A Gene Expression Signature for Relapse of Primary Wilms Tumors

Wenliang Li,1,2 Patricia Kessler,3 Herman Yeger,3 Jennifer Alami,3 Anthony E. Reeve,4 Rosemary Heathcott,4 Jane Skee,5 and Bryan R.G. Williams1,2

1Department of Genetics, Case Western Reserve University; 2Department of Cancer Biology, Cleveland Clinic Foundation, Cleveland, Ohio; 3Department of Pediatric Laboratory Medicine, Hospital for Sick Children, Toronto, Canada; 4Department of Biochemistry, University of Otago, Dunedin, New Zealand; and 5Pediatric Haematology/Oncology, Starship Children’s Hospital, Auckland, New Zealand

Abstract

Anaplastic histology and metastasis are each associated with higher relapse and mortality rates in Wilms tumor patients. However, not all anaplastic tumors relapse and some non-anaplastic tumors relapse unexpectedly. To identify more accurate early prognostic indicators, we analyzed expression of 4,900 cancer-related genes in 26 primary Wilms tumors. This analysis revealed that expression of a set of four genes predicts future relapse of primary Wilms tumors with high accuracy, independent of anaplasia. Random permutation testing of this prognostic gene expression signature yielded P = 0.003. Real-time reverse transcription-PCR analysis of the four genes in an independent primary tumor set resulted in correct prediction of future relapse with an accuracy of 92%. One of the four genes in the prognostic signature, CCAAT/enhancer binding protein β (C/EBPB), is expressed at higher levels in both primary relapsing tumors and metastatic tumors than in primary nonrelapsing tumors. Short interfering RNA–mediated down-regulation of C/EBPB expression in Wt149, a cell line derived from a metastatic Wilms tumor, resulted in spontaneous apoptosis. These findings suggest that C/EBPB is a critical survival factor for Wilms tumor cells and that its expression contributes to the prognosis of Wilms tumor patients. (Cancer Res 2005; 65(7): 2592-601)

Introduction

Wilms tumor (nephroblastoma) is the most common malignant abdominal tumor in children (1). Whereas 80% to 85% of Wilms tumor patients survive the disease, there remain 15% to 20% that experience relapse and low survival rates despite aggressive retreatment (2). On the other hand, those patients cured of their disease frequently experience adverse side effects from therapy (3). Therefore, it is important to identify patients at the time of diagnosis who might benefit from intensified initial treatment to prevent future relapse as well as those who might be cured with limited treatment and reduced toxicity. Anaplasia (unfavorable histology) and metastasis (stage) are each associated with higher relapse and mortality rates in Wilms tumor patients (4, 5). Nevertheless, not all anaplastic tumors relapse and some primary tumors with favorable histology relapse unexpectedly (6, 7). In addition, cytogenetic studies have suggested the existence of a genetically distinct subgroup of Wilms tumors that is resistant to chemotherapy, regardless of histology (8). Because neither histology nor stage accurately predicts the course of many Wilms tumors, better prognostic markers are required.

Potential prognostic markers for Wilms tumors have been indicated by studies demonstrating association of loss of heterozygosity of 16q, 11q, 1p, and possibly 22q with adverse outcome (9–12). In addition, for patients with favorable histology tumors, decreased relapse-free survival times and overall survival rates have been linked to high telomerase activity, gain of chromosome 1q or overexpression of genes on 1q (13–15). Notably, strong p53 immunostaining, high levels of full-length TrkB mRNA or low levels of functionally inactive truncated TrkB mRNA have been suggested as negative prognostic factors independent of histology (7, 16).

In this study, we analyzed the expression of 4,900 cancer-related genes in 26 primary Wilms tumors and identified a gene expression pattern that is correlated with future relapse and independent of anaplastic histology. Moreover, one of the prognostic genes, CCAAT/enhancer binding protein β (C/EBPB), was found to be overexpressed in both primary relapsing and metastatic tumors compared with primary nonrelapsing tumors and to control cell survival in a Wilms tumor cell line model.

Materials and Methods

Cancer cDNA array construction. The cancer cDNA array used in this study contained probes for 9,240 cDNA clones corresponding to 4,900 unigenes or unclustered expressed sequence tags (EST) relevant to cancer or kidney development. These include (a) 1,800 genes implicated in metastasis or cancer development in general from the literature (17, 18) and Affymetrix cancer G110 array (http://www.affymetrix.com/analysis/index.affx); (b) 1,809 genes involved in kidney development (http://golgi.ana.ed.ac.uk/kidhome.html) or showing >3-fold variation of expression among four primary Wilms tumors and fetal kidney in our previous Affymetrix GeneChip experiments2 (c) 1,400 potential downstream targets of p53 transcriptional regulation; (d) 950 genes (19) with AU-rich elements in their 3’ untranslated region; (e) 1,245 cytokine responsive genes, genes for zinc finger proteins, or genes implicated in apoptosis. Construction of the array was similar to that described previously (19).

