Elucidation of the molecular signatures of renal cell carcinoma by gene expression profiling

Masayuki Takahashi, Bin T. Teh*, and Hiro-omi Kanayama

Department of Urology, Institute of Health Biosciences, The University of Tokushima Graduate School, Tokushima, Japan; *Laboratory of Cancer Genetics, Van Andel Research Institute, 333 Bostwick Avenue, N.E., Grand Rapids, MI 49503; and Department of Urology, Institute of Health Biosciences, The University of Tokushima Graduate School, Tokushima, Japan

Abstract: Renal cell carcinoma (RCC) is the 10th most common cancer in United States. It is a heterogeneous disease with various histologic types. Since high-throughput technologies such as microarrays have been introduced, molecular confirmation of previously known findings in RCC has been made and new molecular findings have emerged. We review the accumulating advances in this field and their clinical implications. The published data so far have proved to be significant and promising, and numerous microarray studies with larger number of cases are currently ongoing or being planned. Although various clinical parameters are being refined for diagnosis and prognosis, these data obtained by microarray studies will undoubtedly contribute to both and eventually impacts the treatment of RCC. J. Med. Invest. 53 : 9-19, February, 2006

Keywords: gene expression profiling, renal cell carcinoma, prognosis

INTRODUCTION

There are approximately 32,000 new cases of renal cell carcinoma (RCC) diagnosed each year in the United States, accounting for 3% of all adult malignancies(1). There are approximately 12,000 deaths each year related to this cancer. One-third of the patients present with metastatic disease and have a median survival of 7-11 months and 5-year survival of 0-10%. The incidence of RCC has been increasing, a phenomenon that cannot be accounted for by the wider use of imaging procedures(2). RCC is more common in men than women, especially in men over 55 years of age. Risk factors include genetic predisposition, hypertension, obesity(3) and occupational exposures(4). Clinically flank pain, a palpable mass, grosshematuria are considered to be classical triad and there are other symptoms of weight loss, pain from a metastatic lesion or accompanied by paraneoplastic syndrome. However as the diagnostic imaging such as helical CT and MRI are widespread and much more improved than before, more and more cases of RCC diagnosed are asymptomatic. Recently, Motzer et al. (5) identified five prognostic factors that correlated with overall survival in patients with metastatic RCC, which are Karnofsky performance status, time from diagnosis of RCC to treatment with interferon alfa, serum lactate dehydrogenase, corrected serum calcium, and hemoglobin. These prognostic factors were very recently validated by another group(6). In addition to that, incorporation of molecularly identified prognostic markers would undoubtedly lead to

Received for publication December 25, 2005 ; accepted January 20, 2006.

Address correspondence and reprint requests to Hiro-omi Kanayama, M.D., Ph.D., Department of Urology, Institute of Health Biosciences, The University of Tokushima Graduate School, Kuramoto-cho, Tokushima 770-8503, Japan and Fax : +81-88-633-7160
better stratification of RCC patients in regard with prognosis.

MICROARRAY TECHNOLOGY

To date, numerous studies of gene expression profiling in RCC have been reported with focus on molecular classification or diagnosis and prognosis. Different microarray platforms have been used. One of them is the two-colored spotted cDNA microarrays, relatively large cDNA probes (that have been generated from the reverse transcription of RNA from a variety of sources) are spotted on a support (usually glass) with an arrayer or inkjet printer. RNA is obtained from RCC and is labeled with a fluorescently tagged nucleotide (typically green or red fluorescent cytosine triphosphate [CTP]) that is incorporated during a reverse transcription reaction to generate labeled cDNA. Simultaneously, RNA from normal kidney tissue is labeled in the same fashion with a fluorescent nucleotide of a different color. The two labeled cDNA pools (from tumor and normal) are then mixed and allowed to hybridize to the various target probes on the microarray. After washing, the microarrays are scanned to measure the fluorescence at each wavelength, such that the relative degree of fluorescence can be calculated (expressed as a fluorescence ratio) and used to determine changes in RNA expression between the reference and the sample.

An alternative technology using short sequences of oligonucleotides (typically 20-60 base pairs long) physically builds these sequences onto the array surface using photolithographic technology(7). The sequence of these probes is pre-selected from genomic sequencing databases. Using this technology, it is possible to exceed hundreds of thousands of oligonucleotides on a small chip. RNA is then isolated from the samples of interest, labeled with fluorescently tagged nucleotides in a similar fashion to the spotted cDNA arrays, and hybridized to the chip surface. For those genes that are found to be present, the intensity of fluorescence is a direct measure of gene expression. Computer-assisted software then allows the researcher to compare gene expression changes among samples. In addition to these technologies, there are a number of variations on the theme, each of which offers some advantages (as well as disadvantages) relative to the others.

