American Journal of Pathology, Vol. 158, No. 5, May 2001

Expression Profiling of Renal Epithelial Neoplasms

A Method for Tumor Classification and Discovery of Diagnostic Molecular Markers

Andrew N. Young,*§ Mahul B. Amin,*† Carlos S. Moreno,‡ So Dug Lim,* Cynthia Cohen,* John A. Petros,†§ Fray F. Marshall,† and Andrew S. Neish*

From the Departments of Pathology and Laboratory Medicine, Urology,† and Biochemistry,‡ Emory University School of Medicine, Atlanta; and the Atlanta VA Medical Center,§ Decatur, Georgia*

The expression patterns of 7075 genes were analyzed in four conventional (clear cell) renal cell carcinomas (RCC), one chromophobe RCC, and two oncocytomas using cDNA microarrays. Expression profiles were compared among tumors using various clustering algorithms, thereby separating the tumors into two categories consistent with corresponding histopathological diagnoses. Specifically, conventional RCCs were distinguished from chromophobe RCC/oncocytomas based on large-scale gene expression patterns. Chromophobe RCC/oncocytomas displayed similar expression profiles, including genes involved with oxidative phosphorylation and genes expressed normally by distal nephron, consistent with the mitochondrion-rich morphology of these tumors and the theory that both lesions are related histogenetically to distal nephron epithelium. Conventional RCCs underexpressed mitochondrial and distal nephron genes, and were further distinguished from chromophobe RCC/oncocytomas by overexpression of vimentin and class II major histocompatibility complex-related molecules. Novel, tumor-specific expression of four genes—vimentin, class II major histocompatibility complex-associated invariant chain (CD74), parvalbumin, and galectin-3—was confirmed in an independent tumor series by immunohistochemistry. Vimentin was a sensitive, specific marker for conventional RCCs, and parvalbumin was detected primarily in chromophobe RCC/oncocytomas. In conclusion, histopathological subtypes of renal epithelial neoplasia were characterized by distinct patterns of gene expression. Expression patterns were useful for identifying novel molecular markers with potential diagnostic utility. *(Am J Pathol 2001, 158:1639–1651)*

Renal epithelial neoplasms are increasingly common, making up 2% of all adult malignancies today.[1] Traditionally, these neoplasms have been grouped into histopathological subtypes that correlate with distinct clinical features and cytogenetic abnormalities.[2] The most common subtype is conventional, or clear cell, renal cell carcinoma (RCC), which tends to grow as a solid or cystic mass of neoplastic cells containing glycogen and lipid-rich cytoplasm (clear cells). The neoplastic cells have antigenic and ultrastructural characteristics of proximal nephron epithelium.[3] Conventional RCCs occur frequently in individuals with von Hippel-Lindau disease, an autosomal dominant tumor susceptibility syndrome caused by germline loss of the von Hippel-Lindau tumor suppressor gene (*VHL*) located at chromosome 3p25.[4] This association has led to the realization that many conventional RCCs, both familial and sporadic cases, exhibit biallelic loss of the *VHL* gene.[5] In addition to *VHL*, other potential tumor suppressor genes for conventional RCC have been mapped to chromosome 3p.[6,7] Loss of chromosome 3p appears to be an early genetic event in the development of conventional RCCs, as it is seen in clear cell tumors at all stages of development. Many advanced cases of conventional RCC contain additional mutations such as loss of chromosome 17p, which encompasses the p53 locus at 17p13, and loss of chromosome 14q.[8,9] Conventional RCCs are highly vascular and immunogenic tumors, making them targets for anti-angiogenic and immunological therapies.[10]

Oncocytoma and chromophobe RCC are renal epithelial neoplasms that exhibit antigenic characteristics of distal nephron intercalated cells.[11–13] Oncocytomas exhibit circumscribed growth of characteristic neoplastic cells (oncocytes), which contain small round nuclei and eosinophilic cytoplasm packed with mitochondria.[14] Chromophobe RCCs span a spectrum of morphology.[15,16] Some tumors exhibit nested architecture and cells with abundant, eosinophilic granular cytoplasm (eosinophilic variety), whereas others are composed of cells

Supported by the Molecular Based Testing Initiative, Emory University Department of Pathology and Laboratory Medicine.

Accepted for publication February 9, 2001.

Address reprint requests to Andrew S. Neish, M.D., Woodruff Memorial Building, Room 2337, 1639 Pierce Drive, Atlanta, GA 30322. E-mail: aneish@emory.edu.

containing clear cytoplasm filled with vesicles that stain with Hale's colloidal iron (typical variety). There is evidence that the Hale's colloidal iron-positive vesicles actually represent abortive mitochondria.[17] Histologically, the eosinophilic variety of chromophobe RCC can be difficult to distinguish from renal oncocytoma, whereas the typical variety can resemble conventional RCC. The different clinical behaviors of oncocytomas and chromophobe RCCs (the former are invariably benign; the latter are indolent yet malignant) warrant their diagnostic separation. However, overlap in the morphological, immunohistochemical, ultrastructural, and cytogenetic features of these two tumor types suggest that they are closely related.[17–19] Cytogenetically, for example, the loss of chromosome 1 is a common finding in both tumor types,[2,20,21] with oncocytomas often exhibiting loss of chromosomes 1 and Y, and chromophobe RCCs often exhibiting concurrent loss of multiple chromosomes, including chromosome 1.[22,23] Translocations involving chromosome 11q13 may define a distinct subset of oncocytomas.[24] Mutations of mitochondrial DNA have also been described in both chromophobe RCCs and oncocytomas.[25] However, it is still unclear if these mutations are consistent findings in either lesion, or if they are directly related to the pathogenesis or mitochondrion-rich morphology of these tumors.

The underlying pathogenetic mechanisms of renal neoplasms, and of neoplasms in general, remain a mystery, but new insights into the pathobiology of neoplasia are emerging as technical advances permit large-scale, parallel analysis of eukaryotic gene expression.[26] Using cDNA microarrays to analyze total cellular mRNA, one can compare the relative expression levels of several thousand genes in different cell types simultaneously. Several groups have used such expression profiling methods to identify gene expression patterns associated with various tumors and other disease states.[27–30] Although this field is still relatively new, it has already succeeded in associating specific gene expression patterns with either neoplastic or non-neoplastic cell types,[27–30] allowing researchers to postulate novel gene regulatory circuits and correlate the expression of previously unsuspected genes with specific cell phenotypes. In this sense, gene expression profiling has proved to be an extremely powerful, high-throughput method for identifying specific molecular markers of disease.[31,32] In this report, we have applied gene expression profiling to a series of renal epithelial tumors including conventional RCC, chromophobe RCC, and oncocytoma. Based solely on patterns of gene expression, the renal neoplasms were clustered into subtypes consistent with the clinical, histological, and molecular understanding of these tumors. For several individual genes—vimentin, class II major histocompatibility complex (MHC)-associated invariant chain (CD74), parvalbumin, and galectin-3—differential expression patterns identified by cDNA microarrays were validated in a larger series of tumors by immunohistochemistry, thus confirming these gene products as promising pathological markers for the differential diagnosis of renal epithelial neoplasia.

Materials and Methods

Clinical Material

For microarray experiments, matched specimens of renal tumor and grossly non-neoplastic kidney from the same patient (100–200 mg per specimen) were obtained from the tissue bank maintained by Dr. Fray Marshall. All specimens were promptly frozen and stored at −80°C. Histopathological diagnoses were rendered by the Johns Hopkins University Department of Pathology (Baltimore, MD). The tumors consisted of four cases of conventional (clear cell) RCC (two Fuhrman grade II, one Fuhrman grade III, and one Fuhrman grade IV), one case of chromophobe RCC, and two cases of renal oncocytoma. The conventional RCC patients were a 62-year-old female, a 47-year-old male, a 67-year-old female, and a 57-year-old male; the chromophobe RCC patient was a 73-year-old male; the oncocytoma patients were a 33-year-old female and 72-year-old male. The patients were not followed clinically in this study. For immunohistochemistry, representative tissue blocks were obtained from the Emory University Department of Pathology. The tissues were derived from radical nephrectomies performed at Emory University, and consisted of 20 conventional RCCs (10 Fuhrman grades I-II and 10 Fuhrman grades III-IV), 6 chromophobe RCCs, and 8 oncocytomas. All tissues were fixed in 10% neutral buffered formalin and embedded in paraffin using standard surgical pathology protocols.

