Prognostic Classification of Relapsing Favorable Histology Wilms Tumor Using cDNA Microarray Expression Profiling and Support Vector Machines

Richard D. Williams,[1]* Sandra N. Hing,[1] Braden T. Greer,[2] Craig C. Whiteford,[2] Jun S. Wei,[2] Rachael Natrajan,[1] Anna Kelsey,[3] Simon Rogers,[4] Colin Campbell,[4] Kathy Pritchard-Jones,[1] and Javed Khan[2]

[1]Section of Paediatric Oncology, Institute of Cancer Research, Sutton, Surrey, United Kingdom
[2]Oncogenomics Section, Pediatric Oncology Branch, National Cancer Institute Advanced Technology Center, Gaithersburg, Maryland
[3]Department of Paediatric Pathology, Royal Manchester Children's Hospital, Manchester, United Kingdom
[4]Department of Engineering Mathematics, Bristol University, Bristol, United Kingdom

Treatment of Wilms tumor has a high success rate, with some 85% of patients achieving long-term survival. However, late effects of treatment and management of relapse remain significant clinical problems. If accurate prognostic methods were available, effective risk-adapted therapies could be tailored to individual patients at diagnosis. Few molecular prognostic markers for Wilms tumor are currently defined, though previous studies have linked allele loss on 1p or 16q, genomic gain of 1q, and overexpression from 1q with an increased risk of relapse. To identify specific patterns of gene expression that are predictive of relapse, we used high-density (30 k) cDNA microarrays to analyze RNA samples from 27 favorable histology Wilms tumors taken from primary nephrectomies at the time of initial diagnosis. Thirteen of these tumors relapsed within 2 years. Genes differentially expressed between the relapsing and nonrelapsing tumor classes were identified by statistical scoring (*t* test). These genes encode proteins with diverse molecular functions, including transcription factors, developmental regulators, apoptotic factors, and signaling molecules. Use of a support vector machine classifier, feature selection, and test evaluation using cross-validation led to identification of a generalizable expression signature, a small subset of genes whose expression potentially can be used to predict tumor outcome in new samples. Similar methods were used to identify genes that are differentially expressed between tumors with and without genomic 1q gain. This set of discriminators was highly enriched in genes on 1q, indicating close agreement between data obtained from expression profiling with data from genomic copy number analyses.

INTRODUCTION

Wilms tumor is the most common renal tumor of childhood, and treatment is generally very successful, with about 85% of affected children expected to be long-term survivors. However, nearly half of these survivors will have received treatments with the potential for serious permanent side effects, such as anthracycline-related cardiotoxicity or the sequelae of radiotherapy on growth, fertility, and risk of second cancers. The results of clinical trials in both Europe and North America have suggested that for the majority of children, tumor control can be achieved using only vincristine and actinomycin D chemotherapy, which have minimal risks of late sequelae. However, successful treatment of relapsing or nonresponding Wilms tumors remains a challenge. Therefore, rational selection of patients at diagnosis for risk-adapted therapy is essential for maintaining or improving cure rates while reducing the burden of treatment for the majority.

Established adverse prognostic factors in Wilms tumor are advanced tumor stage and histological subtype. Approximately 5% of Wilms tumors show unfavorable histology, defined as anaplasia, a morphological diagnosis based on increased abnormal mitoses that is associated with the presence of *TP53* mutations (Bardeesy et al., 1994). The majority of Wilms tumors are classed as having favorable histology and consist of varying proportions of blastemal, epithelial, and stromal cells, the three major cell types seen in this tumor. Several molecular alterations have been suggested as important predictors of adverse outcome in Wilms tumor with favorable histology. Allele loss at 1p and 16q has been implicated in a number of large studies (Dome and Coppes, 2002), and the recently closed

Supported in part by: Cancer Research UK; Royal Marsden Hospital Children's Cancer Unit Fund; David Adams Appeal Fund; donations from the Gobat family.

*Correspondence to: Richard Williams, Section of Paediatric Oncology, Institute of Cancer Research, 15 Cotswold Rd., Sutton, Surrey SM2 5NG, UK. E-mail: richardw@icr.ac.uk

Received 18 December 2003; Accepted 15 April 2004

DOI 10.1002/gcc.20060

Published online 16 June 2004 in
Wiley InterScience (www.interscience.wiley.com).

National Wilms Tumor Study Group (NWTSG) 5 trial confirmed prospectively the adverse prognostic significance of loss at these alleles (Dr. Paul Grundy, National Wilms Tumor Study Group, personal communication). However, allele loss is present in only one third of tumors that subsequently relapse. Further, the regions of common allele loss are large, and no candidate genes have yet been identified in these regions. There is, therefore, a need to identify other biological characteristics of this tumor that are associated with relapse.

Smaller studies have highlighted more specific molecular changes apparently associated with adverse tumor outcome (defined as an increased risk of relapse or a reduced chance of survival). These include gain or overexpression of genes from 1q (Hing et al., 2001; Lu et al., 2002), high expression of the telomerase reverse transcriptase gene *TERT* (Dome et al., 1999), and expression of the neurotrophic tyrosine kinase receptor *NTRK2* (*TRKB*) gene (Eggert et al., 2001). Immunological assays have linked adverse outcome with relative overexpression of TP53 (Sredni et al., 2001), PCNA (Skotnicka-Klonowicz et al., 2002), and FASN proteins (Camassei et al., 2003) and with relative underexpression of HSPA1A (HSP70) and ABCC1 (MRP1) (Efferth et al., 2001a, 2001b). Similarly, poor outcome has been found to be associated with blastemal-specific protein expression of WT1 and EGR1 (Ghanem et al., 2000), TGFA and EGFR (Ghanem et al., 2001a), BCL2 (and the BCL2/BAX ratio; Ghanem et al., 2001b), specific CD44 isoforms (Ghanem et al., 2002), and VEGF and FLT1 (Ghanem et al., 2003). However, larger-scale studies would be required to determine whether any of these observations are of genuine value in the clinical prognosis, and at present no individual genes are in general use as specific prognostic factors for Wilms tumor relapse.

Recently, microarray technology has been used to obtain prognostically significant expression profiles from a variety of tumors. Much of this research has focused on adult tumors, including breast cancer (Sorlie et al., 2001; West et al., 2001; van de Vijver et al., 2002; van 't Veer et al., 2002), prostate cancer (Dhanasekaran et al., 2001; Singh et al., 2002), lung adenocarcinoma (Beer et al., 2002), and bladder carcinoma (Dyrskjot et al., 2003). In one recent article (Pomeroy et al., 2002), clinical outcome prediction was demonstrated in medulloblastoma, a pediatric malignancy. During the preparation of this article, evidence for a clinical outcome (survival)–based classification of a limited number of Wilms tumors was reported (Takahashi et al., 2002), although the sample set contained only 3 patients who died from their tumors, whereas other recent array-based studies (Li et al., 2002; Udtha et al., 2003) did not attempt prognostic classification.

Sophisticated methods of analysis are required to process the large, complex datasets generated by microarray experiments. The analysis of array data using machine learning algorithms, including perceptron-type artificial neural networks (ANNs), has proven to be a particularly powerful approach to tumor classification (Khan et al., 2001). Another application of learning theory to array classification is the use of support vector machines (SVMs). When applied to binary classification of such data (e.g., distinguishing relapsing from nonrelapsing tumors), SVMs determine a "decision function" by finding a separating hyperplane that maximizes the distance between itself and the closest points of each class (on both sides) in an abstract "input space" (the dimensionality of which is determined by the number of array features). In this situation, learning is dependent on the number of samples rather than the (usually much larger) number of array features. Learning is therefore rapid, but classification performance is comparable to that of ANNs (Furey et al., 2000). Here we examine the gene expression profiles of 27 Wilms tumors sampled at diagnosis using high-density cDNA microarrays with more than 30,000 features, and SVM classification. We compare tumors that were presumed cured ($n = 14$) with those that had relapsed ($n = 13$), with the aim of identifying critical genes or pathways that distinguish these two clinical behaviors.

MATERIALS AND METHODS

Tumor Samples

Wilms tumor samples were obtained from the National Wilms Tumor Study Group (Seattle, WA), the United Kingdom Children's Cancer Study Group (Leicester, UK), and the Royal Manchester Children's Hospital (Manchester, UK). All samples were selected and snap-frozen by experienced pediatric pathologists at the time of primary nephrectomy, close to the time of initial diagnosis and prior to any chemotherapy. Sections from each tumor were subjected to histopathological examination, and because Wilms tumors rarely contain interdigitated normal tissue, these can be expected to be representative of the tumor as a whole. This analysis demonstrated that all tumors were of the favorable histological subtype. Four-

TABLE 1. Tumor Samples

Tumor sample ID	Class	Age at 1st diagnosis (months)	Disease-free interval (days)	Follow-up (months)	Stage	Sex	Gain of 1q
D17	D	23	167	NA	II	F	No
D19	D	13	169	NA	IV	F	No
D22	D	48	161	NA	IV	M	Yes
D24	D	51	45	NA	II	F	Yes
D29	D	7	383	NA	I	F	No
D34	D	31	180	NA	II	M	Yes
D38	D	182	197	NA	IV	F	Yes
D40	D	16	174	NA	III	M	No
D45	D	29	150	NA	I	M	Yes
D47	D	5	198	NA	I	M	Yes
D194	D	48	330	NA	II	M	No
D204	D	120	510	NA	I	F	No
D240	D	53	600	NA	III	F	Yes
N168	N	66	NA	58	II	F	No
N169	N	60	NA	79	III	M	No
N200	N	5	NA	43	II	M	No
N203	N	30	NA	119	I	F	No
N208	N	8	NA	94	I	M	No
N210	N	132	NA	44	I	F	No
N214	N	36	NA	102	II	F	No
N215	N	24	NA	42	II	M	No
N216	N	36	NA	67	III	M	No
N221	N	11	NA	93	I	F	Yes
N224	N	9	NA	86	I	M	No
N226	N	48	NA	47	II	M	No
N228	N	84	NA	54	III	F	No
N234	N	20	NA	57	III	F	No

D, relapsing; N, nonrelapsing; NA, not applicable.

teen tumors (designated as class N—nonrelapsing) were from children who did not subsequently suffer a relapse during prolonged follow-up (at least 3 years, median 5.2 years) and who were expected to be cured (>90% of relapses in Wilms tumor occur within 3 years of diagnosis). Thirteen tumors (designated as class D—relapsing) were from children who later suffered relapse ($n = 12$) or whose tumors progressed despite treatment ($n = 1$). All relapse events occurred within 2 years of the original diagnosis. The relapse status and clinical stage of these patients are summarized in Table 1.

