Ontogeny and Oncogenesis Balance the Transcriptional Profile of Renal Cell Cancer

Ena Wang,1 Rudolf Lichtenfels,2 Jurgen Bükur,2,3 Yvonne Ngalame,1 Monica C. Panelli,1 Barbara Seliger,2,3 and Francesco M. Marincola1

1Immunogenetics Section, Department of Transfusion Medicine, Clinical Center, National Institutes of Health, Bethesda, Maryland; 2II Department of Internal Medicine, The Johannes Gutenberg University, Mainz, Germany; and 3Martin Luther University, Institute of Medical Immunology, Halle, Germany

ABSTRACT

Global transcript analysis is increasingly used to describe cancer taxonomies beyond the microscopic reach of the eye. Diagnostic and prognostic portraits are formulated by ranking cancers according to transcriptional proximity. However, the role that distinct biological factors play in defining these portraits remains undefined. It is likely that the transcriptional repertoire of cancers depends, on one hand, on the anamnestic retention of their ontogeny, and, on the other, on the emergence of novel expression patterns related to oncogenesis. We compared the transcriptional profile of primary renal cell cancers (RCCs) with that of normal kidney tissue and several epithelial cancers of nonrenal origin to weigh the contribution that ontogeny and oncogenesis make in molding their genetic profile. Unsupervised global transcript analysis demonstrated that RCCs retain transcriptional signatures related to their ontogeny and cluster close to normal renal epithelium. When renal lineage-associated genes are removed from the analysis and cancer-specific genes are analyzed, RCCs segregate with other cancers with limited lineage specificity underlying a predominance of the oncogenic process over lineage specificity. However, a RCC-specific set of oncogenesis-related genes was identified and surprisingly shared by sarcomas. In summary, the transcriptional portrait of primary RCCs is largely dominated by ontogeny. Genes responsible for lineage specificity may represent poor molecular targets for immune or drug therapy. Most genes associated with oncogenesis are shared with other cancers and may represent better therapeutic targets. Finally, a small subset of genes is associated with lineage-specific oncogenesis, and these may provide information regarding the biological behavior of RCCs and facilitate diagnostic classification of RCCs.

INTRODUCTION

There is substantial evidence that global transcript analysis can differentiate tumors into subclasses beyond the discriminatory power of histopathological observation, and such a distinction may have diagnostic and prognostic value (1–9). Whereas in some cases, molecular subclasses may portray distinct disease taxonomies associated with specific biological behaviors and/or susceptibility to therapeutic intervention, it is possible that temporal changes associated with the dedifferentiation of cancer may segregate taxonomically identical entities into separate groups according to their genetic profile. For instance, we noticed that melanoma metastases can segregate into two molecular subclasses originally believed to represent distinct disease taxonomies (2). However, serial analysis of identical metastases through fine-needle aspiration biopsies suggested that the two subclasses were most likely the result of temporal changes of individual lesion phenotype because material obtained from the same lesions could cluster within either subgroup with a unilateral shift of the later samples toward a less differentiated phenotype (7, 10). We therefore hypothesized that global transcript analysis is influenced by the role that the anamnestic expression of lineage-specific genes may play at different time points of differentiation. However, the previously described study compared melanoma metastases with cultured normal epithelial melanocytes, stretching dedifferentiation analysis to the limit. In addition, the genetic profile of in vitro cultured melanocytes could severely misrepresent “normality” due to nonphysiologic culture conditions.

In this study, we entertained an analysis of the role that ontogeny plays in cancer using as a model primary renal cell cancer (RCCs) tissues paired with normal renal tissue subjected to identical surgical manipulation and experimental preparation. The transcriptional analysis of the paired specimens was subsequently compared with archival frozen samples of melanoma metastases, representing a putative extreme of diversity with regard to ontogenesis and neoplastic progression, and with various primary epithelial cancers to frame the boundaries of similarities and discrepancies in the transcriptional program.

