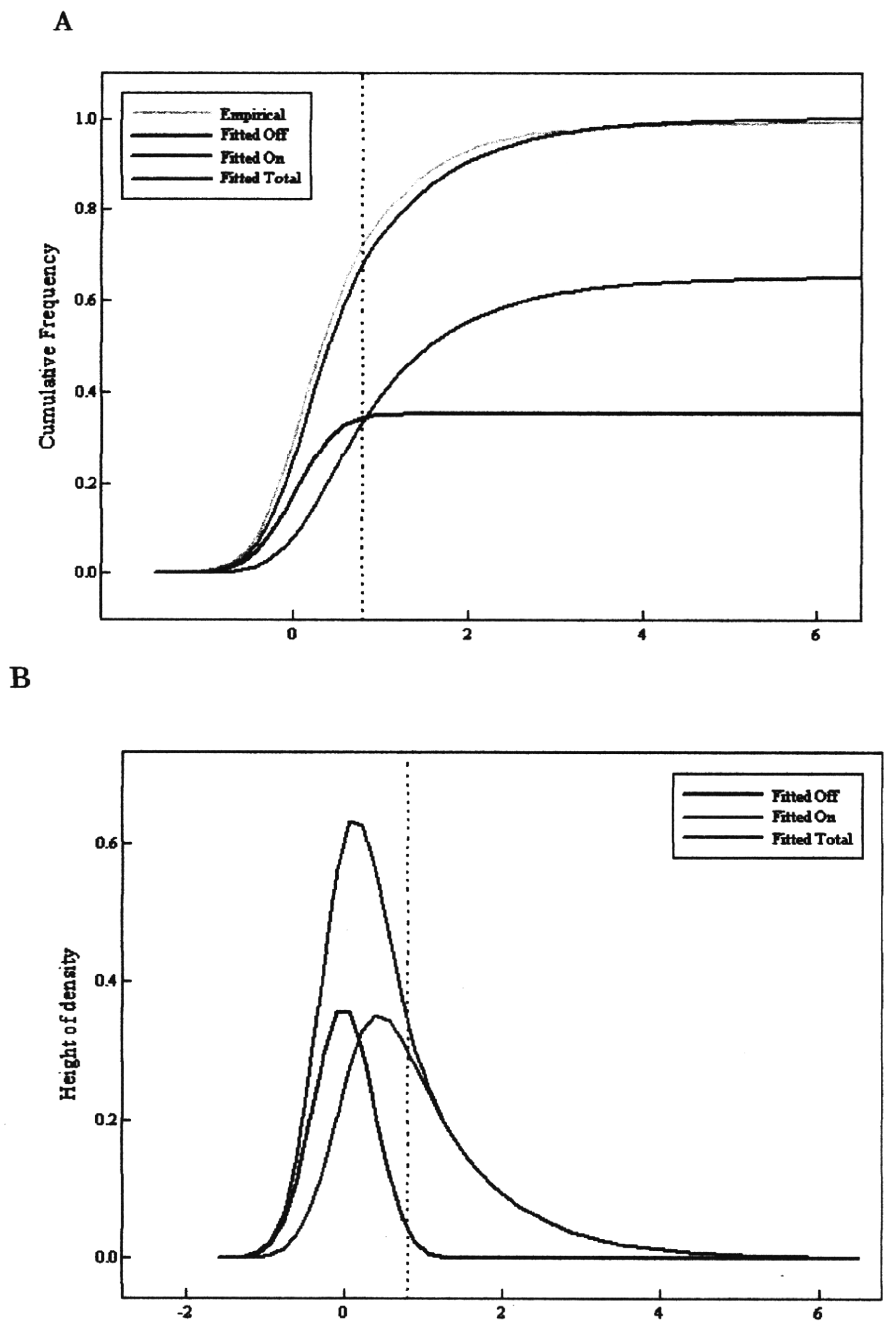



**Figure 3 Results of applying the Weibull-normal mixture model to the Human Genome U95 data sets. (A)** Comparison between the fitted Weibull-normal distribution (green line) with two components ('On' and 'Off') and the empirical distribution (brown line). The pink line shows the component of expression intensities of 'On' genes and the blue line that of 'Off' genes. **(B)** Density function corresponding to the fitted Weibull-normal distribution function (green line) and its components (pink and blue lines).

**Table 1: Comparison between Detection Call and On/Off method using spiked-in genes.**

| | Detection Call | | | On/Off | | |
|---|---|---|---|---|---|---|
| True concentration (pM) | Present | Marginal | Absent | $X \geq 11$ | $X \leq 10$ | Total |
| 0 | 1 | 0 | 58 | 4 | 55 | 59 |
| 0.25 | 10 | 0 | 49 | 27 | 32 | 59 |
| 0.5 | 21 | 4 | 33 | 44 | 14 | 58 |
| 1.0 | 22 | 1 | 36 | 48 | 11 | 59 |
| 2 | 37 | 3 | 19 | 52 | 7 | 59 |
| 4 | 48 | 2 | 9 | 56 | 3 | 59 |
| 8 | 58 | 1 | 0 | 59 | 0 | 59 |
| 16 | 59 | 0 | 0 | 59 | 0 | 59 |
| 32 | 59 | 0 | 0 | 59 | 0 | 59 |
| 64 | 50 | 0 | 0 | 50 | 0 | 50 |
| 128 | 50 | 0 | 0 | 50 | 0 | 50 |
| 256 | 59 | 0 | 0 | 59 | 0 | 59 |
| 512 | 68 | 0 | 0 | 68 | 0 | 68 |
| 1024 | 69 | 0 | 0 | 69 | 0 | 69 |
| Total | 611 | 11 | 204 | 704 | 122 | 826 |

6.8% false positives, although less than 4% (the boundary p-value for defining Present calls) is desirable. The sensitivities of the 'On/Off' and MAS5 methods are shown in Figure 4. The 'On/Off' and MAS5 methods require at least 0.5 pM and 2.0 pM of spike-in genes to achieve around 80% sensitivity, respectively.

### Identification of candidate genes for predicting neuroblastoma prognosis

Neuroblastoma is one of the most common solid tumors in childhood. Its prognosis varies remarkably, ranging from spontaneous regression to fatal progression [21]. We call these outcomes 'favorable' and 'unfavorable', respectively. It is well known that *MYCN* amplification strongly correlates with adverse outcome in neuroblastoma [22]. Nevertheless, whether *MYCN* expression is truly predictive of neuroblastoma outcome remains controversial [23]. We examined the relationship between *MYCN* expression and clinical outcomes. A scatter diagram of $X$ (the number of probe pairs on an *MYCN* probe set satisfying $PM > MM$) versus expression intensity of *MYCN* is shown in Figure 5 for each neuroblastoma case.

**Table 2: False positive and negative rates by MAS5 and On/Off methods.**

| | False positive rate | False negative rate |
|---|---|---|
| MAS5 | 1/59 (1.7%) | 146/767 (19.0%) |
| ON/Off | 4/59 (6.8%) | 67/767 (8.7%) |

Pink points represent cases with unfavorable outcome and blue points represent those with favorable outcome. A cross-tabulation of state of *MYCN* being 'On/Off' and outcome (favorable/unfavorable) is also shown in Figure 5, where we define the gene state as 'On' if $X \geq 7$ and 'Off' if $X \leq 6$ according to the fit of the Weibull-Normal model. The state of *MYCN* is uniformly 'Off' in the favorable group but variable--either 'On' or 'Off'--in the unfavorable group, suggesting that *MYCN* being 'On' is sufficient for unfavorable outcome and that genes other than *MYCN* are associated with poor prognosis. We then introduce a new notation, 'OR_On' type gene, which shows the logical relationship between multiple genes and binary phenotypes.

Results of applying the Weibull-Normal mixture model to 40 cases of neuroblastoma with favorable outcome and 21 cases with unfavorable outcome are shown in Figures 6A and 6B, respectively. The estimated parameter vectors for favorable and unfavorable groups were $(\mu, \alpha, \xi, \widehat{\sigma^2})$ = (1.96, 1.10, 0.17, 0.13) and (1.86, 1.16, 0.16, 0.14), respectively, where $\mu$ and $\alpha$ denote location and power parameters of the Weibull distribution, $\xi$ denotes mixture rate of 'Off' genes, and $\sigma^2$ denotes the variance of the Normal distribution. Figures 6A-1 and 6B-1 show the fitted Weibull-Normal distribution with two components ('On'

