Original Article

Biphasic components of sarcomatoid clear cell renal cell carcinomas are molecularly similar to each other, but distinct from, non-sarcomatoid renal carcinomas

Kanishka Sircar,[1,2] Suk-Young Yoo,[3] Tadeusz Majewski,[1] Khalida Wani,[1] Lalit R. Patel,[1] Horatiu Voicu,[4]
Wandaliz Torres-Garcia,[3] Roel G. W. Verhaak,[3,5] Nizar Tannir,[6] Jose A. Karam,[7] Eric Jonasch,[6]
Christopher G. Wood,[7] Pheroze Tamboli,[1] Keith A. Baggerly,[3] Kenneth D. Aldape[8]* and Bogdan Czerniak[1]*

[1] Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA
[2] Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA
[3] Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA
[4] Department of Medicine, Duncan Cancer Center, Baylor College of Medicine, Houston, Texas, USA
[5] Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA
[6] Department of Genitourinary Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA
[7] Department of Urology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA
[8] Department of Pathology, Toronto General Hospital/Research Institute (UHN), Toronto, Ontario, Canada

*Correspondence to: Kenneth Aldape, Department of Pathology, Toronto General Hospital/Research Institute (UHN), Toronto Medical Discovery Tower, 101 College St., Room 14-601, Toronto, Ontario M5G 1L7, e-mail: kaldape@uhnresearch.ca or Bogdan Czerniak, Department of Pathology, Unit 85, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA. e-mail: bczernia@mdanderson.org

Abstract

Sarcomatoid transformation, wherein an epithelioid carcinomatous tumour component coexists with a sarcomatoid histology, is a predictor of poor prognosis in clear cell renal cell carcinoma. Our understanding of sarcomatoid change has been hindered by the lack of molecular examination. Thus, we sought to characterize molecularly the biphasic epithelioid and sarcomatoid components of sarcomatoid clear cell renal cell carcinoma and compare them to non-sarcomatoid clear cell renal cell carcinoma. We examined the transcriptome of the epithelioid and sarcomatoid components of advanced stage sarcomatoid clear cell renal cell carcinoma (*n*=43) and non-sarcomatoid clear cell renal cell carcinoma (*n*=37) from independent discovery and validation cohorts using the cDNA microarray and RNA-seq platforms. We analyzed DNA copy number profiles, generated using SNP arrays, from patients with sarcomatoid clear cell renal cell carcinoma (*n*=10) and advanced non-sarcomatoid clear cell renal cell carcinoma (*n*=155). The epithelioid and sarcomatoid components of sarcomatoid clear cell renal cell carcinoma had similar gene expression and DNA copy number signatures that were, however, distinct from those of high-grade, high-stage non-sarcomatoid clear cell renal cell carcinoma. Prognostic clear cell renal cell carcinoma gene expression profiles were shared by the biphasic components of sarcomatoid clear cell renal cell carcinoma and the sarcomatoid component showed a partial epithelial-to-mesenchymal transition signature. Our genome-scale microarray-based transcript data were validated in an independent set of sarcomatoid and non-sarcomatoid clear cell renal cell carcinomas using RNA-seq. Sarcomatoid clear cell renal cell carcinoma is molecularly distinct from non-sarcomatoid clear cell renal cell carcinoma, with its genetic programming largely shared by its biphasic morphological components. These data explain why a low percentage of sarcomatoid histology augurs a poor prognosis; suggest the need to modify the pathological grading system and introduce the potential for candidate biomarkers to detect sarcomatoid change preoperatively without specifically sampling the histological sarcomatoid component.

Keywords: sarcomatoid; renal; carcinoma; clear cell; expression; RNA-seq; molecular

Introduction

Sarcomatoid change is a microscopically defined entity; it manifests as a biphasic histological pattern with a better differentiated, parent epithelioid (E) component that resembles typical carcinoma and a dedifferentiated sarcomatoid (S) component with spindled morphological characteristics that resembles a mesenchymally derived sarcoma. It is known to be associated with a poor prognosis in cancers from various organs, including the kidneys [1–4]. Sarcomatoid renal cell carcinoma (RCC) is notable in that it represents the most aggressive, treatment-resistant type of RCC, accounting for almost 20% of stage IV RCCs [5–7] with a median survival duration of less than 1 year. [6–8] Thus, sarcomatoid RCC contributes significantly to RCC-specific mortality [9].

The presence of sarcomatoid change alters clinical decision-making because such tumours are often treated differently from non-sarcomatoid RCC [6,10–12]. As there is no standard systemic treatment protocol for this aggressive variant, patients are encouraged to participate in clinical trials, in which the presence and proportion of the spindled sarcomatoid component is currently used for enrolment and stratification [13].

Despite its clinical importance, sarcomatoid RCC is poorly understood at the genetic level. There have been no genome-wide studies of this biphasic tumour reported to date, largely because of the difficulty in identifying and harvesting frozen epithelioid and sarcomatoid tumour tissue. The molecular characterization of sarcomatoid transformation in RCC thus represents an unmet need of major clinical importance. Our aim was to gain a better molecular understanding of sarcomatoid clear cell RCC (ccRCC) by performing a genome-wide examination of this entity at the transcript and DNA copy number levels. We found the genomic landscape of the E and S components of sarcomatoid ccRCC to be remarkably similar but sharply distinct from non-sarcomatoid ccRCC. These results help explain the aggressive clinical course of ccRCC with even a minor sarcomatoid element, challenge the existing grading system of sarcomatoid ccRCC, and represent an essential first step to developing a panel of biomarkers that can preoperatively detect sarcomatoid change.

Materials and Methods

Patient and tumour characteristics

For this retrospective study, we obtained ccRCC tissue samples from the Department of Pathology at The University of Texas MD Anderson Cancer Center (Houston, Texas) after informed consent and using an institutional review board-approved protocol (IRB# LAB 08-670). The clinicopathological characteristics of the tumour samples and patient cohorts used in this study are summarized in Table 1 and Supplementary Material Table 1.

We included formalin-fixed, paraffin-embedded (FFPE) samples that had been resected from patients with sarcomatoid and non-sarcomatoid ccRCC. Samples from advanced stage non-sarcomatoid ccRCCs (Epithelioid* (E*)) were used as controls. Lesional foci (E, S and E*) were marked on H&E-stained slides and all cases were reviewed by at least two genitourinary pathologists. Only sections with >70% cancer cells were included, with lesional foci from sarcomatoid ccRCC (E/S) and advanced stage non-sarcomatoid ccRCC (E*) macrodissected as indicated in Figure 1.

This study was conducted on sample sets and related patient cohorts divided into three groups: discovery and validation cohorts as well as the clear cell renal carcinoma The Cancer Genome Atlas (TCGA) cohort. We first studied the gene expression profiles, using cDNA microarray, of 58 patients from the discovery cohort composed of 36 sarcomatoid E and S pairs and 22 non-sarcomatoid E* samples. Four sarcomatoid ccRCC pairs from this discovery cohort were also analyzed using the RNA-seq platform. For the validation cohort, we used RNA-seq to interrogate 22 ccRCC patient samples comprising 7 sarcomatoid E and S pairs and 15 non-sarcomatoid E* cases. We next performed a genome-wide DNA copy number analysis of 10 sarcomatoid ccRCC and used TCGA DNA copy number data from 155 advanced stage ccRCC patients as an E* reference set. TCGA gene expression data from 41 MD Anderson patients was also used for *in silico* analysis in this study.

