Tumor-Specific Isoform Switch of the Fibroblast Growth Factor Receptor 2 Underlies the Mesenchymal and Malignant Phenotypes of Clear Cell Renal Cell Carcinomas

Qi Zhao, Otavia L. Caballero, Ian D. Davis, Eric Jonasch, Pheroze Tamboli, W.K. Alfred Yung, John N. Weinstein, Kenna Shaw for the TCGA research network, Robert L. Strausberg, and Jun Yao.

Abstract

Purpose: We aim to identify tumor-specific alternative splicing events having potential applications in the early detection, diagnosis, prognosis, and therapy for cancers.

Experimental Design: We analyzed RNA-seq data on 470 clear cell renal cell carcinomas (ccRCC) and 68 kidney tissues to identify tumor-specific alternative splicing events. We further focused on the fibroblast growth factor receptor 2 (FGFR2) isoform switch and characterized ccRCCs expressing different FGFR2 isoforms by integrated analyses using genomic data from multiple platforms and tumor types.

Results: We identified 113 top candidate alternatively spliced genes in ccRCC. Prominently, the FGFR2 gene transcript switched from the normal IIIb isoform ("epithelial") to IIIc isoform ("mesenchymal") in nearly 90% of ccRCCs. This switch is kidney specific as it was rarely observed in other cancers. The FGFR2-IIIb ccRCCs show a transcriptome and methylation resembling those from normal kidney, whereas FGFR2-IIIc ccRCCs possess elevated hypoxic and mesenchymal expression signatures. Clinically, FGFR2-IIIb ccRCCs are smaller in size, of lower tumor grade, and associated with longer patient survival. Gene set enrichment and DNA copy number analyses indicated that FGFR2-IIIb ccRCCs are closely associated with renal oncocytomas and chromophobe RCCs (chRCC). A reexamination of tumor histology by pathologists identified FGFR2-IIIb tumors as chRCCs and clear cell papillary RCCs ( ccpRCC).

Conclusions: FGFR2-IIIb RCCs represent misdiagnosed ccRCC cases, suggesting FGFR2 isoform testing can be used in the diagnosis of RCC subtypes. The finding of a prevalent isoform switch of FGFR2 in a tissue-specific manner holds promise for the future development of FGFR2-IIIc as a distinct early detection biomarker and therapeutic target for ccRCC.

Clinc Cancer Res; 19(9); 2460–72. ©2013 AACR.
FGFR2 Isoform Switch in Renal Clear Cell Carcinoma

Translational Relevance

Aberration in alternative mRNA splicing is implicated in tumorigenesis and cancer progression. Through in silico analysis of The Cancer Genome Atlas (TCGA) RNA-seq data, we identified fibroblast growth factor receptor 2 (FGFR2) isoform switch from the epithelial "IIIb" isoform to the mesenchymal "IIIc" isoform in ~90% of ccRCCs. This was experimentally validated by PCR sequencing using additional ccRCC cases. FGFR2-IIIb RCCs do not exhibit activated mesenchymal gene expression signature, are less hypoxic, and associated with longer patient survival. Through integrated cancer genomics analyses, we concluded that FGFR2-IIIb RCCs were misdiagnosed clear cell papillary RCC (ccpRCC) and chromophobe RCC (chrRCC) cases. The FGFR2 isoform switch is ccRCC specific, as it is not observed in other cancers examined including those of breast, lung, colon, head and neck, endometrium, and bladder. Our study shows FGFR2-IIIc switch as a hallmark event in ccRCCs and suggests FGFR2-IIIc isoform as a promising candidate biomarker for early detection, diagnosis, and targeted therapy for ccRCC.

Functional studies using mouse models suggest that FGFR2 IIIb plays a critical role in mesenchymal–epithelial signaling during early organogenesis. Studies based on expression of a soluble dominant negative FGFR2 IIIb mutant in mice have implicated FGFR2 in the development of many organs including the limb, lung, skin, kidney, and several glandular tissues (13–15). Results from studies on knockout mice corroborate these findings. FGFR2 IIIb-null mice are viable until birth, with severe defects of limbs, lung, anterior pituitary gland, and dysgenesis of many organs including the kidney, thymus, and pancreas (16–18).

Developmentally, the mammalian kidney is derived from intermediate mesoderm. For kidney formation, both EMT and the reciprocal mesenchymal–epithelial induction between the epithelial ureteric bud and the metanephric mesenchyme occur during embryogenesis (19). Hence, unlike other types of cancers, such as breast and colon cancer, where putative tumor-initiating cells originate from epithelial cells, the stem cell origin of renal cell carcinomas (RCC) is less clear due to the presence of both epithelial and mesenchymal cells during kidney development. Histologically, RCCs are divided into several major subtypes, including clear cell RCC (ccRCC), papillary RCC (pRCC), chromophobe RCC (chrRCC), and renal oncocytoma. Although ccRCC and pRCC have high expression of the mesenchymal marker vimentin, chrRCC and renal oncocytoma express the epithelial marker E-cadherin. This suggests that different subtypes of RCC may develop from cells of different origins and/or experience EMT during tumor progression.

