treatment, compared to only 30–40% in control cells (data not shown). Thus, the overexpression of topors enhances the apoptotic response of U2-OS cells to DNA damage.

This finding suggests that upregulation of endogenous topors might be involved in mediating p53-dependent cellular responses to stress stimuli. We thus examined the expression of endogenous *topors* after cisplatin treatment in U2-OS cells. *topors* expression was obviously upregulated 24 and 48 h after cisplatin treatment consistent with the expression of p53 protein (Figure 6a). The analysis was extended to primary MEFs and C-20 cells, a mouse cell line derived from colon carcinoma. In these cells before treatment, *topors* was constitutively expressed while p53 was barely detectable. The apoptotic response of MEFs and C-20 cells to 20 μM cisplatin was more prominent than that of U2-OS cells (Figure 6b and c; compare to Figure 4d). The expression of the *topors* gene and p53 protein was increased in MEFs and C-20 cells by cisplatin treatment as well (Figure 6b and c). Since topors is a topoisomerase I-binding protein, we further investigated whether the expression of *topors* is also induced after treatment by CPT, a potent antineoplastic inhibitor of topoisomerase I (Haluska *et al.*, 1999). As shown in Figure 6d, when U2-OS cells and NIH 3T3 cells were treated with CPT at the concentration of 10 μM, the expression of *topors* increased more than twofold compared to cells without treatment. However, it is still possible that the upregulation of *topors* mRNA might be a consequence of p53 induced during DNA damage-induced apoptosis. To exclude this possibility, we examined *topors* expression after overexpression of p53 by transient transfection. The overexpression of p53 did not significantly affect *topors* transcript levels (Figure 6e). Taken together, the induction of *topors* by the genotoxic reagents provides a mechanism to facilitate the p53-mediated DNA damage-dependent apoptosis in tumor and primary cells.

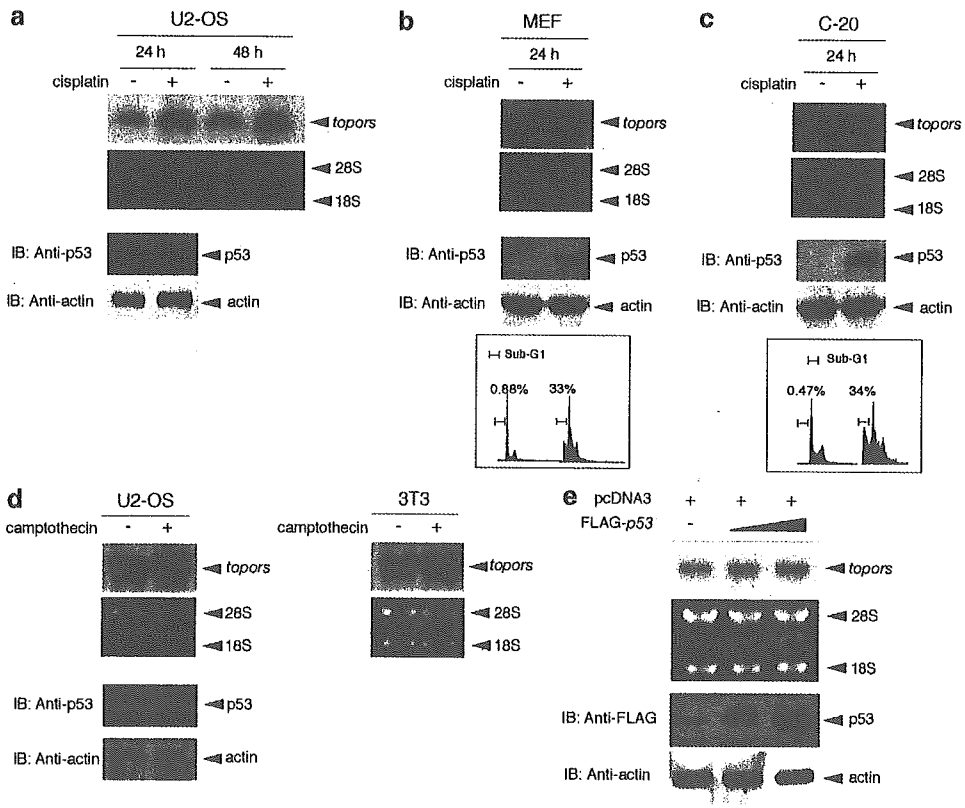


Figure 6 Induction of endogenous *topors* expression by DNA damage in human and mouse cell lines. (**a**) The expressions of *topors* and p53 are significantly induced in U2-OS cells in response to cisplatin treatment as revealed by RNA blot and immunoblotting analysis, respectively. (upper panels) *topors* transcripts detected as 4.0-kb bands (top) and ethidium bromide staining of 28S and 18S ribosomal as a loading control are shown (bottom). (lower panels) The expressions of p53 (top) and actin (lower) as a loading control are shown. (**b, c**) Induction of *topors* (upper panels), p53 expression (middle panels) and FACS analysis-based estimation of apoptosis (lower panels) by cisplatin treatment are also seen in primary MEFs (**b**) and murine colon carcinoma cells (C-20) (**c**). Percentages of cells in the sub-G1 fraction are shown in the lower panels. (**d**) The expressions of *topors* and p53 are significantly induced in U2-OS cells and NIH/3T3 cells in response to 12-h CPT treatment as revealed by Northern blot and immunoblotting, respectively. (upper panels) *topors* transcripts (top) and ethidium bromide staining of 28S and 18S ribosomal as a loading control are shown (bottom). (lower panel) The expressions of p53 (top) and actin (lower) as a loading control in U2-OS cells are shown. (**e**) The expression of *topors* is not affected by the transient overexpression of p53 in U2-OS cells. Increasing amounts of FLAG-*p53* documented by immunoblotting analysis (lower panels) do not significantly affect the expression of endogenous *topors* (upper panels). Ethidium bromide staining of 28S and 18S ribosomal RNA is shown as a loading control

Discussion

In the present study, we identify the mouse counterpart of human topors and document its involvement in p53-regulated cell growth control. Overexpressed topors associates with p53, activates p53-dependent transcription via the stabilization of p53 and, consequently, induces either cell cycle arrest or apoptosis in a dose-dependent manner. Therefore, our observations strongly implicate the topors protein as a novel coactivator of the p53 tumor suppressor protein.

We further show that expression of endogenous *topors*, as well as the p53 protein, is induced by DNA damage in tumor and primary cells. The present observations indicate that topors mediates p53-dependent cellular responses to some forms of DNA damage and, thus, could act as a tumor suppressor that activates p53. Interestingly, the human *topors* gene maps to chromosome 9p21 where the tumor suppressor genes associated with 86% of small cell lung carcinomas are suggested to reside and the above observations provide further evidence for the *topors* gene as a candidate tumor suppressor gene mapped to this region. Our preliminary studies have revealed that *topors* mRNA is expressed in all small cell lung cancer cell lines at levels similar to other tumor cell lines (L Lin and A Hata, unpublished data). Most recently, Oyanagi *et al.* (2004) revealed that the expression of *LUN* (*topors*) gene is downregulated in the development and metastases of lung cancer, suggesting that *LUN* might play important roles in inhibiting the oncogenesis of non-small cell lung cancer. The expression of topors protein in normal lung tissues and small cell lung cancers will require further future study.

The data clearly show that the overexpression of topors may cause either cell cycle arrest or apoptosis, and suggest that which occurs depends both on the topors level and the cellular context, although other variables require further exploration. There is a positive correlation between the amount of p53 protein and overexpressed topors in U2-OS cells in both transient and stable transfectants. Since it has repeatedly been reported that low levels of p53 expression are antiapoptotic while high levels promote apoptosis, high levels of topors could facilitate apoptosis due to the high level of p53 expression (Chen *et al.*, 1996; Lassus *et al.*, 1996). This phenomenon is further supported by another observation made in this study. We identified only two stable transfectants overexpressing myc-topors from 200 G418-resistant colonies and both exhibited a lower level of topors expression than that seen in the transiently transfected cells. This implies that only U2-OS cells overexpressing permissive amounts of exogenous topors might be allowed to survive by escaping apoptotic outbursts and instead exhibiting G1 cell cycle arrest, which is correlated with the induction of the downstream gene of p53, $p21^{Waf1}$. Interestingly, slight but significant apoptotic outbursts coincide with G1 arrest in *topors* sta-1, but not sta-2, cells. Slightly higher expression levels of myc-topors and, consequently, p53 in *topors* sta-1 cells than in sta-2 cells may be the cause for this phenotypical difference.

Our findings demonstrate that the expression level of topors protein may be a rate-limiting factor in the regulation of p53 activity. While this manuscript was in preparation, Saleem *et al.* (2004) also reported that transient expression of h-topors showed antiproliferative activity and was associated with G0/G1 cell cycle arrest in HeLa cells. Our results additionally show that overexpression of topors could activate the expression of p53 and induce either cell cycle arrest or apoptotic response.

In this study, we show that endogenous topors is significantly upregulated by the genotoxic reagents at the level of transcription in tumor and primary cells. A previous study has suggested that post-translational regulation of topors by DNA damage may also be involved (Rasheed *et al.*, 2002). In that study, overexpressed GFP-topors fusion protein immediately relocalized from PML nuclear bodies to the nucleoplasm in cells exposed to CPT or DRB. Since DRB exposure has also been shown to induce the accumulation of diffuse p53 in the nucleoplasm, it is possible that the relocalization of topors into the nucleoplasm might allow its association with and subsequent stabilization of p53 (Klibanov *et al.*, 2001).

The stability of the tumor suppressor p53 is crucial for its function to induce cell cycle arrest and/or apoptosis as a consequence of its ability to bind to specific DNA sequences and activate the transcription of adjacent genes (Vogelstein *et al.*, 2000). Overexpression of topors increased the stability of p53 and correspondingly enhanced the p53-dependent transcriptional activities. topors contains a RING finger motif, a domain that has been implicated in protein–DNA and protein–protein interactions, E2-dependent ubiquitination and SUMO conjugation (Kahyo *et al.*, 2001; Matthews and Sunde, 2002). Indeed, topors proteins have been recently shown to possess evolutionarily conserved E3 ubiquitin ligase activity. A *Drosophila* topors was shown to mediate Hairy polyubiquitination, resulting in Hairy, but not Dmp53 or topoisomerase I (dTopoI), degradation (Secombe and Parkhurst, 2004). Human GFP-topors fusion protein was shown to function as an E3 ligase for p53, albeit to a lesser degree than MDM2 (Rajendra *et al.*, 2004). Therefore, topors could affect the stability of p53 via covalent modification of p53 by ubiquitin isopeptide. This observation, however, is not necessarily consistent with our and others' previous observations suggesting topors as a likely tumor suppressor protein since polyubiquitination of p53 mediated by GFP-topors induces proteasome-dependent downregulation of p53 in U2-OS cells (Chu *et al.*, 2001; Rajendra *et al.*, 2004; Saleem *et al.*, 2004). This implies that regulation of p53 by topors may involve not only ubiquitination but also some other molecular mechanism.

Recently, the activity of p53 has been reported to be positively regulated by SUMO conjugation, which is mediated, at least in part, by another RING finger protein, PIAS (Gostissa *et al.*, 1999; Rodriguez *et al.*, 1999; Kahyo *et al.*, 2001). Sequence similarity between

two regions of topors and human PIASx protein, the RING domains and putative SUMO-1 interaction motifs suggests the possibility that topors may function as a SUMO E3 ligase (Weger *et al.*, 2003). Indeed, h-topors and several components of SUMO-conjugating systems have been isolated together as interacting proteins with Mx1, an interferon-inducible GTPase that is associated with PML nuclear bodies (Engelhardt *et al.*, 2001), which include p53, pRb, DAXX and CBP (Mann and Miller, 2004). Thus, topors may contribute to p53 stabilization or activation via SUMO conjugation processes.

The molecular mechanisms underlying the stabilization and/or activation of p53 by topors may also involve protein–protein interactions. AAV-2 Rep78/68 has been shown to possess a tumor suppressor property and to bind physically to p53 to prevent the adenovirus-mediated degradation of p53 (Batchu *et al.*, 1999). Since topors is capable of interacting not only with p53, but also with Rep78/68, topors may cooperate with Rep78/68 to stabilize p53. It is possible that endogenous cellular regulators of p53 may function in ways analogous to these viral products and cooperate with topors to stabilize p53. These possibilities are not necessarily mutually exclusive. It will be essential to address experimentally the possible catalytic activities of topors.

Materials and methods

Isolation and preparation of an expression vector for a full-length cDNA of topors

Yeast two-hybrid screening was performed to isolate the cDNAs for Mph2-binding proteins from a cDNA library derived from E11.0 mouse embryo as described previously, and 177 independent clones were found to be positive for β-galactosidase activities as examined by filter assay (Vojtek *et al.*, 1993; Kingma and Osheroff, 1998; Yamaki *et al.*, 2002). The nucleotide sequences of the positive cDNA clones were determined. Two clones contained 1.2-kb cDNA fragments that were highly homologous to *h-topors* (Haluska *et al.*, 1999). To isolate a full-length *topors* cDNA, a radiolabeled 1.2-kb cDNA fragment was used as a probe to screen the mouse brain cDNA library. After three rounds of hybridization, six positive clones were obtained and their nucleotide sequences were determined. The longest cDNA, containing a 3.5-kb insert, turned out to encode from the RING finger motif to the stop codon, but lacked a putative *topors* initiation codon. By searching public databases of mouse ESTs, we found a cDNA clone (MNCb-6014) that overlapped with the 3.5-kb cDNA clone. MNCb-6014, which was kindly provided by Dr Katsuyuki Hashimoto (Division of Genetic Resources, National Institute of Infectious Diseases, Japan), contained a *topors* initiation codon, but instead lacked the carboxyl-terminal region. To generate an expression vector encoding a full-length *topors*, the 5′-region of MNCb-6014 was amplified with primers 5′-CGTCGACAAGCTTATGGGGTCGCAGC CGCC-3′ (the *Sal*I/*Hind*III restriction sites are underlined) and 5′-CAGAAGACAGTTCAACAAGTTCCGGGGTCC-3′ using the MNCb-6014 cDNA clone as a template. The PCR product was digested with *Sal*I and *Sph*I and subcloned into the identical restriction sites of the above-mentioned plasmid containing the 3.5-kb insert to produce a full-length cDNA for *topors* (pBluscript-M-*topors*). The pBluscript-M-*topors* was digested with *Hind*III and *Not*I and introduced into the identical restriction sites of pcDNA3-myc (Invitrogen, Carlsbad, CA, USA) to generate an expression vector for myc-tagged topors (pcDNA3-myc-*topors*).