Sample collection, array experiment, and data processing. Fifty-four Wilms tumor specimens were obtained from the Cleveland Clinic Foundation (Cleveland, OH), the Hospital for Sick Children (Toronto, Canada), and the National Wilms Tumor Study Group (NWTS) under Institutional Review Board approval. All tumors are prechemotherapy. Patients were subjected to surgery first and treated similarly under the NWTS protocol. Among these tumors, 26 were primary tumors whose future relapse information was available, 12 from patients who went on to

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org).

Requests for reprints: Bryan R.G. Williams, Department of Cancer Biology, Lerner Research Institute, Cleveland Clinic Foundation, 9500 Euclid Avenue, Cleveland, OH 44195. Phone: 216-445-9652; Fax: 216-444-3164; E-mail: willbr@ccf.org.

© 2005 American Association for Cancer Research.

Cancer Res 2005; 65: (7). April 1, 2005 2592 www.aacrjournals.org
relapse within 4 years (termed "relapsing"), 14 from patients who remained relapse-free after their initial diagnosis for a period of at least 4 years (termed "nonrelapsing"). Pooled RNA from 13- to 24-week-old human fetal kidneys (Stratagene, La Jolla, CA) was used as a common reference in the cDNA array experiments. GeneSpring software (Silicon Genetics, Redwood City, CA) was used for data analyses. Intensities below 100 were raised to 100 and nonlinear normalization (LOWESS) was done for each array to determine the abundance of each transcript in a tumor sample as a fold change with respect to its abundance in the human fetal kidney sample (common reference).

**Relapse prediction.** From the 26 primary tumors evaluable with respect to relapse, 18 tumor samples were randomly selected to form the training set, and the remaining eight tumors formed the testing set. Upon filtering out rarely expressed features (cDNAs) in the training set, signal-to-noise ratios, $S_2N = \frac{\text{Mean}_{\text{relapsing}} - \text{Mean}_{\text{nonrelapsing}}}{\text{SD}_{\text{relapsing}} + \text{SD}_{\text{nonrelapsing}}}$, were calculated for remaining features (Supplementary Information). Welch $t$ test was also used to evaluate the significance for each feature to be associated with relapse. These features were then ranked on the basis of their $S_2N$. The top 1 feature (cDNA clone) was first selected and evaluated for its overall power to correctly classify the training samples using the “leave-one-out” cross-validation with the k-nearest neighbor ($k$-NN) algorithm (21). This performance evaluation procedure was repeated when one or two features were added each time from the top of the candidate list to optimize the number of genes in the molecular signature. An optimal set of genes obtained from the training samples through this cross-validation was evaluated by hierarchical clustering and then used to predict the testing samples.

To test the validity of the molecular signature, two other classification methods (discriminant analysis and classification tree model (5 plus 6.0, http://www.insightful.com) were applied. To further assess the statistical significance of the molecular signature, 1,500 permutation of labels of the 26 samples were done, and for each permuted data set, a similar number of cDNA clones were used (20) were calculated for remaining features (Supplementary Information). Welch $t$ test was also used to evaluate the significance for each feature to be associated with relapse. These features were then ranked on the basis of their $S_2N$. The top 1 feature (cDNA clone) was first selected and evaluated for its overall power to correctly classify the training samples using the “leave-one-out” cross-validation with the k-nearest neighbor ($k$-NN) algorithm (21). This performance evaluation procedure was repeated when one or two features were added each time from the top of the candidate list to optimize the number of genes in the molecular signature. An optimal set of genes obtained from the training samples through this cross-validation was evaluated by hierarchical clustering and then used to predict the testing samples.

**Real-time reverse transcription-PCR of CCL2 and vascular endothelial growth factor.** cDNA reverse transcribed from 2.5 μg of total RNA using random primers was used as the template for amplification of CCL2 and vascular endothelial growth factor (VEGF) transcript sequences using SYBR Green Master Mix (Perkin-Elmer, Wellesley, MA) or ABI 7700 system. Primers for CCL2 forward primer 5′-TGGCTTCCGCAAGTAAAGTG-3′ and reverse primer 5′-ATGAAGTGTGCTCTGATGAC-3′ and VEGF (forward primer 5′-TACCTCCACCCTGCGAAGT-3′ and reverse primer 5′-TGGATGATGACCATTGCGTC-3′) were shown to give a single band of expected size on an agarose gel. Amplification of 18s RNA using Taqman Master Mix (Perkin-Elmer) was used as a normalization control. $\Delta CT$ values were calculated these genes in each primary relapsing and nonrelapsing tumor sample. The Welch $t$ test was used to assess the significance of $\Delta CT$ difference between primary relapsing and nonrelapsing tumors.