Chromosome Arm 1q Status

Gain of chromosome arm 1q was assessed by comparative genomic hybridization (CGH) or by fluorescence in situ hybridization (FISH); a subset of the CGH data has been described previously (Hing et al., 2001). DNA was extracted from the tumor samples by SDS/proteinase K digestion and phenol:chloroform extraction. CGH was performed as described previously (Weber-Hall et al., 1996; Hing et al., 2001). For FISH, 2 yeast artificial chromosomes (YACs) were selected (955E11, 845D8) that mapped between 1q21 and 1q31 [in the previous CGH study (Hing et al., 2001), all detectable 1q gain was in this region]. One microgram of each YAC DNA and of chromosome 1 pericentromere control DNA was labeled by nick translation with either fluorescein–12-dUTP or rhodamine–12-dUTP (Amersham Biosciences, Chalfont St. Giles, UK). Then 300 ng of YAC DNA, 100 ng of pericentromere DNA, and 10 μg of Cot-1 DNA (Invitrogen, Paisley, UK) were cohybridized to touch preparations made from frozen tumor tissue for 24 hr at 37°C. After hybridization, the slides were washed and mounted in Citifluor antifade (Vector Laboratories, Peterborough, UK) with a 0.1 μg/ml 4′,6-diamidino-2-phenylindole counterstain. Images were captured using a cooled charge-coupled device camera (Photometrics, Tucson AZ). Between 50 and 80 interphase cells were captured, and the number of YAC and pericentromere signals in each cell was counted. If both YAC probes detected at least 3 copies of their targets in >50% of these cells, the sample was judged to have genomic 1q gain.

RNA Preparation

Frozen tumors were pulverized under liquid nitrogen, and RNA was extracted using Trizol (Invitrogen) according to the manufacturer's instructions. Following a second round of Trizol purification, the RNA was quantitated and its quality assessed by resolving ribosomal RNA bands on an agarose gel. As a further test of quality, cDNA was synthesized from 1 μg of total RNA using Superscript II (Invitrogen), essentially according to the manufacturer's instructions, and was used to amplify fragments from 3 constitutively expressed genes (*ETV6*—primer 1: 5′-TACATGAACCACATCATGGTC-3′, primer 2: 5′-TTTGTTCATCCAGCTCCTGGG-3′, amplicon: 377 bp; *BCR*—primer 1: 5′-ACAGAATTCCGCTGACCATCAATAAG-3′, primer 2: 5′-ATAGGATCCTTTGCAACCGGGTCTGAA-3′, amplicon: 808 bp; *B2M*—primer 1: 5′-ACCCCCACTGAAAAAGATGA-3′, primer 2: 5′-ATCTTCAAACCTCCATGATG-3′, amplicon: 120 bp). Total reference RNA was similarly extracted from actively dividing HEK-293 cells.

cDNA Microarrays

cDNA microarrays containing 30,720 features were prepared from sequence-verified Research Genetics IMAGE clones as described in the standard National Human Genome Research Institute protocol (http://research.nhgri.nih.gov/microarray/Protocols.pdf), except that the PCR products were printed on double polylysine-coated slides in 50% aqueous dimethyl sulfoxide. Further details are available at the supplementary Web site (http://www.icr.ac.uk/paedonc/Wilmsarraydata.html).

Gene Assignment and Genomic Mapping

IMAGE clone-to-gene assignments were checked and updated using a relational database containing a recent build (161) of National Center for Bioinformatics UniGene [http://www.ncbi.nih-.gov/UniGene (Schuler et al., 1996)], formatted for database input by a Perl script (http://www.icr.ac.uk/paedonc/Wilmsarraydata.html). 19,824 unique UniGene clusters were represented on the array. Array clones were mapped to their positions on a recent repeat-masked assembly of the human genome (hg16) using standalone BLAT [http://www.cse.ucsc.edu/~kent/ (Kent, 2002)]. The start position of the best-quality match of each IMAGE clone sequence to the entire genome was taken as the position of that sequence within the genome.

Sample Labeling and Hybridization

Cyanine-3- or cyanine-5-labeled cDNA was prepared from 30 μg of tumor or reference RNA by anchored oligo(dT)-primed reverse transcription in the presence of an aminoallyl-labeled nucleotide, followed by covalent coupling to the appropriate cyanine dye ester, as described in the The Institute for Genome Research microarray protocol [http://www.tigr.org/tdb/microarray/protocolsTIGR.shtml (Hegde et al., 2000)]. Array slides were prepared for hybridization by both succinic anhydride blocking (NHGRI protocol) and subsequent preincubation with bovine serum albumin (TIGR protocol). Differentially labeled tumor and HEK-293 cell reference cDNA were cohybridized to the arrays as described in the TIGR protocol, but with the addition of 6 μg of yeast tRNA to each hybridization.

Array Imaging and Normalization

All array images were acquired with an Axon 4000B scanner and GenePix Pro 3 software. Features were extracted from the raw image data and quantified using the Scanalytics MicroArray Suite DeArray package (Chen et al., 2002). The DeArray output files were then formatted by a Perl script for input into the com.braju.sma array analysis R package [http://www.maths.lth.se/help/R/com.braju.sma (Bengtsson, 2002)]. Using com.braju.sma, background fluorescence was subtracted from the foreground intensity in each channel, and intensities were adjusted by global lowess (locally weighted regression) normalization to compensate for nonlinear dye biases. Data were then normalized between slides by scaling the log 2 ratios for each so as to obtain the same spread of the absolute median deviation across all slides, and log 2 tumor/control fluorescence ratios were calculated. Prior to further analysis, the data were filtered for quality using the quality metric Q (Chen et al., 2002), calculated by the MicroArray Suite software: all features were rejected when the mean value of Q was less than 0.7 (on a scale of 0.0–1.0) across all 27 experiments.

Identification of Differentially Expressed Genes

The array data were analyzed for differential gene expression by two complementary methods.

First, genes that were differentially expressed between tumor classes in our existing data set were identified by statistical scores from the *t* test (without Bonferroni correction), Mann–Whitney *U* test, and Fisher score. This approach is useful for iden-

tifying genes of potential biological interest that may be relevant to the specific tumor classification, but may not select discriminator gene sets that are specific enough to be used in the predictive classification of new samples.

Second, the ability of our data to predict tumor class from differential gene expression was evaluated using an SVM with a linear kernel. An estimate of the test performance was obtained using leave-one-out testing in which the SVM was trained on 26 data points and tested on the data point left out, with successive rotation through the dataset of the left-out example. A leave-one-out estimate of the test performance gives the most unbiased estimate of the test error (the test error is least influenced by individual examples and most representative of the distribution as a whole), but is not expected to be the best choice for the bias-variance tradeoff (Hastie et al., 2001); consequently, results for *n*-fold cross validation were also calculated. Good predictive ability was found to be dependent on using a small set of genes; thus, feature selection played a critical role. Because the number of high-quality features was large (17,790), we used filter methods (Dudoit and Fridlyand, 2003) and graded the significance of individual features according to statistical scores, again using the Fisher score, the Mann–Whitney *U* test, and the *t* test (without Bonferroni correction). During evaluation of the test error, the 26 examples in the training set change with every leave-one-out rotation. Consequently, we reevaluated these statistical scores for each of the 27 evaluations on the test point without incorporation of the test point in the computation of the score. This approach typically identifies a slightly different set of highest-scoring discriminator genes for each leave-one-out rotation, with partial overlap between the sets (since some genes generally appear in multiple sets). For example, the union of 27 discriminator feature sets using the 5 highest-scoring class D versus class N discriminators for each leave-one-out rotation (see Results section) actually contained 15 features, 3 of which were found in at least 26 of the sets.

The ability of discriminator gene sets to distinguish between classes was visualized by the TIGR Multi-Experiment Viewer (Saeed et al., 2003) or Genesis (Sturn et al., 2002) implementation of hierarchical clustering (using the Euclidean distance metric) and by multidimensional scaling (centered correlation metric) using the NCI BRB ArrayTools (http://linus.nci.nih.gov/BRB-ArrayTools.html).

