

Article

# Gene Expression Analysis of Aggressive Clinical T1 Stage Clear Cell Renal Cell Carcinoma for Identifying Potential Diagnostic and Prognostic Biomarkers

Jee Soo Park <sup>1</sup>, Phillip M. Pierorazio <sup>2,3</sup> , Ji Hyun Lee <sup>4,5</sup>, Hyo Jung Lee <sup>1</sup>, Young Soun Lim <sup>1</sup>, Won Sik Jang <sup>1</sup> , Jongchan Kim <sup>1</sup>, Seung Hwan Lee <sup>1</sup>, Koon Ho Rha <sup>1</sup>, Nam Hoon Cho <sup>6</sup> and Won Sik Ham <sup>1,\*</sup> 

<sup>1</sup> Department of Urology and Urological Science Institute, Yonsei University College of Medicine, Seoul 03722, Korea; sampark@yuhs.ac (J.S.P.); q8341@yuhs.ac (H.J.L.); silverrain@yuhs.ac (Y.S.L.); sindakjang@yuhs.ac (W.S.J.); lumpakcef@yuhs.ac (J.K.); leeseh@yuhs.ac (S.H.L.); khrha@yuhs.ac (K.H.R.)

<sup>2</sup> The James Buchanan Brady Urological Institute and Department of Urology, Johns Hopkins University School of Medicine, 600 North Wolfe Street, Park 213, Baltimore, MD 21287, USA; philpierorazio@jhmi.edu

<sup>3</sup> Department of Oncology, Sydney Kimmel Comprehensive Cancer Center, Johns Hopkins University of Medicine, Baltimore, MD 21287, USA

<sup>4</sup> Department of Clinical Pharmacology and Therapeutics, College of Medicine, Kyung Hee University, Seoul 02447, Korea; hyunihyuni@gmail.com

<sup>5</sup> Department of Biomedical Science and Technology, Kyung Hee Medical Science Research Institute, Kyung Hee University, Seoul 02447, Korea

<sup>6</sup> Department of Pathology, Yonsei University College of Medicine, Seoul 03722, Korea; cho1988@yuhs.ac

\* Correspondence: uroham@yuhs.ac; Tel.: +82-10-6242-7938

Received: 23 December 2019; Accepted: 15 January 2020; Published: 16 January 2020



**Abstract:** The molecular characteristics of early-stage clear cell renal cell carcinomas (ccRCCs) measuring  $\leq 7$  cm associated with poor clinical outcomes remain poorly understood. Here, we sought to validate genes associated with ccRCC progression and identify candidate genes to predict ccRCC aggressiveness. From among 1069 nephrectomies performed on patients, RNA sequencing was performed for 12 ccRCC patients with aggressive characteristics and matched pairs of 12 ccRCC patients without aggressive characteristics. Using a prospective cohort (ClinicalTrials.gov Identifier: NCT03694912), the expression levels of nine genes (*PBRM1*, *BAP1*, *SETD2*, *KDM5C*, *FOXC2*, *CLIP4*, *AQP1*, *DDX11*, and *BAIAP2L1*) were measured by reverse-transcription polymerase chain reaction from frozen tissues, and their relation to Fuhrman grade was investigated in 70 patients with small ccRCC ( $\leq 4$  cm). In total, 251 genes were differentially expressed and presented fold changes with  $p$ -values  $< 0.05$ ; moreover, 10 genes with the greatest upregulation or downregulation in aggressive ccRCC remained significant even after adjustment. We validated previously identified genes that were associated with ccRCC progression and identified new candidate genes that reflected the aggressiveness of ccRCC. Our study provides new insight into the tumor biology of ccRCC and will help stratify patients with early-stage ccRCC by molecular subtyping.

**Keywords:** clear cell renal cell carcinomas; RNA sequencing; gene expression analysis; prognostic biomarkers

## 1. Introduction

Approximately one-third of patients treated for localized clear cell renal cell carcinoma (ccRCC), the most common subtype of RCC, relapse following surgery; moreover, 15% of these patients exhibit metastatic potential, which can lead to death [1–3]. Tumor growth and changes in radiographic images

are time-dependent predictors of poor oncological outcomes, whereas tumor grade, tumor necrosis, and patient performance are subject to inter-observer variability [1–4]. Alternatively, multigene assays may provide prognostic information beyond what is possible with traditional approaches and are now included in standard treatment guidelines for some tumors [4].