Alterations of the FGFR2 gene, especially FGFR2 IIIb, have been reported in many cancers, with evidence for both tumor promoting and suppressing roles within specific cancer contexts (20). For example, consistent with a tumor promoting role, FGFR2 IIIb is amplified in a subset of gastric cancers, which are poorly differentiated and do not have ERBB2 amplification (21, 22). Genome-wide association studies have identified FGFR2 as a breast cancer susceptibility gene, with single-nucleotide polymorphism (SNP) located inside FGFR2 exon 2 correlated with increased risk of breast cancer (23, 24). Transfection of FGFR2 IIIb cDNA into NIH/3T3 fibroblasts induced foci formation (25). In addition, missense activating mutations of FGFR2 were found in endometrial, breast, lung, and ovarian cancer (26, 27). However, FGFR2 IIIb also exhibits cancer-specific properties suggestive of tumor suppressor functions. FGFR2 IIIb expression is downregulated in solid tumors, such as in bladder, prostate, and salivary cancer (26, 27).}

www.aacrjournals.org  Clin Cancer Res; 19(9) May 1, 2013 2461

Downloaded from clincancerres.aacjournals.org on September 5, 2021. © 2013 American Association for Cancer Research.
RNA-seq data on the exon expression level and exon junction reads, we identified specific expression of FGFR2 IIIc in nearly 90% of ccRCCs, as compared with predominant FGFR2 IIIb expression in normal kidneys. This isoform switch of FGFR2 is tissue-specific as it is rarely observed in paired normal and tumor tissues in breast, lung, colon, bladder, and head and neck cancer. Compared with FGFR2 IIIb ccRCCs, FGFR2 IIIc ccRCCs have higher expression of mesenchymal and hypoxic genes, are more malignant, and have worse clinical outcome.

Materials and Methods

Molecular profiling datasets and data preprocessing

Level 3 RNA-seq data (containing data on gene, exon, and junction levels), level 3 SNP array data, level 2 DNA methylation data (Infinium Human Methylation 450), mutation data, and clinical data for multiple cancers were downloaded from The Cancer Genome Atlas (TCGA) data portal (https://tcga-data.nci.nih.gov/tcga/dataAccessMatrix.htm) by January 2012. For DNA methylation data, M values were calculated as log2 ratio of methylated intensity over unmethylated intensity (39). Normalized gene expression datasets of mouse kidney development and RCC subtypes (GSE1983 and GSE15641) were downloaded from NCBI GEO website (http://www.ncbi.nlm.nih.gov/geo). Probesets with matching human orthologs were kept by referencing University of California Santa Cruz (UCSC) blast tables (http://genome.ucsc.edu/cgi-bin/hgTables).

Identification of recurrent tumor-specific alternative splicing events

For each exon, its RPKM (Reads Per Kilobase per Million mapped reads, a measure of expression level by RNAseq) is normalized by dividing the sum RPKM of all exons from a gene. To identify recurrent ccRCC-specific alternative splicing events, we first limited exons to have a more than 1.25-fold difference of median normalized RPKM between normal kidney and ccRCCs. To exclude low-expressing exons, which might represent background or low frequency events, we also excluded exons having a median normal and tumor RPKM below 0.5. A total of 6,820 exons remain after this filtering step. The 'limma' package in R was used to identify differentially expressed exons, which is essentially conducting Student t tests on normalized exon RPKMs between normal and tumor samples with adjusted output P values. We selected exons having more than 3-fold difference in expression and adjusted P value < 1e−5. Candidate exons from the above screen were further validated by junction reads using a similar strategy. For each junction read associated with a candidate exon, we determined whether there was a more than 3-fold difference in median normalized junction reads between normal and tumor. Multiple junctions may be associated with one single exon. We selected candidate exons with more than 50% junctions that passed this test. The frequency of altered splicing events in ccRCC was calculated as the percentage of tumors having exon RPKM values below or above 1.5 times the SD of normal exon RPKM reads. We selected exons with an altered frequency of more than 50%. Candidate exons were annotated using UCSC gene tables (http://genome.ucsc.edu/cgi-bin/hgTables) to designate exons as first (representing alternative transcriptional start), last (representing alternative transcriptional end), or internal within any known transcripts. R scripts for this study are available upon request.