In vitro *transcription/translation of the topors gene product*

The topors protein was generated from the pcDNA3-myc-*topors* vector using the T7-TNT Quick-Coupled Transcription–Translation system (Promega, Madison, WI, USA) in the presence of [^{35}S]methionine (Amersham Biosciences Inc., Piscataway, NJ, USA) according to the manufacturer's instructions. The quality of the synthesized protein was verified by electrophoresis though an 8% SDS–polyacrylamide gel and autoradiography.

Cell culture and transfection

Human osteosarcoma cell lines U2-OS and SAOS-2, COS-7, NIH/3T3 and primary MEFs were maintained in DMEM supplemented with 10% FBS and antibiotics. Human large cell lung carcinoma H1299 cells and human small cell lung cancer SBC3 cells were grown in RPMI-1640 medium supplemented with 10% FBS and antibiotics. All cells were cultured at 37°C in a water-saturated atmosphere of 5% CO_2 in air. COS-7 and U2-OS cells contain wild-type p53, while SAOS-2 and H1299 cells are deficient in p53 expression. For transfection, U2-OS cells were transfected by either the calcium phosphate co-precipitation method or Lipofectamine 2000 transfection reagent (Invitrogen, Grand Island, NY, USA) in accordance with the manufacturer's specifications. COS-7 cells were transfected with the FuGENE6 transfection reagent (Roche Molecular Biochemicals, Indianapolis, IN, USA). H1299 and SAOS-2 cells were transfected with the LipofectAMINE Plus transfection kit (Invitrogen, Grand Island, NY, USA) according to the manufacturer's protocol. To obtain stable transfectants, either pcDNA3 or pcDNA3-myc-*topors* was introduced into the exponentially growing U2-OS cells by the calcium phosphate co-precipitation method. At 48 h post-transfection, the cells were transferred to fresh medium containing G418 at a final concentration of 800 μg/ml. At 1 week after selection, drug-resistant colonies were isolated and the expression of myc-topors was examined by immunofluorescence staining as described below. As a control, a stable cell line of the pcDNA3 vector was established.

Immunoprecipitation and immunoblotting

Transfected cells were washed in ice-cold phosphate-buffered saline (PBS), lysed in lysis buffer (25 mM Tris-Cl, pH 7.5, 137 mM NaCl, 2.7 mM KCl, 1% Triton X-100 and 1 mM PMSF) and the extracts were sonicated briefly and centrifuged at 800 *g* for 5 min to remove insoluble materials. The protein concentrations were determined by the Bradford protein assay (Bio-Rad, Hercules, CA, USA) using BSA as a standard. For immunoprecipitation, the cell lysates prepared from COS-7 cells transfected with pcDNA3-myc-*topors* were precleared using protein G-Sepharose beads at 4°C for 30 min under gentle rotation, and then incubated with either NMS or antibodies against p53 (DO-1, Oncogene Research Products, Cambridge, MA, USA and Pab1801, Santa Cruz Biotechnology Inc., Santa Cruz, CA, USA) at 4°C for 2 h. The immune complexes were then recovered with protein G-Sepharose beads. The immunoprecipitates or supernatants were subjected to SDS–polyacrylamide gel electrophoresis and electrophoretically transferred onto Immobilon P membranes (Millipore

Corp., Bedford, MA, USA). The membranes were blocked with TBS containing 0.1% Tween 20 and 5% nonfat dry milk, probed with antibodies against c-myc epitope (562, Medical and Biological Laboratories, Nagoya, Japan), p53 (DO-1, Oncogene Research Products, Cambridge, MA, USA), $p21^{Waf1}$ (H-164, Santa Cruz), β-actin (20-33, Sigma Chemical Co., St Louis, MO, USA) and FLAG (M2, Sigma Chemical Co., St Louis, MO, USA), and then incubated with horse radish peroxidase-conjugated goat anti-rabbit or anti-mouse secondary antibody (Santa Cruz Biotechnology Inc., Santa Cruz, CA, USA). Immunoreactive bands were visualized with an ECL Western blot detection kit (Amersham Biosciences Inc., Piscataway, NJ, USA).

Immunofluorescence and confocal microscopy

Transfected cells grown on coverslips were fixed with 3.7% formaldehyde for 30 min at room temperature, permeabilized with 0.2% Triton X-100 for 5 min at room temperature and then incubated with 3% BSA in PBS for 2 h to reduce nonspecific antibody binding. Immunostaining was performed by incubating cells with a monoclonal anti-myc antibody (9E10, diluted 1 : 10) for 1 h at room temperature in a humidified chamber, followed by incubation with fluorescein isothiocyanate (FITC)-conjugated goat anti-mouse IgG (diluted 1 : 250) for 1 h at room temperature. The coverslips were washed extensively with PBS, mounted with PermaFluor (Immunon, Pittsburgh, PA, USA) and the labeled cells were examined using a confocal laser scan microscope (LSM510; Carl Zeiss Co., Ltd, Jena, Germany).

Luciferase reporter assay

SAOS-2 or H1299 cells were seeded at a density of 5×10^4 cells/well in a 12-well tissue culture dish and then cotransfected with 100 ng of p53/p73-responsive luciferase reporter constructs carrying *Bax*, $p21^{Waf1}$ or *MDM2* promoter, 10 ng of pRL-TK and 25 ng of the p53 expression plasmid in either the presence or absence of increasing amounts of pcDNA3-myc-*topors* as described previously (Watanabe *et al.*, 2002). The total amounts of DNA used in each transfection were kept constant (510 ng/transfection) using pcDNA3. Luciferase assays were performed 48 h post-transfection with a dual luciferase reporter assay system (Promega) according to the manufacturer's instructions.

Quantitative real-time RT–PCR analysis

Small cell lung cancer cell SBC3 and U2-OS cells were transfected with pcDNA3 or myc-*topors*. At 48 h after transfection, total RNA was extracted with the RNeasy Mini kit (Qiagen Inc., Valencia, CA, USA). Quantitative real-time RT–PCR was performed using the Brilliant SYBR Green QRT-PCR Master Mix Kit, 1-Step (Stratagene, La Jolla, CA, USA) and specific primers for human $p21^{Waf1}$ and human *β-actin*. Quantitative results of $p21^{Waf1}$ mRNA were normalized for the levels of *β-actin* mRNA. Specific primers for $p21^{Waf1}$ are as follows: 5′-ATGAAATTCACCCCCTTTCC-3′ (forward primer) and 5′-CCCTAGGCTGTGCTCACTTC-3′ (reverse primer).

Colony formation assay

U2-OS and H1299 cells were transfected with pcDNA3 or pcDNA3-myc-*topors* in the presence or absence of FLAG-*p53* (a kind gift from Dr Toshiharu Suzuki, University of Tokyo, Tokyo, Japan). After 48 h of culture, the cells were divided into new dishes and cultured for 2 weeks in the presence of 400 μg/ml of G418. The cell dishes were fixed, stained with Giemsa's solution (Merck KgaA, 54271, Darmstadt, Germany, Art. 1.09204) and the colonies were counted.

Analyses of cell cycle and apoptosis

After treating cells with or without cisplatin (Sigma Chemical Co., St Louis, MO, USA) at a final concentration of 20 μM for 24 h, both floating and adherent cells were collected by brief centrifugation, fixed in 70% (v/v) ethanol for more than 4 h at −20°C and stained with PI (Sigma Chemical Co., St Louis, MO, USA). To identify cells with sub-G1 DNA content, the fluorescence of the nuclei was measured by flow cytometry (FACScan, Becton Dickinson, Oxford, UK). At least 5×10^4 events were analysed with Cell Quest software. For the morphological assessment of fragmented nuclei, U2-OS cells grown on coverslips were transfected with either myc-*topors* or pEGFP-N3 (BD Biosciences, CA, USA). At 48 h post-transfection, the cells were fixed with 3.7% formaldehyde, permeabilized with 0.2% Triton X-100 and blocked with 3% BSA in PBS. myc-topors was visualized with 9E10 monoclonal antibody (diluted 1 : 10) followed by FITC-conjugated goat anti-mouse IgG (diluted 1 : 250). DNA was visualized by incubating the cells with 1 mM Hoechst 33258 dye. Cells showing apoptotic morphological changes were analysed under a Leica QFluoro confocal spectral microscope (Leica Microsystems Imaging Solutions Ltd, Cambridge, UK).

Protein half-life determination

At 24 h after culture of the pcDNA3 stable cells (V-1) and *topors* sta-2 cells, or at 24 h post-transfection of p53-deficient human H1299 cells with FLAG-*p53* in combination with either pcDNA3 or pcDNA3-myc-*topors*, cycloheximide (Sigma) was added to the cell culture medium at a final concentration of 100 μg/ml. Cells were collected at the indicated time points and whole-cell extracts were subjected to immunoblot analysis with anti-p53 antibody. Best-fit linear regression analysis of the data points was performed with Prism 4.0 software (GraphPad).

Northern blot analysis

U2-OS cells, MEFs, C-20 cells and NIH/3T3 cells were cultured in the presence of cisplatin (20 μM) for 24 or 48 h or CPT (10 μM) for 12 h. Total cellular RNA was isolated using the Isogen kit (Wako Pure Chemical Industries, Ltd, Osaka, Japan). RNA (10 μg) was denatured in formaldehyde-formamide, separated by electrophoresis in 1.5% agarose gels and transferred to Hybond™-N+ membranes (Amersham Biosciences, Tokyo, Japan). The resulting blots were individually hybridized with radiolabeled probes specific for *topors* and *p53* at 65°C for more than 10 h. The filters were washed twice with 2 × SSC (300 mM NaCl and 30 mM sodium citrate, pH 7.0)/0.1% *N*-lauroyl sarcosine at room temperature for 30 min, and then once with 1 × SSC/0.1% *N*-lauroyl sarcosine at 55°C for 30 min.

Statistical analysis

Statistical analyses were performed using the unpaired Student's *t*-test. The differences between two groups were considered to be statistically significant when $P<0.05$.

Acknowledgements
We are grateful to Dr M Vidal and Dr O Tetsu for critical reading of the manuscript, Dr O Ohara for providing the mouse cDNA library, Dr T Oda for providing the pBTM116

vector, Ms Sanae Takeda, Dr Jie Liu and Dr Tomomi Kaneko for kind assistance and Dr T Akasaka for initial instructions on yeast two-hybrid screening. This project was supported by a grant-in-aid for Scientific Research on Priority Areas and for Scientific Research (B) and Special Coordination Funds for Promoting Science and Technology from the Ministry of Education, Culture, Sports, Science and Technology of the Japanese Government to HK and AN.

References

Barak Y, Juven T, Haffner R and Oren M. (1993). *EMBO J.*, **12,** 461–468.

Batchu RB, Shammas MA, Wang JY and Munshi NC. (1999). *Cancer Res.*, **59,** 3592–3595.

Chen X, Ko LJ, Jayaraman L and Prives C. (1996). *Genes Dev.*, **10,** 2438–2451.

Chu D, Kakazu N, Gorrin-Rivas MJ, Lu HP, Kawata M, Abe T, Ueda K and Adachi Y. (2001). *J. Biol. Chem.*, **276,** 14004–14013.

el-Deiry WS, Tokino T, Velculescu VE, Levy DB, Parsons R, Trent JM, Lin D, Mercer WE, Kinzler KW and Vogelstein B. (1993). *Cell*, **75,** 817–825.

Engelhardt OG, Ullrich E, Kochs G and Haller O. (2001). *Exp. Cell Res.*, **271,** 286–295.

Gostissa M, Hengstermann A, Fogal V, Sandy P, Schwarz SE, Scheffner M and Del Sal G. (1999). *EMBO J.*, **18,** 6462–6471.

Haluska Jr P, Saleem A, Rasheed Z, Ahmed F, Su EW, Liu LF and Rubin EH. (1999). *Nucleic Acids Res.*, **27,** 2538–2544.

Harper JW, Adami GR, Wei N, Keyomarsi K and Elledge SJ. (1993). *Cell*, **75,** 805–816.

Kahyo T, Nishida T and Yasuda H. (2001). *Mol. Cell*, **8,** 713–718.

Kingma PS and Osheroff N. (1998). *Biochim. Biophys. Acta*, **1400,** 223–232.

Klibanov SA, O'Hagan HM and Ljungman M. (2001). *J. Cell Sci.*, **114,** 1867–1873.

Ko LJ and Prives C. (1996). *Genes Dev.*, **10,** 1054–1072.

Lane DP. (1992). *Nature*, **358,** 15–16.

Lassus P, Ferlin M, Piette J and Hibner U. (1996). *EMBO J.*, **15,** 4566–4573.

Levine AJ. (1997). *Cell*, **88,** 323–331.

Mann KK and Miller Jr WH. (2004). *Cancer Cell*, **5,** 307–309.

Matthews JM and Sunde M. (2002). *IUBMB Life*, **54,** 351–355.

Miyashita T and Reed JC. (1995). *Cell*, **80,** 293–299.

Oyanagi H, Takenaka K, Ishikawa S, Kawano Y, Adachi Y, Ueda K, Wada H and Tanaka F. (2004). *Lung Cancer*, **46,** 21–28.

Rajendra R, Malegaonkar D, Pungaliya P, Marshall H, Rasheed Z, Brownell J, Liu LF, Lutzker S, Saleem A and Rubin EH. (2004). *J. Biol. Chem.*, **279,** 36440–36444.

Rasheed ZA, Saleem A, Ravee Y, Pandolfi PP and Rubin EH. (2002). *Exp. Cell Res.*, **277,** 152–160.

Rodriguez MS, Desterro JM, Lain S, Midgley CA, Lane DP and Hay RT. (1999). *EMBO J.*, **18,** 6455–6461.

Saleem A, Dutta J, Malegaonkar D, Rasheed F, Rasheed Z, Rajendra R, Marshall H, Luo M, Li H and Rubin EH. (2004). *Oncogene*, **23,** 5293–5300.

Schlehofer JR. (1994). *Mutat. Res.*, **305,** 303–313.

Secombe J and Parkhurst SM. (2004). *J. Biol. Chem.*, **279,** 17126–17133.

Sionov RV and Haupt Y. (1999). *Oncogene*, **18,** 6145–6157.

Vogelstein B, Lane D and Levine AJ. (2000). *Nature*, **408,** 307–310.

Vojtek AB, Hollenberg SM and Cooper JA. (1993). *Cell*, **74,** 205–214.

Watanabe K, Ozaki T, Nakagawa T, Miyazaki K, Takahashi M, Hosoda M, Hayashi S, Todo S and Nakagawara A. (2002). *J. Biol. Chem.*, **277,** 15113–15123.

