Summary. Kidney neoplasms are classified by light microscopy using the World Health Organization (WHO) system. The WHO system defines histopathologic tumor subtypes with distinct clinical behavior and underlying genetic mutations. In adults, the common malignant subtypes are variants of renal cell carcinoma (RCC). Histopathologic classification is critical for clinical management of RCC, but is becoming more complex with recognition of novel tumor subtypes, development of procedures yielding small diagnostic biopsies, and emergence of molecular therapies directed at tumor gene activity. Therefore, classification systems based on gene expression are likely to become essential for diagnosis, prognosis and treatment of kidney tumors. Recent DNA microarray studies have shown that clinically relevant renal tumor subtypes are characterized by distinct gene expression profiles, which are useful for discovery of novel diagnostic and prognostic biomarkers. In this review, we summarize the WHO classification system for renal tumors, general applications of microarray technology in cancer research, and specific microarray studies that have advanced knowledge of renal tumor diagnosis, prognosis, therapy and pathobiology.

Key words: Kidney Neoplasms, Gene Expression Profiling, Microarrays

Introduction

Malignant kidney tumors are increasing in incidence, accounting for 2-3% of human cancers today. In adults, renal cell carcinomas (RCC) are the most common kidney malignancies (Jemal et al., 2004). Most renal carcinomas arise sporadically, although cases occur in several hereditary cancer syndromes. In developed societies, major risk factors include cigarette smoking, obesity and hypertension (Parkin et al., 1994; Chow et al., 2000). Localized RCC can be cured by surgery (Hommra et al., 1995). However, 25-40% of cases present with extrarenal growth or metastases (Amin et al., 2002), and one-third of apparently localized lesions develop metastases postoperatively (Gieseg et al., 2002; Zisman et al., 2002). Advanced RCC responds poorly to systemic therapy and has a 5-year survival rate of less than 10% (Zisman et al., 2001). The ability to predict outcome after surgical or systemic therapy is limited, underscoring the need for new approaches to clinical management of these malignancies.

Renal neoplasms are classified into specific subtypes using the World Health Organization (WHO) histopathologic system. WHO classification is an important predictor of clinical behavior (Moch et al., 2000; Amin et al., 2002; Eble et al., 2004). In addition, tumor variables (Fuhrman nuclear grade and TNM stage) and clinical variables (including performance status and perioperative thrombocytosis) are used as prognostic factors (Motzer et al., 1999; Kattan et al., 2001; Amin et al., 2002; O’Keefe et al., 2002). Using WHO terminology, over 90% of clinically significant adult lesions are diagnosed as a common subtype of renal epithelial tumor: clear cell RCC (75% of surgically removed renal tumors), papillary RCC (10%), chromophobe RCC (5%) and renal oncocytoma (5%). Angiomyolipoma is the most common adult mesenchymal tumor (1%). The WHO system also includes newly recognized, rare forms of renal

malignancy, which are defined by clinical, morphologic and genetic factors: examples include carcinoma of the collecting ducts of Bellini, renal medullary carcinoma, renal carcinoma associated with Xp11.2 translocations, and mucinous tubular spindle cell carcinoma (Eble et al., 2004). Nephroblastoma (Wilms tumor) is the most common pediatric renal malignancy (Breslow et al., 1993).

Each renal tumor subtype in the WHO system is defined by specific pathologic criteria, which emphasize growth pattern, nuclear morphology and cytoplasmic features (Amin et al., 2002; Eble et al., 2004). Typical examples of adult epithelial tumors are characterized as follows: Clear cell RCC is a solitary infiltrating mass, with solid, alveolar or acinar growth patterns, and “clear” tumor cells associated with anastomosing blood vessels. Papillary RCC is a circumscribed mass with fibrous capsule, containing neoplastic cells in papillary growth pattern, admixed with foam cells and necrosis. Chromophobe RCC exhibits an alveolar or nested growth pattern, with tumor cells containing irregular nuclei, perinuclear halos, and clear or granular eosinophilic cytoplasm. Oncocytoma is a circumscribed mass (often with central scar), with nested or tubulocystic growth of “oncocyes” containing round nuclei and granular eosinophilic cytoplasm. In spite of these gross and microscopic criteria, diagnosis is difficult and subjective because many tumors are highly variegated histopathologically. For example, adult renal epithelial tumors of any subtype may exhibit solid, alveolar or papillary growth patterns, contain neoplastic cells with clear or granular cytoplasm, or display high-grade, sarcomatoid (spindle cell) histology.

**Clinical behavior and management**

Clear cell RCC has the highest rate of local invasion, metastasis and mortality of adult renal tumors (Moch et al., 2000; Amin et al., 2002). Papillary and chromophobe carcinomas are relatively indolent, but may metastasize or transform to high-grade sarcomatoid malignancies. In addition, papillary RCC has a higher rate of multifocality and association with end-stage renal disease than other RCC subtypes (Takahashi et al., 1993; Amin et al., 2002). Oncocytoma is closely related to chromophobe carcinoma (Tickoo et al., 2000), but is consistently benign (Amin et al., 2002). Total nephrectomy is the standard therapy for localized RCC. However, small localized tumors (< 4 cm) are diagnosed increasingly by routine abdominal studies (Homma et al., 1995), spurring development of nephron-sparing tumor resections and ablations, coupled with less-traumatic laparoscopic surgery (Mabjeesh et al., 2004; Novick, 2004). Based on different rates of extrarenal growth or multifocality, the use of laparoscopic nephron-sparing procedures may depend on prior histopathologic subtyping of tumor biopsies (Uzzo and Novick, 2001). For advanced-stage RCC, first-line systemic therapy involves interleukin-2 and interferon-alpha. These immunotherapies are more effective for clear cell RCC than other subtypes (Motzer et al., 2002); however, most clear cell tumors fail to respond markedly (Atkins et al., 2004). Anti-angiogenic agents are under investigation for clear cell RCC, targeting tumor vascularity (Gordon, 2004; Yang, 2004). Systemic therapy is not standardized for RCC with non-clear cell histology (Stadler, 2004).