Recently, three companion articles were published based on multicenter, prospective longitudinal cohort studies on ccRCC patients, referred to as TRACERx Renal (tracking renal cell cancer evolution through therapy). They provide insights into ccRCC, including its origin of complex genetic alterations, evolution, and progression to metastasis [5–8]. More than 90% of ccRCC cases show inactivation of the gene encoding the von Hippel-Lindau tumor suppressor on chromosome 3p. Additionally, mutations in other tumor suppressor genes are associated with ccRCC tumorigenesis, including polybromo 1 (*PBRM1*), BRCA1-associated 1 (*BAP1*), and SET domain-containing 2 (*SETD2*). *PBRM1* and *BAP1* mutations are largely mutually exclusive, and *BAP1* mutations are significantly associated with high-grade, high-stage tumors that result in low survival [9]. Although mutations in lysine-specific demethylase 5C (*KDM5C*) occur at a lower frequency, they are also associated with poor oncological outcomes [10].

Forkhead box protein C2 (*FOXC2*) and cytoskeleton-associated protein-glycine (CAP-Gly) rich domain-containing linker protein family member 4 (*CLIP4*) mutations were associated with early-stage ccRCC and synchronous metastasis [3]. Furthermore, *PBRM1*, *BAP1*, and *FOXC2* were shown to be significantly associated with aggressive early-stage ccRCC through target sequencing and immunohistochemistry in our previous study [11].

Presently, molecular profiling does not influence the decision-making processes for ccRCC, and RNA sequencing (RNA-seq) data for aggressive early-stage ccRCC patients have not been reported. Moreover, the association between candidate gene mutations and aggressiveness has also not been reported. Hence, this study aims to identify the biomarkers associated with aggressive early-stage ccRCC.

## 2. Results

### 2.1. Baseline Characteristics

Clinicopathologic characteristics of the patients included in this study ( $n = 24$ ) have been summarized in Table S2, and specific patient information is listed in Table 1. The mean tumor size was  $4.9 \pm 1.6$  cm. Patients with aggressive ccRCC had no significant differences in age, BMI, tumor size, Fuhrman grade, or invasion status when compared with those with non-aggressive ccRCC. Among the 24 patients, six patients had synchronous metastasis to the lung (57.1%) and bone (42.9%). Seven patients reported tumor recurrence after nephrectomy within a mean follow-up period of 25.3 months. Cancer-specific death was reported for eight patients, with a mean survival time of 44.1 months.

**Table 1.** Clinicopathological characterization of each patient.

Aggressiveness	Patient ID	Sex	Age (yr)	Tumor Size (cm)	Fuhrman Grade	Invasion (Perinephric/Sinus Fat/Vascular)	Survival Time <sup>a</sup> (m)	Outcome <sup>b</sup>	Recurrence Site	Synchronous Metastatic Site
N	Pt0-1	M	67	1.6	3	N	41	ned		
N	Pt0-2	M	56	6.3	3	Y	36	ned		
N	Pt0-3	M	76	3.5	2	Y	32	ned		
N	Pt0-4	F	73	5.3	2	Y	92	ned		
N	Pt0-5	M	81	6.6	3	Y	17	ned		
N	Pt0-6	M	72	5.4	3	Y	77	ned		
N	Pt0-7	M	60	6.9	3	Y	59	ned		
N	Pt0-8	M	59	6.4	3	Y	63	ned		
N	Pt0-9	M	61	4.7	3	Y	58	ned		
N	Pt0-10	M	76	5.0	2	N	46	ned		
N	Pt0-11	M	65	4.4	2	N	32	ned		
N	Pt0-12	M	75	2.7	2	N	34	ned		
Y	Pt1-1	M	73	5.5	3	N	19	cd		Lung
Y	Pt1-2	M	73	6.9	4	Y	17	cd		Lung
Y	Pt1-3	M	60	5.1	3	Y	57	pd/cd	Local recurrence, liver	
Y	Pt1-4	M	58	2.8	3	N	36	cd		Bone
Y	Pt1-5	M	74	1.6	2	N	48	pd/cd	Lung, lymph nodes	
Y	Pt1-6	M	70	6.6	3	Y	38	pd/cd	Liver, bone, lymph nodes	
Y	Pt1-7	M	61	6.7	3	Y	14	pd/cd	Lung, liver	Lung, bone
Y	Pt1-8	M	54	4.8	2	Y	56	pd/cd	Local recurrence	
Y	Pt1-9	M	66	4.2	2	N	20	ned		Bone
Y	Pt1-10	M	60	5.2	4	Y	78	pd	Lung, bone	
Y	Pt1-11	M	70	3.5	2	N	58	pd	Bone	
Y	Pt1-12	M	44	5.5	3	Y	31	ned		Lung