Microarray–based gene expression profiling and analysis

Total cellular RNA was isolated according to the manufacturer's protocol (Epicentre Biotechnologies) after deparaffinization and proteinase K treatment. RNA samples were normalized using the Ribogreen RNA quantitation kit (Life Technologies) for the whole-genome cDNA-mediated annealing, selection, extension and ligation HT assay. Normalized RNA was converted to cDNA and incubated on Illumina HumanHT-12v4 BeadChips. The slides were scanned using a BeadArray Reader and the signal intensities were quantified using GenomeStudio software. We performed 2-sample *t*-tests to compare expression data for E versus

Table 1. Clinicopathological characteristics of sarcomatoid and non-sarcomatoid ccRCC cases

Discovery cohort, gene expression	No.	Platforms	Pretreatment	Age, years	Sarcomatoid histology (%)	Stage				Fuhrman grade				Died of disease	Follow-up, weeks
						I	II	III	IV	1	2	3	4		
Sarcomatoid ccRCC	36	cDNA microarray	None, n=27	56 (40–76)	30 (10–90)	0	1	6	20	0	0	0	27	21 (78%)	56 (2–715)
			C, n=4; S, n=5	61 (40–73)	40 (15–90)	0	0	0	9	0	0	0	9	8 (89%)	42 (8–172)
		RNAseq	None, n=3	67 (48–69)	40 (15–90)	0	0	3	0	0	0	0	3	2 (67%)	49 (52–282)
			S, n=1	55 (55)	40 (40)	0	0	0	1	0	0	0	1	1 (100%)	23 (23)
Non-sarcomatoid ccRCC	22 / **58**	cDNA microarray	None, n=22	69 (56–88)	0 (0)	0	0	7	15	0	4	10	8	16 (73%)	53 (1–653)
Validation cohort, gene expression															
Sarcomatoid ccRCC	7	RNAseq	None, n=6	60 (50–79)	50 (20–80)	0	0	1	5	0	0	0	6	4 (67%)	39 (10–107)
			S, n=1	35 (35)	60 (60)	0	0	0	1	0	0	0	1	1 (100%)	13 (13)
Non-sarcomatoid ccRCC	15 / **22**	RNAseq	None, n=15	64 (46–85)	0 (0)	0	0	7	8	0	0	9	6	8 (53%)	224 (2–452)
DNA copy number															
Sarcomatoid ccRCC	10	SNP array	None, n=9	55 (45–76)	50 (10–90)	0	0	2	7	0	0	0	9	8 (89%)	15 (1–436)
			C, n=1	69 (69)	90 (90)	0	0	0	1	0	0	0	1	1 (100%)	66 (66)
Non-sarcomatoid ccRCC*	155 / **165**	SNP array	None, n=155	62 (33–88)	0 (0)	0	0	100	55	0	44	92	19	72 (46%)	121 (0–412)

C, Chemotherapy; S, Sutent.
Bold indicates total number of case summed.
*TCGA data; Data are expressed as median (range) or n (%).

S, E/S versus E* and E/S versus Fuhrman grade 4 E* (Supplementary Material Methods).

In silico analysis of The Cancer Genome Atlas clear cell renal cell carcinoma gene expression data

We analyzed TCGA expression data from a cohort of stage III and IV non-sarcomatoid ccRCC tumours from MD Anderson using annotated clinical factors, such as tumour grade and survival [14]. Highly over-expressed coding genes found in sarcomatoid tumours from our experiments were studied in this cohort of TCGA non-sarcomatoid ccRCC samples to determine whether the expression of these genes differed according to tumour grade (ie G2, G3 or G4) or patient survival using the Kruskal–Wallis statistical test.

DNA copy number assessment and analysis

Lesional tissues from sarcomatoid ccRCC and normal adjacent kidneys were macrodissected and processed for DNA isolation using the BiOstic FFPE tissue kit. DNA concentration and quality were determined using the Nanodrop spectrophotometer (Thermo Scientific) and the Invitrogen Quant-iT PicoGreen dsDNA assay kit. The DNA copy number was assessed using the high-resolution SNP genotyping array (HumanOmniExpress FFPE-12 v1.0, Illumina).

To perform a DNA copy number analysis of non-sarcomatoid ccRCC cases, we used TCGA data from patients with stage III and IV non-sarcomatoid ccRCC derived from Affymetrix SNP 6 arrays run by the Broad Institute (Supplementary Material Table 1). Further details are provided in the Supplementary Material Methods.

RNA-seq based gene expression profiling and analysis

To better evaluate the E to S transition and confirm our microarray-based data, we used the complementary next-generation RNA-seq platform to generate gene expression data. We first assessed E/S pairs (n=4) for which we had microarray data using paired-end sequencing on an Illumina GA-IIx. We then assayed an independent set of E/S pairs (n=7 pairs) and E* (n=15) cases that was subjected to paired-end sequencing on an Illumina HiSeq2000. Data was analyzed across sample subtypes with detailed methodology provided in Supplementary Material Methods.

Statistical analysis

A *p* value ≤ 0.05 was considered statistically significant except in multiple comparisons, in which the
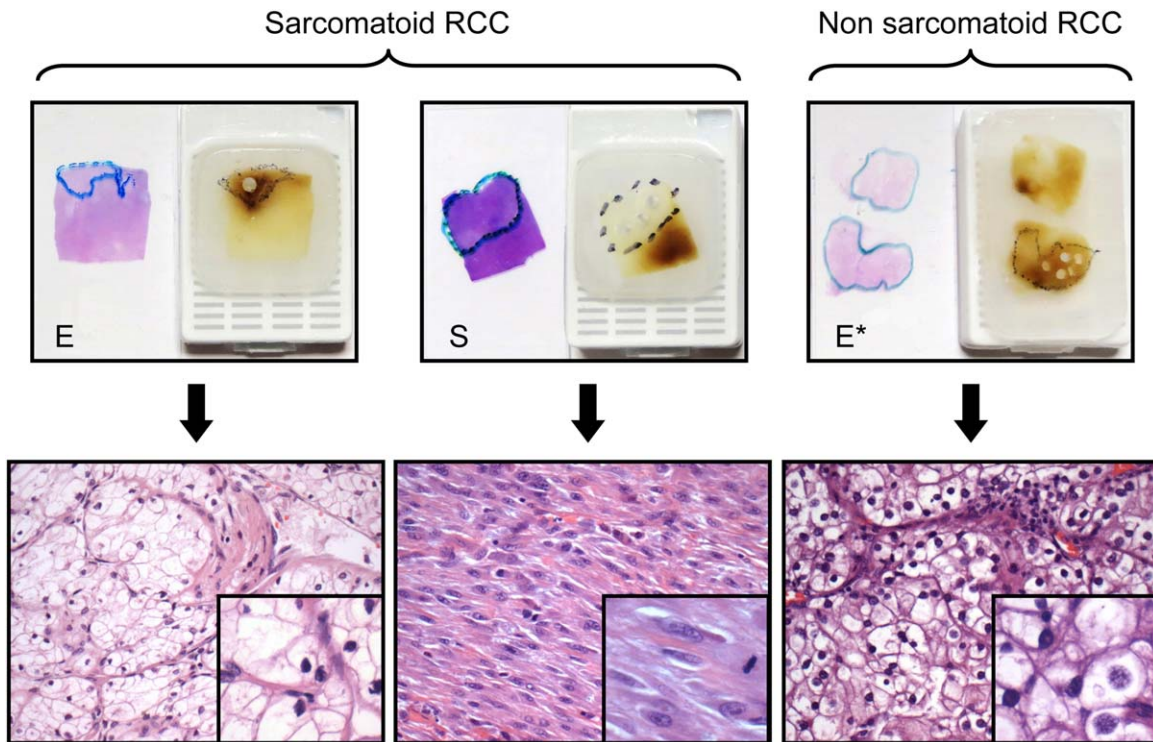
Figure 1. Macrodissected biphasic sarcomatoid and non-sarcomatoid ccRCCs. The paired E and S components of sarcomatoid ccRCC and the E* component of non-sarcomatoid ccRCC, macrodissected as illustrated above (H&E stain, inset).

false discovery rate (FDR) was controlled by requiring $q \leq 0.05$.