Computational prediction of FGFR2 IIIb/c isoforms using RNA-seq data

To assign samples into FGFR2 IIIb/c groups, we calculated expression ratio of FGFR2 exon 8 (hg19:chr10:123,278,343-196) over exon 9 (hg19:chr10:123,276,986-728). Samples with more than 2-fold expression of exon 8, supported by 2-fold more reads from either junction of exon 7–8 over 7–9 or exon 8–10 over 9–10 were assigned to the FGFR2 IIIb group. Samples with more than 2-fold expression of exon 9, and supported by 2-fold more reads from either junction of exon 7–9 or 9–10 were assigned into the FGFR2 IIIc group. To deal with low expressers, for each sample, at least one junction raw reads involved must be equal or more than 3. Samples that met these criteria were called as either "IIIb" or "IIIc." Otherwise, the sample was marked as no call.

Experimental validation of FGFR2 isoform switch in ccRCC

A total RNA of 18 primary ccRCC tumor samples, composed of stage T1 to T3 tumors, were obtained by Dr. Ian Davis from the Victorian Cancer Biobank, which is supported by the Victorian Government, Australia, with appropriate ethics approval. Two ccRCC tumor cell lines (Caki-1 and 786-O) were purchased from American Type Culture Collection (ATCC) and grown according to the ATCC’s recommendations. Total cellular RNA was harvested and subjected to real-time (RT)–PCR using primers specific for the exons upstream and downstream of the alternative exons IIIb and IIIc as shown in (Supplementary Fig. S2). cDNA was amplified with the Omniscript RT kit using 1 μg of total RNA. Primer sequences flanking exon8/9 are: forward- GGAATCAAGCAGCTGAAAAG in exon7 and reverse- TCAGGCAAGTATGCACTTGA in exon10. JumpStart RED-Taq ReadyMix PCR Reaction Mix (Sigma) was used for PCR amplifications after the addition of 5 pmoles of each primer and one μL of the cDNA solution in 25 μL. The PCR conditions were 95°C for 3 minutes followed by 42 cycles at 95°C for 15 seconds and 60°C for 30 seconds and 72°C for 30 seconds, followed by a final 7-minute extension. PCR products were gel-purified and submitted to Sanger sequencing with the PCR primers to confirm the identity of the amplicon.

Clustering analysis and statistical testing

Exon level data were log transformed and clustered using the ‘hclust’ function from R software (http://www.r-project.org). The options used were “canberra” for distance calculation, and “median” for clustering. We chose to use these conditions to better separate samples on both differential exon expression and overall FGFR2 expression. All other clustering analyses were done using the Cluster
in 64% of cases, we observed differential exon usage (exam-
and junction sequencing reads, we were able to predict the
1A) could not be validated by the junction data. Using exon
IIIb). Another tumor-specific change in FGFR2 exon19 (Fig.
mal kidney expresses FGFR2 with exons 7-8-10 (FGFR2
express FGFR2 having exons 7-9-10 (FGFR2 IIIc) and nor-
10, whereas the majority of ccRCCs predominantly have
fraction of ccRCCs cluster with normal kidney and also
normal tissues and exon 9 to ccRCCs. Notably, a small
expression reads of both normal kidney and ccRCCs (Fig. 1B
shown in Fig. 1A, the median normalized expression of
tumor growth, invasion, and stem cell maintenance. As
usage, and because the known important role of FGFR2 in
and TreeView software (http://rana.lbl.gov/EisenSoftware.
htm). Fisher exact test, Wilcoxon test, Cox proportional
hazard regression, and Kaplan–Meier log rank test were
done using the R software.

Gene set enrichment analysis to identify RCC
subtypes most related to FGFR2 IIIb RCCs
Differentially expressed genes between oncocytoma,
chRCC, normal kidney, and ccRCC were identified by sig-
edu/~tibs/SAM) of published microarray data GSE15641.
Genes with more than 3-fold differences in expression
with a false discovery rate less than 0.01 were selected and
were limited to top 250 genes with highest folds of
difference. These RCC subtype-specific up- and downregu-
lated gene signatures were combined with the MSigDB c4
computational gene sets (cancer-orient expression sets, avail-
able at http://www.broadinstitute.org/gsea/index.jsp) and
subjected to gene set enrichment analysis (GSEA; ref. 40)
using eaft conditions on 22 FGFR2-IIIb tumors versus
55 randomly picked FGFR2-IIIc tumors.

Results
Identification of tumor-specific alternative splicing of
FGFR2 mRNA in ccRCC
To search for recurrent tumor-specific alternative splicing
events in ccRCC, which may play a critical role in tumor-
genesis and tumor progression, we examined the RNA-seq
data from the TCGA datasets on 470 ccRCCs and 68 normal
kidney tissues (see Materials and Methods for algorithm
details). This yielded a top candidate list of 234 exons from
113 genes (Supplementary Table S1). About one third of
theses cases involved 5 alternative transcriptional start, and
in 64% of cases, we observed differential exon usage (exam-
examples are shown in Supplementary Fig. S1).

