Removing intensity effects and identifying significant genes for Affymetrix arrays in macrophage migration inhibitory factor-suppressed neuroblastoma cells

Jianqing Fan†‡, Yongxiong Chen§, Hiu Man Chan§, Paul K. H. Tam§, and Yi Ren‡§

†Department of Operation Research and Financial Engineering, Princeton University, Princeton, NJ 08544; and §Department of Surgery, University of Hong Kong, Hong Kong, China

Communicated by Peter J. Bickel, University of California, Berkeley, CA, October 24, 2005 (received for review March 26, 2005)

A semilinear in-slide model is introduced to remove the intensity effect in the scanning process. It is demonstrated that the intensity effect can be estimated accurately and removed effectively. This normalization step is vital for Affymetrix arrays to reveal relevant biological results when comparing gene expression in multiple arrays. The normalized expression ratios are analyzed further by a modified two-sample t test along with a sieved permutation scheme for computing P values. The improved specificity and sensitivity are demonstrated by using a study on the impact of macrophage migration inhibitory factor (MIF) reduction in neuroblastoma cells. With semilinear in-slide model analysis, expression of 166 genes was altered with a P value no greater than 0.001. Among those genes, 44 were altered >2-fold. MIF-regulated genes associated with tumor development including IL-8 and C-met, which are overexpressed in many tumors, were down-regulated in MIF-reduced cells. On the other hand, some tumor-suppressor genes such as EPHB6, visinin-like protein 1 (VSNL-1), and BLU were up-regulated in MIF-reduced cells. In addition, we demonstrated that down-regulation of MIF expression could result in a reduction in cell proliferation and tumor growth in vitro and in vivo. Our data not only demonstrate that targeting MIF expression is a promising therapeutic strategy in human neuroblastoma therapy but also indicate the MIF target genes for additional study.

MIF | normalization | significance analysis | SLIM

Affymetrix GeneChip arrays (1, 2) have been widely used for monitoring mRNA expression in many areas of biomedical research. The high-density oligonucleotide array technology allows researchers to monitor tens of thousands of genes in a single hybridization experiment as they are expressed differently in tissues and cells. The expression profile of an mRNA molecule of a gene is obtained by the combined intensity information from probes in a probe set, which consists of 11–20 probe pairs of oligonucleotides of 25 bp in length, interrogating a different part of the sequence of a gene. Several techniques have been proposed for extracting expression profiles from the information at probe level. These techniques include the detection signals that are prominently featured in Affymetrix Micorarray Suite 5.0 (MAS 5.0) software, the model-based expression index (3), and the robust multichip average (4, 5). A comprehensive review and comparisons of various expression measures are given in refs. 6 and 7.

To account for the overall brightness of scanned images and other experimental variations such as differences in sample preparation and array production, normalization is needed for direct array-to-array comparisons. Several methods (7–10) have been proposed for normalizing data at probe level, and their impacts on the analysis of gene expressions have been examined. However, most researchers use detection signals from MAS 5.0 as starting points of their investigation. Yet, as reported below, the techniques for probe-level normalization are not effective at the detection-signal level. Indubitably, there are strong demands for normalization methods based on the output of detection signals. Indeed, we are not aware of any methods for normalization of Affymetrix arrays at the detection-signal level.

Proper normalization is critical for revealing relevant biological results. Although the detection algorithm for MAS 5.0 was designed carefully, systematic biases such as the intensity effect still exist, and they should be removed before making multiple-array comparisons. One aim of this article is to develop a normalization and analysis system based on sound statistical principles. The observed detection signals are decomposed into three components: the treatment effect, intensity effect, and random noise. With very mild statistical assumptions, the intensity effect can be estimated accurately and removed effectively as long as there are replications in treatment arrays. The normalized log ratios are then analyzed by a modified two-sample t test. The P values are estimated by a permutation technique that aims at reducing biases of the estimates and effective use of data. The resulting comprehensive system will be called the semilinear in-slide model (SLIM) system.