**Immunohistochemical staining.** Immunohistochemical staining was done on an independent set of samples using an indirect immunoperoxidase labeling method on paraffin sections as described previously (22). Rabbit anti-VEGF (A-20, 1:100) and mouse monoclonal anti-C/EBP β (H7, 1:100) from Santa Cruz were used. H&E staining was used to verify histopathological details and identify the most representative regions of each tumor.

**Classification of tumors based on real-time reverse transcription-PCR of the four signature genes.** Total RNA (1.5 μg) was reverse transcribed into cDNA using oligo-DT primer. The PCR primer sequences are as follows: C/EBP β (forward primer 5′-GCCAAGAAGACGCTGACA-3′ and reverse primer 5′-GGCAGTTGGTGTTTGGCCT-3′), cDNA C542255 (forward primer 5′-TATAGTGGGATGACACAGTGT-3′ and reverse primer 5′-GAGACATGTGCTAAGAGGTCTAC-3′), p21 (forward primer 5′-TGAGAGCTCTAAGGCGTCA-3′ and reverse primer 5′-GGCGCTTTGGAGTGTGAAA-3′), H4FG (forward primer 5′-ACAAACCAGCTTGGC-3′ and reverse primer 5′-AAGCGG- GAATGGCTTGGG-3′), and glyceraldehyde-3-phosphate dehydrogenase (GAPDH forward primer 5′-TGGCTCAGCTGAGCAC-3′ and reverse primer 5′-CAAGGCTCAAGTGAGGAG-3′). Before real-time PCR analysis, each primer set was tested to ensure a single PCR product of expected size. Real-time PCR was done using SYBR Green Master mix on an ABI 7700 system. GAPDH expression was used as a normalization control. $\Delta CT$ values = $CT_{\text{gene}} - CT_{\text{GAPDH}}$ were calculated these genes in each primary relapsing and nonrelapsing tumor sample. Discriminant analysis was used to validate this classification and assess the misclassification error rates (5 plus 6.0).

**Generation of short interfering RNAs.** Three short interfering RNA (siRNA) sequences (21 nt) were selected from three positions of C/EBPβ cDNA sequence and checked to ensure the lack of repetitive elements or homologies to other genes. The three siRNA target sequences and their starting positions on NM_005194 were siCB-1, 1AATCCATGGAAGTGGCC-CAAC, 259 (5′); siCB-2, AAGTTGGCACAATTTCTAACCAG, 269 (5′); and siCB-3, AACAAACCGACATCGAGAT, 1417 (3′ untranslated region). The siRNAs were generated using the siRNA construction kit from Ambion according to the manufacturer’s instructions. Control siRNAs against GAPGH (siGAPGH) and SUMO-1 (siSUMO-1) were also generated. The siRNA target sequence for SUMO-1 (AAGTTGAATATATTAAACTCA) was from Akio Kishi et al. (23). The siRNA templates for GAPDH were provided in the construction kit.

**Short interfering RNA transfection and immunoblotting.** siCB-1, siCB-2, siCB-3, siSUMO-1, siGAPDH were transfected into WiT49 cells (30-40% confluent) at a final concentration of 25 nmol/L in Opti-MEM medium using oligofectAMINE as transfection reagent according to the manufacturer’s protocol (Invitrogen, Carlsbad, CA). Mock transfection (oligofectAMINE reagent only) was done as a negative control. Forty-eight hours post-transfection, detached and adherent WiT49 cells were lysed as described (24). Extracts containing equal amounts of protein as determined by Bradford assay (Bio-Rad, Hercules, CA) were resolved by SDS-PAGE (10% gel) and subjected to immunoblotting. Antibodies used were H-7 (1:500, Santa Cruz Biotechnology, Santa Cruz, CA) for human C/EBPβ, 9542 (1:1000, Cell Signaling, Beverly, MA) for poly(ADP-ribose) polymerase (PARP), p56 (1:2,000, provided by Dr. Ganes Sen; ref. 25), 18-2306 (1:1,000, Zymed, South San Francisco, CA) for SUMO-1, 14AA (1:1,000, Santa Cruz Biotechnology) for EBP, and A-5441 (1:5,000, Sigma, St. Louis, MO) for β-actin.