**Genetics**

Renal tumor subtypes in the WHO system are associated with distinct, reoccurring cytogenetic lesions (Meloni-Ehrig, 2002) and hereditary cancer syndromes (Takahashi et al., 2002), underscoring the significance of tumor classification and the potential role of molecular diagnostics in clinical management. Clear cell RCC is strongly associated with von Hippel-Lindau (VHL) disease, an autosomal dominant tumor susceptibility syndrome. This disease is caused by germline loss-of-function mutations of the VHL tumor suppressor gene on chromosome 3p25. In addition, VHL loss-of-function is the most common genetic defect in sporadic clear cell RCC. These mutations promote tumor hypervascularity; VHL protein functions in ubiquitinization and degradation of hypoxia-inducible factor-1 alpha (HIF1A), the primary positive regulator of angiogenesis (Na et al., 2003). Rarely, clear cell RCC arises in other hereditary conditions, such as the syndrome of constitutional chromosome 3 translocations, which has been mapped to various breakpoint sites distinct from the VHL locus (Van Erp et al., 2003). Hereditary papillary renal carcinoma (HPRC) is a rare autosomal dominant syndrome caused by germline, gain-of-function mutation of the hepatocyte growth factor receptor gene (MET) on chromosome 7q31-34 (Schmidt et al., 1998). Familial papillary RCC also occurs in the hereditary leiomyomatosis and renal cell cancer syndrome (HLRCC), caused by germline mutations in the fumarate hydratase gene (FH) on chromosome 1q42.1 (Toro et al., 2003). In sporadic papillary RCC, the most common mutations are trisomy 7 and 17, and loss of chromosome Y (Kovacs et al., 1991; Moch, 2004). Chromophobe RCC and oncocytoma arise in Birt-Hogg-Dubé syndrome, an autosomal dominant, multi-organ system tumor syndrome mapped to chromosome 17p12-q11.2 (Nickerson et al., 2002). Other renal tumor subtypes, including clear cell RCC, arise less frequently in this syndrome. The Birt-Hogg-Dubé gene, termed folliculin, is of unknown function. However, folliculin may be inactivated, either by loss of heterozygosity or promoter hypermethylation, in sporadic renal tumors of all histologic subtypes (Khoobehi et al., 2003). In sporadic chromophobe RCC, the most characteristic genetic lesion is loss of multiple chromosomes, while sporadic oncocytoma tends to exhibit different mutations (Kovacs and Kovacs, 1992; Moch, 2004). Angiomyolipoma is associated with tuberous sclerosis, an autosomal dominant disorder caused by loss-of-function mutations in the TSC1 and
TSC2 tumor suppressors on chromosomes 9q34 and 16p13, respectively (Fryer et al., 1987; Kandt et al., 1992; Chen et al., 2003).

Although these genomic abnormalities are an important component of the WHO classification system, cytogenetic studies are not performed routinely on clinical specimens, due to technical difficulty and limited diagnostic sensitivity. Other approaches for molecular diagnostics, such as genomic-scale expression profiling, may provide new assays with greater clinical utility. In the following sections, we describe current technologies for gene expression profiling and general applications of these technologies in cancer research. We then emphasize the use of microarrays to develop novel molecular systems for renal tumor classification.

**Microarray technology**

Several methods are available to profile mRNA expression on a genomic scale, such as gene expression microarrays, differential display and serial analysis of gene expression; the latter two methods are described elsewhere (Velculescu et al., 1995; Liang, 2002). Expression microarrays are solid matrices containing high-density arrays of nucleic acid hybridization targets. Arrays are probed with labeled cDNA or cRNA, derived from the mRNA of biological samples. Probed arrays are scanned robotically for signal (usually fluorescence) at each hybridization target – or “spot” – in order to quantify expression of each gene. Two distinct microarray platforms are widely available: 1) spotted microarrays, containing purified cDNAs or oligonucleotides printed robotically onto glass slides (Schena et al., 1995; Hughes et al., 2001); and 2) microarrays consisting of short oligonucleotides synthesized directly onto solid substrates using photolithographic or inkjet techniques. The photolithographic approach has been pioneered by Affymetrix (Lockhart et al., 1996; Hughes et al., 2001). Spotted and Affymetrix microarrays can be designed to profile thousands of genes (approximating the entire expressed genome) in a single experiment, although each platform has relative strengths and limitations. For example, spotted arrays tend to be lower in cost, offer greater flexibility in the choice of targets, and are designed for co-hybridization with probes from experimental and reference specimens, labeled with distinct fluorescent dyes. Fluorescence ratios at each spot correspond to differential expression of particular genes. Based on direct co-hybridization, spotted arrays provide excellent pair-wise comparisons of relative expression in experimental versus reference specimens. However, spotted arrays cannot determine absolute expression levels. In contrast, Affymetrix arrays measure absolute expression and require less input RNA, a particular advantage when analyzing limited clinical materials such as tumor biopsies. Affymetrix arrays are synthesized at extremely high density with 25-mer oligonucleotide targets from multiple locations of each gene, and each target is paired with a control 25-mer containing a single-base mismatch, in order to reduce false positive interpretations.