MATERIALS AND METHODS

Tissue Procurement

Seventeen autologous RCCs and normal kidney samples were collected at the Department of Urology of The Johannes Gutenberg University (Mainz, Germany). Specimens were collected as the result of routine operative procedures, and portions were frozen for subsequent analysis, whereas the remaining tissue was used for pathological confirmation. Tissue procurement followed standard ethical procedures according to institutional policy. Details regarding the specimens studied are shown in Table 1. One normal kidney sample (CHTN-5) and nonrenal samples were variably collected during therapy of patients referred at the Surgery Branch, National Cancer Institute (Bethesda, MD) or through the US Tissue Network (Philadelphia, PA). In addition, the following nonrenal specimens were used for the analysis: seven melanoma metastases (Mel), five primary esophageal adenocarcinomas (EO-T); two primary ovarian cancers (OT); three primary soft-tissue sarcomas (CHTN-sarcoma); two primary breast cancers (CHTN-breast); one endometrial cancer (CHTN-endometrial); one colon cancer (CHTN-colon Ad); and one laryngeal cancer (CHTN-larynx).

RNA Preparation, Amplification, and Labeling

Total RNA was extracted from frozen material using Trizol reagent (Invitrogen, Carlsbad, CA) and amplified into antisense RNA (aRNA) as described previously (7, 11–13). Although the quantity of starting total RNA was in most cases sufficient for cDNA array hybridization, we have shown repeatedly that the fidelity of aRNA hybridization is at least equal and likely superior to that of total RNA for transcriptional profiling due to lack of contaminant ribosomal and rRNA (11, 14). Therefore, we used aRNA to increase the consistency of results, particularly when low-quality total RNA was documented by Agilent Bioanalyzer 2000 (Agilent Technologies, Palo Alto, CA). After amplification, aRNA quality was tested with the Agilent Bioanalyzer (12). High-quality aRNA was obtained from 14 RCCs and 12 normal kidney samples that were subsequently subjected to analysis. Of them, 10 constituted RCCs and normal kidney pairs from the same surgical specimen (Table 1).

Similarly, total RNA from peripheral blood mononuclear cells pooled from six normal donors was extracted and amplified to serve as constant reference
INSTITUTE (Bethesda, MD) were printed at the Immunogenetics Section, Department of Transfusion Medicine, Clinical Center, National Institutes of Health (Bethesda, MD) with a configuration of 32 × 24 × 23 and contained 17,500 elements. Clones were used for printing included a combination of the Research Ge-

<table>
<thead>
<tr>
<th>ID</th>
<th>Subtype</th>
<th>Size (cm)</th>
<th>Stage</th>
<th>Grade</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>2733</td>
<td>NN/TU</td>
<td>7.5</td>
<td>pM1</td>
<td>3</td>
<td>a</td>
</tr>
<tr>
<td>2820</td>
<td>NN/TU</td>
<td>8.5</td>
<td>pM1</td>
<td>2</td>
<td>a</td>
</tr>
<tr>
<td>2825</td>
<td>NN/TU</td>
<td>4.5</td>
<td>pM1</td>
<td>2</td>
<td>a</td>
</tr>
<tr>
<td>2826</td>
<td>NN/TU</td>
<td>7</td>
<td>pN0</td>
<td>3</td>
<td>a</td>
</tr>
<tr>
<td>2828</td>
<td>NN/TU</td>
<td>2.8</td>
<td>pM2</td>
<td>2</td>
<td>c</td>
</tr>
<tr>
<td>2829</td>
<td>NN/TU</td>
<td>6</td>
<td>pM1</td>
<td>2</td>
<td>a</td>
</tr>
<tr>
<td>2832</td>
<td>NN/TU</td>
<td>9</td>
<td>pM1</td>
<td>3</td>
<td>NA</td>
</tr>
<tr>
<td>2833</td>
<td>NN/TU</td>
<td>16.5</td>
<td>pM1</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>2834</td>
<td>NN/TU</td>
<td>11</td>
<td>pN0</td>
<td>2</td>
<td>b</td>
</tr>
<tr>
<td>2835</td>
<td>NN/TU</td>
<td>3.5</td>
<td>pM1</td>
<td>3–4</td>
<td>NA</td>
</tr>
<tr>
<td>2837</td>
<td>NN/TU</td>
<td>5.5</td>
<td>pM1</td>
<td>3</td>
<td>a</td>
</tr>
<tr>
<td>2838</td>
<td>NN/TU</td>
<td>ND</td>
<td>pN0</td>
<td>3</td>
<td>b</td>
</tr>
<tr>
<td>2839</td>
<td>NN/TU</td>
<td>Clear</td>
<td>6</td>
<td>pN0</td>
<td>3</td>
</tr>
<tr>
<td>2843</td>
<td>NN/TU</td>
<td>Leptomiosarcoma</td>
<td>ND</td>
<td>3</td>
<td>NA</td>
</tr>
<tr>
<td>2858</td>
<td>NN/TU</td>
<td>Chromophobe</td>
<td>?</td>
<td>pM1</td>
<td>2</td>
</tr>
<tr>
<td>2865</td>
<td>Clear</td>
<td>8</td>
<td>pN1</td>
<td>3</td>
<td>a</td>
</tr>
<tr>
<td>2875</td>
<td>Clear</td>
<td>7</td>
<td>pM1</td>
<td>2</td>
<td>a</td>
</tr>
</tbody>
</table>