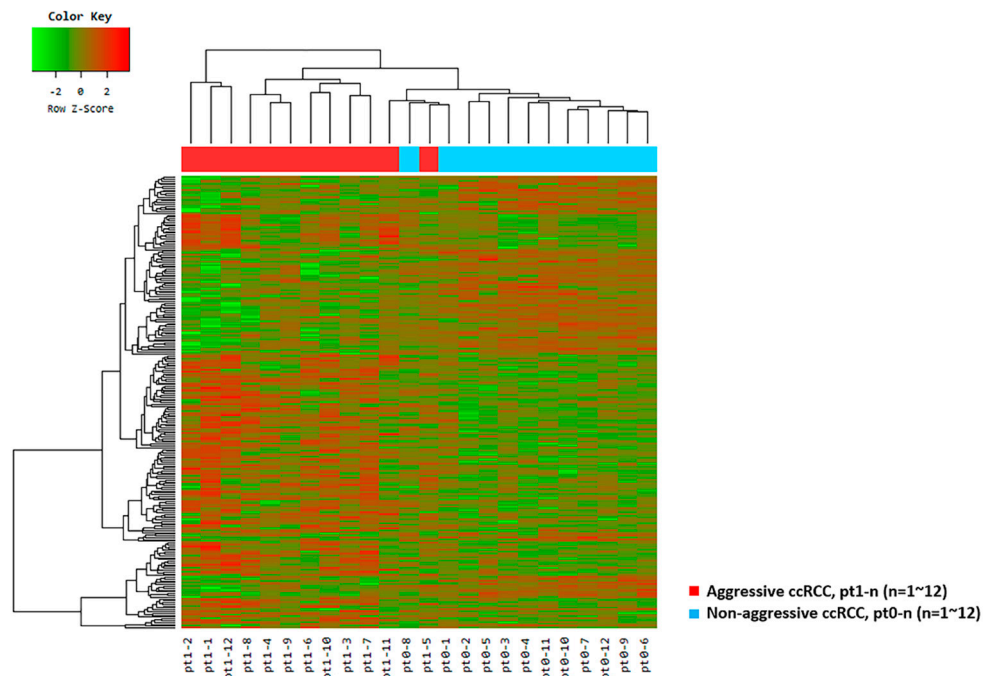
<sup>a</sup> Survival time was defined as the time from nephrectomy until the patient's death or the last time that the patient was known to be alive; <sup>b</sup> cd, cancer death; ned, no evidence of disease; pd, progression of the disease.

## 2.2. Results from the RNA-Seq Analysis

RNA-seq analysis generated  $8266 \times 10^6$  base pairs (bp) from ccRCC with aggressive characteristics and  $7583 \times 10^6$  bp from ccRCC without aggressive characteristics. In total,  $71.68 \times 10^6$  reads in ccRCC with aggressive characteristics and  $66.69 \times 10^6$  reads in ccRCC without aggressive characteristics were mapped. There were no differences in the number of reads between the two groups, among which a total of 12,636 genes were identified.

## 2.3. Differentially Expressed Genes

In total, 251 genes were differentially expressed with fold changes and  $p$ -values  $< 0.05$ ; 153 upregulated genes and 98 downregulated genes had fold changes  $\geq 2$  in ccRCC with aggressive characteristics when compared with matched ccRCC without aggressive characteristics (Figure 1 and Figure S1). The ten genes that were most upregulated or downregulated in patients with aggressive ccRCC and retained significance even after adjustment are summarized in Table 2 and represented as box plots in Figure S2.