**Expression profiling of clinical tumor specimens**

Both spotted and Affymetrix microarrays have been used to classify human cancers with respect to diagnosis, prognosis or therapeutic response (Alizadeh et al., 2000; Takahashi et al., 2001; Shipp et al., 2002; Schuetz et al., 2005). In microarray studies of clinical tumors, variables relating to the specimens themselves assume great importance in experimental design. For example, standard protocols require high quality input mRNA, restricting most solid tumor studies to frozen tissue specimens. Newer protocols are being developed for paraffin-embedded tissue (Ma et al., 2004), which may permit access to the much larger number of well-annotated cases in pathology archives. Still, even in the largest studies, the number of tumor samples is orders of magnitude lower than the number of genes analyzed. Thus, microarrays are expected to correlate many genes with tumor subclasses purely by chance. To increase statistical confidence, technical and biological replicates are incorporated to the extent feasible. Technical replicates are repeated measurements of the same sample, while biological replicates are independent measurements from a single sample class (e.g. different samples of a particular tumor subtype). Biostatisticians have proposed a minimum of 3 technical replicates per specimen and 8-15 biological replicates per sample class (Lee and Whitmore, 2002; Pavlidis et al., 2003). However, few studies have met these criteria fully, due to practical considerations such as numbers of banked specimens and cost per experiment. To overcome these limitations, differential expression must be validated for selected genes using independent methods such as quantitative RT-PCR or immunohistochemistry (Schuetz et al., 2005; Yao et al., 2005). Validation is of particular importance for solid tumors like RCC, which are composed of heterogeneous cell populations; in many studies, immunohistochemistry has served the dual purpose of confirming differential expression and localizing gene products to specific cell types (Young et al., 2001; Higgins et al., 2003; Takahashi et al., 2003b; Schuetz et al., 2005).

**Computational data analysis**

Most cancer-focused microarray studies compare experiments from a series of tumors, in order to identify common gene expression patterns among tumor subgroups. Computational tools for microarray data analysis can be categorized as supervised and unsupervised algorithms. Unsupervised clustering methods are well suited for class discovery (e.g: exploratory identification of novel diagnostic or prognostic subtypes), while supervised methods are ideal for class prediction (e.g: classifying unknown tumors into
previously established diagnostic subtypes) or identification of novel biomarkers.

**Unsupervised analysis**

Average-linkage hierarchical clustering is the dominant unsupervised analytical method in cancer research. The algorithm in Cluster software (Eisen et al., 1998) employs an iterated, agglomerative process to group tumors and genes into hierarchies defined by similarity in expression, as measured by Pearson correlation. Expression hierarchies are visualized as dendrograms. In a single iteration, the two most similar data elements (ie: expression profiles of specific tumors or genes) are joined by a dendrogram node, averaged and replaced by a pseudoelement used in subsequent iterations. The process is repeated until all elements are joined in tumor and gene dendrograms. The algorithm does not constrain dendrogram structures a priori; thus, the number, composition and biological significance of clusters is exploratory in nature. Hierarchically clustered data are often displayed in color-coded grids, or “heat maps”, which list individual tumors and genes in columns and rows, with color indicating relative overexpression or underexpression (see Figure 1). Important limitations of hierarchical clustering must be recognized. For example, large numbers of genes with minimal variation can bias tumor clustering, and mask biologically significant clusters that would be defined by smaller gene sets. Thus, data are often filtered with fold-

![Fig. 1. Unsupervised hierarchical clustering of renal tumors. Similarity measurements are based on Pearson correlation. Median-centered differential gene expression is shown in the color-coded grid, with columns representing individual tumors and rows representing individual genes. Red, green and black indicate expression above, below and at the median of all tumors, respectively. Tumors were clustered into subgroups corresponding to clear cell RCC (CC), papillary RCC (PAP), chromophobe RCC/renal oncocytoma (CHR, ONC) and angiomyolipoma (AML) based on gene expression profiles. The sole outlier was a high-grade papillary RCC with sarcomatoid transformation, which was clustered with clear cell RCC. Roman numerals in the tumor identifications indicate Fuhrman nuclear grade. Adapted from (Schuetz et al., 2005).](image-url)
Molecular classifications of renal tumors

change expression cutoffs or statistical measures. In addition, the clustering algorithm is not based on probabilistic theory. Thus, even randomly generated data will form hierarchical clusters, and statistical confidence cannot be assigned to actual experimental results. In contrast to hierarchical clustering, non-hierarchical methods (eg: quality threshold clustering, k-means clustering and self-organizing maps) redefine the expected numbers of sample clusters; furthermore, the quality threshold algorithm is non-agglomerative. These methods are reviewed elsewhere (Heyer et al., 1999; Quackenbush, 2001). No consensus has been reached on approaches for unsupervised clustering, and many studies use multiple algorithms and metrics to optimize results.

Supervised analysis

In supervised clustering, tumors are classified a priori, based on factors independent of the microarray data (eg: diagnosis, therapy, clinical outcome). Various methods are used to rank genes by correlation with the supervised subclasses. The t test measures differences between subclass means normalized by sample variation, based on the assumption that both subclasses have approximately normal distribution (Roberts et al., 2000). However, this assumption may not hold for small sample numbers. Thus, the SAM algorithm ranks genes by an adjusted t statistic, in which a small positive constant is added to the standard deviation in the denominator to stabilize variance and minimize selection of low-variance genes (Tusher et al., 2001). To estimate statistical significance, SAM uses random sample permutation to calculate a false discovery rate, defined as a ratio of false to total significant genes, averaged over all permutations. If, for example, the gene of a given rank has stronger SAM statistic in the actual data set than similarly ranked genes in all but five of 1000 permutations, the statistical significance could be estimated at p=0.005. A powerful machine learning method known as SVM is also appropriate for supervised analysis (Brown et al., 2000). Results are plotted in n-dimensional gene space (n=number of genes on microarray), and SVM “finds” a hyperplane separating two classes of data with the greatest margin. A major advantage of SVM is ability to employ kernel functions, to map data in different gene spaces for optimum separation. The most basic kernel is linear, but polynomial and radial basis kernels may also be implemented with little effect on computation speed.

