

Molecular subclassification of kidney tumors and the discovery of new diagnostic markers

Masayuki Takahashi^{1,2}, Ximing J Yang^{3,4}, Jun Sugimura^{1,11}, Jesper Backdahl¹, Maria Tretiakova³, Chao-Nan Qian¹, Steven G Gray¹, Robert Knapp⁵, John Anema⁶, Richard Kahnoski⁶, David Nicol⁷, Nicholas J Vogelzang^{8,4}, Kyle A Furge⁹, Hiroomi Kanayama¹⁰, Susumu Kagawa¹⁰, Bin Tean Teh^{1,*}

¹Laboratory of Cancer Genetics, Van Andel Research Institute, 333 Bostwick NE, Grand Rapids, MI 49503, USA; ²Department of Urology, School of Medicine, The University of Tokushima, Tokushima 770-8503, Japan; ³Department of Pathology, University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637, USA; ⁴Department of Surgery/Urology, University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637, USA; ⁵Department of Pathology, Spectrum Health Hospital, Michigan St, NE, Grand Rapids, MI 49503; ⁶Division of Urology, Spectrum Health Hospital, Michigan St NE, Grand Rapids, MI 49503, USA; ⁷Department of Urology, Princess Alexandra Hospital, Queensland, 4102 Australia; ⁸Department of Medicine, University of Chicago, 5841 South Maryland Avenue, Chicago IL 60637, USA; ⁹Bioinformatics Program, Van Andel Research Institute, 333 Bostwick NE, Grand Rapids, MI 49503, USA; ¹⁰Department of Urology, School of Medicine, The University of Tokushima, Tokushima 770-8503, Japan; ¹¹Department of Urology, School of Medicine, Iwate medical University, Morioka 020-8505, Japan

We analysed the expression profiles of 70 kidney tumors of different histological subtypes to determine if these subgroups can be distinguished by their gene expression profiles, and to gain insights into the molecular mechanisms underlying each subtype. In all, 39 clear cell renal cell carcinomas (RCC), seven primary and one metastatic papillary RCC, six granular RCC from old classification, five chromophobe RCC, five sarcomatoid RCC, two oncocytomas, three transitional cell carcinomas (TCC) of the renal pelvis and five Wilms' tumors were compared with noncancerous kidney tissues using microarrays containing 19 968 cDNAs. Based on global gene clustering of 3560 selected cDNAs, we found distinct molecular signatures in clear cell, papillary, chromophobe RCC/oncocytoma, TCC and Wilms' subtypes. The close clustering in each of these subtypes points to different tumorigenic pathways as reflected by their histological characteristics. In the clear cell RCC clustering, two subgroups emerged that correlated with clinical outcomes, confirming the potential use of gene expression signatures as a predictor of survival. In the so-called granular cell RCC (terminology for a subtype that is no longer preferred), none of the six cases clusters together, supporting the current view that they do not represent a single entity. Blinded histological re-evaluation of four cases of 'granular RCC' led to their reassignment to other existing histological subtypes, each compatible with our molecular classification. Finally, we found gene sets specific to each subtype. In order to establish the use of some of these genes as novel subtype markers, we selected four genes and performed immunohistochemical analysis on 40 cases of primary kidney tumors. The results were consistent with the gene expression microarray data: glutathione *S*-transferase α was highly

expressed in clear cell RCC, α methylacyl racemase in papillary RCC, carbonic anhydrase II in chromophobe RCC and K19 in TCC. In conclusion, we demonstrated that molecular profiles of kidney cancers closely correlated with their histological subtypes. We have also identified in these subtypes differentially expressed genes that could have important diagnostic and therapeutic implications.

Oncogene (2003) 22, 6810–6818. doi:10.1038/sj.onc.1206869

Keywords: kidney tumor; microarray; gene expression profiling; molecular classification; diagnosis

Introduction

Renal cell carcinoma (RCC) is the most common malignancy of the adult kidney, representing 2% of all malignancies and 2% of cancer-related deaths. The incidence of RCC is increasing, a trend that cannot be explained by the increased use of abdominal imaging procedures alone (Chow *et al.*, 1999).

RCC is a clinicopathologically heterogeneous disease subdivided into clear cell, papillary, chromophobe, collecting duct and unclassified (WHO International Histological Classification of Kidney Tumors; Mostfi and Davis, 1998). Clear cell RCC is the most common adult renal neoplasm representing 70% of all renal neoplasms, and is thought to originate in the proximal tubules. Papillary RCC accounts for 10–15%, chromophobe RCC 4–6%, collecting duct carcinoma <1% and unclassified lesions 4–5% of RCC. Some of these tumors show sarcomatoid change (sometimes called spindle RCC or sarcomatoid RCC) and are thought to represent the high-grade end of all subgroups. Finally, the granular cell RCC, an old term that is no longer

*Correspondence: BT Teh; E-mail: bin.teh@vai.org
Received 19 February 2003; revised 19 June 2003; accepted 24 June 2003

included in the current classification systems, is still being diagnosed by many pathologists around the world. Instead, granular RCC can often be reclassified into other subtypes (Storkel *et al.*, 1997).

With recent advances in molecular genetics, the subtypes of RCC have been associated with distinct genetic abnormalities. This association has led to a proposal for molecular diagnosis of RCC (Bugert and Kovacs, 1996). The majority of clear cell RCC, for example, has a loss of chromosome 3 and inactivating mutations of the *VHL* gene, whereas papillary RCC are frequently associated with trisomy of chromosomes 3q, 7, 12, 16, 17 and 20, and loss of the Y chromosome. A portion of them also harbor *MET* mutations. It has been proposed that, even in the absence of prominent papillae, these aberrant chromosomal features could support the diagnosis of papillary RCC. Conversely, kidney cancers that do not possess these genetic characteristics should not be designated as papillary RCC even when papillary structures are prominent (Storkel *et al.*, 1997). Frequent loss of sex chromosomes, chromosomes 1 and 14 have been found in renal oncocytoma, a benign tumor composed of large eosinophilic cells arranged in acini (Presti *et al.*, 1996). Accurate subtyping of renal tumors is critical for predicting prognosis and designing appropriate treatment for patients (Motzer *et al.*, 2002).