We focused our study on the FGFR2 gene, because this is
the only gene where we observed mutually exclusive exon
usage, and because the known important role of FGFR2 in
tumor growth, invasion, and stem cell maintenance. As
shown in Fig. 1A, the median normalized expression of
FGFR2 exon 8 (encoding FGFR2 IIb) is statistically signif-
ically higher in normal kidney than in ccRCC, and this
situation is reversed in exon 9 (encoding FGFR2 IIc). The
finding is confirmed by the unsupervised clustering of exon
expression reads of both normal kidney and ccRCCs (Fig. 1B
middle), where exon 8 expression is clearly associated to
normal tissues and exon 9 to ccRCCs. Notably, a small
fraction of ccRCCs cluster with normal kidney and also
express exon 8. These results are further corroborated by the
exon junction sequence reads (Fig. 1B bottom), where
normal tissues have more reads of exon 7–8 and exon 8–
10, whereas the majority of ccRCCs predominantly have
reads of exon 7–9 and exon 9–10. Clearly, most ccRCCs
express FGFR2 having exons 7-9-10 (FGFR2 IIc) and
normal kidney expresses FGFR2 with exons 7-8-10 (FGFR2
IIb). Another tumor-specific change in FGFR2 exon19 (Fig.
1A) could not be validated by the junction data. Using exon
and junction sequencing reads, we were able to predict the
major FGFR2 isoform present in both normal kidney and
cRCCs (see Materials and Methods). This predicted 47 (out
of 68) normal tissues and 22 ccRCCs as IIb type, and 398
ccRCCs as IIc type (Supplementary Table S2). None of the
normal tissues were predicted as major IIc type, although
there were 21 normal (31% of 68) and 50 ccRCCs (11% of
470) whose major FGFR2 isoform type could not be deter-
mined, presumably due to low gene expression or compa-
able expression of both FGFR2 IIc and IIb isoforms. To
validate our findings of FGFR2 IIc as the major isoform
expressed in ccRCCs, we amplified the mRNA coding region
spanning FGFR2 exon8 and exon9 in 18 primary ccRCC
tumors and 2 ccRCC cell lines using RT-PCR followed by
Sanger sequencing of the PCR products (See Materials and
Methods). This showed that exon9 (IIc) is the predominant
form expressed in all 20 ccRCC samples tested (example
result shown in Supplementary Fig. S2).

Isoform IIIc switch of FGFR2 is tissue specific and
correlates with reduced expression of ESRP1 and
GRHL2
To examine whether the FGFR2 IIIc switch is a more
general event occurring in additional cancer types besides
ccRCC, we used TCGA RNA-seq exon and junction level
data to predict major FGFR2 isoforms present in normal
tumor tissues from a panel of cancer types including
breast, lung, colon, head and neck, uterus, kidney, and
bladder. This revealed that FGFR2 mainly exists as the IIb
isoform in all cancer types examined except ccRCC (Sup-
plementary Table S3). Closer examination of paired normal
and ccRCC tumor samples obtained from the same patient
showed prevalent FGFR2 isoform switching from normal
IIb to tumor IIIc. This occurs only in ccRCCs, except for a
single case of breast cancer with a clear IIIc switch (Fig. 2A).
Thus, FGFR2 IIIc isoform switching in cancer has a strict
tissue specificity.

To search for possible regulators controlling this iso-
form switch, we carried out the Wilcoxon test to find associa-
tion of gene expression to FGFR2 isoforms in ccRCC,
endometrial, and bladder cancer. The top 100 genes with the
least P values from all 3 cancer types were collected and
intersected (data not shown). This identified 2 genes in
common, ESRP1 and GRHL2, among which ESRP1 is a
splicing factor previously reported to control alternative
splicing of FGFR2 (12). As shown in Fig. 2B, there is a
good correlation of high ESRP1 or GRHL2 expression with
the prediction of FGFR2 isoform being IIb, where expres-
sion above the yellow lines correlated perfectly well with
the presence of FGFR2 IIb but not IIc isoform. Because
ESRP1 is a known FGFR2 splicing regulator, this result
serves well as an indirect validation of our FGFR2 iso-
form prediction algorithm. When we looked at ESRP1
expression in all cancer types, we observed universally
equivalent high-level expression of ESRP1 in most tumor
tissues, as contrasted by a substantial reduction of ESRP1
expression in ccRCCs. High levels of ESRP1 were cor-
related with FGFR2 IIb expression (Fig. 2C). Because
FGFR1 and FGFR3 have exon structures similar to those
of FGFR2, we also examined whether these receptors were undergoing similar isoform switches as in FGFR2. No isoform IIIc switch was found in FGFR1 and FGFR3 in ccRCCs (Supplementary Fig. S3). Thus, despite similar organization of exons, alternative splicing of FGFR1 and FGFR3 in ccRCC is either not controlled by ESRP1 alone or requires additional regulators. Taken together, these results suggest that FGFR2 IIIc switch occurs in a tissue-specific manner as well as in a FGFR2-specific manner within the FGFR receptor family.