RESULTS

Tumor Samples

A study in one of our laboratories previously showed an association between increased risk of relapse and increased genomic copy number on 1q as measured by CGH (Hing et al., 2001), and an even stronger association with overexpression of genes on 1q analyzed by the novel technique of comparative expressed sequence hybridization to chromosomes (CESH; Lu et al., 2002). Therefore, we ascertained the genomic copy number of 1q for all tumors included in this microarray expression analysis. For the 14 tumors that had not previously been subjected to CGH analysis, we used FISH to determine copy number. The clinical characteristics of tumors in the N and D classes were similar for median age at diagnosis and sex ratio. There were more tumors with distant metastases at diagnosis in the D class; this would be expected to exaggerate rather than obscure any differences in molecular characteristics between the two groups. As expected (Hing et al., 2001), the incidence of 1q gain was higher in the relapsing D-class tumors (7 of 13, Table 1) than in the nonrelapsing N-class tumors (1 of 14).

Array Data

Following lowess normalization and application of the quality filter, data from 17,790 features on 27 arrays were available for further analysis. These features represented nearly 12,000 unique UniGene clusters, of which more than 9,000 contained known unique NCBI LocusLink genes. Approximately 4.5% of the high-quality features mapped to 1q. Full raw and processed data compliant with the MIAME guidelines (http://www.mged.org/Workgroups/MIAME/miame_checklist.html) are available on the supplementary Web site (http://www.icr.ac.uk/paedonc/Wilmsarraydata.html).

Classification by Relapse Category

One hundred and thirty-eight features representing genes that were differentially expressed between relapsing class D (13 tumors) and nonrelapsing class N (14 tumors) samples were identified by *t* test ($P < 0.005$; Fig. 1A). Genes of potential interest that were measurably overexpressed in class D tumors (relative to those in class N) included cyclin *CCNT1*, oncogene *KRAS2*, cytokine *PGLYRP*, and the *BIRC6* apoptotic inhibitor. Relatively underexpressed genes in class D tumors included the *NOTCH1* signaling molecule, apoptotic factors *FAF1*, *TNFRSF1A*, and *GRIM19*, inter-

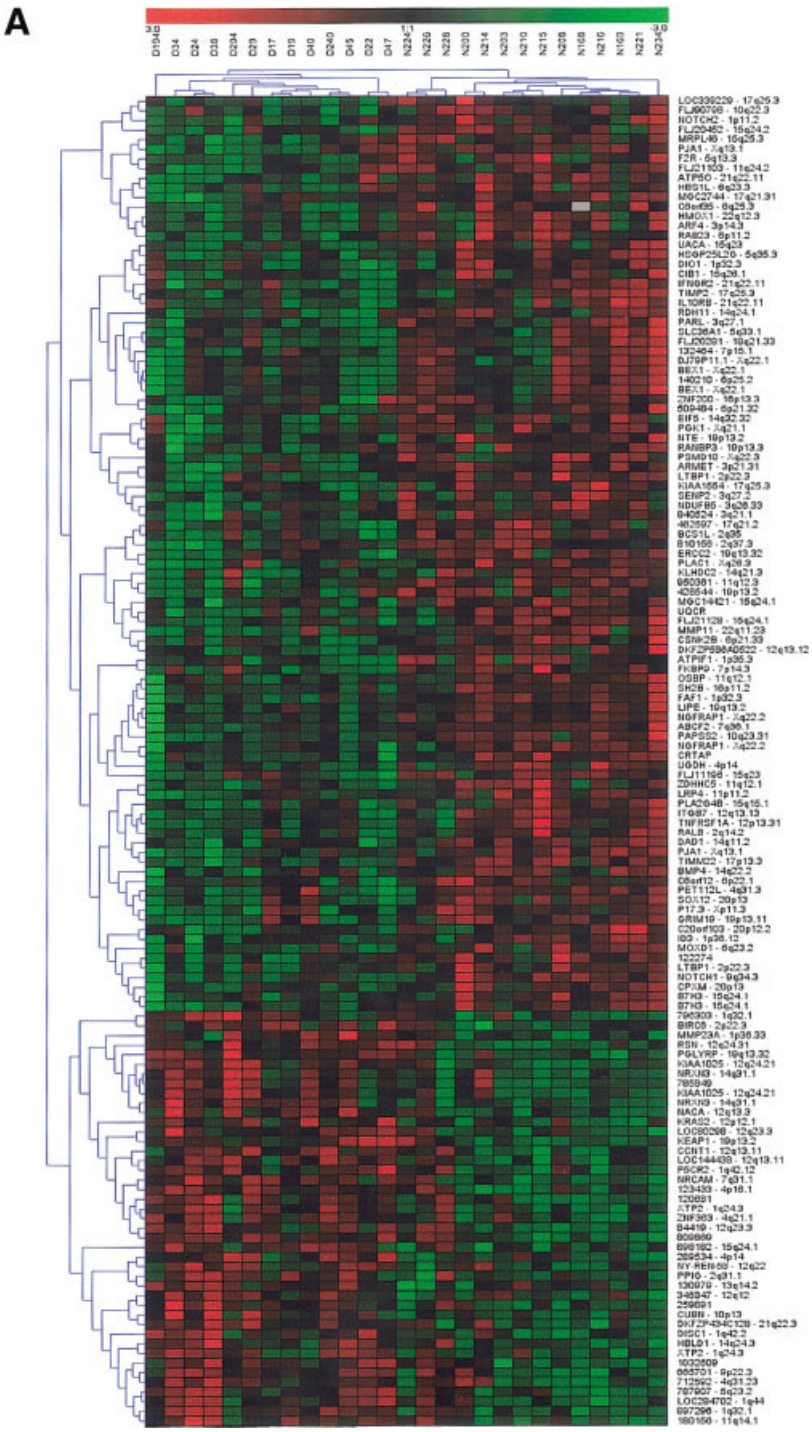

Figure 1. Features representing genes differentially expressed in relapsing (class D) and nonrelapsing (class N) Wilms tumors. **(A)** Heat map with hierarchical clustering dendrograms (Euclidean distance metric) for 138 features identified by *t* test ($P < 0.005$) using normalized log 2–transformed expression ratios. Feature rows were further normalized for clustering and display by subtracting the mean row value from each expression ratio and dividing the result by the standard deviation of the expression ratios in that row. Gene names and cytobands are given where known; unmapped features are identified by their IMAGE IDs. **(B)** Number of leave-one-out test errors (*y* axis) versus number of top-ranked features per leave-one-out rotation (*x* axis) remaining, with feature ranking by *t* test. A minimum error of 1 misclassification from 27 was obtained using the 3–5 highest-scoring features per rotation. **(C)** Heat map with hierarchical clustering dendrograms (as above) for 15 features representing 15 differentially expressed genes with generalizable expression signatures, that is, the union of all 27 partially overlapping discriminator feature sets obtained using the 5 highest-scoring features (by *t* test) per leave-one-out rotation. The 3 features used in 26–27 leave-one-out evaluations are shaded. **(D)** 3-dimensional plot of log 2 expression ratios of the 3 most frequently used features in leave-one-out evaluation for all 27 samples. Gene names and feature IDs are indicated; red diamond is class D, and green cross class N.

leukin receptor *IL10RB*, and the mesodermal/skeletal development factor *BMP4*. The fold-change in expression ratio between classes D and N was modest for the genes detected by all 138 features; few changes more than 2-fold greater were identified by *t* test, and the median fold-change was only ≈1.3-fold. However, these changes were consistently observed, and the expression signature detected by the 138 features in combination was able to separate the two relapse classes clearly (Fig. 1a).

Although genomic gain of 1q is associated with an increased risk of relapse (Hing et al., 2001), only 7 (≈5%) of the 138 features that detected differential expression between relapsing and nonrelaps-