NOTE. Bold type indicates the ability to extract and amplify high-quality mRNA for microarray hybridization. Group refers to the assignment of the RCCs according to the unsupervised analysis shown in Fig. 1B.

Abbreviations: NN, normal kidney; TU, tumor, ND, not determined; ?, size unknown; NA, not applicable (no tumor specimen used for analysis).

(7, 11–13). Test and reference RNAs were labeled with Cy5 (red) and Cy3 (green) and cohybridized to a coset-up made 17.5K cDNA microarray. Microarrays were printed at the Immuneonetics Section, Department of Transfusion Medicine, Clinical Center, National Institutes of Health (Bethesda, MD) with a configuration of 32 × 24 × 23 and contained 17,500 elements. Clones used for printing included a combination of the Research Genetics HG_HsKG_031901, 8k clone set and 9,000 clones selected from the HG_Hs_seq_ver_070700_40k clone set. The 17,500 spots included 12,072 uniquely named genes, 875 duplicated genes, and about 4,000 expression sequence tags.

Data Analysis

Quality Validation. All statistical analyses were performed using the log2-based ratios normalizing the median log2 ratio value across the array equal to zero. Validation and reproducibility were performed using our internal reference concordance system, based on the expectation that results obtained through the hybridization of the same test and reference material in different experiments should perfectly collimate. The level of concordance was measured by periodically rehybridizing the same arbitrarily selected test sample (A375 melanoma cell line) with the reference sample. High concordance in gene expression predicts that ratios in different experimental conditions are highly reproducible. With this goal, we analyzed seven forward and seven reciprocally labeled replicate array experiments that were hybridized periodically every other 25 cDNA array slides within each printing. SDS across those 14 arrays were analyzed after labeling swap, and discordant genes due to consistent labeling bias were excluded. This analysis demonstrated a >95% concordance level. Nonconcordant genes due to reproducible or random biases were excluded from subsequent analysis (15).

Unsupervised Analysis. Principal component analysis (PCA) was performed with the Partek Pro software program (Partek Inc., St. Charles, MO). Unsupervised clustering was performed according to the Pearson correlation method of Eisen et al. (16) and visualized with Treeview software (Stanford University, Stanford, CA). Genomic portraits were depicted according to the Pearson correlation method of Eisen et al. (16). One cluster included nine RCCs specimens (cluster a) and two other subclusters closer to each other. One of them included a combination of three RCCs and five normal kidney samples plus the normal kidney sample obtained from the US Tissue Network (cluster b). A third cluster included eight normal kidney samples and one RCCs (R2828 TU) representing the only chromophobe tumor studied (cluster c). Finally, one of the RCCs samples (R2833 TU) clustered away from renal tissues together with other epithelial cancers.

Although the distribution of samples in the three clusters suggested that cluster a contained a more undifferentiated type of RCCS than cluster b, no correlation was noted between the distribution of the RCCS samples in either cluster and their tumor-node-metastasis (TNM) staging or histologic grading (Table 1). Interestingly, three primary soft-tissue sarcomas were included in cluster a in agreement with the PCA analysis. All other primary cancers of nonrenal lineage segregated from the transcriptional profile of renal samples (Fig. 1B). Therefore, primary RCCS clustered closer to normal kidney tissue than to other epithelial nonurologic cancers, underlining the predominance of ontogeny on their global expression pattern. As predictable, melanoma metastases segregated from all other specimens either because of the dedifferentiation that accompanies the metastatic process or because of their neuroectodermal origin. Melanoma metastases were combined with the data set only in this preliminary analysis to portray a "biological extreme" that could help frame the transcriptional profile of the primary cancers studied here between the highly differentiated profile of normal kidney tissue and the highly undifferentiated profile of metastatic tumors. Melanoma metastases were not further analyzed.
because this study was aimed at analysis of the differentiation pattern of primary RCCs compared with that of other primary epithelial nonurologic cancers.