Microarray Analysis

Total cellular RNA was prepared from frozen specimens by mechanical disruption in TRIzol reagent (Gibco BRL, Gaithersburg MD), followed by chloroform extraction and alcohol precipitation according to the manufacturer's instructions. PolyA+ RNA was isolated with Oligotex oligo-dT beads (Qiagen, Valencia, CA) according to manufacturer's instructions. PolyA+ RNA (500 ng) was shipped frozen to Incyte Genomics (Palo Alto, CA) for labeling and hybridization using proprietary methods described in detail on the company's Internet site, http://www.incyte.com/gem/technology/index.shtml. For each case, matched polyA+ RNA samples from tumor and non-neoplastic kidney from the same patient were reverse-transcribed into cDNA, incorporating deoxynucleotides coupled to distinct fluorescent dyes; cDNAs derived from non-neoplastic kidneys were labeled with the green dye Cy3, and cDNAs derived from tumors were labeled with the red dye Cy5. Differentially labeled cDNAs from matched tumors and controls were pooled and hybridized simultaneously to Incyte UniGEM v.1 microarrays containing single-stranded cDNA molecules covalently bound to modified glass substrates. The UniGEM v.1 microarrays featured targets for 7075 unique human genes, spotted at known positions on the array grids, as well several proprietary non-human gene targets that served as controls for reverse transcription and hybridization efficiency. The arrays were washed after hybridization and scanned by a specialized fluorescent confocal microscope to detect bound, Cy3-labeled, and

Cy5-labeled cDNAs. Fluorescence intensities at each target position on the array were balanced to the intensities of internal control targets, for which known amounts of cognate mRNAs were added to the reverse transcription reactions. The ratios of balanced Cy5/Cy3 fluorescence intensities at each target represented the ratios of specific gene expression in the tumor *versus* the uninvolved kidney. Reproducibility data generated by Incyte, available on the company's Internet site, indicated that the sensitivity of mRNA detection with UniGEM microarrays is 2 pg, the dynamic range for mRNA detection is 2–2000 pg, the level of detectable differential expression is ≥1.8-fold, and the average coefficient of variation for Cy5/Cy3 ratios is 15%.

Informatics

Differential expression data were analyzed using Incyte's proprietary GemTools software. We used a minimum absolute fluorescence intensity cutoff of 700 units in either the Cy3 or Cy5 channel, in at least two hybridization experiments, to select 4906 expressed genes representing 69% of the array targets. The use of this absolute fluorescence cutoff was determined through personal communications with Incyte. Subsequent data analysis was restricted to genes overexpressed or underexpressed ≥1.8- to 2.0-fold in tumors relative to matched non-neoplastic kidneys. Differential expression profiles were analyzed using the hierarchical average linkage clustering algorithm supplied with Cluster[33] software (Michael Eisen, Stanford University, Stanford, CA). This algorithm used an iterated, agglomerative process of similarity measurements based on the Pearson correlation. In each iteration of the algorithm, the two most similar data elements (ie, expression profiles) were joined by a node of a dendrogram, after which the joined elements were averaged and replaced by a pseudo-element to be used in all subsequent iterations. Hierarchically clustered gene expression and tumor data were analyzed graphically using the TreeView[33] program bundled with Cluster software.

Differential expression profiles were also clustered non-hierarchically using the Quality Threshold (QT) clustering algorithm as originally described.[34] The advantage of this non-agglomerative approach over hierarchical clustering was that all of the data were compared without the generation of pseudo-elements. In addition, QT clustering is not designed to separate the data into a predetermined, user-defined number of clusters, as would occur with other commonly used clustering algorithms such as self-organizing maps or K-means clustering. Instead, the user input for QT clustering is limited to the QT, which represents how highly correlated each of the members of a given cluster must be. User input QT can range from −1 (completely inversely correlated) to 1 (perfectly correlated). In general, the QT value needed for significance is inversely related to the number of elements to be clustered. For example, a QT of 0.1 to 0.2 may be adequate to cluster tumors reliably using 1000 genes, whereas a higher value of 0.3 to 0.4 may be required to cluster tumors using 100 genes. On the other hand, a QT of 0.6 to 0.7 is likely to be necessary to cluster genes using only 10 or fewer tumors. At QT values higher than 0.7, elements tend not to be clustered even though they are significantly correlated.

For all analyses using the hierarchical and QT algorithms, differential expression values were transformed to $\log_2$ before clustering, so that overexpressed and underexpressed genes would have values of opposite sign. In addition, for analyses using the QT algorithm, $\log_2$-transformed values were normalized so that every gene had a mean differential expression of 0 and a variance of 1 across the seven experiments.[34] Thus, the expression patterns of individual genes were relatively independent of absolute differential expression levels.

Immunohistochemistry

Representative formalin-fixed, paraffin-embedded tissue sections were dewaxed and subjected to antigen retrieval in citrate buffer, pH 6, using an electric pressure cooker set at 120°C for 5 minutes.[35] Sections were incubated for 5 minutes in 3% hydrogen peroxide to quench endogenous tissue peroxidase. Immunohistochemistry was performed using primary antibodies directed against vimentin (mouse monoclonal M0725, 1:80 dilution: DAKO Corp., Carpinteria, CA), CD74 (mouse monoclonal LN2, 1:8 dilution; ICN Biomedicals, Costa Mesa, CA), parvalbumin (goat polyclonal Sc7447, 1:40 dilution; Santa Cruz Biotechnology, Santa Cruz, CA), and galectin-3 (mouse monoclonal GALECT3abm, 1:100 dilution; Research Diagnostics, Inc., Flanders, NJ). After 25-minute incubations with appropriate primary antibody, sections were washed and treated with commercial biotinylated secondary anti-immunoglobulin, followed by avidin coupled to biotinylated horseradish peroxidase, according to manufacturer's instructions (LSAB2 kit for mouse primary antibodies and LSAB+ kit for goat primary antibody, DAKO). The immunohistochemical reactions were visualized using diaminobenzidine as the chromogenic peroxidase substrate. Sections were counterstained with hematoxylin after immunohistochemistry. Strong positive immunohistochemical staining was defined as ≥3+ intensity in at least 30% of tumor cells. Specificity of the procedure was verified by negative control reactions without primary antibody and by appropriate staining of positive control tissues.

Results

Gene expression was compared in seven matched pairs of renal epithelial tumors and non-neoplastic kidneys using Incyte UniGEM v.1 cDNA microarrays. The tumors consisted of four conventional RCCs (two low-grade and two high-grade lesions), one chromophobe RCC, and two oncocytomas. The microarrays featured 7075 unique cDNA targets representing 15 to 25% of the predicted human genome.[36] Approximately 3000 of the targets corresponded to expressed sequence tags of unknown function. The microarrays detected 4906 unique mRNA species, representing 69% of the array targets, in at least

Table 1. Genes Overexpressed in Conventional RCCs Compared to Chromophobe RCC/Oncocytomas

Gene name	GenBank no.	Locus
α-2 Macroglobulin	M11313	12p13.3-p12.3
Adipophilin	NM001122	9
Angiopoietin 2	NM001147	8p23.1
Caldesmon 1	M64110	7q33
Class II MHC-associated invariant chain (CD74)	M13560	5q32
Collagen IV-α1	M26576	13q34
Complement component 1qβ	AA953314	1p36.3-p34.1
Complement component 3	NM000064	19p13.3-p13.2
Cytochrome P450, subfamily IIJ polypeptide 2	U37143	1p31.3-p31.2
Delta sleep-inducing peptide	AL110191	X
EST (GenBank no. AA664156)	AA664156	3
EST (GenBank no. AI018324)	AI018324	Unmapped
Fc γ receptor IIIa (CD16)	J04162	1q23
Galectin 1	AA568129	22q13.1
HLA-B	U29057	6p21.3
HLA-DR α	AA360644	6p21.3
HLA-DR β 1	M32578	6p21.3
HLA-SB(DP) α	X03100	6p21.3
IFN-induced transmembrane protein 3 (1–8U)	X57352	Unmapped
IFN-induced transmembrane protein 1 (9–27)	J04164	11
Lysyl Oxidase	W77730	5q23.3-q31.2
Monocyte chemotactic protein 1	M26683	17q11.2-q21.1
Nidogen (Enactin)	NM002508	1q43
Serin protease 11 (IGF binding)	D87258	10q25.3-q26.2
TGF β-induced, 68 kD	AC004503	5q31
Tissue inhibitor of metalloproteinase 3	AI245471	22q12.3
Triosephosphate isomerase 1	X69723	12p13
Vascular cell adhesion Molecule 1	AL037831	1p32-p31
VEGF-related protein	AI004656	14q24-q31
Versican	X15998	5ql4.3
Vimentin	X56134	10p13
von Willebrand factor	NMOO0552	12p13.3

The microarray data for these genes, including absolute differential expression levels in each of the seven tumors analyzed, are presented as Supplemental Data at www.amjpathol.org. The 22 genes in bold type make up the second largest Quality Threshold cluster (Figure 2B).

two of the seven hybridization experiments (with expression detected in tumor, matched non-neoplastic kidney, or both). Initially, we defined significant differential expression between a tumor and its matched non-neoplastic kidney as ≥1.8-fold overexpression or underexpression, based on Incyte's internal quality control data (see Materials and Methods). By this definition, 8% of the 4906 detected genes (385 genes) were overexpressed or underexpressed in at least two tumors. Subsequently, we increased the differential expression cutoff to ≥2.0-fold to reduce the chance of false positive signals. At this higher stringency, 189 (4%) of the detected genes were differentially expressed in at least two tumors. Most of these 189 genes could be grouped into functional categories such as cell growth and differentiation, cell adhesion, immune regulation, energy metabolism, cytoskeleton, vascular biology, and extracellular matrix, whereas 28 sequences corresponded to uncharacterized expressed sequence tags. A complete listing of the 189 differentially expressed genes, including microarray fluorescent expression data and differential expression values in each tumor, is presented as Supplemental Data on our departmental Internet site www.amjpathol.org.Inspection of this microarray data revealed that several genes were expressed preferentially in specific renal tumor subtypes, as listed in Tables 1 and 2 and discussed further below.