**Figure 1.** Unsupervised hierarchical clustering analysis (red, high relative expression; green, low relative expression) in aggressive clear cell renal cell carcinomas (ccRCC) patients ( $n = 12$ , red) versus matched pairs of non-aggressive ccRCC patients ( $n = 12$ , blue) based on 251 differentially expressed genes.

**Table 2.** Significantly upregulated/downregulated genes in clinical T1 stage with or without aggressive characteristics (metastasis, recurrence, or cancer-specific death).

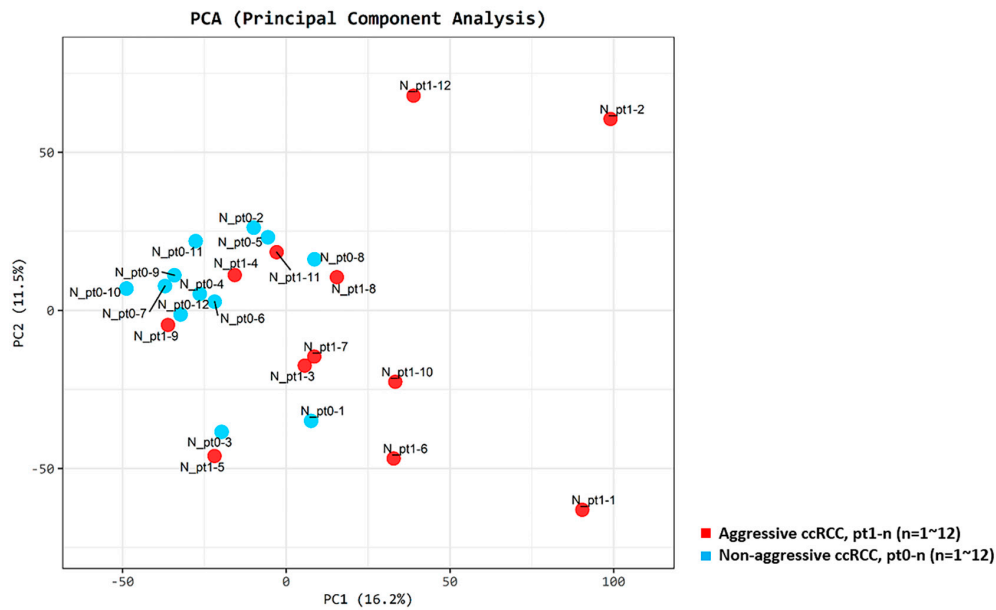
<i>Upregulated</i>						
Gene Symbol	Gene Title	FPKM (Mean ± SD)		Log <sub>2</sub> Fold Change (With Aggressive/without Aggressive Characteristics)	<i>p</i> -Value <sup>a</sup>	Adjusted <i>p</i> <sup>b</sup>
		RCC with Aggressive Characteristics	RCC without Aggressive Characteristics			
MOCOS	Molybdenum cofactor sulfurase	16.76 ± 15.41	1.40 ± 1.29	+3.77	9.29 × 10 <sup>-6</sup>	1.66 × 10 <sup>-2</sup>
RGPD8	RANBP2-like and GRIP domain containing 8	41.49 ± 35.62	6.51 ± 5.38	+2.78	3.67 × 10 <sup>-5</sup>	4.59 × 10 <sup>-2</sup>
BAIAP2L1	BAI1 associated protein 2 like 1	7.05 ± 3.39	2.18 ± 1.28	+1.87	2.18 × 10 <sup>-6</sup>	7.98 × 10 <sup>-3</sup>
DDX11	DEAD/H-box helicase 11	31.95 ± 13.09	14.25 ± 5.18	+1.32	2.55 × 10 <sup>-6</sup>	7.98 × 10 <sup>-3</sup>
<i>Downregulated</i>						
Gene Symbol	Gene Title	FPKM (Mean ± SD)		Log <sub>2</sub> Fold Change (With Aggressive/without Aggressive Characteristics)	<i>p</i> -Value <sup>a</sup>	Adjusted <i>p</i> <sup>b</sup>
		RCC with Aggressive Characteristics	RCC without Aggressive Characteristics			
SLC16A9	Solute carrier family 16 member 9	2.13 ± 2.51	11.90 ± 8.48	-2.41	5.91 × 10 <sup>-6</sup>	1.48 × 10 <sup>-2</sup>
FRAS1	Fraser extracellular matrix complex subunit 1	3.15 ± 3.80	13.66 ± 6.60	-2.12	4.37 × 10 <sup>-8</sup>	5.47 × 10 <sup>-4</sup>
NPR3	Natriuretic peptide receptor 3	3.22 ± 3.37	12.53 ± 7.84	-1.88	3.61 × 10 <sup>-5</sup>	4.59 × 10 <sup>-2</sup>
AQP1	Aquaporin 1 (Colton blood group)	11.02 ± 11.86	36.80 ± 17.38	-1.63	1.73 × 10 <sup>-5</sup>	2.71 × 10 <sup>-2</sup>
TMEM38B	Transmembrane protein 38B	2.74 ± 1.81	8.07 ± 2.70	-1.47	1.21 × 10 <sup>-6</sup>	7.57 × 10 <sup>-3</sup>
PRUNE2	Prune homolog 2	19.10 ± 9.42	49.39 ± 15.65	-1.24	8.48 × 10 <sup>-6</sup>	1.66 × 10 <sup>-2</sup>