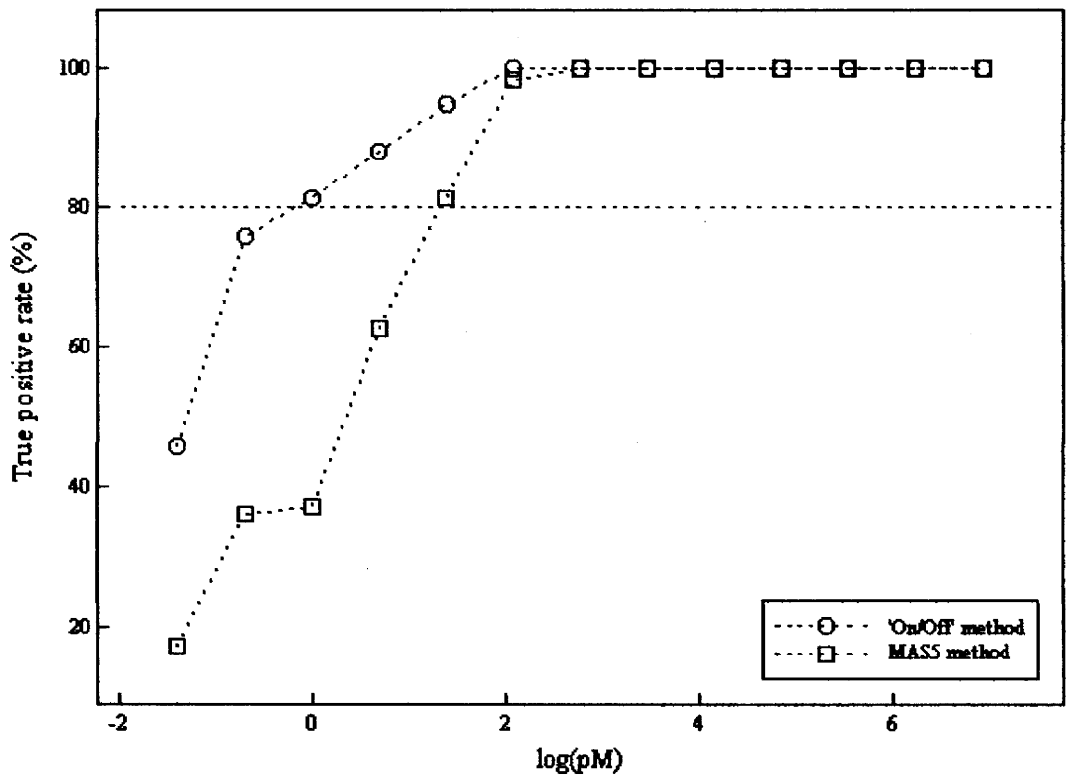


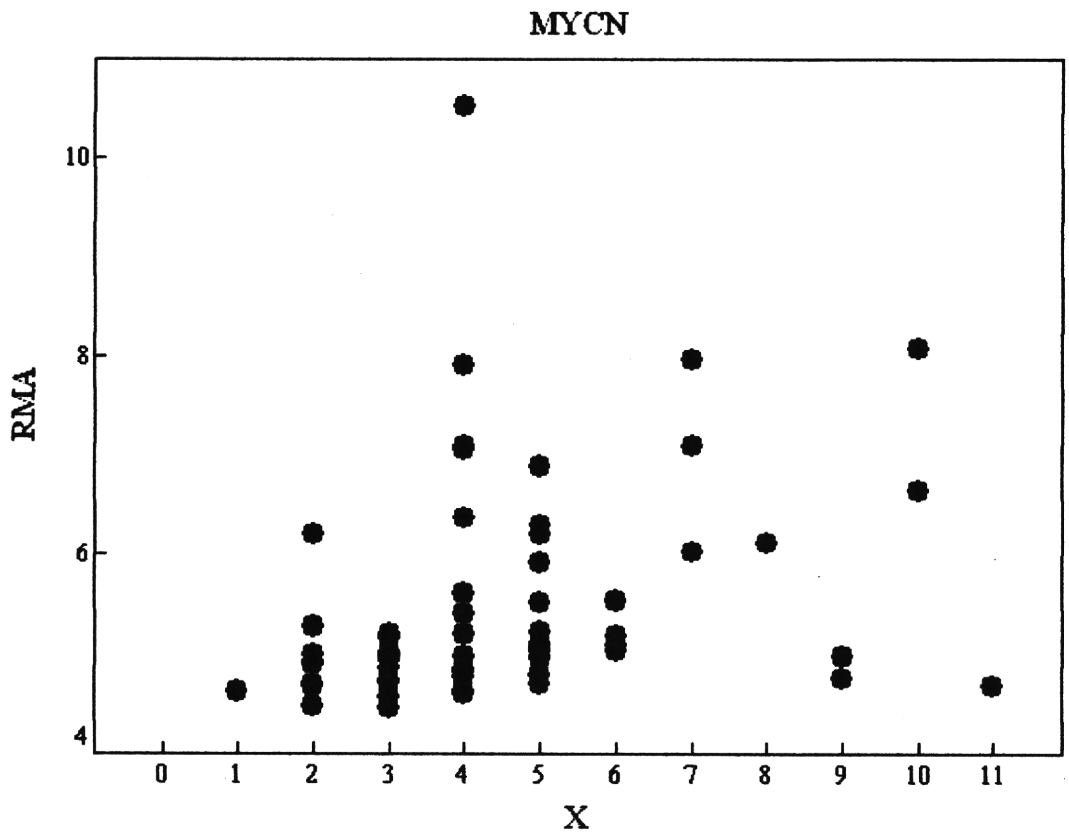

**Figure 4 Sensitivities of On/Off calls and MAS5 calls using the spike in genes.** The X-axis indicates log-transformed concentrations of spike-in genes; the Y-axis indicates true positive rates. Circles represent sensitivities by On/Off calls; triangles represent those by MAS5 calls.

and 'Off') and the empirical distribution. Figures 6A-2 and 6B-2 show the corresponding density functions and their components. The cutoff point dividing genes into 'On' and 'Off' states, $X = 7$, is denoted by the vertical dotted lines in Figures 6A and 6B.

We set out to identify OR_On type genes involved in neuroblastoma prognosis using the estimate of the probability of a gene being 'On' or 'Off' (see 'Methods' section). A hundred genes were selected as candidates from a total of 54,109 genes. The five genes identified as candidate genes involved in neuroblastoma progression were: *MYCN* (neuroblastoma derived), *NPW* (neuropeptide W), *SLC30A3* (solute carrier family 30, member 3), *MYCNOS* (neuroblastome derived opposite strand), and *MYCN** (v-myc myelocytomatosis viral related oncogene). *MYCN* and *MYCN** are the same genes detected by different probes. *MYCNOS* and *SLC30A3* were confirmed to be correlated with the status of expression of *MYCN* in neuroblastoma [24,25]. For each of the selected genes, the probability of being 'On' in the favorable group, that in the unfavorable group, and the difference in probability of being 'On' between the unfavorable and favorable groups and its ranking, are listed in Table 3.

To assess the advantage of the 'On/Off' method, we calculated 'relative difference' statistics for the difference of average of gene expression intensities between favorable and unfavorable groups (the 'relative difference' statistic was proposed by Tusher et al [26] to stabilize the t-value). For each of the selected genes, average expression intensities obtained by the RMA methods and their standard errors in the favorable and unfavorable groups, as well as the ranking of 'relative difference' statistics in descending order, are listed in Table 4. Accordingly, the OR_On type genes are difficult to select based on the ranking of gene expression intensities. The method based on the 'On/Off' state of a gene performs better than the method based on gene expression intensity.

Real-time RT-PCR was employed to verify whether the four distinct selected genes were OR_On type genes. Gene-expression features obtained by microarray data analysis and Ct values from real-time RT-PCR for each of these genes are shown in Figures 7A-D. The displayed features are scatter diagrams of $X$ and expression intensity (RMA summarized value), parallel box plots of Ct




| | unfavorable | favorable |
|---|---|---|
| MYCN: On | 9 | 0 |
| Off | 12 | 40 |

**Figure 5 Scatter diagram of *X* and expression intensities of *MYCN* for each neuroblastoma case.** The pink points represent cases with unfavorable outcome and the blue points those with favorable outcome. The relationship between *MYCN* state ('On/Off') and outcome (favorable/unfavorable) is listed, where the gene state 'On' is defined as $X \geq 7$ and 'Off' is defined as $X \leq 6$.

values in real-time RT-PCR for three groups -- favorable, unfavorable 'On', and unfavorable 'Off' (the latter two groups abbreviated as 'unfavorable_On' and 'unfavorable_Off ') -- and parallel box plots of Ct values from real-time RT-PCR of *GAPDH* according to the same three groups. As mentioned above, the state of a gene was defined as 'On' if $X \geq 7$ and 'Off' if $X \leq 6$. Average Ct value was significantly lower in the unfavorable_On group compared with that in the favorable and unfavorable_Off groups. This confirms that the selected genes were not expressed in the favorable and unfavorable_Off groups but were in the unfavorable_On group.

## Discussion

MAS5 P/M/A calls are based on a nonparametric statistical test, in which the default state of a gene is 'absent'. Therefore, it inevitably yields many false negatives which, we think, is its main disadvantage. For example, *BIRC5* (baculoviral IAP repeat-containing 5), also called *survivin*, which is a human gene that is a member of the inhibitor of apoptosis (IAP) family, is expressed at high levels in most human tumors but is completely absent in terminally differentiated cells [27]. Figure 8 shows a diagram of *X* and expression RMA intensities of *BIRC5* for each neuroblastoma case. *BIRK5* was judged as 'On' in all

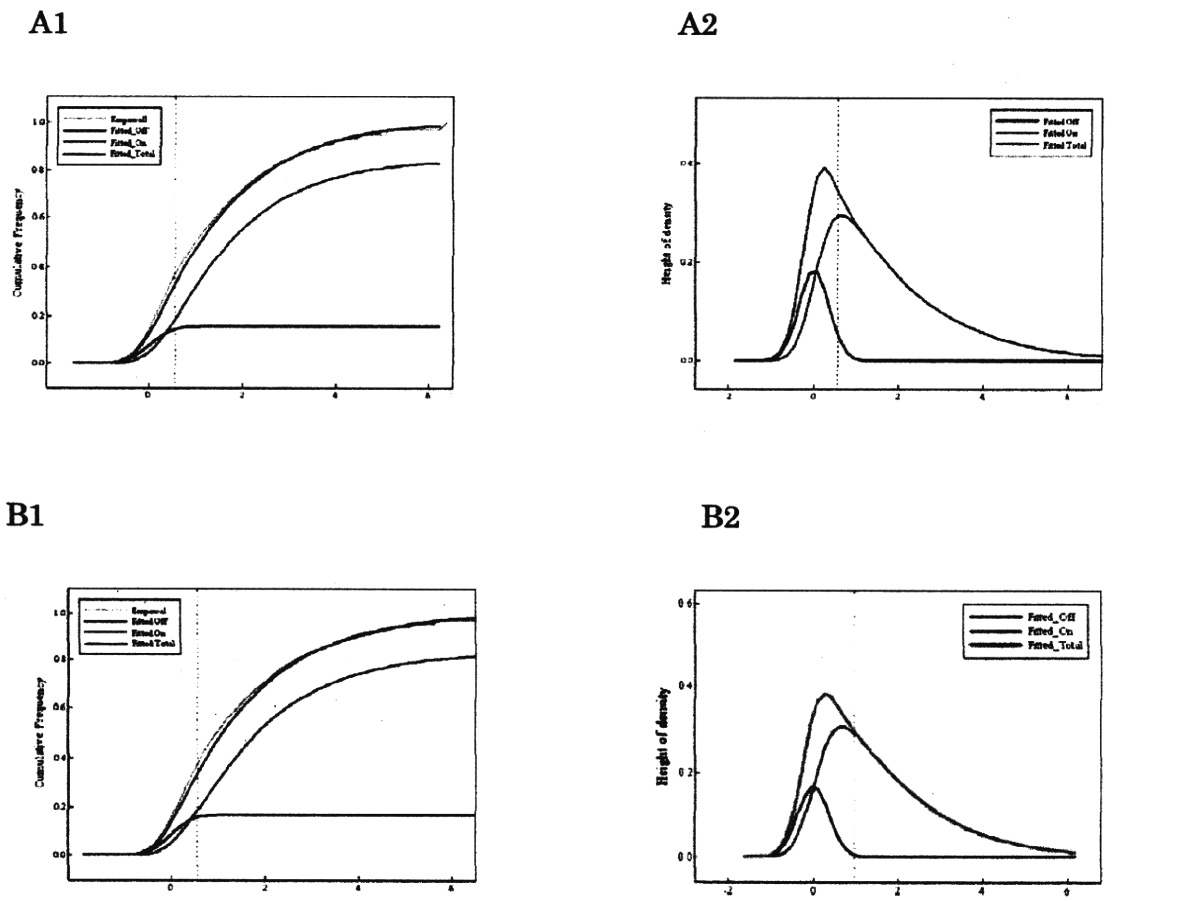



**Figure 6 Results of applying the Weibull-normal mixture model to the neuroblastoma data sets - A and B show the favorable and unfavorable group, respectively. (A1, B1)**. Comparison between the fitted Weibull-normal distribution (green line) with two components ('On' and 'Off') and the empirical distribution (brown line). The pink line shows the component of expression intensities of 'On' genes, the blue line that of 'Off' genes. **(A2, B2)** Density function corresponding to the fitted Weibull-normal distribution function (green line) and its components (pink and blue lines).

of the 61 cases by our method but 12 cases, which are circled in Figure 8, were classified as 'absent' by MAS5. As is shown in Table 2 or Figure 8, 'On/Off' calls generated few false negatives compared to *MAS5* calls. Although the poor separation of 'On' and 'Off' components of the Weibull-Normal mixture would result in false positives or negatives, we think that it is due to a limitation of microarray performance.