Results

E and S components of sarcomatoid clear cell renal cell carcinoma show a similar gene expression pattern that differs from that of non-sarcomatoid clear cell renal cell carcinoma

An unsupervised clustering analysis of gene expression microarray data across the biphasic E and S components of sarcomatoid ccRCC and advanced stage non-sarcomatoid ccRCC showed few significant differences in global gene expression between the E and S groups in sarcomatoid ccRCC. However, non-sarcomatoid ccRCC (E*) had many more differentially expressed genes (n=873 genes, FDR<0.001) compared to sarcomatoid ccRCC (E/S), as shown in Figure 2A and Supplementary Material Figure 1A. The contrast between sarcomatoid ccRCC and non-sarcomatoid ccRCC was maintained when considering only Fuhrman grade 4 non-sarcomatoid ccRCC (n=263 genes, FDR<0.01, Figure 2B and Supplementary Material Figure 1B).

We evaluated pretreated sarcomatoid ccRCC samples to determine whether neoadjuvant therapy had different effects on the E or S components. We performed a supervised analysis of sunitinib-pretreated tumours with respect to hypoxia-inducible factor (HIF) pathway genes and of chemotherapy-pretreated tumours with respect to proliferation and apoptosis genes. No significant differences were seen between the E or S components of sunitinib-pretreated ccRCC with respect to the expression of HIF pathway genes (Supplementary Material Figure 2A), suggesting that antiangiogenic therapy did not exert a differential effect on the biphasic histological elements. Chemotherapy-pretreated ccRCC likewise showed no significant gene expression differences between the E and S components with respect to apoptosis-related genes (data not shown); however, E and S elements clustered separately when evaluating proliferation genes whereas untreated sarcomatoid ccRCC did not show such spontaneous clustering (Supplementary Material Figures 2B, C).

Next-generation sequencing performed on 4 E/S pairs that had already been interrogated by gene expression microarray did not reveal any spontaneous clustering among E or S samples across RefSeq-annotated genes (Supplementary Material Figure 3A).
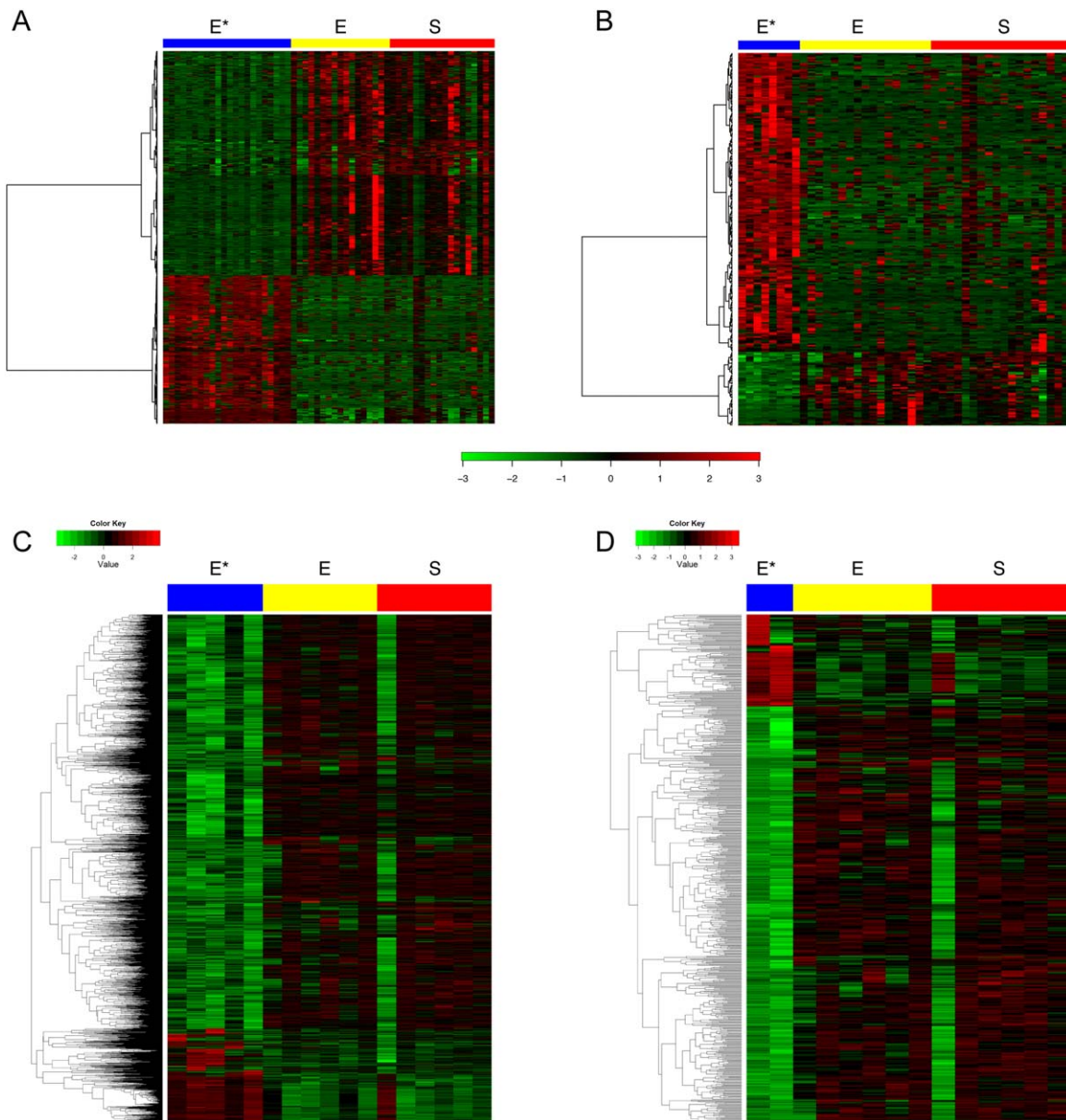
Figure 2. E and S components of sarcomatoid ccRCC show a similar gene expression signature that differs from that of non-sarcomatoid ccRCC. (A) Microarray analysis shows the distinctive expression profile of sarcomatoid ccRCC compared with that of advanced-stage non-sarcomatoid ccRCC, with a heatmap of the 873 significant probes in the non-sarcomatoid (E*) and sarcomatoid (E and S) samples contrasting at a FDR of 0.001. (B) Microarray analysis shows the distinctive expression profile of sarcomatoid ccRCC versus Fuhrman grade 4 non-sarcomatoid ccRCC, with a heatmap of the 263 significant probes in the Fuhrman grade 4 non-sarcomatoid (E*) and sarcomatoid (E and S) samples contrasting at a FDR of 0.01. (C) An RNA-seq analysis shows the distinctive expression profile of sarcomatoid ccRCC versus advanced-stage non-sarcomatoid ccRCC, with a heatmap of the 2549 significant probes in the non-sarcomatoid (E*) and sarcomatoid (E and S) samples contrasting at a FDR of 0.05. (D) An RNA-seq analysis shows the distinctive expression profile of sarcomatoid ccRCC versus Fuhrman grade 4 non-sarcomatoid ccRCC, with a heatmap of the 657 significant probes in the Fuhrman grade 4 non-sarcomatoid (E*) and sarcomatoid (E and S) samples contrasting at a FDR of 0.05. Gene expression values were centered. Samples are ordered by subtype in columns, and genes are in rows.