Weger S, Hammer E and Engstler M. (2003). *Exp. Cell Res.*, **290,** 13–27.

Weger S, Hammer E and Heilbronn R. (2002). *J. Gen. Virol.*, **83,** 511–516.

Yamaki M, Isono K, Takada Y, Abe K, Akasaka T, Tanzawa H and Koseki H. (2002). *Gene*, **288,** 103–110.

Zhou R, Wen H and Ao SZ. (1999). *Gene*, **235,** 93–101.

Expression profiling using a tumor-specific cDNA microarray predicts the prognosis of intermediate risk neuroblastomas

Miki Ohira,[1,8] Shigeyuki Oba,[2,8] Yohko Nakamura,[1] Eriko Isogai,[1] Setsuko Kaneko,[3] Atsuko Nakagawa,[4] Takahiro Hirata,[5] Hiroyuki Kubo,[5] Takeshi Goto,[5] Saichi Yamada,[6] Yasuko Yoshida,[6] Misa Fuchioka,[7] Shin Ishii,[2] and Akira Nakagawara[1,*]

[1]Division of Biochemistry, Chiba Cancer Center Research Institute, Chiba 260-8717, Japan
[2]Graduate School of Information Science, Nara Institute of Science and Technology, Ikoma 630-0192, Japan
[3]Department of Pediatric Surgery, University of Tsukuba School of Medicine, Tsukuba 305-8575, Japan
[4]Second Department of Pathology, Aichi Medical University, Nagakute 480-1195, Japan
[5]Hisamitsu Pharmaceutical Co. Inc., Tokyo 100-622, Japan
[6]Micro Ceramics Laboratory, R & D Center, NGK Insulators, LTD, Nagoya 467-8530, Japan
[7]Center for Molecular Biology and Cytogenetics SRL Inc., Tokyo 191-0002, Japan
[8]These authors contributed equally to this work.
*Correspondence: akiranak@chiba-cc.jp

Summary

To predict the prognosis of neuroblastoma patients and choose a better therapeutic protocol, we developed a cDNA microarray carrying 5340 genes obtained from primary neuroblastomas and examined 136 tumor samples. We made a probabilistic output statistical classifier that provided a high accuracy in prognosis prediction (89% at 5 years) and a highly reliable method to validate it. Kaplan-Meier analysis indicated that the patients in an intermediate group defined by existing markers are divided by microarray into two further groups with 5 year survivals for 36% and 89% of patients ($p < 10^{-4}$), i.e., with unfavorably and favorably predicted neuroblastomas, respectively. According to these results, we developed a gene subset chip for a clinical tool, for which our classifier exhibited 88% prediction accuracy.

Introduction

Neuroblastoma is one of the most common solid tumors in children and originates from the sympathoadrenal lineage of the neural crest (Bolande, 1974). Its clinical behaviors are heterogeneous. The tumor, when developed in infants, frequently regresses spontaneously by inducing differentiation and/or programmed cell death. When developed in children over 1 year of age, however, the tumor is often aggressive and acquires resistance to intensive chemotherapy. Although recent progress in therapeutic strategies against advanced neuroblastoma has improved patient survival, long-term outcomes still remain very poor. Furthermore, part of neuroblastomas categorized to the intermediate group (stage 3 or 4 tumors that possess a single copy of the *MYCN* gene) often recurs after complete response to initial therapy. Such differences in the final outcomes of the tumor are considered presumably attributable to differences in genetic and biological abnormalities, which are reflected in the gene and protein expression profiles of the tumor.

The prediction of cancer prognosis is one of the most urgent demands to initiate the treatment of neuroblastoma. As expected from the natural course of neuroblastoma, patient age at diagnosis (over or under 1 year of age) is an important prognostic factor (Evans et al., 1971). Disease stage is also a powerful indicator for neuroblastoma prognosis (Brodeur et al., 1993). Moreover, recent advances in basic research have discovered several molecular markers that are useful in clinical practice, including amplification of the *MYCN* oncogene (Schwab et al., 1983; Brodeur et al., 1984), DNA ploidy (Look et al., 1984; Look et al., 1991), deletion of chromosome 1p (Brodeur et al., 1988), and *TrkA* expression (Nakagawara et al., 1992; Nakagawara et al., 1993). Other indicators also include *telomerase* (Hiyama et al., 1995), *CD44* (Favrot et al., 1993), *pleiotrophin* (Nakagawara et al., 1995), *N-cadherin* (Shimono et al., 2000), *CDC10* (Nagata et al., 2000), and *Fyn* (Berwanger et al., 2002). However, the combinations thereof still frequently fail to predict patient outcome. In the post-genome sequence era, therefore, the advent of new diagnostic tools has been ex-

SIGNIFICANCE

Neuroblastoma is an enigmatic tumor with heterogeneous clinical behaviors including maturation, regression, and growth. Despite recent improvements in the cure rate of many pediatric tumors, the prognosis of advanced neuroblastoma is still poor. In addition, it is usually difficult to predict the prognosis of the intermediate risk group in advanced stages without *MYCN* amplification. Through our supervised machine learning and highly reliable statistical validation procedure with the 5 year prognosis of the patients, we established a simple, low-cost microarray system carrying top-ranked genes, which exhibited high accuracy (88%) to predict the neuroblastoma prognosis and is highly feasible as a clinical tool.

pected. Recently, the DNA microarray method, applied to comprehensively demonstrate expression profiles of primary neuroblastomas and cell lines, has already identified the following: (1) differences in gene expression between favorable and unfavorable subsets (Yamanaka et al., 2002; Berwanger et al., 2002); and (2) differences in gene expression that occur during retinoic acid-induced neuronal differentiation (Ueda, 2001). However, a study to predict neuroblastoma prognosis with a microarray using a large number of neuroblastoma samples has never been reported. We have recently isolated 5500 genes from the cDNA libraries, which were generated from primary neuroblastomas, part of which has previously been reported (Ohira et al., 2003a; Ohira et al., 2003b). In this study, to identify genes strongly associated with neuroblastoma prognosis and to apply them to make a really practical cDNA microarray for neuroblastoma diagnosis, we constructed an in-house, ink-jet-printed cDNA microarray carrying 5340 genes proper to neuroblastoma and applied it to analyze 136 samples. After selecting genes significantly related to patient prognosis, we made a mini-chip carrying 200 top-ranked genes to apply for the clinic.

There have been many attempts to predict cancer outcome using microarray. A reliable prediction for outcomes of cancer patients naturally demands its reproducibility, and it is quite important to use sound and highly reliable statistical methodologies; a complete crossvalidation analysis without introducing any information leakage and an independent test using new samples are necessary. As Ntzani and Ioannidis (2003) pointed out, however, such a careful methodology has often been ignored in most microarray studies. We here developed a supervised classification method without any information leakage as a statistic tool and demonstrated that the probabilistic output of the analysis defines the molecular signature of neuroblastoma to predict its prognosis. Although the construction of the statistical tool was based on one of the most reliable statistical tests, we also consulted a validation test for an independent experiment examining 50 samples (whose RNAs were prepared in an independent laboratory) by using the mini-chip. The high performance for the outcome prediction by the mini-chip system suggests the high feasibility of developing a clinical tool based on molecular signature.

Results

Neuroblastoma proper cDNA microarray

The whole scheme of our study is summarized in Figure 1. We first constructed a neuroblastoma proper cDNA microarray harboring the spots of 5340 genes on a slide glass by using a ceramics-based ink-jet printing system (the 5340 genes system). This in-house cDNA microarray appeared to have overcome the previous problems caused by pin-spotting, e.g., uneven quantity or shape of individual spots on an array. Ten micrograms each of the total RNA extracted from 136 frozen tissues of primary neuroblastomas were labeled with Cy3 dye. As a common reference, the mixture of the total RNA obtained from four neuroblastoma cell lines with a single copy of *MYCN* (NB69, NBLS, SK-N-AS, and SH-SY5Y) was labeled with Cy5 dye.

We first evaluated the quality of our cDNA microarray, the 5340 genes system. The log Cy3/Cy5 fluorescence ratio of each gene spot was normalized to eliminate intensity-dependent biases. Since the 5340 genes array contains 260 duplicated or multiplicated genes, the expression ratio of such a duplicated gene was represented by the average of multiple spots. Based on estimation performance for missing values (see the Supplemental Data available with this article online) and on reproduction variance of the duplicated genes, the standard deviation for the log ratio of a single gene was sufficiently small, ranging between about 0.2 and about 0.3 (Figure S1A). The scatter plots of the log Cy3/Cy5 fluorescence ratio between duplicated gene spots in the 136 experiments and those between repeated experiments also indicated high reproducibility of spotting and experiment (Figures S1B and S1C). These suggest that the production of and experiments by our cDNA microarray are highly reproducible.

Supervised classification

To develop a statistical tool that predicts the prognosis of a new patient with neuroblastoma, we introduced a supervised classification. In the development, we used 136 neuroblastomas, randomly selected tumor samples from the neuroblastoma tissue bank, consisting of 41 stage 1 tumors, 22 stage 2 tumors, 33 stage 3 tumors, 28 stage 4 tumors, and 12 stage 4s tumors. The follow-up duration ranged between 3 and 241 months (median, 56 months, mean, 57.3 months) after diagnosis. The left panel in Figure 2 compiles summary information of each sample, including survival time and important prognosis markers (see Experimental Procedures for details). Since variations in follow-up duration generated noises in the supervised classification, we used patient outcome (dead or alive) at 5 years after diagnosis as the target label to be predicted. Since the outcomes of 40 of 136 samples were unknown at 5 years after diagnosis, data for 96 remaining samples were used subsequently. When we were interested in short-term outcome prediction, the target label was set at 2 years after diagnosis, for which purpose 126 samples out of the 136 samples were used.

We constructed the weighted voting as a supervised classifier after important genes were selected according to pairwise *F* scores. To estimate the prediction accuracy for new data, we consulted leave two out (LTO) analysis, which obtains almost unbiased estimation of prediction accuracy for new data while avoiding overestimation due to information leakage (Figure S2A). Although it is known that the prediction accuracy of a supervised classifier depends on the number of genes to be used (Figure S3), the LTO procedure enables us to optimize it without introducing information leakage, by using a sample left out at the outer loop of the double-loop procedure (see Experimental Procedures). The crossvalidation accuracy for the 5 year prognosis prediction was as high as 88.5% (sensitivity of 86.7% and specificity of 89.4%) (Table 1, "Whole cases"). In the LTO analysis, we selected genes and constructed the corresponding classifier individually for the outcome prediction of each sample. The average number of the selected genes, n, was 30.7. If we applied the same procedure to the short-term (2 year) prediction, the accuracy, sensitivity, and specificity were 89.8%, 88.0%, and 90.2%, respectively (data not shown).

Construction of a probabilistic output

According to the LTO analysis, we can obtain weighted vote values and the corresponding survival rates. After approximat-