In this study, we only identified genes that switched 'On' and 'Off' between phenotypes using an indicator based on the probability of being 'On'. Of course, some genes exhibit differences between phenotypes in terms of their quantitative expression intensities, going from normal to abnormally increased or decreased. For example, *BIRC5* was judged as 'On' in all cases and, in addition, 'On' with 'abnormally increased intensity' predicted poor diagnosis (unfavorable) except for one case that had been detected by mass-screening at stage I followed by radical resection. We further explored the genes whose intensity levels changed quantitatively from normal to abnormal between two different phenotypes, such as *BIRC5*, by calculating the likelihoods under the null hypothesis that gene expression intensities obey a normal distribution. As a result, more than a thousand genes were selected as candidates.

A method to create a gene expression barcode - genes expected to be expressed are coded with ones and those expected to be unexpressed are coded with zeros - was developed by Zilliox and Irizarry [28]. Furthermore, an algorithm for estimating expression states was described

**Table 3: Top five genes selected using the probability of a gene being 'On'.**

| | Based on the probability being 'On' | | | |
|---|---|---|---|---|
| Gene | FavorablePr_On | UnfavorablePr_On | Pr_On (Unfav.) - Pr_On (Fav.) | Ranking |
| *SIC30A3* | 0.26 | 0.89 | 0.63 | 1 |
| *MYCN*[a] | 0.38 | 0.94 | 0.56 | 2 |
| *MYCNOS* | 0.38 | 0.88 | 0.50 | 3 |
| *NPW* | 0.16 | 0.58 | 0.42 | 4 |
| *MYCN* | 0.19 | 0.58 | 0.39 | 5 |

by McCall et al. [29]. They state that the magnitude of the unexpressed observed intensities differs among genes; accordingly the distribution of observed intensities must be estimated for each gene. Their method therefore requires a large database of observed intensities across many different tissues. We consider it natural that the magnitude of the unexpressed observed intensities differs by gene because unexpressed observed intensities include gene dependent cross-hybridization. On the other hand, our method can correct gene dependent cross-hybridizations by using MM probes. Therefore, it is unnecessary to be concerned with differences in distribution of unexpressed observed intensities among genes. Several tens of samples are enough to estimate it. Another characteristic of our method is robustness, because it is based on the order relationship between PM and MM values. Our assumptions are just (1) the expected PM value is larger than that of the MM value when a gene is expressed, and (2) the expected PM value equals that of the MM value when a gene is unexpressed. Although availability of MM probes may be in doubt [10], our approach provides good justification for the use of MMs.

## Conclusion

The qualitative evaluation 'probability of a gene being expressed' provides a useful indicator for improving the performance of microarray data analysis. When expression intensity of a gene is not high, it is difficult to determine its real intensity after removing non-specific binding. Especially in this case, 'probability of a gene being expressed' gives useful qualitative information complementing its true intensity. In regards to a practical problem in expression array analysis, genes that switch between 'On' and 'Off' with different phenotypes can be found with greater confidence. Our proposed method of estimating 'probability of a gene being expressed' is robust because it is not based on expression intensities but rather is based on the order relationship between PM and MM values.

## Methods

### Human Genome U95 data sets

The human genome U95 data sets consist of a series of 14 genes spiked-in at known concentrations (0, 0.25, 0.5, 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, and 1024 pM) and arranged in a Latin square format. Each subsequent experiment rotates the spike-in concentrations by one experimental group. The data consist of 14 spiked-in genes in 14 experimental groups. Replicates within each experimental group result in a total 59 CEL files.

### Neuroblastoma samples

Total RNA was extracted from 61 neuroblastoma samples. Ages at diagnosis and stages at surgery according to the INSS (International Neuroblastoma Staging System) are shown in Additional file 1. All patients were diag-

**Table 4: Ranking using 'relative difference' statistics.**

| | RMA method | | |
|---|---|---|---|
| Gene | Favorable average (s.d.) | Unfavorable average (s.d.) | Ranking |
| *SIC30A3* | 4.93 (0.190) | 5.40 (0.521) | 100 |
| *MYCN*[aa] | 4.59 (0.186) | 5.21 (0.617) | 17 |
| *MYCNOS* | 4.59 (0.152) | 5.12 (0.523) | 24 |
| *NPW* | 4.17 (0.106) | 4.63 (0.670) | 1016 |
| *MYCN* | 4.57 (0.161) | 5.02 (0.508) | 151 |

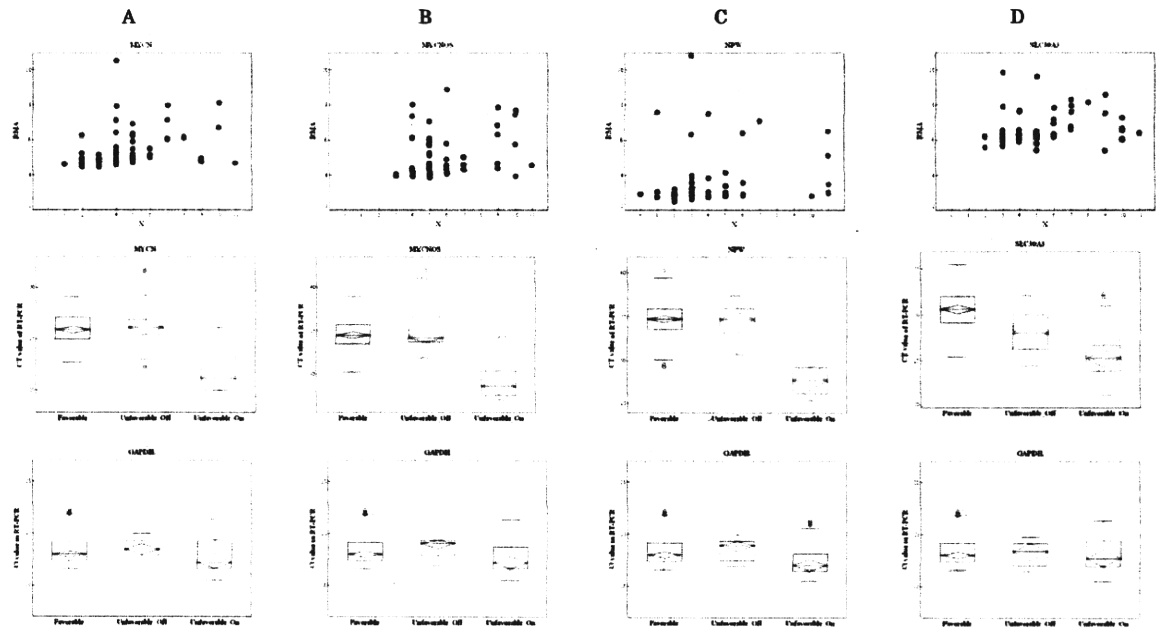



**Figure 7 Gene-expression features assessed by microarray and Ct values in real-time RT-PCR for four unique selected genes.** The genes selected were: A: *MYCN*, B: *MYCNOS*, C: *NPW*, and D: *SLC30A3*. The first row shows scatter diagrams of *X* versus expression intensities for each case. Pink points represent cases with unfavorable outcome and blue points cases with favorable outcome. The second row shows parallel boxplots of Ct values in real-time RT-PCR by three groups: favorable, unfavorable_On, and unfavorable_Off. The third row shows parallel boxplots of Ct values in real-time RT-PCR of *GAPDH* for the same three groups.

nosed as having neuroblastoma between 1991 and 2005 at Hiroshima University Hospital or affiliated hospitals. Most of the patients were treated according to the Japanese neuroblastoma protocols for infants or advanced stage NB (A1, new A1, or A3) [30]. The follow-up period was more than 5 years for all patients. This research was approved by the Ethics Committee of Hiroshima University (Hiro-Rin-20). Written informed consent was obtained from parents of all patients. None of the patients had therapy prior to surgery or biopsy.

### Affymetrix microarray analysis

Microarray experiments were conducted according to standard protocols for Affymetrix Genome U133 Plus 2.0 arrays (Affymetrix, Inc., Santa Clara, CA) [31]. Briefly, using 1 μg of total RNA, cDNA and biotinated cRNA synthesis was performed using the GeneChip expression 3' amplification reagents (one-cycle cDNA synthesis, and IVT labeling) kits of Affymetrix following the manufacturer's protocols. Fragmented cRNA was applied to the hybridization and scanning of the array was performed following the manufacturer's protocols. Experimental details and all results are available at the Gene Expression Omnibus, http://www.ncbi.nlm.nih.gov/geo/ (GEO accession number GSE16237).

### Quantitative RT-PCR

In each tumor sample, cDNA was synthesized from 5 μg total RNA using a High Capacity cDNA Archive™ kit (Applied Biosystems), and then a five-hundredth aliquot of the cDNA (equivalent to 10 ng total RNA) was subjected to real-time RT-PCR using Universal Probe Library (UPL, Roche Diagnostics, Tokyo, Japan) for each target gene, or an internal control *GAPDH (glyceraldehyde-3-phosphatedehydrogenase)* TaqMan™ probe (Applied Biosystems) on an ABI PRISM™ 7900HT sequence detection system (Applied Biosystems) with 384-well plates. The relative gene expression levels were calculated as a ratio relative to *GAPDH expression level.*

### Quantification of the likelihood that a gene is 'On'

Define a random variable $U_j^{(g)}$, the specific value of which indicates the order relationship between PM and MM: i.e.,

$$U_j^{(g)} = \begin{cases} 0\ (PM_j^{(g)} \le MM_j^{(g)}) \\ 1\ (PM_j^{(g)} > MM_j^{(g)}) \end{cases}, \ j = 1, \cdots, J.$$

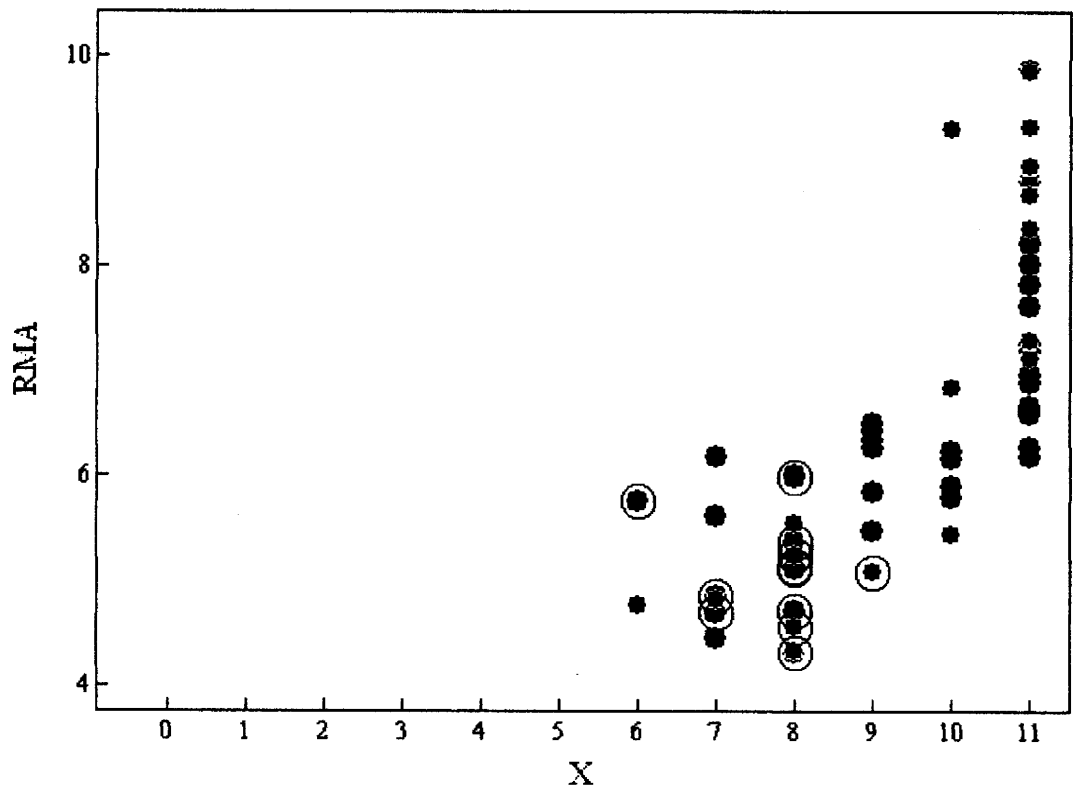



**Figure 8 Scatter diagram of *X* versus expression intensities of *BIRC5* for each neuroblastoma case.** The pink points show cases with unfavorable outcome and the blue points show cases with favorable outcome. The cases called 'absent' by MAS5 are circled.

