Table 2
Differentially expressed genes identified by gene expression profiling using microarray

	Genes identified	Genes on microarray	Sample number	Category
NBL diagnosis				
Khan et al. [21]	15 genes	6,567cDNAs	16 NBLs	For SRBCTs diagnosis *DPYSL4, CDH2, AF1Q, CRMP1, KIF3C, GAP43, MAP1B, RCV1, SFRP1, GATA2, PFN2, FHL1* (highly and specifically expressed in NBLs)
Expression profiling				
Yamanaka et al. [22]	78 genes	23,040 cDNAs	14 NBLs	Differentiating NBLs vs. poorly differentiated NBLs *ITGE, SYP, OLIG2, MADH2, DFFB, CASP8, CASP9* (up-regulated in differentiating NBLs) *CLDN5, CCND1, NFKBIL2* (up-regulated in poorly differentiated NBLs)
Berwanger et al. [23]	36 genes	4,608 cDNAs	94 NBLs	Stage 1 vs. stage 4 *MYCN* non-amplified tumors *FYN, AFAP, CTNNA1, NRCAM, tropomodulin, MARCKS* (down-regulated in advanced stage NBL)
Takita et al. [24]	3 genes	1,700 genes	20 NBLs	Stage 1 vs. stage 4 *BIRC3, CDKN2D* (up-regulated in the early-stage group) *SMARCD3* (down-regulated in the early-stage group)
Hiyama et al. [25]	123 genes	6,272 cDNAs	20 NBLs	Unfavorable vs. favorable (regressing and maturing) 43 genes including *MYCN, hTERT, NME1, CCND1, CCNE1, E1, BIRC5, BIRC1* (up-regulated in unfavorable NBLs) 80 genes (up-regulated in favorable NBLs) *CD44, IGF2, TRKA, ANK1* (highly expressed in maturing NBLs) *CASP8, CASP9, TNFSF10, NGFA, GDF10* (highly expressed in regressing NBLs)
McArdle et al. [26]		14,500 genes	20 NBLs	Differentially expressed in the 11q-, MYCN non-amplified and hyperdiploid subtypes of NBL
Janoueix-Lerosey et al.[27]		320 genes (on 1p35-36)	43 NBLs	1p loss vs. 1p-normal *CDC42, VAMP3, CLSTN1, GNB1, STMN1, RPA2, RBAF600, FBXO6, MAD2L2* (decreased expression in NBLs with 1p deletion)
Prognosis prediction				
Wei et al. [28]	19 genes	42,578 cDNA	56 NBLs	To develop an accurate predictor of survival for patient with NBL *DLK1, PRSS3, ARC, SLIT3, MYCN, JPH1* (up-regulated in the poor-outcome group) *ARHI, CNR1, CD44, ROBO2, BTBD3, KLRC3* (down-regulated in the poor-outcome group)
Ohira et al. [29]	70 genes	5,340 cDNAs	136 NBLs	To develop an accurate predictor of survival for patient with NBL

primary tumors (stage 1 versus 4) and identified that the expression of *BIRC3* and *CDKN2D* genes were significantly higher in the early-stage group than in the advanced-stage group and that the expression of the *SMARCD3* gene was significantly reduced in the early-stage group. The *BIRC3*, *CDKN2D*, and *SMARCD3* genes have been reported to be associated with apoptosis, cell cycles, and the transcriptional activator, respectively. Hiyama et al. [25] analyzed 20 NBLs with 6272 cDNAs microarray, and revealed that 43 genes, including *MYCN*, *hTERT*, *NM23-H1*, cell cycle regulatory protein-coding genes (*CCND1*, *CCNE1*, *E1*), and apoptosis-escape genes (*BIRC5*, *BIRC1*) were highly expressed in unfavorable neuroblastomas, while another 80 genes, including neuronal differentiating genes and apoptotic inducing genes (*CD44*, *IGF2*, *TRKA*, *ANK1* in maturing NBLs, *CASP8*, *CASP9*, *TNFSF10*, *NGFA*, *GDF10* in regressing NBLs) were detected as highly expressed in favorable tumors.

There have been several studies by focusing the certain genomic aberrations and corresponding gene expression profiles. McArdle et al. [26] identified transcripts that are differentially expressed in the 11q-, *MYCN* nonamplified and hyperdiploid subtypes of NBL. Janoueix-Lerosey et al. [27] compared the expression profiles between the tumors with 1p loss and those with normal 1p status and identified the genes with decreased expression in NBLs with 1p deletion (Table 2).

5. Microarray-based system for predicting prognosis of neuroblastoma patients

We now have in our hands the increasing number of information for genes which can distinguish prognosis of the patient with NBL as described above. These data should subsequently be integrated and organized to construct a simple prediction system which is practical for the clinic. Such efforts are now ongoing.

Wei et al. [28] have performed gene expression profiling of 56 NBLs using cDNA microarrays containing 42,578 cDNA clones and used artificial neural networks (ANNs) to develop an accurate predictor of survival for each individual patient with NBL. ANN-based prognosis prediction has been accomplished by using expression levels of only 19 genes including *MYCN* and *CD44*. In addition, these 19 predictor genes were able to additionally classify high-risk patients into two subgroups according to their survival status. Among these predictor genes, *DLK1*, *PRSS3*, *ARC*, *SLIT3*, *MYCN* and *JPH1* were up-regulated in the poor-outcome group included, whereas *ARHI*, *CNR1*, *CD44*, *ROBO2*, *BTBD3* and *KLRC3* were down-regulated.