We used a hierarchical average linkage clustering algorithm[33] to group the seven tumors and 189 differentially expressed genes in two dimensions based on similarities of expression patterns. In the first dimension, the tumors were clustered by similarity of their expression profiles of the 189 genes; in the second dimension, the genes were clustered by similarity of their expression patterns across the seven tumors. These clustering experiments separated the seven tumors into two main gene expression classes that correlated exactly with the histopathological categories of conventional (clear cell) RCCs and chromophobe RCC/oncocytomas (Figure 1; note the dendrogram on the *x* axis). The same clustering of tumors was obtained when we used a cutoff of 1.8-fold to define differential expression or when we eliminated the differential expression cutoff and included all 4906 detected genes in the analysis (data not shown). The chromophobe RCC and oncocytomas expressed similar patterns of genes, as shown by their being grouped together in these clustering experiments (Figure 1). Of particular note, the chromophobe RCC/oncocytomas overexpressed genes coding for β-defensin-1 (an antibacterial peptide) and parvalbumin (a calcium-binding protein), the products of which have been localized in previous studies[37,38] to mammalian distal nephron epithelium (Table 2 and Supplemental Data). In addition, the chromophobe RCC/oncocytomas overexpressed many sequences related to mitochondrial biogenesis and oxidative phosphorylation, a finding consistent with the large numbers of mitochondria per cell in these particular tu-

Table 2. Genes Overexpressed in Chromophobe RCC/Oncocytomas Compared to Conventional RCCs

Name	Genbank no.	Locus
3-Oxoacid CoA transferase	U62961	5p13
Adenine nucleotide translocator 2	AW160902	Xq24-q26
Adenine nucleotide translocator 3	J03592	Xp22.32 Yp
Aspartate aminotransferase 1	NM_002079	10q24.1-q25.1
Carbonic anhydrase II	NM_000067	8q22
Carbonic anhydrase XII	NM_001218	15q22
CDC-like kinase 2	AF023268	1q21
Chloride channel Kb	Z30644	1p36
Creatine/kinase 2, mitochondrial	J05401	5q13.3
Cytochrome c oxidase subunit IV	NM_001861	16q24.1
Cytochrome c oxidase subunit VIIb	Z14244	X
Defensin, β 1	U50931	8p23.2-p23.1
Desmoglein 2	NM_001943	18q12.1
Desmoplakin	J05211	6p24
Dihydrolipoamide dehydrogenase	U03620	7q31-q32
EGF domain-containing extracellular matrix protein 1	U03877	2p16
EGF receptor pathway substrate 8	AI679737	12q23-q24
Epidermal growth facto	NM_001963	4q25
EST (GenBank no. AA034414)	AA034414	12
EST (GenBank no. AA284067)	AA284067	Unmapped
EST (GenBank no. AA595322)	AA595322	15
EST (GenBank no. AI090186)	AI090186	3
EST (GenBank no. AI816358)	A1816358	7
EST (Chloride channel Kb homolog) (GenBank no. AI311009)	AI311009	1p36
EST/KIAA0439 (GenBank no. AB007899)	AB007899	18
EST/KIAA0517 (GenBank no. AB011089)	AB011089	4
Galectin 3	AB006780	14q21-q22
H+ Transporting ATPase (lysosomal vacuolar proton pump, β 1)	M25809	2cen-q13
H+ Transporting ATP synthase (mitochondrial F0 complex, C 3)	AA479643	2
H+ Transporting ATP synthase (mitochondrial F1 complex, α 1)	AW161540	18q12-q21
Hydroxysteroid (11-beta) dehydrogenase 2	U26726	16q22
Isocitrate dehydrogenase 2 (NADP+), mitochondrial	X69433	15q26.1
LIM/Senescent cell antigen-like domains 1	NM_004987	2
Mal, T-cell differentiation protein	NM_002371	2cen-q13
NADH-CoQ reductase	X61100	2q33-q34
NADH dehydrogenase (Ubiquinone) 1 α	AA813106	Xq24
Nicotinamide nucleotide transhydrogenase	Z5O101	5p13.1–5cen
Parvalbumin	AI022812	22q13.1
Plastin 3 (T isoform)	M22299	X
Propionyl CoA carboxylase α	NM_000282	13q32
Ribosomal protein L17	X53777	18q
Sialyltransferase 1	AF007133	3q27-q28
Single minded (*drosophila*) homologue 2	U80456	21q22.2
Syntaxin 3A	U32315	11
TB1, human	AA723646	5
Tetraspan 5	AF065389	4
Thioredoxin homologue	U60873	18
Ubiquinol-cytochrome c reductase hinge	AA568387	1

The microarray data for these genes, including absolute differential expression levels in each of the seven tumors analyzed, are presented as Supplemental Data at www.amjpathol.org. The 43 genes in boldface make up the largest Quality Threshold cluster (Figure 2A).

mors. Included among these sequences were genes encoding mitochondrial adenine nucleotide translocator isoforms, proton-transporting ATP synthase subunits, carbonic anhydrase isoenzymes, and cytochrome c oxidase subunits (Table 2 and Supplemental Data).

We next used QT clustering software[34] to recluster the gene expression data from the set of 189 differentially expressed genes, again in two dimensions. In the first dimension, the seven tumor samples were analyzed using a QT of 0.4, producing two clusters of tumor types identical to those identified with hierarchical clustering (data not shown). Thus, the renal tumor categories of conventional RCC and chromophobe RCC/oncocytoma could be distinguished by their patterns of gene expression using independent clustering algorithms to analyze the expression data. In the second dimension, the set of 189 genes was clustered using a QT of 0.6 to generate 19 individual clusters of similarly expressed genes. The largest cluster contained 43 genes that were overexpressed in chromophobe RCC/oncocytomas and underexpressed in conventional RCCs. The second largest cluster contained 22 genes that were overexpressed in conventional RCCs and underexpressed in chromophobe RCC/oncocytomas; interestingly, these 22 genes tended to be more strongly overexpressed in the two low-grade conventional RCCs (Fuhrman grade II) than in the two high-grade conventional tumors (Fuhrman grades III and IV). The means and standard deviations of the two gene clusters are shown in Figure 2, A and B. Overall, the two largest QT clusters produced a set of 65 genes that were