FPKM, fragments per kilobase of exon per million fragments mapped; RCC, renal cell carcinoma; SD, standard deviation; <sup>a</sup> *p*-values by nbiomTest function in DESeq2; <sup>b</sup> Adjusted *p*-values by the Benjamini–Hochberg algorithm.

Four genes were upregulated: molybdenum cofactor sulfurase (*MOCOS*), RANBP2-like and GRIP domain containing 8 (*RGPD8*), brain-specific angiogenesis inhibitor 1 associated protein 2 like 1 (*BAIAP2L1*), and DEAD/H-box helicase 11 (*DDX11*). The six downregulated genes were solute carrier family 16 member 9 (*SLC16A9*), fraser extracellular matrix complex subunit 1 (*FRAS1*), natriuretic peptide receptor 3 (*NPR3*), aquaporin 1 (*AQP1*), transmembrane protein 38B (*TMEM38B*), and prune homolog 2 (*PRUNE2*).

For the ten selected genes, a supervised hierarchical clustering analysis of patients with aggressive ccRCC versus matched pairs of patients with non-aggressive ccRCC was performed and revealed clustering of Pt1-4 and Pt1-9 in the non-aggressive ccRCC group (Figure S3).

We observed clear clusters distinguishing patients from the aggressive and non-aggressive ccRCC groups, as determined by the principal component analysis (PCA; Figure 2).



**Figure 2.** Principal component score plots. First and second principal component scores in patients with aggressive ccRCC (n = 12, red) versus matched patient pairs with non-aggressive ccRCC (n = 12, blue) are plotted.

Pt1-1, Pt1-2, and Pt1-12 (lung metastasis group) were grouped separately from other patients, while Pt1-4 and Pt1-9 (bone metastasis group) were not clearly separated from patients with non-aggressive ccRCC. The expression of ten genes was compared between the lung metastasis and bone metastasis groups and other patients. Significant overexpression of *MOCOS* and *BAIAP2L1* was observed in the lung metastasis group ( $p < 0.001$  and  $p = 0.012$ , respectively) (Figure S4). Mutations were found in the six candidate genes (*PBRM1* (11/24, 45.8%), *BAP1* (6/24, 25.0%), *SETD2* (24/24, 100.0%), *KDM5C* (9/24, 37.5%), *FOXO2* (6/24, 25.0%), and *CLIP4* (8/24, 33.3%)) in more than 25% of the patients from our cohort; Figure 3 provides a visual depiction of the frequency of target gene mutations in our cohort.



**Figure 3.** Frequency of candidate gene mutations in the patient cohort, and comparison of those in aggressive and non-aggressive ccRCC.

There were no significant differences in the frequencies of the six candidate aggressiveness-associated mutations in patients with aggressive ccRCC and those without aggressive ccRCC (Table S3); however, for the patients with aggressive ccRCC, *BAP1*, *KDM5C*, and *FOXC2* mutations were enriched by two-fold, and *CLIP4* mutations were enriched by three-fold, compared to that in patients without aggressive ccRCC.