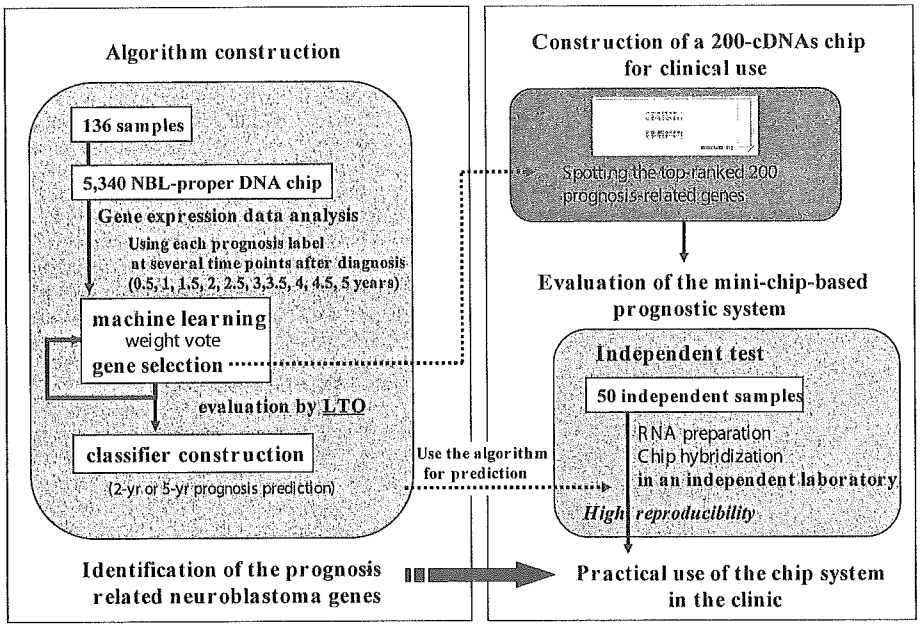


Figure 1. Schematic diagram of this study

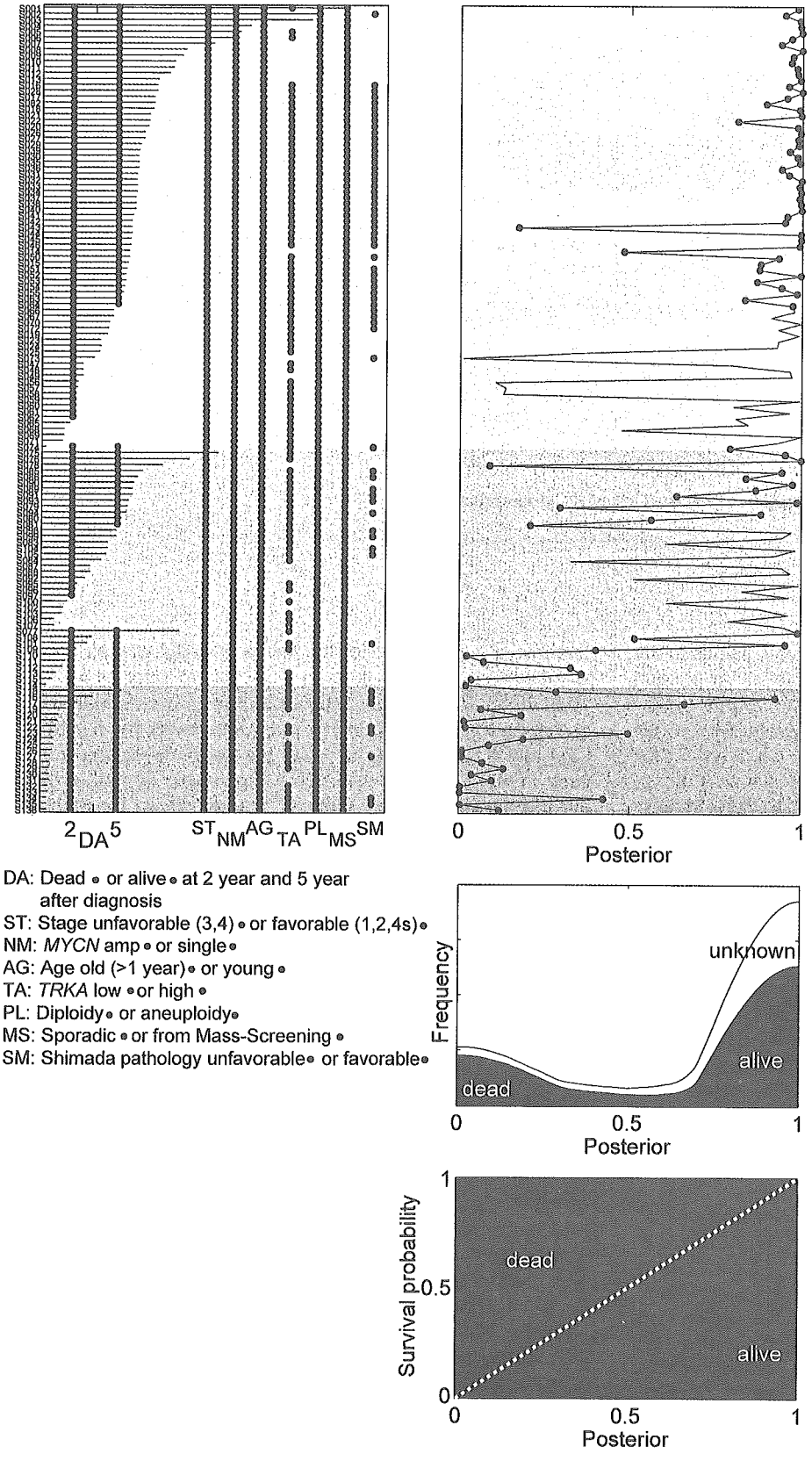
ing a nonlinear transformation from weighted vote values to the survival rates, the transformation outputs the reliability of each sample's outcome prediction as a probabilistic output, posterior probability. We suppose each posterior probability, a real number between 0 and 1, corresponds to the expected 5 year survival rate. The right upper panel of Figure 2 shows the predictions for the 136 samples as posterior probabilities. Most of the samples alive at 5 years after diagnosis (blue mark) are found to have posterior values near 1, while most of the dead samples (red mark) have those near 0. It is known that it is difficult to predict the prognosis of neuroblastoma patients of the intermediate risk group (the type II subset: stage 3 or 4, without amplification of *MYCN*), denoted by green area. The posterior values are likely to take intermediate values near 0.5; however, their binarization after being separated by threshold 0.5 shows good accordance with the actual outcome. Frequencies of posterior values for alive and dead samples are shown in the right middle panel. The rate of alive samples among the whole samples, which denotes the actual survival rate, is plotted against each posterior value in the right bottom panel in Figure 2; this panel shows the good correspondence between the posterior value and the survival rate.

Probabilistic outputs are considered to be advantageously useful as compared with conventional binary outputs when used in making a clinical assessment and may be considered identical to them if establishing an appropriate threshold value. The real-valued posterior can be used for categorization into arbitrary number of groups. For example, dividing the posterior values into three by setting thresholds 0.3 and 0.7, we obtain three groups whose survival curves are significantly different from each other; this tertiary categorization provides another definition of intermediate risk group based only on expression patterns (Figure S4).

Comparing the survival curves

Figure 3A shows survival curves for favorable and unfavorable patients predicted by the classifier with a binary threshold (0.5). The 5 year survival rate for the former (n = 98) was as good as 94%, while that for the latter (n = 38) was as poor as 33% ($p < 10^{-5}$). When 70 sporadic neuroblastomas were evaluated after excluding the tumors found by mass screening, the 5 year survival rate for the former (n = 40) was 85%, while that for the latter (n = 30) was 20% ($p < 10^{-5}$) (Figure 3B). To further evaluate the efficiency of our system, we calculated the posterior value for the intermediate subset of neuroblastoma (type II), whose prognosis is usually difficult to predict. As shown in Figure 3C, the survival curves were significantly categorized into two groups. The 5 year survival rate of patients who were predicted as favorable was 89%, while that for unfavorable patients was 36% ($p = 0.000067$). Since the age at diagnosis (≥ 1 year) is currently used as a poor prognostic factor for the type II tumors, we examined the ability of the classifier for the older patients with type II tumors. Even for such patients whose prognosis is difficult to predict, the survival rate (45%) of all 18 patients was divided solely by gene expression into the group with favorable prognosis (n = 10; 73%) and that with poor outcome (n = 8; 13%) (Figure 3D). In addition, if the intermediate risk group was further separated into stage 3 tumor group and stage 4 tumor group, the posterior value was significantly related to the survival, especially for stage 3 tumors (Figure S5). These results suggest that the posterior value obtained by our statistical analysis highly efficaciously allows the classification of patient outcomes, even when the tumor is of the intermediate type.

We further compared our results to existing prognosis markers in Table 1 and found that the supervised microarray analysis showed the best sensitivity-specificity balance among the prognostic factors for predicting the outcome of neuroblastoma. When the classifier is combined with the age at diagnosis, the disease stage (stage 1, 2, or 4s versus stage 3 or 4) and the *MYCN* amplification, accuracy, sensitivity, and specificity increased up to 95.8%, 93.3%, and 97.0%, respectively. Although the currently used markers (age, stage, and *MYCN*)



DA: Dead • or alive • at 2 year and 5 year after diagnosis
ST: Stage unfavorable (3,4) • or favorable (1,2,4s) •
NM: *MYCN* amp • or single •
AG: Age old (>1 year) • or young •
TA: *TRKA* low • or high •
PL: Diploidy • or aneuploidy •
MS: Sporadic • or from Mass-Screening •
SM: Shimada pathology unfavorable • or favorable •

Figure 2. Posterior probability of survival at 5 years

Posterior probability of survival at 5 years for 136 training data samples, output by the leave two out (LTO) crossvalidation without any information leakage. Left panel: Neuroblastoma samples. A red or blue horizontal line denotes survival period after diagnosis for a dead or alive patient, respectively. Red and blue marks denote various clinical properties of patients; see text below the panel for detailed explanation. Background colors show groups determined by stage and *MYCN* amplification status: red, type III, with *MYCN* amplification; green, type II, with single copy of *MYCN* at unfavorable stage (3 or 4); and blue, type I, with single copy of *MYCN* and at favorable stage (1, 2, or 4s). Right upper panel: The LTO crossvalidated prediction (posterior) for each patient; a red or a blue mark denotes that the patient is dead or alive at 5 years, respectively. Right middle panel: Cumulative smooth histogram of posterior probabilities for patients of dead (red), alive (blue), and unknown (white) at 5 years after diagnosis. Right lower panel: The horizontal and vertical axes denote the posterior and the empirical probability of 5 year survival, i.e., the ratio of the smooth histogram values between dead and alive patients, shown in the middle panel, respectively. Because the border between dead and alive samples is close to the white broken line (x = y), the posterior can be regarded as a 5 year survival chance rate.

Table 1. Accuracy of each marker for prognosis prediction (5 years after diagnosis)

	Whole cases				Sporadic cases		Intermediate and old age[a]	
	n	accuracy	sensitivity	specificity	n	accuracy	n	accuracy
Microarray classifier	136	89%	87%	89%	56	82%	14	86%
Age (less than 1 year old)	136	81%	83%	80%	56	71%	14	64%
Stages (1, 2, and 4s)	136	83%	97%	77%	56	84%	14	64%
Shimada classification (unfavorable)	62	87%	78%	89%	25	72%	(n < 10)	—
Hyperdiploidy (aneuploidy)	62	72%	67%	73%	27	56%	(n < 10)	—
MYCN amplification	136	89%	67%	99%	56	80%	14	36%
Microarray + age + stages + MYCN*	136	96%	93%	97%	56	93%	14	86%

Sensitivity/specificity is the rate of unfavorably/favorably predicted samples, i.e., LTO posterior <0.5/>0.5, among actually unfavorable/favorable samples. Microarray classifier, supervised classification based on the microarray data. *By this classifier, all patients with the *MYCN* amplification are predicted as unfavorable, and all patients with a single copy of *MYCN* and at stage 1, 2, or 4s are predicted as favorable. In the remaining intermediate samples (with a single copy of *MYCN* and at stage 3 or 4), the patients with age <1 year are predicted as favorable, and the microarray predictions are applied for those with age >1 year.
[a] Age at diagnosis >1 year.

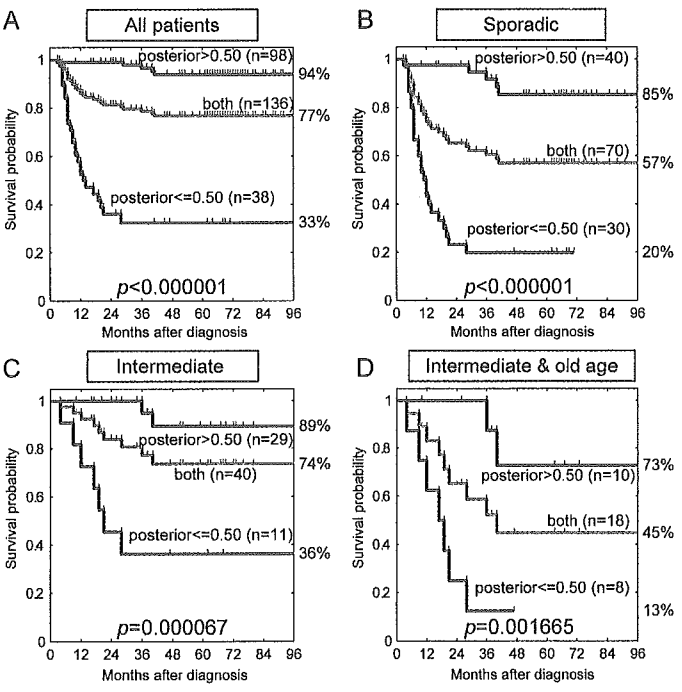


Figure 3. Disease-free survival of patients who were stratified based on the gene expression patterns

For each of the four figures, whole objective patients (green) are divided into favorable (blue) or unfavorable (red) based on the posterior values with threshold 0.5, which are calculated from gene expression patterns, and statistical features of their survival times are denoted by the Kaplan-Meier survival curves. The differences of the survival curves between the favorable (blue) and unfavorable (red) groups are evaluated by p values of the log rank test.

A and B: Survival analysis of whole and sporadic patients, respectively, divided by the supervised classifier based on microarray data.

C and D: Survival analysis of patients in the intermediate risk group with different definitions, divided by the supervised classifier. The intermediate risk group shown in (C) is defined as *MYCN* single and stage 3 or 4 (type II), and that in (D) is defined as *MYCN* single, stage 3 or 4, and older than 1 year of age.

also showed good potential to predict generally but less than the microarray, these exhibited only 64% accuracy of prediction for the type II tumors with $\geq$1 year of age (Table 1). Together with the results of survival analysis, the microarray classifier is revealed to be a powerful predictor to classify such group of neuroblastomas (86% accuracy; Table 1).

Practical application of 200 cDNAs microarray and independent test

For the practical use in the clinic, a cDNA microarray system that contains cDNA spots of a relatively small number and hence is easy to treat is expected. According to our gene selection based on the pairwise *F* score, the numbers of genes that were appropriate for the 5 year and 2 year prognosis prediction for all available samples were 10 and 70, respectively. In order for the system to reserve the applicability to short-term and long-term outcome prediction simultaneously, we selected 200 top-ranked genes according to the pairwise *F* scores in the 2 year prediction, because the 2 year prediction required larger variety of genes, and then made a smaller cDNA microarray system carrying the 200 genes. The newly designed microarray system (the mini-chip system) was evaluated by being hybridized with 5 μg total RNA obtained from 50 independent test samples. To preserve the independence of experimental procedure, these RNAs were prepared and hybridized in a different laboratory from the original experiments of 136 samples with the 5340 genes system (see Experimental Procedures). Although the weight values in the weighted voting classifier were determined by the 5340 genes system without any information leakage from the 50 independent samples, the result was as good as that obtained by the 5340 cDNA microarray analysis (90% [45/50] for 2 year, and 87.8% [43/50] for 5 year prognosis prediction; Figure 4B). This test validated not only the prediction robustness of our classifier constructed by the 5340 genes system, but also the construction procedure of the mini-chip system according to our gene selection based on pairwise *F* scores. When we reconstructed another supervised classifier by applying the LTO analysis to the 50 samples measured by the mini-chip system, the accuracy of the 5 year prediction was 91.8% (45/49) (Figure 4C). These results suggest

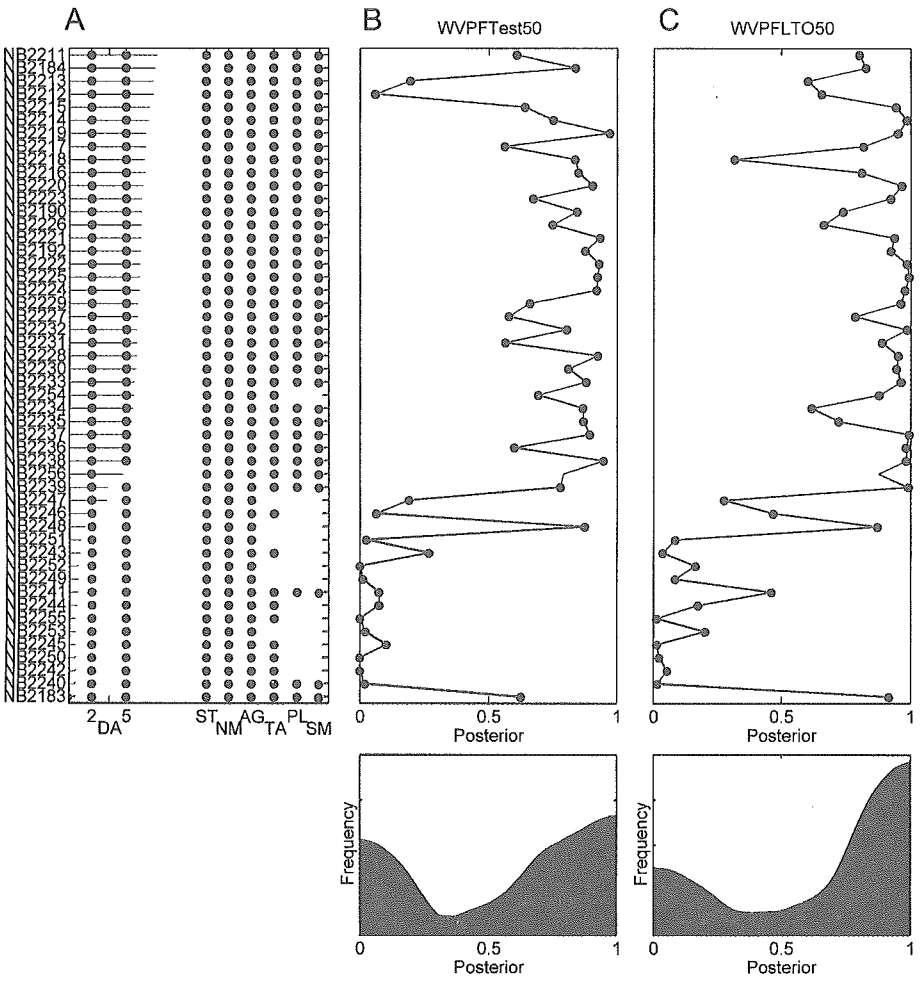


Figure 4. Posterior probability of survival at 5 years for test samples

Posterior probability of survival at 5 years for 50 independent test samples measured by newly developed 200 genes chip (the mini-chip system). Left panel: Neuroblastoma samples; see also Figure 2 legend. Center panel: Prediction results when the supervised classifier constructed from 96 training samples is applied to the 50 independent samples (independent test for the classifier's reproducibility). Right panel: LTO crossvalidation analysis using the new 50 samples (test for the procedure's reproducibility). Both tests do not introduce any information leakage. Lower panels: Smooth histograms of posterior probabilities for dead (red) and alive (blue) patients.

three things. (1) The supervised classifier obtained by the statistical analysis by the 5340 genes system is reproducible even if it is applied to the data measured by the reduced 200 genes system. Note that the 50 samples were completely new data for the classifier in this case. (2) Our procedure to construct a supervised classifier according to the LTO analysis is also reproducible, because the same procedure was successful in making another classifier with a high prediction accuracy when applied to the data taken by the mini-chip system. (3) A simple, low-cost microarray system, the mini-chip system, is highly feasible for predicting the prognosis of neuroblastoma.