We also have constructed an in-house, microceramic pump-based ink-jet-printed cDNA microarray carrying 5,340 genes obtained from primary NBL cDNA libraries and applied it for the analysis of 136 tumors. A probabilistic output computational analysis using learning samples has selected 70 genes which constructed a classifier for patient outcome, and provided a correct prognosis of test samples with high efficiency [29]. Of clinical interest, Kaplan-Meier analysis indicated that the classifier can divide significantly 5-year survivals of NBL patients, even for the intermediate risk type. Furthermore, our microarray prediction exhibited the best balance between sensitivity and specificity among prognostic factors including *MYCN* amplification and *TrkA* expression.

These findings provide evidence of a gene expression signature that can predict prognosis independent of currently known risk factors and could assist physicians in the individual management of patients with high-risk NBL. Such gene expression-based diagnosis system should be highly accurate and reproducible, as well as simple, easy to analyze, and with low cost. Based on this, we subsequently made a mini-chip carrying top-ranked 200 genes for clinical use. We are now investigating its (1) reproducibility from the original 5,340 genes chip, (2) potential of predicting prognosis for the newly added test samples, (3) potential of predicting prognosis when whole experimental process (RNA isolation, sample labeling, hybridization, and calculation of the probability of patient survival by using the classifier) is conducted in an independent laboratory. After confirming these, the microarray system will come to prove its practical use to be feasible in the clinic to predict the prognosis of the patient with NBL.

6. Conclusion

Using the gene expression profiling, we are now able to distinguish a group of high-risk patients with high efficiency who will not respond to conventional therapy and therefore require alternative treatment strategies. Although further prospective studies will be necessary, we may also be able to reduce the toxicity of treatment regime for the patients who have been predicted to survive according to gene expression profile. Recently, in addition to the gene expression profiling, certain genome aberrations as well as epigenetic alterations have been reported to be strongly related to the patient prognosis with NBL. Abe et al. [30] indicated that poor NBLs suffer the increased methylation pressure in their tumor genome, and that the methylation of certain CpG island, such as *Protocadherin beta* family, can predict poor NBL prognosis with high sensitivity. Furthermore, screening of prognosis-related proteins secreted into serum of NBL patients has been started by many groups. Combination of such independent systems to predict

prognosis will further improve the accuracy of the diagnostic system for NBL.

Acknowledgements

This work was supported in part by a fund from Hisamitsu Pharmaceutical Co., Inc. and by Grant-in-Aids 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 Aids for Cancer Research and for the 2nd Term Comprehensive 10-year Strategy for Cancer Control from the Ministry of Health, Labour and Welfare of Japan.

References

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

[2] J.L. Weinstein, H.M. Katzenstein, S.L. Cohn, Advances in the diagnosis and treatment of neuroblastoma, Oncologist 8 (2003) 278–292.

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

[4] A. Nakagawara, C.G. Azar, N.J. Scavarda, G.M. Brodeur, Expression and function of TRK-B and BDNF in human neuroblastomas, Mol. Cell Biol. 14 (1994) 759–767.

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

[6] M.C. Favrot, V. Combaret, C. Lasset, CD44–a new prognostic marker for neuroblastoma, N. Engl. J. Med. 1993; 329.

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

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

[9] T. Tanaka, T. Sugimoto, T. Sawada, Prognostic discrimination among neuroblastomas according to Ha-ras/trk A gene expression: a comparison of the profiles of neuroblastomas detected clinically and those detected through mass screening, Cancer 83 (1998) 1626–1633.

[10] T. Nagata, Y. Takahashi, S. Asai, Y. Ishii, H. Mugishima, T. Suzuki, et al., The high level of hCDC10 gene expression in neuroblastoma may be associated with favorable characteristics of the tumor, J. Surg. Res. 92 (2000) 267–275.

[11] A. Islam, H. Kageyama, N. Takada, T. Kawamoto, H. Takayasu, E. Isogai, et al., High expression of Survivin, mapped to 17q25, is significantly associated with poor prognostic factors and promotes cell survival in human neuroblastoma, Oncogene 19 (2000) 617–623.

[12] N. Hailat, D.R. Keim, R.F. Melhem, X.X. Zhu, C. Eckerskorn, G.M. Brodeur, et al., High levels of p19/nm23 protein in neuroblastoma are associated with advanced stage disease and with N-myc gene amplification, J. Clin. Invest. 88 (1991) 341–345.

[13] A. Leone, R.C. Seeger, C.M. Hong, Y.Y. Hu, M.J. Arboleda, G.M. Brodeur, et al., Evidence for nm23 RNA overexpression, DNA amplification and mutation in aggressive childhood neuroblastomas, Oncogene 8 (1993) 855–865.