We further verified the expression of 16 genes from the TCGA ccRCC database using UALCAN (Figure S5). Our results showed that 15 of the 16 genes (all except *RGPD8*) were differentially expressed in ccRCC samples, compared to the normal kidney samples. Moreover, the expression levels of 13 out of 16 genes (except *RGPD8*, *NPR3*, and *KDM5C*) differed according to the cancer stage.

#### 2.4. Association between Oncological Outcomes and Expression of the 10 Newly Selected Genes

Cancer-specific survival (CSS) and recurrence-free survival (RFS) and their association with the ten newly selected genes were analyzed. For CSS, patients with cancer-specific death presented higher expression patterns of *RGPD8*, *BAIAP2L1*, and *DDX11* and lower expression patterns of *SLC16A9*, *FRAS1*, *AQP1*, *TMEM38B*, and *PRUNE2*, in both the univariate and multivariate analyses. For RFS, those with cancer recurrence presented higher expression patterns of *DDX11* and lower expressions of *TMEM38B* and *PRUNE2*, in both the univariate and multivariate analyses. Genes that were significantly associated with both CSS and RFS included *DDX11*, *TMEM38B*, and *PRUNE2* (Table 3).

We have dichotomized the expression of *DDX11*, *TMEM38B*, and *PRUNE2* as negative vs. positive and evaluated their association with CSS and RFS. Patients who had *DDX11*-positive and *TMEM38B*-negative tumors had significantly higher chances of cancer-specific death (odds ratio [OR] 49.000; 95% confidence interval [CI] 3.765–637.794) and recurrence (OR 28.000, 95% CI 2.399–326.735; Table 4 and Table S4).

**Table 3.** Comparison of target gene expression according to oncological outcomes (cancer-specific death and recurrence).

Cancer-Specific Death						
FPKM (Mean ± SD)	RCC with Cancer-Specific Death (n = 8)	RCC without Cancer-Specific Death (n = 16)	<i>p</i> -Value <sup>a</sup>	Multivariate OR (95% CI)	<i>p</i> -Value <sup>b</sup>	
MOCOS	16.4 ± 12.5	5.4 ± 12.4	0.054	-	0.093	
RGPD8	47.6 ± 38.4	12.2 ± 17.5	0.036	1.048 (1.006–1.093)	0.026	
BAIAP2L1	7.5 ± 4.0	3.2 ± 2.2	0.019	1.553 (1.088–2.215)	0.015	
DDX11	34.3 ± 13.4	17.5 ± 9.3	0.002	1.130 (1.025–1.245)	0.014	
SLC16A9	1.3 ± 2.3	9.9 ± 8.2	0.001	0.585 (0.362–0.946)	0.029	
FRAS1	2.4 ± 2.5	11.4 ± 7.4	<0.001	0.741 (0.567–0.967)	0.028	
NPR3	3.1 ± 3.4	10.3 ± 8.0	0.006	-	0.058	
AQP1	11.3 ± 12.8	30.2 ± 19.7	0.023	0.926 (0.859–0.997)	0.042	
TMEM38B	1.9 ± 1.2	7.2 ± 2.9	<0.001	0.277 (0.091–0.841)	0.024	
PRUNE2	20.1 ± 8.5	41.3 ± 20.5	0.002	0.921 (0.854–0.993)	0.032	
Recurrence						
FPKM (Mean ± SD)	RCC with Recurrence (n = 7)	RCC without Recurrence (n = 17)	<i>p</i> -Value <sup>a</sup>	Multivariate OR (95% CI)	<i>p</i> -Value <sup>b</sup>	
MOCOS	9.5 ± 6.6	8.9 ± 15.4	0.921	-	0.917	
RGPD8	65.5 ± 25.9	6.9 ± 6.0	0.001	-	0.996	
BAIAP2L1	6.8 ± 3.0	3.7 ± 3.4	0.048	-	0.066	
DDX11	36.0 ± 10.2	17.8 ± 10.6	0.001	1.140 (1.029–1.264)	0.012	
SLC16A9	1.5 ± 2.1	9.3 ± 8.3	0.002	-	0.056	
FRAS1	3.1 ± 2.8	10.6 ± 7.8	0.002	-	0.051	
NPR3	2.9 ± 2.5	9.9 ± 8.1	0.004	-	0.073	
AQP1	12.7 ± 13.0	28.5 ± 20.3	0.071	-	0.090	
TMEM38B	2.5 ± 1.5	6.6 ± 3.4	0.001	0.579 (0.355–0.944)	0.028	
PRUNE2	14.5 ± 7.9	42.4 ± 17.6	<0.001	0.788 (0.623–0.996)	0.047	

Data are shown as mean ± SD; <sup>a</sup> *p*-value calculated using the *t*-test; <sup>b</sup> *p*-value calculated using logistic regression for multivariate analysis.