Genes selected for prognosis prediction

To assess the relationship between the clinically defined subsets of neuroblastoma and the expression of 70 genes that were selected as top scored with 2 year prognosis according to the pairwise *F* score, we conducted an unsupervised clustering analysis (Figure 5). As expected, part of the type II (intermediate) tumors of patients with a poor prognosis showed an expression pattern that was similar to that of the type III (unfavorable) tumors, and many of them died. On the other hand, expression profiles of the remaining type II tumors seemed to be heterogeneous similarly to those of the type I (favorable) tumors with a good outcome. Most of the tumors with highly expressed *TrkA* and hyperdiploidy, as well as tumors detected by mass screening, were included in the latter group. Table 2 shows a list of 41 genes that corresponded to the 70 top-scored genes and their p and q values (Storey and Tibshirani, 2003) in the log rank test, since we found that several genes were duplicated in the selected 70 genes. Based on the above clustering, these genes were categorized into two groups (group F and group UF; the gene groups strongly correlated with favorable and unfavorable prognosis, respectively) (Figure 5 and Table 2).

The genes in group F tended to show high levels of expression in the type I tumors, while those in group UF were highly expressed in the type III tumors. The former contained genes that were related to neuronal differentiation (*tubulin* α, *peripherin*, *neuromodulin* [*GAP43*], and *HMP19*) and genes that were related to catecholamine metabolism (*dopa decarboxylase* [*DDC*], *dopamine* β-*hydroxylase* [*DBH*], and *tyrosine hydroxylase* [*TH*]). On the other hand, the latter involved many members of genes that are related to protein synthesis (ribosomal protein genes such as *RPL18A*, *RPLP0*, *RPL5*, *RPL4*, and *RPL7A* as well as translation initiation and elongation factor genes *EEF1G* and *EIF3S5*) and genes that are related to me-

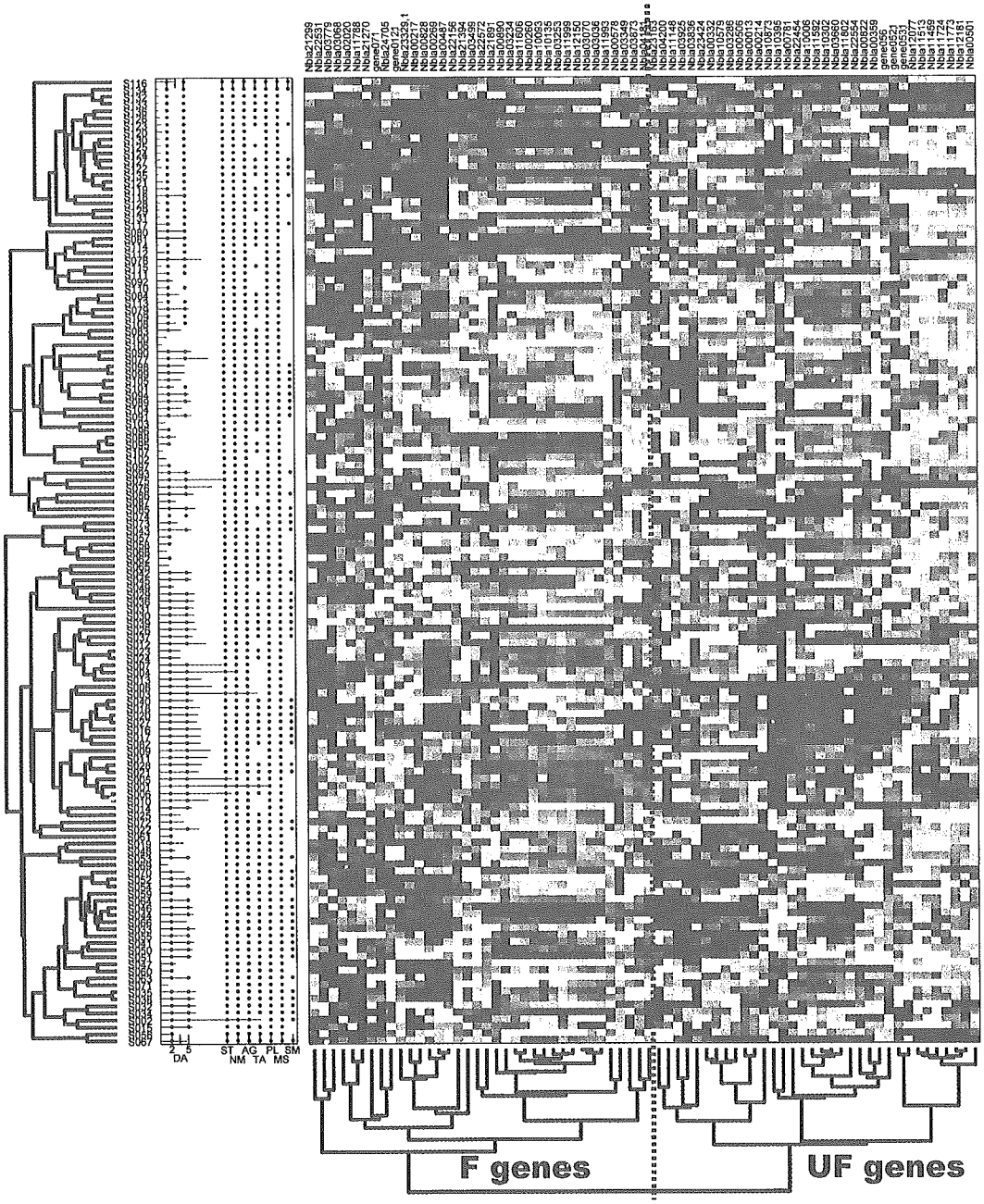


Figure 5. Expression profiles of 70 genes selected for predicting neuroblastoma prognosis at 2 years

Note that 10 genes for predicting prognosis at 5 years are also included in the 70 genes. The left and lower trees depict hierarchical clustering of the 136 neuroblastoma samples and the 70 genes selected in the present study, respectively. In the left tree, blue, green, and red colors denote "*MYCN* single and stage 1, 2 or 4s tumor" (type I, favorable), "*MYCN* single and stage 3, 4 tumor" (type II, intermediate), and "*MYCN* amplified tumor" (type III, unfavorable), respectively. The blue and red colors in the expression matrix show the high and low expression, respectively. A gene showing high expression level likely for unfavorable samples belongs to the group "UF" (red subtree in the lower tree), while one showing high expression likely for favorable samples belongs to the group "F" (blue subtree in the lower tree).

tabolism (*enolase 1* [*ENO1*] and *transketolase* [*TKT*]). The top 10 genes selected for the 5 year outcome prediction were *RPL18A*, *ENO1*, *EEF1G*, *TUBA3*, *GNB2L1*, *ARHGEF7*, *GCC2*, *DDX1* (duplicated), and *PRPH*. The *MYCN* gene was also a member of 70 genes (group UF) as expected; however, it was outside of the top 10 genes for the 5 year label. Instead, *DDX1*, which is frequently coamplified with *MYCN* on chromosome 2p24, was a member of the top 10 genes (UF group) for both of the 2 year and 5 year labels. Confirmation of the differential expression of the selected genes was further conducted by

Table 2. Top-ranked genes used for prediction of 2 year and 5 year prognosis of neuroblastoma

	Spot name	Accession number	Gene code	Chromosome map	Pattern	Log rank p	q value
F group							
	Nbla11606	NM_006009	*TUBA3*	12q13.12	F > UF	0	0.000674
	Nbla00890	NM_003899	*ARHGEF7*	13q34	F > UF	0.000001	0.000743
	Nbla00260	NM_006082	*K-ALPHA-1*	12q13.12	F > UF	0.000003	0.000926
	Nbla21891	U87309	*VPS41*	7p14.1	F > UF	0.000006	0.001096
	Nbla03873	NM_006054	*RTN3*	11q13.1	F > UF	0.00001	0.001282
	Nbla11788	NM_006262	*PRPH*	12q13.12	F > UF	0.000017	0.001522
	Nbla10093	NM_000183	*HADHB*	2p23.3	F > UF	0.000018	0.001541
	Nbla22572	NM_000790	*DDC*	7p12.2	F > UF	0.000035	0.00213
	Nbla21270	NM_001915	*CYB561*	17q23.3	F > UF	0.00016	0.00495
	gene071	NM_000360	*TH*	11p15.5	F > UF	0.000787	0.012173
	Nbla03499	NM_002074	*GNB1*	1p36.33	F > UF	0.000795	0.012237
	Nbla04181	AK55112	*AK55112*	5q13.2	F > UF	0.001425	0.017462
	Nbla00487	AB075512	*C6orf134*	6p21.33	F > UF	0.002751	0.025273
	Nbla00269	NM_000787	*DBH*	9q34.2	F > UF	0.00362	0.030407
	Nbla22531	NM_002045	*GAP43*	3q13.31	F > UF	0.004394	0.034175
	Nbla22156	NM_014944	*CLSTN1*	1p36.22	F > UF	0.005233	0.038274
	Nbla00578	NM_006818	*AF1Q*	1q21.3	F > UF	0.009397	0.05354
	Nbla00217	NM_032638	*GATA2*	3q21.3	F > UF	0.010245	0.056301
	Nbla21394	NM_000743	*CHRNA3*	15q25.1	F > UF	0.072464	0.162629
	Nbla11993	NM_015980	*HMP19*	5q35.2	F > UF	0.204274	0.282486
UF group							
	Nbla00214	NM_000980	*RPL18A*	19p13.11	F < UF	0.000002	0.001107
	Nbla00013	NM_006098	*GNB2L1*	5q35.3	F < UF	0.000006	0.001051
	Nbla11459	NM_004939	*DDX1*	2p24.3	F < UF	0.000024	0.001795
	Nbla11148	NM_001002	*RPLP0*	12q24.23	F < UF	0.000049	0.002549
	Nbla00332	NM_001404	*EEF1G*	11q12.3	F < UF	0.000055	0.002696
	Nbla10395	NM_002593	*PCOLCE*	7q22.1	F < UF	0.000164	0.005009
	Nbla03286	NM_020198	*GK001*	17q23.3	F < UF	0.000175	0.005204
	Nbla23163	NM_003754	*EIF3S5*	11p15.4	F < UF	0.000341	0.007105
	Nbla10579	NM_181453	*GCC2*	2q12.3	F < UF	0.000962	0.01407
	Nbla00359	NM_003550	*MAD1L1*	7p22.3	F < UF	0.00112	0.01525
	gene052-1	NM_005378	*MYCN*	2p24.3	F < UF	0.001253	0.016367
	Nbla03925	NM_002295	*LAMR1*	3p22.2	F < UF	0.001773	0.01931
	Nbla23424	NM_001404	*EEF1G*	11q12.3	F < UF	0.003579	0.030326
	Nbla22554	NM_000687	*AHCY*	20q11.22	F < UF	0.003946	0.032409
	gene056	NM_000546	*TP53*	17p13.1	F < UF	0.004087	0.032829
	Nbla10873	NM_005762	*TRIM28*	19q13.43	F < UF	0.004984	0.037476
	Nbla00501	NM_000969	*RPL5*	1p22.1	F < UF	0.005786	0.04012
	Nbla10302	NM_001428	*ENO1*	1p36.23	F < UF	0.007702	0.048179
	Nbla04200	NM_000968	*RPL4*	15q22.31	F < UF	0.04097	0.120453
	Nbla03836	NM_000972	*RPL7A*	9q34.2	F < UF	0.048031	0.132345
	Nbla00781	NM_001064	*TKT*	3p21.1	F < UF	0.048075	0.132342

Although 70 clones were selected as important genes for the supervised classifier, duplicated and multiplicated clones are omitted in this table. The 41 genes are classified into two groups, "F > UF" and "F < UF," when the expression in favorable samples is higher than that in unfavorable samples, and vice versa, respectively. In each group, genes are sorted by log rank p values. The log rank p value for each gene was calculated by comparing survival curves of two patient groups, in which the expression of the gene is higher and lower, respectively, than the median over the samples. A "q value" of a gene denotes the estimated false discovery rate among the genes whose p value is the same or smaller than that of the gene, and is a p-like value while incorporating multiplicity of the statistical test.

using representative 16 favorable and 16 unfavorable tumor samples that were independent of the 136 samples used in the present analysis, by semiquantitative RT-PCR (Figure S6; refer also to Ohira et al., 2003a). We also conducted immunohistochemical analysis for peripherin antibody using tissue sections prepared from primary neuroblastoma with favorable and unfavorable histology, since peripherin gene is a member of the top 10 genes for both 2 year and 5 year outcome prediction (Table 2). Peripherin protein was positively detected in the cytoplasm of neuroblastic cells as well as neuritis in all three favorable histology tumors (Figure S7, FH&NA). Two unfavorable histology tumors with poorly differentiated subtype, regardless of *MYCN* status, showed sporadic staining (less than 20% of the favorable histology tumor) for peripherin protein in neurites. Peripherin was completely negative in the unfavorable histology tumor of undifferentiated subtype (Figure S7, UF&NA). These results indicate the reliability of our gene selection. In the log rank test, p values of 18 of 20 genes in group F and of all 21 genes in group UF were less than 0.05 (Table 2), indicating that these 39 genes can be independent prognostic factors for primary neuroblastomas.