[14] F. Saito-Ohara, I. Imoto, J. Inoue, H. Hosoi, A. Nakagawara, T. Sugimoto, J. Inazawa, PPM1D is a potential target for 17q gain in neuroblastoma, Cancer Res. 63 (2003) 1876–1883.

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

[16] M. Ohira, A. Morohashi, H. Inuzuka, T. Shishikura, T. Kawamoto, H. Kageyama, et al., Expression profiling and characterization of 4200 genes cloned from primary neuroblastomas: identification of 305 genes differentially expressed between favorable and unfavorable subsets, Oncogene 22 (2003) 5525–5536.

[17] T. Kawamoto, M. Ohira, S. Hamano, T. Hori, A. Nakagawara, High expression of the novel endothelin-converting enzyme genes, Nbla03145/ECEL1alpha and beta, is associated with favorable prognosis in human neuroblastomas, Int. J. Oncol. 22 (2003) 815–822.

[18] S. Hamano, M. Ohira, E. Isogai, K. Nakada, A. Nakagawara, Identification of novel human neuronal leucine-rich repeat (hNLRR) family genes and inverse association of expression of Nbla10449/hNLRR-1 and Nbla10677/hNLRR-3 with the prognosis of primary neuroblastomas, Int. J. Oncol. 24 (2004) 1457–1466.

[19] C. Kato, K. Miyazaki, A. Nakagawa, M. Ohira, Y. Nakamura, T. Ozaki, et al., Low expression of human tubulin tyrosine ligase and suppressed tubulin tyrosination/detyrosination cycle are associated with impaired neuronal differentiation in neuroblastomas with poor prognosis, Int. J. Cancer 112 (2004) 365–375.

[20] S.S. Roberts, M. Mori, P. Pattee, J. Lapidus, R. Mathews, J.P. O'Malley, et al., GABAergic system gene expression predicts clinical outcome in patients with neuroblastoma, J. Clin. Oncol. 22 (2004) 4127–4134.

[21] J. Khan, J.S. Wei, M. Ringner, L.H. Saal, M. Ladanyi, F. Westermann, et al., Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks, Nat. Med. 7 (2001) 673–679.
[22] Y. Yamanaka, Y. Hamazaki, Y. Sato, K. Ito, K. Watanabe, T. Heike, et al., Maturational sequence of neuroblastoma revealed by molecular analysis on cDNA microarrays, Int. J. Oncol. 21 (2002) 803–807.
[23] B. Berwanger, O. Hartmann, E. Bergmann, S. Bernard, D. Nielsen, M. Krause, et al., Loss of a FYN-regulated differentiation and growth arrest pathway in advanced stage neuroblastoma, Cancer Cell 2 (2002) 377–386.
[24] J. Takita, M. Ishii, S. Tsutsumi, Y. Tanaka, K. Kato, Y. Toyoda, et al., Gene expression profiling and identification of novel prognostic marker genes in neuroblastoma, Genes Chromosomes Cancer 40 (2004) 120–132.
[25] E. Hiyama, K. Hiyama, H. Yamaoka, T. Sueda, C.P. Reynolds, T. Yokoyama, Expression profiling of favorable and unfavorable neuroblastomas, Pediatr. Surg. Int. 20 (2004) 33–38.
[26] L. McArdle, M. McDermott, R. Purcell, D. Grehan, A. O'Meara, F. Breatnach, et al., Oligonucleotide microarray analysis of gene expression in neuroblastoma displaying loss of chromosome 11q, Carcinogenesis 25 (2004) 1599–1609.
[27] I. Janoueix-Lerosey, E. Novikov, M. Monteiro, N. Gruel, G. Schleiermacher, B. Loriod, et al., Gene expression profiling of 1p35-36 genes in neuroblastoma, Oncogene 23 (2004) 5912–5922.
[28] J.S. Wei, B.T. Greer, F. Westermann, S.M. Steinberg, C.G. Son, Q.R. Chen, et al., Prediction of clinical outcome using gene expression profiling and artificial neural networks for patients with neuroblastoma, Cancer Res. 64 (2004) 6883–6891.
[29] M. Ohira, S. Oba, Y. Nakamura, E. Isogai, S. Kaneko, A. Nakagawa, et al., Expression profiling using a tumor-specific cDNA microarray prediols the prognosis of intermediate-risk neuroblastomas, Cancer Cell 7 (2005) 337-350.
[30] M. Abe, M. Ohira, A. Kaneda, Y. Yagi, S. Yamamoto, Y. Kitano, et al., CpG island methylator phenotype is a strong determinant of poor prognosis in neuroblastomas, Cancer Res., 65 (2005) 828–834.