To date, microarray technology has provided comprehensive insights into the underlying molecular mechanisms of many types of cancers. These gene expression profiles can serve as the molecular signatures of cancer, and may be used to distinguish among histological subtypes as well as to aid in the discovery of novel clinical subtypes such as those related to drug response. These distinctions may reflect the heterogeneity in transformation mechanisms, cell types, and behavior of the tumors. For example, several studies have identified distinct gene sets that distinguish serous and mucinous ovarian cancer (Ono *et al.*, 2000), acute myeloid and acute lymphoblastic leukemias (Golub *et al.*, 1999), *BRCA1* and *BRCA2* hereditary breast cancer (Hedenfalk *et al.*, 2001), hepatitis-B and hepatitis-C hepatocellular carcinomas (Okabe *et al.*, 2001) and diffuse large B-cell lymphomas with good and poor prognosis (Rosenwald *et al.*, 2002).

In kidney tumors, distinctive molecular signatures for clear cell RCC and chromophobe RCC/oncocytoma were previously shown (Boer *et al.*, 2001; Young *et al.*, 2001). More recently, we identified a prognostic gene set in clear cell RCC that could distinguish between patients with good or poor outcome (Takahashi *et al.*, 2001). Obviously, there is a need to extend the profiling to include other subtypes of kidney tumors.

In this study, we examined the molecular signatures of 70 kidney tumors from seven different subgroups: 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. Our aims were twofolds. First, we intended to establish a molecular subclassification based on expression profiles. We then tested its

potential use by including six cases of 'granular cell RCC' which were diagnosed prior to the recommendation by the work group of UICC (Union Internationale Contre le Cancer) and AJCC (American Joint Committee on Cancer) (Storkel *et al.*, 1997). This is to determine if their recommendation was supported by gene expression-based molecular subclassification. Our second aim was to identify novel genes that may serve as novel diagnostic markers. The strength of the candidacy as new diagnostic markers lay in the number of subtypes included in the studies. Using statistical method, we could select the genes that were unique to one subgroup and thus were able to distinguish against other subgroups. We then established their potential use in diagnosis by performing immunohistochemical staining on paraffin-embedded blocks of primary kidney tumors using antibodies of four selected genes which have corresponding monoclonal antibody available.

Results

We used hierarchical clustering (Eisen *et al.*, 1998) to classify kidney tumors based on their gene expression profiles using the expression ratios of a selected 3560 cDNA set (see Materials and methods). The clustering algorithm groups both genes and tumors by similarity in expression pattern. The patient dendrogram, which is based on expression profile of all 3560 cDNAs, is shown in Figure 1. The gene expression pattern below the dendrogram (Figure 1) was based on 1309 genes that were significantly unique to each subtype. Two broad clusters emerged: (1) one consisting of 35 clear cell RCC and four granular RCC, and (2) all other types of kidney tumors and four clear cell RCC. In the latter group, there are different subclusters; five chromophobe RCC and two oncocytoma, seven primary and one metastatic papillary RCC, five Wilms' tumors, and three TCC. In the first main cluster of clear cell RCC, there are two subclusters: one including all patients (except one) who died of cancer (Figure 1e) and the other the survivors of cancer without evidence of metastasis (Figure 1d). Two papillary RCC, one primary tumor and a metastasized lymph node, from the same patient were also examined (papillary 7P, 7M) and were found to have very similar expression pattern, pointing to the genealogical relationship between the primary and metastatic tumor (Haddad *et al.*, 2002). A set of more highly expressed genes in each subtype of tumors compared to all other types of kidney tumors studied is indicated by side bars with different colors on the right-hand side of Figure 1 (A, chromophobe RCC; B, papillary RCC; C, Wilms' tumors; D, clear cell RCC with good outcome; E, all cases of clear cell RCC).

Six granular cell RCC did not cluster by themselves but instead were located in a 'random' fashion, indicating that they are not a single entity (Figure 1). The diagnoses of these six cases were made in Japan prior to the recommendation of the work group of UICC and AJCC for RCC diagnosis. The old histology slides of five of these cases were available which were

