internal control was confirmed by the comparison of the amount of total RNA and real-time RT-PCR results of *ACTB*, *GAPD* and *TBP* in each strain. mRNA expression level was calculated by the ratio in the value of target gene to *ACTB*, and relative quantity was expressed compared to that of control cDNA consisting of the same amount of cDNA of the 8 cell strains examined. Each reaction was carried out in duplicate using ABI PRISM 7900 or triplicate using ABI PRISM 7700 Sequence Detection System (Applied Biosystems).

Microarray analysis. Biotin-labeled cRNA was synthesized from 5 μg total RNA of each MJ90 or TF1 (PDL 216) strain using First and Second Strand cDNA Synthesis Kit™ (CodeLink), *In vitro* Transcription (IVT) Kit™ (CodeLink), and Biotin-11-UTP (Perkin-Elmer), according to the manufacturer's protocols. After purification, each cRNA sample was fragmented, and hybridized with a CodeLink UniSet Human 20K I Bioarray™ using Parallel Processing Kit™ (CodeLink) for 18 h at 37°C in an Innova™ 4080 incubator (New Brunswick Scientific) at 300 rpm. The hybridized arrays were rinsed with 0.75 x TNT buffer at 46°C for 1 h, labeled with Streptavidin-Cy5 at room temperature for 30 min, rinsed with TNT buffer and 0.05% Tween-20, dried by centrifugation, and then scanned using GenePix 4000B Array Scanner™ (Amersham Biosciences, Tokyo, Japan) and/or Agilent DNA Microarray Scanner™ (Agilent, Palo Alto, CA, USA). The experiment was carried out twice independently.

Statistical analysis. Statistical significance was analyzed using the software package Statview™ 5.0 (Abacus Concepts, Berkeley, CA, USA) for Student *T* analysis on the mRNA expression levels between the strains before and after the *hTERT* transfection.

Results

Cell growth. After the transfection of the *hTERT*-expression plasmids into the MJ90 cells at PDL 41 and the HUE142-2 cells at PDL 35, the isolated clones TF1 and nTE4-5 were serially passaged until 216 PDL. The growth curve of the TF1 strain was slow until ~PDL 120 and turned to rapid growth thereafter, while it was straight and rapid from just after the *hTERT* transfection in the nTE4-5 cells (Fig. 1).

Telomerase activity. By TRAP assay, the 6-bp ladders were observed in both TF1 and nTE4-5 strains (Fig. 2, left), indicating the existence of telomerase activity in *hTERT* transfected cells.

TRF length. The TRF length of each strain was measured by Southern blot analysis. The mean TRF lengths determined by the peak sizes of the TRF smear signals were 8.1, 6.1, 6.8, and 8.4 kb for MJ90, TF1 PDL 104, 148 and 216 cells, respectively (Fig. 3, left), while that in HUE142-2 was 8.0 kb and no significant change was observed after the *hTERT* transfection. The senescent HUE142-2 cells without *hTERT* transfection showed telomere shortening to 4.9 kb at 73 PDL (Fig. 3, right).

Differential gene expressions. In the experiment of real-time RT-PCR, *ACTB* was used as an internal control, after the confirmation of its feasibility comparing the amounts of *ACTB*, *GAPD*, and *TBP* mRNA estimated by the real-time RT-PCR with the amount of total RNA (data not shown). The *hTERT* mRNA expression was detected in TF1 PDL 104, 148 and 216 cells and nTE4-5 PDL 105, 145 and 216 cells but not in MJ90 and HUE142-2 cells, by real-time RT-PCR using originally designed TaqMan probe which is specific for the full length *hTERT* mRNA distinguishing the α- and ß-splice variants (18) (Fig. 2, right). While the *hTERT* expression levels in the nTE4-5 strains were comparable among the PDL 105, 145 and 216 cells, those of TF1 cells at PDL 148 and 216

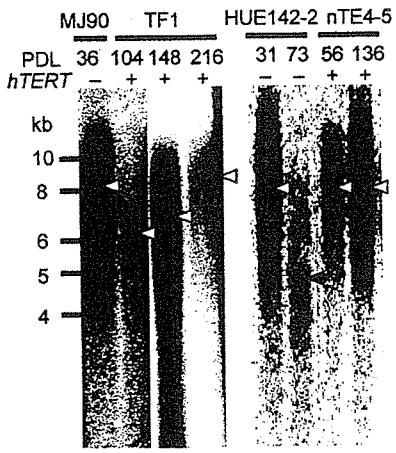

Figure 3. Sourthern blot analysis for TRF length. The peaks of the TRF smear signals are indicated by white arrowheads at 8.1, 6.1, 6.8 and 8.4 kb in MJ90, TF1 PDL 104, 148 and 216 cells, respectively (left), and 8.0 kb in HUE142-2 cells before and after *hTERT* transfection (right). That of the senescent HUE142-2 cells at 73 PDL is shown with black arrowhead at 4.9 kb.

were somewhat higher than that at PDL 104. This pattern was similar to the expression levels of PCNA (Fig. 4) and the growth curve (Fig. 1). Also, the expression level of TP53 was dramatically repressed at PDL 104 and gradually recovered thereafter in the TF1 strains, while such repression was not observed in the nTE4-5 strains (Fig. 4).

The expression levels of *RB1*, *BAX* and *MYC* were somewhat repressed in TF1 cells, while they were upregulated in general in nTE4-5 cells (Fig. 4). On the contrary, expression level of *BCL2* was apparently upregulated in the TF1 cells at PDL 104 and gradually decreased to the baseline, while no such transient upregulation was observed in nTE4-5 cells. Though, the upregulation of *BCL2L1* was only observed in nTE4-5 cells. The expression level of *CDKN2A*, encoding p16, was once upregulated at PDL 104 and dramatically down-regulated after then in TF1 strains. The expression levels of *FAS* and *APAF1* were comparable during cellular immortalization in both strains, fibroblasts and endothelial cells. Progressive down-regulation of *CDKN1A* was observed in TF1 cells, but not in nTE4-5 cells.

Sequence of RB1 and TP53. By the sequence analysis of the *RB1* cDNA from exon 10 to the 5'-site of exon 27 and *TP53* cDNA in exons 5-8 of both TF1 and nTE4-5 strains, no sequence abnormalities were observed at any PDL. The sequences of the *TP53* exons 4-9 and corresponding splice sites were also examined in the genomic DNA of the TF1 strains, which showed dramatic change in the *TP53* mRNA expression levels, and no abnormalities were detected.