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-

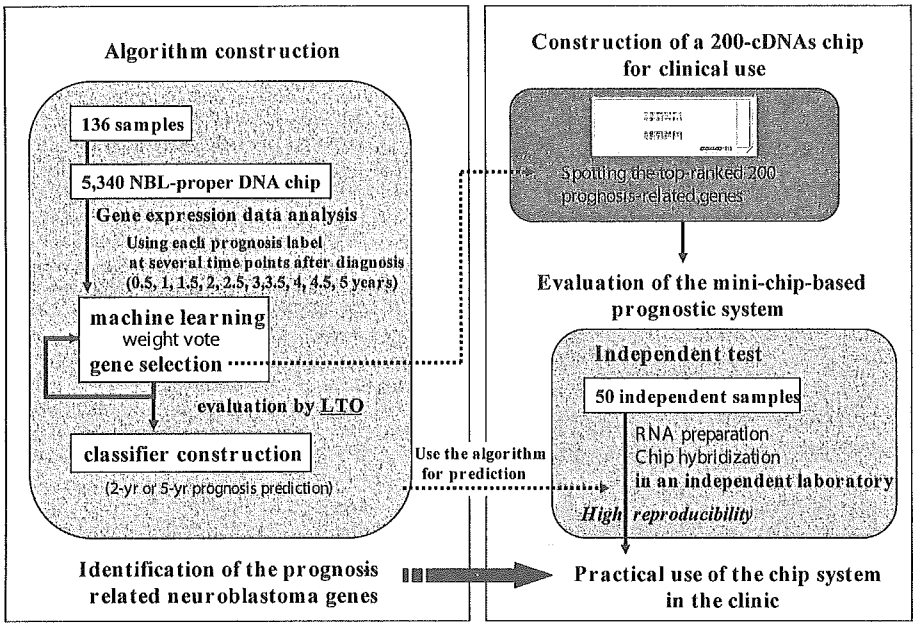

Figure 1. Schematic diagram of this study

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*)

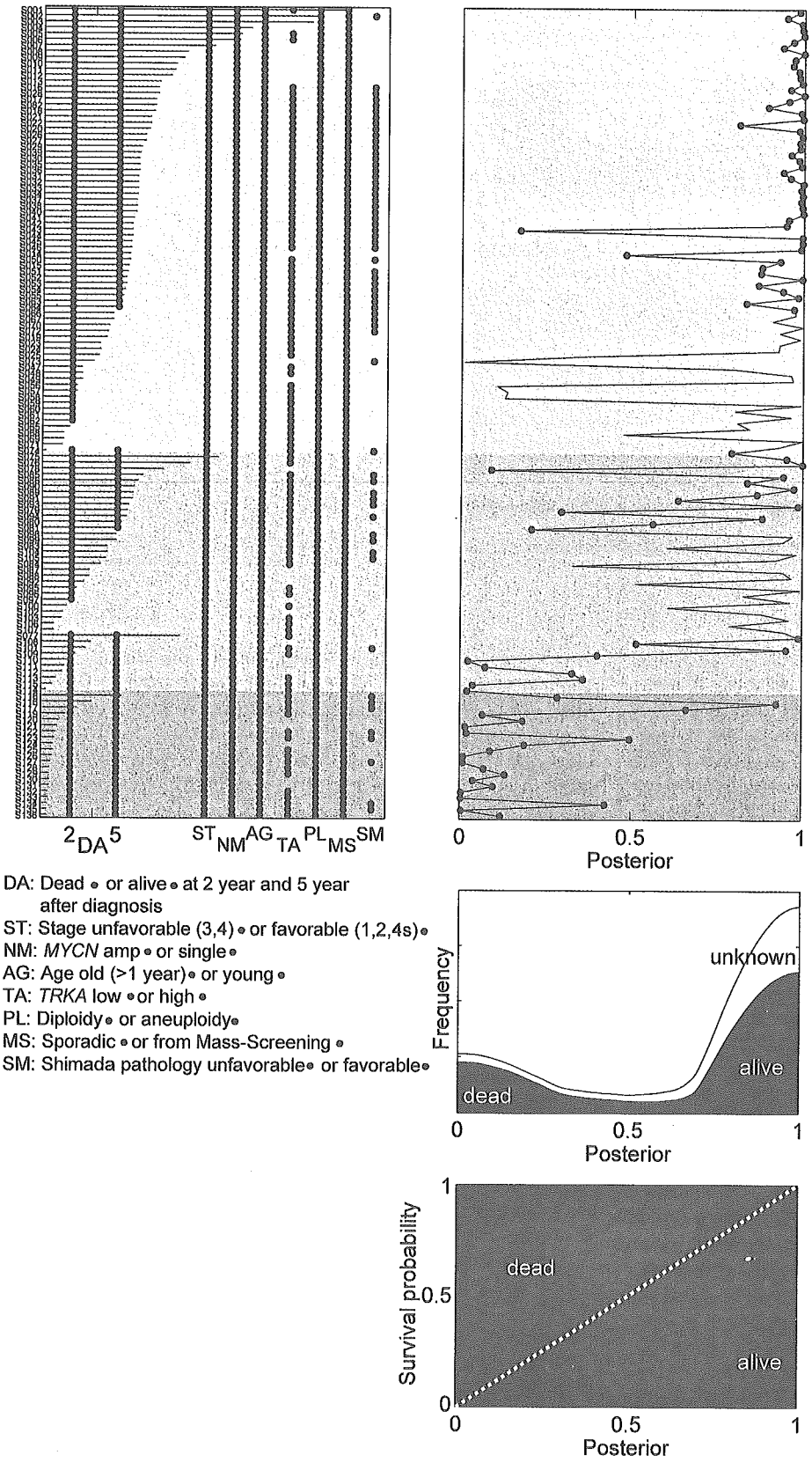

Figure 2. Posterior probability of survival at 5 years

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.