Supervised analyses such as SAM and SVM are useful for defining smaller gene sets (“molecular classifiers”) that correlate most strongly with tumor categories and predict identity of unknown tumor samples (Rosenwald et al., 2002; Shipp, et al., 2002). To develop molecular classifiers, gene sets are varied experimentally with respect to correlation rank cutoff and relative weights of each gene for classification. Potential classifiers are tested for predictive power using computational methods such as “leave-one-out cross validation”. Leave-one-out algorithms remove individual tumors from the data set iteratively, and predict their identity by expression of genes in candidate classifiers. Molecular classifiers are optimized for rates of correct classification of left-out samples, and may be tested prospectively in an independent set of microarray data.

Knowledge-based analysis

Knowledge-based analysis frameworks (Brown et al., 2000) are useful to correlate high-volume expression data with knowledge from independent research. These frameworks focus attention on data most likely to be valid, and enhance discovery of novel relationships among genes or specimens with similar expression profiles. Comparison of independent microarray studies has been facilitated by standardized protocols for annotating and formatting data (Brazma et al., 2001; Spellman et al., 2002), and creation of large public data repositories (Edgar et al., 2002; Brazma et al., 2003; Ball et al., 2005). Knowledge-based analysis is optimized by detailed functional annotation of arrays, using hierarchical controlled vocabulary structures such as Gene Ontology (Ashburner et al., 2000) or the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database terminology (Kanehisa and Goto, 2000). Several tools correlate functional annotation with differential expression profiles from microarray experiments. For example, GOStat (Beissbarth and Speed, 2004) and GOMiner (Zeeberg et al., 2003) use the Fisher exact test to determine if expression profiles are significantly enriched or depleted for genes with particular Gene Ontology terms. In addition, several programs upload and analyze gene lists for possible involvement of specific biochemical pathways defined by KEGG (or similar terminology systems). One commercial example is the Affymetrix NetAffx tool at http://www.affymetrix.com/index.affx.

Diagnostic classification of renal tumors

Diagnosis of renal tumors is becoming more complex with the recognition of novel tumor subtypes and the growing use of laparoscopic procedures, which result in small, distorted tumor biopsies with limited histopathologic information (Mabjeesh et al., 2004). Several studies have explored the potential utility of gene expression profiling as a complement to histopathology for renal tumor classification (Boer et al., 2001; Young, et al., 2001; Gieseg et al., 2002; Skubitz and Skubitz, 2002; Higgins et al., 2003; Takahashi et al., 2003b; Yamazaki et al., 2003; Furge et al., 2004; Schuetz et al., 2005; Sultmann et al., 2005; Yao et al., 2005). While these studies have varied with respect to sample cohorts, microarray platforms and analysis algorithms (Table 1), their principal findings have been consistent. In particular, the major histopathologic
subtypes of renal tumor have been distinguished reproducibly by patterns of gene expression. The unique expression profiles are relevant to tumor pathobiology and potentially significant for clinical diagnosis and management. In our recent study, for example, gene expression was measured in clear cell RCC, papillary RCC, chromophobe RCC, oncocytoma and angiomyolipoma using Affymetrix oligonucleotide microarrays (Schuetz et al., 2005). Unsupervised hierarchical algorithms clustered the tumors into groups that correlated strongly with histopathologic subtypes (Fig. 1). Other researchers have reported similar findings (Higgins et al., 2003; Takahashi et al., 2003b; Sultmann et al., 2005). From study to study, individual genes distinguishing the tumor subtypes have varied substantially (Lenburg et al., 2003). However, knowledge-based analyses have revealed consistent themes. For example, supervised SAM analysis of our microarray data, followed by functional classification using GOStat, revealed that expression profiles of clear cell RCC were markedly enriched for immune response and angiogenesis genes, while those of chromophobe

Table 1. Gene expression profiling studies of renal cell carcinoma.