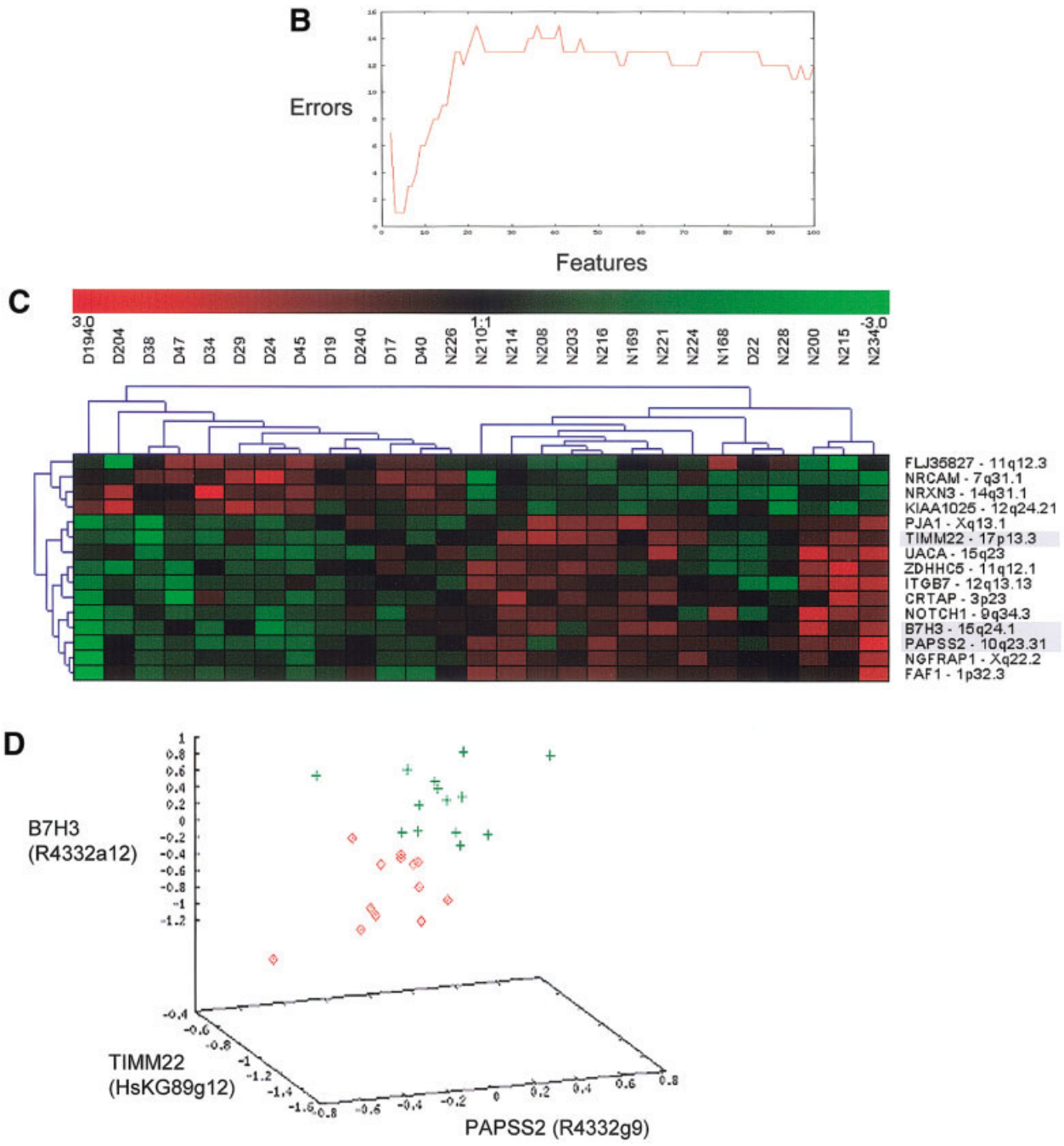

Figure 1. (Continued)

ing tumors mapped to 1q, a proportion similar to the ≈4.5% of total high-quality array features that mapped to this region. None of the genes detected by these 7 features (including 2 representing *XTP2*; single clones for *P5CR2*, *DISC1*, and *LOC284702*; and 2 additional clones—IMAGE IDs 897296 and 796303—that do not represent known genes) has a defined biological function immediately suggestive of a role in tumor relapse.

To test whether we could identify genes whose expression is predictive of relapse rather than merely associated with this outcome, we conducted a more rigorous analysis using an SVM with linear kernel and leave-one-out testing. For this analysis, we experimented with 3 filter methods to rank the significance of the genes. If no feature selection was used (i.e., the entire data set of high-quality genes was included in the analysis), generalization

was found to be weak, with a leave-one-out test error of 10 misclassifications from 27. However, when feature selection was implemented, the prediction ability improved substantially. With ranking of features by Fisher score, the lowest test error achieved was 5 errors from 27 (using the top 5 features), whereas using the Mann–Whitney score, the minimum test error fell to 4 (not shown). Optimum performance was obtained with the *t* test (Fig. 1B), where the minimum error was reduced to 1 misclassification in 27 using the 3–5 highest-scoring features per leave-one-out rotation (this was to be expected because the *t* test is generally superior to nonparametric methods such as the Mann–Whitney test for cDNA array feature selection when an approximately normal distribution can be assumed). Using the 5 highest-scoring features per leave-one-out test, the union of all 27 (partially overlapping) discriminator feature sets contained 15 features, each representing a unique gene (see Materials and Methods section). All but one (*FLJ35827*) of these features also were present in the 138-feature set identified by *t* test alone (above). None of the 15 genes mapped to 1q, and expression fold-change between relapse classes D and N was again very modest, although consistent. The previously defined functions of these 15 genes were not known to be associated with tumor relapse, although it can be speculated that reduced expression of two apoptotic factors (*FAF1* and *NGFRAP1*) in the relapsing tumors may be of some significance. The only gene with a specific link to Wilms tumor is *NOTCH1*, an intercellular signaling receptor that interacts with *NOV* [an overexpressed factor in Wilms tumor (Sakamoto et al., 2002)]. However, the significance of its reduced expression in relapsing tumors is unclear. Interestingly, two genes (*NRXN3* and *NRCAM*) with previously defined functions in neural cell adhesion were overexpressed in the relapsing tumors, whereas another cell adhesion molecule (*ITGB7*) was underexpressed in these tumors.

Hierarchical clustering (Fig. 1C) and multidimensional scaling (not shown) indicated that this set of features distinguished well, but not perfectly, between the relapsing and nonrelapsing tumors. However, we noted that all 15 features were not used equally in the 27 leave-one-out evaluations; 3 features (representing *PAPSS2*, *TIMM22*, and *B7H3*) were used in 26 or all 27 evaluations (Fig. 1C, shaded), whereas the remainder were used in fewer than half of the evaluations (6 features were used only once). Interestingly, *PAPSS2*, which encodes a developmentally important protein with sulfurylase and kinase activities, was reported to be expressed at much higher levels in a nonmetastatic colon carcinoma cell line than in an isogenetic highly metastatic line (Franzon et al., 1999); its expression was also elevated in our nonrelapsing tumors (Fig. 1C). *TIMM22* encodes an inner mitochondrial membrane protein translocase with no documented role in tumor biology, though its differential expression might reflect some difference in the metabolic or apoptotic states of relapsing and nonrelapsing tumors. The *B7H3* gene encodes an immune cell costimulator; decreased expression of this factor in the relapsing tumors may also be significant, as transfection of this gene into melanoma cells was shown to increase antitumor cytotoxic (cocultured) lymphocyte activity (Chapoval et al., 2001).

When the SVM classifier was restricted to using only these 3 prominent features, a zero leave-one-out test error was obtained, whereas both hierarchical clustering (not shown) and a 3-dimensional plot of the expression ratios (Fig. 1D) indicated that class D and class N tumors were clearly separable using this small subset of features. As a random baseline comparison, the leave-one-out test errors using 100 random permutations of the class labels were computed. This gave 12.95 ± 0.26 test errors from 27, using the 3 top-ranked features. We also performed 9-fold cross-validation (24 training and 3 test points) averaged over 1,000 random reshufflings of the order. This gave percentage test errors of 10.9 ± 1.2 (5 features), 14.5 ± 1.1 (4 features), 6.4 ± 1.2 (3 features), and 15.1 ± 1.2 (2 features). These results therefore suggest that an apparently reliable prediction of relapse can be achieved, although the limited number of high-ranking features involved, and the modest size of the changes in expression between tumor classes, mean that the data need to be interpreted cautiously.

Classification by Chromosome Arm 1q Status

Although most genes identified by relapse category analysis did not map to chromosome arm 1q, we reasoned that it still would be useful to identify discriminators that classify these tumors by their 1q status. Both genomic gain of 1q (Hing et al., 2001) and overexpression from this region (Lu et al., 2002) have been found to be associated with an increased risk of relapse. Thus, 1q status discriminator genes may be of some importance in relapse, even if they are not the primary predictors of clinical outcome. The results of this comparison should also give some insight into the general association