### Results and Discussion

**Prediction of Wilms tumor relapse based upon gene expression signature.** We generated gene expression profiles for 26 primary Wilms tumors for which we had information regarding future relapse. Of these, 12 tumors were from patients who went on to relapse within 4 years (termed “relapsing”), seven of which had favorable histology. The remaining 14 were from patients who remained relapse-free after their initial diagnosis for a period of at least 4 years (termed “nonrelapsing”), all having favorable histology. Relapse of Wilms tumor typically occurs 2 to 3 years after original diagnosis. We did not observe a correlation between relapse and stage of these primary tumors (Fisher’s exact test $P = 0.93$, Supplementary Table 1). To identify genes whose expression in the primary tumor is associated with its tendency towards future relapse, we randomly selected 18 tumors from the 26 primary tumors to form a training set. The remaining eight tumors constituted the test set. In the training set, we used both the signal-to-noise (S2N) ratio matrix $[S2N=\frac{\text{Mean}_{\text{relapsing}} - \text{Mean}_{\text{nonrelapsing}}}{\text{SD}_{\text{relapsing}} + \text{SD}_{\text{nonrelapsing}}}]$; ref. 20] and the Welch $t$ test to search for signature genes associated with relapse. Among the 7,462 clones that remained in the data set after removal of rarely expressed clones (Supplementary Information), 199 clones (182 unique genes or ESTs) were
found to be strongly associated with relapse (absolute S2N ratio (|S2N|) > 0.45 and \( P < 0.04 \)). Expression of multiple clones for the same genes/clusters correlated nicely (Supplementary Table 2). The correlation coefficient is \( R^2 = 0.985 \). The 199 clones were ordered by decreasing |S2N| ratio. We then optimized the molecular signature by sequentially adding one to two genes from the rank-ordered list and evaluating their overall power for correct classification in the leave-one-out cross-validation with the \( k \)-nearest neighbor (\( k \)-NN) algorithm. We stopped this process when an optimal gene set consisting of a minimal number of genes yet having maximal accuracy in cross-validation of the training set was reached. The top four genes in the rank-ordered list (\( C/EBPB \), cDNA CF542255, \( p21 \), and \( H4FG \)) were capable of distinguishing between the primary relapsing and nonrelapsing tumors in the training set with 100% accuracy. Classification of the 18 training samples by the four-gene signature was confirmed by unsupervised hierarchical clustering analysis, where the primary relapsing tumors were clearly separated from primary nonrelapsing tumors (Fig. 1A). The trained model, based on the expression patterns of \( C/EBPB \), cDNA CF542255, \( p21 \), and \( H4FG \) in the 18 training samples, was then used to predict the relapse status of the eight tumors in the test set using the \( k \)-NN algorithm. All eight samples were correctly classified as relapsing or nonrelapsing and this again was confirmed by unsupervised hierarchical clustering (Fig. 1B). We tested the performance of the four-gene signature in the 18-sample training set and eight-sample test set with two other classification methods, discriminant analysis and classification tree model, both of which yielded similarly high accuracy (data not shown). To ensure that the predictive power of the four-gene signature was not biased by the particular composition of the training and testing sets, we randomly formed several other training sets (range, 13-19 samples) from the 26 samples with different relapsing/nonrelapsing ratios and used the expression patterns of the four genes in the training sets to predict the corresponding testing sets. The average accuracies in cross-validation of the training sets and prediction of the testing sets were 98.2% and 94.1%, respectively.

To further assess the statistical significance of the four-gene prognostic model, we did 1,500 permutations of labels of the 26 samples, and for each permuted data set, similar numbers of genes with the greatest |S2N| values were used in discriminant analysis.

---

**Figure 1.** Hierarchical clustering of the 18 training samples (A) and eight testing samples (B) using the four genes. Sample (column) and gene (row). Color scale, expression level for a gene in a sample relative to normal fetal kidney. Sample trees (x-axis) and gene trees (y-axis). Bottom, sample types: relapsing (blue) or nonrelapsing (red). Signal-to-noise ratio S2N = (Mean relapsing − Mean nonrelapsing) / (SD relapsing + SD nonrelapsing).
Only 5 of the 1,500 permutations yielded models whose cross-validation error rates or Hotelling's t test F statistics matched or exceeded those of the four-gene model. Thus, the likelihood of matching the success of the observed four-gene model by chance was estimated at \( P = 0.003 \). The four-gene signature that we have identified seems to clearly distinguish relapsing tumors from nonrelapsing tumors and show predictive power.

In general, C/EBPB and H4FG (HIST1H4C) were up-regulated in relapsing tumors compared with nonrelapsing tumors, whereas p21 and cDNA CF542255 were down-regulated. It is important to note that it is the expression pattern of all four genes taken together that provides the predictive power described above. Expression of any single gene is not absolutely correlated with prognosis. The up-regulation of C/EBPB in tumors that are destined to relapse is intriguing given that C/EBPB is a critical component of a Ras-dependent transformation pathway (26). Moreover, transcriptional activity of C/EBPB leads to induction of a number of prometastatic genes, including hepatocyte growth factor HGF, IL-6, and IL-8, which in turn induce C/EBPB (27, 28). As further support for its potential role in tumor progression in general, increased expression of C/EBPB has been observed in solid cancers other than Wilms tumor, such as breast, ovarian, and colorectal cancer (29). In tumors that go on to relapse, the activities of C/EBPB are combined with an elevated proliferative rate as indicated by increased histone H4FG expression and decreased levels of the cyclin-dependent kinase inhibitor p21. The contribution of cDNA CF542255 to the progression of Wilms tumor is unclear since its protein and biological function remain to be determined.