<table>
<thead>
<tr>
<th>STUDY</th>
<th>SPECIMENS</th>
<th>PLATFORM</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Young et al., 2001; Young et al., 2003)</td>
<td>4 CC-RCC, 3 P-RCC, 2 CH-RCC, 1 ONC, matched kidney</td>
<td>cDNA microarray (7000 clones)</td>
<td>Diagnostic classification; discovery of immunohistochemical markers</td>
</tr>
<tr>
<td>(Takahashi et al., 2001)</td>
<td>29 CC-RCC, matched kidney</td>
<td>cDNA microarray (22000 clones)</td>
<td>Prognostic classification</td>
</tr>
<tr>
<td>(Boer et al., 2001)</td>
<td>37 renal tumors, matched kidney</td>
<td>Nylon cDNA microarray (31500 clones)</td>
<td>Diagnostic classification; correlation of gene expression with tumor stage</td>
</tr>
<tr>
<td>(Skubitz and Skubitz, 2002)</td>
<td>8 CC-RCC, 11 normal and 8 diseased kidney, 256 other tissues</td>
<td>Affymetrix U95 (63000 probe sets)</td>
<td>Expression patterns that discriminate RCC from non-neoplastic kidney and other tissues</td>
</tr>
<tr>
<td>(Gieseg et al., 2002)</td>
<td>9 CC-RCC, 2 CH-RCC, 1 UC, 1 MA, 8 matched kidney</td>
<td>Affymetrix HuGene FL (5600 genes)</td>
<td>Diagnostic classification; discrimination of RCC from non-neoplastic kidney</td>
</tr>
<tr>
<td>(Higgins et al., 2003)</td>
<td>28 conventional (CC-RCC), 4 P-RCC, 3 CH-RCC, 2 ONC, 1 AML</td>
<td>cDNA microarray (23000 clones)</td>
<td>Diagnostic classification; discovery of immunohistochemical markers</td>
</tr>
<tr>
<td>(Takahashi et al., 2003b)</td>
<td>39 CC-RCC, 8 P-RCC, 5 CH-RCC, 6 G-RCC, 5 S-RCC, 2 ONC, 3 UC, 5 WT, non-neoplastic kidney.</td>
<td>cDNA microarray (23000 clones)</td>
<td>Diagnostic classification; prognostic classification; discovery of immunohistochemical markers</td>
</tr>
<tr>
<td>(Lenburg et al., 2003)</td>
<td>18 CC-RCC, 9 matched kidney</td>
<td>Affymetrix U133A/B (45000 probe sets)</td>
<td>Discrimination of RCC from non-neoplastic kidney; comparison with expression profiles from other studies</td>
</tr>
<tr>
<td>(Copland et al., 2003)</td>
<td>Localized or metastatic RCC, matched kidney, RCC cell lines</td>
<td>Affymetrix U95A (10000 genes)</td>
<td>Discovery of aberrant transforming growth factor beta signaling in RCC carcinogenesis and progression</td>
</tr>
<tr>
<td>(Vasselli et al., 2003)</td>
<td>58 Stage IV RCC, 8 matched kidney.</td>
<td>cDNA microarray (64000 clones)</td>
<td>Prognostic classification</td>
</tr>
<tr>
<td>(Yamazaki et al., 2003)</td>
<td>10 CC-RCC, 2 P-RCC, 3 CH-RCC, matched kidney.</td>
<td>Affymetrix U95A (10000 genes)</td>
<td>Diagnostic classification; discovery of immunohistochemical markers</td>
</tr>
<tr>
<td>(Liou et al., 2004)</td>
<td>Pooled CC-RCC, pooled matched kidney, pooled RCC cell lines</td>
<td>Affymetrix HuGene FL (5600 genes)</td>
<td>Discrimination of RCC from non-neoplastic kidney</td>
</tr>
<tr>
<td>(Furge et al., 2004)</td>
<td>60 CC-RCC, 5 P-RCC, 16 CH-RCC, matched kidney.</td>
<td>cDNA microarray (23000 clones)</td>
<td>Diagnostic classification; inference of cytogenetic abnormalities</td>
</tr>
<tr>
<td>(Sultmann et al., 2005)</td>
<td>65 CC-RCC, 13 P-RCC, 9 CH-RCC, 25 kidney.</td>
<td>cDNA microarray (42000 clones)</td>
<td>Diagnostic classification; prognostic classification; correlation with cytogenetics</td>
</tr>
<tr>
<td>(Yao et al., 2005)</td>
<td>29 CC-RCC, 3 CH-RCC, 1 WT, 9 kidney</td>
<td>Affymetrix U95A (10000 genes)</td>
<td>Diagnostic classification; discovery of quantitative RT-PCR and immunohistochemical markers</td>
</tr>
<tr>
<td>(Schuetz et al., 2005)</td>
<td>13 CC-RCC, 5 P-RCC, 4 CH-RCC, 3 ONC, 6 AML</td>
<td>Affymetrix HG-Focus (8700 genes)</td>
<td>Diagnostic classification; discovery of quantitative RT-PCR and immunohistochemical markers</td>
</tr>
</tbody>
</table>

CC-RCC: Clear cell renal cell carcinoma (RCC); P-RCC: Papillary RCC; CH-RCC: Chromophobe RCC; G-RCC: Granular RCC; S-RCC: Sarcomatoid RCC; ONC: Oncocytoma; AML: Angiomyolipoma; UC: Urothelial carcinoma; WT: Wilms tumor; MA: Metanephric adenoma.
RCC and oncocytoma were enriched for energy pathway genes, including many genes associated with mitochondrial biology (Schuetz et al., 2005). Angiogenic gene expression is likely related to tumor vascularity in clear cell RCC (Eble et al., 2004), while energy pathway gene expression may correlate with the abundant mitochondria in chromophobe RCC and oncocytoma (Tickoo et al., 2000). As shown in Table 2, we obtained very similar results using publicly available data from Stanford University (Higgins et al., 2003). Other groups have produced concordant findings (Takahashi et al., 2003a,b; Sulmann et al., 2005).

Microarray data are consistent with current models of renal tumor histogenesis. These models relate clear cell and papillary RCC to proximal nephron epithelium, and chromophobe RCC and oncocytoma to distal nephron intercalated cells (Wallace and Nairn, 1972; Ortman et al., 1988; Storkel et al., 1989; Sulmann et al., 2005). In our experiments, clear cell RCC overexpressed the proximal nephron markers megalin and cubulin; papillary RCC strongly overexpressed the proximal nephron marker alpha methylacyl CoA racemase; and chromophobe RCC and oncocytoma overexpressed the distal nephron markers beta defensin-

Table 2. Gene ontologies enriched in clear cell RCC and chromophobe RCC/oncocytoma.