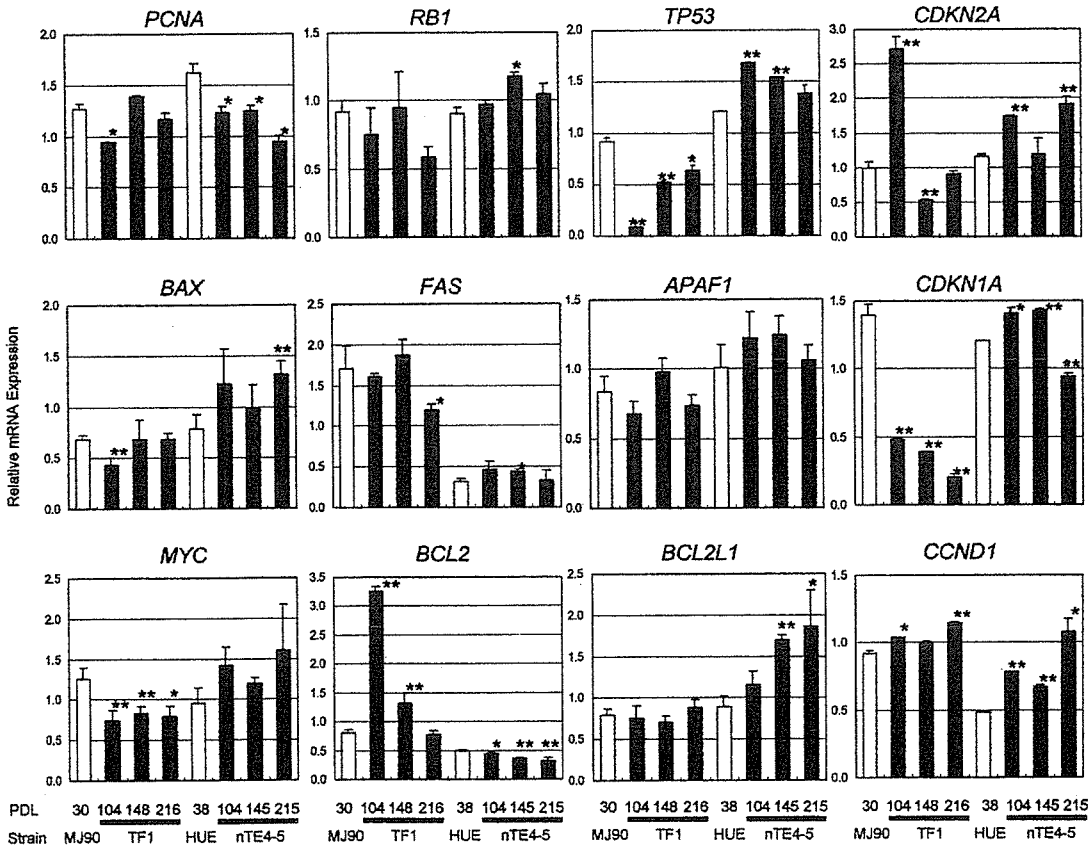

Figure 4. Relative levels of the mRNA expression in each strain. The expression level of each gene was evaluated by real-time RT-PCR using *ACTB* as an internal control, and expressed as mean + SD. The levels in a strain after *hTERT* transfection (closed bar) were statistically compared with those before transfection (open bar). $^{*}P<0.05$; $^{**}P<0.01$.

Table I. Upregulated genes in TF1 at PDL 216 compared with MJ90 by microarray analysis.

Gene	UniGene ID	Cytogenetic
RAGE	Hs.104119	14q32
CLIC2	Hs.54570	Xq28
MCOLN3	Hs.49344	1p22.3
PODXL	Hs.16426	7q32-q33
SPP1	Hs.313	4q21-q25
CEB1	Hs.26663	4q22.1-q23
FOLR3	Hs.352	11q13
HLA-DRB1	Hs.308026	6p21.3
AK3	Hs.10862	9pter-p13
RAB33A	Hs.56294	Xq26.1
LGALS3BP	Hs.79339	17q25
LOC91752	Hs.159528	2q32.1
MX2	Hs.926	21q22.3

Microarray analysis. To investigate intrinsic responses required in the immortalization of fibroblast strain MJ90 other than the upregulation of cell-cycle promoting genes and down-regulation of apoptosis-inducing genes, microarray analysis using oligoarrays were carried out between the MJ90 cells, normal mortal cells, and the TF1 cells at PDL 216, definitely immortalized cells. After global normalization of the spot intensity, the TF1/MJ90 ratio was calculated for each spot. The genes whose TF1/MJ90 ratios of the normalized intensity were ≥4 or ≤0.25 are listed in Tables I (13 genes) and II (101 genes), respectively.

Discussion

By the transfection of *hTERT* gene into fibroblasts MJ90 and endothelial cells HUE142-2, immortal cell strains TF1 and nTE4-5 were obtained. During this immortalization course, no abnormalities were detected in the core regions of *RB1* or *TP53* genes, but the differential expression of the *TP53* was characteristic. It dramatically decreased at PDL 104 in the TF1 cells, but showed recovery after that, indicating that this down-regulation may not be due to a mutational event in the promoter region. The p53 is widely accepted to inhibit cell growth through activation of cell-cycle arrest and apoptosis, and its disruption with inactivation of Rb would provide normal human somatic cells to escape the Hayflick limit, the replicative senescence checkpoint, endowing the cells with elongated lifespan (2).

The TF1 at PDL 104 cells, where telomere maintenance mechanism seems insufficient with shortened telomeres, also showed transient down-regulation of *BAX* and upregulation of *BCL2*, which were not observed in nTE4-5 strains. *BAX* is a p53 primary-response gene (19) and was reported to be essential for death receptor-mediated apoptosis in cancer cells (20). *BCL2* was reported to control a cell-death pathway independently of Apaf-1, which was not differentially expressed either in the present study, in the upstream of initiator caspases in mice (21). Thus, the repression of *TP53* and *BAX* expression and transient upregulation of *BCL2* around PDL 104 might have promoted the resistance to apoptosis in the TF1 cells.

Down-regulation of *MYC* and *CDKN1A* was observed in all TF1 strains, which was absent in nTE4-5 strains. Whereas Myc is known to be a transcriptional activator of the *hTERT* promoter (22) and promotes cell growth and proliferation, it also induces apoptosis under certain conditions (23). More recently Myc was proved to switch the p53-dependent response of cancer cells to DNA damage from cytostatic to apoptotic, inhibiting p21(cip1/waf1) induction by p53 and other activators (24). The p21, encoded by *CDKN1A*, is the effector of p53 cell cycle control and both appeared to be essential for maintaining the G2 checkpoint (25).

The p16, encoded by *CDKN2A*, was once upregulated at PDL 104 and dramatically down-regulated thereafter in TF1 strains. The expression of p16 was previously found to be repressed during cellular immortalization in Epstein-Barr virus-transformed human B-lymphoblastoid cell lines (26). The dramatic down-regulation of the *CDKN2A* expression between PDL 104 and 148 in TF1 strains may indicate that cellular immortalization may have occurred between these PDLs, but the repression of this gene does not seem to be universally required in immortalization, e.g., in nTE4-5 strains.