A set of probe level data for genes in an array $i$ may be described as

$$\begin{pmatrix} U_{1i}^{(1)} & \cdots & U_{1i}^{(g)} & \cdots & U_{1i}^{(G)} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ U_{ji}^{(1)} & \cdots & U_{ji}^{(g)} & \cdots & U_{ji}^{(G)} \\ \vdots & \cdots & \vdots & \ddots & \vdots \\ U_{Ji}^{(1)} & \cdots & U_{Ji}^{(g)} & \cdots & U_{Ji}^{(G)} \end{pmatrix}, \quad i = 1, \cdots, N.$$

We introduce a random variable $Z^{(g)}$ defined as

$$Z^{(g)} = \frac{1}{J}\sum_{j=1}^{J} \text{logit}\{\frac{1}{N+1}(\sum_{i=1}^{N} U_{ji}^{(g)} + \frac{1}{2})\}$$

to quantify the likelihood of a gene being expressed, where $\frac{1}{2}$ is added to avoid discontinuity at $\sum_{i=1}^{N} U_{ji}^{(g)} = N$ or 0 and $N + 1$ enhances the model fit. When a gene $g$ is sampled randomly, the random variable $Z = Z^{(g)}$ is assumed to follow a mixture distribution with two components corresponding to the separate states 'On' and 'Off'. We assume that $Z$ may be expressed as the sum of random variables $T$ and $X$, where $T$ expresses the likelihood of a gene being expressed when in the 'On' state and X expresses a random error having a normal density function $\varphi$ with mean zero and variance $\sigma^2$. We further assume that the density function of $T$ is given by $f(t|\mu, \alpha, \xi) = \xi\delta(t) + (1-\xi) f_w(t|\mu, \alpha)$ where $f_w$ denotes the Weibull density function with location parameter $\mu$ and power parameter $\alpha$, $\delta$ denotes the Dirac function, and $\xi$ denotes the mixture rate of 'Off' genes. Then the density function of $Z$ can be expressed as

**Table 5: Association between gene status and response with 'OR' and 'AND' type genes.**

| | OR_On | | OR_Off | | AND_On | | AND_Off | |
|---|---|---|---|---|---|---|---|---|
| | $R^+$ | $R^-$ | $R^+$ | $R^-$ | $R^+$ | $R^-$ | $R^+$ | $R^-$ |
| $g^+$ | a | 0 | $N_1$- b | $N_2$ | $N_1$ | c | 0 | d |
| $g^-$ | $N_1$- a | $N_2$ | b | 0 | 0 | $N_2$- c | $N_1$ | $N_2$- d |
| Total | $N_1$ | $N_2$ | $N_1$ | $N_2$ | $N_1$ | $N_2$ | $N_1$ | $N_2$ |

$$h(z \mid \theta) = \int_0^{+\infty} f(t \mid \mu, \alpha, \xi)\phi(z - t \mid \sigma^2)dt$$
$$= \xi\phi(z \mid \sigma^2) + (1 - \xi)\int_0^{+\infty} f_W(t \mid \mu, \alpha)\phi(z - t \mid \sigma^2)dt.$$

Given a set of samples $\{z^{(g)} \mid g = 1, \cup, G\}$, the maximum likelihood estimate $\hat{\theta} = (\hat{\mu}, \hat{\alpha}, \hat{\xi}, \widehat{\sigma^2})$ of $\theta$ is obtained by maximizing $l(\theta) = \sum_{g=1}^{G} \log\{h(z^{(g)} \mid \theta)\}$. The posterior probabilities with respect to the status of gene expression can be written as

$$\Pr(\tau_g = 0 \mid z, \theta) = \frac{\xi\phi(z|\sigma^2)}{h(z|\theta)},$$

$$\Pr(\tau_g = 1 \mid z, \hat{\theta}) = 1 - \frac{\xi\phi(z|\sigma^2)}{h(z|\theta)}. \quad (1)$$

A cutoff value $z_c$ dividing gene states 'On' and 'Off' is determined as the lowest value of $z$ satisfying $\Pr(y \geq z \mid y \sim \varphi(0, \sigma^2))$ smaller than $\alpha$ $(0 \leq \alpha \leq 1)$.

### 'AND' and 'OR' type genes

We presume the existence of 'AND' and 'OR' type genes that show a logical relationship between gene status 'On/Off' and binary phenotype. They are defined as follows:

*Definition 1*: A gene $g$ is defined as 'OR_On' in the case that $g^+$ leads to $R^+$ (the gene being *on* is sufficient for unfavorable outcome), 'OR_Off' in the case that $g^-$ leads to $R^+$ (the gene being *off* is sufficient for unfavorable outcome), 'AND_On' in the case that $R^+$ implies $g^+$ (the gene being *on* is necessary for unfavorable outcome), and 'AND_Off' in the case that $R^+$ implies $g^-$ (the gene being *off* is necessary for unfavorable outcome), where the symbols $R^+$ and $R^-$ indicate the two outcome phenotypes and $g^+$ and $g^-$ indicate a gene being 'On' and 'Off', respectively.

The frequency distribution of cases according to these four types of genotype/phenotype relationship can be expressed by a set of two-by-two tables (Table 5).

### Identification of 'OR_On type genes involved in neuroblastoma

We set out to identify OR_On type genes involved in neuroblastoma prognosis using the estimate of the probability of a gene being 'On' or 'Off'. OR_On type genes were identified by the following procedure.

***Step 1***

Calculate the probability of a gene being 'On' in the favorable and unfavorable groups using formula (1).

***Step 2***

A gene set $G_1$ satisfying two conditions--(1) uniformly 'Off' in the favorable group and (2) varying 'On' and 'Off' in the unfavorable group--is defined as

$$G_1 = \{g \mid \Pr(\text{g is 'Off'} \mid \text{favorable group}) > 0.5\}$$
$$\cap \{g \mid \Pr(\text{g is 'On'} \mid \text{unfavorable group})$$
$$- \Pr(\text{g is 'On'} \mid \text{favorable group}) > 0.2\}.$$

***Step 3***

Arrange the values of Pr($g$ is 'Off' in the favorable group | $g$ $G_1$) in descending order and select the top 100 genes. Then rearrange according to

$$\Pr(g \text{ is 'On' in the unfavorable group})$$
$$- \Pr(g \text{ is 'On' in the favorable group})$$

in descending order.

### Software

MAS5 and RMA expression indices were calculated using the package *affy* [32,33] provided by BioConductor [34]. Fortran was used to perform all of the analyses.

## Additional material

**Additional file 1 Neuroblastoma cases.** The file contains the table including ages at diagnosis and stages at surgery according to the INSS (International Neuroblastoma Staging System) of 61 neuroblastoma cases.

**Abbreviations**

MAS5: Affymetrix Microarray Suite version 5; MM: mismatch probe; PM: perfect match probe; RMA: robust multi-array analysis.

**Authors' contributions**

MO participated in the design of the study, statistical analysis, and drafting of the manuscript. KO participated in the statistical analysis and drafting of the manuscript. KH and EH were involved in conducting the microarray experiment and assisted with manuscript preparation. NK participated in conducting the microarray experiment. KS was involved in microarray analysis. All authors edited and approved the final version of the manuscript.

**Acknowledgements**

The present study was supported in part by a grant from the Ministry of Education, Culture, Sports, Science and Technology (No. 18300095) and in part by a grant from the Ministry of Health, Labor and Welfare.

**Author Details**

[1]Department of Environmetrics and Biometrics, Research Institute for Radiation Biology and Medicine, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima, 734-8551, Japan, [2]Department of Translational Cancer Research, Research Institute for Radiation Biology and Medicine, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima, 734-8551, Japan and [3]Natural Science Center for Basic Research and Development, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima, 734-8551, Japan