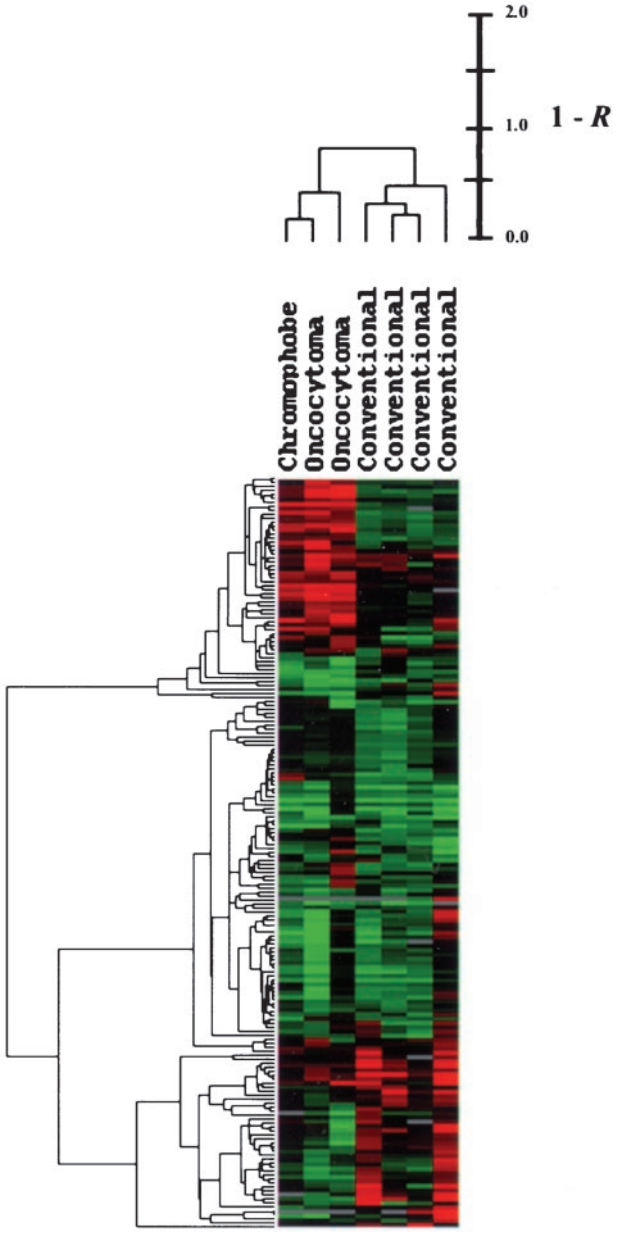

Figure 1. Hierarchical average linkage clustering of the seven renal tumors, based on expression patterns of the 189 differentially expressed genes. Expression patterns were analyzed using Cluster and TreeView software.[33] In the color-coded grid, individual tumors and genes have been listed in the *x*- and *y*-axis, respectively. Each grid block shows the differential expression value determined by microarray for a specific gene in a particular tumor (relative to the tumor's matched non-neoplastic kidney), with red indicating relative overexpression, green indicating relative underexpression, and black indicating similar expression levels in the tumor and non-neoplastic kidney. The brightness of red or green correlates with the degree of differential expression. The average linkage clustering algorithm grouped the tumors and genes by similarity of expression patterns and displayed the results in a dendrogram format. Items with similar expression profiles were placed into adjacent dendrogram nodes, connected by branches proportional in length to 1 minus the corresponding Pearson correlation coefficients (R). In the tumor dendrogram (*x* axis), the conventional RCCs and chromophobe RCC/oncocytomas were separated into two clusters represented by the two major dendrogram branches.

highly discriminatory between conventional RCCs and chromophobe RCC/oncocytomas. To demonstrate this point, the 65 genes were reanalyzed with the hierarchical average linkage clustering algorithm (shown graphically in Figure 2C). Comparison of the tumor dendrograms in Figures 1 and 2C, in which the relative height of each node is proportional to 1 minus the Pearson correlation coefficient *R*,[33] shows that analysis of the 65 genes identified by the QT algorithm produced a greater distinction between the two renal tumor categories than analysis of the total group of 189 differentially expressed genes. Thus, the subset of 65 differentially expressed genes may be particularly useful for classification of renal epithelial neoplasms, eg, by development of customized lower-density cDNA microarrays.

An important application of clustering algorithms for microarray data analysis is the identification of genes with highly correlated expression patterns across a series of experiments. Such coexpressed genes may be functionally related or may be regulated by common factors.[33,34] In our analysis, several genes identified by the average linkage and QT algorithms were grouped into functionally related clusters by virtue of similar expression profiles across the seven tumors (Figure 2C, *y* axis). For example, several class II MHC-related genes, including HLA-DRα, HLA-DRβ1, and class II MHC-associated invariant chain (CD74), were placed into adjacent nodes by the hierarchical clustering algorithm (Figure 2C, highlighted with a blue bar), based on strikingly similar expression patterns with relative overexpression in the conventional RCCs. Similarly, the genes encoding two interferon-induced transmembrane proteins were coclustered on the basis of similar differential expression profiles across the seven tumors, again with relative overexpression in the conventional RCCs (Figure 2C, highlighted with a black bar). In certain cases, the clustering of genes was predictive of function. An expressed sequence tag corresponding to GenBank accession number AI311009 was clustered into a highly correlated node with chloride channel Kb (Figure 2C, highlighted with an orange bar). Sequence comparison using the BLAST algorithm[39] showed that these two sequences are highly homologous.

The protein products of genes overexpressed in conventional RCCs or chromophobe RCC/oncocytomas, but not in both tumor categories, represented a set of potential immunohistochemical markers for renal tumor diagnosis. As described above, the QT algorithm revealed two large clusters of genes that were overexpressed in only one of the two tumor categories. The products of several of these genes, including vimentin, class II MHC-associated invariant chain (CD74), and galectin-3, have been identified previously in neoplastic and non-neoplastic renal tissue by immunohistochemistry, making these antigens candidate pathological markers.[31,40,41] Because of the small number of tumor samples analyzed by microarray, the variability introduced by a single outlier value for parvalbumin prevented this gene from being included in the largest QT cluster, even though it was overexpressed in chromophobe RCC/oncocytomas and underexpressed in conventional RCCs, similar to most members of that gene cluster (Table 2 and Supplemental Data). Parvalbumin has been localized to distal nephron epithelium by immunohistochemistry,[38] making it an interesting candidate marker for chromophobe RCC/oncocytomas, which appear to be antigenically related to

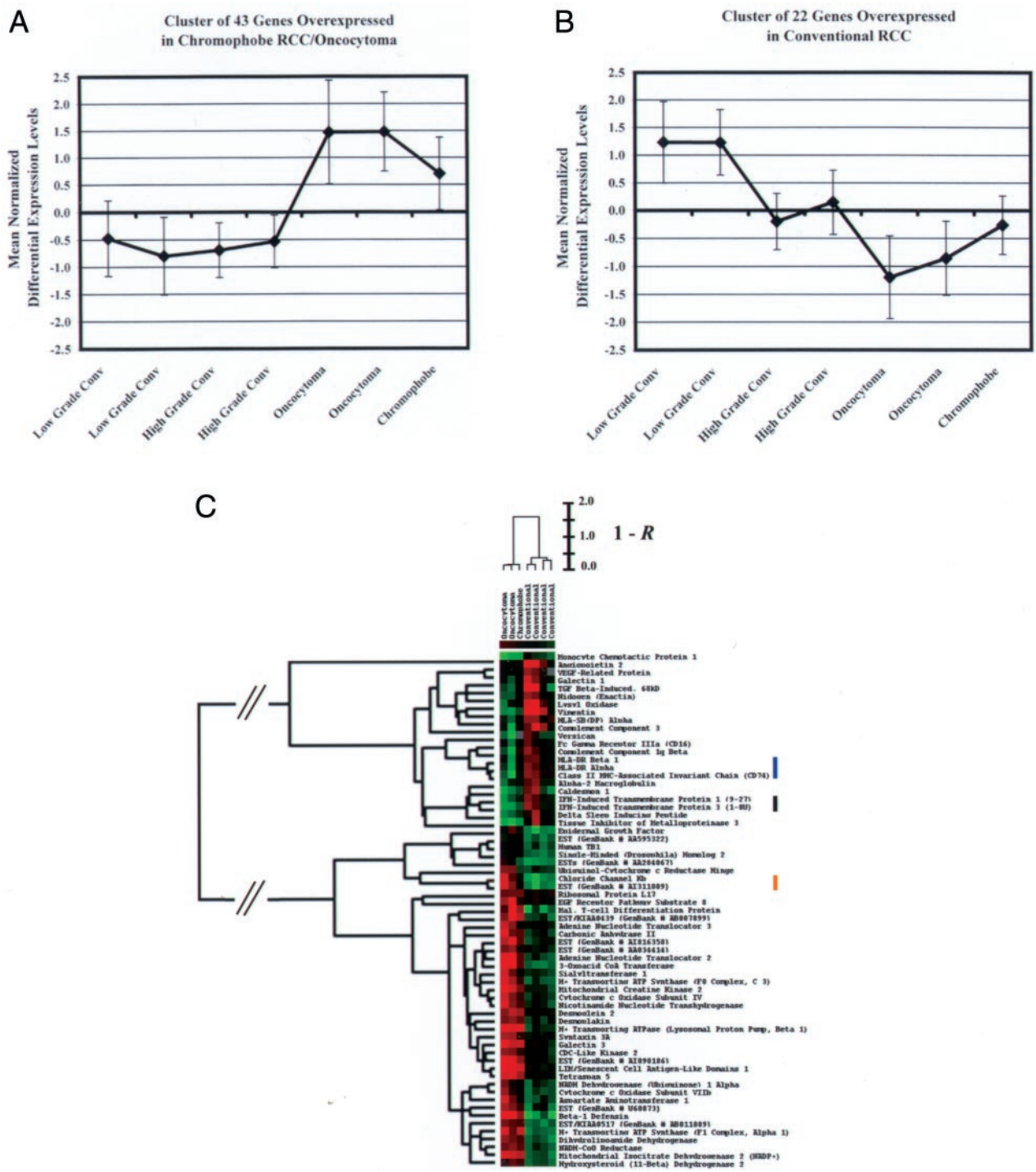

Figure 2. Quality Threshold (QT) clustering of the seven renal tumors and 189 differentially expressed genes in two dimensions.[34] In the first dimension, the tumors were analyzed with a QT of 0.4, producing two clusters corresponding to conventional RCCs and chromophobe RCC/oncocytomas (data not shown). In the second dimension, the genes were analyzed with a QT of 0.6, producing 19 clusters of genes with similar expression patterns in the renal tumors. The two largest gene clusters were used for the data in this figure. **A:** The largest QT cluster contained 43 genes, which were characterized by overexpression in the chromophobe RCC/oncocytomas and underexpression in the conventional RCCs. The graph shows the means of the normalized, $\log_2$-transformed, differential expression levels for these 43 genes in each of the seven tumors, with standard deviations indicated by the error bars. As seen from this graph, the 43 genes showed highly correlated expression patterns in the series of renal tumors. **B:** The second largest QT cluster contained 22 genes, which tended to be overexpressed in the conventional RCCs and underexpressed in the chromophobe RCC/oncocytomas. The graph shows the means of the normalized, $\log_2$-transformed, differential expression levels for these 22 genes in each of the seven tumors, with standard deviations indicated by the error bars. The mean overexpression of these 22 genes was greater in the two low-grade conventional RCCs (Fuhrman grade II) than in the two high-grade tumors (Fuhrman grades III and IV). **C:** Hierarchical average linkage clustering of the seven renal tumors, based on expression patterns of the 65 genes in the two largest QT clusters. In the color-coded grid, the low-grade conventional RCCs are represented in the fourth and fifth columns and the high-grade conventional RCCs are represented in the sixth and seventh columns. Class II MHC-related genes (blue bar), genes encoding interferon-induced transmembrane proteins (black bar) and genes encoding homologous chloride channels (orange bar) were grouped into functionally related clusters (*y*-axis).