<table>
<thead>
<tr>
<th>SYMBOL UNIGENE</th>
<th>GENE NAME</th>
<th>SYMBOL UNIGENE</th>
<th>GENE NAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear cell RCC - GO:0006952 (Defense response)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1QB Hs.8986</td>
<td>complement component 1q beta</td>
<td>G1P2 Hs.458485</td>
<td>interferon alpha-inducible protein (IFI-15K)</td>
</tr>
<tr>
<td>C1QR1 Hs.97199</td>
<td>complement component 1q receptor 1</td>
<td>GEM Hs.79022</td>
<td>GTP binding protein overexpressed in skeletal muscle</td>
</tr>
<tr>
<td>C1S Hs.458355</td>
<td>complement component 1s</td>
<td>HLA-DPB1 Hs.368409</td>
<td>major histocompatibility complex, class II, DP beta 1</td>
</tr>
<tr>
<td>CD69 Hs.82401</td>
<td>CD69 antigen</td>
<td>HLA-DRB5 Hs.534322</td>
<td>major histocompatibility complex, class II, DR beta 5</td>
</tr>
<tr>
<td>CLECSF2 Hs.85201</td>
<td>C-type lectin, superfamily member 2</td>
<td>IFITM1 Hs.458414</td>
<td>interferon induced transmembrane protein 1 (9-27)</td>
</tr>
<tr>
<td>CSF1R Hs.171412</td>
<td>colony stimulating factor 1 receptor (c-fms)</td>
<td>IL4R Hs.75545</td>
<td>interleukin 4 receptor</td>
</tr>
<tr>
<td>F8 Hs.413083</td>
<td>coagulation factor VIII</td>
<td>ISGF3G Hs.1706</td>
<td>interferon-stimulated transcription factor 3 gamma</td>
</tr>
<tr>
<td>FCER1G Hs.433300</td>
<td>Fc fragment of IgE receptor; gamma polypeptide</td>
<td>ITGB2 Hs.375957</td>
<td>integrin, beta 2 (CD18 antigen)</td>
</tr>
<tr>
<td>FCGRT Hs.111903</td>
<td>Fc fragment of IgG, receptor, transporter, alpha</td>
<td>RNASE6 Hs.23262</td>
<td>ribonuclease, RNase A family, k6</td>
</tr>
<tr>
<td>FOS Hs.25647</td>
<td>c-fos</td>
<td>UBD Hs.44532</td>
<td>ubiquitin D</td>
</tr>
<tr>
<td>Chromophobe RCC/oncocytoma - GO:0006091 (Generation of precursor metabolites and energy)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACO2 Hs.202505</td>
<td>aconitase 2, mitochondrial</td>
<td>IVD Hs.410396</td>
<td>isovaleryl Coenzyme A dehydrogenase</td>
</tr>
<tr>
<td>ATP6V1A Hs.409131</td>
<td>ATPase, H+ transporting, lysosomal 70kDa, V1 subunit A</td>
<td>NDUFS4 Hs.10758</td>
<td>NADH dehydrogenase (ubiquinone) Fe-S protein 4</td>
</tr>
<tr>
<td>ATP6V1F Hs.78089</td>
<td>ATPase, H+ transporting, lysosomal 14kDa, V1 subunit F</td>
<td>PFKM Hs.75160</td>
<td>phosphofructokinase, muscle</td>
</tr>
<tr>
<td>BPGM Hs.198365</td>
<td>2,3-bisphosphoglycerate mutase</td>
<td>PPARA Hs.534037</td>
<td>peroxisome proliferative activated receptor, alpha</td>
</tr>
<tr>
<td>CKMT2 Hs.80691</td>
<td>creatine kinase, mitochondrial 2</td>
<td>PPARG Hs.387667</td>
<td>peroxisome proliferative activated receptor, gamma</td>
</tr>
<tr>
<td>COX17 Hs.16297</td>
<td>COX17 homolog, cytochrome c oxidase assembly protein</td>
<td>UQCR Hs.8372</td>
<td>ubiquinol-cytochrome c reductase (6.4kD) subunit</td>
</tr>
<tr>
<td>COX7C Hs.430075</td>
<td>cytochrome c oxidase subunit VIIc</td>
<td>UQCRB Hs.131255</td>
<td>ubiquinol-cytochrome c reductase binding protein</td>
</tr>
<tr>
<td>IDH3A Hs.277543</td>
<td>isocitrate dehydrogenase 3 (NAD+) alpha</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data are from Schuetz et al, 2005 and Higgins et al, 2003 (granular RCC was not included). Differential expression was determined with SAM and GOSTat. Clear cell RCC overexpressed Gene Ontology GO:0006952 (Emory: p=4.04E-27; Stanford: p=3.85E-07). Chromophobe RCC/oncocytoma overexpressed GO:0006091 (Emory: p=6.16E-42; Stanford: p=1.16E-02)
Molecular classifications of renal tumors

1, parvalbumin, chloride channel Kb, claudin-7, claudin-8 and epidermal growth factor. These findings have been validated in our laboratory by quantitative RT-PCR and immunohistochemistry (Young et al., 2001, 2003; Schuetz et al., 2005), and are consistent with research from other laboratories (Martignoni et al., 2001; Tretiakova et al., 2004). Thus, proximal and distal nephron markers may be useful for diagnostic classification of renal tumors.

To further develop microarray-based methods for renal tumor classification, the group from Van Andel Research Institute has developed an intriguing method termed comparative genomics microarray analysis (CGMA), which is used to infer cytogenetic abnormalities from regional expression biases between tumors and adjacent non-neoplastic tissue (Takahashi et al., 2003a). In a recent study, abnormalities inferred by CGMA correlated with recurring cytogenetic lesions in renal tumor subtypes. In addition, molecular diagnostic classification of clear cell, papillary and chromophobe RCC was strengthened by combining CGMA data with differential gene expression profiles (Furge et al., 2004). The German Cancer Research Center group compared gene expression profiles with direct cytogenetic data from a large cohort of renal tumors, and confirmed that expression levels of specific genes correlated with the copy number of respective chromosomal arms (Sultmann et al., 2005).

Unsupervised microarray analyses have potential to discover novel renal tumor subclasses. The Stanford University group described a subset of conventional RCC with granular cytoplasm, which did not cluster together or with any particular subtype in the overall study cohort (Higgins et al., 2003). Similar findings were obtained by the Van Andel Research Institute group (Takahashi et al., 2003b). Historically, the term “granular cell RCC” was used for high-grade tumors with eosinophilic cytoplasm, and was intended to distinguish these cases from conventional clear cell RCC. However, it is now appreciated that some lesions with “granular cells” are clinically and molecularly consistent with the current concept of clear cell RCC, while other cases should be diagnosed as papillary, chromophobe or other types of RCC. Thus, the term granular cell RCC is not included in the current WHO classification system (Eble et al., 2004). Nevertheless, it remains possible that certain granular tumors will eventually be classified as a distinct RCC subtype by gene expression profiling.