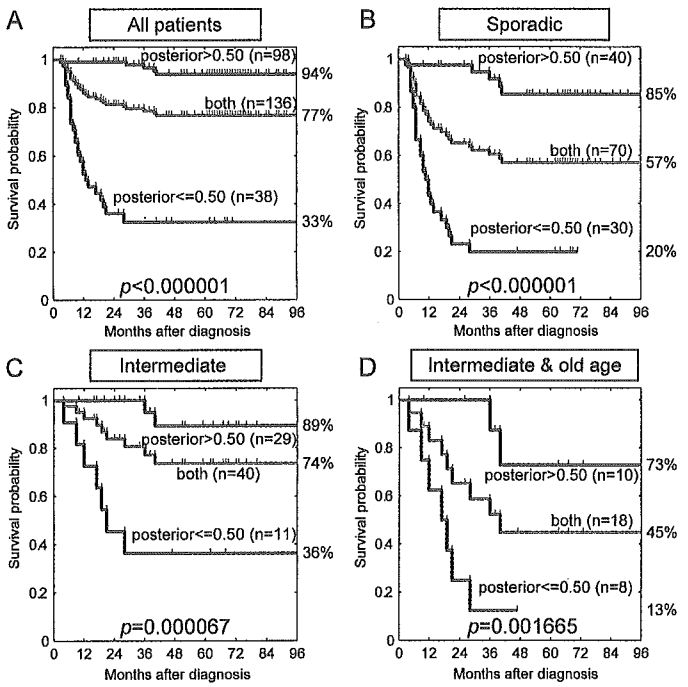

Figure 3. Disease-free survival of patients who were stratified based on the gene expression patterns