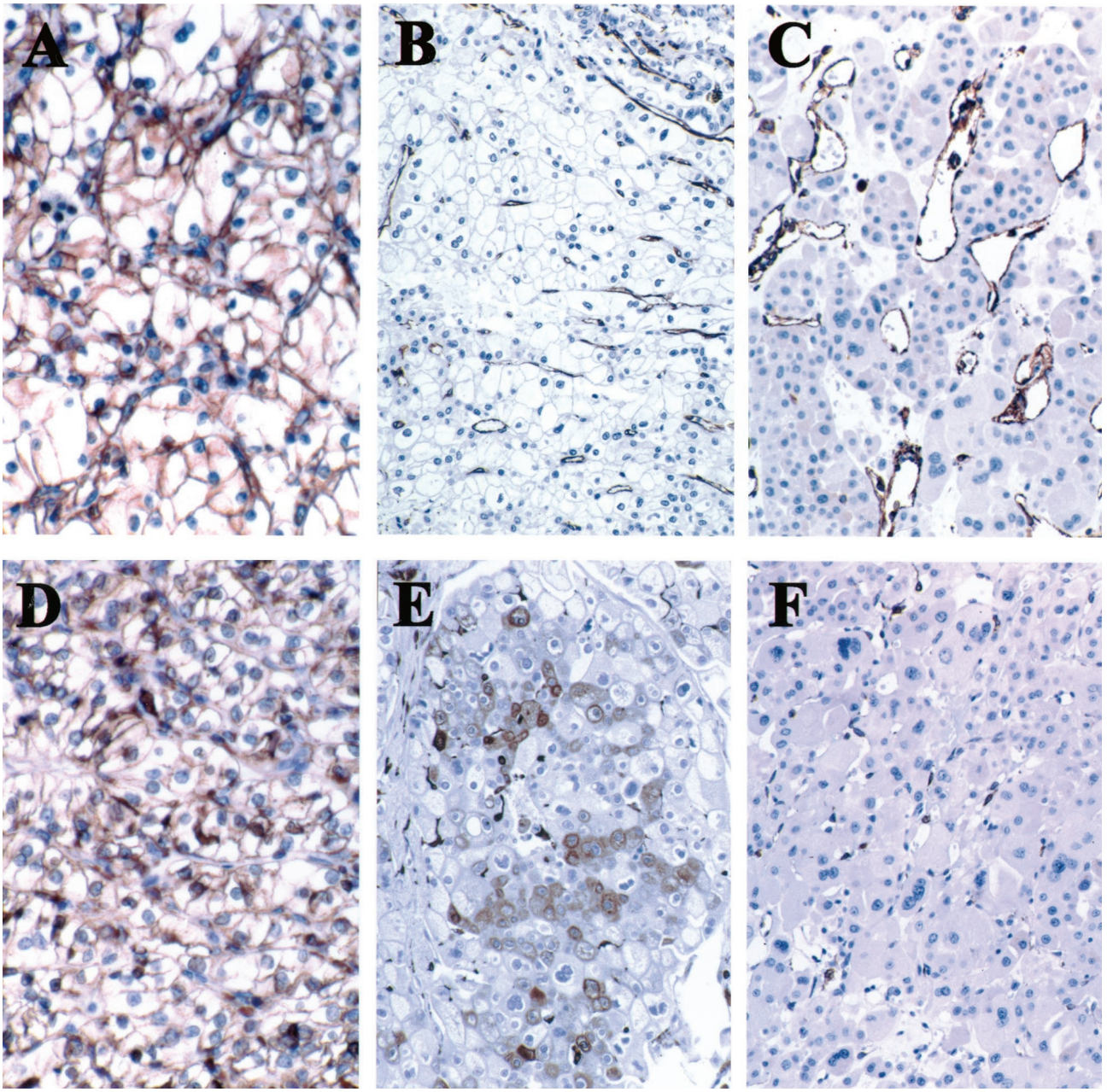

Figure 3. Immunoperoxidase staining of renal tumor subtypes for vimentin (**A–C**) and CD74 (**D–F**). Reactions were visualized with the brown peroxidase substrate diaminobenzidine and counterstained with hematoxylin. **A** and **D:** Conventional RCC. **B** and **E:** Chromophobe RCC. **C** and **F:** Oncocytoma. Positive cytoplasmic staining for vimentin distinguished conventional RCC from chromophobe RCC/oncocytoma. CD74 staining was seen in conventional and chromophobe RCC but not oncocytoma. For both antigens, cytoplasmic staining was observed in tumors cells as well as tumor-associated stromal and endothelial cells. Original magnifications, ×200.

distal nephron intercalated cells.[11–13] To evaluate the diagnostic utility of vimentin, CD74, parvalbumin, and galectin-3, as well as to test the validity of the microarray data, we measured the expression of these proteins in an independent series of 34 renal tumors by immunohistochemistry. In agreement with the microarray data, which showed vimentin mRNA overexpression to be specific to conventional tumors, vimentin protein was detected in the tumor cells of 17/20 (85%) conventional RCCs *versus* 0/6 chromophobe RCCs and 0/8 oncocytomas (Figure 3, A–C, and Table 3; $P \leq 0.001$). Also consistent with the microarray data, CD74 was detected immunohistochemically in the tumor cells of 13/20 (65%) conventional carcinomas *versus* 0/8 oncocytomas. Interestingly, 4/6 (67%) chromophobe RCCs stained positively for CD74, a finding not seen in the single chromophobe tumor analyzed by microarray. These data suggest that CD74 might be a useful marker in certain cases to discriminate chromophobe tumors from oncocytomas (Figure 3, D–F, and Table 3; $P \leq 0.01$). The immunohistochemical detection of parvalbumin correlated very closely with the microarray results, with strong tumor cell expression in 6/6 chromophobe RCCs and 8/8 oncocytomas *versus* only 5/20 (20%) conventional RCCs (Figure 4, A–C, and Table

Table 3. Summary of Immunohistochemical Data

	Fraction of cases with strong positive staining in tumor cells			
Marker	Conventional RCC	Chromophobe RCC	Oncocytoma	*P*
Vimentin	17/20	0/6	0/8	≤0.001
CD74	13/20	4/6	0/8	≤0.01
Parvalbumin	5/20	6/6	8/8	≤0.001
Galectin-3	13/20	5/6	8/8	>0.05*

*Galectin-3-positive conventional RCCs included 9/10 low-grade tumors *versus* 4/10 high-grade tumors ($P \leq 0.025$)

3; $P \leq 0.001$). In the tissue sections, non-neoplastic kidney showed high background staining, probably related to high endogenous peroxidase content, which prevented us from confirming if parvalbumin was indeed expressed selectively by distal nephron epithelium. Also in agreement with the microarray data, galectin-3 was consistently detected by immunohistochemistry in chromophobe RCCs (5/6 tumors strongly positive) and oncocytomas (8/8 tumors strongly positive). Data were discordant, however, for galectin-3 expression in conventional RCCs. Whereas 0/4 conventional tumors overexpressed galectin-3 mRNA by microarray, 13/20 (65%) expressed the protein at high levels by immunohistochemistry (Figure 4, D–F, and Table 3). Interestingly, immunohistochemical detection of galectin-3 in conventional tumors was correlated with low histological grade (9/10 low-grade tumors *versus* 4/10 high-grade tumors; $P \leq 0.025$).