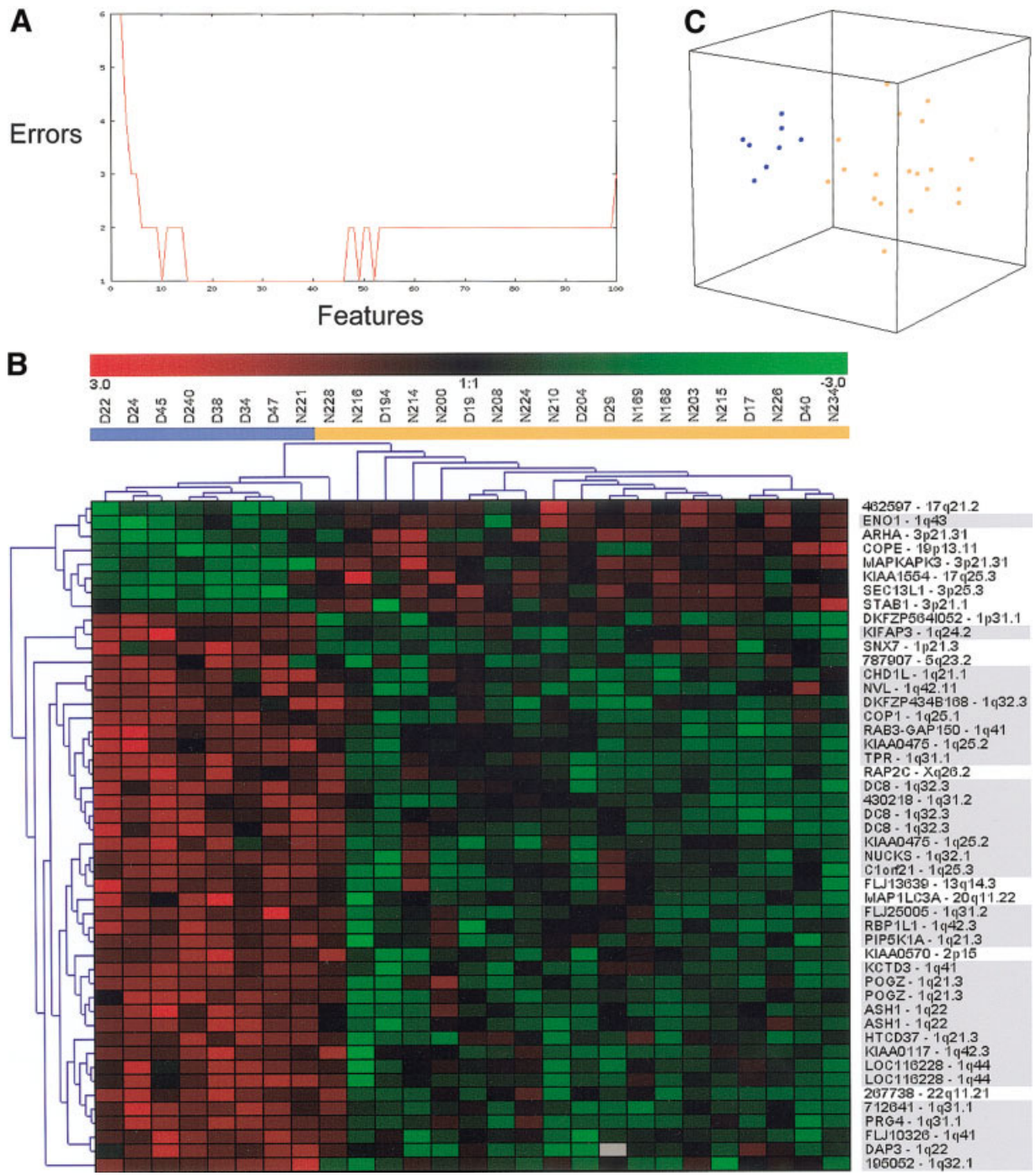

Figure 2. Features representing genes differentially expressed in 1q-gain and no-1q-gain classes of Wilms tumors. (**A**) Number of leave-one-out test errors (*y* axis) versus number of top-ranked features per leave-one-out rotation (*x* axis) remaining, with feature ranking by Fisher score. A minimum error of 1 misclassification from 27 was obtained across a wide range of highest-scoring features per rotation. **(B)** Heat map with hierarchical clustering dendrograms (Euclidean distance metric) for 48 features representing differentially expressed genes with generalizable expression signatures, that is, the union of all 27 partially overlapping discriminator feature sets obtained using the 20 highest-scoring features (by Fisher score) per leave-one-out rotation. Normalized log 2–transformed expression ratios were used as input. Feature rows were further normalized for clustering and display by subtracting the mean row value from each expression ratio and dividing the result by the standard deviation of the expression ratios in that row. Gene names and cytobands are given where known, unmapped features are identified by their IMAGE IDs, and features that map to chromosome arm 1q are shaded. Samples with 1q gain are indicated by a blue bar, and samples with no 1q gain by an orange bar. **(C)** Multidimensional scaling plot derived from the expression ratios of the genes differentially expressed between samples with 1q gain (blue) and no 1q gain (orange), identified using leave-one-out validation with 20 features per rotation (48 unique features in total).

between genomic gain (as measured by CGH or FISH) and microarray expression profiles in Wilms tumor.

We therefore applied the SVM and leave-one-out testing to the problem of distinguishing between 1q-gain and no-1q-gain tumors, irrespective of relapse category. Without the use of feature selection, a better-than-random (6 from 27) test error was obtained. Feature selection by Mann–Whitney *U* test or *t* test reduced the minimum test error to 4 misclassifications from 27. A significantly improved minimum test error of 1 misclassification from 27 was obtained using Fisher score feature selection, possibly because this method is more influenced by the magnitude of differences between values than the rank-ordering of values as in the Mann–Whitney and *t* tests. Low test errors (1–2 misclassifications) were obtained across a broad numerical range of selected features (Fig. 2A). The 1q-gain discriminator gene sets identified by the SVM were greatly enriched in genes that mapped to 1q. Of the 48 features identified as discriminators by the SVM when using 20 features in each leave-one-out iteration (test error = 1, Fig. 2A), 33 (69%) mapped to 1q, 32 of which represented genes overexpressed rather than underexpressed in the tumors with gain of 1q (Fig. 2B). This gene set discriminated clearly between tumors with and without 1q gain, as illustrated by multidimensional scaling analysis (Fig. 2C). Genes on 1q also tended to be the most frequently selected features in leave-one-out validation (9 of 10 features used in at least 26 of 27 leave-one-out iterations in the 20-feature example mapped to 1q). Moreover, a reduction in the number of features used in classification led to further enrichment in the proportion that mapped to 1q; of the 23 features identified as discriminators when using 10 features in each leave-one-out iteration, 19 (≈83%) mapped to 1q. In comparison, only about 4.5% of all the high-quality features on the array with reliable matches to the genome mapped to 1q. A plot comparing median expression ratios across chromosome 1 in the 1q-gain and no-1q-gain classes also suggested significant overexpression of genes mapping to the 1q arm in the 1q-gain class (Fig. 3a), whereas expression ratios across chromosome 2 were essentially random (Fig. 3b). Thus, there was a clear association between genomic 1q gain measured by molecular cytogenetic techniques and 1q overexpression in this set of tumors. Our ability to identify individual overexpressed genes on 1q as the best discriminators of genomic 1q status also provides significant validation of our experimental and analytic methodology.

In a previous study from one of our laboratories (Lu et al., 2002), overexpression from 1q relative to the level of expression in normal lymphocyte control cells was found to be associated with subsequent relapse, even in the absence of underlying genomic gain. The current microarray data cannot be directly compared with the previous CESH data, since both the sample cohort and the control RNA differ. However, an assessment of the potential association between relapse category and 1q status can be made from the present data by comparing median class D with median class N expression ratios across chromosome 1 (Fig. 3c). Although a modest increase in expression from 1q was observed in class D tumors compared with those of class N (Fig. 3c), this effect was not as pronounced as the increase in expression from 1q in the 1q-gain tumors compared with the no-1q-gain samples (Fig. 3a). Furthermore, when tumors with known 1q gain were excluded from the analysis, no increase in expression from 1q was observed in class D samples compared with class N samples (Fig. 3d). This suggests that, in general, overexpression from 1q in our sample series was related to the corresponding genomic gain and not specifically to relapse class. The overrepresentation of those with 1q gain in relapsing class D samples (Table 1) appears to have been largely or completely responsible for the apparent increase in the 1q expression ratios (Fig. 3c) in these samples. We also note that there was little overlap between relapse-class and 1q-gain status discriminator gene sets. Although a number of the 138 features initially identified as relapse-class discriminators by *t* test (without cross-validation) mapped to 1q (Fig. 1A), none of the genes identified by the SVM as relapse-class discriminators with the ability to generalize (using leave-one-out validation) were 1q-gain class discriminators. The reason for this apparent discrepancy with the earlier study (Lu et al., 2002) is unclear. Three of the five relapsing tumors previously found by CESH to have 1q overexpression (relative to normal lymphocytes) in the absence of 1q gain were included in the present study (D17, D19, and D40; see http://www.icr.ac.uk/paedonc/Wilmsarraydata.html) and contributed to the class comparisons shown in Figure 3a–d. These tumors did not individually appear to exhibit 1q overexpression by array analysis relative to the HEK-293 control RNA (not shown). However, comparisons between CESH and microarray data are not straightforward, especially when (as here) the ref-