Thus, repression of *MYC*, *CDKN1A*, *TP53* and/or *CDKN2A* in TF1 strains would have promoted their continuous proliferation and the resistance to apoptosis. Thus, the upregulation of anti-apoptotic genes and down-regulation of pro-apoptotic or apoptosis-related genes were observed in the TF1 cells in general, especially in early phase after *hTERT* transfection where telomere-maintenance mechanism has not been completed, while no such findings were observed in the nTE4-5 cells. Only the *CCND1*, which encodes cyclin D1 and is frequently overexpressed in a broad range of human tumor types as a positive regulator of progression through the G1 phase of the cell cycle interacting with cdk4, pRb, and some transcription factors (27), was overexpressed in both strains.

To investigate other intrinsic responses required in the immortalization of the fibroblast strain, microarray analysis was carried out between the MJ90 cells and the TF1 cells at PDL 216, and the 13 genes and 101 genes were found to be upregulated or down-regulated more than 4-fold (Tables I and II). The *CDKN1A*, whose expression was found to be dramatically repressed after immortalization in real-time RT-PCR analysis was also included as the down-regulated gene in microarray analysis (Table II), supporting the reliability of the analysis. The sequence of the *hTERT*-oligo spotted on the array located near the poly A site, which hybridizes both full length and splice variant *hTERT* mRNAs. So it is reasonable that the *hTERT* expression evaluated by oligoarray analysis was only slightly upregulated (data not shown), whereas it revealed to be clearly upregulated from none to positive after *hTERT* transfection by real-time RT-PCR analysis using full-length *hTERT*-specific primers.

Table II. Down-regulated genes in TF1 at PDL 216 compared with MJ90 by microarray analysis.

Gene	UniGene ID	Cytogenetic
NDN	Hs.50130	15q11.2-q12
PSMB3	Hs.82793	17q12
GDF15	Hs.296638	19p13.1-13.2
SEPP1	Hs.275775	5q31
CYBA	Hs.68877	16q24
KRT19	Hs.309517	17q21.2
PRG1	Hs.1908	10q22.1
PTPN12	Hs.62	7q11.23
FOXF1	Hs.155591	16q24
PPP1R14A	Hs.348037	19q13.1
LAMA4	Hs.437536	6q21
PDLIM4	Hs.424312	5q31.1
ALDH1A1	Hs.76392	9q21.13
TNA	Hs.65424	3p22-p21.3
COL3A1	Hs.443625	2q31
PTGIS	Hs.302085	20q13.11-q13.13
GABARAPL1	Hs.336429	12p13.31
IFITM1	Hs.458414	11p15.5
CCRL1	Hs.310512	3q22
AEBP1	Hs.439463	7p13
GALNT6	Hs.528445	12q13
F3	Hs.62192	1p22-p21
ACAA2	Hs.172506	18q21.1
BST1	Hs.169998	4p15
EFEMP1	Hs.76224	2p16
TAGLN	Hs.410977	11q23.2
CASC3	Hs.350229	17q11-q21.3
CDKN1A	Hs.370771	6p21.2
HAK	Hs.388674	18q21.31-q21.32
WNT2	Hs.89791	7q31
AQP10	Hs.259048	1q22
GPX7	Hs.43728	1p32
SERPINB2	Hs.75716	18q21.3
COX7A1	Hs.421621	19q13.1
BEX1	Hs.334370	Xq21-q23
TPK1	Hs.127548	7q34-q35
EPB41L3	Hs.103839	18p11.32
HAPLN1	Hs.2799	5q14.3
GAL	Hs.278959	11q13.1
T1A2	Hs.468675	1p36
GW112	Hs.273321	13q14.2
SERPING1	Hs.384598	11q12-q13.1
ROBO2	Hs.13305	3p13
POSTN	Hs.136348	13q13.3
SULF1	Hs.409602	8q13.2
SUCLA2	Hs.182217	13q12.2-q13.3
SLC12A7	Hs.172613	5p15
BGN	Hs.821	Xq28
DOK5	Hs.127751	20q13.2
KCNE4	Hs.348522	2q36.3

Table II. Continued.

Gene	UniGene ID	Cytogenetic
HOXD11	Hs.421136	2q31.1
NCAG1	Hs.124673	18q22.1
CDCP1	Hs.146170	3p21.32
RAC2	Hs.301175	22q13.1
PTGER1	Hs.159360	19p13.1
TRO	Hs.434971	Xp11.22-p11.21
CCDC8	Hs.97876	19q13.33
ARHGDIB	Hs.292738	12p12.3
MRVI1	Hs.251385	11p15
NET02	Hs.6823	16q11
LTC4S	Hs.456	5q35
QPCT	Hs.79033	2p22.3
MGC3036	Hs.284135	7q31-q35
POFUT1	Hs.178292	20q11
RAMP1	Hs.32989	2q36-q37.1
C20orf59	Hs.512686	20q13.33
GPR51	Hs.198612	9q22.1-q22.3
ISL1	Hs.505	5q11.2
DKFZP434B044	Hs.262958	16q24.1
FLJ11175	Hs.33368	15q26.1-q26.2
SH2D2A	Hs.103527	1q21
DF	Hs.155597	19p13.3
ZD52F10	Hs.417795	19q13.13
NLGN4	Hs.21107	Xp22.33
SLC38A5	Hs.195155	Xp11.23
FLJ10847	Hs.48403	17p11.2
MST4	Hs.23643	Xq26.2
C9orf67	Hs.134292	9q34.2-q34.3
NFATC1	Hs.96149	18q23
PTGDS	Hs.446429	9q34.2-q34.3
HSD17B2	Hs.155109	16q24.1-q24.2
STMN2	Hs.90005	8q21.11-q21.12
GAS1	Hs.65029	9q21.3-q22
GGTLA1	Hs.437156	22q11.23
STAP2	Hs.194385	19p13.3
IL17RB	Hs.5470	3p21.1
SUSD2	Hs.131819	22q11-q12
TSPAN-2	Hs.234863	1p13.1
ELN	Hs.252418	7q11.23
KISS1	Hs.95008	1q32
CXCL6	Hs.164021	4q21
LRRN1	Hs.126085	3p26.2
PTPRD	Hs.323079	9p23-p24.3
RGMA	Hs.271277	15q26.1
AGXT2L1	Hs.106576	4q25
C12orf14	Hs.356223	12p11
HOXD10	Hs.123070	2q31.1
EPSTI1	Hs.343800	13q13.3
ACVRL1	Hs.410104	12q11-q14
TU3A	Hs.8022	3p21.1
NTNG1	Hs.111224	1p13.3

Interestingly, among the 36 reportedly ≥4-fold dysregulated genes, including 11 upregulated and 25 down-regulated genes, during immortalization of normal human foreskin fibroblasts BJ cells (8), only *LAMA4*, *COL3A1* and *T1A2* were commonly down-regulated ≥4-fold and one third was oppositely dysregulated. Although both MJ90 and BJ cells are derived from the same organ, acquisition of telomere maintenance mechanism after *hTERT* transfection seems different between them, straightforward in the BJ cells acquiring telomere maintenance mechanism (1) and not straight in the MJ90 cells with telomere shortening (Fig. 3).