Discussion

Our study has disclosed the molecular signature of neuroblastoma that predicts patient outcomes by using RNA ob-

tained from 136 primary neuroblastomas. The highly reliable statistical analysis by using a neuroblastoma proper cDNA microarray harboring 5340 genes based on an electrically controlled ceramics-based ink-jet method led us to design a cDNA microarray system harboring 200 genes, which is applicable to short-term (2 year) and long-term (5 year) prognosis predictions for neuroblastoma.

Our study demonstrated that the supervised classifier produced by the 5340 genes system provided a high accuracy (88.5%) for the 5 year outcome prediction, with a good balance between sensitivity (86.7%) and specificity (89.4%). Although age at diagnosis, disease stage, *MYCN* amplification, and patients found by mass screening have been useful prognostic markers currently used at the bedside, most of them have either high sensitivity or high specificity (Table 1). The microarray analysis showed the best sensitivity-specificity balance among the prognostic factors for predicting the outcome of neuroblastoma. When the classifier is combined with the age at diagnosis, the disease stage (stage 1, 2, or 4s versus stage 3 or 4) and the *MYCN* amplification, accuracy, sensitivity, and specificity increased up to 95.8%, 93.3%, and 97.0%, respectively. Furthermore, the intermediate subset of neuroblastomas (type II), for which a long-term prognosis is usually difficult to make, was also categorized by microarray analysis into groups of patients with a favorable prognosis and those with an unfavorable prognosis. These successful results led us to produce a more practical tool at the bedside, the mini-chip system, whose accuracy, sensitivity, and specificity were 87.8%, 76.5%, and 93.8%, respectively, when the classifier constructed by the 5340 genes system was applied to 50 independent samples measured by the mini-chip system, and were 91.8%, 82.4% and 96.9%, respectively, when another classifier was constructed by applying the LTO procedure to the same data (Figure 4).

It is well recognized now that gene expression analyses for cancer prognosis prediction should pay close attention to the reproducibility of obtained results. A complete crossvalidation analysis without introducing any information leakage and an independent test using new samples are necessary. Although the determination of the appropriate number of genes used in supervised classifiers should be included in the validation procedure, it has often been ignored in most microarray studies. van 't Veer et al. (2002) applied the supervised classification to the breast cancer gene signature, which is predictive of a short interval to distant metastases in 78 patients who were initially devoid of local lymph node metastasis. Although their crossvalidation analysis without the validation of the number of genes correctly predicted the actual outcome of disease for 63 of 78 patients (80.7%), the accuracy was worse when a complete validation was applied (73.1%). This difference suggests that even small information leakage may lead to overestimation of the accuracy. Beer et al. (2002) applied the supervised classification to the outcome prediction of lung adenocarcinoma. Their statistical analysis was complete without any information leakage. They did not report the prediction accuracy, but we estimated the accuracy to be about 70% from the data in their paper and found that the prediction by their supervised classifier was not very superior to that by existing prognosis markers. Iizuka et al. (2003) applied the supervised classification to the prediction of intrahepatic recurrence within 1 year after curative surgery for hepatocellular carcinoma patients. Although their predictor showed sufficiently high accuracy in an independent test with 27 samples, their crossvalidation procedure excluded the validation of the determination process of the number of optimum genes (steps 5 and 6 in their algorithm). The high crossvalidation accuracy of 100% may be an overestimation due to the information leakage.

According to the recent study that evaluated statistical methodologies used by microarray studies published between 1995 and April 2003, the three papers above were the only ones that reported both fairly sound crossvalidation analyses and independent tests (Ntzani and Ioannidis, 2003). Our LTO procedure includes the validation process of the number of genes used in the classifier and hence is a complete crossvalidation process. In addition, the obtained classifier was applied to the 50 independent samples that were measured by the reduced 200 genes system. This is a stronger test than usual independent tests but is important for the development of a practical system at the bedside. In addition, our LTO analysis achieved an almost unbiased estimation of the accuracy. Our crossvalidation analysis using the LTO procedure, the independent test of the classifier, and the validation of the procedure itself within a new experimental environment using the mini-chip system exhibited one of the most conservative and reliable statistical methodologies. In addition, our gene selection procedure according to the pairwise F score tries to extract correlation structures among genes, based on an idea similar to the exhaustive optimization method used in Iizuka et al. (2003), is beneficial in enhancing the applicability of the mini-chip system to various prediction problems, namely, short-term and long-term outcome predictions.

In addition to high accuracy, another advantage of our method is to provide a type of predictive information beyond the conventional binary prediction like favorable and unfavorable, which is ambiguous. The probabilistic output based on the hypothetical distribution obtained by the LTO analysis, the posterior probability, was found to show good accordance with actual survival rate (right bottom panel in Figure 2); this enables us to make a simple interpretation of the output: a patient with a posterior value of 0.8 has 80% chance for the 5 year survival, for example. Moreover, by calculating posterior probabilities for various future time points, a survival chance curve for each patient can be depicted (Figure 6). Although the follow-up period of patient "S057" is 2 years, and the patient is alive at this time, the individual survival chance curve says that his/her survival chance estimated from the gene expression pattern at diagnosis will get smaller than 50% at about 3 years after diagnosis. Such an individual survival chance curve can be used in choosing a suitable therapeutic protocol.

Another advantage of our method is that the probabilistic output is very stable in the presence of noise. Even when an artificial noise, whose variance is as large as the estimated noise variance of microarray, was added to expression profile data, prognosis prediction did not degrade very much (Figure S8). This robustness was confirmed when the noise variance went up to 1.0, which was sufficiently greater than the actual reproduction noise level of 0.4 (Figures S1A–S1C).

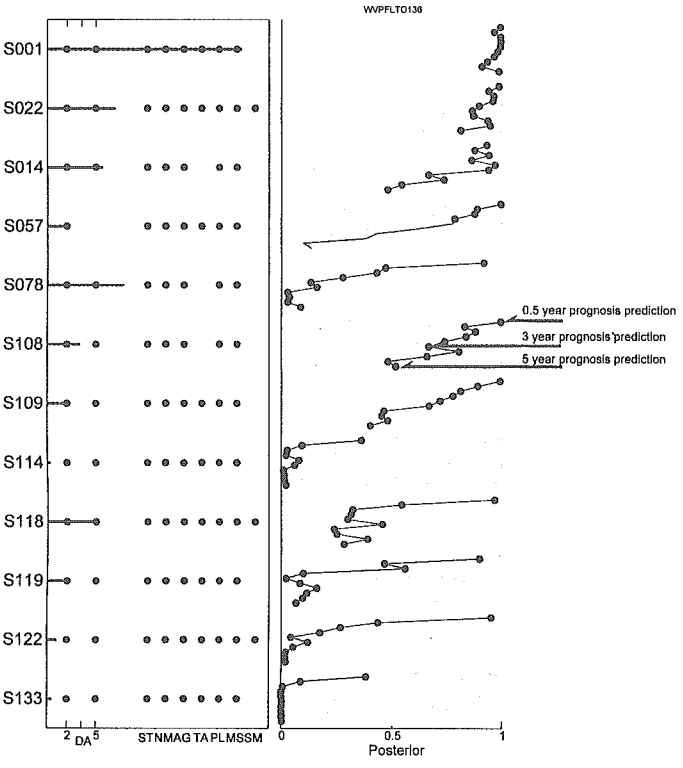


Figure 6. Individual survival chance analysis based on posterior probabilities

LTO estimation of survival probabilities at 0.5, 1.0, 1.5, ..., 5.0 years after diagnosis for 12 typical patients. Left panel: Information of patients (see caption of Figure 2). Right panel: Estimated posterior probabilities at 0.5, 1.0, 1.5, ..., 5.0 years after diagnosis, which predict the time course of patient's survival chance. A blue or a red mark denotes that the patient is alive or dead at that time after diagnosis, respectively. For example, the patient "S108," who died at 40 months, is predicted as 100% alive at 0.5 year and 52% alive at 5 year, solely from the microarray analysis at the diagnosis

The high outcome predictability of our system is attributable to multiple reasons. The quality of tumor samples is high because (1) an appropriate system was established for our neuroblastoma tissue bank, and (2) handling of tumor tissues is rather uniform at each hospital, in which informed consent was obtained. An array, produced by a new apparatus equipped with a piezo microceramic pump, generates highly reproducible signals. The noncontact spotting method makes the spot shape almost a perfect circle. Consequently, the spot excels in signal uniformity. We did not conduct microdissection of the 136 tumor samples, because the stromal components of the tumor, e.g., Schwannian cells, are already known to be very important to characterize its biology (Ambros and Ambros, 1995; Ambros and Ambros, 2000). Therefore, a good combination or selection of these procedures may have provided high outcome predictability. In addition, the high predictability was reliably confirmed by the complete crossvalidation analysis and the independent test. The probabilistic output based on the LTO analysis can provide a new type of information that will improve the therapeutic decision at the bedside. In addition, the probabilistic output is highly robust against noises that may be involved in test samples (described above); this can be the major reason for the high prediction accuracy when the classifier constructed by the 5340 genes system was applied to the data taken by the mini-chip system.

The impact of the selected genes is strong. The genes with the highest score in F group genes (F > UF) were *tubulin* α members (*TUBA3* and *K-ALPHA-1*, which corresponds to *TUBA1*), which have never been reported to be prognostic factors in neuroblastoma. Their prognostic significance has also been confirmed by RT-PCR in primary tumors (data not shown). The high expression of *TUBA1* in neuronal cells is associated with axonal outgrowth during development as well as with axonal degeneration after axotomy in adult animals (Knoops and Octave, 1997). The expression of *TUBA3* has been reported to be restricted to adherent, morphologically differentiated neuronal and glial cells (Hall and Cowan, 1985). We have also found that high expression of *tubulin tyrosine ligase* and enhanced tubulin tyrosination/detyrosination cycle are associated with neuronal differentiation in neuroblastomas with favorable prognosis (Kato et al., 2004). Thus, high mRNA expression of *TUBA* genes in favorable neuroblastoma may reflect differentiated status of tumors. ARHGEF7, Rho guanine nucleotide exchange factor 7, activates Rho proteins by exchanging bound GDP for GTP and can induce membrane ruffling. In our previous paper, we found that many family members of such G protein-related genes are highly expressed in favorable neuroblastomas compared to unfavorable ones (Ohira et al., 2003a). This may also imply a neuronal maturity nature of favorable tumors. Peripherin, a type III intermediate filament protein, was initially found as a cytoskeletal protein in the peripheral nervous system and in cultured cells of neuronal origin. This protein is known to be a marker of terminal neuronal differentiation; however, its functional role in neuroblastoma has been elusive. The previous evidence indicates that peripherin is transcriptionally upregulated by treatment with NGF, an important neurotrophin in neuroblastoma, and that the protein product is directly phosphorylated by NGF receptor, TrkA (Aletta et al., 1989). Thus, peripherin may play an important role as one of the signal transduction components involved in elaboration and maintenance of neuronal differentiation. In the UF gene group, many ribosomal protein-related genes are selected. *GNB2L1*, a receptor for activated C-kinase *RACK1*, is implicated in linking between *PKC* signaling and ribosome activation (Ceci et al., 2003). The *DDX1* gene, which is frequently coamplified with the *MYCN* gene in advanced neuroblastomas (Godbout and Squire, 1993; Noguchi et al., 1996), is also a member of this group. Its protein product is a putative RNA helicase and is implicated in a number of cellular processes involving alteration of RNA secondary structure such as translation initiation, nuclear and mitochondrial splicing, and ribosome and spliceosome assembly. *DDX1* is ranked at a higher score than the *MYCN* gene, which is concordant with the previous reports describing that *MYCN* mRNA expression is a weaker prognostic marker than its genomic amplification (Slavc et al., 1990). Another important prognostic factor, *TrkA*, is not included in the top 70 genes but in the 90 (in the top 20 genes when the 5 year label was used) (data not shown), probably due to its relatively low levels of mRNA expression as compared with those of other genes. The prognos-

tic effect of *TrkA* expression may be compensated by other genes which are affected or regulated by *TrkA* intracellular signaling. Similarly, *MYCN*-regulated genes such as ribosomal genes, translation initiation and elongation factors, and laminin receptor may compensate the effect of *MYCN* gene expression in aggressive tumors. It is intriguing that high mRNA expression of *p53* gene is also strongly related to unfavorable outcome. Although *p53* mutation is rare in primary neuroblastomas, and its gene product frequently accumulated in cytoplasm, an unknown mechanism that upregulates *p53* expression in aggressive tumors may exist.

Our results showed that the decision by majority by the genes selected based on microarray data alone can be a prognostic indicator comparative to the existing prognostic markers, and that the addition of the microarray data to the prognosis markers improved the outcome prediction (Table 1). The outcomes of patients belonging to the intermediate subset, whose prognosis prediction had been very difficult by existing prognosis markers, were effectively separated into favorable group and unfavorable group ($p < 10^{-4}$). The posterior value will help the decision of therapeutic modalities, and outcome prediction based on the posterior value is extremely robust against a possible noise. In addition, our practical, low-cost microarray carrying only 200 genes should make its clinical use possible. Our further validation by hybridizing RNA obtained from 50 fresh neuroblastomas on the 200 cDNAs microarray in a completely independent laboratory indicated that our prediction system is consistent and feasible. Therefore, the application of a highly qualified cDNA microarray at the bedside may bring tailored medicine that allows better treatment of neuroblastoma patients.