Molecular Subclassification of Kidney Tumors

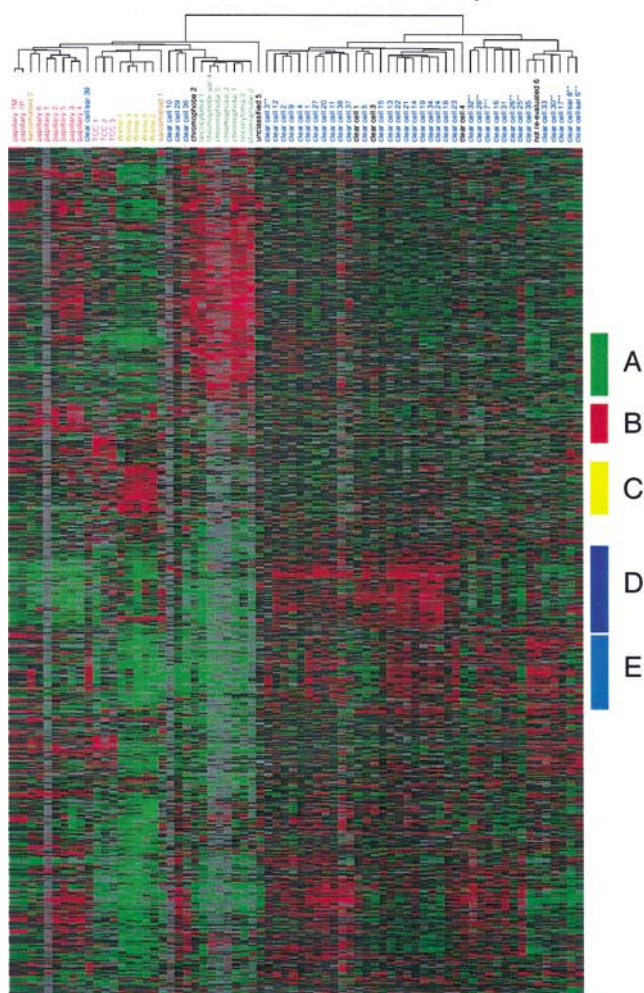


Figure 1 Clustering of 70 kidney tumors. The clustering of patients (using Pearson's correlation) is based on global gene expression profiles consisting of median polished data of 3560 selected spots. Rows represent individual cDNAs and columns represent individual tumor samples. The color of each square represents the median-polished, normalized ratio of gene expression in a tumor relative to reference. Expression levels greater than the median are shaded in red, those below the median in green; black, equal to the median; and gray, inadequate or missing data. The color saturation indicates the degree of divergence from the median. The tumors clustered into two broad groups with one group consisting of primarily clear cell RCC and the other consisting of all other kidney tumors. Five chromophobe RCC and two oncocytoma are clustered close together. Each group of eight papillary RCC, five Wilms' tumors, or three TCC is clustered together. The six cases in black represent those cases that were originally diagnosed as granular cell RCC. The diagnosis after re-evaluation is indicated here except case 6 (histological slide not available for re-evaluation). Case A set of the most highly expressed genes in each subtype of tumors compared to all other types of kidney tumors studied is shown by using colored side bars to the right of the image (a, chromophobe RCC; b, papillary RCC; c, Wilms' tumors; d, clear cell RCC with good outcome; e, clear cell RCC)

re-evaluated in a blinded fashion by an expert urologic pathologist (XJY). 'Granular RCC 1, 3 and 4', which were clustered in clear cell RCC group, were reclassified as clear cell RCC. 'Granular 2', which was closely clustered with chromophobe RCC, was reassigned as a

chromophobe RCC. 'Granular 5', which has a distinct histology was not clustered with any RCC group by genetic profiles. It is labeled as unclassified which may represent a novel subtype of RCC and warrant further investigation (Figure 1).

We then used multidimensional scaling (MDS) to visualize the relationship among the profiles of all tumors. Three-dimensional (3d) visualization of the MDS data demonstrated how each RCC subtype clustered, for example, chromophobe RCC/oncocytoma, papillary RCC, Wilms' tumors and TCC (Figure 2a). 'Granular 5', which was of aggressive type and could not be reclassified, was placed next to the sarcomatoid RCC. In addition, a large majority of clear cell RCC with poor outcome clustered to one side suggesting that they shared similar expression profiles (Figure 2b).

We then identified differentially expressed genes which are unique in each subtype (Tables 1, 2 and 3) and performed immunohistochemical staining on selected genes from the lists. Routine Hematoxylin and eosin (H&E) staining (Figure 3a–d), immunostaining for glutathione *S*-transferase alpha (GST-alpha, Figures 3E–H), alpha methylacyl Coenzyme racemase (AMACR, Figures 3i–l) and carbonic anhydrase II (CAII, Figures 3m–p) were performed with benign kidney tissues (Figure 3a, e, i and m), clear cell (Figure 3b, f, j and n), papillary (Figure 3c, g, k and o) as well as chromophobe RCC (Figure 3d, h, l and P).

In the non-neoplastic renal cortex (Figure 3a), GST- α and AMACR were present in the proximal convoluted tubules (Figure 3e, i, respectively), but absent in distal convoluted tubules, glomeruli as well as stromal cells. CAII, however, was detected in the distal convoluted tubules, but very minimal in distal tubules, glomeruli, medullary tubules and stromal cells (Figure 3m).

Nine of 10 clear cell RCCs (90%) showed immunoreactivity to GST- α (Figure 3f), while none of 10 papillary RCC (0%) and only one of 10 chromophobe RCCs (10%) were reactive to GST- α . All 10 papillary RCCs (100%) were strongly positive for AMACR (Figure 3k), while one of 10 (10%) clear cell RCCs and none of 10 chromophobe RCCs were positive for AMACR. All 10 chromophobe RCC (100%) showed reactivity to CAII, while only three of 10 clear cell RCCs, one of 10 papillary RCCs and two of 10 transitional cell carcinomas of the renal pelvis showed reactivity to CAII staining. All five of five transitional cell carcinomas were positive for keratin 19, while none of 10 clear cell RCCs, one of 10 papillary RCCs and none of chromophobe RCCs were positive for keratin 19 staining. The results are summarized in Table 4. The findings confirmed the presence of these overexpressed genes located in tumor cells not in stromal cells.

Discussion

We established that different subtypes of kidney tumors are indeed characterized by distinct molecular signatures as are many other types of cancer. These reflect the