The inconsistency of the dysregulated genes during *hTERT*-induced immortalization between the cell strains derived from the same origin with differences in the acquisition of telomere maintenance mechanism may also support the idea that some specific intrinsic responses are necessary for *hTERT*-induced immortalization in some kinds of non-cancerous cells. Since no report has found that once telomerase-activated cancer cells lose their proliferation capacity without artificial manipulation, the intrinsic responses required in the immortalization of *hTERT*-induced non-cancerous cells may be the specific characteristics that distinguish cancerous and non-cancerous cells.

It was reported that the telomerase catalytic subunit *hTERT* was expressed at low level in cycling human fibroblasts, which were previously believed to lack *hTERT* expression and telomerase activity (28), indicating that expression of *hTERT* is not necessarily the rate-limiting step for cellular immortalization. Moreover, it was reported that there are at least three tumor suppressor pathways regulating *hTERT* expression (29), suggesting that it is not straightforward but there are different settings for human somatic cells to acquire immortality through telomerase activation.

In conclusion, we propose that immortalization of normal human somatic cells by *hTERT* expression is not straightforward but requires some specific intrinsic responses including upregulation of cell-cycle promoting genes and down-regulation of apoptosis-inducing genes, not in a mutational event manner, in early phase after transfection in some non-cancerous cells. These requirements seem to be decreased after continuous culture concomitant with acquisition of telomere-elongation capacity, and may be the specific characteristics that distinguish non-cancerous cells from cancer cells.

Acknowledgements

We thank Dr F. Ishikawa for providing *hTERT* constructs and Dr J.R. Smith for MJ90 strain. Part of this work was supported by grants from The Science Promotion Fund of the Ministry of Education, Culture, Science, Sports and Technology of Japan.

References

1. Bodnar AG, Ouellette M, Frolkis M, Holt SE, Chiu CP, Morin GB, Harley CB, Shay JW, Lichtsteiner S and Wright WE: Extension of life-span by introduction of telomerase into normal human cells. Science 279: 349-352, 1998.
2. Harley CB: Telomerase is not an oncogene. Oncogene 21: 494-502, 2002.
3. Hiyama E and Hiyama K: Telomerase as tumor marker. Cancer Lett 194: 221-233, 2003.
4. Hiyama K, Hirai Y, Kyoizumi S, Akiyama M, Hiyama E, Piatyszek MA, Shay JW, Ishioka S and Yamakido M: Activation of telomerase in human lymphocytes and hematopoietic progenitor cells. J Immunol 155: 3711-3715, 1995.
5. Bryan TM, Englezou A, Dalla PL, Dunham MA and Reddel RR: Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. Nat Med 3: 1271-1274, 1997.
6. Lundberg AS, Randell SH, Stewart SA, Elenbaas B, Hartwell KA, Brooks MW, Fleming MD, Olsen JC, Miller SW, Weinberg RA and Hahn WC: Immortalization and transformation of primary human airway epithelial cells by gene transfer. Oncogene 21: 4577-4586, 2002.
7. Kiyono T, Foster SA, Koop JI, McDougall JK, Galloway DA and Klingelhutz AJ: Both Rb/p16INK4a inactivation and telomerase activity are required to immortalize human epithelial cells. Nature 396: 84-88, 1998.
8. Lindvall C, Hou M, Komurasaki T, Zheng C, Henriksson M, Sedivy JM, Bjorkholm M, Teh BT, Nordenskjold M and Xu D: Molecular characterization of human telomerase reverse transcriptase-immortalized human fibroblasts by gene expression profiling: activation of the epiregulin gene. Cancer Res 63: 1743-1747, 2003.
9. Smith LL, Coller HA and Roberts JM: Telomerase modulates expression of growth-controlling genes and enhances cell proliferation. Nat Cell Biol 5: 474-479, 2003.
10. Tominaga K, Olgun A, Smith JR and Pereira-Smith OM: Genetics of cellular senescence. Mech Ageing Dev 123: 927-936, 2002.
11. Kumazaki T, Sasaki M, Nishiyama M, Teranishi Y, Sumida H and Mitsui Y: Effect of BCL-2 down-regulation on cellular life span. Biogerontology 3: 291-300, 2002.
12. Kumazaki T, Sasaki M, Nishiyama M, Teranishi Y, Sumida H, Eboshida A and Mitsui Y: Life span shortening of normal fibroblasts by overexpression of BCL-2: a result of potent increase in cell death. Exp Cell Res 285: 299-308, 2003.
13. Schwarze SR, Shi Y, Fu VX, Watson PA and Jarrard DF: Role of cyclin-dependent kinase inhibitors in the growth arrest at senescence in human prostate epithelial and uroepithelial cells. Oncogene 20: 8184-8192, 2001.
14. Kumazaki T, Kobayashi M and Mitsui Y: Enhanced expression of fibronectin during *in vivo* cellular aging of human vascular endothelial cells and skin fibroblasts. Exp Cell Res 205: 396-402, 1993.
15. Nakayama J, Tahara H, Tahara E, Saito M, Ito K, Nakamura H, Nakanishi T, Tahara E, Ide T and Ishikawa F: Telomerase activation by hTRT in human normal fibroblasts and hepatocellular carcinomas. Nat Genet 18: 65-68, 1998.
16. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PLC, Coviello GM, Wright WE, Weinrich SL and Shay JW: Specific association of human telomerase activity with immortal cells and cancer. Science 266: 2011-2015, 1994.
17. Lee JO, Russo AA and Pavletich NP: Structure of the retinoblastoma tumour-suppressor pocket domain bound to a peptide from HPV E7. Nature 391: 859-865, 1998.
18. Ulaner GA, Hu JF, Vu TH, Giudice LC and Hoffman AR: Telomerase activity in human development is regulated by human telomerase reverse transcriptase (hTERT) transcription and by alternate splicing of hTERT transcripts. Cancer Res 58: 4168-4172, 1998.
19. Miyashita T and Reed JC: Tumor suppressor p53 is a direct transcriptional activator of the human bax gene. Cell 80: 293-299, 1995.
20. LeBlanc H, Lawrence D, Varfolomeev E, Totpal K, Morlan J, Schow P, Fong S, Schwall R, Sinicropi D and Ashkenazi A: Tumor-cell resistance to death receptor-induced apoptosis through mutational inactivation of the proapoptotic Bcl-2 homolog Bax. Nat Med 8: 274-281, 2002.
21. Marsden VS, O'Connor L, O'Reilly LA, Silke J, Metcalf D, Ekert PG, Huang DC, Cecconi F, Kuida K, Tomaselli KJ, Roy S, Nicholson DW, Vaux DL, Bouillet P, Adams JM and Strasser A: Apoptosis initiated by Bcl-2-regulated caspase activation independently of the cytochrome c/Apaf-1/caspase-9 apoptosome. Nature 419: 634-637, 2002.
22. Takakura M, Kyo S, Kanaya T, Tanaka M and Inoue M: Expression of human telomerase subunits and correlation with telomerase activity in cervical cancer. Cancer Res 58: 1558-1561, 1998.