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

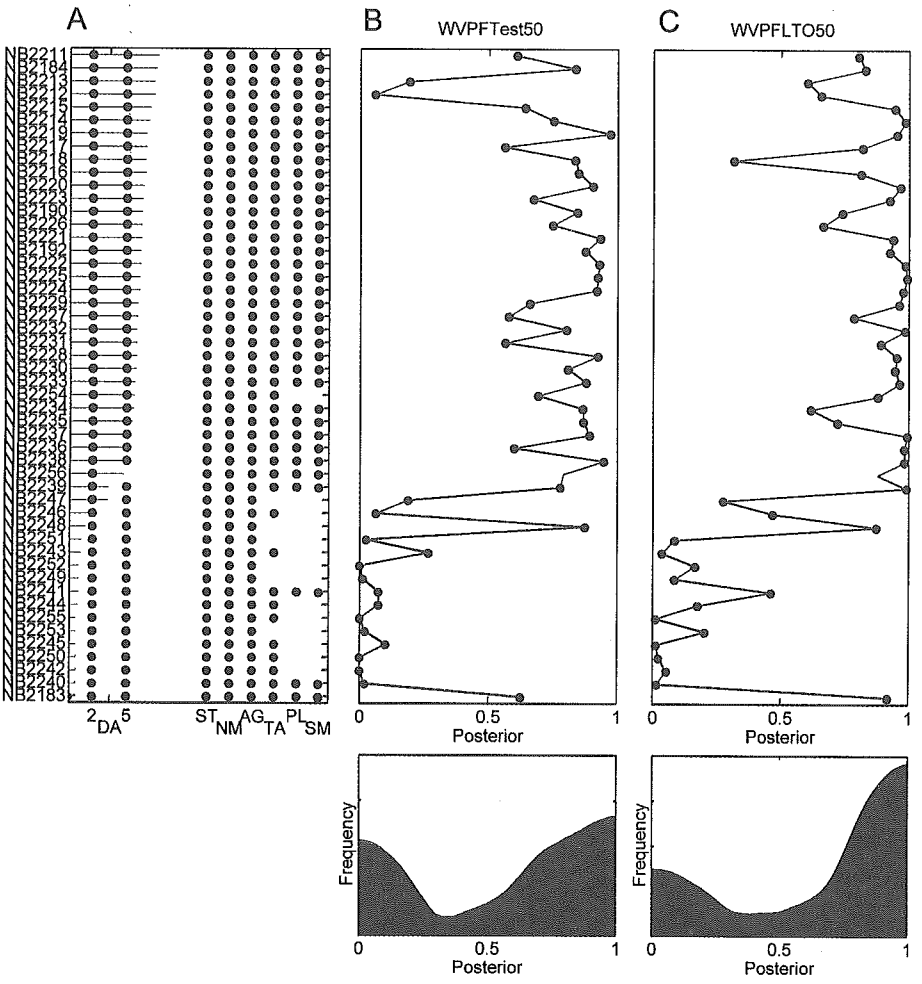

Figure 4. Posterior probability of survival at 5 years for test samples

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-

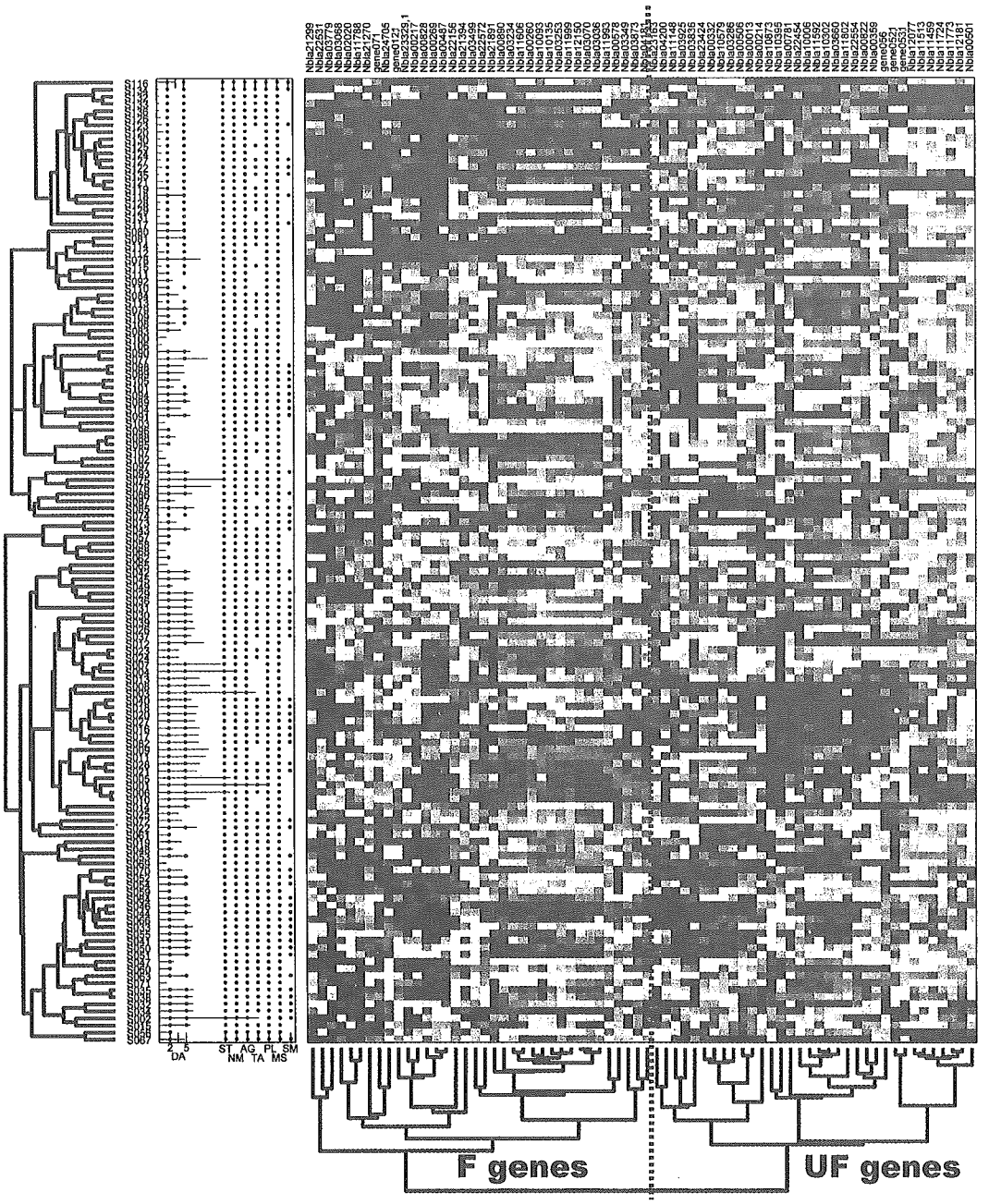

Figure 5. Expression profiles of 70 genes selected for predicting neuroblastoma prognosis at 2 years

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).

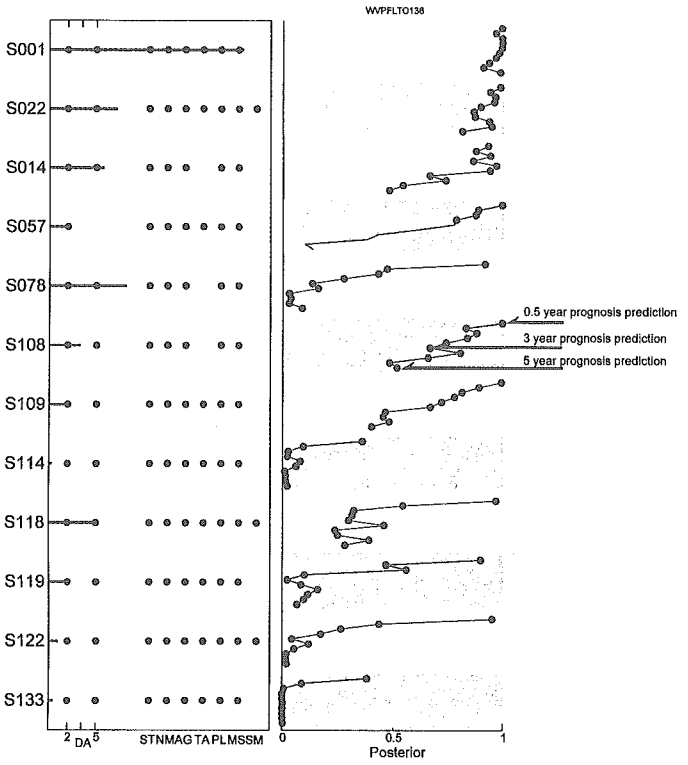

Figure 6. Individual survival chance analysis based on posterior probabilities

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).