The complete list of 199 clones (182 genes) that are differentially expressed between primary relapsing and primary nonrelapsing Wilms tumors includes many additional genes implicated in processes relevant to tumor progression such as proliferation, metastasis, angiogenesis, apoptosis, transcriptional regulation, and signal transduction. Examples of these genes are those expressed higher in relapsing tumors: VEGF, STAT1, IL-8, integrin β4 binding protein, coagulation factor VII (F7), collagen type VIII α2, retinoblastoma binding protein 8, hepatocyte growth factor, early growth response 2, neumedin B, Ras homologue enriched in brain (RHEB2), C-C motif chemokine receptor 1 (CCR1), and C-C motif chemokine ligand 2 (CCL2); and those expressed lower in relapsing tumors: IGFB2, ZNF23, ZNF33a, caspase 6 (CASP6), mothers against decapentaplegic homologue SMAD3, E2F transcription factor 4 (E2F4), transforming acidic coiled coil containing protein 1, active BCR-related gene, 90-kDa heat shock protein, BCL2-associated X protein (BAX), damage-specific DNA binding protein 2, and suppressor of cytokine signaling 5 (see Supplementary Table 2 for a complete list). Together with our expression data, previously published functional information for a number of these genes suggests potential mechanisms for Wilms tumor progression. For example, our data show that induction of a number of proangiogenic/metastatic factors is associated with Wilms tumor relapse. If these factors include CCL2, CCR1, IL-8, VEGF, F7, HGF and of course, C/EBPB. Interestingly, we recently identified significant coexpression of HGF and its receptor, c-Met, in Wilms tumors (22). Furthermore, abnormal expression of adhesion proteins regulated by HGF/c-Met signaling was found to correlate with Wilms tumor invasiveness and metastasis (30). It is also likely that some of the relapse-associated gene products cooperate to promote angiogenesis/metastasis. C/EBPB and HGF, VEGF, and HGF induce each other, forming positive feedback loops. In addition, both F7 and HGF positively regulate the expression of VEGF and IL-8 (31, 32). Whereas proangiogenic/metastasis genes are up-regulated in relapsing Wilms tumors, a number of genes that encode proapoptotic or growth suppressive factors (such as CASP6, BAX, IGFB2, E2F4, and p21) are down-regulated. Studies examining gene expression in a number of other types of cancer have also identified many of the genes on our Wilms tumor relapse-associated list (29, 33–35). Thus, our data add support to the idea that a common molecular mechanism underlies metastasis of many different human solid cancers (36).

Cytoplasmic immunopositivity for p53, which indicates a deficiency in its activities, has been suggested as a poor prognostic marker for Wilms tumors, independent of histology (7). It is possible that the gene expression differences between nonrelapsing and relapsing tumors are linked to inactivation of p53 pathway. This is supported by the down-regulation of p21 we detected in relapsing tumors. There are little overlap between the 199 clones we defined and a partial list of 80 cDNA clones differentially expressed between nine nonrelapsing tumors with favorable histology and four relapsing tumors with favorable histology described in a recent report (37). This is probably due to incompleteness of their gene list, different cDNA array platforms and their focus on tumors with favorable histology. Although insights into the aggressiveness of favorable histology tumors may be provided in their study, restricting gene expression profiling to favorable histology tumors can only identify prognostic markers that are supplementary to histology.

Confirmation of microarray gene expression patterns. The protein expression pattern for C/EBPB was examined by immunohistochemical staining on an independent set of samples. Four cases of primary relapsing, three cases of metastatic and four cases of primary nonrelapsing Wilms tumors were examined. C/EBPB protein staining was detected in primary relapsing (2 of 4) and metastatic tumors (2 of 3) but not in primary nonrelapsing tumors (0 of 4; Fig. 2A). Positive C/EBPB staining was nuclear, concurring with its function as a transcription factor. The expression levels of VEGF, another gene expressed higher in primary relapsing tumor than in nonrelapsing tumors, was also analyzed. Its staining was much stronger in primary relapsing tumors (4 of 4) and marginally stronger in metastatic cases (2 of 3) than in nonrelapsing cases (1 of 4). Positive VEGF staining is cytoplasmic and staining of blood vessels serves as an internal positive control. Thus, even in this relatively small sample set, the protein level differences of C/EBPB and VEGF in primary nonrelapsing, primary relapsing, and metastatic Wilms tumors are notable.

We also arbitrarily selected two genes down the list (in addition to the top four genes, see next) to validate the cDNA array expression patterns using real-time reverse transcription-PCR (RT-PCR). Similar to VEGF, CCL2 is also expressed at a higher levels in primary relapsing tumors than in primary nonrelapsing tumors based on the cDNA array data. When analyzed by real-time PCR, expression of CCL2 and VEGF were significantly higher in primary relapsing tumors \( (n = 12) \) than in primary nonrelapsing tumors \( (n = 11) \); Welch's t test \( P = 0.0067 \) and 4.26-fold for CCL2, and \( P = 0.021, 2.95\text{-fold for VEGF}; Fig. 2B and C \). Both the immunohistochemical staining and real-time RT-PCR results were in agreement with the cDNA array data.