FGFR2 IIIb ccRCC has gene expression and FGFR2 promoter methylation patterns resembling that of normal kidney but not IIIc ccRCC

To study the difference at the molecular level between FGFR2 IIIb and IIIc ccRCC, we carried out an unsupervised clustering of the top 3,600 variably expressed genes in 470 ccRCCs and 68 normal kidney tissues (Fig. 3A, similar results can be obtained using 2,500–5,000 genes). All normal tissues were clustered together as expected. The majority of ccRCCs were clustered into 2 major groups with one cluster of tumors being more malignant when assessed by tumor grade, stage, and size, although VHL mutation was observed to occur at similar frequency between the 2 subtypes. Interestingly, all FGFR2-IIIb ccRCCs were outside of these 2 major ccRCC clusters, clustered next to normal tissues, and possessed a global gene expression pattern very similar to that of normal samples, indicating these tumors have a different transcription program from FGFR2 IIIc ccRCCs. Notably, with FGFR2-IIIb ccRCCs, tumors can further be clustered into 2 subgroups based on their transcriptome differences (designated FGFR2-IIIb-1 and 2, Fig. 3A). We further carried out an unsupervised cluster analysis of promoter DNA methylation in 227 ccRCCs and 21 normal kidney tissues which have both RNA-seq and methylation data. This was done on 82 methylation probes.
within a region encompassing 20kbp of the FGFR2 gene. Again, normal tissues and most IIIb ccRCCs were clustered together (Fig. 3B). These results argue that FGFR2 IIIb ccRCC represents a small subset of renal clear cell carcinomas (22/470, 5%) constituting a distinct molecular subtype. In addition to isoform switch, we also observed a significant reduction of FGFR2 gene expression in ccRCC compared with that in normal kidney. This downregulation can be at least partially attributed to the hypermethylation of exon1 and intron1 region of the FGFR2 gene (Supplementary Fig. S4).

**FGFR2 IIib and IIIc ccRCCs show distinct gene expression signatures and signaling pathways**

Although EMT is thought to be a critical event implicated in cancer progression and metastasis, the prevalence of mesenchymal FGFR2 IIIc ccRCCs among all ccRCCs studied (~85%), together with results from recent studies on RCC stem cells, raises the possibility that FGFR2 IIIc RCCs may have a mesenchymal stem cell origin (41). To this end, we compared the gene expression differences between FGFR2 IIib and IIIc ccRCCs with previously published data on mouse embryonic kidney development (42), in which gene
expression of ureteric bud and metanephric mesenchyme, integral components for kidney development with epithelial and mesenchymal origins, respectively, were profiled. This confirmed that FGFR2 IIIb ccRCCs displayed an epithelial phenotype, with elevated expression of epithelial markers E-cadherin (CDH1), keratin 7 (KRT7) and claudins (CLDN4, 7, 8), and reduced expression of mesenchymal markers N-cadherin (CDH2), collagen 23 A1 (COL23A1), and vimentin (VIM; ref. Fig. 4A).

In addition, we observed other gene expression changes in signaling pathways between FGFR2 IIIb and IIIc ccRCCs, which are similar to those between mouse ureteric bud and metanephric mesenchyme tissues. For example, both FGFR2 IIIb RCCs and ureteric bud tips had elevated FGF9 gene expression. The HOXB genes are overexpressed and HOXA4 downregulated in FGFR IIIb RCCs as well as ureteric bud tips as compared with IIIc RCCs and metanephric mesenchyme. This is also true for a panel of cell surface markers (e.g., CD82). Because HOX gene expression perfectly matches the timing and route of development, combination of these results suggested that FGFR2 IIIc and IIIb ccRCC have originated from stem cells of mesenchymal and epithelial origins, respectively, during tumorigenesis.

Downregulation of VHL and hypoxia are hallmarks of renal cell carcinomas. When we examined the hypoxia response in FGFR2 IIIb and IIIc ccRCCs using a 15-gene hypoxia gene expression signature (43), we observed elevated expression of the hypoxic signature in IIIc ccRCCs (Fig. 4B) and FGFR2-IIIb-1 tumors. There is no activation of the 15-gene signature in normal kidney tissues, as expected. However, activation of this hypoxic signature in FGFR2-IIIb-2 ccRCCs is much lower compared with IIIc ccRCCs, suggesting they are less hypoxic (Fig. 4B). This result can be further confirmed using a different set of 36-gene hypoxia-
inducible factor alpha target signature (data not shown). When VHL gene copy number was examined, we observed prevalent copy number loss in FGFR2-IIIc ccRCCs, as expected. However, VHL copy numbers are largely intact in IIIb-1 tumors despite they also exhibit elevated hypoxia activation, and VHL copy gains are observed in FGFR2-IIIb-2 tumors. In contrast, FGFR2-IIib-2 tumors exhibit loss of p53 and PTEN copy numbers, which is correlated with reduced expression of these tumor suppressors in these tumors (Fig. 4B). Taken together, these findings argue that FGFR2 IIib and IIIc ccRCCs are two or more distinct molecular subtypes, and might represent different disease entities.