Further, no hierarchical clustering or differential expression was seen between the E and S components of these tumours with respect to a prognostic signature (Supplementary Material Figures 3B–D). These findings support the microarray-based data we had obtained, now using the orthogonal RNA-seq platform.

We next performed an RNA-seq analysis on an independent validation set of sarcomatoid and

non-sarcomatoid ccRCC tumours (Table 1). We again found few significant expression differences between the biphasic E and S components of sarcomatoid ccRCC. However, sarcomatoid ccRCC had numerous differentially expressed genes compared to non-sarcomatoid ccRCC (n=2549 genes, FDR<0.05) (Figure 2C and Supplementary Material Figure 1C), including Fuhrman grade 4 non-sarcomatoid ccRCC (n=657 genes, FDR<0.05) (Figure 2D and Supplementary Figure 1D). Druggable pathways (mTOR, HIF, MAPK/ERK) in ccRCC also showed differential regulation in sarcomatoid versus non-sarcomatoid ccRCC (Supplementary Material Figures 4A–C).

In evaluating the contrast between sarcomatoid and non-sarcomatoid ccRCC in independent samples using both microarray and RNA-seq methods, we obtained differential gene lists that we ranked according to the results of our pathway analysis (Supplementary Material Table 2). An enrichment P value based on a hypergeometric distribution was used to determine that the number of common pathways out of the top 200 from each experiment was greater than that expected by chance (p <0.001) (Figure 3A). Pathways that were consistently enriched in sarcomatoid ccRCC compared to non-sarcomatoid ccRCC using both platforms are shown in Figure 3B.

Candidate biomarkers show abrupt upregulation with sarcomatoid change and do not correlate with tumour grade or survival in non–sarcomatoid clear cell renal cell carcinoma

We next examined the most highly overexpressed individual coding genes in sarcomatoid ccRCC versus non-sarcomatoid ccRCC (fold change >5) because these can function as potential biomarkers of sarcomatoid change. We sought to better understand the behavior of these candidate genes in the non-sarcomatoid ccRCC setting. For this purpose, we interrogated TCGA expression data in an advanced stage cohort of non-sarcomatoid MD Anderson ccRCC tumours.

Interestingly, we found that only 2 among 23 of our top candidate genes showed significantly increased expression with grade progression or vital status (p<0.05). The large majority of our candidate genes did not correlate with tumour grade or survival in non-sarcomatoid ccRCC (Supplementary Material Table 3). Together, our results and analysis of TCGA data suggest that an abrupt upregulation occurs in the expression of these markers in sarcomatoid ccRCC rather than a gradual change with increasing tumour grade, making them suitable for use as biomarkers for detecting sarcomatoid change.

Among the gene candidates derived from microarray data, a subset (RUNX2, IGSF6, RPB1, ALDH1A3, PTPRC, PILRA and BAT5) were also significantly overexpressed in sarcomatoid ccRCC on the basis of an RNA-seq evaluation of independent samples. For example, Runt-related transcription factor 2 (RUNX2), a transcription factor and a target of the transforming growth factor-β1 pathway, [15] was a promising candidate gene for sarcomatoid change on the basis of our expression data and an in silico analysis of TCGA samples (Figure 3C).

E and S components of sarcomatoid clear cell renal cell carcinoma have a similar DNA copy number profile that is distinct from that of non-sarcomatoid clear cell renal cell carcinoma

In assessing DNA copy number changes, we did not find any genes with significant recurrent copy differences between the E/S pairs of sarcomatoid ccRCC. We did, however, find many genes that were significantly associated with copy number variations between the non-sarcomatoid (E*) and sarcomatoid (E/S) groups (1648 genes, FDR<0.05). A subset analysis of Fuhrman grade 4 non-sarcomatoid ccRCCs from MD Anderson (n=19) also revealed significant copy number differences compared to sarcomatoid ccRCC (347 genes, FDR<0.05).

A schematic representation of genome-wide copy number aberrations among the E, S and E* subtypes is illustrated in Figures 4A–C. There are multiple regions of losses and gains across the genome, attesting to the greater karyotypic complexity and distinctiveness of sarcomatoid ccRCC. The extent and proportion of samples showing losses at 3p is markedly lower in the sarcomatoid group. Known poor-prognosis changes (ie 9p loss and 14q loss) are also much less prominent among sarcomatoid ccRCC tumours. The 5q gain that has been reported to confer a better prognosis in ccRCC [16] is largely absent in our sarcomatoid cohort.

Sarcomatoid clear cell renal cell carcinoma has a poor-prognosis gene expression signature that is shared by its E and S components