**Received: 17 February 2009 Accepted: 12 April 2010**
**Published: 12 April 2010**

**References**

1. MAQC Consortium: **The Microarray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements.** *Nature Biotechnology* 2006, **24**:1151-1161.
2. Affymetrix: **Affymetrix Microarray Suite User Guide, Version 4 ed.** Affymetrix Santa Clara, CA; 1999.
3. Li C, Wong WH: **Model-based analysis of oligonucleotide arrays: Expression index computation and outlier detection.** *Proc Natl Acad Sci USA* 2001, **98(1)**:31-36.
4. Affymetrix: **Statistical Algorithms Description Document.** [http://media.affymetrix.com/support/technical/whitepapers/sadd_whitepaper.pdf].
5. Affymetrix: **Guide to probe logarithmic intensity error (PLIER) estimation.** [http://media.affymetrix.com/support/technical/technotes/plier_technote.pdf].
6. Li C, Wong WH: **Model-based analysis of oligonucleotide arrays: model valiation, design issues and standard error application.** *Genome Biol* **2**:1-11.
7. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP: **Summaries of Affymetrix GeneChip probe level data.** *Nucleic Acids Res* 2003, **31(4)**:e15.
8. Wu Z, Irizarry R, Gentleman R, Martinez-Murrillo F, Sprencer F: **A model-based background adjustment for oligonucleitide expression arrays.** *Journal of the American Statistical Association* 2004, **99**:909-917.
9. Wu C, Carta R, Zang L: **Sequence dependence of cross-hybridization on short oligo microarrays.** *Nucleic Acids Res* 2005, **33(9)**:e84.
10. Wu BZ, Irizarry RA: **A statistical frameworks for the analysis of microarray probe-level data.** *The Annals of Applied Statisstics* 2007, **1(2)**:333-357.
11. Millenaar FF, Okyere J, May ST, Zanten MV, Voesenek LACJ, Peeters AJM: **How to decide? Different methods of calculating gene expression from short oligonucleotide array data will give different results.** *BMC Bioinformatics* 2006, **7**:137.
12. Calza C, Raffelsberger W, Polner A, Sahel J, Leveillard T, Pawitan Y: **Filtering genes to improve sensitivity in oligonucleotide microarray data analysis.** *Nucleic Acids Res* 2007, **35(16)**:e102.
13. Su AI, Cooke MP, Ching K, Hakak Y, Walker JR, Wilshire T, Orth AP, Vega RG, Sapinoso LM, *et al.*: **Large-scale analysis of the human and mouse transcriptions.** *PNAS* 2002, **99**:4465-4470.
14. Jongeneel CV, Iseli C, Stevevson BJ, Riggies GJ, Lal A, Mackey A, Harris RA, O'Hare MJ, Neville AM, *et al.*: **Comprehensive sampling of gene expression in human cell lines with massively parallel signature sequencing.** *PNAS* 2003, **100**:4702-4705.
15. McClintik JN, Edenberg HJ: **Effects of filtering by Present call on analysis of microarray experiments.** *BMC Bioinformatics* 2006, **7(49)**:.
16. Ohtaki M, Otani K, Satoh K, Kawamura T, Hiyama K, Nishiyama M: **Model-based analysis of microarray data: expression of differentially expressed genes between two cell types based on a two-dimensional mixed normal model.** *Jpn J of Biometrics* 2005, **2(1)**:31-48.
17. Otani K, Hiyama K, Satoh K, Shimamoto T, Mohamad D, Andoh M, Tonda T, Kohda M, Ohazaki Y, Nishiyama M, Hiyama E, Ohtaki M: **A Mathematical Model for Affymetrix GeneChip Probe Level Data.** *JP Journal of Biostatistics* 2007, **1(3)**:283-306.
18. Hiyama K, Otani K, Ohtaki M, Satoh K, Kumazaki T, Takahashi T, Mitsui Y, Okazaki Y, Hayashizaki Y, Omatsu H, Noguchi T, Tanimoto K, Nishiyama M: **Differentially expressed genes throughout the cellular immortalization prcesses are quite different between normal human fibroblasts and endothelial cells.** *Int J of Oncl* 2005, **27**:87-95.
19. Komatsu M, Hiyama K, Tanimoto K, Yunokawa M, Otani K, Ohtaki M, Hiyama E, Kigawa J, Ohwada M, Suzuki M, Nagai N, Kudo Y, Nishiyama M: **Prediction of individual response to platinum/paclitaxel combination using novel marker genes in ovarian cancers.** *Mol Cancer Ther* 2006, **5**:767-775.
20. Fumoto S, Shimokuni T, Tanimoto K, Hiyama K, Otani K, Ohtaki M, Hihara J, Yoshida K, Hiyama E, Nogutchi T, Nishiyama M: **Selection of a novel drug-response predictor in esophageal cancer: A novel screening method using microarray and identification of IFITM1 as a potent marker gene of CDDP response.** *International Journal of oncology* 2007, **32**:423-423.
21. Hiyama E, Iehara T, Sugimoto T, Fukazawa M, Hayashi Y, Sakaki F, Sugiyama M, Kondo S, Yoneda A, Yamaoka H, Tajiri T, Akazawa K, Ohtaki M: **Effectiveness of screening for neuroblastoma at 6 months of age: a retrospective population-based cohort study.** *Lancet* 2008, **371**:1173-80.
22. Segar RC, Brodeur GM, Sather H, Dalton A, Siegel SE, Wong KY, Hammond D: **Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastoma.** *New Engl J Med* 1985, **313**:1111-1116.
23. Tang XX, Zhao H, Kung B, Kim DY, Hicks SL, Dohn SL, Cheng NK, Seeger RC, Evans AE, Ikegaki N: **The *MYCN* Enigma: Significance of *MYCN* Expression in Neuloblastoma.** *Cancer Res* 2006, **66**:2826-2833.
24. Alaminos M, Mora J, Cheung NKV, Smith A, Qin J, Chen L, Gerald WL: **Genome-wide Analysis of Gene Expression Associated with MYCN in Human Neuloblastoma.** *Cancer Research* 2003, **63**:4538-4546.
25. Lastowska M, Viprey V, Santibanez-Koerf M, Wapper I, Peters H, Cullinane C, Roberts P, Hall AG, Tweddle DA, Pearson ADJ, Lewis I, Burchill SA, Jackson MS: **Identification of candidate genes involved in neuroblastoma progression by combining genomic and expression microarray with survival data.** *Oncogene* 2007:1-13.
26. Tusher VG, Tibshirani R, Chu G: **Significance analysis of microarrays applied to the ionizing radiation response.** *PNAS* 2001, **98**:5116-5121.
27. Sah NK, Khan Z, Khan GJ, Bisen PS: **Structural, functional and therapeutic biology of survivin.** *Cancer Lett* 2006, **244(2)**:164-171.
28. Zillox MJ, Irizarry RA: **A gene expression bar code for microarray data.** *Nat Methods* 2007, **4(11)**:911-913.
29. McCall MN, Zilliox MJ, Irizarry RA: **Gene Expression Barcodes Based Data from 8,277 Microarrays.** Johns Hopkins University, Dept of Biostatistics Working paper; 2009. Paper 200
30. Kaneko M, Tsuchida Y, Uchino J, Takeda T, Iwafuchi M, Ohnuma M, Mugishima H, Yokoyama J, Nishihara H, Nakada K, Sasaki S, Sawada T, Kawa K, Nagahawa N, Suita S, Sawaguchi S: **Treatment results of advanced neuroblastoma with the first Japanese study group protocol. Study Group of Japan for Treatment of Advanced Neuroblastoma.** *J Pediatr Hematol Oncol* 1999, **21**:190-197.
31. Martinez T, Pascual A: **Gene expression profile in beta-amyloid-treated SH-SY5Y neuroblastoma cells.** *Brain Res Bull* 2007, **72**:225-231.
32. Irizarry RA, Hobbs B, Colin F, Beazer-Barcley YD, Antonellis KJ, Scherf U, Speed TP: **Exploration, Normarization, and Summaries of High Density Oligonucleotide Array Probe Level Data.** *Biostatistics* 2003, **4**:249-264.
33. Hubbell E, Liu WM, Mei R: **Robust estimators for expression analysis.** *Bioinformatics* 2002, **18**:1585-1592.
34. **Bioconductor** [http://www.bioconductor.org]

**doi: 10.1186/1471-2105-11-183**
**Cite this article as:** Ohtaki *et al.*, A robust method for estimating gene expression states using Affymetrix microarray probe level data *BMC Bioinformatics* 2010, **11**:183

PEDIATRICS INTERNATIONAL Official Journal of the Japan Pediatric Society



# High-dose lorazepam for convulsive status epilepticus in an infant with holoprosencephaly

Kaoru Okazaki,[1] Masatoshi Kondo,[1] Masaya Kubota,[2] Ryota Kakinuma,[1] Ai Hoshino,[3] Hirokazu Kimura[4] and Susumu Itoh[5]

[1]*Department of Neonatology, Tokyo Metropolitan Hachioji Children's Hospital,* [2]*Department of Neurology, National Center for Child Health and Development,* [3]*Department of Pediatrics, Tokyo Metropolitan Neurological Hospital,* [4]*Infectious Diseases Surveillance Center, National Institute of Infectious Diseases, Tokyo, and* [5]*Department of Pediatrics, Kagawa University, Kagawa, Japan*

**Key words** lorazepam, refractory seizure, semilobar holoprosencephaly.

Holoprosencephaly (HPE) is characterized by severe brain malformation due to early arrest in brain cleavage and rotation, and the prevalence is only about 1 in 10 000.[1] HPE has various central nervous system (CNS) manifestations, including epilepsy and mental retardation.[2] In particular, approximately half of the children with the semilobar type of HPE have at least one seizure, and about half of these have difficult-to-control seizures.[3] In Western countries, intravenous lorazepam is frequently used as a first-line anti-epileptic drug for the control of status epilepticus due to congenital CNS abnormalities, however this drug is not available in Japan.[4,5] Lorazepam has a longer duration of action than other benzodiazepines such as diazepam and has few side-effects (i.e. respiratory depression).

We encountered a case of refractory status seizures due to semilobar HPE in a 6-month-old Japanese girl. The patient was first treated with conventional anticonvulsants, including phenobarbital, diazepam, clobazam, and valproic acid. However, we failed to control the severe daily seizures with these drugs. The patient also had life-threatening complications, including status epilepticus, hypoxia, apnea and tachycardia. Treatment with high-dose oral lorazepam drastically improved these complications. Here, we report on the patient's clinical course and discuss the beneficial effects of high-dose oral lorazepam when treating intractable seizures.

## Case report

The patient was born at 36 weeks and 2 days of gestation by natural vaginal delivery. A prenatal cranial ultrasound study revealed ventriculomegaly. She was small for dates (birthweight 2068 g). The parents were non-consanguineous and had one healthy girl. At birth, features of incomplete midline facial development such as hypotelorism and absence of a nasal septum were found. At 16 h after birth, she had generalized tonic–clonic seizures with some mild abnormal vital signs, such as low-grade fever, slight hypoxia, and tachycardia. On electroencephalography (EEG), frequent convulsive waves were observed. A brain computed tomography (CT) scan suggested semilobar HPE (Fig. 1). In addition, chromosome analysis showed 45, XX, der (15; 18) (q10; q10), and add (19)(p13.3). These present as deletions of 18p and dup(19)(q13.3). HPE is the most severe congenital malformation with deletions of 18p (18p-syndrome).[1] The dup(19)(q13.3) is associated with some congenital anomalies, including cleft lips, cardiac defects, and cerebral atrophy.[6] Thus, we diagnosed the condition as semilobar HPE accompanied by convulsive status epilepticus. First, phenobarbital (loading dose, 20 mg/kg) was given as a suppository. After the loading dose, the patient received maintenance doses (8 mg/kg/24 h in two divided doses). However, even on day 4 after initiation of treatment, the seizures could not be controlled completely by phenobarbital alone. Another anticonvulsant, midazolam (0.25 mg/kg/h), was

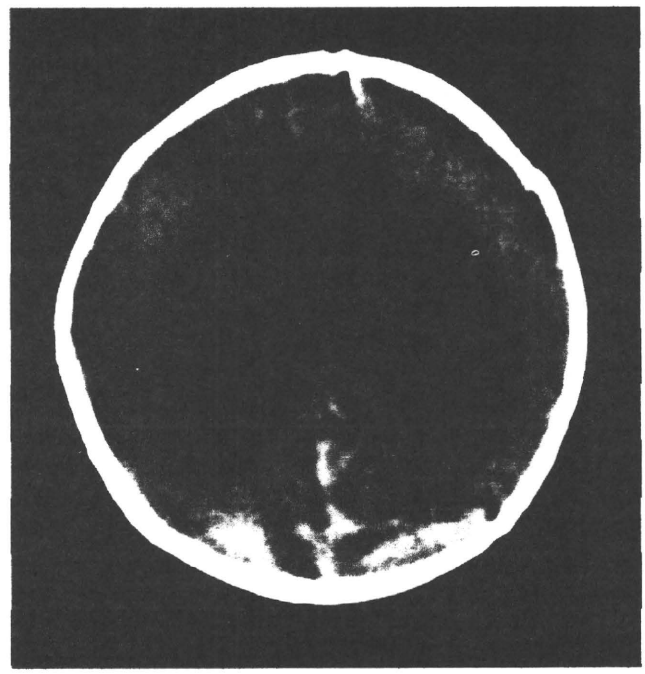


**Fig. 1** Computed tomography (CT) study of the brain. An axial CT scan obtained 8 h after birth shows separation of the hemispheres posteriorly but not anteriorly. Anterior horns of the lateral ventricles are absent. There is also an incomplete separation of the basal ganglia.