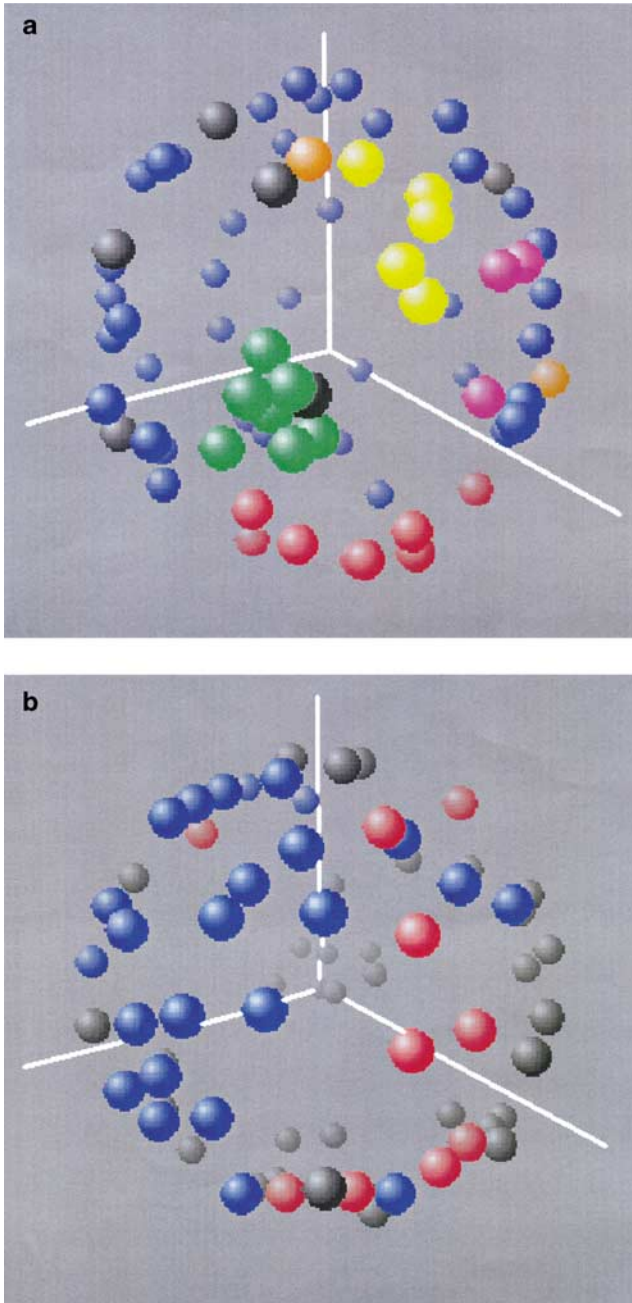


Figure 2 3D imaging. (a) Each subtype of kidney tumors is represented by a different color: chromophobe RCC/oncocytoma, green; papillary RCC, red; clear cell RCC, blue; TCC, purple; sarcomatoid RCC, orange; granular cell RCC, black. Five chromophobe RCC and two oncocytoma clustered close together. The eight papillary RCC, five Wilms' tumors and three TCC, clustered close together. Clear cell RCC on the other hand looked more scattered than in 2D clustering by TreeView. (b) All tumors with a focus on clear cell RCC whose outcome data were available. Patient who survive more than 5 years after surgery are represented by blue, whereas those who died of cancer within 5 years after surgery are represented by red. All the rest of samples are assigned black. Clear cell RCCs with poor outcome are located on one side and those with good outcome on the other

difference in the cell type, biology and underlying molecular mechanisms in these subtypes. This is very important especially from the point of management.

Immunohistochemistry of Renal Cell Carcinoma

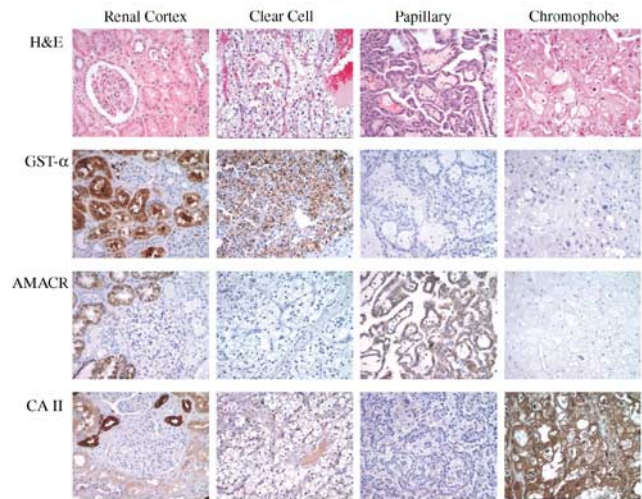


Figure 3 Immunohistochemistry of renal tumors. Hematoxylin and eosin stain (H&E, a–d) and immunostaining for glutathione S-transferase alpha (GST- α , f–h), α methylacyl CoA racemes (i–l), and carbonic anhydrase II (CAII, m–p) are demonstrated in normal renal cortex (a, e, i and m), clear cell RCC (b, f, j and n), papillary RCC (c, g, k and o) and chromophobe RCC (d, h, l and p). Clear cell RCC (f) showing strong immunoreactivity for GST- α , papillary RCC showing strong immunoreactivity for AMACR in (k), chromophobe RCC showing strong immunoreactivity for CA II in (p)

First of all, based on their distinct molecular signatures, the prognostic factors, which again can be reflected by gene expression profiles (see below), have to be determined separately in each subtype instead of by lumping all subtypes together. This most likely hold true in treatment too. The low rate of response (5–10%) for some of the trailed chemotherapy like interferon and IL-2 point to the heterogeneous nature of kidney cancers treated. It is clear that, for example, non-clear-cell metastatic RCC has a high resistance to systemic therapy and poor survival (Motzer *et al.*, 2002). And even within the subtype such as the clear cell RCC, in the present studies there are at least two major subgroups: one group is characterized by deaths within 5 years and metastasis while the other consist of patients surviving more than 5 years without evidence of metastasis. This has further supported our previous study describing the use of expression profiles as an indicator of survival in clear cell RCC (Takahashi *et al.*, 2001). We also demonstrated the potential gene expression-based subclassification by studying six cases of the so-called granular cell RCC. These were old cases whose diagnoses were made prior to 1997, where the WHO classification was introduced. Although no longer supposed to be used, this diagnosis is still widely made by some pathologists. Our findings clearly support that it is not an entity. Our gene expression-based subclassification designated these cases to either chromophobe RCC or clear cell RCC. In five cases whereby old histological subtypes were available for re-evaluation, the diagnosis was in agreement with our subclassification in four of them. In one case (granular 5), it was not clustered with any subgroups