Discussion

Our first objective in these experiments was to demonstrate the feasibility of gene expression profiling for the classification of renal epithelial neoplasms. Recent studies have shown that various tumors can be classified by expression profiling in clinically and biologically meaningful ways. However, most of these studies have used a large number of samples to classify tumors with subtle phenotypic differences.[27–29] In our study, we were faced with a limited number of frozen tumor samples suitable for mRNA analysis. Therefore, we confined these initial studies to tumor types with very distinct features, to increase the likelihood of resolving significant gene expression clusters using relatively few samples. We focused the experiments on conventional (clear cell) RCC, chromophobe RCC, and oncocytoma, which are related to different renal epithelial cell types and exhibit markedly different cytogenetics and clinical behavior.[2] In this preliminary study, we were indeed able to identify specific, reproducible patterns of gene expression that correlated very strongly with the major histopathological classifications of conventional RCC and chromophobe RCC/oncocytoma, using multiple independent algorithms to analyze the microarray data.[33,34] Several aspects of this expression data provided novel insights into the histogenesis and pathobiology of renal tumors. For example, the chromophobe RCC and oncocytomas were characterized by strikingly similar gene expression profiles, providing compelling evidence that they are related tumors. Specifically, the chromophobe RCC/oncocytomas overexpressed the distal nephron markers parvalbumin and β-defensin-1,[37–38] suggesting that distal nephron-like patterns of gene expression may be characteristic of these tumor types. This finding supports current histogenetic models that relate these tumors to distal nephron intercalated cells.[11–13] The chromophobe RCC/oncocytomas also overexpressed a large cluster of genes related to mitochondrial biology, consistent with the mitochondrion-rich morphology of these tumors. Given this feature, it is perhaps somewhat surprising that the chromophobe RCC/oncocytomas did not overexpress several genes related to mitochondrial biology that were represented on the cDNA microarrays (Supplemental Data). Possibly, the regulation of these particular gene products occurred post-translationally, making differential expression transparent to the microarray analysis. The conventional RCCs underexpressed epidermal growth factor, whereas the chromophobe RCC/oncocytomas overexpressed epidermal growth factor receptor pathway substrate 8 (Figure 2C). Thus, certain aspects of epidermal growth factor signaling may be important for the biology of distal nephron epithelium and/or distal nephron-related epithelial tumors. The conventional RCCs overexpressed several class II MHC-related genes, as discussed in greater detail below. In addition, these tumors were characterized by specific overexpression of several genes related to vascular biology, including vascular cell adhesion molecule 1, vascular endothelial growth factor-related protein, von Willebrand factor, and angiopoietin 2 (Table 1 and Supplemental Data). This finding may be related to the anastomosing vascular network that distinguishes conventional RCC from other renal tumors[2,42] and may have implications for anti-angiogenic therapies directed against this tumor subtype.[10] Finally, the conventional RCCs overexpressed at least three genes on chromosome 5q (CD74, transforming growth factor-β, and lysyl oxidase; Table 1 and Supplemental Data), consistent with previous comparative genomic hybridization data showing that DNA gains in this tumor type most often involve chromosome 5q.[43]

Examination of Figure 1 (*y* axis) shows that several genes were either overexpressed or underexpressed in the all of the tumor subtypes examined in this study. This gene group included ribosomal proteins, HLA-B and insulin-like growth factor binding protein 3 (overexpressed in conventional RCCs, chromophobe RCC, and oncocytomas), and c-*fos*, cathepsin H, and uromodulin (underexpressed in conventional RCCs, chromophobe RCC, and oncocytomas). Specific changes in the expression of these genes may be common in the neoplastic transformation of renal epithelial cells. Indeed, each of these

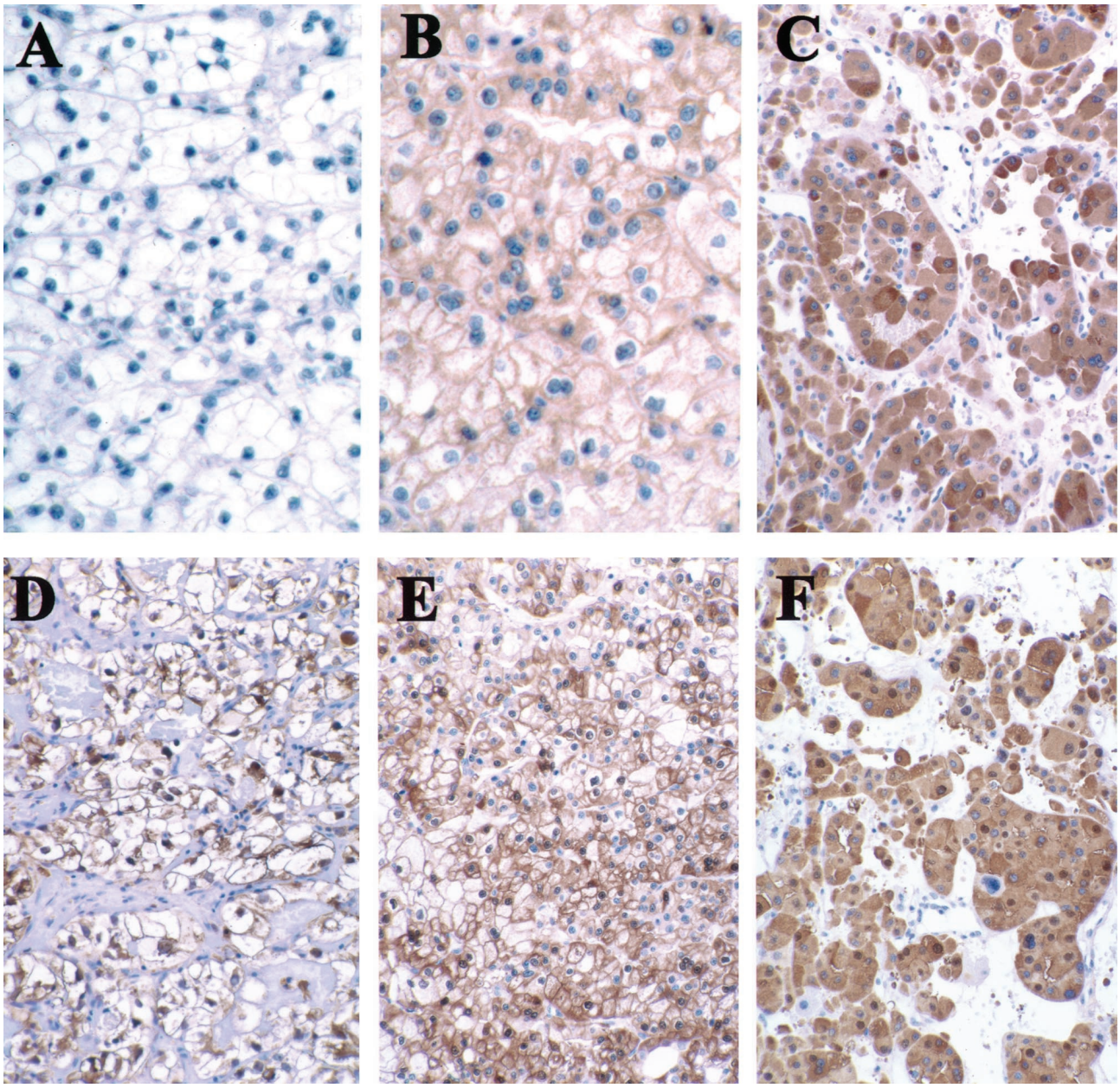

Figure 4. Immunoperoxidase staining of renal tumor subtypes for parvalbumin (**A–C**) and galectin-3 (**D–F**). Reactions were visualized with the brown peroxidase substrate diaminobenzidine and counterstained with hematoxylin. **A** and **D:** Conventional RCC. **B** and **E:** Chromophobe RCC. **C** and **F:** Oncocytoma. Positive cytoplasmic staining for parvalbumin distinguished chromophobe RCC/oncocytoma from conventional RCC. Cytoplasmic galectin-3 staining was seen in low-grade conventional RCC, chromophobe RCC, and oncocytoma. Original magnifications, ×200.

gene products has been associated with processes relevant to tumor development and spread.[40,44–48]

Because our microarray study was based on a limited number of grossly dissected renal tumor samples consisting of heterogeneous cell populations, two obvious concerns were (i) whether the microarray data reflected differential gene expression specific to tumor cells and (ii) if so, whether the data could be generalized to conventional RCCs and chromophobe RCC/oncocytomas as a whole. To try to address these concerns, we validated microarray data for several genes with the cell-specific technique of immunohistochemistry in a larger tumor series. For each of the antigens tested—vimentin, CD74, parvalbumin, and galectin-3—the immunohistochemical signals were detected mainly in neoplastic epithelium, suggesting that the microarray data primarily reflected differential gene expression in tumor cells. Vimentin and CD74 antigen expression was also detected in tumor-associated stroma and vasculature, although these non-neoplastic cells typically comprised a small portion of the total renal tumor cellularity.