Quantitative mRNA expression changes detected with microarrays do not necessarily reflect changes at the protein level. Also, microarrays do not detect post-transcriptional modifications such as protein phosphorylation. Therefore, characterization of RCC by gene expression profiling will likely evolve toward integrated analysis of the transcriptome and proteome. A recent study identified differentially expressed proteins in tissue and primary cell cultures derived from RCC (clear cell and papillary) and matched non-neoplastic kidney, using two-dimensional SDS-polyacrylamide gel electrophoresis. Selected proteins were characterized by liquid chromatography and mass spectroscopy (Shi et al., 2004). The proteins overexpressed most consistently in RCC were alpha-beta crystallin, manganese superoxide dismutase and annexin IV. In the study cohort, clear cell and papillary carcinomas were not distinguished by expression of these gene products. In addition, the biological significance of this expression profile, or the relationship to mRNA data in RCC, remains unclear.

Novel molecular assays for clinical diagnosis of renal tumors

The United States Food and Drug Administration has not yet approved microarrays for clinical use, largely because methods for quality control and interpretation have not been standardized (Hackett and Lesko, 2003). Still, microarray analyses have produced a direct impact on clinical diagnosis of RCC, through discovery of novel immunohistochemical markers for each major tumor subtype (Zhou and Rubin, 2001; Zhou et al., 2005). Several candidate markers have been validated in subsequent independent research (Young et al., 2003; Pan et al., 2004; Tretiakova et al., 2004; Chuang et al., 2005). Table 3 summarizes renal tumor immunomarkers identified with microarrays. Continued work is needed to combine these markers into panels, for diagnosis of renal tumors that are difficult to classify by light microscopy. For example, we have shown that a panel of beta defensin-1, parvalbumin and vimentin is superior to any single marker used alone for classification of clear cell RCC, papillary RCC and chromophobe RCC/oncocytoma (Young et al., 2003).

Quantitative RT-PCR is established in many clinical diagnostic laboratories, and offers certain advantages over immunohistochemistry, such as wider dynamic range. We have developed assays for megalin, alpha methylacyl CoA racemase, beta defensin-1 and other markers; these assays distinguish the major subtypes of renal tumor reliably and can utilize RNA from fixed archival tissue (Schuetz et al., 2005). The German Cancer Research Center group developed a promising 80-gene predictor for renal tumor classification from their microarray data, and validated several genes by quantitative RT-PCR (Sultmann et al., 2005). Larger prospective studies are needed to establish the clinical potential of RT-PCR for renal tumor diagnosis.

Prognostic classification of renal tumors

The group from University of Texas and MD Anderson Cancer Center conducted microarray experiments on patient-matched non-neoplastic kidney, localized RCC and metastatic RCC (Copland et al., 2003). Aberrations in the signaling pathway of transforming growth factor beta were discovered, with consistent loss of type III receptor (TBR3) expression in all RCC samples, and further loss of type II receptor...
Markers for Chromophobe RCC and Oncocytoma

<table>
<thead>
<tr>
<th>MARKER</th>
<th>SYMBOL</th>
<th>FUNCTION</th>
<th>LOCALIZATION</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta defensin-1</td>
<td>DEFB1</td>
<td>Antimicrobial/antitumor agent</td>
<td>Cytoplasm</td>
<td>(Young, et al., 2001; Young, et al., 2003)</td>
</tr>
<tr>
<td>Parvalbumin</td>
<td>PVALB</td>
<td>Calcium-binding protein</td>
<td>Cytoplasm</td>
<td>(Martignoni, et al., 2001; Young, et al., 2003)</td>
</tr>
<tr>
<td>Stem cell factor receptor</td>
<td>KIT</td>
<td>Cell differentiation</td>
<td>Membrane</td>
<td>(Higgins, et al., 2003; Yamazaki, et al., 2003)</td>
</tr>
<tr>
<td>Carbonic anhydrase II</td>
<td>CA2</td>
<td>Zinc metalloenzyme</td>
<td>Cytoplasm</td>
<td>(Takahashi, et al., 2003b)</td>
</tr>
<tr>
<td>Cytokeratin 7</td>
<td>CK7</td>
<td>Cytoplasm</td>
<td>Membrane</td>
<td>(Schuetz, et al., 2005)</td>
</tr>
</tbody>
</table>
of the study cohort, in order to develop an expression classifier for distinguishing RCC patients from healthy volunteers. This classifier was validated with leave-one-out algorithms on the training subset, followed by predictive assays on the remainder (testing subset) of the study cohort (Twine et al., 2003). A similar approach was used to identify gene expression markers for predicting patient outcome after CCI-779 therapy (Burczynski et al., 2005b). The performance of gene expression was reported to be comparable or better than standard methods for clinical risk classification (Motzer et al., 1999). These studies included patients with different subtypes of RCC; however, the ability to predict tumor histology by PBMC gene expression was not described. Nevertheless, this research highlights the possibility of minimally invasive testing for RCC diagnosis and therapeutic management, based on gene expression profiling of peripheral blood specimens.

**Therapeutic implications of gene expression profiling**

The preceding PBMC experiments illustrate a potential role of microarrays in clinical trials for RCC (Burczynski et al., 2005a). Also, many overexpressed genes in RCC tumor tissue have therapeutic implications. For example, immune response genes, which distinguish clear cell RCC from other subtypes, may be important for the relative responsiveness of clear cell RCC to immunotherapy compared to non-clear cell tumors (Motzer et al., 2002). Expression levels of immune response genes vary among individual tumors (Young et al., 2001), and future studies should determine if variability correlates with response to immunotherapy. Similarly, angiogenesis genes, which are overexpressed in clear cell RCC, may be relevant to anti-angiogenic therapies being evaluated in clinical trials (Gordon, 2004; Yang, 2004). Many overexpressed angiogenic genes are regulated by HIF1A, itself activated by loss of VHL in clear cell RCC (Na et al., 2003). Curiously, expression levels of HIF1A and VHL have not varied reproducibly in microarray studies, possibly due to post-transcriptional regulation. Nevertheless, HIF1A-dependent angiogenesis markers such as vascular endothelial growth factor (VEGF) and endothelin-1, which are overexpressed consistently in clear cell RCC (Lenburg et al., 2003), may be useful for therapeutic monitoring and outcome prediction. HIF1A and immune response regulators both increase expression of carbonic anhydrase IX, which is the target of G250 monoclonal antibody therapy for RCC (Brouwers et al., 2003; Mulders et al., 2004). In large microarray studies of multiple cancers, it appears that RCC may be distinguished from other tumor types by overexpression of angiogenesis genes and coregulation of VEGF and carbonic anhydrase IX (Amatschek et al., 2004; Jubb et al., 2004). Microarray studies have established that stem cell factor receptor (KIT) is overexpressed in chromophobe RCC (Higgins et al., 2003; Yamazaki et al., 2003), leading several experts to suggest tyrosine kinase inhibitors such as Gleevec for advanced carcinomas of this subtype (Potti and George, 2004; Rini et al., 2004).