Table 1 Relatively more highly expressed genes in clear cell RCC

Accession ID	Gene name	Fold change	P-value
T54298	PPAR (γ) angiopoietin-related protein (PGAR)	18.3	0.0001
H95633	Crystallin, α A	16.5	0.0001
T73468	glutathione S-transferase A2	11.4	0.0001
N59772	ESTs	9.9	0.0001
AA664406	Complement component 4A	9.7	0.0001
AA668470	Regulator of G-protein signaling 5	8.8	0.0001
AA169469	Pyruvate dehydrogenase kinase, isoenzyme 4	8.4	0.0001
AA700054	Adipose differentiation-related protein	8.0	0.0001
H18608	ESTs, highly similar to organic anion transporter 3	7.9	0.0001
AA150532	Keratin 6A	7.6	0.0001
H09076	Cytochrome P450, subfamily III polypeptide 2	7.4	0.0001
AA136707	Procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2	7.2	0.0001
W72294	Small inducible cytokine subfamily B, member 14	7.1	0.0001
N30096	Glutathione S-transferase A3	6.6	0.0002
AA454159	Homo sapiens HRBPiso mRNA, complete cds	6.4	0.0001
AA017544	Regulator of G-protein signaling 1	6.3	0.0001
AA102107	Glutamyl aminopeptidase (aminopeptidase A)	6.3	0.0001
AA488070	Immunoglobulin κ constant	6.2	0.0002
N92646	Colony stimulating factor 2 receptor, α , low-affinity	6.2	0.0001
N93191	Homo sapiens cDNA: FLJ22811 fis, clone KAIA2944	6.1	0.0001
R50354	Leukemia inhibitory factor (cholinergic differentiation factor)	5.9	0.0001
AA432292	Hypothetical protein DKFZp434F0318	5.8	0.0001
T67053	Immunoglobulin λ locus	5.7	0.0001
AA486082	Serum/glucocorticoid-regulated kinase	5.6	0.0001
AA598601	Insulin-like growth factor binding protein 3	5.6	0.0001
N58170	Kidney- and liver-specific gene	5.6	0.0002
H15366	ESTs	5.3	0.0001
H88329	Calbindin 1 (28 kDa)	5.2	0.0001
H38650	Solute carrier family 2, member 5	5.1	0.0001
R45059	Vascular endothelial growth factor (VEGF)	5.1	0.0001

Top 30 differential expressed cDNA in clear cell RCC are listed. They are significantly more highly expressed in clear cell RCC compared to all other types of kidney tumors studied by 10 000 times of permutation test. Fold change indicates clear cell RCC have relatively higher expression of this fold change compared to all other types of kidney tumors studied

and indeed re-evaluation of the slide by our expert pathologist also could not pinpoint its diagnosis. Therefore, this case may represent a novel subtype of RCC yet to be properly characterized and described. This also illustrates the potential benefits and future direction of pathology whereby the corroboration of conventional pathologist expertise/skills and expression profiling may complement each other and will be most helpful in equivocal or non-clear-cut cases.

In conjunction with this, we also used statistical method to identify potential diagnostic markers in each subtype of tumor from the expression profiles. Here, we are interested in genes that are not only differentially expressed in each subtype but also have the highest power to distinguish from one another (Tables 1, 2 and 3). For example, the genes that we select here for clear cell RCC are only unique to clear cell subtype and not to the other subtypes. So some of the genes that appear in our previous study on solely clear cell RCC (Takahashi *et al.*, 2001) are not present here. Using antibodies corresponding to the four selected genes, corresponding to the antibody available, we performed immunohistochemical staining on 40 cases of primary tumors: 10 clear cell RCC, 10 papillary RCC, 10 chromophobe RCC, five oncocytoma and five transitional cell carcinoma. The results were consistent with

microarray expression profiles (Table 4; Figure 3), confirming their potential use as novel diagnostic markers. Naturally, higher specificity and sensitivity can probably be found in other candidates that are yet to have corresponding antibodies and would require further studies to establish their use as diagnostic markers.

The functions of the majority of these genes and how they are associated with tumorigenesis in each subtype remain largely unknown. For example, the first gene on which we performed immunohistochemical staining is GST- α , which functions to protect cell by catalysing the detoxification of xenobiotics and carcinogens. Previous immunohistochemical studies have demonstrated strong expression in normal kidney, especially in the proximal tubules as well as in kidney cancer (Grignon *et al.*, 1994). However, its role in clear cell RCC remains unknown. The second gene studied is the *AMACR* gene, which plays a critical role in peroxisomal beta oxidation of branched-chain fatty acid molecules. *AMACR* has recently been shown overexpressed in prostate cancers at both the transcript level by microarray experiments and at the protein level (Xu *et al.*, 2000; Luo *et al.*, 2002; Rubin *et al.*, 2002). Further studies by immunohistochemistry have demonstrated the elevation of *AMACR* protein in more than 90% of prostate cancer cases but not

Table 2 Relatively more highly expressed genes in papillary RCC

Accession ID	Gene name	Fold change	P-value
R60170	Guanine deaminase	18.0	0.0002
W85851	Homo sapiens chromosome 16 BAC clone	10.6	0.0002
H86812	Heparan sulfate (glucosamine) 3-O-sulfotransferase 1	7.9	0.0001
AA496334	Dynamin 1	7.7	0.0001
AA873159	Apolipoprotein C-I	6.8	0.0003
AA459296	Solute carrier family 34, member 2	6.5	0.0001
AA451904	Epididymis-specific, whey-acidic protein type	6.4	0.0004
R93124	Aldo-keto reductase family 1, member C1	5.7	0.0003
AA135886	Homo sapiens mRNA; cDNA DKFZp434F053	5.5	0.0001
AA127965	Integrin, β 8	5.3	0.0002
AA453310	α -methylacyl-CoA racemase	5.2	0.0001
AA916325	Aldo-keto reductase family 1, member C3	5.0	0.0004
AA478724	Insulin-like growth factor binding protein 6	4.9	0.0001
AA416585	Angiotensin I-converting enzyme 2	4.8	0.0002
R51836	Homo sapiens clone CDABP0036 mRNA sequence	4.6	0.0002
AA430665	Claudin 4	4.5	0.0002
AA456022	Fibronectin leucine-rich transmembrane protein 3	4.5	0.0003
AA664101	Aldehyde dehydrogenase 1 family, member A1	3.9	0.0096
R35051	ESTs	3.9	0.0001
AA704995	Putative glycine-N-acyltransferase	3.8	0.0066
AA757672	ESTs	3.8	0.0001
AA464688	ESTs, weakly similar to unnamed protein product	3.7	0.0001
AA292226	Accessory proteins BAP31/BAP29	3.6	0.0055
AA437099	ESTs	3.6	0.0002
AA406126	Nit protein 2	3.5	0.0001
AA489246	Suppression of tumorigenicity 14	3.5	0.0029
H69786	Homo sapiens MAIL mRNA, complete cds	3.5	0.0018
T94781	Potassium inwardly -rectifying channel, subfamily J, member 15	3.5	0.0040
AA455632	Chromosome 3p21.1 gene sequence	3.4	0.0070
AA644088	Cathepsin C	3.3	0.0006