Overall, we obtained strong correlations between microarrays and immunohistochemistry for differential gene/protein expression in renal tumor subtypes, although the correlations were not perfect. Most published expression-profiling studies have described good but

imperfect correlations between the data from microarrays and gene-specific validation assays.[49] This inability to confirm every aspect of microarray data is not surprising, given the number of genes interrogated by typical high-density microarrays and the fact that most array experiments (including ours) have been unable to perform extensive data replication due to limitations in sample size. A recent, comprehensive statistical analysis of microarray experiments has shown that non-replicated expression data are prone to numerous misclassifications, particularly false positive results.[50] Thus, appropriate data validation, either by microarray replication or an independent technique such as immunohistochemistry, is clearly essential for accurate interpretation of gene expression profiles.

The cDNA microarray and immunohistochemical experiments in our study both showed vimentin to be a sensitive and specific marker for conventional RCC. In agreement with our results, an independent microarray analysis that compared gene expression between a renal cancer cell line and benign kidney tissue also singled out vimentin as a useful diagnostic and prognostic marker for RCC.[31] We also obtained excellent concordance between microarray and immunohistochemical data for the calcium-binding protein parvalbumin, which emerged as a promising marker for chromophobe RCC/oncocytomas. Thus, our study supports the idea that expression profiling of tumors is a potentially powerful method for identifying new pathological markers for tumor diagnosis.[31,32]

Our data validations were less exact for galectin-3 and CD74. For example, although the microarrays suggested that galectin-3 mRNA overexpression was specific to chromophobe RCC/oncocytomas, the corresponding protein was detected by immunohistochemistry in several conventional tumors as well. It is possible that galectin-3 mRNA and protein levels do not correlate precisely in certain tissues. Although our findings may exclude galectin-3 as a useful marker for differentiating renal tumor subtypes, the expression of this gene product appeared to correlate with tumor indolence, being expressed predominantly in low-grade conventional RCCs, indolent chromophobe RCCs, and benign oncocytomas. This finding is consistent with several studies showing reduced galectin-3 expression in clinically aggressive tumors and may be relevant to the function of galectin-3 as an adhesion molecule that inhibits metastasis.[41,51–53] Interestingly, our microarray data suggested that the related adhesion molecule galectin-1 was also expressed differentially in the renal tumor subtypes, albeit with relative overexpression in conventional tumors (Table 1 and Supplemental Data). Though we did not confirm galectin-1 data immunohistochemically, other studies have shown that expression of this lectin may be prognostically or diagnostically relevant to tumor biology, either alone or in combination with galectin-3.[41]

Taken together, the microarray and immunohistochemical experiments suggested that class II MHC-associated invariant chain (CD74) was expressed in conventional and chromophobe RCCs but not in oncocytomas. Differential CD74 expression in chromophobe RCCs and oncocytomas is a finding of potential importance for surgical pathology, since reliable immunomarkers are not available to distinguish these histologically related lesions and the benign nature of oncocytomas, compared with the potential of chromophobe RCCs for metastasis and sarcomatoid transformation,[54,55] makes this an important differential diagnosis. Intriguing therapeutic implications are also raised by the expression of class II MHC-related genes in RCCs, given the responsiveness of many cases to interleukin-2 and/or interferon-α.[56] Our results are in general agreement with a recent report showing class II MHC-associated invariant chain expression by immunohistochemistry in 53/60 renal carcinomas.[40] In that study, CD74 expression was correlated with lymphocytic infiltration of tumor, and the authors speculated that class II MHC-related gene expression might be relevant to the overall responsiveness of RCC to immunotherapy. Future microarray studies of primary and metastatic RCC should help determine whether overall expression patterns of class II MHC-related molecules, including CD74, are predictive of immunotherapeutic response.

At the time this project began, Incyte UniGEM v.1 microarrays were the most complete cDNA arrays available for gene expression profiling. Their cDNA targets were chosen from expression libraries of all major organ groups and represented genes involved in many biological activities such as cell growth and development, cytoskeletal structure, cell motility, molecular recognition, membrane transport, protein and nucleic acid biosynthesis, and energy metabolism. Despite this, the UniGEM microarrays interrogated a small fraction of the total expected genome.[36] Denser arrays are becoming available, both commercially and from individual laboratories, as genome projects and microarray fabrication technologies continue to progress. Thus, expression profiling of the entire genome is likely to be possible in the near future. However, until this becomes a reality, large-scale expression profiling studies will suffer from the somewhat ironic problem of enormous, yet incomplete, data sets. Given these limitations, we expect that we have identified only a fraction of the genes that are relevant to the pathobiology of renal epithelial neoplasms, which might explain why relatively few growth regulatory genes were overexpressed or underexpressed consistently in the tumor subtypes we studied. Even though the UniGEM arrays contained more than 600 genes related to cell growth and development, this still did not represent a complete survey. Notably, for example, the von Hippel-Lindau tumor suppressor gene was not included on the microarrays.

In conclusion, we studied the gene expression profiles of conventional RCCs, chromophobe RCC, and oncocytomas and separated the tumors into reproducible gene expression classes that correlated with histopathological diagnoses. Several functionally related gene clusters were informative for distinguishing conventional RCCs from chromophobe RCC/oncocytomas (eg, class II MHC-related and vascular genes in conventional RCCs; distal nephron and oxidative phosphorylation genes in chromophobe RCC/oncocytomas). These gene clusters offered insights into tumor pathobiology and histogenesis, and highlighted several possible gene regulatory networks

that could be important in renal epithelial neoplasia. Several individual genes showed potential utility as immunohistochemical markers for the pathological diagnosis of renal tumor subtypes. Based on these results, we are encouraged to expand our expression profiling studies, using a larger number of specimens that include not only conventional RCCs and chromophobe RCC/oncocytomas, but also other renal tumor subtypes, such as papillary RCCs. We predict that these gene expression profiling experiments will lead to improvements in the basic understanding of renal tumor pathogenesis and will promote the discovery of novel molecular markers for renal tumor diagnosis.