Application of the four-gene prognostic signature to an independent sample set using real-time RT-PCR. To determine whether the four-gene prognostic model might be of practical use in stratification of patients, we designed real-time RT-PCR assays to
quantify expression of the four genes. Gene expression as determined by these assays was used to classify an independent set of 12 primary Wilms tumors (6 relapsing and 6 nonrelapsing). Eleven of the 12 samples (92%) were correctly classified with respect to future relapse using the RT-PCR results in either discriminant analysis or tree model classification. When the RT-PCR results were visualized in a three-dimensional scatter plot using the three most significant genes from cDNA array data (C/EBPB, cDNA CF 542255, and p21), six primary nonrelapsing tumors were separated from five primary relapsing tumors, with one primary relapsing tumor behaving as outlier (Fig. 3). Importantly, among these 12 tumors, only one was known to have unfavorable histology. Thus, this sample set provides further support of the improved accuracy of the four-gene signature as a prognosticator compared with histology. These results not only confirm our cDNA array data on an independent sample set using a different approach but also indicate that the four-gene prognostic model may be applied in practice using a simpler technique like RT-PCR.

**Identification of genes associated with Wilms tumor metastasis.** In addition to identifying genes associated with Wilms tumor relapse, we also mined our array data for genes that are differentially expressed in both metastatic tumors and primary tumors that display future metastasis compared with primary tumors that do not metastasize. Among the 12 primary relapsing
tumors in our original sample set, one later relapsed only in the original tumor bed, whereas 11 later relapsed outside of the primary site, mostly in the lung or abdomen. Relapse of primary Wilms tumors typically occurs in the lung or abdomen (38). Because several studies (including this work) have shown that the metastatic potential of human cancers is embodied in the gene expression profiles of the primary tumors (39–41), it is presumed that some relevant gene expression changes necessary to initiate the metastatic process occur in primary metastasizing tumors. Some of these gene expression changes might be preserved in the final metastasis, which would suggest roles in maintenance of the phenotype as well. Thus, comparison of both primary metastasizing tumors and metastatic tumors to primary nonmetastasizing tumors may identify gene expression changes involved in both initiation and maintenance of the metastatic phenotype. Accordingly, we identified 95 clones that were persistently up-regulated or down-regulated in both the 11 primary metastasizing tumors and six metastatic (relapsed) tumors compared with the 14 primary nonmetastasizing (nonrelapsing) tumors (|S2N| > 0.4 and Welch t test two-sided P < 0.05).

A subset of the 95 clones are shown in Table 1, including clones whose expression changes at least 50% with a Welch t test two-sided P < 0.05 in both the metastasizing/nonmetastasizing and metastatic/nonmetastasizing comparisons (see Supplementary Table 3 for a complete list). We found that C/EBPB, one of our relapse signature genes, is also among the genes up-regulated in both primary metastasizing tumors and metastatic tumors compared with primary nonmetastasizing tumors. This finding is consistent with other studies implicating C/EBPB in tumor metastasis as described above. Increased expression of C/EBPB in metastatic Wilms tumors was confirmed at the protein level by immunohistochemical staining (Fig. 2A).

To discover enriched biological themes within the list of 95 clones associated with Wilms tumor metastasis, we used the EASE program (http://david.niaid.nih.gov/david/ease.htm; 42) and noticed that five general categories were significantly overrepresented: calcineurin signaling, cell cycle regulation, protein/nucleus import, metabolism, and apoptosis regulation (Supplementary Table 4). It was noted that biological themes relating to calcineurin signaling ranked highest among the descriptive categories. Three calcineurin isoforms, including calcineurin Aα, Aβ, and B type I, were all up-regulated in both primary metastasizing and metastatic tumors compared with primary nonmetastasizing tumors (Table 1; Supplementary Table 3). Calcineurin activates NFAT (nuclear factor of activated T cells), which mediates VEGF signaling and results in transcriptional activation of proangiogenic genes such as tissue factor and cyclooxygenase 2 (43, 44). This result suggests that calcineurin-NFAT signaling plays an important roles in Wilms tumor metastasis.