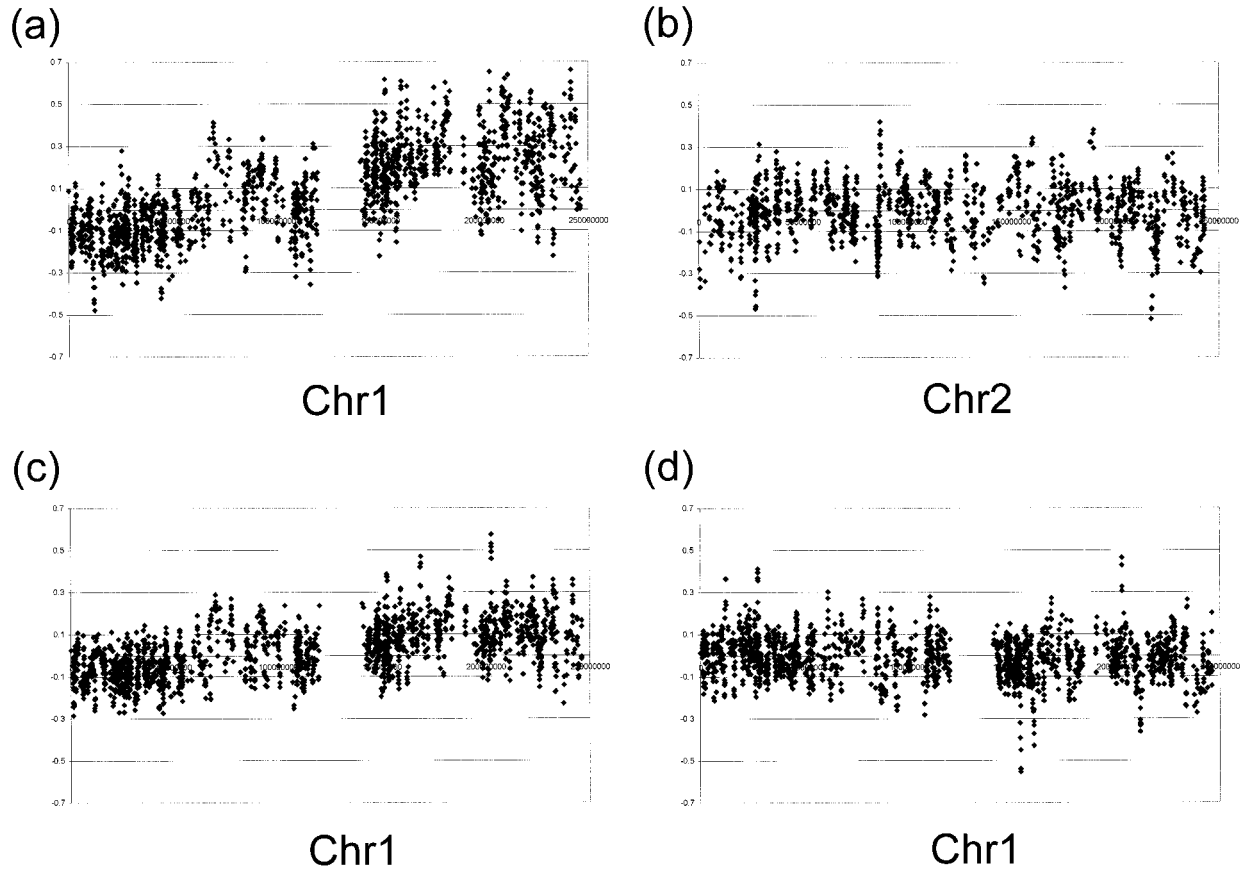

Figure 3. Comparisons of expression ratio versus chromosomal position (all ratios are overlapping moving averages of data from 5 adjacently mapped clones, log 2 scale): (**a**) [median expression ratio in all 1q-gain tumors/median expression ratio in all no-1q-gain tumors] (*y* axis) versus position on chromosome 1 (*x* axis); (**b**) median expression ratio in all 1q-gain tumors/median expression ratio in all no-1q-gain tumors (*y* axis) versus position on chromosome 2 (*x* axis); (**c**) median expression ratio in all relapsing class D tumors/median expression ratio in all nonrelapsing class N tumors (*y* axis) versus position on chromosome 1 (*x* axis); (**d**) median expression ratio in class D tumors without 1q gain/median expression ratio in class N tumors without 1q gain (*y* axis) versus position on chromosome 1 (*x* axis).

erence RNA samples differ. The interpretation of CESH results is more dependent on the choice of control RNA than is array analysis; averaged class comparisons of the type presented in Figure 3 (where the control RNA expression pattern is effectively cancelled out by the ratio comparison) are not currently possible with CESH. The CESH protocol also has an RNA amplification step that could allow the detection of differential expression in genes that are expressed only at low levels, but this may make the technique more liable to experimental artifact. CESH surveys gene expression from the complete genome (albeit at a relatively low resolution), whereas microarray analysis, although capable of higher resolution, is limited to the specific features that are spotted on the array. At present, we can only say that the array data do not provide further confirmation of any association between tumor relapse and overexpression from 1q in the absence of an underlying genomic gain, but neither do they invalidate the previous study. Any potential role for 1q genes in tumor relapse therefore remains to be elucidated.

Classification by Clinical Stage

We were not able to define a statistically significant molecular classification for clinical stage (not shown). The number of tumors at each stage was insufficient to detect significance in comparisons between individual stages (Table 1), and grouping the tumors according to stage as either low (stages I and II) or high (stages III and IV) did not reveal any discriminator genes that could be verified by the SVM with leave-one-out testing.

Wilms Tumor Marker Genes

For the purpose of tumor classification, our use of HEK-293 cells as a source of reference RNA was somewhat arbitrary, since the reference was included simply to provide a common standard to

which all the samples could be compared. In principle, any reference that hybridized to a high proportion of the array features (as did labelled HEK-293 cDNA) could have been used. Although HEK-293 cells were originally derived from human embryonal kidney tissue and could thus be regarded as a physiologically relevant control, they may not be representative of the authentic cellular origin of Wilms tumor, and it cannot be assumed that genes differentially expressed between a tumor and this (or any available) single control are necessarily important to tumor biology. Such direct comparisons between tumor and control are therefore of limited value but may serve to highlight genes or pathways worthy of further investigation. With these caveats, we examined the data to identify genes that were differentially expressed in the Wilms tumor samples relative to the HEK-293 cell control, regardless of tumor classification.

Median tumor/control gene expression ratios of at least 2.0 were detected by 750 features, whereas 728 features detected genes with median ratios less than or equal to 0.5 (the full tumor/HEK-293 ratios are available in the supplementary data, http://www.icr.ac.uk/paedonc/Wilmsarraydata.html). Tumor overexpressed genes of particular interest, several of which have previously been described in Wilms tumor, included *WT1* (Pritchard-Jones and Fleming, 1991), *MYCN* (Nisen et al., 1986), *ABCC1* (CFTR/MRP), *MEOX1* (MOX1), *CDKN1C* (p57/Kip2), *JUNB*, *NTRK2* (TRKB) (Donovan et al., 1994; Eggert et al., 2001), *IGF2* (Reeve et al., 1985; Scott et al., 1985), and other insulin-like growth factor binding protein genes (*IGFBP2, IGFBP7, CTGF*); the cadherins and proto-cadherins *CDH11* (Schulz et al., 2000), *CDH12, CDH13, PCDH16, PCDH18, PCDHA1,* and *PCDHB2*; and the glypicans *GPC3* (Saikali and Sinnett, 2000; Toretsky et al., 2001), and *GPC5*. Tumor underexpressed genes included *MYC* (Nisen et al., 1986), *CDC6*, *CDKN2C* (p18) (Arcellana-Panlilio et al., 2000), *RBBP7*, and MDR/TAP family members *ABCB1* (Re et al., 1997) and *ABCB10*. Our set of overexpressed genes had some overlap with gene sets reported in recent comparisons between Wilms tumor and various controls. In one study (Takahashi et al., 2002), overexpression of 267 genes in a small group of Wilms tumors in comparison to a pooled adult kidney reference was reported. Of the top 40 listed array clones, 15 (representing *C5orf13*, *CDH13*, *COL6A3, CPXM, CRABP2, ENC1, IGF2, LAMA4*, "*LOC143914,*" *MEST, MGC29643*, *RBP1*, *STMN1, TMEFF1*, and the unnamed cluster Hs.443132 in UniGene 161) had accession numbers that mapped to overexpressed genes in our comparison with HEK-293 cells. Seven of the overexpressed genes we detected (*CRABP2*, *ETV4* (*E1AF*), *MEOX1* (*MOX1*), *MN1*, *NNAT*, *SALL2*, and *FZD7*) have also been reported as significantly overexpressed in Wilms tumor when compared to a panel of heterologous tumors and control tissues (Li et al., 2002). We suggest that those genes found to be differentially expressed between Wilms tumor and multiple references by independent methods (such as *CRABP2*) are most likely to be of genuine physiological importance.

DISCUSSION

We have examined the expression profiles of 27 favorable histology Wilms tumors and identified sets of discriminator genes that can potentially be used to classify these tumors by relapse category and genomic 1q-gain status. We have also identified genes that are overexpressed in Wilms tumors as a group relative to a human embryonal kidney cell line.

Two approaches were used to examine the potential association between the gene expression profile of Wilms tumor at the time of initial diagnosis and clinical outcome. First, statistical scoring (e.g., the *t* test) was used to identify genes that were differentially expressed between class D (subsequently relapsing) and class N (nonrelapsing) tumors in our existing sample set. These genes encode proteins with diverse molecular functions, including transcription factors, developmental regulators, apoptotic factors, and signaling molecules, and their expression may be relevant to the process of relapse. Second, we used a more sophisticated approach to determine whether these results generalize, that is, whether an expression signature can be defined that differentiates between the two classes and also can be used to classify new data correctly. We therefore applied machine learning algorithms and feature selection using three different statistical scores to find a subset of genes expressed in a pattern that could be learned by a classifier, and we tested for the ability to predict tumor outcome by cross validation. Interestingly, we found that the predictive test error was minimized using a very small number of features. Several of these features represent known apoptotic, anti-metastatic, and antitumor cytotoxic factors that were relatively underexpressed in the relapsing tumors. This result appears encouraging but should be treated with caution. The fold-change in expression of these genes between class N and class D tumors was modest (<2-fold), although

both expression within each tumor class and the degree of differential expression between classes were consistent. It also should be noted that, given the large number of features analyzed and the statistical scoring used (e.g., a standard *t* test without Bonferroni correction), the expression profiles of some genes might be expected to align with the class labels by chance, even when the combined expression pattern of a set of discriminators enables statistically significant tumor classification. Thus, the ability to predict relapse may reside in an ensemble of features and not in any individual feature taken in isolation. Given the small number of features involved, the (albeit imperfect) classification achieved with a total of 15 features (Fig. 1C) may prove more robust when applied to independent data than the complete separation obtained with a total of 3 features (Fig. 1D). Moreover, discriminator genes that mapped to more than one array feature, where differential expression was not statistically significant at the prescribed level for every feature (observed in some cases), should be treated with particular caution (*t*-test *P* values for individual features are available on the supplementary Web site).

The potential ability of subtle differences in the expression of a small number of genes to predict outcome in Wilms tumor in any case needs to be tested in a large, independent set of samples before it can be regarded as clinically relevant. This would be facilitated by custom arrays on which possible genes of interest are represented by many replicate probes to enhance the accuracy of measurements, or alternatively by similarly replicated quantitative RT-PCR assays for each gene. These studies may also benefit from the use of RNA amplification techniques, which could allow the identification of any additional discriminators that are expressed at very low levels (perhaps with larger-fold changes between tumor classes than those so far detected). If a robust prognostic classifier can be defined, it should prove clinically useful, enabling physicians to refine therapeutic strategies at an early stage.