Experimental procedures

Patients and tumor specimens

Fresh, frozen tumor tissues were sent to the Division of Biochemistry, Chiba Cancer Center Research Institute, from a number of hospitals in Japan (1996–2002). Informed consent was obtained at each institution or hospital. We randomly selected tumor samples from this neuroblastoma tissue bank and then successfully conducted hybridization in 136 neuroblastomas consisting of 41 stage 1 tumors, 22 stage 2 tumors, 33 stage 3 tumors, 28 stage 4 tumors, and 12 stage 4s tumors. Among the 136 fresh neuroblastomas, seventeen tumors were obtained at the delayed primary surgery after giving chemotherapy, but the other 119 tumors were resected by biopsy or surgery without giving any therapy. After surgery, patients were treated according to the previously described common protocols (Kaneko et al., 1998). Biological information on each tumor, including *MYCN* gene copy number, *TrkA* gene expression, and DNA ploidy, was analyzed in our laboratory, as described previously (Hishiki et al., 1998). All the tumors were classified according to the International Neuroblastoma Staging System (INSS) (Brodeur et al., 1993). The stage 4s neuroblastoma shows a special pattern of clinical behaviors, and the tumor itself, as well as its widespread metastases to the skin, liver, or bone marrow, usually regresses spontaneously. For a better understanding of statistical results, we introduced Brodeur's classification of neuroblastoma subsets: type I (stages 1, 2, or 4s; a single copy of *MYCN*; blue marks in Figure 2), type II (stage 3 or 4; a single copy of *MYCN*; green marks in Figure 2), and type III (all stages; amplification of *MYCN*; red marks in Figure 2) (Brodeur and Nakagawara, 1992). Among 136 tumors that we analyzed, 66 were found by mass screening of urinary catecholamine metabolites at the age of 6 months, which has been performed nationwide in Japan from 1984 to 2004 (Sawada et al., 1984). The follow-up duration ranged between 3 and 241 months (median, 56 months; mean, 57.3 months) after diagnosis. All diagnoses of neuroblastoma were confirmed by the histological assessment of a surgically resected tumor specimen at each hospital. Shimada's classification (Shimada et al., 1984) was performed in 62 out of 136 cases. The macroscopic necroses in the tumor were excluded from the tissue sampling for molecular analysis. We used for the microarray analysis only the tumor samples whose adjacent tissues contained more than 70% tumor cells in the thin sections stained with hematoxylin-eosin. For independent test, 50 (19 were found by mass screening and 31 were clinically found) tumors (15 of stage 1, 6 of stage 2, 9 of stage 3, 14 of stage 4, and 6 of stage 4s) were used.

Total RNA was extracted from each frozen tissue according to the AGPC method (Chomczynski and Sacchi, 1987). RNA integrity, quality, and quantity were then assessed by electrophoresis on the Agilent RNA 6000 nanochip using Agilent 2100 BioAnalyzer (Agilent Technologies, Inc.).

cDNA microarray experiments

We previously obtained approximately 5,000 genes after selecting from 10,000 clones randomly picked up from the mixture of oligo-capping cDNA libraries, which were generated from three primary neuroblastomas with a favorable outcome (stage 1; high *TrkA* expression and a single copy of *MYCN*), three tumors with a poor prognosis (stage 3 or 4; low expression of *TrkA* and amplification of *MYCN*), and a stage 4s tumor (Ohira et al., 2003a; Ohira et al., 2003b). Using these isolated genes together with 80 known cDNAs that were thought to be neuroblastoma-related genes, we first constructed a neuroblastoma proper cDNA microarray (named CCC-NB5000-Chip) carrying 5340 cDNA spots (the 5340 genes system). Insert DNAs (average size, approximately 2.5kb) were amplified by polymerase chain reaction (PCR) from these cDNA clones, purified by ethanol precipitation, and spotted onto a glass slide at a high density with an ink-jet printing tool (NGK Insulators, Ltd.).

Ten micrograms each of total RNA were labeled with the CyScribe RNA labeling kit in accordance with the manufacturer's manual (Amersham Pharmacia Biotech), followed by probe purification with the Qiagen MinElute PCR purification kit (Qiagen). We used a mixture of equal amounts of RNA from each of four neuroblastoma cell lines (NB69, NBL-S, SK-N-AS, and SH-SY5Y) as a reference. RNAs extracted from primary neuroblastoma tissues and RNAs of the reference mixture were labeled with Cy3 and Cy5 dye, respectively, and were used as probes together with yeast tRNA and polyA for suppression. Subsequent hybridization and washing were conducted as described previously (Takahashi et al., 2002; Yoshikawa et al., 2000). Hybridized microarrays were scanned using the Agilent G2505A confocal laser scanner (Agilent Technologies, Inc.), and fluorescent intensities were quantified using the GenePix Pro microarray analysis software (Axon Instruments, Inc.). The procedure of this study was approved by the Institutional Review Board of the Chiba Cancer Center.

After selecting genes strongly related to the prognosis of patients with neuroblastoma (at 2 years and at 5 years after diagnosis), we constructed a 200 cDNAs microarray on glass slides by the same procedure described above (the mini-chip system). For the independent test using 50 samples, tumor RNA preparation, probe labeling, and hybridization were conducted in a completely different laboratory from the original 136 hybridization. In this independent test, 5 μg each of total RNA were used for labeling.

Data preprocessing

To remove chip-wise biases of a microarray system, we used the LOWESS normalization (Cleveland, 1979). When the Cy3 or Cy5 strength for a clone was smaller than 3, strength was regarded as abnormally small, and the log expression ratio of the corresponding clone was treated as a missing value. The rate of such missing entries was less than 1%. After normalizing the 5340 (genes) by 136 (samples) log expression matrix and removing missing values, each missing entry was imputed to an estimated value by Bayesian principal component analysis, which was developed previously (Oba et al., 2003).

Supervised machine learning and LTO crossvalidation

The 96 samples, whose prognosis at 5 years after diagnosis had been successfully checked, were used to train a supervised classifier that predicts the 5 year prognosis of a new patient. When we considered the short-term prediction, 126 samples whose 2 year prognosis is known were used. Selection of the genes that are related to the classification is an important preprocess for reliable prediction. We omitted the genes whose standard

deviation of the log ratios for the genes obtained over 136 experiments was smaller than 0.36, so that 1000 genes remained, because the background noise level was about 0.2–0.3. After the gene screening, the genes were scored by the pairwise *F* score, which is a modification of a pairwise correlation method (Bo and Jonassen, 2002), to conduct gene ranking in an attempt not only to obtain higher discrimination accuracy by using a smaller number of genes but also to reserve the applicability to various outcome prediction by the set of selected genes (see the Supplemental Data).

We used a well-established technique in the supervised classification (prognosis prediction), that is, weighted voting with linear discriminators, where each weight value was calculated as the signal-to-noise ratio (Golub et al., 1999). In the weighted voting, only n genes with the largest pairwise *F* score were used. The number of top genes, n, strongly affects the prediction accuracy (Figure S3) as found in various microarray studies and hence should be determined such to maximize the leave one out (LOO) crossvalidation accuracy. However, a naive determination process of n may introduce information leakage, and the accuracy optimized by the LOO crossvalidation involves overestimation. To avoid such an overestimation, we consulted a LTO analysis. The LTO analysis was constituted of inner and outer loops of LOO (Figure S2A); the gene number n was optimized by the LOO crossvalidation repeating the inner loops, and the optimized classifier was evaluated by an independent test for a single sample left out at a single step in the outer loop. During repetition of such steps, the test results of the outer loop were never fed back to the classifier's optimization process in the inner loops, and hence the tests in the outer loop did not include any overestimation, and the estimated accuracy involved the smallest bias as possible.

The posterior value for a single sample was calculated based on the distribution of the weighted vote (decision by majority by the genes that join the vote) *f* within the LTO analysis. We regard a real-valued weighted vote as carrying two kinds of information: its sign predicts the label (favorable or unfavorable) of the corresponding sample, and its absolute value shows the prediction strength. The posterior probability p for this sample being favorable (alive at 5 years) was evaluated as the logit transformation $p = \exp(\beta_0 + \beta_1 f)/[1 + \exp(\beta_0 + \beta_1 f)]$, where parameters β_0 and β_1 were estimated by the maximum likelihood method, in each step in the outer loop of LTO using the remaining 95 samples and the corresponding labels (5 year prognosis). Then, the posterior probability of the sample left out in the outer loop was predicted by the weighted vote *f* by the classifier constructed in the inner LOO loops and the parameters β_0 and β_1 obtained above. There is therefore no information leakage in this calculation process of the posterior of the sample left out.

Independent test

Using the 50 independent samples, we performed two kinds of tests. The first one is an independent test to validate the classifier obtained by our method and the applicability of our classifier to the mini-chip system, which has been developed as a clinical tool at the bedside (Figure S2B). According to the LTO analysis, the supervised classifier was finally constructed by using all of the 96 training samples measured by the 5340 genes system. This classifier was evaluated by being directly applied to the 50 samples measured by the mini-chip system without any information from measurements by the mini-chip system and the 50 test samples. In this test, tumor RNA preparation, probe labeling, and hybridization were conducted in a completely different laboratory from that for the 5340 genes system. The second one is to validate the LTO analysis to construct a supervised classifier by applying the procedure to the data taken by the mini-chip system.

Survival analysis

The Kaplan-Meier survival analysis was also programmed and used to compare patient survival. To assess the association of selected gene expression with patient clinical outcome, the statistical p and q values were calculated based on the log rank test.

Immunohistochemistry

Immunostaining with the antibody against peripherin protein (Santa Cruz Biotechnology; 1:400) was performed on six human neuroblastoma tumors selected from the surgical pathology file at the Department of Pathology, Aichi Medical University. They were all neuroblastoma (Schwannian stroma–poor) and included three favorable histology tumors (poorly differentiated subtype without *MYCN* amplification [one case]; differentiated subtype without *MYCN* amplification [two cases]) and three unfavorable histology tumors (undifferentiated subtype without *MYCN* amplification [one case]; poorly differentiated subtype with *MYCN* amplification [one case]; poorly differentiated subtype without *MYCN* amplification). All tumor tissues were obtained prior to chemotherapy and irradiation therapy. Four micron thick sections from the formalin-fixed, paraffin-embedded samples of these tumors were treated according to the protocol described previously (Kato et al., 2004). As for the negative controls, normal goat immunoglobulins (1:500 dilution; Vector Laboratories) were applied as the primary antibody.

Supplemental data

The Supplemental Data include Supplemental Experimental Procedures and ten supplemental figures and can be found with this article online at http://www.cancercell.org/cgi/content/full/7/4/337/DC1/.

Acknowledgments

We are grateful to the hospitals and institutions that provided us with surgical specimens (see the Supplemental Data). We also thank Shigeru Sakiyama and John K. Cowell for reading the manuscript; Naohiko Seki, Tsutomu Yoshikawa, and Masaki Kato for their valuable suggestions; and Natsue Kitabayashi, Tomonori Saito, Naoko Sugimitsu, Yuki Nakamura, Naoko Shibano, Emiko Kojima, Hisae Murakami, and Kazumi Yagyu for their technical support. This work was supported in part by a fund from Hisamitsu Pharmaceutical Co., Inc.; by Grants-in-Aid for Scientific Research on Priority Areas (C) "Medical Genome Science" and "Genome Information Science" and for Scientific Research (B) from the Ministry of Education, Culture, Sports, Science and Technology of Japan; and by Grant-in Aid for Cancer Research from the Ministry of Health, Labor and Welfare of Japan.

Received: November 17, 2003
Revised: January 8, 2005
Accepted: March 11, 2005
Published: April 18, 2005

References

Aletta, J.M., Shelanski, M.L., and Greene, L.A. (1989). Phosphorylation of the peripherin 58-kDa neuronal intermediate filament protein. J. Biochem. (Tokyo) *264*, 4619–4627.

Ambros, I.M., and Ambros, P.F. (1995). Schwann cells in neuroblastoma. Eur. J. Cancer *4*, 429–434.

Ambros, I.M., and Ambros, P.F. (2000). The role of Schwann cells in neuroblastoma. In Neuroblastoma, G.M. Brodeur, T. Sawada, Y. Tsuchida, and P.A. Voute, eds. (Amsterdam: Elsevier), pp. 229–243.

Beer, D.G., Kardia, S.L., Huang, C.C., Giordano, T.J., Levin, A.M., Misek, D.E., Lin, L., Chen, G., Gharib, T.G., Thomas, D.G., et al. (2002). Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat. Med. *8*, 816–824.

Berwanger, B., Hartmann, O., Bergmann, E., Bernard, S., Nielsen, D., Krause, M., Kartal, A., Flynn, D., Wiedemeyer, R., Schwab, M., et al. (2002). Loss of a FYN-regulated differentiation and growth arrest pathway in advanced stage neuroblastoma. Cancer Cell *2*, 377–386.

Bo, T., and Jonassen, I. (2002). New feature subset selection procedures for classification of expression profiles. Genome Biol. *3*, RESEARCH0017.

Bolande, R.P. (1974). The neurocristopathies: a unifying concept of disease arising in neural crest maldevelopment. Hum. Pathol. *5*, 409–429.

Brodeur, G.M., and Nakagawara, A. (1992). Molecular basis for clinical heterogeneity in neuroblastoma. Am. J. Pediatr. Hematol. Oncol. *14*, 111–116.

Brodeur, G.M., Seeger, R.C., Schwab, M., Varmus, H.E., and Bishop, J.M.

(1984). Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage. Science *224*, 1121–1124.

Brodeur, G.M., Fong, C.T., Morita, M., Griffith, R., Hayes, F.A., and Seeger, R.C. (1988). Molecular analysis and clinical significance of N-myc amplification and chromosome 1p monosomy in human neuroblastomas. Prog. Clin. Biol. Res. *271*, 3–15.

Brodeur, G.M., Pritchard, J., Berthold, F., Carlsen, N.L., Castel, V., Castelberry, R.P., De Bernardi, B., Evans, A.E., Favrot, M., Hedborg, F., et al. (1993). Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. J. Clin. Oncol. *11*, 1466–1477.

Ceci, M., Gaviraghi, C., Gorrini, C., Sala, L.A., Offenhauser, N., Marchisio, P.C., and Biffo, S. (2003). Release of eIF6 (p27BBP) from the 60S subunit allows 80S ribosome assembly. Nature *426*, 579–584.

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

Cleveland, W.S. (1979). Robust locally weighted regression and smoothing scatterplots. J. Am. Stat. Assoc. *74*, 829–836.

Evans, A.E., D'Angio, G.J., and Randolph, J. (1971). A proposed staging for children with neuroblastoma. Children's cancer study group A. Cancer *27*, 374–378.

Favrot, M.C., Combaret, V., and Lasset, C. (1993). CD44—a new prognostic marker for neuroblastoma. N. Engl. J. Med. *329*, 1965.

Godbout, R., and Squire, J. (1993). Amplification of a DEAD box protein gene in retinoblastoma cell lines. Proc. Natl. Acad. Sci. USA *90*, 7578–7582.

Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., et al. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science *286*, 531–537.

Hall, J.L., and Cowan, N.J. (1985). Structural features and restricted expression of a human α-tubulin gene. Nucleic Acids Res. *13*, 207–223.