C/EBPB is required for survival of Wilms tumor cells. C/EBPB is an essential member of the molecular signature for relapse of primary Wilms tumor that we identified through microarray expression analysis. Three distinct polypeptides are produced from the C/EBPB gene through use of different translation initiation sites. The LAP-1 (46 kDa) and LAP-2 (42 kDa) isoforms differ at the NH2 terminus, whereas translation from a third downstream in-frame AUG results in a 20-kDa LIP isoform (ref. 45; Fig. 4A). Previous studies suggest that C/EBPB may be involved in tumor development, with elevated levels of C/EBPB being found in human breast (reviewed in ref. 45), ovarian (46), colorectal (29), renal cell (47), and other cancers. Furthermore, in a mouse model of carcinogen-induced skin tumorigenesis involving oncogenic Ras signaling, C/EBPB-null mice were resistant to tumor development (26). Here the role of C/EBPB in Wilms tumor cell growth and survival was investigated by siRNA-mediated suppression of C/EBPB expression in Wit-49, a cell line derived from a metastatic Wilms tumor (48). Three different siRNA target sequences specific for the C/EBPB transcript were designed (Fig. 4A). As controls, siRNAs against GAPDH and SUMO-1 were also generated.

As measured by immunoblotting, all three isoforms of C/EBPB protein were significantly down-regulated following transfection of Wit-49 cells with C/EBPB siRNAs, with siCB-2 being the most...
potent (Fig. 4B). Notably, levels of the closely related protein CCAAT/enhancer binding protein α (C/EBPA) were not reduced by siRNAs directed against C/EBPB. Rather, C/EBPA protein levels increased as C/EBPB protein expression was down-regulated by C/EBPB-specific siRNAs. This observation is in agreement with some other circumstances where C/EBPA expression decreases when C/EBPB expression increases or vice versa (49–51). Transfection of siRNA directed against the SUMO-1 transcript resulted in reduced levels of SUMO-1 protein, but not C/EBPB protein (Fig. 5). These observations show the specificity of gene targeting by the different siRNAs.

It has been reported by our lab and others that siRNAs can activate double-stranded RNA (dsRNA) responsive pathways in cells (52–54). Indeed, as shown in Fig. 5, transfection of the different

### Table 1. Genes persistently up-regulated or down-regulated in both primary metastasizing and metastatic tumors compared with primary nonmetastasizing tumors

<table>
<thead>
<tr>
<th>Change* (metastasizing versus non metastasizing), %</th>
<th>Change* (metastatic versus nonmetastasizing), %</th>
<th>Change* (metastatic versus nonmetastasizing), %</th>
<th>Name</th>
<th>Symbol</th>
<th>Genbank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>78.1</td>
<td>0.0002</td>
<td>99.6</td>
<td>0.008</td>
<td>CCAAT/enhancer binding protein (C/EBP), beta</td>
<td>C/EBPB</td>
</tr>
<tr>
<td>85.7</td>
<td>0.039</td>
<td>109.1</td>
<td>0.017</td>
<td>Cyclin A1</td>
<td>CCNA1</td>
</tr>
<tr>
<td>117.1</td>
<td>0.001</td>
<td>208.1</td>
<td>0.0001</td>
<td>Early growth response 2</td>
<td>EGR2</td>
</tr>
<tr>
<td>108.9</td>
<td>0.001</td>
<td>104.0</td>
<td>0.018</td>
<td>EST</td>
<td>R67218</td>
</tr>
<tr>
<td>101.1</td>
<td>0.003</td>
<td>118.6</td>
<td>0.004</td>
<td>EST</td>
<td>A688512</td>
</tr>
<tr>
<td>99.2</td>
<td>0.007</td>
<td>106.9</td>
<td>0.0005</td>
<td>EST</td>
<td>AA460254</td>
</tr>
<tr>
<td>75.8</td>
<td>0.002</td>
<td>71.7</td>
<td>0.003</td>
<td>EST</td>
<td>H92571</td>
</tr>
<tr>
<td>72.0</td>
<td>0.005</td>
<td>110.4</td>
<td>0.021</td>
<td>EST</td>
<td>R01118</td>
</tr>
<tr>
<td>71.2</td>
<td>0.011</td>
<td>53.9</td>
<td>0.009</td>
<td>EST</td>
<td>AA873507</td>
</tr>
<tr>
<td>66.3</td>
<td>0.031</td>
<td>89.4</td>
<td>0.014</td>
<td>EST</td>
<td>AA448396</td>
</tr>
<tr>
<td>57.4</td>
<td>0.022</td>
<td>50.8</td>
<td>0.007</td>
<td>EST</td>
<td>N32201</td>
</tr>
<tr>
<td>53.2</td>
<td>0.013</td>
<td>54.8</td>
<td>0.015</td>
<td>EST</td>
<td>R11398</td>
</tr>
<tr>
<td>74.0</td>
<td>0.031</td>
<td>59.9</td>
<td>0.008</td>
<td>Histone 1, H1c</td>
<td>HIST1H1C</td>
</tr>
<tr>
<td>86.2</td>
<td>0.048</td>
<td>101.7</td>
<td>0.043</td>
<td>Hypothetical protein DKFZp434D2328</td>
<td>LOC91526</td>
</tr>
<tr>
<td>72.7</td>
<td>0.025</td>
<td>74.1</td>
<td>0.014</td>
<td>IFN induced transmembrane protein 1</td>
<td>IFITM1</td>
</tr>
<tr>
<td>64.0</td>
<td>0.010</td>
<td>133.3</td>
<td>0.029</td>
<td>Interleukin 1, beta</td>
<td>IL1B</td>
</tr>
<tr>
<td>68.6</td>
<td>0.012</td>
<td>94.2</td>
<td>0.041</td>
<td>Matrix metalloproteinase 12</td>
<td>MMP12</td>
</tr>
<tr>
<td>66.0</td>
<td>0.001</td>
<td>63.3</td>
<td>0.022</td>
<td>Neuromedin B</td>
<td>NMB</td>
</tr>
<tr>
<td>69.4</td>
<td>0.017</td>
<td>123.0</td>
<td>0.002</td>
<td>Receptor-interacting serine-threonine kinase 2</td>
<td>RIPK2</td>
</tr>
<tr>
<td>112.1</td>
<td>0.005</td>
<td>101.6</td>
<td>0.008</td>
<td>Serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 3</td>
<td>SERPINB3</td>
</tr>
<tr>
<td>50.6</td>
<td>0.025</td>
<td>58.4</td>
<td>0.024</td>
<td>Sorcin</td>
<td>SRI</td>
</tr>
<tr>
<td>113.4</td>
<td>0.001</td>
<td>104.7</td>
<td>0.004</td>
<td>Transketolase-like 1</td>
<td>TKTL1</td>
</tr>
<tr>
<td>55.9</td>
<td>0.011</td>
<td>50.2</td>
<td>0.021</td>
<td>Tumor necrosis factor, α-induced protein 3</td>
<td>TNFAIP3</td>
</tr>
<tr>
<td>−59.9</td>
<td>0.030</td>
<td>−53.0</td>
<td>0.008</td>
<td>Adenylate cyclase 2 (brain)</td>
<td>ADCY2</td>
</tr>
<tr>
<td>−55.0</td>
<td>0.026</td>
<td>−54.6</td>
<td>0.005</td>
<td>Insulin-like growth factor 2</td>
<td>IGF2</td>
</tr>
<tr>
<td>47.5</td>
<td>0.042</td>
<td>85.0</td>
<td>0.0048</td>
<td>Calcinurin B, type 1</td>
<td>PPP3R1</td>
</tr>
<tr>
<td>40.2</td>
<td>0.012</td>
<td>47.5</td>
<td>0.0089</td>
<td>Calcinurin A α</td>
<td>PPP3CA</td>
</tr>
<tr>
<td>37.8</td>
<td>0.021</td>
<td>49.7</td>
<td>0.0037</td>
<td>Calcinurin A β</td>
<td>PPP3CB</td>
</tr>
</tbody>
</table>