To separate ontogeny from oncogenesis, a supervised analysis was performed to identify genes among the complete data set that were differentially expressed between RCCs included in cluster a and normal kidney. Genes were identified at an arbitrarily set significance threshold of $P < 0.001$ (two-tailed, unpaired Student’s t test). RCCs samples included in groups b and c were excluded with the assumption that a partial overlap of expression of RCCs-specific and renal tissue-specific genes was at the basis of their clustering in proximity to normal kidney samples (see Results section). Renal-related clusters according to B are shown as horizontal bars over the clustering portrait. Vertical bars underline signature of genes overexpressed in RCCs compared with normal kidney (red bars) and underexpressed (green bars).

In an independent analysis, the same 9 RCCs samples in cluster a were compared with all other epithelial primary cancers using a significance cutoff identical to the one used for their comparison against the 13 normal kidney specimens ($P < 0.001$, two-tailed unpaired Student’s t test) that identified 350 differentially expressed genes. This analysis identified 716 transcripts differentially expressed between RCCs in group a and other epithelial cancers. This difference was not attributable to different statistical power between the two comparisons because the number of samples in both groups compared

Fig. 1. A, PCA depicting the special distribution of normal kidney samples (R NN and CHTN-R NN), RCCs (R TU), sarcomas (CHTU Sarcoma), breast cancers (CHTU Breast Ca), colon cancers (CHTU Colon Ca), endometrial cancers (CHTU Endom. Ca), laryngeal carcinomas (CHTU Laryng. Ca), and ovarian cancers (OT TU). PCA analysis was performed on the complete data set. Melanoma metastases were excluded to simplify the graph because they clustered far away from all other tissues. B, Eisen’s clustering of the same samples including melanoma metastases based on 5,098 genes filtered from the complete data set as described in Results. The three subclasses of renal samples are shown by the blue (group a), yellow (group b), and red (group c) vertical bars. C, supervised description of the same cluster presenting the 350 genes responsible for the segregation of RCCs in group a from normal kidney samples (see Results section). Renal-related clusters according to B are shown as horizontal bars over the clustering portrait. Vertical bars underline signature of genes overexpressed in RCCs compared with normal kidney (red bars) and underexpressed (green bars).
with RCCs in group a was similar (14 normal kidney samples and 12 nonrenal epithelial cancers compared with a constant number of RCCs). The number of genes differentially expressed between RCCs and other epithelial cancers (716 genes) was twice the number of genes differentially expressed between the RCCs lesions in group a and normal kidney epithelium (350 genes). This finding underlines the closer proximity of the least differentiated primary RCCs to their lineage compared with other cancers when their global transcriptional profile is monitored.

An analysis using the same significance cutoff was performed on 10 paired RCCs lesions and normal kidney samples derived from the same surgical specimens. This paired analysis identified 170 differentially expressed genes, mostly overlapping those identified with the previous analyses, and included a majority of genes overexpressed in normal renal tissue (data not shown).