Correspondence: Kaoru Okazaki, MD, Tokyo Metropolitan Hachioji Children's Hospital, 4-33-13 Daimachi, Hachioji, Tokyo 193-0931, Japan. Email: okazaki@chp.hachioji.tokyo.jp

Received 1 July 2008; revised 1 August 2009; accepted 20 August 2009.
doi: 10.1111/j.1442-200X.2010.03077.x

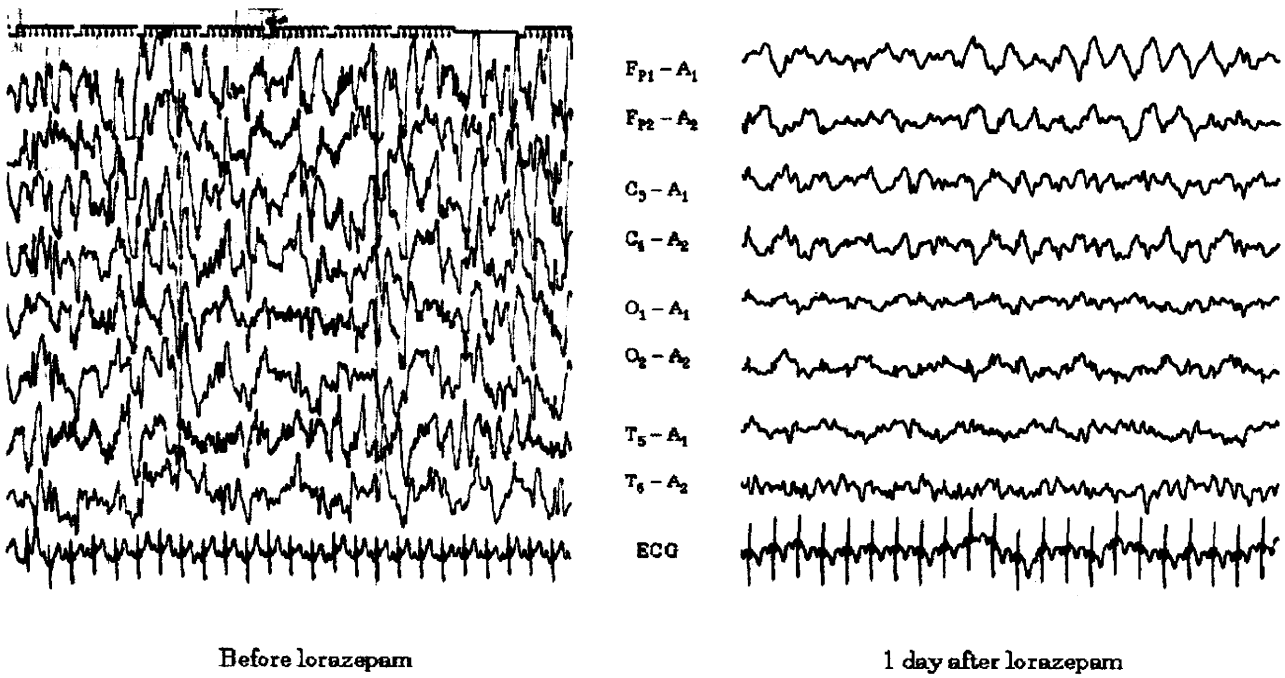



**Fig. 2** Electroencephalograms (ECG) before (left panel) and after (right panel) lorazepam administration.

given intravenously. This treatment was effective and the seizures transiently disappeared. Oral administration of clobazam, the same benzodiazepine as midazolam, was started, and the dose of midazolam was tapered slowly over a period of 1 month. After stopping treatment with midazolam, seizures (such as myoclonic seizures) sometimes appeared, but they could be controlled by many anticonvulsants, including phenobarbital, clobazam, diazepam, and valproic acid.

However, at 6 months of age, the severe generalized tonic–clonic seizures recurred with various severe symptoms such as tachycardia (200 b.p.m.), high fever, rolling of the eyes, and hypoxic attacks. All anticonvulsants were maintained at therapeutic serum levels, including phenobarbital (serum concentration 53.4 μg/mL) and clobazam (serum concentration 50 ng/mL). Treatment with midazolam was also resumed at a maximum dose of 0.3 mg/kg/h but was not effective. The patient was then treated with intravenous thiopental sodium (maximum dose: 2.0 mg/kg/h). However, she continued to have repeated, severe seizures and was in a life-threatening condition. Accordingly, as an ultimate treatment, we started high-dose lorazepam therapy (2 mg/kg/24 h) by oral administration, after approval by the local ethics committee of the Tokyo Metropolitan Hachioji Children's Hospital and with written informed consent from the patient's parents. However, lorazepam is available only as a tablet form for anxiety disorders in Japan. Thus, we made a finely milled powder from the tablets, mixed this with distilled water, and administered it through a gastric tube. Thirty minutes after lorazepam administration, seizure activity immediately ceased. The heart rate decreased to 120 b.p.m. and the fever reduced. On EEG, diffuse bilateral spike-wave complexes (hypsarrhythmia-like) were observed during status seizures before lorazepam administration. After lorazepam administration, the spike-wave complexes completely disappeared (Fig. 2). No status seizure was seen, but tonic seizures occurred a few times a day. Thus, the dose of lorazepam was gradually increased to a maximum dose of 5 mg/kg/24 h at 7 months of age. Finally, we continued lorazepam for 2 months for the treatment of myoclonus.

The patient is now 2 years old. Her seizures are controlled by phenobarbital (8 mg/kg/24 h) and clobazam (0.6 mg/kg/24 h). On an as-needed basis, lorazepam (0.5 mg/kg/dose) is used about 3 times per month only for refractory myoclonus. No clinically significant side-effects of lorazepam, such as excessive sedation or respiratory depression, have been observed. Thus, high-dose oral lorazepam was far more effective and safe for controlling seizures than both intravenous midazolam and thiopental.

In Japan, lorazepam is not officially approved for treatment of seizures in infants, children and adults. To avoid adverse events, we measured serum concentrations of lorazepam by using high-performance liquid chromatography (SCG, Saitama, Japan). We checked lorazepam serum concentrations at 0, 30, 60, 120 and 180 min after administration of each dose (0.2 mg/kg) (Fig. 3). The peak value was at 30 min. The elimination rate constant and distribution volume were 0.18/h and 0.67 L/body, respectively.

## Discussion

HPE cases generally show various CNS signs. About 50% of HPE cases with convulsions manifest with refractory seizures resulting in multiple drug resistance to anticonvulsants.[3] We showed that a case of refractory seizures due to semilobar HPE was controlled with lorazepam, while various other anticonvul-

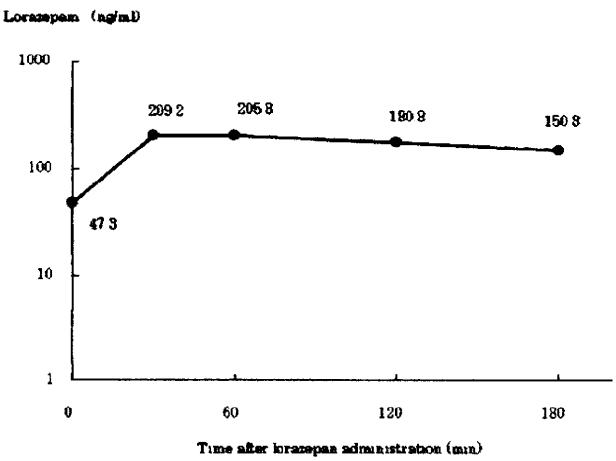



**Fig. 3** Change in serum concentration after administration of lorazepam. Lorazepam was not administrated for 24 h before measuring concentrations.

sant drugs were not effective. Lorazepam is a benzodiazepine and used as an anti-anxiety drug in general. This medication is also frequently used in Europe and the USA to control status epilepticus due to various CNS diseases, including congenital CNS abnormalities.[4,7] Although lorazepam is not generally used to control status epilepticus caused by CNS diseases in children and adults in Japan,[5] this drug may be a first-line anti-epileptic drug when other conventional anticonvulsants are not effective for refractory seizures in semilobar HPE.

We first used phenobarbital, diazepam and valproic acid to control the myoclonic seizures in our patient. However, these were not effective. At the beginning phase of the refractory seizures, we also used another benzodiazepine drug, midazolam. Midazolam was effective for controlling status epilepticus and was administered for 1 month. At 6 months of age, we resumed administration of midazolam to control the refractory seizures; however the treatment was not effective. We did not know why midazolam was no longer effective. Our patient also had life-threatening complications, including status epilepticus, apnea, and tachycardia. Therefore, we used a relatively high dose of lorazepam. She was successfully treated with this anticonvulsant and had no side-effects. Previous reports demonstrated that lorazepam is effective in status epilepticus and refractory seizures.[4] Lorazepam has a longer duration of action than diazepam and readily crosses the blood–brain barrier after rapid absorption. These pharmaceutical properties might be associated with the efficacy of lorazepam. In the present case, we assumed the suitable oral dose of lorazepam for status epilepticus as previously described.[8] We thought that the oral equivalent (1.2 mg/kg/dose) of the intravenous dose used elsewhere (0.1 mg/kg/dose) was too high compared with the recommended intravenous dose (0.1 mg/kg/dose). Thus, we administrated 2 mg/kg/24 h (1.0 mg/kg/dose every 12 h). This dose was very effective for the suppression of status epilepticus and did not produce any side-effects, such as respiratory depression or hypotension. At 30 min and at 7 h after administration of this dose, the serum concentrations of lorazepam were 715.4 and 415.6 ng/mL, respectively. This concentration may be high (therapeutic range: 50–240 ng/mL). However, relatively high doses of various anticonvulsants may be used to suppress the status epilepticus.[9,10], Next, the ordinary therapeutic dose of lorazepam must be determined. We measured serum concentration of the drug after oral administration. Thirty minutes after oral administration of 0.2 mg/kg/dose, the serum drug concentration was 209.2 ng/mL, suggesting that this was a suitable oral dose (Fig. 3). Few reports on high-dose lorazepam have been published.[11] Additional large studies regarding oral lorazepam may be needed, although our dose was effective in the present case.

The recommended treatment for convulsive status epilepticus in children in Japan includes intravenous diazepam as a first-line treatment, followed by intranasal or intravenous midazolam, intravenous phenytoin, and intravenous barbiturates.[5] Previous experiences have shown that intravenous diazepam was ineffective in these cases. In accordance with these regulations and findings, we administered midazolam initially but found it to be ineffective. In addition, in the present case, we could only administer the drugs with a small digital vein, because it was difficult to prick other peripheral veins with a needle. Induced vasculitis may be necessary to administer phenytoin via a small vein route. Thus, we did not use this drug. Other anticonvulsants, such as paraldehyde and lidocaine (xylocaine) may be applicable. However, paraldehyde is not available in Japan. In addition, lidocaine is not definitively recommended, although it is included in the guidelines for the medical treatment of pediatric status epilepticus.[5]

It has been suggested that HPE involves midline structural abnormalities that may include endocrinological disturbance. In such cases, it is possible that lorazepam works against intractable seizures through a unique mechanism. However, the detailed mechanisms of anticonvulsants are not yet known. A large number of case studies involving the administration of lorazepam in refractory seizures with cerebral dysgenesis, including HPE, may be needed to address these issues.