The specificity and the statistical power of the SLIM system are demonstrated by our application to microarray analysis of neuroblastoma cells with macrophage migration inhibitory factor (MIF) being reduced. Three treatment and three control arrays were analyzed by the SLIM. Among ~200 genes that are identified significantly differently expressed, 12 important ones are selected for additional biological confirmation by using real-time quantitative RT-PCR, Western blot analysis, and ELISA. All of them were biologically confirmed. On the other hand, without normalization, two important genes were missed, giving a missed-discovery rate of 17%. By using the MAS 5.0 system, the missed-discovery rate was 29% (see Discussion for additional details).

Neuroblastoma is a malignant tumor of neural crest origin that may arise anywhere along the sympathetic ganglia or within the adrenal medulla. Neuroblastoma is the most frequent solid extracranial neoplasia in children and is responsible for ~15% of all pediatric cancer deaths (11). Spontaneous regressions and differentiation are common in infants and in early-stage tumors, whereas neuroblastoma is extremely aggressive in older children with late-stage tumors (12). MIF has been observed recently as a protumorigenic factor (13) in addition to its effects on proinflammatory and immune responses. Although a growing body of data have been gathered on the expression of MIF in several tumors, the exact mechanism of its function is unknown. Our previous results revealed that MIF was highly expressed in neuroblastoma and MIF could stimulate oncogene N-myc expression and up-regulate the expression of angiogenic factors (14). Furthermore, MIF is able to regulate the expression of genes that are related to tumor cell proliferation, migration, and antiapoptosis (15). These results sug-

Conflict of interest statement: No conflicts declared.

Abbreviations: MAS 5.0, Affymetrix Micorarray Suite, version 5.0; SLIM, semilinear in-slide model; MIF, macrophage migration inhibitory factor; FDR, false-discovery rate; VSNL-1, visinin-like protein 1.

© 2005 by The National Academy of Sciences of the USA
gested that MIF may play an important role in the development of neuroblastoma.

The biological aim of this study was to investigate whether MIF would be a target for limiting neuroblastoma development, that is, whether reduction of MIF expression could control tumor proliferation and tumorigenicity in neuroblastoma. A decrease in MIF expression in neuroblastoma cells was observed after transfection with MIF antisense expression vector. We demonstrated that down-regulation of MIF expression could result in a reduction in cell proliferation and tumor growth in vitro and in vivo. To better understand the MIF molecular mechanism in neuroblastoma, we used Affymetrix GeneChip to identify genes in MIF-reduced neuroblastoma cells. As reported below, gene-expression profiling analyzed by SLIM, validated at mRNA and protein levels for several gene products, provides a comprehensive view of molecular changes in MIF-reduced tumor cells and reveals possible underlying molecular unidentified pathogenesis.

Materials and Methods

Cells. Human neuroblastoma cell line SK-N-DZ was obtained from the American Type Culture Collection. Cells were maintained in DMEM supplemented with 10% FBS (GIBCO/BRL-Life Technologies, Grand Island, NY) and 1% penicillin-streptomycin in a humidified incubator (95% air

Preprocessing of Detection Signals. The detection signals of 22,283 probe sets were extracted from the Affymetrix MAS 5.0 software, which resulted in three expression profiles, respectively, from the control and treatment groups (three cell lines 3, 5, and 24). The repeatability, assessed by the coefficient of variation, is low for genes with small detection signals. Thus, only genes with all detection signals >50 were considered, which filtered 37.3% of genes, with 13,980 genes remaining for additional investigation. The percentage is actually smaller than those called “absent” by the Affymetrix MAS 5.0 software (somewhat >50%). The remaining genes had much smaller coefficients of variations (better repeatability). The repeatability in the control group is somewhat better but is in the same bulk.

SLIM Normalization. Let \( x_{gi} \) and \( y_{gi} \) be the log-detection signal of the \( g \)th probe set in the \( i \)th control and treatment array, respectively. Denote by \( \bar{x}_g \) the average of the log-detection signals over the control arrays for the \( g \) probe set. Similar to that in ref. 16, we first computed the log intensities and log ratios, respectively, as follows:

\[
A_{gi} = \frac{y_{gi} + \bar{x}_g}{2} \quad \text{and} \quad M_{gi} = y_{gi} - \bar{x}_g, \tag{1}
\]

for \( i = 1, \ldots, n \) and \( g = 1, \ldots, G \). In our particular application, \( n = 3 \) and \( G = 13,980 \). Fig. 1A depicts the log ratios versus log intensities for a treatment array, along with the lowest fit for the conditional mean and conditional SD curves (17, 18) for the array.