**Influence of the General Process of Oncogenesis on the Transcriptional Program of RCCs.** To identify genes associated with oncogenesis independently of lineage, we looked for genes expressed differentially between all primary tumors (RCCs and non-RCCs) and normal kidney with the same arbitrary significance cutoff \( P < 0.001 \) (two-tailed Student’s \( t \) test). This process identified 1,347 genes. Some of the differences identified could be attributed to a predominant effect of genes differentially expressed between nonrenal tumors and normal kidney tissue still due to lineage diversity. Therefore, we separated the recovered 1,347 genes into those that were differentially expressed between RCCs and other cancers (possibly because RCCs retained lineage-specific expression patterns) and those that were not differentially expressed between the same two groups. This process identified 276 genes that were differentially expressed at the \( P < 0.001 \) level between RCCs and other primary cancers, whereas the remainder were above that threshold. Further analysis of these genes demonstrated that they were related to renal lineage specificity, and their inclusion in the data set was due to the strong differences between epithelial cancers and normal kidneys. However, few genes were truly specific for renal oncogenesis, and those were rescued with a separate analysis discussed later. Because the range of significance between \( P < 0.001 \) and 0.01 was felt to be ambiguous for the separation of genes belonging to either group, these genes were also excluded from the analysis. Thus, another 233 genes were excluded from the data set, and the remaining 850 genes were left as putative oncogenesis-specific genes. No information about normal tissues from the other tumors was available. Therefore, some of the differences found might still be due to the presence of lineage-specific genes related to ontogeny of other cancers; however, this is unlikely because these cancers originated from different epithelial tissues.

Among the 850 clones differentially up-regulated in cancer independently of histologic background, 176 clones included expression...
tag sequences or other genes with unknown function. Therefore, we eliminated them from further analysis and focused our attention on the remaining 674 clones with functional annotations (Fig. 2B). As expected, unsupervised reshuffling of tissue samples based on these genes (see the dendrogram in Fig. 2C) segregated all normal renal samples into a separate cluster that also included one RCCs (R2838 TU; red bar). Primary epithelial tumors segregated into two main subclusters, one exclusively including seven RCCs (blue bar) and a second one that included all other tumors including sarcomas and the remaining six RCCs lesions (maroon bar). Therefore, even the elimination of genes differentially expressed between RCCs and other epithelial cancers did not completely abrogate the lineage specificity, at least for some RCCs. The first cluster exclusive for RCCs included all of the RCCs samples from groups b and c (R2834 TU, R2839 TU, and R2828 TU) except for R2838 TU that clustered together with normal renal tissue. The other cluster included exclusively RCCs that belonged to the least differentiated cluster a and R2833 TU that also clustered in the previous analyses with tumors other than RCCs, emphasizing the lower level of differentiation of these samples.

Of the 674 genes specific for oncogenesis, 329 were down-regulated in cancer tissues compared with normal kidney, and most of them represented lineage-specific genes associated with renal function. In the end, 345 genes were identified that were specifically up-regulated in cancerous tissues compared with normal renal tissue.

Interestingly, we noted substantial heterogeneity in the expression of these genes in different cancers independent of lineage. Thus, we examined which of the oncogenesis-related genes were consistently expressed in all cancers or more frequently expressed in RCCs lesions or cancers of different histology. This was achieved by identifying within each group the genes with the highest median ratio. Due to their ambiguous biological behavior, sarcomas were excluded from the analysis. Those genes with a log₂ ratio between median and mean value ≥ 1 within each category were consistently expressed. This parameter identified genes that were expressed above the threshold level in at least 50% of cancers within a given category.

This exercise identified 132 genes whose median expression was similar or higher than their mean level of expression. Of those, 41 were categorized as expressed in RCCs samples, 43 were concordantly expressed by most cancers, and 48 were preferentially expressed by nonurologic epithelial tumors (Supplementary Tables 1–3, respectively).

Influence of RCCs-specific Oncogenesis on the Transcriptional Program of RCCs. A separate analysis was performed to identify genes uniquely expressed by RCCs and unrelated to their renal heritage. This separate analysis was performed to rescue genes possibly missed by the previous analysis due to their down-regulation in RCCs or due to the dilution of their significance level by combining in the analysis the total number of primary cancers. The complete data set was reanalyzed by comparing 14 RCCs samples against the other 12 epithelial cancers. First, this was performed by studying all RCCs cancers independently of the relationship among themselves according to the unsupervised clustering shown in Fig. 1B. The comparison between the two sets of cancers was performed based on a two-tailed unpaired Student’s t test. This analysis identified 1,114 clones differentially expressed by the two sets of cancers. Ontogenesis-related genes were subsequently excluded by subtracting genes coexpressed between RCCs and normal renal tissue (P ≥ 0.001). The large majority of genes dissipated with this test, suggesting that differences between RCCs and other cancers are predominantly due to lineage specificity. However, 43 clones remained that were highly specific for RCCs and differentially expressed between RCCs and other cancer as well as normal renal tissue.