Top 30 differential expressed cDNA in papillary RCC are listed. They are significantly more highly expressed in papillary RCC compared to all other types of kidney tumors studied by 10 000 times of permutation test. Fold change indicates papillary RCC have relatively higher expression of this fold change compared to all other types of kidney tumors studied

in benign prostatic tissues, suggesting that *AMACR* may be a more specific marker than prostate-specific antigen (PSA) for prostate cancer (Jiang *et al.*, 2001; Luo *et al.*, 2002; Rubin *et al.*, 2002). This gene has never been associated with any subtype of RCC. It was 5.3 times more highly expressed in papillary RCC than in normal kidney. In addition, immunohistochemical analysis demonstrated immunoreactivity in 100% of papillary RCC cases, and in less than 10% of other subtypes of RCC (Figure 3e–h). Obviously its role in papillary RCC remains to be elucidated and warrant further studies.

The third gene we studied is CA II, which belongs to a family of zinc metalloenzymes. Our immunostaining demonstrated positivity in all chromophobe RCC (10/10) and oncocytomas (5/5). Several studies have looked at its expression in kidney as well as the phenotypes in CAII-null mice, but its role in RCC is yet to be elucidated (Brechue *et al.*, 1991; Lai *et al.*, 1998; Lewis *et al.* 1988). Interestingly, the profiles we found for this group of tumors are similar to those previously described (Young *et al.*, 2001). The chromophobe RCC/oncocytomas contain abundant mitochondria and had overexpression of genes related to mitochondrial biology and oxidative phosphorylation (Table 3), suggesting the high

specificity of these genes to chromophobe RCC/oncocytomas.

TCC arising in the renal pelvis may invade the entire kidney, and as such it may be difficult to distinguish TCC from RCC. Finding new markers for TCC may facilitate its diagnosis. By immunohistochemistry, we found K19 expression in some renal tubules, benign transitional epithelium and in all five cases of TCC (Table 4). Several genes that were highly specific for TCC which are related to skin. K19 is one of them and it has been found in the periderm, the transient superficial layer that envelops the developing epidermis (Van Muijen *et al.*, 1987). However, its association with TCC has never been described and further studies will be required to understand its role in RCC.

Obviously, it is beyond the scope and aims of this study to discuss all the genes we have identified in each subtype of these tumors. However, it is very obvious that more value and potential of gene expression profiling are yet to be discovered. Larger number of tumors, especially in some subtypes, with detailed clinical information will be very valuable and worthwhile for future studies. For example, it may be possible to identify the genes that can differentiate between oncocytoma, a benign tumor and chromophobe RCC. It

Table 3 Relatively more highly expressed genes in chromophobe RCC/oncocytoma

Accession ID	Gene name	Fold change	P-value
H57180	Phospholipase C, $\gamma 2$	19.6	0.0001
H23187	Carbonic anhydrase II	13.8	0.0001
AA399633	ESTs	9.9	0.0001
N89673	PPAR, γ , coactivator 1	9.2	0.0001
W95082	Hydroxysteroid (11- β) dehydrogenase 2	9.0	0.0001
N93505	Transmembrane 4 superfamily member 2	8.9	0.0001
R59722	Hypothetical protein FLJ 10851	8.3	0.0011
T60160	Homo sapiens mRNA; cDNA	7.6	0.0001
H17036	DHHC1 protein	7.6	0.0001
AA446650	Homo sapiens mRNA; cDNA DKFZp586M0723	7.5	0.0001
R16134	Plasmalipin	7.2	0.0001
AA406233	ESTs, highly similar to GTPase-activating proteins	7.1	0.0001
T49816	ESTs	7.0	0.0001
H22944	Nicotinamide nucleotide transhydrogenase	6.9	0.0001
R43873	Human chromosome 16 BAC clone CIT987SK-A-101F10	6.8	0.0001
AA463445	Homolog of yeast ubiquitin-protein ligase Rsp5	6.7	0.0001
N54401	Rag D protein	6.5	0.0001
H22856	Glutamic-oxaloacetic transaminase 1, soluble	6.3	0.0001
R09053	ESTs	6.1	0.0001
AA406362	Prostaglandin E receptor 3 (subtype EP3)	6.1	0.0001
H97921	ESTs	6.0	0.0001
W31540	KIAA 1450 protein	5.9	0.0001
AA427619	1,2- α -mannosidase IC	5.9	0.0001
W47387	Ecotropic viral integration site 5	5.7	0.0004
N29800	Hypothetical protein FLJ20783	5.7	0.0001
H99738	Rag D protein	5.7	0.0001
AA894557	Creatine kinase, brain	5.7	0.0001
AA452566	Peroxisomal membrane protein 3 (35 kDa)	5.7	0.0001
AA504265	LIM and senescent cell antigen-like domains 1	5.6	0.0001
AA682684	Protein tyrosine phosphatase, nonreceptor type 3	5.5	0.0001

Top 30 differential expressed cDNA in are listed. They are significantly more highly expressed in chromophobe RCC/oncocytoma compared to all other types of kidney tumors studied by 10 000 times of permutation test. Fold change indicates chromophobe RCC/oncocytoma have relatively higher expression of this fold change compared to all other types of kidney tumors studied

Table 4 Immunostainoreactivity of four markers in 40 primary kidney tumors

Marker	Clear cell (10 cases)	Papillary (10 cases)	Chromophobe (10 cases)	Oncocytoma (five cases)	TCC (five cases)
GST- α	90%	0%	10%	0%	ND
AMACR	10%	100%	0%	0%	ND
CA II	30%	10%	100%	100%	20%
K19	0%	10%	0%	0%	100%

will also be feasible to identify the prognostic molecular signatures for other subtypes of RCC, besides clear cell RCC.