**FGFR2 IIIb ccRCCs are closely related to clear cell papillary RCC and chromophobe RCC and have better clinical outcomes**

The fact that FGFR2 IIib ccRCCs possess gene expression and methylation patterns resembling normal kidney tissues, express epithelial markers, and do not carry VHL loss of heterozygosity (LOH) events, suggests that these tumors may be of a more differentiated state and less malignant in nature. We tested this hypothesis by examining the clinical parameters from 22 FGFR2 IIib ccRCCs and 398 IIIc ccRCCs (available on the TCGA website). Remarkably, all 22 FGFR2 IIib ccRCCs do not show distant metastasis (Table 1, Fisher exact test 2-sided P value = 0.034) and for those tumors having lymph node status, none of the IIib ccRCCs had spread of tumor into regional lymph nodes. Compared with FGFR2 IIIc ccRCCs, IIib tumors are statistically significantly smaller in tumor size and of lower tumor grade and stage (Table 1). Specifically, 18 of 22 FGFR IIib tumors were diagnosed as stage T1 tumors. Only one out of 22 FGFR IIib tumors was diagnosed as a stage T3 tumor (i.e., tumors extending into major veins or perinephric tissue and not beyond Gerota’s fascia). This frequency (~5%) is much lower than the overall T3/T4 frequency found in ccRCCs (42% for the TCGA dataset).

Because microarray gene expression profiles of different RCC subtypes are available (44; GSE15641), we further compared differential expression between FGFR2 IIib and IIIc ccRCCs with known changes in gene expression between ccRCCs and other RCC subtypes, using GSEA analysis (see Material and Methods). This analysis indicated that FGFR2 IIib ccRCC is most related to renal oncocytoma, which is benign in nature, and to chRCC (Fig. 5A). This result was further supported by examining the gene copy number alterations in FGFR2-IIIb and -IIIc tumors (Fig. 5B). On the basis of SNP array-derived copy number data, FGFR2-IIIb-1 tumors had minimum copy number alterations, whereas FGFR2-IIIb-2 tumors had copy number gains and losses on almost every autosomal chromosome. Copy number changes in FGFR2-IIIb-2 tumors are recurrent at an extremely high frequency (~90%) and this differentiates them from FGFR2-IIIc tumors, which had consistent VHL LOH on chromosome 3 (Fig. 3B, Supplementary Fig. S5). The cytogenetic abnormalities observed in FGFR2-IIIb-2 RCCs matched perfectly well with those reported for chRCCs (45), where LOH on chromosomes 1, 2, 6, 10, 13, 17, and 21 were reported. This strongly suggests FGFR2-IIIb-2 RCCs are in fact misdiagnosed chRCC cases, which is confirmed by pathologic reexamination of tumor histology. With respect to FGFR2-IIIb-1 tumors, although these tumors resemble renal oncocytomas on the transcriptome level, pathologically they were redeagnosed as clear cell papillary RCCs (ccpRCC), a unique RCC subtype known not to metastasize and which lacks known copy number alterations found in other RCC subtypes (46). Consistent with all these findings, Kaplan–Meier analysis on patient survival showed a trend that patients with FGFR2 IIib RCCs had a better clinical outcome (Fig. 5C, Cox proportional hazard regression likelihood ratio test P = 0.047).

**Discussion**

In this study, we carried out an in silico analysis of the ccRCC transcriptome to identify recurring tumor-specific alternative splicing events. Two major types of tumor-specific alternative splicing events identified were alternative transcriptional start sites and exon skipping. For one gene (FGFR2), we identified a mutually exclusive usage of exons. Although previous studies suggested that pyruvate kinase isoform switch from PKM1 to PKM2 in cancers is important for cancer metabolism and tumor growth (3), we did not observe such a switch in paired normal kidney and ccRCCs (data not shown). However, this is consistent with a recent study reporting that such isoform shift does not occur in kidney, lung, liver, and thyroid cancers (47). In another interesting case, we observed disrupted expression of the AP1M2 gene at the 5′ end exons in ccRCC but not in normal kidney (Supplementary Fig. S1). The basis of this phenomenon is