Hishiki, T., Nimura, Y., Isogai, E., Kondo, K., Ichimiya, S., Nakamura, Y., Ozaki, T., Sakiyama, S., Hirose, M., Seki, N., et al. (1998). Glial cell line-derived neurotrophic factor/neurturin-induced differentiation and its enhancement by retinoic acid in primary human neuroblastomas expressing c-Ret, GFR α-1, and GFR α-2. Cancer Res. *58*, 2158–2165.

Hiyama, E., Hiyama, K., Yokoyama, T., Matsuura, Y., Piatyszek, M.A., and Shay, J.W. (1995). Correlating telomerase activity levels with human neuroblastoma outcomes. Nat. Med. *1*, 249–255.

Iizuka, N., Oka, M., Yamada-Okabe, H., Nishida, M., Maeda, Y., Mori, N., Takao, T., Tamesa, T., Tangoku, A., Tabuchi, H., et al. (2003). Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection. Lancet *361*, 923–929.

Kaneko, M., Nishihira, H., Mugishima, H., Ohnuma, N., Nakada, K., Kawa, K., Fukuzawa, M., Suita, S., Sera, Y., and Tsuchida, Y. (1998). Stratification of treatment of stage 4 neuroblastoma patients based on N-myc amplification status. Study Group of Japan for Treatment of Advanced Neuroblastoma, Tokyo, Japan. Med. Pediatr. Oncol. *31*, 1–7.

Kato, C., Miyazaki, K., Nakagawa, A., Ohira, M., Nakamura, Y., Ozaki, T., Imai, T., and Nakagawara, A. (2004). High expression of human tubulin tyrosine ligase and enhanced tubulin tyrosination/detyrosination cycle are associated with neuronal differentiation in neuroblastomas with favorable prognosis. Int. J. Cancer *112*, 365–375.

Knoops, B., and Octave, J.N. (1997). α1-tubulin mRNA level is increased during neurite outgrowth of NG 108-15 cells but not during neurite outgrowth inhibition by CNS myelin. Neuroreport *8*, 795–798.

Look, A.T., Hayes, F.A., Nitschke, R., McWilliams, N.B., and Green, A.A. (1984). Cellular DNA content as a predictor of response to chemotherapy in infants with unresectable neuroblastoma. N. Engl. J. Med. *311*, 231–235.

Look, A.T., Hayes, F.A., Shuster, J.J., Douglass, E.C., Castleberry, R.P., Bowman, L.C., Smith, E.I., and Brodeur, G.M. (1991). Clinical relevance of tumor cell ploidy and N-myc gene amplification in childhood neuroblastoma: a Pediatric Oncology Group study. J. Clin. Oncol. *9*, 581–591.

Nagata, T., Takahashi, Y., Asai, S., Ishii, Y., Mugishima, H., Suzuki, T., Chin, M., Harada, K., Koshinaga, S., and Ishikawa, K. (2000). The high level of hCDC10 gene expression in neuroblastoma may be associated with favorable characteristics of the tumor. J. Surg. Res. *92*, 267–275.

Nakagawara, A., Arima, M., Azar, C.G., Scavarda, N.J., and Brodeur, G.M. (1992). Inverse relationship between trk expression and N-myc amplification in human neuroblastomas. Cancer Res. *52*, 1364–1368.

Nakagawara, A., Arima-Nakagawara, M., Scavarda, N.J., Azar, C.G., Cantor, A.B., and Brodeur, G.M. (1993). Association between high-levels of expression of the TRK gene and favorable outcome in human neuroblastoma. N. Engl. J. Med. *328*, 847–854.

Nakagawara, A., Milbrandt, J., Muramatsu, T., Deuel, T.F., Zhao, H., Cnaan, A., and Brodeur, G.M. (1995). Differential expression of pleiotrophin and midkine in advanced neuroblastomas. Cancer Res. *55*, 1792–1797.

Noguchi, T., Akiyama, K., Yokoyama, M., Kanda, N., Matsunaga, T., and Nishi, Y. (1996). Amplification of a DEAD box gene (DDX1) with the MYCN gene in neuroblastomas as a result of cosegregation of sequences flanking the MYCN locus. Genes Chromosomes Cancer *15*, 129–133.

Ntzani, E.E., and Ioannidis, J.P. (2003). Predictive ability of DNA microarrays for cancer outcomes and correlates: an empirical assessment. Lancet *362*, 1439–1444.

Oba, S., Takemasa, N., Monden, M., Matsubara, K., and Ishii, S. (2003). A Bayesian missing value estimation method. Bioinformatics *19*, 2088–2096.

Ohira, M., Morohashi, A., Inuzuka, H., Shishikura, T., Kawamoto, T., Kageyama, H., Nakamura, Y., Isogai, E., Takayasu, H., Sakiyama, S., et al. (2003a). Expression profiling and characterization of 4200 genes cloned from primary neuroblastomas: identification of 305 genes differentially expressed between favorable and unfavorable subsets. Oncogene *22*, 5525–5536.

Ohira, M., Morohashi, A., Nakamura, Y., Isogai, E., Furuya, K., Hamano, S., Machida, T., Aoyama, M., Fukumura, M., Miyazaki, K., et al. (2003b). Neuroblastoma oligo-capping cDNA project: toward the understanding of the genesis and biology of neuroblastoma. Cancer Lett. *197*, 63–68.

Sawada, T., Hirayama, M., Nakata, T., Takeda, T., Takasugi, N., Mori, T., Maeda, K., Koide, R., Hanawa, Y., Tsunoda, A., et al. (1984). Mass screening for neuroblastoma in infants in Japan. Interim report of a mass screening study group. Lancet *2*, 271–273.

Schwab, M., Alitalo, K., Klempnauer, K.H., Varmus, H.E., Bishop, J.M., Gilbert, F., Brodeur, G., Goldstein, M., and Trent, J. (1983). Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumour. Nature *305*, 245–248.

Shimada, H., Chatten, J., Newton, W.A., Sachs, N., Hamoudi, A.B., Chiba, T., Marsden, H.B., and Misugi, K. (1984). Histopathologic prognostic factors in neuroblastic tumors; definition of subtypes of ganglioneuroblastoma and an age-linked classification of neuroblastomas. J. Natl. Cancer Inst. *73*, 405–416.

Shimono, R., Matsubara, S., Takamatsu, H., Fukushige, T., and Ozawa, M. (2000). The expression of cadherins in human neuroblastoma cell lines and clinical tumors. Anticancer Res. *20*, 917–923.

Slavc, I., Ellenbogen, R., Jung, W.H., Vawter, G.F., Kretschmar, C., Grier, H., and Korf, B.R. (1990). myc gene amplification and expression in primary human neuroblastoma. Cancer Res. *50*, 1459–1463.

Storey, J.D., and Tibshirani, R. (2003). Statistical significance for genome-wide studies. Proc. Natl. Acad. Sci. USA *100*, 9440–9445.

Takahashi, M., Seki, N., Ozaki, T., Kato, M., Kuno, T., Nakagawa, T., Watanabe, K., Miyazaki, K., Ohira, M., Hayashi, S., et al. (2002). Identification of the p33(ING1)-regulated genes that include cyclin B1 and proto-oncogene DEK by using cDNA microarray in a mouse mammary epithelial cell line NMuMG. Cancer Res. *62*, 2203–2209.

Ueda, K. (2001). Detection of the retinoic acid-regulated genes in a RTBM1 neuroblastoma cell line using cDNA microarray. Kurume Med. J. *48*, 159–164.

van 't Veer, L.J., Dai, H., van de Vijver, M.J., He, Y.D., Hart, A.A., Mao, M., Peterse, H.L., van der Kooy, K., Marton, M.J., Witteveen, A.T., et al. (2002). Gene expression profiling predicts clinical outcome of breast cancer. Nature *415*, 530–536.

Yamanaka, Y., Hamazaki, Y., Sato, Y., Ito, K., Watanabe, K., Heike, T., Nakahata, T., and Nakamura, Y. (2002). Maturational sequence of neuroblastoma revealed by molecular analysis on cDNA microarrays. Int. J. Oncol. *21*, 803–807.

Yoshikawa, T., Nagasugi, Y., Azuma, T., Kato, M., Sugano, S., Hashimoto, K., Masuho, Y., Muramatsu, M., and Seki, N. (2000). Isolation of novel mouse genes differentially expressed in brain using cDNA microarray. Biochem. Biophys. Res. Commun. *275*, 532–537.

Accession numbers

Microarray data are available at NCBI Gene Expression Omnibus (accession number GSE2283).

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 280, No. 17, Issue of April 29, pp. 16665–16675, 2005
Printed in U.S.A.

Identification of Protein Kinase A Catalytic Subunit β as a Novel Binding Partner of p73 and Regulation of p73 Function*

Received for publication, December 20, 2004, and in revised form, January 31, 2005
Published, JBC Papers in Press, February 21, 2005, DOI 10.1074/jbc.M414323200

Takayuki Hanamoto‡§, Toshinori Ozaki‡, Kazushige Furuya‡, Mitsuchika Hosoda‡, Syunji Hayashi‡, Mitsuru Nakanishi‡, Hideki Yamamoto‡, Hironobu Kikuchi‡, Satoru Todo§, and Akira Nakagawara‡¶

From the ‡Division of Biochemistry, Chiba Cancer Center Research Institute, Chiba 260-8717, Japan and the §Department of General Surgery, Hokkaido University School of Medicine, Kita-ku, Sapporo 060-8638, Japan

Post-translational modifications play a crucial role in regulation of the protein stability and pro-apoptotic function of p53 as well as its close relative p73. Using a yeast two-hybrid screening based on the Sos recruitment system, we identified protein kinase A catalytic subunit β (PKA-Cβ) as a novel binding partner of p73. Co-immunoprecipitation and glutathione *S*-transferase pull-down assays revealed that p73α associated with PKA-Cβ in mammalian cells and that their interaction was mediated by both the N- and C-terminal regions of p73α. In contrast, p53 failed to bind to PKA-Cβ. *In vitro* phosphorylation assay demonstrated that glutathione *S*-transferase-p73α-(1–130), which has one putative PKA phosphorylation site, was phosphorylated by PKA. Enforced expression of PKA-Cβ resulted in significant inhibition of the transactivation function and pro-apoptotic activity of p73α, whereas a kinase-deficient mutant of PKA-Cβ had no detectable effect. Consistent with this notion, treatment with H-89 (an ATP analog that functions as a PKA inhibitor) reversed the dibutyryl cAMP-mediated inhibition of p73α. Of particular interest, PKA-Cβ facilitated the intramolecular interaction of p73α, thereby masking the N-terminal transactivation domain with the C-terminal inhibitory domain. Thus, our findings indicate a PKA-Cβ-mediated inhibitory mechanism of p73 function.

p73 has been identified as a structural and functional homolog of the tumor suppressor p53 (1). p53 and p73 share the same domain organization, consisting of an N-terminal transactivation domain, a central sequence-specific DNA-binding domain, and a C-terminal oligomerization domain. As expected, several pieces of evidence suggest that p73 can bind to the p53-responsive element and transactivate an overlapping set of p53 target genes, thus leading to induction of G_1/S cell cycle arrest and apoptosis (1–6). In marked contrast to p53, p73 is expressed as multiple isoforms arising from alternative splicing of the primary p73 transcript (p73α, p73β, p73γ, p73δ, p73ε, p73η, and p73ζ) termed the TA variant (1, 3, 7–9). These alternatively spliced isoforms vary in their C termini and display different transcriptional and biological properties. Additionally, the ΔN variant (ΔNp73α and ΔNp73β), which is generated by alternative promoter utilization, lacks the N-terminal transactivation domain and exhibits dominant-negative behavior toward wild-type p73 as well as p53 (10–12). Recently, we (14) and others (13, 15) demonstrated that p73 directly transactivates the expression of its own negative regulator (ΔNp73), creating an autoregulatory feedback loop in which both the activity of p73 and the expression of ΔNp73 are regulated. Thus, the pro-apoptotic activity of p73 is determined by the relative expression levels of its TAp73 and dominant-negative ΔNp73 variants in cells.

In sharp contrast to p53, it was initially reported that p73 was not induced by DNA damage (1). However, recent studies demonstrated that, in response to a subset of DNA-damaging agents, p73 is positively regulated by multiple post-translational modifications, including phosphorylation and acetylation. During cisplatin-mediated apoptosis, phosphorylation of p73 at Tyr-99 by the non-receptor tyrosine kinase c-Abl results in an increase in its stability and pro-apoptotic activity (16–18). In addition to c-Abl, the protein kinase Cδ catalytic fragment has the ability to phosphorylate p73 at Ser-289 and contributes to the accumulation of p73 during the apoptotic response to cisplatin treatment (19). It is worth noting that the physical and functional interaction between c-Abl and protein kinase Cδ leads to the cross-activation of their kinase functions (20, 21). Furthermore, the enzymatic activity of Chk1 (checkpoint kinase-1) is enhanced in response to DNA damage (22–24), and Chk1 has the ability to phosphorylate p73 at Ser-47 upon DNA damage, thereby enhancing its transactivation ability and pro-apoptotic activity without affecting the level of total p73 protein, whereas Chk2 has no detectable effect on p73 (25). Alternatively, Zeng *et al.* (26) found that the acetyltransferase p300/CBP (cAMP-responsive element-binding protein-binding protein) interacts with the N-terminal region of p73 and stimulates p73-mediated transcriptional activation and apoptosis. Recently, Costanzo *et al.* (27) reported that doxorubicin treatment induces the p300-mediated acetylation of p73 at Lys-321, Lys-327, and Lys-331 in a c-Abl-dependent manner, which is associated with the efficient recruitment of p73 to the promoter of the apoptotic target gene *p53AIP1*. Additionally, it has been shown that p300-mediated acetylation of p73 results in its significant stabilization in a prolyl isomerase Pin1-dependent manner (28).

To identify cellular protein(s) that could interact with full-length p73α and regulate its function, we screened a human fetal brain cDNA library using a yeast two-hybrid method

* This work was supported in part by a grant-in-aid from the Ministry of Health, Labor, and Welfare for Third Term Comprehensive Control Research for Cancer; a grant-in-aid for scientific research on priority areas from the Ministry of Education, Culture, Sports, Science, and Technology of Japan; and a grant-in-aid for scientific research from the Japan Society for the Promotion of Science. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed: Div. of Biochemistry, Chiba Cancer Center Research Inst., 666-2 Nitona, Chuoh-ku, Chiba 260-8717, Japan. Tel.: 81-43-264-5431; Fax: 81-43-265-4459; E-mail: akiranak@chiba-cc.jp.