References

1. Landis SH, Murray T, Bolden S, Wingo PA: Cancer statistics, 1999. CA Cancer J Clin 1999, 49:8–31
2. Zambrano NR, Lubensky IA, Merino MJ, Linehan WM, Walther MM: Histopathology and molecular genetics of renal tumors toward unification of a classification system. J Urol 1999, 162:1246–1258
3. Wallace AC, Nairn RC: Renal tubular antigens in kidney tumors. Cancer 1972, 29:977–981
4. Poston CD, Jaffe GS, Lubensky IA, Solomon D, Zbar B, Linehan WM, Walther MM: Characterization of the renal pathology of a familial form of renal cell carcinoma associated with von Hippel-Lindau disease: clinical and molecular genetic implications. J Urol 1995, 153:22–26
5. Gnarra JR, Tory K, Weng Y, Schmidt L, Wei MH, Li H, Latif F, Liu S, Chen F, Duh FM: Mutations of the VHL tumour suppressor gene in renal carcinoma. Nat Genet 1994, 7:85–90
6. Lott ST, Lovell M, Naylor SL, Killary AM: Physical and functional mapping of a tumor suppressor locus for renal cell carcinoma within chromosome 3p12. Cancer Res 1998, 58:3533–3537
7. van den Berg A, Buys CHCM: Involvement of multiple loci on chromosome 3 in renal cell cancer development. Genes Chromosomes Cancer 1997, 19:59–76
8. Oda H, Nakatsuru Y, Ishikawa T: Mutations of the p53 gene and p53 protein overexpression are associated with sarcomatoid transformation in renal cell carcinomas. Cancer Res 1995, 55:658–662
9. Wu S, Hafez GR, Xing W, Newton M, Chen X, Messing E: The correlation between the loss of chromosome 14q with histologic tumor grade, pathologic stage, and outcome of patients with nonpapillary renal cell carcinoma. Cancer 1996, 77:1154–1160
10. Berg WJ, Divgi CR, Nanus DM, Motzer RJ: Novel investigative approaches for advanced renal cell carcinoma. Semin Oncol 2000, 27:234–239
11. Ortmann M, Vierbuchen M, Fischer R: Renal oncocytoma. II. Lectin and immunohistochemical features indicating an origin from the collecting duct. Virchows Arch B (Cell Pathol incl Mol Pathol) 1988, 56:175–184
12. Ortmann M, Vierbuchen M, Fischer R: Sialylated glycoconjugates in chromophobe cell renal carcinoma compared with other renal cell tumors. Indication of its development from the collecting duct epithelium. Virchows Arch B (Cell Pathol incl Mol Pathol) 1991, 61:123–132
13. Storkel S, Steart PV, Drenckhahn D, Thoenes W: The human chromophobe cell renal carcinoma: its probable relation to intercalated cells of the collecting duct. Virchows Arch B (Cell Pathol incl Mol Pathol) 1989, 56:237–245
14. Amin MB, Crotty TB, Tickoo SK, Farrow GM: Renal oncocytoma: a reappraisal of morphologic features with clinicopathologic findings in 80 cases. Am J Surg Pathol 1997, 21:1–12
15. Thoenes W, Storkel S, Rumpelt H-J, Moll R, Baum HP, Werner S: Chromophobe cell renal carcinoma and its variants: a report on 32 cases. J Pathol 1988, 155:277–287
16. Durham JR, Keohane M, Amin MB: Chromophobe renal cell carcinoma. Adv Anat Pathol 1996, 3:336–342
17. Tickoo SK, Lee MW, Eble JN, Amin M, Christopherson T, Zarbo RJ, Amin MB: Ultrastructural observations on mitochondria and microvesicles in renal oncocytoma, chromophobe renal cell carcinoma, and eosinophilic variant of conventional (clear cell) renal cell carcinoma. Am J Surg Pathol 2000, 24:1247–1256
18. Tickoo SK, Reuter VE, Amin MB, Srigley JR, Epstein JI, Min K-W, Rubin MA, Ro JY: Renal oncocytosis: a morphologic study of fourteen cases. Am J Surg Pathol 1999, 23:1094–1101
19. Dijkhuizen T, van den Berg E, Storkel S, de Vries B, van der Veen AY, Wilbrink M, Geurts van Kessel A, de Jong B: Renal oncocytoma with t(5;12;11), der(1)1;8) and add(19): "true" oncocytoma or chromophobe adenoma? Int J Cancer 1997, 73:521–524
20. Crotty TB, Lawrence KM, Moertel CA, Bartelt DH Jr, Batts KP, Dewald GW, Farrow GM, Jenkins RB: Cytogenetic analysis of six renal oncocytomas and a chromophobe cell renal carcinoma. Evidence that -Y, -1 may be a characteristic anomaly in renal oncocytomas. Cancer Genet Cytogenet 1992, 61:61–66
21. Bugert P, Gaul C, Weber K, Herbers J, Akhtar M, Ljungberg B, Kovacs G: Specific genetic changes of diagnostic importance in chromophobe renal cell carcinomas. Lab Invest 1997, 76:203–208
22. Brown JA, Takahashi S, Alcaraz A, Borell TJ, Anderl KL, Qian J, Persons DL, Bostwick DG, Lieber MM, Jenkins RB: Fluorescence in situ hybridization analysis of renal oncocytoma reveals frequent loss of chromosomes Y and 1. J Urol 1996, 156:31–35
23. Kovacs A, Kovacs G: Low chromosome number in chromophobe renal cell carcinomas. Genes Chromosomes Cancer 1992, 4:267–268
24. Fuzesi L, Gunawan B, Braun S, Bergmann F, Brauers A, Effert P, Mittermayer C: Cytogenetic analysis of 11 renal oncocytomas: further evidence of structural rearrangements of 11q13 as a characteristic chromosomal anomaly. Cancer Genet Cytogenet 1998, 107:1–6
25. Kovacs A, Storkel S, Thoenes W, Kovacs G: Mitochondrial and chromosomal DNA alterations in human chromophobe renal cell carcinomas. J Pathol 1992, 167:273–277
26. Eisen MB, Brown PO: DNA arrays for analysis of gene expression. Methods Enzymol 1999, 303:179–205
27. Bittner M, Meltzer P, Chen Y, Jiang Y, Seftor E, Hendrix M, Radmacher M, Simon R, Yakhini Z, Ben-Dor A, Sampas N, Dougherty E, Wang E, Marincola F, Gooden C, Lueders J, Glatfelter A, Pollock P, Carpten J, Gillanders E, Leja D, Dietrich K, Beaudry C, Berens M, Alberts D, Sondak V, Hayward N, Trent J: Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature 2000, 406:536–540
28. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM: Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 2000, 403:503–511
29. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999, 286: 531–537
30. Heller RA, Schena M, Chai A, Shalon D, Bedilion T, Gilmore J, Woolley DE, Davis RW: Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. Proc Natl Acad Sci USA 1997, 94:2150–2155
31. Moch H, Schraml P, Bubendorf L, Mirlacher M, Kononen J, Gasser T, Mihatsch MJ, Kallioniemi OP, Sauter G: High-throughput tissue microarray analysis to evaluate genes uncovered by cDNA microarray screening in renal cell carcinoma. Am J Pathol 1999, 154:981–986
32. Shibata D: Pattern recognition and microarrays: the times are a-changing. Am J Pathol 1999, 154:979–980
33. Eisen MB, Spellman PT, Brown PO, Botstein D: Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 1998, 95:14863–14868
34. Heyer LJ, Kruglyak S, Yooseph S: Exploring expression data: identification and analysis of coexpressed genes. Genome Res 1999, 9:1106–1115
35. Norton AJ, Jordan S, Yeomans P: Brief, high-temperature heat denaturation (pressure cooking): a simple and effective method of antigen retrieval for routinely processed tissues. J Pathol 1994, 173:371–379
36. International Human Genome Sequencing Consortium: Initial sequencing and analysis of the human genome. Nature 2001, 409:806–921

37. Valore EV, Park CH, Quayle AJ, Wiles KR, McCray PB Jr, Ganz T: Human beta-defensin-1: an antimicrobial peptide of urogenital tissues. J Clin Invest 1998, 101:1633–1642
38. Bindels RJ, Timmermans JA, Hartog A, Coers W, van Os CH Calbindin-D9k and parvalbumin are exclusively located along basolateral membranes in rat distal nephron. J Am Soc Nephrol 1991, 2:1122–1129
39. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. J Mol Biol 1990, 215:403–410
40. Saito T, Kimura M, Kawasaki T, Sato S, Tomita Y: MHC class II antigen-associated invariant chain on renal cell cancer may contribute to the anti-tumor immune response of the host. Cancer Lett 1997, 115:121–127
41. Francois C, van Velthoven R, De Lathouwer O, Moreno C, Peltier A, Kaltner H, Salmon I, Gabius HJ, Danguy A, Decaestecker C, Kiss R: Galectin-1 and galectin-3 binding pattern expression in renal cell carcinomas. Am J Clin Pathol 1999, 112:194–203
42. Reuter VE: Renal tumors exhibiting granular cytoplasm. Semin Diagn Pathol 1999, 16:135–145
43. Moch H, Presti JC Jr, Sauter G, Buchholz N, Jordan P, Mihatsch MJ, Waldman FM: Genetic aberrations detected by comparative genomic hybridization are associated with clinical outcome in renal cell carcinoma. Cancer Res 1996, 56:27–30
44. Tanaka T, Kondo S, Iwasa Y, Hiai H, Toyokuni S: Expression of stress-response and cell proliferation genes in renal cell carcinoma induced by oxidative stress. Am J Pathol 2000, 156:2149–2157
45. Zumkeller W, Schofield PN: The role of insulin-like growth factors and IGF-binding proteins in the physiological and pathological processes of the kidney. Virchows Arch B (Cell Pathol incl Mol Pathol) 1992, 62:207–220
46. Liebermann DA, Gregory B, Hoffman B: AP-1 (Fos/Jun) transcription factors in hematopoietic differentiation and apoptosis. Int J Oncol 1998, 12:685–700
47. Castren JP, Kamel DE, Nurmi MJ, Collan YU: Cathepsin H expression distinguishes oncocytomas from renal cell carcinomas. Anticancer Res 2000, 20:537–540
48. Virlon B, Cheval L, Buhler JM, Billon E, Doucet A, Elalouf JM: Serial microanalysis of renal transcriptomes. Proc Natl Acad Sci USA 1999, 96:15286–15291
49. Eckmann L, Smith JR, Housley MP, Dwinell MB, Kagnoff MF: Analysis by high density cDNA arrays of altered gene expression in human intestinal epithelial cells in response to infection with the invasive enteric bacteria Salmonella. J Biol Chem 2000, 275:14084–14094
50. Lee M-LT, Kuo FC, Whitmore GA, Sklar J: Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. Proc Natl Acad Sci USA 2000, 97:9834–9839
51. Idikio H: Galectin-3 expression in human breast carcinoma: correlation with cancer histologic grade. Int J Oncol 1998, 12:1287–1290
52. Pacis RA, Pilat MJ, Pienta KJ, Wojno K, Raz A, Hogan V, Cooper CR: Decreased galectin-3 expression in prostate cancer. Prostate 2000, 44:118–123
53. Iurisci I, Tinari N, Natoli C, Angelucci D, Cianchetti E, Iacobelli S: Concentrations of galectin-3 in the sera of normal controls and cancer patients. Clin Cancer Res 2000, 6:1389–1393
54. Renshaw AA, Henske EP, Loughlin KR, Shapiro C, Weinberg DS: Aggressive variants of chromophobe renal cell carcinoma. Cancer 1996, 78:1756–1761
55. Akhtar M, Tulbah A, Kardar AH, Ali MA: Sarcomatoid renal cell carcinoma: the chromophobe connection. Am J Surg Pathol 1997, 21:1188–1195
56. Bukowski RM: Cytokine combinations: therapeutic use in patients with advanced renal cell carcinoma. Semin Oncol 2000, 27:204–212