The SLIM in ref. 15 and two-way semilinear model in refs. 19 and 20 were extended to remove the intensity effect. Let \( \mu_g \) be the treatment effect on gene \( g \) and \( f_i(\cdot) \) represent the array-dependent intensity effect. The model

\[
M_{gi} = \mu_g + f_i(A_{gi}) + e_{agi}, \quad g = 1, \ldots, G; i = 1, \ldots, n \tag{2}
\]

was used to assess the intensity effect \( f_i(\cdot) \) and treatment effect \( \mu_g \) on genes. Once the intensity effect is estimated, the normalized log ratios are given by

\[
m_{gi} = M_{gi} - \hat{f}_i(A_{gi}), \tag{3}
\]

for an estimated intensity effect \( \hat{f}_i(\cdot) \).

The parameters in SLIM can be estimated by the profile least-squares method. Let \( M, \mu, \) and \( f \) denote, respectively, the vector of length \( G \) for the log ratios, treatment effect, and intensity effect. For the given estimated treatment effect \( \hat{\mu} \), the intensity effect is estimated by smoothing \( \hat{M} - \hat{\mu} \) on the intensities \( \{A_{gi}\} \) separately for each array, resulting in

\[
\hat{f}_i = S_i(M_i - \hat{\mu}), \tag{4}
\]

for a smoothing operator \( S_i \) (see, e.g., ref. 21), whereas for given estimated intensity effect functions \( \hat{f}_i \),

\[
\hat{\mu} = n^{-1} \sum_{i=1}^{n} (M_i - \hat{f}_i). \tag{5}
\]

Solving Eqs. 3 and 4 yields the explicit solution

\[
\hat{\mu} = (I_G - S) \sum_{i=1}^{n} (I_G - S_i)M_i, \tag{6}
\]

and \( \hat{f}_i \) in Eq. 3, where \( I_G \) is the \( G \times G \) identity matrix and \( S = n^{-1} \sum_{i=1}^{n} S_i \). Inverting the matrix in the order of tens of thousands is not practically feasible, and the solution can be computed by the G-Seidel or back-fitting algorithm (21, 22), which iteratively computes Eqs. 3 and 4.
A Theoretical Analysis of SLIM. For the normalization purpose, the number of nuisance parameters \( \mu_g \) is large, consisting of \( 1/n \) of the total sample size \( nG \). Can the functions \( f_j \) be estimated accurately? In a different context and somewhat different model, it was demonstrated (20, 22) that the rate of convergence for estimating \( f_j \) is \( G^{-1/2} \). Because \( G \) is large, the intensity functions can be estimated accurately.

To appreciate the estimability of the intensity effect in Eq. 1, consider the case that \( n = 2 \). Let \( z_g = M_{g1} - M_{g2} \). Then,
\[
z_g = f_1(A_{g1}) - f_2(A_{g2}) + e_g^* \quad g = 1, \ldots, G,
\]
where \( e_g^* = e_{g1} - e_{g2} \) is the new random noise. Eq. 6 is an additive model (21), and its components \( f_1 \) and \( f_2 \) can be estimated as though one of them were known (23–26). The cost for estimating parameters \( \mu_g \) reflects in the increase of the variability of the noise \( \{e_g^*\} \).

High correlation between the intensities of arrays \{\text{A}\} and \{\text{H}\}, however, reduces the accuracy of the estimated intensity effect.

To gain an idea of the degree of accuracy in the estimation of the intensity effect, we simulated 10 data sets from the model in Eq. 1, with true parameters and functions taken from those estimated from our data. The new random noises in Eq. 1 were generated from the normal distribution with a mean of 0 and SD of 0.12. The mean absolute deviation error for the estimation of the intensity effect is \( 0.0134 \pm 0.0003 \), which is negligible for our application. See ref. 22 for more intensive simulation studies.