**Microarray studies of renal development and tumor biology**

To study the effect of VHL expression and hypoxia in RCC, microarray data were compared from a human RCC cell line with mutant versus transfected wild-type VHL, under hypoxic and normoxic conditions (Wykoff et al., 2004). Significant concordance was observed in the gene response to hypoxia and genes suppressed by VHL. The gene most upregulated by hypoxia and VHL mutation was cyclin D1.

Systematic microarray analyses of rat and mouse kidney development have characterized expression profiles in metanephric mesenchyme, embryonic and adult kidney. Embryonic E12.5 kidney overexpressed a large number of transcription factor, growth factor, signal transduction and cell cycle genes (Stuart et al., 2001; Schwab et al., 2003). Renal gene expression has also been profiled in the Eker rat, which bears a heterozygous germline mutation in one allele of the tuberin (TSC2) tumor suppressor, and is predisposed to spontaneous and chemical-induced RCC. In one study, microarrays were used to compare non-neoplastic kidney from TSC2 (+/-) and wild type TSC2 (+/+ ) rats (Sen et al., 2004). Differentially expressed genes in this system were related to the phosphatidylinositol-3 kinase/AKT pathway, cell cycle regulation, cell proliferation, cell adhesion and endocytosis. Microarrays have also been applied to chemically-induced neoplastic cells derived from Eker rat renal epithelium (Patel et al., 2003). Differential expression was pronounced in genes related to signal transduction (protein kinases and ras-pathway genes); electron transport and energy homeostasis (cytochrome c oxidase subunits); stress response, tissue remodeling, and DNA repair (glutathione-S-transferases; plasminogen activator; tissue inhibitor of metalloproteinase 3). Interestingly, several of these genes have been overexpressed in clinical microarray studies of adult RCC. In a study of pediatric nephroblastoma (Wilms tumor), tumor cells resembled an early committed stage in metanephric development, overexpressing transcription factors essential for cell survival and proliferation at this stage (Li et al., 2002). Based on expression profiles, Wilms tumor blastema appeared to be arrested at a committed stage of epithelial-mesenchymal transition. Expression databases of normal human kidney development will be useful to correlate this finding with Wilms tumor pathobiology. Also, since epithelial-mesenchymal transition is critical for neoplastic and non-neoplastic pathology of the adult human kidney, as well as for the Eker rat model (Rastaldi et al., 2002; Cook and Walker, 2004), this research should help to elucidate the complex gene expression profiles in adult RCC subtypes.
Future directions

Microarray experiments have advanced the understanding of renal tumor pathobiology considerably, although need for further discovery clearly still exists. Clinically, current trends point to an increased role of expression profiling for diagnosis, prognosis and treatment of RCC. For diagnosis, the growing frequency of small tumor biopsies has underscored the need to discover molecular markers that complement histopathology. A critical challenge is classification of tumors with granular eosinophilic cytoplasm (referred to as oncocytic neoplasms), based on limited biopsy material – particularly the related lesions chromophobe RCC and oncocytoma (Abrahams and Tamboli, 2005). Due to morphologic heterogeneity of renal tumors, and the resulting concern of sampling bias, few urologists or pathologists accept a benign diagnosis such as oncocytoma from biopsies. Unfortunately, while benign lesions might be managed appropriately by watchful waiting in some patients, resections or ablations are rarely deferred based on biopsy results, due to concern over missed diagnosis of cancer. In our recent study, we identified claudin-7 and claudin-8 (tight junction proteins expressed normally in distal nephron epithelium) as preliminary candidate expression markers for chromophobe RCC (Schuetz et al., 2005). Cytokeratin 7 and the RON oncogene product (macrophage stimulating protein receptor) have also been suggested as markers for chromophobe carcinoma and oncocytoma, respectively (Rampino et al., 2003; Zhou et al., 2005), although independent studies have been inconclusive (Abrahams et al., 2004; Patton et al., 2004). Due to the extreme similarity in gene expression between chromophobe RCC and oncocytoma (Schuetz et al., 2005), the identification of reliable molecular marker panels will likely be difficult. However, new technologies for high-density microarray analysis of fixed tissue (Ma et al., 2004), coupled with highly quantitative RT-PCR or nanoparticle-based immunoassays (Nie and Emory, 1997; Han et al., 2001; Mulvaney et al., 2003; Schuetz et al., 2005), may allow the precise study of sufficiently large sample cohorts to develop methods with clinical utility. Similarly, these new technologies should permit the study of well-documented RCC patient cohorts with long-term survival and therapeutic outcome data, in order to develop novel assays for managing patients with these very challenging cancers. Overall, molecular classification is a very promising approach to improve clinical diagnosis and treatment of renal tumors.

References


Molecular classifications of renal tumors

ArrayExpress--a public repository for microarray gene expression data at the EBI. Nucleic Acids Res. 31, 68-71.


Molecular classifications of renal tumors

Roberts C.J., Nelson B., Marton M.J., Stoughton R., Meyer M.R.,
Molecular classifications of renal tumors


Molecular classifications of renal tumors


Accepted October 19, 2005