### Table 1. Fisher exact tests of clinical parameters in FGFR2 IIib and IIIc ccRCCs

<table>
<thead>
<tr>
<th>Parameter</th>
<th>IIib Yes</th>
<th>IIib No</th>
<th>IIIc Yes</th>
<th>IIIc No</th>
<th>Fisher exact test, P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distant metastasis (M1)</td>
<td>0</td>
<td>22</td>
<td>70</td>
<td>345</td>
<td>0.034</td>
</tr>
<tr>
<td>Lymph node spread (N1)</td>
<td>0</td>
<td>9</td>
<td>14</td>
<td>203</td>
<td>1</td>
</tr>
<tr>
<td>Tumor size (T1)</td>
<td>18</td>
<td>4</td>
<td>203</td>
<td>212</td>
<td>0.0035</td>
</tr>
<tr>
<td>Tumor grade (G1/2)</td>
<td>13</td>
<td>5</td>
<td>180</td>
<td>233</td>
<td>0.027</td>
</tr>
<tr>
<td>Tumor stage (stage I)</td>
<td>18</td>
<td>4</td>
<td>199</td>
<td>216</td>
<td>0.0019</td>
</tr>
</tbody>
</table>

NOTE: Each row represents data for a 2 × 2 contingency table used for Fisher exact test.
unclear but suggests possible gene translocation, alternative transcription start, or deletion of part of the AP1M2 gene.

Through integrated analysis using data from 3 types of cancers (ccRCC, endometrial, and bladder cancer), we identified ESRP1 and GRHL2 as putative regulators of the FGFR2 IIIc isoform switch (Fig. 2B and C). Identification of ESRP1 is proof-of-principle of our approach as previous in vitro studies have established a role for ESRP1 in controlling the fate of FGFR2 splicing in epithelial cells (12). Our analysis suggests that GRHL2 is also a likely regulator of FGFR2. This gene encodes a transcription factor with an emerging role as a master regulator in maintaining the epithelial phenotype and suppression of EMT (48, 49). In addition, our data revealed a good correlation between ESRP1/GRHL2 expression level with FGFR2 expression level in normal kidney (Pearson’s r of 0.78 and 0.80, respectively), which is lost in ccRCCs (Pearson r of −0.013 and 0.017, respectively) in both IIIc and IIIb tumors (Supplementary Fig. S3). One possible explanation for this phenomenon is that in tumors an additional mechanism(s) such as promoter methylation contributes to FGFR2 transcriptional regulation. Indeed, ccRCCs are subjected to increased DNA methylation proximal to FGFR2 intron 1, which correlates with downregulation of FGFR2 expression (Supplementary Fig. S6).

Figure 5. FGFR2 IIIb ccRCCs are related to renal oncocytoma and chRCC. A, GSEA enrichment score plots showing oncocytoma and chRCC-specific genes are highly enriched in FGFR2-IIIb tumors, whereas ccRCC-specific genes are negatively correlated. B, average gene copy number alterations across the genome in FGFR2-IIIb-1/2 and FGFR2-IIIc ccRCCs. Average gene copy number alterations were plotted for 9 FGFR2-IIIb-1, 9 FGFR2-IIIb-2, and 395 FGFR2-IIIc tumors across the genome. Red and blue represent copy number gain and loss regions. The y-axis ranges from −0.5 to 0.5 of log2 SNP array segment values. C, Kaplan–Meier plot of ccRCCs stratified by FGFR2 isoforms.
EMT is a process that potentiates cell migration and invasion, and is thought to be a critical step in tumor metastasis. Consistent with this notion, we found FGFR2 IIIb ccRCCs, which retain an epithelial phenotype, do not have distant metastasis and lymph node spread (Fig. 4, Table 1). Surprisingly, while EMT is a known phenomenon in cancer, the FGFR2 IIc switch, which is linked to the mesenchymal phenotype and EMT, was rarely detected in cancers except in ccRCC in our analysis of paired normal and tumor samples (Fig. 2A). Possible explanations are either FGFR2 IIc expression per se is not important for EMT or EMT and FGFR2 IIc switch occurs transiently or in a small percentage of tumor cells and thus cannot be detected by our analysis, which used bulk tumor tissues.

About 20% of FGFR2 IIc ccRCCs from the TCGA dataset are recorded to have distant metastasis, a percentage much lower than IIc isoform switch observed in these tumors, which is approximately 90% (Supplementary Table S3). In the context of embryonic kidney development during which both epithelial and mesenchymal stem cells are involved, we suggest that the prevalence of FGFR2 IIc isoform switch points to a mesenchymal stem cell origin for this subtype of ccRCC. This is partially supported by data in Fig. 4A, and provides a plausible explanation for why the FGFR2 IIc isoform switch is kidney-specific. Studies on tumor-initiating population of renal carcinomas also support this idea (41). Three types of EMT have been described (50). Type 1 EMT occurs during embryonic development and is associated with processes, such as implantation and gastrulation. Type 2 EMT is involved in response to inflammation, wound healing, and tissue regeneration. Type 3 EMT occurs in neoplastic cells and is attributed to tumor metastasis. Although we could not exclude the possibility that type 3 EMT occurs in ccRCC during tumor dissemination, we propose that FGFR2 IIc expression in most ccRCCs may be linked to a type 1 EMT in the early developmental stage, when ccRCC first originated in cells having a mesenchymal phenotype.