23. Grandori C, Cowley SM, James LP and Eisenman RN: The Myc/Max/Mad network and the transcriptional control of cell behavior. Annu Rev Cell Dev Biol 16: 653-699, 2000.
24. Seoane J, Le HV and Massague J: Myc suppression of the p21(Cip1) Cdk inhibitor influences the outcome of the p53 response to DNA damage. Nature 419: 729-734, 2002.
25. Bunz F, Dutriaux A, Lengauer C, Waldman T, Zhou S, Brown JP, Sedivy JM, Kinzler KW and Vogelstein B: Requirement for p53 and p21 to sustain G2 arrest after DNA damage. Science 282: 1497-1501, 1998.
26. Takahashi T, Kawabe T, Okazaki Y, Itoh C, Noda K, Tajima M, Satoh M, Goto M, Mitsui Y, Tahara H, Ide T, Furuichi Y and Sugimoto M: *In vitro* establishment of tumorigenic human B-lymphoblastoid cell lines transformed by Epstein-Barr virus. DNA Cell Biol 22: 727-735, 2003.
27. Lamb J, Ramaswamy S, Ford HL, Contreras B, Martinez RV, Kittrell FS, Zahnow CA, Patterson N, Golub TR and Ewen ME: A mechanism of cyclin D1 action encoded in the patterns of gene expression in human cancer. Cell 114: 323-334, 2003.
28. Masutomi K, Yu EY, Khurts S, Ben-Porath I, Currier JL, Metz GB, Brooks MW, Kaneko S, Murakami S, De Caprio JA, Weinberg RA, Stewart SA and Hahn WC: Telomerase maintains telomere structure in normal human cells. Cell 114: 241-253, 2003.
29. Lin SY and Elledge SJ: Multiple tumor suppressor pathways negatively regulate telomerase. Cell 113: 881-889, 2003.

COMMUNICATIONS IN STATISTICS
Theory and Methods
Vol. 33, No. 10, pp. 2409–2422, 2004

A Note on Multiple Regression for Single Index Model

Kenichi Satoh* and Megu Ohtaki

Department of Environmetrics and Biometrics, Research Institute for Radiation Biology and Medicine, Hiroshima University, Hiroshima, Japan

ABSTRACT

A simple method based on sliced inverse regression (SIR) is proposed to explore an effective dimension reduction (EDR) vector for the single index model. We avoid the principle component analysis step of the original SIR by using two sample mean vectors in two slices of the response variable and their difference vector. The theories become simpler, the method is equivalent to the multiple linear regression with dichotomized response, and the estimator can be

*Correspondence: Kenichi Satoh, Department of Environmetrics and Biometrics, Research Institute for Radiation Biology and Medicine, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8553, Japan; Fax: 81-82-256-7106; E-mail: ksatoh@hiroshima-u.ac.jp.

DOI: 10.1081/STA-200031363 0361-0926 (Print); 1532-415X (Online)

expressed by a closed form, although the objective function might be an unknown nonlinear. It can be applied for the case when the number of covariates is large, and it requires no matrix operation or iterative calculation.

Key Words: Correlation coefficient; Effective dimension reduction; Semiparametric regression; Single index model; Sliced inverse regression.

Mathematics Subject Classification: Primary 62G08; Secondary 62J05.

1. INTRODUCTION AND NOTATION

The family of single index models includes many widely used regression models, for example, the multiple linear regression model and the generalized linear regression model, which is expressed as

$$y = f(\boldsymbol{\beta}_0'\mathbf{x}, \epsilon),$$

where y is a response variable, f is an unknown but globally monotone function, $\boldsymbol{\beta}_0$ is a K-dimensional unknown coefficient vector that is the true EDR unit vector, $\mathbf{x}$ is a covariate vector, and ϵ is a random variable that is independent of $\mathbf{x}$. Duan and Li (1991) and Li (1991) suggested a sliced inverse regression for the single index model and for the model having K-dimensional EDR space, respectively, but the allowable covariate distributions are restricted and the method requires a matrix operation of size $K \times K$, which is computationally intensive when the number of covariates is large. Some estimators of the EDR vector have been proposed using smoothing procedure (e.g., Härdle et al., 1993). More recently, Xia et al. (2002) developed an alternative methodology to find the EDR space, but it requires nonlinear optimization with a smoothing procedure, which also involves intensive computations for the case with many covariates.

In this article, we propose a simple method to explore the EDR vector, by dividing the response variables into two slices and calculating the two corresponding means of the covariates and their difference vector. Cook and Lee (1999) also discussed the statistic for the case when the original response is binary, of which the natural slice number is two. In Sec. 2, the method is explained, then the theoretical properties of the suggested estimator and its interpretation are presented in Sec. 3. Its numerical behavior is investigated in Sec. 4. In Sec. 5, we discuss some problems faced in real situations.

2. METHOD

Given a set of data $(y_i, \mathbf{x}_i)$, $i = 1, \ldots, N$, the process flow of our method can be described as follows.

Step 0. Standardize each covariate $x^{(j)}$ by its sample mean $\hat{\mu}^{(j)}$ and variance $(\hat{\sigma}^{(j)})^2$, i.e., let $z^{(j)} = \frac{x^{(j)} - \hat{\mu}^{(j)}}{\hat{\sigma}^{(j)}}$ for all $x^{(j)}$ in $\mathbf{x} = (x^{(1)}, \ldots, x^{(K)})'$.

Step 1. Divide the range of y into two slices, $I_H = \{i \mid y_i \geq t, i \in I\}$ and $I_L = \{i \mid y_i < t, i \in I\}$, where t is the median of y and $I = \{1, \ldots, N\}$.

Step 2. Within each slice, compute the sample mean of $\mathbf{z}_i's$, denote by $\hat{\mathbf{m}}_H$ and $\hat{\mathbf{m}}_L$ for I_H and I_L, respectively, so $\hat{\mathbf{m}}_H = \frac{1}{N_H}\sum_{i \in I_H} \mathbf{z}_i$ and $\hat{\mathbf{m}}_L = \frac{1}{N_L}\sum_{i \in I_L} \mathbf{z}_i$, here N_H is the number of indices in I_H, $N_L = N - N_H$, and $\mathbf{z} = (z^{(1)}, \ldots, z^{(K)})'$.