Genes Commonly Expressed by Cluster a RCCs and Sarcomas. To explain the close association between RCCs in cluster a and sarcomas, genes differentially expressed by these tissues were compared with those of other tumors excluding melanoma metastases and other RCCs that did not cluster in cluster a. This analysis identified 597 genes differentially expressed at the 0.001 level (two-tailed t test). The large majority of these genes were coordinately expressed by normal renal tissue, suggesting that the proximity of sarcomas to RCCs in cluster a reflects similar or higher than their mean level of expression. Of those, 41 were categorized as expressed in RCCs samples, 43 were concordantly expressed by most cancers, and 48 were preferentially expressed by nonurologic epithelial tumors (Supplementary Tables 1–3, respectively).

Influence of RCCs-specific Oncogenesis on the Transcriptional Program of RCCs. A separate analysis was performed to identify genes uniquely expressed by RCCs and unrelated to their renal heritage. This separate analysis was performed to rescue genes possibly missed by the previous analysis due to their down-regulation in RCCs or due to the dilution of their significance level by combining in the analysis the total number of primary cancers. The complete data set was reanalyzed by comparing 14 RCCs samples against the other 12 epithelial cancers. First, this was performed by studying all RCCs cancers independently of the relationship among themselves according to the unsupervised clustering shown in Fig. 1B. The comparison between the two sets of cancers was performed based on a two-tailed unpaired Student’s t test. This analysis identified 1,114 clones differentially expressed by the two sets of cancers. Ontogenesis-related genes were subsequently excluded by subtracting genes coexpressed between RCCs and normal renal tissue (P ≥ 0.001). The large majority of genes dissipated with this test, suggesting that differences between RCCs and other cancers are predominantly due to lineage specificity. However, 43 clones remained that were highly specific for RCCs and differentially expressed between RCCs and other cancer as well as normal renal tissue.

Genes Commonly Expressed by Cluster a RCCs and Sarcomas. To explain the close association between RCCs in cluster a and sarcomas, genes differentially expressed by these tissues were compared with those of other tumors excluding melanoma metastases and other RCCs that did not cluster in cluster a. This analysis identified 597 genes differentially expressed at the 0.001 level (two-tailed t test). The large majority of these genes were coordinately expressed by normal renal tissue, suggesting that the proximity of sarcomas to RCCs...
RCCs is due to a particular expression profile in common with normal tissues (compared with other cancer) or specifically to normal renal tissue. For identification of genes specifically coordinately expressed by RCCs in cluster a and sarcomas, the expression profile of both tissue types was compared with normal renal tissue at the $P$ value of 0.001 (two-tailed $t$ test). In this fashion, 48 clones were identified, of which only a few were up-regulated (Fig. 3). The genes almost totally overlapped the RCCs-specific genes identified with the previous analyses.

**Expression of Glutathione S-Transferase and Carbonic Anhydrase in RCCs.** Because of their established relevance in the context of RCCs, we specifically analyzed the expression of the glutathione $S$-transferase (GST) and carbonic anhydrase (CA) gene products (8, 18; Fig. 4). As expected, GST A 2 and 3 were expressed in some RCCs and in normal kidney tissue. Similarly, CA II and IX were preferentially expressed by RCCs, including the only chromophobe RCCs (R2828 TU) that expressed high levels of CA II, as suggested by others (8). Interestingly, whereas CA IX expression was shared by other primary epithelial cancers (Supplementary Table 2), CA II appeared to be selectively expressed by RCCs and sarcomas and minimally expressed by normal renal parenchyma.

**Expression of Genes Reported by Others to Have Prognostic Significance.** Takahashi et al. (19) have described two molecular phenotypes of clear cell RCCs associated with divergent prognosis. We therefore searched our database for genes identified as prognostic markers by that study. A significant proportion of them were present in our arrays and found to be significantly up-regulated or down-regulated in some clear cell RCCs. As reported, these genes were RCCs specific and, even in this small series, separated clear cell RCCs into two subclasses (Fig. 5). Interestingly, we observed that the clear cell RCCs subclass expressing good prognostic indicators contained all of the RCCs that had appeared as most differentiated clustering in group b or c, suggesting that this relatively benign phenotype combines the expression of oncogenesis-related good indicators with a less-differentiated transcriptional program.