This notion is further supported by our integrated GSEA and copy number analysis to suggest that FGFR2-IIb ccRCCs are likely to be a mixture of ccpRCCs and chRCCs. Because the cancer genome of FGFR2-IIb-2 ccRCCs possesses distinct marker LOH events on multiple chromosomes (1, 2, 6, 10, 13, 17, and 21) found in chRCCs, this suggests that conversion between chRCC and ccRCC tumors would not happen. This further indicates that FGFR2 isoform switch is an early event during RCC tumorigenesis, and may play an important role in regulation of tumor growth and invasion. Recently, TCGA has initiated construction of a chRCC dataset. We have checked the existing RNAseq data from this new dataset and confirmed that both normal and tumor tissues dominantly express the FGFR2 IIb isoform (data not shown).

Although GSEA analysis suggested that FGFR2-IIb-1 tumors possess a transcriptome resembling that of renal oncocyoma, a reexamination of the tumor histology revealed that these were actually ccpRCCs. This raises the possibility that ccpRCCs might develop from oncocyto-

mas. Future study on gene expression profiling of ccpRCCs will answer the question as whether these tumors are similar to each other on the transcriptional level. Our study is based on significant changes in tumors as a whole, and would likely miss some splicing alterations occurring in only a subset of tumors. For example, an examination of CD44 splicing in ccRCC revealed that loss of expression of CD44 variant exons was enriched in a subset of tumors with better prognosis (corresponding to the tumors clustered on the right side in Fig. 3A, data not shown). Such tumor subtype-specific alterations would be missed in our current study.

Our identification of FGFR2 isoform switching in ccRCC has several potential clinical implications. First, detection of FGFR2 IIIb in RCC not only suggests the presence of ccpRCC or chRCC but also predicts a better prognosis. Second, given the high specificity of FGFR2 IIIc isoform switch, it is important to explore its potential use as an early detection biomarker for ccRCC in patient urine and serum profiles. Third, FGFR2 IIIc and IIIb have different ligand specificity. Although FGFR2 IIIb binds specifically to FGF-1, 3, 7, 10, FGFR2 IIc binds to FGF-1, 2, and 9. Therefore, FGFR2 IIc or its specific ligands (FGF-2 and -9) potentially represent specific targets for ccRCC intervention. Our results show the power of integrated analyses of cancer genomics data toward improved understanding of potential molecular drivers of specific cancers, and suggest new approaches for targeted intervention of human cancers.

Disclosure of Potential Conflicts of Interest

W.K. Alfred Yung has a commercial research grant from Daiichi, has honoraria from speakers’ bureau from Novartis and Merck, and is a consultant/advisory board member of Novartis, Merck, and Actelion. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions

Conception and design: Q. Zhao, W.K. Alfred Yung, R.L. Strausberg, J. Yao Development of methodology: Q. Zhao, O.L. Caballero, E. Jonasch, J. Yao Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): O.L. Caballero, I.D. Davis, J. Yao Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): Q. Zhao, E. Jonasch, P. Tamboli, W.K. Alfred Yung, J. Yao Writing, review, and/or revision of the manuscript: Q. Zhao, I.D. Davis, E. Jonasch, P. Tamboli, W.K. Alfred Yung, J.N. Weinstein, R.L. Strausberg, J. Yao Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): J.N. Weinstein, K. Shaw Study supervision: J. Yao

Acknowledgments

The authors thank Dts. Ralf Krahe and Hongwu Zheng for critical reading of the manuscript and insightful discussions.

Grant Support

This study is supported in part by the TCGA grant U24CA143483 from NCI/NIH, and funds from Ludwig Institute for Cancer Research. I.D. Davis was supported by a Victorian Cancer Agency Clinician Researcher Fellowship.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 2, 2012; revised January 25, 2013; accepted February 4, 2013; published OnlineFirst February 26, 2013.
References


Tumor-Specific Isoform Switch of the Fibroblast Growth Factor Receptor 2 Underlies the Mesenchymal and Malignant Phenotypes of Clear Cell Renal Cell Carcinomas

Qi Zhao, Otavia L. Caballero, Ian D. Davis, et al.


Access the most recent version of this article at: doi:10.1158/1078-0432.CCR-12-3708

Access the most recent supplemental material at: http://clincancerres.aacrjournals.org/content/suppl/2013/02/26/1078-0432.CCR-12-3708.DC1

This article cites 48 articles, 20 of which you can access for free at: http://clincancerres.aacrjournals.org/content/19/9/2460.full#ref-list-1

This article has been cited by 9 HighWire-hosted articles. Access the articles at: http://clincancerres.aacrjournals.org/content/19/9/2460.full#related-urls

Sign up to receive free email-alerts related to this article or journal.

To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

To request permission to re-use all or part of this article, use this link http://clincancerres.aacrjournals.org/content/19/9/2460. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.