GENE EXPRESSION PROFILING WITH MICROARRAY ANALYSIS

Given even well conducted clinical trials and careful clinical studies of clinical parameters, there are still challenges in diagnosis, prognosis, and treatment. One main reason is the lack of understanding of the biology and pathogenesis. The recent breakthrough biotechnology is the development of microarray analysis, which allows a comprehensive and high-throughput approach to view the molecular signatures of cancers. Many of these molecular signatures are distinct and highly correlated with pathology, clinical outcome, drug response and undoubtedly more of such correlated are being discovered. For example, Kakikuchi et al analyzed 33 biopsy samples of advanced non-small cell lung cancer (NSCLC) with a genome-wide cDNA microarray to predict the response of NSCLC patients to gefitinib, which is an inhibitor of epidermal growth factor receptor-tyrosine kinase(8). It is reported that the scoring system with the expressions of the 12 selected genes clearly differentiated the responders from non-responders without any overlap, and accurately predicted responses to gefitinib in 16 additional NSCLC cases.

MOLECULAR CLASSIFICATION AND HISTOLOGICAL FINDINGS OF KIDNEY TUMORS

To date several reports showed that gene expression profiling itself could distinguish different subtypes of RCC. We examined the molecular signatures of 70 kidney tumors, consisting of 7 different subtypes: clear cell, papillary, granular, chromophobe, sarcomatoid RCC, oncocytoma, transitional cell carcinoma (TCC) of the renal pelvis and pediatric Wilms tumors by using cDNA microarray analysis(9). Hierarchical clustering revealed several clusters which corresponded to histological appearances (Figure 1). What is interesting and reassuring is that our findings appeared to be consistent with the chromosomal and genetic changes previously reported in RCC. In an effort to correlate gene expression data with chromosomal changes, a software was created called comparative genomic microarray analysis (CGMA) which identify chromosomal loss or gain based on the chromosomal location of genes and their differential expression in different subtypes of RCC (10). For example, CGMA detected the loss of
chromosome 3p and gain of chromosome 5q based on the relative under-expression and over-expression of genes in 3p and 5q respectively in clear cell RCC. Previous chromosomal and genetic studies, for example loss of heterozygosity (LOH) and comparative genomic hybridization (CGH) studies have indeed shown frequent loss of 3p and gain of 5q in clear cell RCC. Similarly, our gene expression signatures of papillary RCC also identify the chromosomal gain of 7, 16 and 17 as previously reported. Interestingly, our gene expression microarray profiling showed that chromophobe RCC and oncocytoma constituted one cluster in a small subset of tumors which share a high degree of similarity in expression of mitochondria-related genes (both tumors are characterized by abundance of mitochondria). However, a subsequent larger study with a special focus on these two entities, have identified the distinguishing set of genes which is further supported by CGMA which identify relatively few chromosomal changes in the oncocytoma group. It is clear that gene expression profiling have served as a powerful tool in diagnosis of tumors and accurate diagnosis of histological subtypes in RCC is important not only in predicting prognosis (11) but also in designing appropriate treatment for patients. For example, several current clinical trials are designed only for clear cell RCC.