**DISCUSSION**

The molecular subclassification of kidney tumors and the discovery of new diagnostic markers were described previously by others and are beyond the purpose of this work (4, 5, 8, 19). Our goal, rather, was to explain the common observation of cancers of similar histology segregating onto different molecular subgroups. For instance, Takahashi et al. (19) described molecular subgroups of RCCs that could prognosticate the natural history of the RCCs. That study, however, did not distinguish whether the two subclasses represented distinct disease taxonomies or different stages of a genetically similar disease. In a previous study, we postulated that melanomas could be subcategorized into different molecular classes based on their transcriptional profile (2). Subsequently, analysis of serial fine-needle aspirates obtained from identical lesions demonstrated a unidirectional shift of molecular profiles of the same lesions from a more to less differentiated phenotype along with disease progression, raising questions about a rigid classification of morphologically identical diseases based on one time point observations (7). In this study, therefore, we measured the role that the anamnestic expression of lineage-specific
genes plays in determining the genetic classification of morphologically similar RCCs. We relied predominantly on clear cell RCCs specimens (12 of the 14 specimens analyzed) representing the most common type of RCCs, resulting mostly from the relatively homogenous effects of VHL gene inactivation. In addition, a chromophobic RCCs (R2828 TU) and a chromophilic RCCs (R2858 TU) were included. The first finding was that primary RCCs segregate into at least two molecular subgroups. Such subgroups are portrayed in Fig. 1A by cluster a containing only RCCs samples and the closely related clusters b and c with a mixture of RCCs and normal renal specimens including the chromophobe RCCs (R2828 TU) consistent with the relatively high level of differentiation of this tumor (4). Interestingly, both normal kidney and tumor tissues of the samples 2838 and 2839 clustered into group b, whereas other normal and tumor samples are quite distant in Fig. 1B. We truly have no exact explanation for the specific pairing of these two tumor/normal tissue samples, although we and others have seen it before in other experimental situations. The best explanation we have is that in conditions of relatively good differentiation, primary tumors tend to align genetically with the normal tissue of origin because of the individual’s genetic background. This hypothesis has never been properly tested, and a much larger population would be necessary to demonstrate this point. We are, however, quite confident that the similarities noted are not artifactual, based on RNA quality analysis and pathological testing (the samples from which RNA was extracted were composed almost exclusively of tumor or normal cells, respectively, in a proportion similar to other samples). The genes responsible for the differences noted between the RCCs of cluster a and those of the other two clusters belonged predominantly to the renal lineage and were mostly coexpressed by the latter groups of RCCs and normal renal tissues, suggesting that global transcript separated RCCs based on their level of differentiation. Conversely, genes overexpressed by RCCs of group a represented a minority and were coordinately expressed by the other RCCs and other primary epithelial cancers. Thus, one could conclude that if no information was available about the transcriptional profile of paired normal tissues, it would have been difficult to explain the separation of the two major classes of RCCs (cluster a compared with the close subcluster b or c), and it might have been tempting to attribute such differences to separate taxonomies of disease as described previously in the context of metastatic melanoma (2) before serially analyzing identical lesions during disease progression (7). In conclusion, the identification of molecular subcategories of disease may be influenced by the level of dedifferentiation of a given cancer rather than portray a distinct disease taxonomy.

The second observation suggests that removal of lineage-specific genes segregated primary RCCs together with other nonurological cancers because at least some mechanisms of oncogenesis are shared by most cancers. Thus, this study suggests that most genes responsible for segregating cancers of various histology into molecular subclasses are related to the ontogeny of the individual cancers, whereas only a small proportion is responsible for the subclassification of cancers into different molecular entities when ontogeny is removed.

The third important observation demonstrated that although the exclusion of lineage-specific genes closely approximates RCCs with other nonurologic epithelial cancers, several genes (signatures) could
be identified as specific for RCCs and, surprisingly, soft-tissue sarcomas. Several of the genes have been described previously by suppression subtractive hybridization (20) and functional genomics (19) to be strongly associated with RCCs oncogenesis, such as, for instance, lysyl oxidase and ceruloplasmin. Ceruloplasmin is secreted by human clear cell carcinoma cells in patients (21) and relevant instance, lysyl oxidase and ceruloplasmin. Ceruloplasmin is secreted by macrophages and endothelial cells, has been previously demonstrated to be strongly increased in RCCs tissues compared with normal renal tissue and benign renal tumors (31). In addition, insulin-like growth factor-binding protein 3 was coexpressed specifically by RCCs in group a and sarcomas (Fig. 3) has been described previously by gene expression profiling (8).