We performed a supervised analysis of the good prognosis (ccA) and poor prognosis (ccB) gene sets to determine the contrast between sarcomatoid (E/S) and advanced stage non-sarcomatoid (E*) ccRCC. Most of the ccA and ccB genes did not show differential expression, which was expected since both cohorts represent poor-prognosis ccRCC tumours and both would be expected to cluster toward the ccB
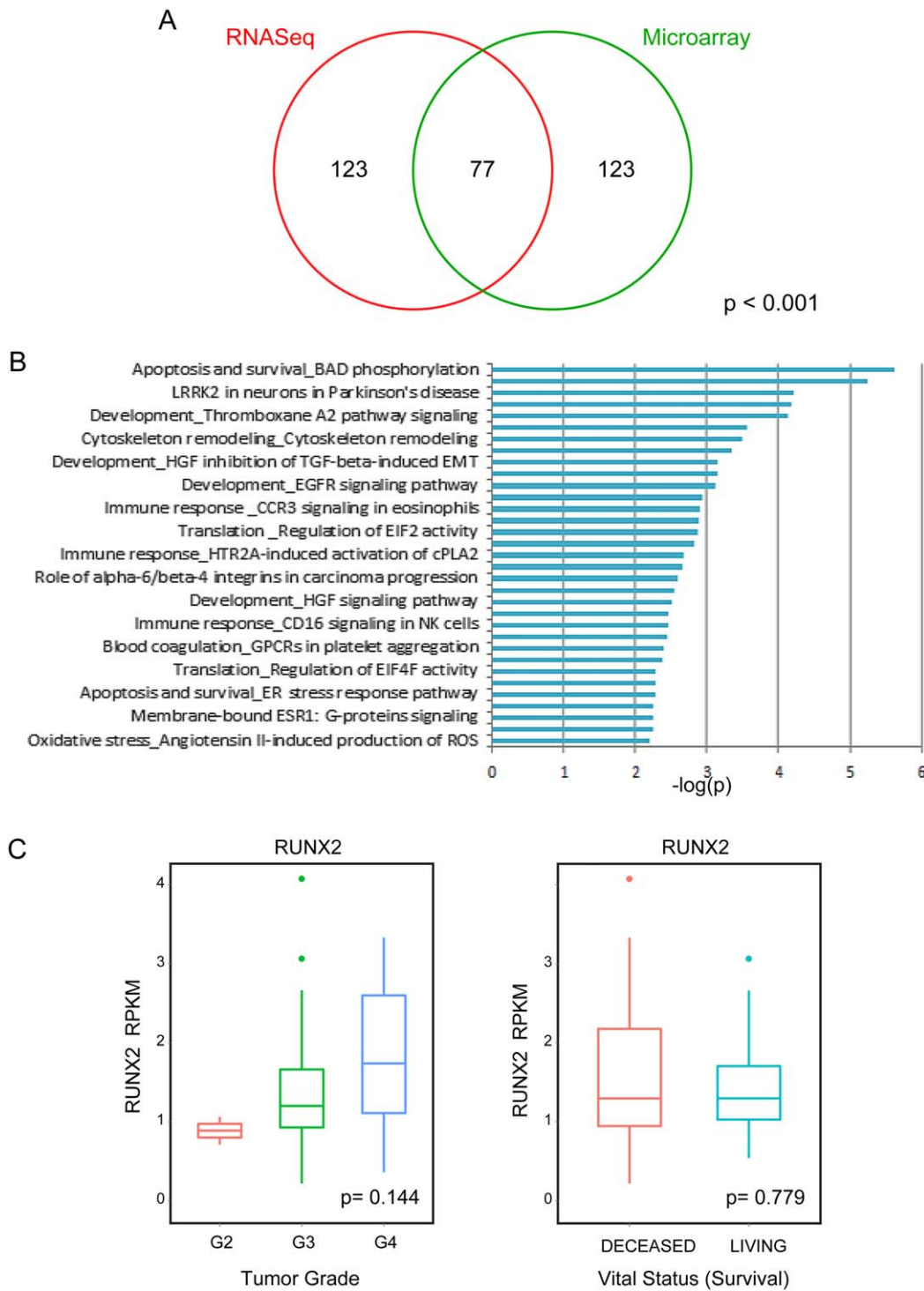
Figure 3. Sarcomatoid ccRCC shows distinct pathway alterations, whereas biomarkers of sarcomatoid change do not correlate with tumour grade or survival in non-sarcomatoid ccRCC. (A) Pathway alterations in sarcomatoid ccRCC compared to those in non-sarcomatoid ccRCC show significant overlap when evaluated by microarray and RNA-seq analyses. A Venn diagram showing commonly altered pathways between sarcomatoid and non-sarcomatoid ccRCC, as evaluated by microarray and RNA-seq analyses using independent samples ($p < 0.001$). (B) The commonly altered pathways between sarcomatoid ccRCC and non-sarcomatoid ccRCC at a FDR < 0.05. (C) RUNX2 expression values by tumour grade and survival. Boxplots for RUNX2 gene expression are shown for tumour grades G2, G3, G4 and stratified by deceased versus living patients. No significant differences were observed in the expression of RUNX2 between different tumour grades ($p=0.144$) or between deceased versus living patients ($p=0.779$).
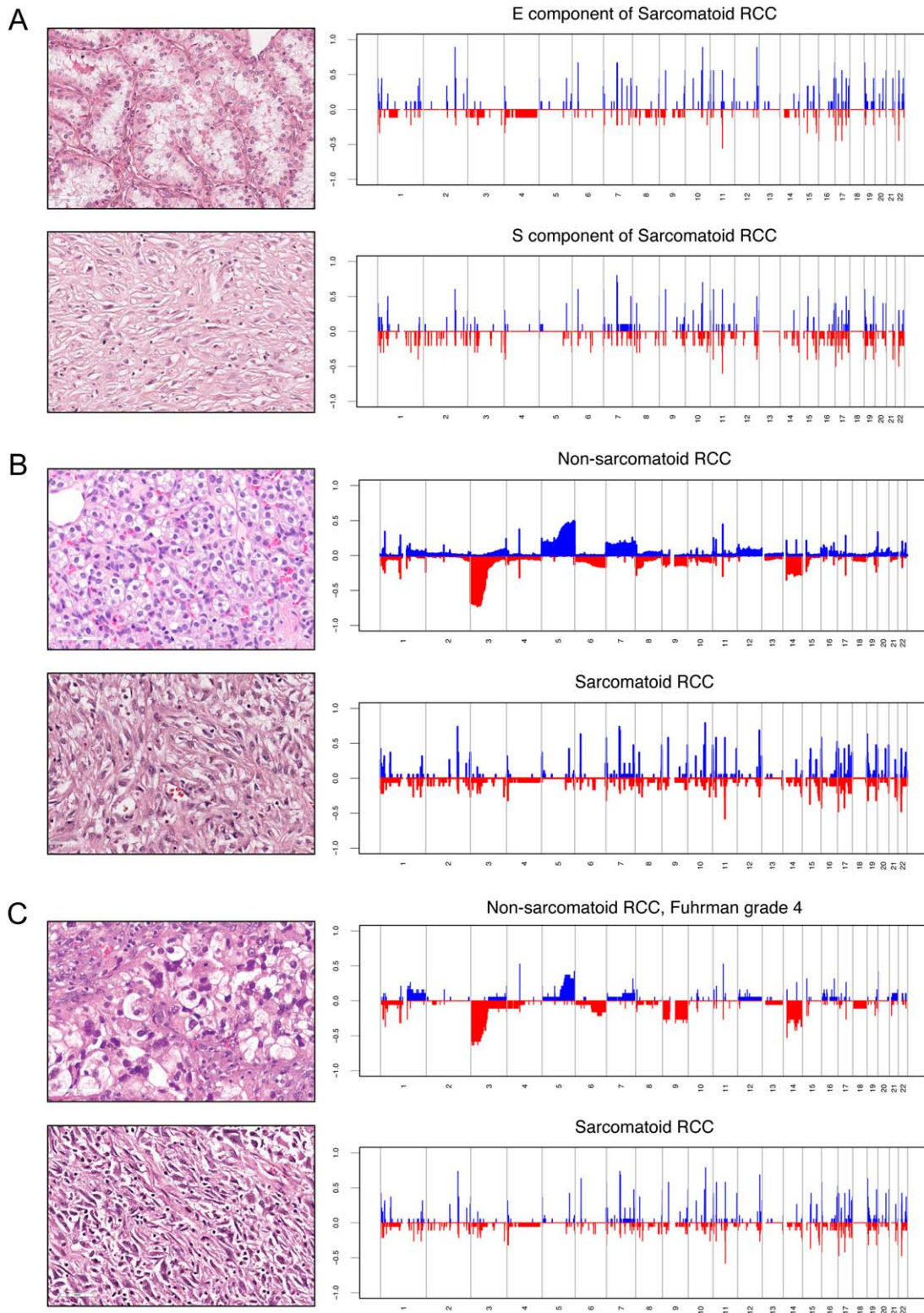
Figure 4. E and S components of sarcomatoid ccRCC show similar DNA copy number aberrations that differ from those of non-sarcomatoid ccRCC. (A) Similar genome-wide DNA copy number signature of E and S components of sarcomatoid ccRCC. (B) Distinct genome-wide DNA copy number signature of sarcomatoid ccRCC versus advanced-stage non-sarcomatoid ccRCC (E*) TCGA cases. (C) Distinct genome-wide DNA copy number signature of sarcomatoid ccRCC versus Fuhrman grade 4 non-sarcomatoid ccRCC (E*) TCGA cases. Copy number alterations are mapped according to their chromosomal location on the x-axis, with the y-axis showing the percentage of cases that harbour these changes.

subtype. However, among those genes that did show differential expression, sarcomatoid ccRCC tumours were associated with ccB genes and non-sarcomatoid ccRCC tumours were associated with the ccA set (p=0.002). This finding supports the clinical observation that sarcomatoid ccRCC represents the extreme end of the poor-prognosis spectrum in ccRCC.

When we restricted our analysis of microarray data to the sarcomatoid ccRCC group, however, we found that the poor-prognosis ccB genes did not significantly differ between the E and S histological elements (Supplementary Material Table 4), indicating that the ccB signature was already embedded in the E phenotype. Similarly, most ccA genes did not differ in expression between the E and S phenotypes. Interestingly, 15 of the 16 ccA genes that showed differential expression had higher expression in the E histological type (Supplementary Material Table 5).