Analysis of Treatment Effect. After obtaining the normalized ratio \( m_{g1} \), significant genes were obtained by testing the following hypothesis separately:
\[
H_0: \mu_g = 0 \quad \text{vs.} \quad H_1: \mu_g \neq 0.
\]
As to be demonstrated, it is reasonable to assume that the errors in Eq. 1 are homoscedastic across arrays, because the intensity \( A_g \) does not vary much across arrays. Indeed, we adapted the weighted \( t \) statistic in ref. 15 to the current problem and did not find significantly different results.

Let \( s_x^2 \) and \( s_y^2 \) represent the sample variance of \( \{x_g\} \) and \( \{y_g\} \), which are of sample sizes \( n_1 \) and \( n_2 \), respectively. Regarding the log intensities \( \{x_g\} \) and \( \{y_g\} \) as two random samples from two populations with the same variance, a test statistic (15, 27) for Eq. 7 is
\[
t_g = \bar{m}_g/(SE_g + s_0),
\]
where \( \bar{m}_g \) is the average of the normalized log ratios in Eq. 2, and
\[
SE_g = \hat{\sigma}_g \sqrt{1/n_1 + 1/n_2} \quad \text{with} \quad \hat{\sigma}_g^2 = (n_1 + n_2 - 2)^{-1}(n_1 - 1)s_{x_g}^2 + (n_2 - 1)s_{y_g}^2.
\]

The constant \( s_0 \) is used here to guard against the zero denominator in Eq. 8. Note that \( m_g \) is the difference between the averages of treatment and control arrays with the intensity effect removed.

The constant \( s_0 \) plays two important roles in identifying significant genes. First, it reduces the false-discovery rate (FDR) (28, 29) by an appropriate choice of \( s_0 \). Because tens of thousands of genes are involved, there is a high chance that some of the SDs \( \hat{\sigma}_g \) will be extremely small, which leads to a large test statistic and a possible falsely discovered gene. Second, a gene needs to satisfy a minimum fold-change requirement to be identified as a statistically significant gene. For example, if a gene is considered significant when \( |t_g| > 3 \), then the fold change \( |\bar{m}_g| > s_0 \). Thus, with the tuning constant \( s_0 \), the \( P \) values and the fold changes are both taken into consideration in one single statistic. In our implementation, \( s_0 = 0.07 \) was used.

Fig. 2. Effect of MIF suppression in vitro and in animal model. (A) Expression of MIF in control and MIF-suppressed neuroblastoma cells assessed by Western blot. (Upper) MIF has successfully been reduced compared with control. (Lower) Confirmation of the equivalent protein loading by using the housekeeping gene "\( \beta \)-actin" as internal control. (B) The average tumor growth curves for the MIF-reduced and control groups. The black curve indicates the control group (\( n = 10 \)), and the three colored curves indicate the MIF-reduced group (total sample size, \( n = 30 \)).

Results

MIF Expression in Cells Transfected with MIF Antisense Expression Vector. MIF expression in control (vector alone) and MIF antisense expression vector-transfected cells (clones 3, 5, and 24) was detected by Western blot analysis. The result showed that the amount of MIF was substantially reduced in three clones with MIF antisense vector-transfected cells compared to control cells (empty vector transfection; Fig. 2A). The protein expression levels of \( \beta \)-actin (used as internal control to confirm equivalent protein loading) in control and MIF antisense vector-transfected cells were nearly identical.

Tumor Growth in Nude Mice. The effect of suppressing MIF expression on tumor growth was further investigated in athymic nude mice. Six-week-old BALB/c nude mice were injected s.c. with \( 2 \times 10^6 \) SN-K-DZ cells transfected with vector alone (\( n = 10 \), control group) or MIF-reduced cells (i.e., clones 3, 5, and 24; \( n = 10 \) for each clone). The mice were monitored for tumor growth after 20 days of inoculation for 10 days. Tumor growth in two groups was monitored everyday by two-dimensional measurements of individual tumors from each mouse. Mice in the treatment group showed an obvious reduction in tumor growth compared to control mice (Fig. 2B).