VOLUME 23 · NUMBER 22 · AUGUST 1 2005

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Prediction of *MYCN* Amplification in Neuroblastoma Using Serum DNA and Real-Time Quantitative Polymerase Chain Reaction

Takahiro Gotoh, Hajime Hosoi, Tomoko Iehara, Yasumichi Kuwahara, Shinya Osone, Kunihiko Tsuchiya, Miki Ohira, Akira Nakagawara, Hiroshi Kuroda, and Tohru Sugimoto

From the Department of Pediatrics, Kyoto Prefectural University of Medicine, Graduate School of Medical Science, Kyoto; and Division of Biochemistry, Chiba Cancer Center Research Institute, Chiba, Japan.

Submitted October 6, 2004; accepted April 13, 2005.

Supported by Grants-in-Aid for Scientific Research grant Nos. 15659248, 15659249, and 14370250 from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Authors' disclosures of potential conflicts of interest are found at the end of this article.

Address reprint requests to Takahiro Gotoh, MD, PhD, Department of Pediatrics, Kyoto Prefectural University of Medicine, Graduate School of Medical Science, Kamigyo-ku, Kyoto 602-8566, Japan; e-mail: takahiro_email:@yahoo.co.jp.

0732-183X/05/2322-5205/$20.00

DOI: 10.1200/JCO.2005.02.014

ABSTRACT

Purpose

MYCN amplification (MNA) indicates a poor prognosis in neuroblastoma (NB) and is routinely assayed for therapy stratification. We aimed to develop a diagnostic tool to predict *MYCN* status using serum DNA, which, in cancer patients, predominantly originates from tumor-released DNA.

Patients and Methods

Using DNA-based real-time quantitative polymerase chain reaction, we simultaneously quantified *MYCN* (2p24) and a reference gene, *NAGK* (2p12), and evaluated *MYCN* copy number as an *MYCN/NAGK* (*M/N*) ratio in 87 NB patients whose *MYCN* status had been determined by Southern blotting. Of these patients, 17 had *MYCN*-amplified NB, and 70 had nonamplified NB.

Results

The serum *M/N* ratio in the MNA group (median, 199.32; range, 17.1 to 901.6; 99% CI, 107.0 to 528.7) was significantly ($P < .001$) higher than the ratio in the non-MNA group (median, 0.87; range, 0.25 to 4.6; 99% CI, 0.82 to 1.26; Mann-Whitney *U* test). The sensitivity and specificity of the serum *M/N* ratio as a diagnostic test were both 100% when the serum *M/N* ratio cutoff was set at 10.0. Among six MNA patients whose clinical courses were followed, the serum ratios decreased to the normal range in the patients in remission (n = 3), whereas the ratios increased to high levels in the patients who relapsed (n = 2) or failed to achieve remission (n = 1).

Conclusion

Measurement of the serum *M/N* ratio seems to be a promising method for accurately assessing *MYCN* status in NB, although a larger set of patients needs to be examined to confirm this result.

J Clin Oncol 23:5205-5210. © 2005 by American Society of Clinical Oncology

INTRODUCTION

Neuroblastoma (NB) is the most common extracranial solid tumor in children and is characterized by a wide range of clinical behaviors, from spontaneous regression to rapid progression with a fatal outcome. The clinical heterogeneity has been reported to be associated with a variety of biologic features of NB. One such aberration, *MYCN* amplification (MNA; ie, creation of multiple copies of the *MYCN* gene in the nuclei of tumor cells), is strongly associated with rapid tumor progression and a poor outcome. MNA is detected in 4% of patients in the early stages of NB, 8% of patients in stage 4S, and approximately 30% of patients in advanced stages. Currently, assessment of *MYCN* status is essential for determining therapy stratification in NB.[1-6] Having rapid access to selected biologic data for each tumor has become increasingly important in

routing patients to appropriate therapies. Several years ago, fluorescence in situ hybridization (FISH) replaced Southern blotting as the most accurate and timely way of evaluating tumors for MNA. Using FISH, the turnaround time for results was shortened from weeks to days, making its use in clinical trials realistic.

In this study, we describe a real-time polymerase chain reaction (PCR) method for evaluating *MYCN* status that shortens the turnaround for results to just a few hours. Furthermore, to facilitate the evaluation of *MYCN* status of tumors, we used serum DNA for the PCR template, which, in cancer patients, predominantly consists of tumor-released DNA.[7] Quantification of serum DNA has also been proposed as a screening tool for early detection of lung cancer.[8] Several groups were able to detect tumor-related aberrations, such as loss of heterozygosity and mutations in the *p53* gene, using the serum DNA of patients with a malignant tumor.[9-11]

Recently, Combaret et al[12] reported that high levels of *MYCN* DNA were present in the peripheral blood of patients with *MYCN*-amplified NB. However, they evaluated serum *MYCN* (2p24) dosage based on PCR without a reference gene, so their assay could be influenced by the quality of the template DNA or a numerical change of chromosome 2. To avoid these problems, we used DNA-based real-time quantitative PCR and a single copy reference gene, the *N-acetylglucosamine kinase* gene (*NAGK*; 2p12), so that *MYCN* copy number per chromosome 2 could be evaluated as the *MYCN/NAGK* (*M/N*) ratio. *NAGK* was chosen because it is on the same chromosome as *MYCN* but sufficiently distant from the region spanned by the *MYCN* amplicon (2p12 *v* 2p24)[13] that a numerical change in chromosome 2 would not affect the *M/N* ratio. The diagnostic performance of the test was evaluated in patients with an NB whose *MYCN* status had been determined by Southern blotting.