The findings of our RNA-seq analysis agreed with those of the microarray analysis: although a difference was seen between the expression of prognostic genes in non-sarcomatoid and sarcomatoid ccRCC, no difference was seen between the E and S components of sarcomatoid cases (Figure 5A, Supplementary Material Figure 3B). Only 3 ccA genes showed differential expression between the E and S phenotypes, with all 3 genes overexpressed in the E component (Figure 5B). No ccB genes showed significant differential expression between E and S (Figure 5C, Supplementary Material Figure 3D). Thus, RNA-seq confirmed the overall similarity in prognostic gene expression between E/S pairs in the initial discovery set and the independent validation set.

A subset of epithelial to mesenchymal transition genes correlates with the sarcomatoid morphological phenotype of ccRCC

EMT is thought to contribute to cancer progression, as epithelial cells acquire a more mesenchymal phenotype and are able to migrate and invade surrounding stroma. As sarcomatoid ccRCC is generally thought to be a classic example of a histological 'mesenchymal' phenotype, we interrogated the E and S phenotypes with respect to 51 EMT-related genes. On microarray analysis, 42 of 51 EMT genes exhibited no expression difference between the E and S histological types. The 9 EMT genes that showed significant expression differences correlated with the S phenotype (Supplementary Material Table 6A). An RNA-seq analysis of independent samples confirmed these findings (Supplementary Material Table 6B). Differentially expressed EMT genes between the biphasic components of sarcomatoid ccRCC across
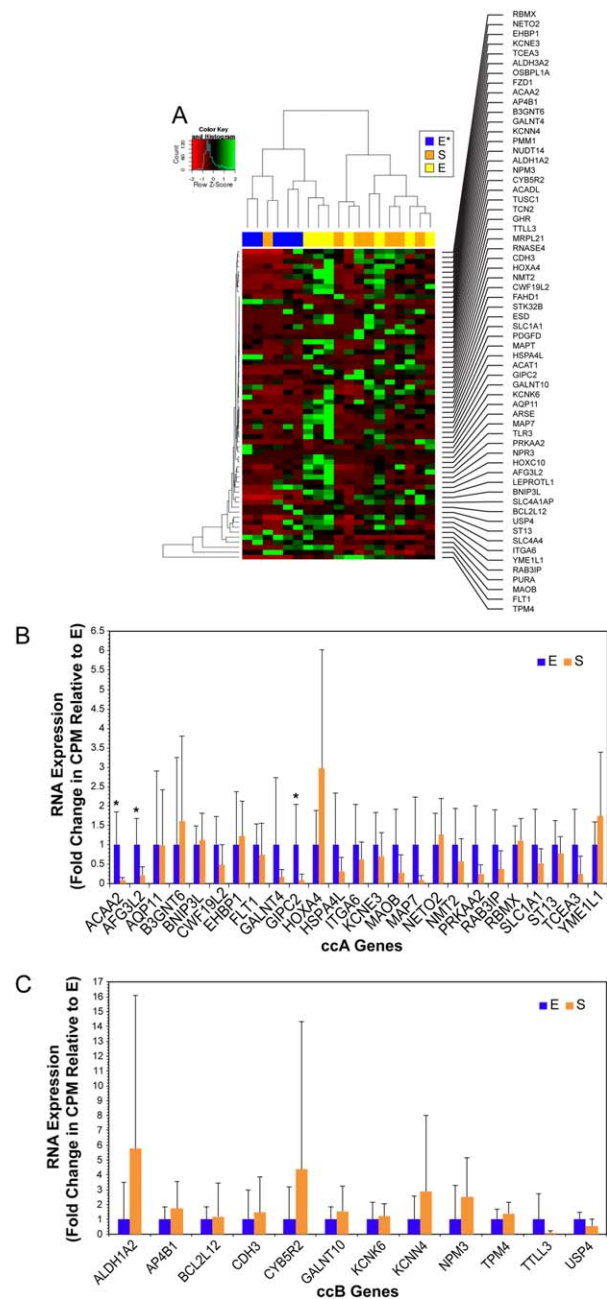

Figure 5. Prognostic gene expression signature of sarcomatoid ccRCC is embedded in its E component. (A) Hierarchical clustering by gene expression profile across a panel of known good-prognosis (ccA) and poor-prognosis (ccB) genes demonstrates a difference in gene expression between non-sarcomatoid (E*) and sarcomatoid (E and S) cases. (B) Relative expression of ccA genes in E and S components demonstrates no significant difference between components for most good-prognosis genes. (C) Relative expression of ccB genes in E and S components demonstrates no significant difference between components in the expression of poor-prognosis genes (*p<0.05).

the microarray and RNA-seq platforms are plotted in Supplementary Material Figures 5A, B.

Discussion

Our core finding that the biphasic E and S components of sarcomatoid ccRCC are similar at the copy number and transcript levels but are markedly different from non-sarcomatoid ccRCC suggests a model in which the major division occurs between the broad categories of sarcomatoid and non-sarcomatoid ccRCC and where most of the genetic programming for sarcomatoid ccRCC is embedded in its E component. The morphological S phenotype would appear to be a histological marker for the overall tumour and not necessarily the most aggressive clone. The observation [17] that both E and S histological tumour components are present in metastatic RCC lesions, sometimes in the same patient, is in line with this concept. Our finding that the poor-prognosis gene expression signature is shared by the E and S elements suggests that even a small amount of sarcomatoid differentiation is dangerous and may help to explain why the percentage of S component has historically not been shown to be correlated with clinical outcome measures [7,8,18,19].

The presence of a similar expression signature in the E and S components of sarcomatoid RCC raises the question of whether sarcomatoid change is a secondary phenomenon related to the progression of tumours that were initially truly epithelial or if there is a subset of tumours that are sarcomatoid from the outset. The fact that the vast majority of sarcomatoid renal cell carcinomas present at high stage, with the median size of the primary tumour being approximately 10 cm [8,20,21], suggests a secondary event. However, the description of isolated small tumours measuring less than 2 cm in large series of sarcomatoid RCC [20,21], implies that the possibility of a *de novo* sarcomatoid phenotype cannot be excluded.

In support of our data, prior studies have shown that the sarcomatoid component maintains the immunohistochemical phenotype and some genetic features of its parent epithelioid RCC tumour component [22–25], while a common cell of origin was postulated [26] based on X-chromosome inactivation data. DNA copy number evaluation of sarcomatoid RCC has been scant, with one study of 12 patients with sarcomatoid carcinoma derived from various RCC subtypes and using an older, lower resolution comparative genomic hybridization (CGH) platform showing a high frequency of copy number aberrations [24,25]. Among the foregoing studies, both the epithelioid and sarcomatoid components were assessed in only two cases—where they showed a similar profile. The similarity of chromosomal aberrations within a given tumour despite sampling from areas with different regional Fuhrman grades was also demonstrated by a Fluorescence *in situ* Hybridization (FISH) based analysis of RCC tumours [27]. Moreover, a FISH study of sarcomatoid chromophobe RCC revealed that the epithelioid and sarcomatoid pairs harbored multiple chromosomal copy number gains, which were different from the genome wide copy number losses that characterized non-sarcomatoid chromophobe RCC [28].

Nevertheless, some differences between the biphasic components have been reported: the S component displayed more mitoses [29] and had a higher proliferation index [30]. We also found subtle differences, with the sarcomatoid histological element showing a partial EMT molecular fingerprint, in line with work by Conant *et al* [31] and Bostrom *et al* [32] who also related sarcomatoid transformation to EMT.