Besides molecular classification, differentially expressed genes in each RCC subtype also provides insights into tumor biology and allows dis-
covery of novel molecular markers. The latter usually involves the immunohistochemical study which was performed using antibody of the differentially expressed genes identified in each subtype of RCC. Like other groups, we performed such studies in tissue microarray containing dozens to hundreds of RCC tissues. The majority of results by immunohistochemical staining were consistent with the gene expression data obtained by microarray study. For example, glutathione S-transferase alpha, was highly expressed in clear cell RCC by tissue microarray (clear cell 90% vs others <10% positive). Insulin-like growth factor binding protein-3 (IGFBP-3) was also strongly positive in clear cell RCC (clear cell 100% vs others <10% positive) in later study(12). A correlation between increased VEGF and IGFBP-3 has recently been identified in lung adenocarcinomas. Considering known high expression of VEGF, high expression of IGFBP-3 might be an interesting finding in clear cell RCC and may have therapeutic implications. Alpha-methylacyl racemase (AMACR) was first found by our study to be specifically positive in papillary RCC (papillary 100% vs others <10%). AMACR has been shown over-expressed in the microarray study of prostate cancer (13). High expression of AMACR at the protein level was also confirmed in more than 90% of prostate cancer but not benign prostatic tissues by immunohistochemistry, suggesting that AMACR may be a specific marker for prostate cancer tissues (14, 15). Subsequently studies on a larger number of papillary RCC have confirmed that AMACR is a specific marker for papillary RCC although its functional role in papillary RCC remains unknown (16, 17). In addition, we included in our study six “granular cell RCC”, which according to the recommendation by the work group of UICC and AJCC, was no longer used as an entity. Our microarray showed that they were scattered in the dendrogram, clustering with other subtypes except one. We further proceeded to examine H&E slides of these granular cell RCC in a blinded fashion by an expert urologic pathologist. Interestingly, the pathologist diagnosed granular cell RCC which clustered with clear cell RCC as clear cell RCC and one which clustered with chromophobe RCC as chromophobe RCC. All these support the recommendation by the UICC and AJCC recommendation. More interestingly, one particular granular cell RCC, which was not clustered with any RCC subgroups by genetic profiles, was found particularly difficult to diagnose after careful review by several experts. A possible diagnosis of pheochromocytoma was also raised. Further immunohistochemical studies showed positive staining for neuroendocrine markers but not for keratin, which are consistent with pheochromocytoma. A germline missense mutation, D 119 E, in the familial pheochromocytoma related gene succinate dehydrogenase subunit D (SDHD), was subsequently identified to strengthen this diagnosis. The treatment modality was revised and several courses of radiotherapy were given to the patient. For more than two and half years, the patient has had stable disease(18). As alteration of diagnosis based on microarray study was also reported in another type of malignancy in the same way(19), it seems obvious that the gene expression profiling definitely adds some more information over clinical parameters in the clinical settings. In other words, if a case shows quite different gene expression patterns from other cases with the same histological diagnosis, re-consideration for the diagnosis might be needed.

One more interesting finding in this study was that two papillary RCC specimens, one primary tumor and the other a metastasized lymph node from the same patient were also examined and found to cluster closely together, showing very similar gene expression pattern. It suggests that there could be a distinct molecular signature for metastasis but more paired samples from primary and metastatic site from the same patients should be studied.

**DIFFERENTIALLY EXPRESSED GENES IN CLEAR CELL RCC**

As clear cell RCC accounts for about 80% of all renal tumors, we particularly gave some insights to altered gene expression in this type of RCC. Several reports showed over-or under-expressed genes in clear cell RCC. The listed genes are somewhat different probably because of technical problems of microarray experiments and different gene selection criteria for altered genes. Some of the genes that have the highest differential expression ratio in the tumors have not reported enough in cc-RCC. For example, ceruloplasmin, a protein involved in iron and copper homeostasis, has the highest increase in expression in clear cell RCC. Its serum level increases markedly in anemia of iron deficiency, hemorrhage, renal failure, sickle cell disease, pregnancy, and inflamma-
tion. To date, only one study has reported its secretion by RCC (20) and another study described its elevation in the serum of RCC patients (21). Another copper-related protein, lysyl oxidase, was also up-regulated. It is an extracellular enzyme involved in the connective tissue maturation pathway. It is highly expressed in invasive breast cancer cell lines (22), but it has heretofore not been studied in RCC. The majority of tumors also have high expression of vascular endothelial growth factor (VEGF), a well-known angiogenesis factor. Other elevated genes in clear cell RCC include fibronectin, caveolin-1, IGFBP-3, and regulator of G-protein signaling. It will be interesting to perform further functional studies to elucidate their roles in clear cell RCC.