Estimated Intensity Effect. Fig. 3A depicts the estimated intensity effects for three treatment arrays that are contrasted with the average of the control arrays. The intensity effect tends to be smaller in the middle than at tails. When the algorithm converges, the residuals \( \{e_g\} \) can be obtained. To examine the degree of heteroscedasticity for the significant analysis of the treatment effect, the squared partial residuals \( \{e_g^2\} \) were smoothed against the log intensity \( \{A_g\} \) for each treatment array, which results in three SD curves (Fig. 3B).
permutation test when the numbers of treatment and control arrays are even and large.

Genes with \( P \) values no greater than \( \alpha = 0.001 \) were selected as statistically significant genes. The FDR among the selected genes was estimated by \( \tilde{p} = (Ga)/N \), where \( N \) is the number of genes “discovered,” having \( P \) values no greater than \( \alpha \), and \( Ga \) is the expected number of falsely discovered genes. For example, at \( \alpha = 0.001 \), 166 genes were discovered, and false discovery of \( 0.001 \times 13,980 = 13.98 \) genes is expected, giving an FDR of 8.4%. This simple method is somewhat more conservative but is almost equivalent to the Benjamini–Hochberg method (28) for controlling FDR at the data-dependent level \( \tilde{p} \). In other words, if the Benjamini–Hochberg method is used with the FDR controlled at \( \tilde{p} \), the same set of genes will be selected frequently. Our method provides an intuitive understanding to the Benjamini–Hochberg method and also gives statistical justification (28, 32) for our method.

To appreciate this connection, recall that to control FDR at level \( \tilde{p}\ ), the Benjamini–Hochberg method selects

\[
k = \max\{i:p_{(i)} \leq ip/G\}
\]

genes with the smallest \( k \) \( P \) values, where \( N \leq p_{(1)} \leq p_{(2)} \leq \ldots \leq p_{(G)} \) are the order statistic of the \( P \) values. Now, if we take \( p = \tilde{p} = Ga/N \), we would select

\[
k = \max\{i:p_{(i)} \leq i \alpha /N \} \geq N
\]

genes with the smallest \( P \) values. If \( p_{(n+1)} > \alpha (n+1)/N \), then \( k = N \); namely, we have selected the same set of genes as the Benjamini–Hochberg method with FDR controlled at \( \tilde{p} \). This simple connection can also provide deeper understanding to the probabilistic properties of the Benjamini–Hochberg method. For example,

\[
N = \sum_{i \leq S_0} I(p_i \leq \alpha) + \sum_{i \in S_0} I(p_i \leq \alpha), \tag{13}
\]

where \( S_0 = \{g: \mu_g = 0\} \) is the set of genes that do not express differently under treatment. The probabilistic aspect \( N \) can be understood easily. Assume ideally that there are \( G_1 \) very differently expressed genes. Under very mild conditions, \( N = Ga + G_1 \), where \( G_1 \) is the number of significantly differently expressed genes. Hence, \( \tilde{p} = Ga/(Ga + G_1) \). Hence, the smaller the \( \alpha \), the lower the FDR.

Table 1 summarizes the results by using the SLIM, the quantile normalization in refs. 5 and 7, and no normalization.

### Gene-Expression Profiling in MIF-Reduced Cells.

Of a total of \( \sim 22,283 \) probe sets, we discovered 166 genes showing significantly \( (P \leq 0.001) \) changed expression in MIF-reduced cells compared to control cells, 44 of which were altered >2-fold. The genes that were found to be up- or down-regulated are shown in Table 2.