PATIENTS AND METHODS

Subjects

Eighty-seven patients diagnosed with NB at the Hospital of the Kyoto Prefectural University of Medicine and Chiba Cancer Center Research Institute were enrolled onto this study with the informed consent of their parents. The studies were conducted under research protocols approved by each institutional review board. At the time of diagnosis, 44 patients were younger than 1 year, and 43 were between 1 and 13 years of age. Seventeen of the patients had MNA, and 70 patients did not have MNA, as determined by Southern blotting. According to the International Neuroblastoma Staging System,[4] the 17 children with MNA included one patient each in stage 1 and 2B, two in stage 3, and 13 in stage 4, whereas the 70 children without MNA included 22 in stage 1, 18 in stage 2A and 2B, five in stage 4S, seven in stage 3, and 18 in stage 4.

Twelve of the 17 patients with MNA and 33 of the 70 nonamplified patients were also analyzed by dual-color FISH technique, as previously described,[14] using an *MYCN* probe (pNb101) and a chromosome 2 centromere probe (D2Z). FISH results of these patients were consistent with the Southern blotting results, although three of the patients who were diagnosed as non-MNA by Southern blotting were found to have one to four extra copies of the *MYCN* gene relative to the chromosome 2 centromere number by FISH. This low level of amplification has been defined as *MYCN* gain, which is an intermediate stage between MNA and non-MNA.[15] Because the prognostic significance of *MYCN* gain is still unclear, these patients were classified as non-MNA according to the Southern blotting results.

Sample Preparation

Tumor specimens were surgically resected and immediately stored at −80°C. Peripheral blood was obtained from each patient before any therapy and surgery. To avoid contamination of serum DNA by the DNA from WBCs, we prepared serum exclusively from the liquid fraction of clotted blood after centrifugation at 1,000 × *g* for 10 minutes and stored it at −20°C until DNA extraction.

DNA Isolation

DNA was extracted from tissues and serum samples by using the QIAmp tissue and blood kits (Qiagen, GmbH, Hilden, Germany), respectively, according to the manufacturer's protocols. For each patient, 200 μL of the stored serum was used for extraction of free DNA.

Real-Time Quantitative PCR

TaqMan PCR was performed using the ABI Prism 5700 Sequence Detection System (Applied Biosystems, Foster City, CA). The PCR mixture contained TaqMan universal PCR master mix (Applied Biosystems), 200 nmol/L of each primer, and 100 nmol/L of fluorogenic probe. The principle of the TaqMan analysis has been described previously in detail.[16-18] In addition to the *MYCN* sequence, *NAGK* (GenBank accession No. NM 017567) located at 2p12 was simultaneously measured as a single-copy reference gene. The sequence of primers and the TaqMan probe used for *MYCN* and *NAGK* are as follows: *MYCN* forward, 5′-GTGCTCTCCAATTCTCGCCT-3′; *MYCN* reverse, 5′-GATGGCCTAGAGGAGGGCT-3′; *MYCN* probe, 5′-FAM-CACTAAAGTTCCTTCCACCCTCTCCT-TAMRA-3′; *NAGK* forward, 5′-TGGGCAGACACATCGTAGCA-3′; *NAGK* reverse, 5′-CACCTTCACTCCCACCTCAAC-3′; and *NAGK* probe, 5′-VIC-TGTTGCCCGAGATTGACCCGGT-TAMRA-3′. All PCR reactions were performed with one cycle of 95°C for 5 minutes, followed by PCR amplification with 50 cycles of 95°C for 15 seconds and 60°C for 1 minute. Standard curves were constructed in each PCR run with four-fold serial dilutions containing 20, 5, 1.25, 0.3125, and 0.078125 ng/μL of a healthy donor's DNA in addition to 20 ng/μL of salmon sperm DNA, and the dosages of the target genes in each sample were interpolated using these standard curves. The *MYCN* copy number of a sample of DNA was determined by the ratio of the *MYCN* dosage to the *NAGK* dosage (*M/N* ratio). Copy numbers were expressed as the average of two measurements.

Effect of WBC Contamination

To assess the effect of WBC contamination in serum samples on the serum *M/N* ratio, we measured the serum *M/N* ratio using DNA extracted from a series of WBC-contaminated serum samples. The samples were prepared by adding 0, 1×10, 1×10^2, 1×10^3, 1×10^4, and 1×10^5 of WBCs from a healthy donor to 200 μL of serum from a *MYCN*-amplified patient.