Some of the down-regulated genes, on the other hand, may be tumor suppressors or involved in the tumorigenesis of clear cell RCC. Those that are highly down-regulated are kinase, fatty acid binding protein 1, phenylalanine hydroxylase, epidermal growth factor (EGF), plasminogen and aldolase B. Most strikingly, kinase was found to be more than 27-fold under-expressed in the tumors. Kinase, a molecule involved in the activation of the cellular contact system, recently has been shown to be an inhibitor of angiogenesis (23). Its down-regulation may concur with up-regulation of VEGF, resulting in hypervascularization, a characteristic of clear cell RCC. We also found the metallothionein (MT) family to be coordinately down-regulated. Metallothionein (MT) is known to modulate the release of gaseous mediators such as hydroxyl radicals or nitric oxide and the binding and exchange of heavy metals such as zinc, cadmium, or copper. Differential expression of this family of genes has been reported in many cancers (24), and several subtypes (MT-1A, MT-1G and MT-1H) were reported to be down-regulated in RCCs (25, 26). Our study supports these reports, and additionally, found MT-1L and MT-1E to be down-regulated. Aldolase B, one of the three aldolase glycolytic enzyme catalyzing the reversible conversion of fructose-1, 6-bisphosphate to glyceraldehyde 3-phosphate and dihydroxyacetone phosphate, is another gene under-expressed in RCC. It has been found in abundance in normal renal cortex, suggesting that it plays a physiological role in normal kidney (27). The relatively lower expression in the cancer, as confirmed by our study, may be due to displacement of normal tissue. One of the heparan sulfate proteoglycans, glypican 3, also stood out in our analysis but has never been associated with kidney cancer. Its down-regulation has been reported in mesotheliomas, ovarian cancer, and breast cancer (28). Glypican 3-deficient mice have been shown to exhibit several clinical features including developmental overgrowth and dysplastic kidneys (29).

GENE EXPRESSION PROFILES AND PATIENT OUTCOME OF CLEAR CELL RCC

Several studies of gene expression profiles of clear cell RCC have been reported in conjunction with clinical parameters, especially patient outcome. We conducted the gene expression profiling of 29 clear cell RCC with diverse clinical outcome using home-made microarrays with 21,632 cDNA spots (30). Hierarchical clustering (31) to group both genes and tumors by similarity in expression pattern without any clinical data, which is called an unsupervised method, showed two main clusters of clear cell RCC, largely correlated with cause-specific survival at five years. Close correlation of gene expression patterns of clear cell RCC with survival encouraged us to identify gene sets to distinguish between clear cell RCC subgroups with different outcome. We used the program Cluster Identification Tool (CIT) (32) to identify and rank sub-clusters of genes. A discrimination score (DS) based on variation between the two groups and variation within each group was used to identify the best gene cluster in that program (19). The best gene cluster was comprised of genes over-expressed in the tumors with a good clinical outcome and genes under-expressed in tumors with a poor clinical outcome (Figure 2). The diversity in the expressions of the best gene cluster largely defined two patient groups that were distinguishable by cause-specific survival at five years. These findings may reflect the existence of distinct subclasses of clear cell RCC that differ in clinical behavior, which are an aggressive and a non-aggressive class.

Kaplan-Meier survival analysis was also examined with parameters of stage, grade, and the gene expression profile and tested by the log-rank test. Classification by grade (p<0.0001) was better than that by stage (p=0.0049). Just as significant as grade, the gene expression profile is also more accurate in predicting clinical outcome than staging alone (Figure 3). Correlation of histological grade or
Figure 2. Clustering of the 51 clones (about 40 genes) using non-median centered values. In this case, the color of each square corresponds to actual normalized gene expression level relative to normal kidney tissue. (A) Genes mostly over-expressed in tumors with the good outcome. (B) Genes mostly under-expressed in tumors with the poor outcome. This figure is cited from reference (30).

Figure 3. Cause specific survival curve with staging (a), histological grading (b), gene expression profiling in all patients (c), and patients with stage III/IV (d) by Kaplan-Meier method. This figure is cited from reference (30).
stage with the gene expression profile was analyzed as the Spearman correlation coefficient by the exact test with the SAS/STAT analysis package (version 8.0, SAS Institute, Inc., Cary, NC). It turned out that histological grade and the classification by the gene expression profile were highly correlated (correlation coefficient = 0.7703, p < 0.0001), indicating that grading is the clinical parameter most closely linked to the gene expression profile.

As looking at each gene in the best gene cluster that could distinguish the outcome of clear cell RCC patients, many of them provide insights into the biology of the clear cell RCC with different outcome. For example, transforming growth factor β receptor II (TGFβRII) and its down-stream effector, tissue inhibitor of metalloproteinase 3 (TIMP3), were exclusively down regulated in the poor outcome group. Loss of the TGFβII signaling pathway previously has been shown to be important for the development of aggressive cancers (33), and loss of TIMP3 expression by promoter methylation has been shown to increase tumorigenicity due to unregulated MMPs (34). A recent study demonstrated the inhibition of invasion in melanoma cell lines by over-expressing TIMP 3 by adenovirus-mediated gene delivery (35). The regulator of G-protein signaling 5 was exclusively over-expressed in the good outcome tumors and likewise may be important for the control of cancer progression.