Materials and methods

Tumor samples and RNA preparation

A total of 69 frozen primary kidney tumors (39 clear cell RCC, seven papillary RCC, six granular RCC, five chromophobe RCC, two sarcomatoid RCC, two oncocytomas, three TCC and five Wilms' tumors), one metastatic papillary RCC and matched or unmatched noncancerous kidney tissue were obtained from the University of Tokushima, the University of Chicago, Spectrum Health Urologic Group and Cooperative Human Tissue Network (CHTN). Internal review board (IRB) approvals were obtained

from these institutions and the Van Andel Research Institute to study these tumor samples. All tissues were accompanied by pathology reports with or without clinical outcome information. The samples were anonymized prior to the study. Part of each tumor sample was frozen in liquid nitrogen immediately after surgery and stored at -80°C . Total RNA was isolated from the frozen tissues using ISOGEN solution (Nippon Gene, Toyama, Japan) or Trizol reagent (Invitrogen, Carlsbad, CA, USA). For the first 45 samples, poly(A)+RNA was isolated from the total RNA using the Oligotex mRNA Mini Kit (Qiagen, Valencia, CA, USA). For the remaining 25 samples, total RNA was purified with 2.5 M final concentration of lithium chloride. The WHO International Histological Classification of Tumors was used for histological evaluation of the specimens (Mostfi and Davis, 1998). Union Internationale Contre le Cancer (UICC) TNM classification and stage groupings were used (Sobin and Fleming, 1997).

cDNA microarray fabrication and procedures

Microarray production was performed as described (Eisen and Brown, 1999; Hegde *et al.*, 2000) with slight modifications. Briefly, 19 968 cDNA clones were PCR amplified directly from bacterial stocks purchased from Research Genetics (Huntsville, AL, USA). Following ethanol precipitation and transfer to 384 well plates, clones were printed onto aminosilane-coated glass slides using a custom-built robotic microarrayer. Slides were chemically blocked using succinic anhydride after UV light crosslinking. When available, cancers were hybridized against patient-matched noncancerous kidney tissue. For tumors without their matched noncancerous kidney tissue available, RNA from five noncancerous kidney tissues was mixed and pooled to serve as a common reference. For the first 45 samples, 2 μ g of poly(A) + RNA from tumors and reference were reverse transcribed with oligo (dT) primer and Superscript II (Invitrogen, Carlsbad, CA, USA) in the presence of Cy5-dCTP and Cy3-dCTP (Amersham Pharmacia Biotech, Peapack, NJ, USA). For the remaining 25 samples, 50 μ g of total RNA from tumors and reference were used for reverse transcription. The Cy5- and Cy3-labeled cDNA probes were mixed with probe hybridization solution containing formamide and hybridized to prewarmed (50°C) slides for 20 h at 50°C. Following hybridization, slides were washed in 1 \times SSC, 0.1% SDS at 50°C for 5 min followed by 0.2 \times SSC, 0.1% SDS at room temperature (RT) for 5 min, 0.2 \times SSC at RT for 5 min twice and 0.1 \times SSC at RT for 5 min. Slides were dried immediately by centrifugation and scanned using a Scan Array Lite scanner at 532 and 635 nm wavelengths (GSI Lumonics, Billerica, MA, USA).

Data analysis

Images were analysed using the software Genepix Pro 3.0 (Axon, Union City, CA, USA). The local background was subtracted for all spots. Spots whose background-subtracted intensities in either Cy5 or Cy3 channel were less than 150 were excluded from the analysis. The ratio of Cy5 intensity to Cy3 intensity was calculated for each spot, representing tumor RNA expression relative to noncancerous kidney tissue. Ratios were log transformed (base 2) and normalized so that the median log-transformed ratio equaled zero. Genes with the following criteria (3560 genes in total) were selected for the global clustering analysis: (1) expression values present in at least 70% of the tumors; (2) expression ratios that varied at least twofold in at least two tumors; and (3) maximum ratio minus minimum ratio values greater than twofold. The gene expression ratios were median polished across all samples. Gene expression values were manipulated and visualized using the CLUSTER and TREEVIEW software (M.B. Eisen, [http://](http://rana.lbl.gov)

rana.lbl.gov). The correlation distances were calculated as $1-r$, where r indicates the Pearson correlation coefficient (Eisen *et al.*, 1998).

The in-house software program, CIT, was used to find genes that were differentially expressed (using a Student's t -test) between one histological subtype and the others (Rhodes *et al.*, 2002). To find significant discriminating genes, 10 000 t -statistics were calculated by randomly placing patients into two groups (Hedenfalk *et al.*, 2001). A 99.9% significance threshold ($\alpha \leq 0.01$) was used to identify genes that could significantly distinguish between two patient groups versus the random patient groupings.

Immunohistochemistry

A total of 50 cases of benign ($N=10$) and neoplastic renal tissues ($N=40$) were analysed using immunohistochemistry. Kidney tumors included clear cell RCC ($N=10$), papillary RCC ($N=10$), chromophobe RCC ($N=10$), oncocytoma ($N=5$) and TCC ($N=5$). A section from each kidney tissue was stained for H&E to verify histology. The following antibodies were obtained commercially: GST- α (Dako, Carpinteria, CA, USA), α methylacyl racemase (Corixa, Seattle, WA, USA), CAII (Santa Cruz Biotechnology, Santa Cruz, CA, USA) and keratin 19 (Dako, Carpinteria, CA, USA). Standard biotin-avidin complex immunohistochemistry was performed. Briefly, tissue sections were then incubated with primary antibodies for 30 min at 20°C. Then, the slides were incubated with biotinylated anti-mouse IgG or anti-rabbit IgG (Vector Laboratories, Burlingame, CA, USA) at 27°C for 30 min and the antigen-antibody complexes were detected with avidin-biotinylated horseradish peroxidase system (Vector, Burlingame, CA, USA) using DAB as a chromogen and hematoxylin as a counterstain. The slides were evaluated as either negative or positive by an expert urologic pathologist (XJY).