#### Table 1. Number of genes that are up- or down-regulated after MIF being inhibited

<table>
<thead>
<tr>
<th>Methods</th>
<th>( \alpha = 5% ) Up</th>
<th>( \alpha = 1% ) Up</th>
<th>( \alpha = 0.5% ) Up</th>
<th>( \alpha = 0.1% ) Up</th>
<th>( \alpha = 0.05% ) Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLIM</td>
<td>1452</td>
<td>1754</td>
<td>21.8</td>
<td>321</td>
<td>101</td>
</tr>
<tr>
<td>Q-norm</td>
<td>1037</td>
<td>1046</td>
<td>33.6</td>
<td>212</td>
<td>73</td>
</tr>
<tr>
<td>No-norm</td>
<td>944</td>
<td>1149</td>
<td>33.4</td>
<td>190</td>
<td>62</td>
</tr>
<tr>
<td>CommGene(1)</td>
<td>718</td>
<td>894</td>
<td>217</td>
<td>135</td>
<td>62</td>
</tr>
<tr>
<td>CommGene(2)</td>
<td>914</td>
<td>1017</td>
<td>290</td>
<td>182</td>
<td>62</td>
</tr>
<tr>
<td>CommGene(3)</td>
<td>732</td>
<td>1006</td>
<td>219</td>
<td>137</td>
<td>45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Methods</th>
<th>( \alpha = 5% ) Down</th>
<th>( \alpha = 1% ) Down</th>
<th>( \alpha = 0.5% ) Down</th>
<th>( \alpha = 0.1% ) Down</th>
<th>( \alpha = 0.05% ) Down</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLIM</td>
<td>1452</td>
<td>1754</td>
<td>21.8</td>
<td>321</td>
<td>101</td>
</tr>
<tr>
<td>Q-norm</td>
<td>1037</td>
<td>1046</td>
<td>33.6</td>
<td>212</td>
<td>73</td>
</tr>
<tr>
<td>No-norm</td>
<td>944</td>
<td>1149</td>
<td>33.4</td>
<td>190</td>
<td>62</td>
</tr>
<tr>
<td>CommGene(1)</td>
<td>718</td>
<td>894</td>
<td>217</td>
<td>135</td>
<td>62</td>
</tr>
<tr>
<td>CommGene(2)</td>
<td>914</td>
<td>1017</td>
<td>290</td>
<td>182</td>
<td>62</td>
</tr>
<tr>
<td>CommGene(3)</td>
<td>732</td>
<td>1006</td>
<td>219</td>
<td>137</td>
<td>45</td>
</tr>
</tbody>
</table>

Number of common genes (CommGene) selected by the following combinations: SLIM normalization and the quantile normalization (Q-norm) \( (1) \), quantile normalization and no normalization (No-norm) \( (2) \), and SLIM normalization and no normalization \( (3) \).

---

Fig. 3. Effect of normalization of detection signals from \( \text{MAS}5.0 \) and summary of treatment effects on gene expressions. (A) Estimated intensity effects for three treatment arrays. They were removed from the log ratios for significant analysis. (B) Estimated SD curves for the residuals in the model in Eq. 1. They were used to judge whether weighted \( t \) statistics are needed. (C) Summary of fold changes \( (\mu_p) \) caused by the treatment. A reduced range is used to get a better view of the histogram. (D) The average log intensities for the treatment and control arrays along with genes (magenta) being selected at a level of \( \alpha = 0.001 \).
Validation of a Subset of Differentially Expressed Genes. To examine the reliability of the microarray data and identify molecules important in the pathogenesis of MIF, 12 sequences (six up-regulated and six down-regulated) were chosen from the lists of up- or down-regulated genes for additional validation. These 12 genes were chosen on the basis of their association with tumor growth, angiogenesis, and metastasis. Each gene or its protein products were validated by real-time quantitative RT-PCR, Western blot, or ELISA. All validation results for these 12 genes on changes in gene expression or protein products are consistent with Affymetrix microarray data. The most interesting genes confirmed were c-met, visinin-like protein 1 (VSNL-1), EPB6, BLU, and interleukin 8 (IL-8).

Discussion

Comparison with Other Methods. The effectiveness of SLIM was compared with no normalization and the quantile normalization (although it was proposed for normalization at probe level). The distributions for the differences between the log ratios after and before normalizations are summarized in Fig. 4. It is clear that the quantile normalization does not much alter the log ratios of detection signals; they are approximately the same as those without normalization. From Table 1, the quantile normalization and no normalization result in higher FDRs. They also discovered some different genes from those with normalization. For example, among 12 confirmed ones, two important genes (BLU and RAS) were accorded statistical significance if the SLIM normalization was used, but not under the quantile normalization or no normalization, which provides additional evidence to support the SLIM-normalization method.