Recent studies suggested that GST A is expressed in the proximal convoluted tubules of normal renal parenchyma and that expression of GST A is related mostly by clear cell RCCs (8). In fact, GST A 2 and 3 were coordinately expressed in most clear cell RCCs (8 of 13) but were not expressed by the chromophobes (R2828 TU) and chromophilic (R2858) RCCs as predicted by others (8). GST was also expressed by normal renal tissue but not by sarcomas, in line with the anamnetic expression of these genes.

CA IX is an independent predictor of survival in RCCs (18), whereas CA II is a reliable marker specific for chromophobe RCCs and erratically expressed by clear cell RCCs (30%) (8). CA II is also expressed in a small proportion of distal convoluted tubules of normal renal parenchyma. Indeed, CA II transcript could be found in normal kidney tissue, suggesting that expression of CA II transcript is related to the ontogenesis of RCCs. Although the highest level of CA II expression was noted in chromophobe RCCs, several clear cell RCCs expressed this gene, suggesting that, at least at the RNA level, there is overlap between the two types of RCCs (8). Interestingly, CA IX expression was specific for RCCs compared with normal renal tissue, confirming that this gene may be a better target for drug or immune therapy (32). Interestingly, the expression of CA IX was shared by other primary epithelial cancers but not melanoma metastases, suggesting that the expression of this gene is at least partly associated with oncogenesis and that it may not be solely responsible for the immune responsiveness of these diseases to immune therapy.

Other genes such as phosphofructokinase have been associated with childhood nephroblastomas, but not with adult RCCs (33). In addition, scavenger receptors have been identified on RCCs cell lines in the past, although no further work has been performed on this subject (34). Surprisingly, tapasin appeared to be overexpressed in RCCs, which is in contrast with immunohistochemical findings using paraffin-embedded tumor material (35). It is possible that the increased mRNA tapasin levels are not associated with enhanced protein expression or that the mRNA is produced by normal cells infiltrating RCCs lesions.

Other genes that have been identified in clear cell RCCs were found in this study to also be coexpressed by primary tumors of other histology such as pro-collagen-lysine, 2-oxoglutarate 5-dioxygenase 2 (8).

Others have compared the transcriptional profile of clear cell RCCs with clinical staging and outcome (19). Surprisingly, transcriptional profiling did not differentiate between stage I/II and stage III/IV lesions. On the contrary, and most relevantly, it was found that genetic profiling could segregate clear cell RCCs into two categories, one with a very poor 5-year survival, and the other with a good prognosis. This a very interesting finding suggesting that tumor biology is better represented at the transcriptional level. We also found no correlation between clinical staging and transcriptional profiling; however, analysis of genes whose expression is associated with good prognosis in RCCs (Fig. 5) suggested that such genes are preferentially expressed by cancers with a genetic profile closer to normal kidney.

In summary, this study is largely confirmatory of other studies in which RCCs-associated markers have been described. However, the three-way comparison between RCCs, normal renal tissue, and assorted nonurologic primary epithelial cancers allowed the differentiation of genes expressed by RCCs as a remnant of their origin from normal kidney from those directly related to the general process of oncogenesis. This differentiation is important because different categories may have diverse biological, diagnostic, and therapeutic value. Genes responsible for lineage specificity may represent poor molecular targets for immune or drug therapy. Most genes associated with oncogenesis are shared with other cancers and may represent better therapeutic targets. Finally, a subset of gene associated with oncogenesis is lineage specific and may provide information regarding the specific biological behavior of individual cancers and facilitate their diagnostic classification. More generally, this study suggests that in large proportion, the molecular portraits of primary tumors are dependent on their ontology, and therefore caution should be applied when defining distinct molecularly defined subcategories of cancer.

REFERENCES

Ontogeny and Oncogenesis Balance the Transcriptional Profile of Renal Cell Cancer


Cancer Res 2004;64:7279-7287.