The close similarity between relapse classes may reflect complex patterns of regulation that could be difficult to dissect by expression profiling alone. It is also possible that tumor cell populations are not homogenous at this stage of the disease, so that the expression signatures of those cells that will ultimately be responsible for relapse are partially masked by gene expression in neighboring cells with different transcriptional programs. This point is of particular relevance to Wilms tumor, in which diverse histological subtypes of tumor cells coexist. Examining the expression profiles of tumors sampled at the time of relapse will help to elucidate this possibility, whereas array analysis of microdissected cell populations and measurement of discriminator gene expression in tissue sections at the time of diagnosis could provide further insight into any early nonhomogeneity in expression patterns.

Only one previous study has attempted to compare Wilms tumor samples by array profiling using a classification based on clinical outcome (Takahashi et al., 2002). In this study, a *t*-test-based statistical analysis was used to identify genes differentially expressed between poor outcome (death from disease) and good outcome (survival) subgroups of a small cohort of tumors. However, the sample set contained only 3 tumors in the poor outcome group (in which relapse could only be confirmed in 2 cases), and it is not clear whether such a small set is statistically meaningful, or whether the differential expression observed would be evident in a larger dataset. We did not observe any overlap between the 40 outcome-related genes reported by this group and the relapse class discriminators identified in the present study by *t* test or SVM analysis.

It is perhaps surprising that there was little overlap between the relapse class and 1q-gain class discriminators we identified, since genomic 1q gain is itself a statistically significant prognostic factor for relapse. However, although 1q gain was largely confined to the class D tumors (only 1 class N tumor had measurable 1q gain, Table 1), 6 of the 13 class D tumors did not have measurable 1q gain, a proportion consistent with a previous study from one of our laboratories (Hing et al., 2001). It is possible that 1q gain is secondary to other, more important molecular events in the process of tumor relapse. Alternatively, 1q genes not represented in our set of 17,790 high-quality array features may be the critical relapse factors; small chromosomal gains including such genes may also be undetectable by metaphase CGH but could potentially be identified by, for example, tiling path array CGH. Nevertheless, the 1q classification results provide a clear demonstration of the relationship between genomic gain (as defined by CGH or FISH) and detectable overexpression across a large chromosomal region in Wilms tumor, and the high proportion of genes that map to 1q in the 1q-gain discriminator gene set is a striking validation of our experimental and analytic methodology.

It was not possible to define a molecular classification for clinical stage from the current data. The number of tumors at each stage was too limited to

detect statistical significance in comparisons of individual stages, and grouping the samples into low- and high-stage classes did not reveal a discriminator gene set that could be verified by the SVM with leave-one-out validation. In any case, staging is based on clinicopathologic criteria that may have no simple relationship to underlying changes in gene expression. In this context, it may be significant that the latest National Wilms Tumor Study Group study found that an adverse molecular marker (1p loss of heterozygosity) occurred at equal frequencies across all tumor stages, with the exception of small stage I tumors in children less than 2 years old (Dr. Paul Grundy, NWTSG, personal communication).

Finally, our identification of Wilms tumor "marker" genes that are over- or underexpressed relative to a single control must be treated very cautiously. We note that several genes that were differentially expressed between our tumor set and the human embryonal kidney cell line control have previously been identified as Wilms tumor markers in single-gene studies, or as differentially expressed genes in earlier expression profiling studies (albeit relative to equally arbitrary controls), or are otherwise of particular physiological interest. However, the actual significance of these genes to Wilms tumor biology remains to be evaluated. Genes that are now known to be differentially expressed relative to a variety of controls in different studies (e.g., *CRABP2*) may be the most attractive candidates for further analysis. Such cross-study comparisons will be greatly facilitated by the public availability of complete microarray data sets. To this end, we have made our full data available on a supplementary Web site (http://www.icr.ac.uk/paedonc/Wilmsarraydata.html).

In summary, we have used microarray analysis of a series of Wilms tumor samples to identify a set of putative tumor marker genes that were overexpressed relative to a common control, demonstrated close agreement between chromosomal gain on 1q and gene overexpression from this region, and defined a generalizable expression signature for tumor relapse with potential clinical prognostic value.

ACKNOWLEDGMENTS

This study received local research ethical approval (RMH protocol 1494). For access to specimens, we thank the United Kingdom Children's Cancer Study Group Tumour Bank, which is funded by Cancer Research UK; the National Wilms Tumor Study Group and the Cooperative Tissue Network (Columbus, OH), which are funded by the National Cancer Institute (United States), and the pediatric pathologists and surgeons at St. George's Hospital, London, and at the Royal Manchester Children's Hospital. Multidimensional scaling analysis was performed using BRB ArrayTools developed by Dr. Richard Simon and Amy Peng. We thank Dr. Chris Jones for helpful comments on the manuscript.

REFERENCES

Arcellana-Panlilio MY, Egeler RM, Ujack E, Pinto A, Demetrick DJ, Robbins SM, Coppes MJ. 2000. Decreased expression of the INK4 family of cyclin-dependent kinase inhibitors in Wilms tumor. Genes Chromosomes Cancer 29:63–69.

Bardeesy N, Falkoff D, Petruzzi MJ, Nowak N, Zabel B, Adam M, Aguiar MC, Grundy P, Shows T, Pelletier J. 1994. Anaplastic Wilms' tumor, a subtype displaying poor prognosis, harbours p53 gene mutations. Nat Genet 7:91–97.

Beer DG, Kardia SL, Huang CC, Giordano TJ, Levin AM, Misek DE, Lin L, Chen G, Gharib TG, Thomas DG, Lizyness ML, Kuick R, Hayasaka S, Taylor JM, Iannettoni MD, Orringer MB, Hanash S. 2002. Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat Med 8:816–824.

Bengtsson H. 2002. com.braju.sma–Object-oriented microarray analysis in 100% R. (http://www.maths.lth.se/help/R/com.braju.sma/HBengtsson-com.braju.sma.pdf).

Camassei FD, Jenkner A, Rava L, Bosman C, Francalanci P, Donfrancesco A, Alo PL, Boldrini R. 2003. Expression of the lipogenic enzyme fatty acid synthase (FAS) as a predictor of poor outcome in nephroblastoma: An interinstitutional study. Med Pediatr Oncol 40:302–308.

Chapoval AI, Ni J, Lau JS, Wilcox RA, Flies DB, Liu D, Dong H, Sica GL, Zhu G, Tamada K, Chen L. 2001. B7–H3: a costimulatory molecule for T cell activation and IFN-gamma production. Nat Immunol 2:269–274.

Chen Y, Kamat V, Dougherty ER, Bittner ML, Meltzer PS, Trent JM. 2002. Ratio statistics of gene expression levels and applications to microarray data analysis. Bioinformatics 18:1207–1215.

Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurachi K, Pienta KJ, Rubin MA, Chinnaiyan AM. 2001. Delineation of prognostic biomarkers in prostate cancer. Nature 412: 822–826.

Dome JS, Coppes MJ. 2002. Recent advances in Wilms tumor genetics. Curr Opin Pediatr 14:5–11.

Dome JS, Chung S, Bergemann T, Umbricht CB, Saji M, Carey LA, Grundy PE, Perlman EJ, Breslow NE, Sukumar S. 1999. High telomerase reverse transcriptase (hTERT) messenger RNA level correlates with tumor recurrence in patients with favorable histology Wilms' tumor. Cancer Res 59:4301–4307.

Donovan MJ, Hempstead B, Huber LJ, Kaplan D, Tsoulfas P, Chao M, Parada L, Schofield D. 1994. Identification of the neurotrophin receptors p75 and trk in a series of Wilms' tumors. Am J Pathol 145:792–801.

Dudoit S, Fridlyand J. 2003. Classification in microarray experiments. In: Speed T, editor. Statistical analysis of gene expression microarray data. Boca Raton: Chapman and Hall/CRC. p 115.

Dyrskjot L, Thykjaer T, Kruhoffer M, Jensen JL, Marcussen N, Hamilton-Dutoit S, Wolf H, Orntoft TF. 2003. Identifying distinct classes of bladder carcinoma using microarrays. Nat Genet 33:90–96.

Efferth T, Schulten HG, Thelen P, Bode ME, Beniers AJ, Granzen B, Ringert RH, Mertens R, Gefeller O, Jakse G, Fuzesi L. 2001a. Differential expression of the heat shock protein 70 in the histological compartments of nephroblastomas. Anticancer Res 21: 2915–20.

Efferth T, Thelen P, Schulten HG, Bode ME, Granzen B, Beniers AJ, Mertens R, Ringert RH, Gefeller O, Jakse G, Fuzesi L. 2001b. Differential expression of the multidrug resistance-related protein MRP1 in the histological compartments of nephroblastomas. Int J Oncol 19:367–371.

Eggert A, Grotzer MA, Ikegaki N, Zhao H, Cnaan A, Brodeur GM, Evans AE. 2001. Expression of the neurotrophin receptor TrkB is associated with unfavorable outcome in Wilms' tumor. J Clin Oncol 19:689–696.