The pairwise comparisons were also conducted by using the Affymetrix MAS 5.0 software. A total of nine possible pairwise comparisons were conducted. Among 12 confirmed ones, three important genes (VSNL-1, BLU, and c-met) were called “NC,” indicating no change consistently across all nine comparisons.

Among $9 \times 12 = 108$ comparisons, 29 were falsely called “NC,” which gives a missed-discovery rate of 27%. The lesson is very clear: even with the most advanced Affymetrix microarrays to date, normalization is required, and replications are needed for the discovery of slightly and moderately differently expressed genes. Replications are needed for both treatment and control arrays to reduce sampling variability.

Biological Implications. c-met was identified as an activated oncogene (33, 34). A large number of studies reveal that c-met is frequently expressed in carcinomas, in other types of human solid tumors, and in their metastases. Overexpression of c-met is often associated with poor prognosis (35). Our unpublished data showed that MIF expression significantly correlated with c-met protein expression in human neuroblastoma specimens ($r = 0.635; P = 0.002; n = 32$), which suggests that c-met and MIF may be closely related. In the present study, c-met was decreased in MIF-reduced cells, which strongly indicates that...
MIF may have a function in the up-regulation of C-met expression. VSNL-1 is expressed in the central nervous system, where it plays a crucial role in regulating cAMP levels, cell signaling, and differentiation. High-level VSNL-1 expression has been found in less aggressive and high-grade squamous cell carcinoma (36). These results indicate that VSNL-1 plays an important role in regulating tumor cell invasiveness and that its loss could help to enhance the advanced malignant phenotype.

EPHB6 is a favorable neuroblastoma gene. EPHB6 is down-regulated in the most aggressive neuroblastoma cell lines, and a prognostic indicator in neuroblastoma and other tumors such as melanoma and breast cancer (37–40). High-level expression of EPHB6 predicts favorable neuroblastoma outcome, and the expression of this gene inhibits growth of unfavorable neuroblastoma cells. This result suggests a valuable line of advance for additional study of the relationship between MIF and EPHB6.

BLU is abundantly expressed in normal lung tissue. However, its expression is markedly reduced in a subset of lung tumor cell lines (41). BLU is also inactivated in neuroblastoma. Methylation of the BLU promoter region (inactivation of BLU) in neuroblastoma was inversely correlated with tumor stage (41). These data suggest that BLU is one of the candidate tumor-suppressor genes.

IL-8 is a multifunctional CXC chemokine. It was identified as an angiogenesis-regulating molecule that induced angiogenesis (42, 43). The expression of IL-8 has been found in various human cancers (44). Recent studies have demonstrated that IL-8 regulates tumor cell growth and metastasis in many tumors such as melanoma and carcinoma of the breast, stomach, pancreas, and liver. Our previous results indicate that serum IL-8 from patients with hepatocellular carcinoma significantly correlates with tumor size, tumor stage, and venous invasion, which suggests that IL-8 may be involved directly or indirectly in the progression of hepatocellular carcinoma (45). We also demonstrated that MIF was able to stimulate neuroblastoma cells to express IL-8 (45). The present results revealed that IL-8 gene expression was markedly decreased in MIF-reduced cells (>13-fold), and this was confirmed further by the ELISA technique, which showed that IL-8 protein produced by MIF-reduced cells was 10 times less than that of control.

In summary, we have demonstrated that the use of antisense transfection when coupled with microarray and SLIM analysis provides a very useful system for defining the role of specific genes that are dysregulated in cancer. With this system, we are able to show that reduction of MIF results in tumor growth inhibition, and this suppression may be achieved by inhibiting angiogenic-related genes such as IL-8, oncogenes (C-met), and prometastatic genes. On the other hand, increasing the expression of tumor-suppressor genes (BLU, VSNL-1) and neuroblastoma favorite gene (EPHB6) may also contribute to tumor reduction. Therefore, these genes could potentially be key players in the development of neuroblastoma. Our data not only demonstrate that targeting MIF expression is a promising therapeutic strategy in human neuroblastoma therapy but also indicate the MIF target genes for additional study.

This research was supported by National Science Foundation Grant DMS-0354223 and University of Hong Kong Grant CRCG 10204242/35644/21700/323/01.