Acknowledgements

We thank the Laboratory of DNA and Protein Microarray technology at Van Andel Research Institute (VARI) and Can Gong at the University of Chicago for their technical assistance. We want to acknowledge the contribution of some of the tumors studied by the Cooperative Human Tissue Network (CHTN). We also thank members of the Grand Rapids Urology Study Group including Ken Shockley, John Ludlow, David Kracklau, Philip Wise, Brian Roelof, Jon Curry and pathologists at the Spectrum Health and Metropolitan Hospitals. Finally, we thank Vivve Howell and Rick Hay for kindly reviewing the manuscript.

References

- Boer JM, Huber WK, Sultmann H, Wilmer F, von Heydebreck A, Haas S, Korn B, Gunawan B, Vente A, Fuzesi L, Vingron M and Poustka A. (2001). *Genome Res.*, **11**, 1861–1870.
- Brechue WF, Kinne-Saffran E, Kinne RK and Maren TH. (1991). *Biochim. Biophys. Acta*, **1066**, 201–207.
- Bugert P and Kovacs G. (1996). *Am. J. Pathol.*, **149**, 2081–2088.
- Chow WH, Devesa SS, Warren JL and Fraumeni Jr JF. (1999). *JAMA*, **281**, 1628–1631.
- Eisen MB and Brown PO. (1999). *Methods Enzymol.*, **303**, 179–205.
- Eisen MB, Spellman PT, Brown PO and Botstein D. (1998). *Proc. Natl. Acad. Sci. USA*, **95**, 14863–14868.
- Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD and Lander ES. (1999). *Science*, **286**, 531–537.
- Grignon DJ, Abdel-Malak M, Mertens WC, Sakr WA and Shepherd RR. (1994). *Mod. Pathol.*, **7**, 186–189.
- Haddad R, Furge KA, Miller J, Haab BB, Schoumans J, Teh BT, Barr LL and Web CP. (2002). *Appl. Genom. Proteom.*, **1**, 51–62.
- Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, Meltzer P, Gusterson B, Esteller M, Kallioniemi OP, Wilfond B, Borg A and Trent J. (2001). *N. Engl. J. Med.*, **344**, 539–548.

- Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R, Hughes JE, Snesrud E, Lee and Quackenbush J. (2000). *Biotechniques*, **29**, 548–562.
- Jiang Z, Woda BA, Rock KL, Xu Y, Savas L, Khan A, Pihan G, Cai F, Babcook JS, Rathanaswami P, Reed SG, Xu J and Fanger GR. (2001). *Am. J. Surg. Pathol.*, **25**, 1397–1404.
- Lai LW, Chan DM, Erickson RP, Hsu SJ and Lien YH. (1998). *J. Clin. Invest.*, **101**, 1320–1325.
- Lewis SE, Erickson RP, Barnett LB, Venta PJ and Tashian RE. (1988). *Proc. Natl. Acad. Sci. USA*, **85**, 1962–1966.
- Luo J, Zha S, Gage WR, Dunn TA, Hicks JL, Bennett CJ, Ewing CM, Platz EA, Ferdinandusse S, Wanders RJ, Trent JM, Isaacs WB and De Marzo AM. (2002). *Cancer Res.*, **62**, 2220–2226.
- Mostfi FK and Davis CJ, (eds) (1998). *WHO International Histological Classification of Tumours*, Springer: Berlin.
- Motzer RJ, Bacik J, Mariani T, Russo P, Mazumdar M and Reuter V. (2002). *J. Clin. Oncol.*, **20**, 2376–2381.
- Okabe H, Satoh S, Kato T, Kitahara O, Yanagawa R, Yamaoka Y, Tsunoda T, Furukawa and Nakamura Y. (2001). *Cancer Res.*, **61**, 2129–2137.
- Ono K, Tanaka T, Tsunoda T, Kitahara O, Kihara C, Okamoto A, Ochiai K, Takagi T and Nakamura Y. (2000). *Cancer Res.*, **60**, 5007–5011.
- Presti Jr JC, Moch H, Reuter VE, Huynh D and Waldman FM. (1996). *Genes Chromosomes Cancer*, **17**, 199–204.
- Rhodes DR, Miller JC, Haab BB and Furge KA. (2002). *Bioinformatics*, **18**, 205–206.
- Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, Gascoyne RD, Muller-Hermelink HK, Smeland EB, Giltnane JM, Hurt EM, Zhao H, Averett L, Yang L, Wilson WH, Jaffe ES, Simon R, Klausner RD, Powell J, Duffey PL, Longo DL, Greiner TC, Weisenburger DD, Sanger WG, Dave BJ, Lynch JC, Vose J, Armitage JO, Montserrat E, Lopez-Guillermo A, Grogan TM, Miller TP, LeBlanc M, Ott G, Kvaloy S, Delabie J, Holte H, Krajci P, Stokke T and Staudt LM. (2002). *N. Engl. J. Med.*, **346**, 1937–1947.
- Rubin MA, Zhou M, Dhanasekaran SM, Varambally S, Barrette TR, Sanda MG, Pienta KJ, Ghosh D and Chinnaiyan AM. (2002). *JAMA*, **287**, 1662–1670.
- Sobin LH and Wittekind C, (eds). (2002). *TNM Classification of Malignant Tumours*, 6th edn. Wiley: New York.
- Storkel S, Eble JN, Adlakha K, Amin M, Blute ML, Bostwick DG, Darson M, Delahunt B and Iczkowski K. (1997). *Cancer*, **80**, 987–989.
- Takahashi M, Rhodes DR, Furge KA, Kanayama H, Kagawa S, Haab BB and Teh BT. (2001). *Proc. Natl. Acad. Sci. USA*, **98**, 9754–9759.
- Van Muijen GN, Warnaar SO and Ponc M. (1987). *Exp. Cell Res.*, **171**, 331–345.
- Xu J, Stolk JA, Zhang X, Silva SJ, Houghton RL, Matsumura M, Vedvick TS, Leslie KB, Badaro R and Reed SG. (2000). *Cancer Res.*, **60**, 1677–1682.
- Young AN, Amin MB, Moreno CS, Lim SD, Cohen C, Petros JA, Marshall FF and Neish AS. (2001). *Am. J. Pathol.*, **158**, 1639–1651.