Franzon VL, Gibson MA, Hatzinikolas G, Woollatt E, Sutherland GR, Cleary EG. 1999. Molecular cloning of a novel human PAPS synthetase which is differentially expressed in metastatic and non-metastatic colon carcinoma cells. Int J Biochem Cell Biol 31:613–626.

Furey TS, Cristianini N, Duffy N, Bednarski DW, Schummer M, Haussler D. 2000. Support vector machine classification and validation of cancer tissue samples using microarray expression data. Bioinformatics 16:906–914.

Ghanem MA, Van der Kwast TH, Den Hollander JC, Sudaryo MK, Oomen MH, Noordzij MA, Van den Heuvel MM, Nassef SM, Nijman RM, Van Steenbrugge GJ. 2000. Expression and prognostic value of Wilms' tumor 1 and early growth response 1 proteins in nephroblastoma. Clin Cancer Res 6:4265–4271.

Ghanem MA, Van Der Kwast TH, Den Hollander JC, Sudaryo MK, Mathoera RB, Van den Heuvel MM, Noordzij MA, Nijman RJ, van Steenbrugge GJ. 2001a. Expression and prognostic value of epidermal growth factor receptor, transforming growth factor-alpha, and c-erb B-2 in nephroblastoma. Cancer 92:3120–3129.

Ghanem MA, Van der Kwast TH, Den Hollander JC, Sudaryo MK, Van den Heuvel MM, Noordzij MA, Nijman RJ, Soliman EH, van Steenbrugge GJ. 2001b. The prognostic significance of apoptosis-associated proteins BCL-2, BAX and BCL-X in clinical nephroblastoma. Br J Cancer 85:1557–1563.

Ghanem MA, Van Steenbrugge GJ, Van Der Kwast TH, Sudaryo MK, Noordzij MA, Nijman RJ. 2002. Expression and prognostic value of CD44 isoforms in nephroblastoma (Wilms tumor). J Urol 168:681–686.

Ghanem MA, van Steenbrugge GJ, Sudaryo MK, Mathoera RB, Nijman JM, van der Kwast TH. 2003. Expression and prognostic relevance of vascular endothelial growth factor (VEGF) and its receptor (FLT-1) in nephroblastoma. J Clin Pathol 56:107–113.

Hastie T, Tibshirani R, Friedman JH. 2001. The elements of statistical learning. Basel: Springer Verlag. p 193–224.

Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R, Hughes JE, Snesrud E, Lee N, Quackenbush J. 2000. A concise guide to cDNA microarray analysis. Biotechniques 29:548–50, 552–554, 556 passim.

Hing S, Lu YJ, Summersgill B, King-Underwood L, Nicholson J, Grundy P, Grundy R, Gessler M, Shipley J, Pritchard-Jones K. 2001. Gain of 1q is associated with adverse outcome in favorable histology Wilms' tumors. Am J Pathol 158:393–398.

Kent WJ. 2002. BLAT—the BLAST-like alignment tool. Genome Res 12:656–664.

Khan J, Wei JS, Ringner M, Saal LH, Ladanyi M, Westermann F, Berthold F, Schwab M, Antonescu CR, Peterson C, Meltzer PS. 2001. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nat Med 7:673–679.

Li CM, Guo M, Borczuk A, Powell CA, Wei M, Thaker HM, Friedman R, Klein U, Tycko B. 2002. Gene expression in Wilms' tumor mimics the earliest committed stage in the metanephric mesenchymal-epithelial transition. Am J Pathol 160:2181–2190.

Lu YJ, Hing S, Williams R, Pinkerton R, Shipley J, Pritchard-Jones K. 2002. Chromosome 1q expression profiling and relapse in Wilms' tumor. Lancet 360:385–386.

Nisen PD, Zimmerman KA, Cotter SV, Gilbert F, Alt FW. 1986. Enhanced expression of the N-myc gene in Wilms' tumors. Cancer Res 46:6217–6222.

Pomeroy SL, Tamayo P, Gaasenbeek M, Sturla LM, Angelo M, McLaughlin ME, Kim JY, Goumnerova LC, Black PM, Lau C, Allen JC, Zagzag D, Olson JM, Curran T, Wetmore C, Biegel JA, Poggio T, Mukherjee S, Rifkin R, Califano A, Stolovitzky G, Louis DN, Mesirov JP, Lander ES, Golub TR. 2002. Prediction of central nervous system embryonal tumor outcome based on gene expression. Nature 415:436–442.

Pritchard-Jones K, Fleming S. 1991. Cell types expressing the Wilms' tumor gene (WT1) in Wilms' tumors: implications for tumor histogenesis. Oncogene 6:2211–20.

Re GG, Willingham MC, el Bahtimi R, Brownlee NA, Hazen-Martin DJ, Garvin AJ. 1997. Anaplasia and drug selection-independent overexpression of the multidrug resistance gene, MDR1, in Wilms' tumor. Mod Pathol 10:129–136.

Reeve AE, Eccles MR, Wilkins RJ, Bell GI, Millow LJ. 1985. Expression of insulin-like growth factor-II transcripts in Wilms' tumor. Nature 317:258–260.

Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J. 2003. TM4: a free, open-source system for microarray data management and analysis. Biotechniques 34:374–378.

Saikali Z, Sinnett D. 2000. Expression of glypican 3 (GPC3) in embryonal tumors. Int J Cancer 89:418–422.

Sakamoto K, Yamaguchi S, Ando R, Miyawaki A, Kabasawa Y, Takagi M, Li CL, Perbal B, Katsube K. 2002. The nephroblastoma overexpressed gene (NOV/ccn3) protein associates with Notch1 extracellular domain and inhibits myoblast differentiation via Notch signaling pathway. J Biol Chem 277:29399–405.

Schuler GD, Boguski MS, Stewart EA, Stein LD, Gyapay G, Rice K, White RE, Rodriguez-Tome P, Aggarwal A, Bajorek E, Bentolila S, Birren BB, Butler A, Castle AB, Chiannilkulchai N, Chu A, Clee C, Cowles S, Day PJ, Dibling T, Drouot N, Dunham I, Duprat S, East C, Hudson TJ, et al. 1996. A gene map of the human genome. Science 274:540–546.

Schulz S, Becker KF, Braungart E, Reichmuth C, Klamt B, Becker I, Atkinson M, Gessler M, Hofler H. 2000. Molecular analysis of E-cadherin and cadherin-11 in Wilms' tumors. J Pathol 191:162–169.

Scott J, Cowell J, Robertson ME, Priestley LM, Wadey R, Hopkins B, Pritchard J, Bell GI, Rall LB, Graham CF, et al. 1985. Insulin-like growth factor-II gene expression in Wilms' tumor and embryonic tissues. Nature 317:260–262.

Singh D, Febbo PG, Ross K, Jackson DG, Manola J, Ladd C, Tamayo P, Renshaw AA, D'Amico AV, Richie JP, Lander ES, Loda M, Kantoff PW, Golub TR, Sellers WR. 2002. Gene expression correlates of clinical prostate cancer behavior. Cancer Cell 1:203–209.

Skotnicka-Klonowicz G, Kobos J, Los E, Trejster E, Szymik-Kantorowicz S, Daszkiewicz P. 2002. Prognostic value of proliferating cell nuclear antigen in Wilms' tumor in children. Eur J Surg Oncol 28:67–71.

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lonning P, Borresen-Dale AL. 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA 98:10869–10874.

Sredni ST, de Camargo B, Lopes LF, Teixeira R, Simpson A. 2001. Immunohistochemical detection of p53 protein expression as a prognostic indicator in Wilms tumor. Med Pediatr Oncol 37:455–458.

Sturn A, Quackenbush J, Trajanoski Z. 2002. Genesis: cluster analysis of microarray data. Bioinformatics 18:207–208.

Takahashi M, Yang XJ, Lavery TT, Furge KA, Williams BO, Tretiakova M, Montag A, Vogelzang NJ, Re GG, Garvin AJ, Soderhall S, Kagawa S, Hazel-Martin D, Nordenskjold A, Teh BT. 2002. Gene expression profiling of favorable histology Wilms tumors and its correlation with clinical features. Cancer Res 62:6598–6605.

Toretsky JA, Zitomersky NL, Eskenazi AE, Voigt RW, Strauch ED, Sun CC, Huber R, Meltzer SJ, Schlessinger D. 2001. Glypican-3 expression in Wilms tumor and hepatoblastoma. J Pediatr Hematol Oncol 23:496–499.

Udtha M, Lee SJ, Alam R, Coombes K, Huff V. 2003. Upregulation of c-MYC in WT1-mutant tumors: assessment of WT1 putative transcriptional targets using cDNA microarray expression profiling of genetically defined Wilms' tumors. Oncogene 22:3821–3826.

van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, Witteveen A, Glas A, Delahaye L, van der Velde T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH, Bernards R. 2002. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 347:1999–2009.

van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH. 2002. Gene expression profiling predicts clinical outcome of breast cancer. Nature 415:530–536.

Weber-Hall S, Anderson J, McManus A, Abe S, Nojima T, Pinkerton R, Pritchard-Jones K, Shipley J. 1996. Gains, losses, and amplification of genomic material in rhabdomyosarcoma analyzed by comparative genomic hybridization. Cancer Res 56:3220–3224.

West M, Blanchette C, Dressman H, Huang E, Ishida S, Spang R, Zuzan H, Olson JA, Jr., Marks JR, Nevins JR. 2001. Predicting the clinical status of human breast cancer by using gene expression profiles. Proc Natl Acad Sci USA 98:11462–11467.