## References

1 Jones KL. Holoprosencephaly sequence. In: Jones KL. (ed). *Smith's Recognizable Patterns of Human Malformation*, 6th edn. W.B. Saunders, Philadelphia, 2005; 701–3.
2 Hahn JS, Plawner LL. Evaluation and management of children with holoprosencephaly. *Pediatr. Neurol.* 2004; **31**: 79–88.
3 Plawner LL, Delgado MR, Miller VS *et al.* Neuroanatomy of holoprosencephaly as predictor of function: Beyond the face predicting the brain. *Neurology* 2002; **59**: 1058–66.
4 Appleton R, Macleod S, Martland T. Drug management for acute tonic-clonic convulsions including convulsive status epilepticus in children. *Cochrane Database Syst. Rev.* 2008; (3): CD001905.
5 Osawa M. *Research committee on clinical evidence of medical treatment for status epilepticus in childhood in 2004 Japan: National Institutes of Health.* [Accessed 26 January 2006.] 2004. Available from URL http://mhlw-grants.niph.go.jp/niph/search/NIDD00.do
6 Schinzel A. Chromosome 19. In: Schinzel A. (ed). *Catalogue of Unbalanced Chromosome Aberrations in Man*, 2nd edn. Walter de Gruyter, Berlin, 2001; 783–90.

7 Appleton R, Choonara I, Martland T, Phillips B, Scott R, Whitehouse W. The treatment of convulsive status epilepticus in children. The Status Epilepticus Working Party, Members of the Status Epilepticus Working Party. *Arch. Dis. Child.* 2000; **83**: 415–9.
8 Yaster M, Kost-Byerly S, Berde C, Billet C. The management of opioid and benzodiazepine dependence in infants, children, and adolescents. *Pediatrics* 1996; **98**: 135–40.
9 Noerr B. Lorazepam. *Neonatal Netw.* 2000; **19**: 65–7.
10 Riviello JJ Jr. Status Epilepticus. In: Wyllie E, Gupta A, Lachhwani DK, eds. *The Treatment of Epilepsy: Principles & Practice*, 4th edn. Lippincott Williams & Wilkins, Philadelphia, 2005; 605–22.
11 Reincke HM, Gilmore RL, Kuhn RJ. High-dose lorazepam therapy for status epilepticus in a pediatric patient. *Drug Intell. Clin. Pharm.* 1988; **22**: 889–90.

# Intestinal amebiasis with Henoch–Schönlein purpura

Young Ok Kim,[1] Young Seok Choi,[1] Young Ho Won,[2] Young Dae Kim,[3] Young Jong Woo,[1] Hee Jo Back,[1] Young Kuk Cho,[1] Dong Kyun Han[1] and Eun Song Song[1]

*Departments of [1]Pediatrics, and [2]Dermatology, School of Medicine, Chonnam National University and [3]Division of Gastroenterology, Department of Internal Medicine, College of Medicine, Chosun University, Gwangju, Korea*

**Key words** Henoch–Schönlein purpura, intestinal amebiasis.

Henoch–Schönlein purpura (HSP), which is characterized by non-thrombocytopenic, purpuric and systemic vasculitis of the small vessels of the skin, joints, gastrointestinal tract, and kidney, is often reported to be associated with various infections such as β-hemolytic streptococcus, varicella, rubella, rubeola, hepatitis B, mycoplasma, mumps, measles, human parvovirus B19, coxsackie virus, adenovirus, salmonella, clostridium, tuberculosis, and HIV.[1–4] HSP, however, is rarely reported in association with amebiasis.[3–6]

Only three reports have been found in a MEDLINE search: two in the English-language literature[4,5] and one in the Japanese-language literature.[6] The two in the English-language literature were both from Turkey.[4,5] Here, we report a very rare case of intestinal amebiasis associated with HSP; the patient had abdominal pain and bloody diarrhea preceding the skin lesions, leading to diagnostic difficulties.

## Case report

An 8-year-old girl, who was previously healthy, presented with a 3 day history of abdominal pain with watery diarrhea. The abdominal pain was colicky and diffuse, but more severe in the epigastric and periumbilical areas. Nausea, vomiting and fever, however, did not develop. Initial physical examination indicated a lethargic and irritable child, who had tenderness but no rebound tenderness on palpation of the abdomen. The medical and dietary history was unremarkable.

Laboratory studies of blood and urine included a white blood cell count, platelet count, coagulation profiles, serum electrolytes, acute phase reactants, serum transaminases, renal function tests, total bilirubin, amylase, the Widal test, and urinary analysis; all of the test results were normal but the stool was positive for occult blood and pus cells. Additional laboratory studies indicated positive ameba antibodies in the serum, and the wet mount confirmed amebic cysts in the stool. Other serology, however, such as serum anti-nuclear antibody, IgG, IgA, IgM, IgE, and gastrin were normal. The tests for other infectious agents were negative and included a tuberculin skin test; stool cultures, antigens (on enzyme-linked immunosorbent assay) or toxins for salmonella species, shigella species, pathogenic *Escherichia coli*, Vibrio species, *Yersinia enterocolitica*, *Clostridium difficile*, *C. perfringens*, rotavirus, norovirus, adenovirus, and astrovirus; non-amoebic ova or parasites in stool; blood or urine culture.

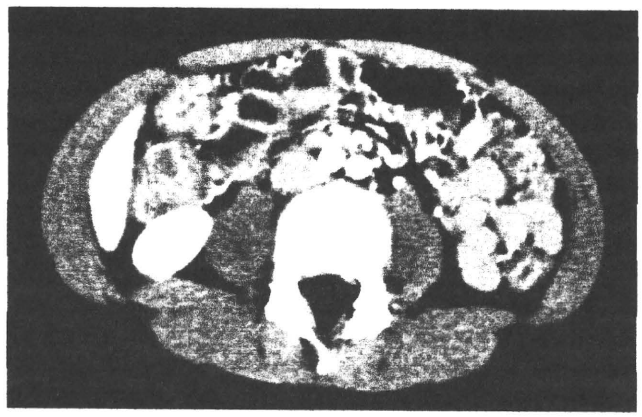


**Fig. 1** Abdominal computed tomography on admission showing diffuse circumferential wall thickening of the colon with enhanced mucosa and edematous submucosa, which was suggestive of colitis.

Abdominal computed tomography on admission indicated diffuse circumferential wall thickening of the colon with enhanced mucosa, edematous submucosa and a small amount of ascites, which was suggestive of colitis (Fig. 1). Even though an inflammatory etiology was suspected, empirical antibiotics

Correspondence: Young Ok Kim, MD, PhD, Department of Pediatrics; School of Medicine: Chonnam National University; 8 Hak-dong, Dong-gu, Gwangju 501-757, Korea. Email: ik052@unitel.co.kr

Received 15 October 2008; revised 7 September 2009; accepted 14 October 2009.
doi: 10.1111/j.1442-200X.2010.03116.x

Transworld Research Network
37/661 (2), Fort P.O.
Trivandrum-695 023
Kerala, India

Clinical Application of Molecular Diagnosis -Cancer, Radiation Effects, and Human Diseases-
ISBN: 978-81-7895-408-0 Editors: Eiso Hiyama and Keiko Hiyama

# 8. Diagnostic and prognostic molecular markers in neuroblastoma

Eiso Hiyama[1] and Keiko Hiyama[2]
[1]*Natural Science Center for Basic Research and Development*, [2]*Department of Translational Cancer Research, Research Institute for Radiation Biology and Medicine Hiroshima University, Hiroshima, Japan*

**Abstract.** Neuroblastoma, one of the common malignant childhood tumors, arises from neuroblast cells derived from the neural crest and destined for the adrenal medulla and the sympathetic nervous system and shows remarkable biological heterogeneity, resulting in favorable or unfavorable outcomes. Some tumors make rapid progress with a fatal outcome. In other instances, the tumors regress spontaneously in infants or differentiate into a benign ganglioneuroma in older patients. This heterogeneity within neuroblastoma depends on the molecular characteristics of tumor cells. Several distinct genomic alterations have been found in neuroblastoma, including *MYCN* amplification, DNA ploidy, deletion of the short arm of chromosome 1, gain of chromosome 17q, and deletion of 11q. Also, dysregurated expression was found in genes related to cellular growth, differentiation, and apoptosis of neural network including *NTRK1* tyrosine kinase receptor and telomerase activity. This chapter demonstrates diagnostic and prognostic molecular makers that would distinguish and explain the extensive heterogeneity of neuroblastoma. It would lead to more risk-adapted therapies according to such molecular markers by which individual neuroblastomas are biologically characterized.

Correspondence/Reprint request: Dr. Eiso Hiyama, Natural Science Center for Basic Research and Development Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima, 734-8551, Hiroshima, Japan
E-mail: eiso@hiroshima-u.ac.jp

## Introduction

Neuroblastoma, which is derived from the neuroblast in the neural crest, is the most common malignant solid tumor in children. The incidence of neuroblastoma is about 1 case per 7,000 born babies a year [1]. More than 90% of children with neuroblastoma are diagnosed within the first until 15 years of age. The tumors exhibit three distinct patterns of clinical behavior: life-threatening progression, spontaneous regression and maturation to ganglioneuroma. Many patients who are diagnosed at more than one year of age have advanced neuroblastomas with metastasis and usually show poor outcome despite multimodal therapies. On the other hand, some tumors undergo complete regression without any treatment. Spontaneous regression usually occurs as part of a clinically identified entity designated as stage 4S which has a primary tumor localized in the adrenal gland and metastasis restricted in the liver, bone marrow, and/or skin [2-4]. In favorable tumors, maturation to benign ganglioneuroma is usually observed after chemotherapy [5]. Since more than 80% of neuroblastomas produce catecholamine, their metabolites (vanillylmandelic acid and homovanillic acid) are detectable in the urine and mass-screening projects to detect earlier stage-neuroblastomas have been carried out in some countries including Japan [6-9]. Our recent report revealed that the incidence of this disease increased more than two fold during the period of mass-screening, whereas both of the advanced neuroblastoma incidence in elder children and the cumulative mortality rate of total children reduced significantly. These findings indicate that a large number of neuroblastomas occur in infants and spontaneously regress or mature behind the scenes without clinical detection, while some of them may progress into malignant phenotypes [6]. These projects also gave us much for solving the biological problems in neuroblastoma. Transition from favorable type to unfavorable type has been clearly suggested [10], leaving the mechanism of malignant alteration of this tumor to be elucidated.

Clinically, biological heterogeneity of neuroblastoma showed contradictory outcomes: some tumors regress or mature and other instances advance aggressively. It is essential to distinguish progressive tumors from favorable tumors, because multimodal appropriate therapies are necessary to improve the prognosis of the patients with progressive tumors, while aggressive therapy should be avoided in the patients with favorable tumors to reduce side effects, late complications and medical expense. Clinical evaluation for this differential diagnosis is based on INSS (International Neuroblastoma Staging System), age at diagnosis, and histological classification by INPC (International Neuroblastoma Pathological Classification), and only more recently several

biological markers have been incorporated [11]. Histological classification by INPC according to the degree of ganglionic differentiation and Schwannian stroma has been widely accepted as an important prognostic factor [12, 13]. Recently, image defined risk factors (IDRF) were reported to be useful in predicting risk and completeness of operation in localized tumors [14]. Serum neuron-specific enolase (NSE) is nonspecific for neuroblastomas, however, increased levels of NSE (>100 ng/ml) have been shown to significantly correlate with advanced-stage neuroblastomas and poor patient survival rates [15]. Lactate dehydrogenase (LDH) is also not specific for neuroblastoma, but high levels (>1500 U/ml) are associated with rapid cellular turnover and poor prognosis [16, 17]. Several patients with neuroblastoma may have increased levels of iron-binding protein ferritin. High ferritin levels (>142 ng/ml) are commonly seen in patients with advanced-stage disease and correlate with unfavorable outcome [18]. Still we cannot distinguish unfavorable and favorable tumors using these markers alone. Thus the development of more efficient markers is essential.