From a pathological standpoint, the distinctiveness of sarcomatoid ccRCC suggests that it should not be grouped with or graded in the manner of non-sarcomatoid ccRCC. The current convention is to assign a Fuhrman nuclear grade 4 for sarcomatoid ccRCC. Our analyses, which draw on our own data as well as TCGA data, show that sarcomatoid ccRCC has a distinct copy number and transcriptomic signature compared to non-sarcomatoid Fuhrman grade 4 ccRCC. On the basis of this molecular distinctiveness and given that sarcomatoid RCC has historically shown more aggressive clinical behavior than has Fuhrman grade 4 RCC without a sarcomatoid component[7,33], it would appear that sarcomatoid change lies outside the Fuhrman grading system. Categorizing sarcomatoid change as Fuhrman grade 4 underestimates its aggressiveness and does not accurately account for its distinct biological characteristics. We therefore suggest reconsidering grading for sarcomatoid ccRCC; at a minimum, the presence of a sarcomatoid component should always be noted in ccRCC.

There are no established, effective therapies for sarcomatoid RCC [13,34]. Chemotherapy and antiangiogenic therapies have been explored, given sarcomatoid ccRCC's high proliferation rate and the continued expression of HIF pathway markers [23]. It has been suggested that a lower percentage of sarcomatoid component predicts for a better therapeutic response to antiangiogenic targeted therapy [35]. Our data, however, showed that sunitinib antiangiogenic therapy did not differentially affect the E or S elements with respect to the expression of HIF pathway genes. Thus, our results call into question the rationale for

stratifying patients in clinical trials or otherwise managing them differently on the basis of the percentage of the sarcomatoid histological component.

The clinical particularity of sarcomatoid RCC has long been appreciated, but sarcomatoid change has been difficult to detect preoperatively using imaging modalities. Examination of renal biopsy samples has also not been effective, with sarcomatoid change diagnosed in only 12% of sarcomatoid RCC cases by biopsy [36]. In this context, our candidate genes represent promising biomarkers that can be applied to biopsy material, and given the similarity in expression between E and S components, they would obviate the need to specifically sample the sarcomatoid or S morphological element. Sarcomatoid histological characteristics may represent a minor component of the overall tumour and can be difficult for a pathologist to diagnose in scant biopsy material that shows a distorted architecture and cytology. The preoperative detection of sarcomatoid change would thus be more feasible in individual patients and would, in turn, alter clinical decision-making in terms of the type of surgical procedure offered in the nonmetastatic setting and whether to proceed with upfront systemic therapy—without cytoreductive nephrectomy—in those with metastatic RCC.

Finally, it must be acknowledged that we have examined the parameters of gene expression and DNA copy number in this report. Other events, including genomic mutation, microRNA expression, epigenetic regulation, post-translational modification and microenvironmental influences may offer further insights into the pathobiology of sarcomatoid change in ccRCC.

Acknowledgments

Author contributions

KS, KA and BC conceived and designed the study. KS, KB, KA and BC developed the methodology. KS, TM, KW and PT collected data and performed experiments. KS, KA, KB, SY, LP, HV and WG analyzed and interpreted data. EJ, NT, CW, JK and RV provided advice, reviewed and edited the manuscript. All authors read and approved the final manuscript.

References

1. Martin LW, Correa AM, Ordonez NG, et al Sarcomatoid carcinoma of the lung: a predictor of poor prognosis. *Ann Thorac Surg* 2007; **84**: 973–980.

2. Choi YY, Jeen YM, Kim YJ. Sarcomatoid carcinoma of colon: extremely poor prognosis. *J Korean Surg Soc* 2011; **80** (Suppl 1): S26–S30.

3. Wang J, Wang FW, Lagrange CA, et al. Clinical features of sarcomatoid carcinoma (carcinosarcoma) of the urinary bladder: analysis of 221 cases. *Sarcoma* 2010; **2010**: 454792.

4. Carter MR, Hornick JL, Lester S, et al. Spindle cell (sarcomatoid) carcinoma of the breast: a clinicopathologic and immunohistochemical analysis of 29 cases. *Am J Surg Pathol* 2006; **30**: 300–309.

5. Margulis V, Matin SF, Tannir N, et al. Surgical morbidity associated with administration of targeted molecular therapies before cytoreductive nephrectomy or resection of locally recurrent renal cell carcinoma. *J Urol* 2008; **180**: 94–98.

6. Shuch B, Said J, La Rochelle JC, et al. Cytoreductive nephrectomy for kidney cancer with sarcomatoid histology–is up-front resection indicated and, if not, is it avoidable? *J Urol* 2009; **182**: 2164–2171.

7. Cheville JC, Lohse CM, Zincke H, et al. Sarcomatoid renal cell carcinoma: an examination of underlying histologic subtype and an analysis of associations with patient outcome. *Am J Surg Pathol* 2004; **28**: 435–441.

8. Shuch B, Bratslavsky G, Shih J, et al. Impact of pathological tumour characteristics in patients with sarcomatoid renal cell carcinoma. *BJU Int* 2012; **109**: 1600–1606.

9. Ferlay J, Shin HR, Bray F, et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010; **127**: 2893–2917.

10. Shuch B, La Rochelle JC, Wu J, et al. Performance status and cytoreductive nephrectomy: redefining management in patients with poor performance. *Cancer* 2008; **113**: 1324–1331.

11. Moch H, Gasser T, Amin MB, et al. Prognostic utility of the recently recommended histologic classification and revised TNM staging system of renal cell carcinoma: a Swiss experience with 588 tumors. *Cancer* 2000; **89**: 604–614.

12. Shuch B, Bratslavsky G, Linehan WM, et al. Sarcomatoid renal cell carcinoma: a comprehensive review of the biology and current treatment strategies. *Oncologist* 2012; **17**: 46–54.

13. Pagliaro LC, Tannir N, Sircar K, et al. Systemic therapy for sarcomatoid renal cell carcinoma. *Expert Rev Anticancer Ther* 2011; **11**: 913–920.

14. Cancer Genome Atlas Research Networks. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* 2013; **499**: 43–49.

15. Lee KS, Kim HJ, Li QL, et al. Runx2 is a common target of transforming growth factor beta1 and bone morphogenetic protein 2, and cooperation between Runx2 and Smad5 induces osteoblast-specific

gene expression in the pluripotent mesenchymal precursor cell line C2C12. *Mol Cell Biol* 2000; **20**: 8783–8792.

16. Gunawan B, Huber W, Holtrup M, et al. Prognostic impacts of cytogenetic findings in clear cell renal cell carcinoma: gain of 5q31-qter predicts a distinct clinical phenotype with favorable prognosis. *Cancer Res* 2001; **61**: 7731–7738.

17. Shuch B, Said J, LaRochelle JC, et al. Histologic evaluation of metastases in renal cell carcinoma with sarcomatoid transformation and its implications for systemic therapy. *Cancer* 2010; **116**: 616–624.

18. Bertoni F, Ferri C, Benati A, et al. Sarcomatoid carcinoma of the kidney. *J Urol* 1987; **137**: 25–28.

19. de Peralta-Venturina M, Moch H, Amin M, et al. Sarcomatoid differentiation in renal cell carcinoma: a study of 101 cases. *Am J Surg Pathol* 2001; **25**: 275–284.

20. Merrill MM, Wood CG, Tannir NM, et al. Clinically nonmetastatic renal cell carcinoma with sarcomatoid dedifferentiation: natural history and outcomes after surgical resection with curative intent. *Urologic oncology* 2015.

21. Zhang BY, Thompson RH, Lohse CM, et al. A novel prognostic model for patients with sarcomatoid renal cell carcinoma. *BJU Int* 2015; **115**: 405–411.