The identification of this and other pathways, which is over-expressed or under-expressed in aggressive clear cell RCC, might lead to the discovery of potential molecular targeting medicine.

THE VON HIPPEL-LINDAU DISEASE (VHL) GENE ALTERATION AND CLEAR CELL RCC PATIENT OUTCOME

Intragenic mutation and hypermethylation of the VHL gene are found in about 80% of sporadic clear cell RCC (36-38). Alteration of the VHL gene can cause accumulation of hypoxia-inducible factors (HIF), leading to up-regulation of angiogenesis-related genes. Generally, high expressions of these genes are correlated to poor prognosis. Recently it has been reported that VHL alterations were strongly associated with better cancer-specific survival for clear cell RCC patients with stage I-III treated by radical nephrectomy. Although this finding seems to be interesting and thought provoking, additional studies are warranted to confirm the findings.

MOLECULAR SUBCLASSIFICATION OF PAPILLARY RCC

Papillary RCC is histologically characterized by the presence of fibrovascular cores with tumor cells arranged in a papillary configuration. The majority of papillary RCC show indolent behavior and have a limited risk of progression and mortality while a distinct subset displays highly aggressive behavior (39). Delahunt have proposed that papillary RCC can be morphologically classified into two subtypes (40). Type 1 is characterized by the presence of small cuboidal cells covering thin papillae, with a single line of small uniform nuclei and basophilic cytoplasm. Type 2 is characterized by the presence of large tumor cells with eosinophilic cytoplasm and pseudostratification. Generally types 2 papillary RCCs have a poorer prognosis than type 1 tumors. Yang et al. conducted the gene expression profiling of 34 cases of papillary RCC to elucidate the molecular background of papillary RCC with diverse outcome. Gene expression profiling of papillary RCC showed the two highly distinct subclasses with morphologic correlation. Class 1 corresponded to three histological subtypes, type 1, low grade type 2 and mixed type of type 1/low grade type 2, having excellent survival. Class 2 corresponded to high grade type 2 with poorer outcome. A seven transcript predictor was identified to classify papillary RCC into their new classification of class 1 and class 2 with 97% accuracy. They also selected two genes for further immunohistochemical study and confirmed high expression of cytokeratin 7 in class 1 papillary RCC and of topoisomerase IIα in class II tumors. These markers of immunohistochemistry for papillary RCC could be used to identify papillary RCC with poor outcome in the clinical settings. Interestingly, topoisomerase IIα is reported to be highly expressed in Wilms tumors (41). Considering the successful chemotherapy using topoisomerase II inhibitors in treating Wilms tumors, the similar chemotherapy regimens for class 2 papillary RCC might prove effective.

OTHER SUBTYPES OF RCC

It has been reported that c-Kit is highly ex-
pressed in chromophobe RCC, but not in normal kidney or clear cell RCC (42). As a tyrosine kinase inhibitor, Gleevec has been developed and effective for the treatment of chronic leukemia and gastrointestinal stromal tumors (43), which have high expression of c-Kit. Gleevec might also have some effect for the treatment for chromophobe RCC. In medullary RCC, which is very rare entity and has poor prognosis, it has been shown that topoisomerase IIα is highly expressed (44). Topoisomerase inhibitor might be effective for medullary RCC like Wilms tumors.

CONCLUSION

With increasing number of microarray studies of RCC, the molecular signature of each subtype of RCC has been gradually elucidated. The idea of molecular sub-classification of kidney tumors based on previous gene expression profiling is shown in Figure 4. Each of the signatures appears to be different and the future treatment strategy for RCC will not be uniform and one needs to consider the histological subtype and the genotype. Even in the same histological type of RCC, such as clear cell RCC and papillary RCC, further molecular subclassification to identify different prognosis can be achieved by gene expression profiling. Based on its molecular background, RCC management should move toward individualized therapies in the future. However, since genetic alterations in cancer are complex and usually involves multiple pathways, combined therapy including immunotherapy will be needed instead of monotherapy targeting one particularly altered gene in each subtype of RCC.

![Molecular sub-classification of kidney tumors](image)

**Figure 4.** The idea of molecular sub-classification of kidney tumors is shown.
REFERENCES

In vivo HFOF FYQSFTTJPO QSPGJMJOH PG DMFBS DFMM SFOBM DFMM DBSDJOPNBHFOF JEFOUJGJDBUJPO BOE QSPHOPTUJD DMBT

10: 537-544, 1997