To distinguish the molecular and biological heterogeneity of neuroblastoma clearly, numerous multilateral approaches have been performed and several distinct differences have been found including *MYCN* amplification, DNA ploidy, chromosomal loss and gain, expression of *NTRKA*, telomerase activation, and others. In this review, molecular heterogeneity of neuroblastoma is summarized and clinical application of these molecular markers for appropriate treatments is presented. This approach will lead to the identification of the molecular markers that distinguish progressive tumors at diagnosis as well as the molecular targets to treat these tumors. In addition, it will provide insights into mechanisms of malignant transformation, progression, spontaneous regression and maturation in neuroblastoma.

## Biological and molecular markers

### 1. *MYCN* Amplification

Amplification of the *MYCN* gene is one of the most distinguished genetic aberrations in neuroblastomas. Cytogenetic analysis of neuroblastoma cells with amplified *MYCN* gene shows chromosomal abnormalities such as homogeneously staining region (HSR) or double minutes (DMs). DMs are usually detected in primary tumors (Fig. 1, top left), and HSR, which is generally located on a chromosome different from the *MYCN* resident site 2p24, is found in cell lines [19, 20].

Amplified *MYCN* was observed in 30-50% of advanced neuroblastomas, but was rarely detected in early-stage tumors [19, 20]. In patients with non-

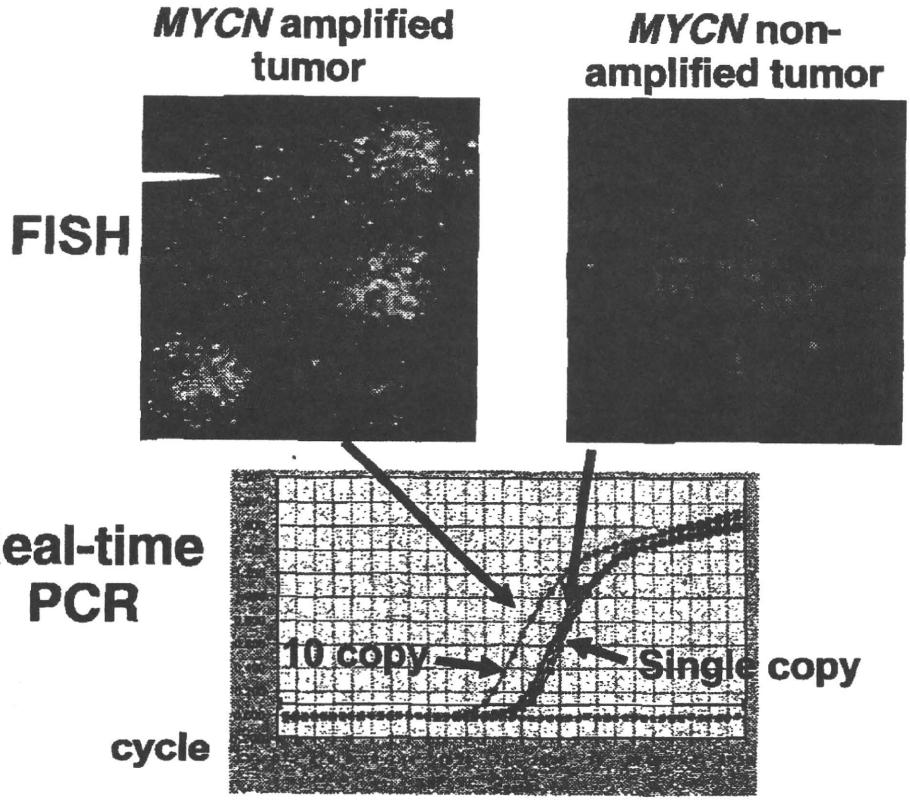



**Figure 1.** Detection of *MYCN* Amplification in Neuroblastoma. Copy number of the *MYCN* gene can be determined by fluorescence *in situ* hybridization (FISH, upper), and quantitative real-time PCR (TaqMan®, lower).

amplified *MYCN* tumors without metastasis, overall survival was approximately 90% over a 5-year period, but less than 30% of patients survived a 2-year period when *MYCN* was amplified over ten copies [10, 11]. Thus, amplified *MYCN* is established as a powerful clinical marker of high-risk neuroblastoma and is the sole genetic marker that has been used for treatment stratification in neuroblastoma clinical trials [21-24]. Recent studies revealed that the patients who had tumors with amplified *MYCN* gene between four and nine copies also showed poor outcomes [25, 26]. Moreover, evaluation of primary tumors at single cell level by FISH (fluorescence *in situ* hybridization) analysis revealed that the copy numbers in individual cells from *MYCN*-amplified tumors typically stray widely from that estimated by molecular analyses of the tumor as a whole [25, 27]. These suggested that amplified *MYCN* gene characterizes a subset of neuroblastoma with extremely aggressive growth potential regardless of the copy number of *MYCN* gene. Now, evaluation for *MYCN* amplification is performed by FISH and/or quantitative PCR for DNA copies (Fig. 1), while it used to be done by Southern blotting. Recently, using serum DNA-based real-time quantitative

PCR with a single-copy reference gene, detection of *MYCN* amplification in serum DNA is a valuable diagnostic tool to discriminate the patients with *MYCN* amplified tumors from other patients [28]. This method might become a powerful diagnostic tool as well as a promising indicator of therapeutic efficacy and relapse in the follow-up of patients with *MYCN* amplified tumors.

The copy number of amplified *MYCN* gene, ranging from three-fold to more than 500-fold, has been considered to be consistent within a tumor: not only at different tumor sites, but also at different times *in vivo* [29]. This suggests that amplified *MYCN* is generally present at the time of diagnosis. However, there have been some reports showing that the *MYCN* gene was amplified during progression of neuroblastoma [30]. Our unique case which had a rearrangement in the half copies of amplified *MYCN* gene in one metastatic lesion remaining another without rearrangement suggested that the *MYCN* amplification had occurred during tumor progression [31]. Epidemiological analysis on the large cohort that underwent mass-screening revealed that advance tumors in older children decreased and favorable tumors in infants without *MYCN* amplification increased [6]. In the majority of neuroblastomas with amplified *MYCN* is accompanied by 1p deletion and/or 17q gain in the same tumors. These phenomena imply that *MYCN* amplification is a later event in the sequence of genetic aberrations underlying neuroblastoma progression [32].

The size of amplicon with amplified DNA encompassing *MYCN* ranges from 100 kb to more than 1 Mb [33], suggesting the possibility that additional genes are co-amplified. The *DDX1* gene and neuroblastoma-amplified protein (*NAG*) gene were reported to be coamplifed in 50-70% of the tumors with *MYCN* amplification but not in those without concomitant *MYCN* amplification, indicating that *MYCN* is functionally responsible for the maintenance of the amplicon. Recently, *ALK* (anaplastic lymphoma receptor tyrosine kinase) gene located near the *MYCN* gene was found to be activated by amplification as well as by mutations in the tyrosine kinase domain in neuroblastoma [34]. The germline mutation in *ALK* gene is considered to be the major cause of familial neuroblastoma predisposition, and amplification of this gene is the second hit that promote neuroblastoma development [35]. Thus, the amplification of the 2p23-24 loci is a critical event in distinguishing neuroblastoma biologies.

High levels of *MYCN* expression are observed in *MYCN*-amplified tumors and seem to contribute to tumorigenesis. However, there is conflicting experience regarding the potential prognostic significance of *MYCN* expression without *MYCN* amplification: prognosis of patients with higher *MYCN* expression without amplification was better than those with lower

expression, because *MYCN* and *NTRK1* (encoding TrkA) expression showed a linear correlation in neuroblastoma lacking *MYCN* amplification [36]. While one study demonstrated no significant correlation of *MYCN* mRNA expression and N-myc protein levels with prognosis in patients without *MYCN* amplification [37], we and others reported that high levels of *MYCN* expression were associated with poor outcomes in patients with *MYCN* non-amplified tumors, especially in patients older than 1 year [38, 39].

## 2. DNA ploidy

Cytogenetic and flow cytometric analyses have been used for evaluating DNA ploidy in neuroblastoma. Flow cytometric analysis revealed that hyperdiploidy, mostly the near-triploidy, is mainly observed in favorable tumors of younger patients, whereas diploid is usually detected in advanced tumors with unfavorable outcomes [40] (Fig. 2). In addition, in children 12-24 months of age, diploidy predicted resistance to chemotherapy, whereas half of the patients with hyperdiploidy achieved long-term disease-free survival. The DNA ploidy did not have its prognostic significance for patients over 2 years of age [41].

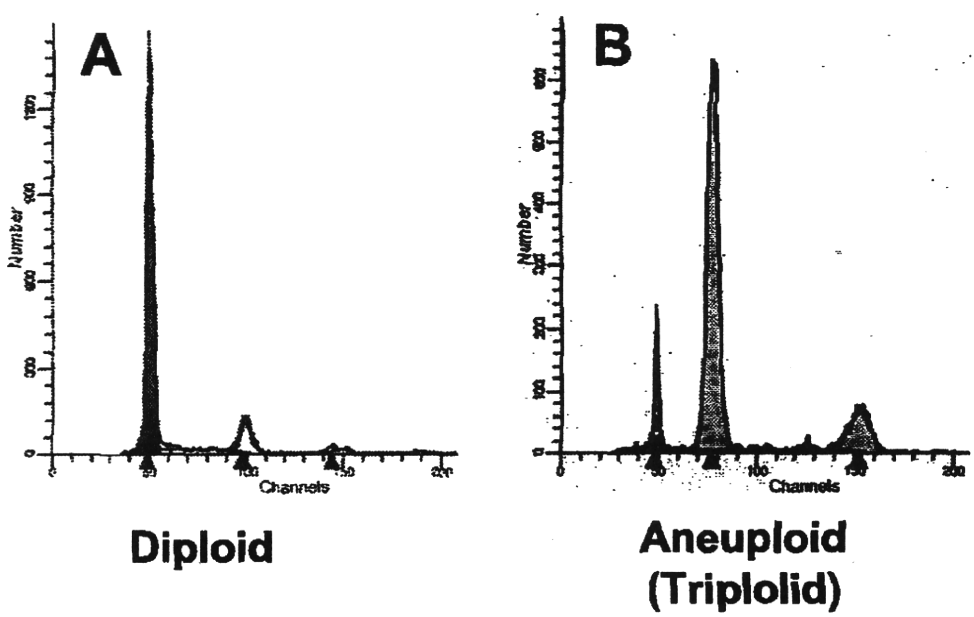



**Figure 2.** DNA pliody evaluated by flow cytometric analysis. Diploid cells demonstrate a G0/G1 large peak at 50 in X axis (DNA content) and a G2 small peak at 100 (A). Normal diploid cells contaminated in tumor tissues demonstrate a small G0/G1 peak at 50 and the aneuploid tumor cells demonstrate a large G0/G1 peak at 75 and a small G2 peak at 150 (B). The ratio of tumor G0/G1 DNA content / normal G0/G1 DNA content was near 1.5, defining this tumor as triploid.