Step 3. Compute the difference vector of the two sample mean vectors, $\hat{\boldsymbol{\eta}} = \frac{1}{2}(\hat{\mathbf{m}}_H - \hat{\mathbf{m}}_L)$.

Step 4. If the sample correlation matrix $\hat{\Omega}$ is defined and there exists the inverse matrix $\hat{\Omega}^{-1}$, compute $\hat{\boldsymbol{\eta}}_N = \hat{\Omega}^{-1}\hat{\boldsymbol{\beta}}$.

Thus, the normalized unit vectors $\widehat{\Sigma}^{-\frac{1}{2}}\hat{\boldsymbol{\eta}}$ and $\widehat{\Sigma}^{-\frac{1}{2}}\hat{\boldsymbol{\eta}}_N$, i.e., $\widehat{\Sigma}^{-\frac{1}{2}}\hat{\boldsymbol{\eta}}/|\widehat{\Sigma}^{-\frac{1}{2}}\hat{\boldsymbol{\eta}}|$ and $\widehat{\Sigma}^{-\frac{1}{2}}\hat{\boldsymbol{\eta}}_N/|\widehat{\Sigma}^{-\frac{1}{2}}\hat{\boldsymbol{\eta}}_N|$ are estimators of the true EDR vector $\boldsymbol{\beta}_0$, where $\widehat{\Sigma} = diag\{(\hat{\sigma}^{(1)})^2, \ldots, (\hat{\sigma}^{(K)})^2\}$ and $\widehat{\Sigma}^{-\frac{1}{2}} = diag\{\frac{1}{\hat{\sigma}^{(1)}}, \ldots, \frac{1}{\hat{\sigma}^{(K)}}\}$.

3. PROPERTIES

Although the original sliced inverse regression uses a principle component analysis on the sample mean vectors, our method uses only the difference vector. This simplified procedure allows some easy interpretations. In this section, we investigate the theoretical properties of the proposed estimator $\hat{\boldsymbol{\eta}}$ and it is not consistent for the true EDR vector as discussed in this section. Therefore, if it is available, we should use the consistent estimator $\hat{\boldsymbol{\eta}}_N = \widehat{\Omega}^{-1}\hat{\boldsymbol{\eta}}$, whose theoretical properties are the same with that of $\hat{\boldsymbol{\eta}}$ having independent covariates.

Because we use the median of y_is for t, then it yields that $N_H = N_L = N/2$ and the true model is rewritten as

$$y = f(\boldsymbol{\eta}_0'\mathbf{z}, \epsilon),$$

with the standardized covariate vector $\mathbf{z}$, each element of which has mean zero and variance one where $\boldsymbol{\eta}_0 \propto \Sigma\boldsymbol{\beta}_0$, $\Sigma = diag\{(\sigma^{(1)})^2, \cdots, (\sigma^{(K)})^2\}$ and $\sigma^{(j)} = Var(x^{(j)})^{\frac{1}{2}}$. Let δ_i be the random variable that takes 1 if $y_i \geq t$ or -1 if $y_i < t$ with the conditional probability function

$$p_i = Pr(\delta_i = 1 \,|\, \mathbf{z}_1, \ldots, \mathbf{z}_N) = Pr(y_i \geq t \,|\, \mathbf{z}_1, \ldots, \mathbf{z}_N).$$

Then the two sample mean vectors and their difference vector can be given by

$$\hat{\mathbf{m}}_H = \frac{1}{N_H}\sum_{i=1}^{N}\frac{1+\delta_i}{2}\mathbf{z}_i, \quad \hat{\mathbf{m}}_L = \frac{1}{N_L}\sum_{i=1}^{N}\frac{1-\delta_i}{2}\mathbf{z}_i, \quad \text{and} \quad \hat{\boldsymbol{\eta}} = \frac{1}{N}\sum_{i=1}^{N}\delta_i\mathbf{z}_i.$$

Proposition 1. *Let* $\boldsymbol{\eta}_* = E(\hat{\boldsymbol{\eta}}) = (\eta_*^{(1)}, \ldots, \eta_*^{(K)})'$. *Then it holds that* $\eta_*^{(j)} = Cor(\delta, z^{(j)})$.

According to Proposition 1, $\hat{\boldsymbol{\eta}}$ is regarded as a robust correlation coefficient between the order statistics of the dichotomized response $\{-1, 1\}$ and each covariate, which catches the global monotone trend and might not be sensitive to outliers and complex variations in the response.

Because δ_i depends on $t = t(y_1, \ldots, y_N)$, $\delta_i(y_i)$, and $\mathbf{z}_j$ for $(i \neq j)$ are not independent. However, we have the following lemma 1 because the sample median of y_is converges to the true value as the sample size gets larger.

Lemma 1. *Asymptotically,* $Pr(y_i \geq t \,|\, \mathbf{z}_1, \ldots, \mathbf{z}_N) = Pr(y_i \geq t \,|\, \mathbf{z}_i)$ *for all* $i = 1, \ldots, N$.

Proposition 2. *Let* Ω *is the variance matrix of the covariate vector* $\mathbf{z}$. *Then the variance matrix of* $\hat{\boldsymbol{\eta}}$ *is asymptotically expressed as*

$$Var(\hat{\boldsymbol{\eta}}) = \frac{1}{N}(\Omega - \boldsymbol{\eta}_*\boldsymbol{\eta}_*').$$

Next, we discuss the asymptotic distribution of the estimator. It shows that the confidence interval width of $\hat{\eta}^{(j)}$ depends on its coefficient size.

Proposition 3. *Let* $\hat{\boldsymbol{\eta}} = (\hat{\eta}^{(1)}, \ldots, \hat{\eta}^{(K)})'$. *Then asymptotically holds*

$$N^{\frac{1}{2}}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_*) \sim N(0_K, \Omega - \boldsymbol{\eta}_*\boldsymbol{\eta}_*').$$

Therefore,

$$N^{\frac{1}{2}}(\hat{\eta}^{(j)} - \eta_*^{(j)}) \sim N\{0, 1 - (\eta_*^{(j)})^2\}.$$

Proposition 4. *Consider the following hypothesis test:*

$$H_0 : \eta_*^{(j_1)} = \cdots = \eta_*^{(j_Q)} = 0 \quad (Q \leq K) \text{ vs. } H_1 : \text{not } H_0.$$

Under the null hypothesis, the test statistic $T = N\sum_{l=1}^{Q}\{\hat{\eta}^{(j_l)}\}^2$ *has asymptotically a* χ^2 *distribution with* Q *degrees of freedom. In particular,* $T = N\hat{\boldsymbol{\eta}}'\hat{\boldsymbol{\eta}} \sim \chi^2(K)$ *for* $Q = K$.

Because our method is simple and does not require intensive calculations, the bootstrap method by Efron (1979) should be suitable for constructing confidence intervals and performing the test.