22. DeLong W, Grignon DJ, Eberwein P, et al. Sarcomatoid renal cell carcinoma. An immunohistochemical study of 18 cases. *Arch Pathol Lab Med* 1993; **117**: 636–640.

23. Tickoo SK, Alden D, Olgac S, et al. Immunohistochemical expression of hypoxia inducible factor-1alpha and its downstream molecules in sarcomatoid renal cell carcinoma. *J Urol* 2007; **177**: 1258–1263.

24. Jiang F, Moch H, Richter J, et al. Comparative genomic hybridization reveals frequent chromosome 13q and 4q losses in renal carcinomas with sarcomatoid transformation. *J Pathol* 1998; **185**: 382–388.

25. Dal Cin P, Sciot R, Van Poppel H, et al. Chromosome changes in sarcomatoid renal carcinomas are different from those in renal cell carcinomas. *Cancer Genet Cytogenet* 2002; **134**: 38–40.

26. Jones TD, Eble JN, Wang M, et al. Clonal divergence and genetic heterogeneity in clear cell renal cell carcinomas with sarcomatoid transformation. *Cancer* 2005; **104**: 1195–1203.

27. Dagher J, Dugay F, Verhoest G, et al. Histologic prognostic factors associated with chromosomal imbalances in a contemporary series of 89 clear cell renal cell carcinomas. *Hum Pathol* 2013; **44**: 2106–2115.

28. Brunelli M, Gobbo S, Cossu-Rocca P, et al. Chromosomal gains in the sarcomatoid transformation of chromophobe renal cell carcinoma. *Mod Pathol* 2007; **20**: 303–309.

29. Ro JY, Ayala AG, Sella A, et al. Sarcomatoid renal cell carcinoma: clinicopathologic. A study of 42 cases. *Cancer* 1987; **59**: 516–526.

30. Kanamaru H, Li B, Miwa Y, et al. Immunohistochemical expression of p53 and bcl-2 proteins is not associated with sarcomatoid change in renal cell carcinoma. *Urol Res* 1999; **27**: 169–173.

31. Conant JL, Peng Z, Evans MF, et al. Sarcomatoid renal cell carcinoma is an example of epithelial–mesenchymal transition. *J Clin Pathol* 2011; **64**: 1088–1092.

32. Bostrom AK, Moller C, Nilsson E, et al. Sarcomatoid conversion of clear cell renal cell carcinoma in relation to epithelial-to-mesenchymal transition. *Hum Pathol* 2012; **43**: 708–719.

33. Delahunt B, McKenney JK, Lohse CM, et al. A novel grading system for clear cell renal cell carcinoma incorporating tumor necrosis. *Am J Surg Pathol* 2013; **37**: 311–322.

34. Steinberg PL, Ghavamian R. Robotic-assisted radical cystectomy: current technique and outcomes. *Expert Rev Anticancer Ther* 2012; **12**: 913–917.

35. Golshayan AR, George S, Heng DY, et al. Metastatic sarcomatoid renal cell carcinoma treated with vascular endothelial growth factor-targeted therapy. *J Clin Oncol* 2009; **27**: 235–241.

36. Abel EJ, Culp SH, Matin SF, et al. Percutaneous biopsy of primary tumor in metastatic renal cell carcinoma to predict high risk pathological features: comparison with nephrectomy assessment. *J Urol* 2010; **184**: 1877–1881.

SUPPLEMENTARY MATERIAL ON THE INTERNET

Additional Supporting Information may be found in the online version of this article.

The following supplementary material may be found online.

Detailed methods for Microarray-based gene expression profiling and analysis, DNA copy number assessment and analysis and RNA-seq based gene expression profiling and analysis.

Figure S1. E and S components of sarcomatoid ccRCC had a similar gene expression signature that differed from that of non-sarcomatoid ccRCC (clustered data). (A) Microarray analysis shows a distinctive expression profile of sarcomatoid versus advanced-stage non-sarcomatoid ccRCC, with a heatmap of the 873 significant probes in non-sarcomatoid (E*) and sarcomatoid (E and S) samples contrasting at a FDR of 0.001. (B) Microarray analysis shows a distinctive expression profile of sarcomatoid versus Fuhrman grade 4 non-sarcomatoid RCC, with a heatmap of the 263 significant probes contrasting at a FDR of 0.01. (C) RNA-seq analysis shows a distinctive expression profile of sarcomatoid versus advanced-stage non-sarcomatoid ccRCC, with a heatmap of the 2549 significant probes in non-sarcomatoid (E*) and sarcomatoid (E and S) samples contrasting at a FDR of 0.05. (D) RNA-seq analysis shows distinctive expression profile of sarcomatoid versus Fuhrman grade 4 non-sarcomatoid ccRCC, with a heatmap of the 657 significant probes in Fuhrman grade 4 non-sarcomatoid (E*) and sarcomatoid (E and S) samples contrasting at a FDR of 0.05. Gene expression values were centred before clustering. Samples are ordered by hierarchical clustering (subtype, columns; genes, rows).

Figure S2. Supervised analyses of biphasic E and S components of pretreated ccRCC samples. (A) Sarcomatoid ccRCCs that had been pre-treated with Sutent had no spontaneous clustering of their biphasic E or S components with respect to HIF pathway genes. (B) The E and S components of chemotherapy-pretreated ccRCCs clustered separately with respect to proliferation genes, whereas untreated sarcomatoid ccRCCs did not show spontaneous clustering (C).

Figure S3. RNA-seq analysis confirmed microarray-based results in biphasic E/S paired samples. (A) Unsupervised biclustering by expression profile, as assayed using RNA-seq across RefSeq-annotated transcripts for 4 tumour-matched pairs of E and S components from sarcomatoid ccRCC. No spontaneous association was found by histological type or differential gene network, between E and S. (B) We observed no hier-archical clustering by RNA expression in E and S components across a known panel of good- (ccA) and poor- (ccB) prognosis genes, as pro-filed by RNA-seq, between sarcomatoid ccRCC types. Differential expression was also not seen in (C) ccA or (D) ccB genes between 4 tumour-matched pairs of E and S components.

Figure S4. Druggable targets in ccRCC show differential gene expression in sarcomatoid and non-sarcomatoid ccRCC. Supervised analysis of mTOR (A), HIF (B) and MAPK/ERK (C) pathway genes with respect to sarcomatoid and non-sarcomatoid ccRCC show differential regulation by RNA-seq.

Figure S5. Sarcomatoid histological phenotype shows a partial EMT signature. Boxplots of relative expression of 13 EMT-associated genes in E and S components of sarcomatoid ccRCC, as assayed by (A) microarray and (B) RNAseq analysis. Gene expression levels of independent samples, measured using the 2 platforms show that most differentially expressed genes had the expected directionality for EMT between E and S components (Supplementary Table S6a,b).

Table S1. Clinicopathologic characteristics of advanced stage non-sarcomatoid ccRCC dervied from TCGA

Table S2. Differentially regulated pathways between sarcomatoid and non-sarcomatoid RCC evaluated by microarray and RNA-seq

Table S3. Correlation of sarcomatoid biomarkers with tumour grade and survival in non-sarcomatoid RCC

Table S4. Microarray based expression of E and S components of sarcomatoid RCC with respect to poor prognosis ccB genes

Table S5. Microarray based expression of E and S components of sarcomatoid RCC with respect to good prognosis ccA genes

Table S6a. Microarray based expression of E and S components of sarcomatoid RCC with respect to EMT genes

Table S6b. RNA-seq based expression of E and S components of sarcomatoid RCC with respect to EMT genes