*Fold change = (primary metastasizing tumors / primary nonmetastasizing tumors − 1) × 100%.

† Two-sided Welch t test P in primary metastasizing versus primary nonmetastasizing tumor comparison.

‡ Fold change = (metastatic tumors / primary nonmetastasizing tumors − 1) × 100%.

§ Two-sided Welch t test P in metastatic versus primary nonmetastasizing tumor comparison.
siRNAs targeting C/EBPB, GAPDH, and SUMO-1 led to induction of the classic dsRNA-responsive gene product p56 (25). Moreover, the GAPDH and SUMO-1 siRNAs induced C/EBPB transcript and protein expression (Figs. 4 and 5), suggesting that C/EBPB is also a dsRNA-responsive gene. Importantly, an increase in spontaneous apoptosis was observed in C/EBPB siRNA-treated cells as shown by proteolytic cleavage of PARP. The amount of PARP cleavage correlated with the extent of down-regulation of C/EBPB proteins.

Figure 4. A, schematics of C/EBPB cDNA sequence, three protein isoforms, and siRNA target positions. B, effects of siRNAs on the protein expression of C/EBPB and the three protein isoforms of C/EBPB. Immunoblotting was performed 48 hours after siRNA transfection.

Figure 5. Down-regulation of C/EBPB protein by siRNAs correlates with an increase in PARP cleavage in Wit49 cells. Immunoblotting was done with antibodies against human C/EBPB (H-7, Santa Cruz Biotechnology), PARP (9542, Cell Signaling), p56 (25), SUMO-1 (18-2306, Zymed), and β-actin (A-5441, Sigma).
References


A Gene Expression Signature for Relapse of Primary Wilms Tumors

Wenliang Li, Patricia Kessler, Herman Yeger, et al.


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/65/7/2592

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2005/04/06/65.7.2592.DC1

Cited articles
This article cites 53 articles, 21 of which you can access for free at:
http://cancerres.aacrjournals.org/content/65/7/2592.full#ref-list-1

Citing articles
This article has been cited by 9 HighWire-hosted articles. Access the articles at:
http://cancerres.aacrjournals.org/content/65/7/2592.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link
http://cancerres.aacrjournals.org/content/65/7/2592.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.