The accuracy of the suggested estimator $\hat{\boldsymbol{\eta}}$ for the true EDR vector $\boldsymbol{\eta}_0$ can be evaluated by the canonical correlation coefficient between $\hat{\boldsymbol{\eta}}'\mathbf{z}$ and $\boldsymbol{\eta}_0'\mathbf{z}$ (see Hooper, 1959 Li, 1991), which is given by

$$Cor(\hat{\boldsymbol{\eta}}'\mathbf{z}, \boldsymbol{\eta}_0'\mathbf{z}) = \frac{\hat{\boldsymbol{\eta}}'\Omega\boldsymbol{\eta}_0}{(\hat{\boldsymbol{\eta}}'\Omega\hat{\boldsymbol{\eta}} \cdot \boldsymbol{\eta}_0'\Omega\boldsymbol{\eta}_0)^{\frac{1}{2}}}.$$

Thus, the true EDR vector $\boldsymbol{\eta}_0$ is characterized by the vector of correlation coefficients $Cor(\delta, z^{(j)})$ or $\boldsymbol{\eta}_*$ given in Proposition 1.

Proposition 5. *Consider the following correlation coefficient:*

$$Cor(\boldsymbol{\eta}'\mathbf{z}, \delta) = \frac{E(\boldsymbol{\eta}'\mathbf{z} \cdot \delta)}{(\boldsymbol{\eta}'\Omega\boldsymbol{\eta})^{\frac{1}{2}}}.$$

Because the mean function f *has only one EDR vector* $\boldsymbol{\eta}_0$, *it maximizes the correlation coefficient,*

$$\boldsymbol{\eta}_0 = \arg\max_{\boldsymbol{\eta}'\boldsymbol{\eta}=1} Cor(\boldsymbol{\eta}'\mathbf{z}, \delta).$$

Then, $\boldsymbol{\eta}_0 \propto \Omega^{-1}\boldsymbol{\eta}_*$ *and* $Cor(\boldsymbol{\eta}_0'\mathbf{z}, \delta) = (\boldsymbol{\eta}_*'\Omega^{-1}\boldsymbol{\eta}_*)^{\frac{1}{2}}$. *The discrepancy between* $\boldsymbol{\eta}_*'\mathbf{z}$ *and* $\boldsymbol{\eta}_0'\mathbf{z}$ *is evaluated by*

$$Cor(\boldsymbol{\eta}_*'\mathbf{z}, \boldsymbol{\eta}_0'\mathbf{z}) = \frac{\boldsymbol{\eta}_*'\boldsymbol{\eta}_*}{(\boldsymbol{\eta}_*'\Omega\boldsymbol{\eta}_* \cdot \boldsymbol{\eta}_*'\Omega^{-1}\boldsymbol{\eta}_*)^{\frac{1}{2}}} = \frac{Cor(\boldsymbol{\eta}_*'\mathbf{z}, \delta)}{Cor(\boldsymbol{\eta}_0'\mathbf{z}, \delta)}.$$

Let $\lambda_1 \leq \cdots \leq \lambda_K$ *be the eigenvalues of* Ω, *which satisfy inequality,*

$$\frac{\lambda_1}{\lambda_K} \leq Cor(\boldsymbol{\eta}_*'\mathbf{z}, \boldsymbol{\eta}_0'\mathbf{z})^2 \leq \frac{\lambda_K}{\lambda_1}.$$

When $\Omega = I_K$, $\lambda_1 = \cdots = \lambda_K$ *and* $Cor(\boldsymbol{\eta}_*'\mathbf{z}, \boldsymbol{\eta}_0'\mathbf{z}) = 1$.

The resultant estimator given in Proposition 6 is equivalent to the least-square estimator on the multiple linear regression model with the binary response δ. Therefore, the binary response instead of raw data are sufficient to find the true EDR vector for the single index model with a globally monotone mean function.

Proposition 6. *Consider the following sample correlation coefficient,*

$$Cor_N(\boldsymbol{\eta}'\mathbf{z}, \delta) = \frac{\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\eta}'\mathbf{z}_i \cdot \delta_i}{(\boldsymbol{\eta}'\widehat{\Omega}\boldsymbol{\eta})^{\frac{1}{2}}},$$

where $\widehat{\Omega}$ *is the sample correlation matrix of covariates. Let* $\hat{\boldsymbol{\eta}}_N$ *be the estimator that maximizes the quantity given by*

$$\hat{\boldsymbol{\eta}}_N = \arg\max_{\boldsymbol{\eta}'\boldsymbol{\eta}=1} Cor_N(\boldsymbol{\eta}'\mathbf{z}, \delta).$$

Then $\hat{\boldsymbol{\eta}}_N \propto \widehat{\Omega}^{-1}\hat{\boldsymbol{\eta}}$. *Further, the estimator is asymptotically consistent for the true EDR vector, in that* $Cor(\hat{\boldsymbol{\eta}}_N'\mathbf{z}, \boldsymbol{\eta}_0'\mathbf{z}) \xrightarrow{P} 1$.

Proposition 7. *Let* $\hat{\boldsymbol{\eta}}_N = \widehat{\Omega}^{-1}\hat{\boldsymbol{\eta}}$ *without loss of generality and* $\hat{\boldsymbol{\eta}}_0 = \frac{1}{|\hat{\boldsymbol{\eta}}_N|}\hat{\boldsymbol{\eta}}_N$. *Then it asymptotically holds that*

$$N^{\frac{1}{2}}(\hat{\boldsymbol{\eta}}_0 - \boldsymbol{\eta}_0) \sim N\left\{0_K, \frac{1}{(\boldsymbol{\eta}_*'\Omega^{-2}\boldsymbol{\eta}_*)^{\frac{1}{2}}}\Omega^{-1} - \boldsymbol{\eta}_0\boldsymbol{\eta}_0'\right\},$$

here $\boldsymbol{\eta}_0$ *is the true EDR unit vector.*

Proposition 8. *The discrepancy between the suggested estimator and the true EDR vector is expressed by*

$$Cor(\hat{\boldsymbol{\eta}}'\mathbf{z}, \boldsymbol{\eta}_0'\mathbf{z}) = \frac{Cor(\hat{\boldsymbol{\eta}}'\mathbf{z}, \delta)}{Cor(\boldsymbol{\eta}_0'\mathbf{z}, \delta)} \cdot \frac{\hat{\boldsymbol{\eta}}'\boldsymbol{\eta}_*}{\hat{\boldsymbol{\eta}}'\hat{\boldsymbol{\eta}}} \leq 1.$$

Although the quantity given in Proposition 8 itself is not available, we can estimate $Cor(\hat{\boldsymbol{\eta}}'\mathbf{z}, \delta)$ and, if it is close to one, the criterion is also close to one.

4. SIMULATIONS

We evaluate our method numerically by using the logistic function,

$$y = \frac{1}{1 + \exp(-5 \cdot \boldsymbol{\eta}_0'\mathbf{z})} + \epsilon, \quad \epsilon \sim N(0, 0.05^2),$$

where $\boldsymbol{\eta}_0 = 5^{-\frac{1}{2}}(1, \ldots, 1, 0)'$, $\mathbf{z} = (z^{(1)}, \ldots, z^{(6)})' \sim N\{0, \Omega(\rho)\}$, and

$$\Omega(\rho) = \begin{pmatrix} 1 & \rho & 0 & 0 & 0 & 0 \\ \rho & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\rho & 0 & 0 \\ 0 & 0 & -\rho & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The dimension K is six, but $z^{(6)}$ is a dummy variable that does not affect the response and the $z^{(j)}$ and ϵ are independent. Sample sizes of $N = 50$, 100, and 500 are used, and the correlation coefficients $\rho = 0.0$ and 0.8 between the covariables are evaluated.

An example data set from the previous model at $N = 50$ and $\rho = 0.8$ is shown in Figs. 1 to 4. The true EDR direction $\boldsymbol{\eta}_0'\mathbf{z}$ versus the response is shown in Fig. 1. Figure 2 shows the scatter plot of $z^{(1)}$ and $z^{(2)}$ whose correlation is 0.8, and the EDR vector has the same direction as the correlation. Therefore, the sets of covariate vectors $\{\mathbf{z}_i \mid i \in I_H\}$ and $\{\mathbf{z}_i \mid i \in I_L\}$ are clearly separated and the sizes of $\hat{\boldsymbol{\eta}}^{(1)}$ and $\hat{\boldsymbol{\eta}}^{(2)}$ are larger than the other $\hat{\boldsymbol{\eta}}^{(j)}$ (Table 1). The EDR direction estimated by $\hat{\boldsymbol{\eta}}'\mathbf{z}$ versus the response is shown in Fig. 3. Although $\hat{\boldsymbol{\eta}}$ does not have the property of consistency for $\boldsymbol{\eta}_0$, we can see a similar shape to that with the true vector $\boldsymbol{\eta}_0$ (Fig. 1) in terms of the direction. Using the estimated correlation matrix $\hat{\Omega}$, the proposed estimator is improved by $\hat{\boldsymbol{\eta}}_N = \hat{\Omega}^{-1}\hat{\boldsymbol{\eta}}$ and Fig. 4 shows its shape vs. the response, which is quite close to that of the true EDR direction.

Table 1 shows the accuracy of the asymptotic variance in Propositions 2 and 3. The approximations are very close to the true values, although the sample size $N = 50$ is not so large. Table 2 gives the average

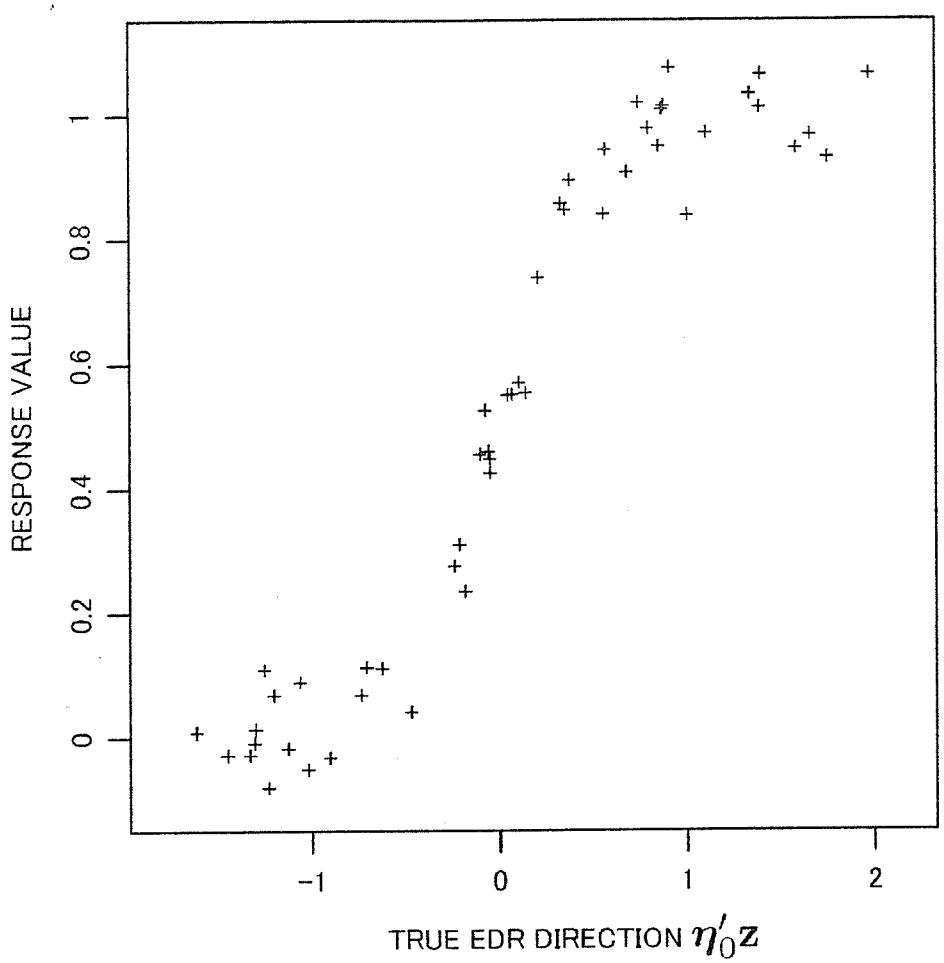

Figure 1. Scatter plot of the true EDR direction $\eta_0'\mathbf{z}$ and the response y for $N = 50$ and $\rho = 0.8$.

of the canonical correlation coefficient $Cor(\hat{\eta}'\mathbf{z}, \eta_0'\mathbf{z})$ and the correlation coefficient between the projected value $\hat{\eta}'\mathbf{z}$ and the binary response δ, $Cor(\hat{\eta}'\mathbf{z}, \delta)$. Note that the entries under $\rho = 0.0$ demonstrate the good performance of $\hat{\eta}_N$. It can be seen that the canonical correlation coefficients approach 1 and its standard deviation is small, although it does not converge when $\rho = 0.8$. The expectation of estimated $Cor(\hat{\eta}'\mathbf{z}, \delta)$ does not change as the sample size becomes larger for either case $\rho = 0.0$ or 0.8, which implies that we do not need a large sample size to catch the binary response. Therefore, the canonical correlation coefficient,

$$Cor(\hat{\eta}'\mathbf{z}, \eta_0'\mathbf{z}) \simeq \frac{Cor(\hat{\eta}'\mathbf{z}, \delta)}{Cor(\hat{\eta}_N'\mathbf{z}, \delta)},$$

has expectation around $0.76/0.80 = 0.95$, and the accuracy of $\hat{\eta}$ is reasonably good despite the simple calculations.