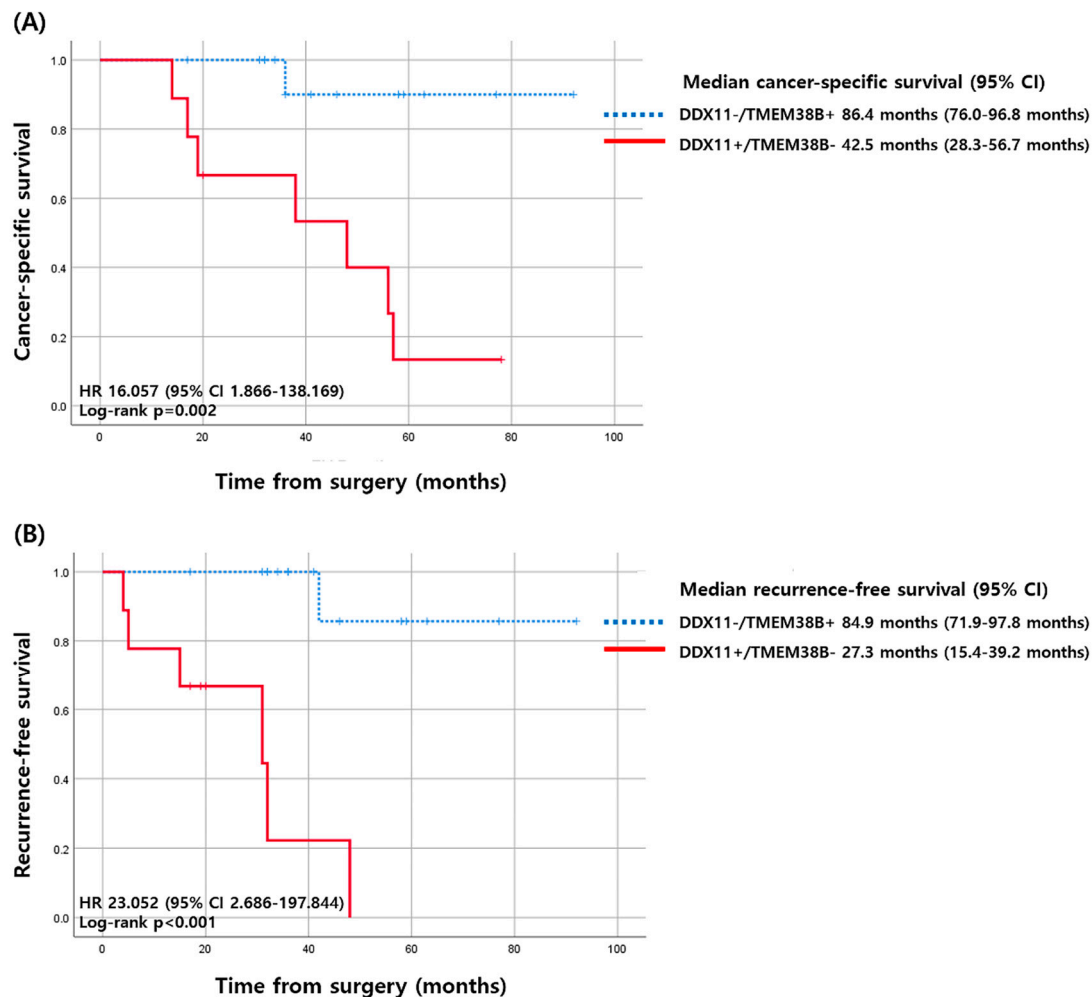


**Table 4.** Expression of DDX11, TMEM38B, and PRUNE2 by oncological outcomes (cancer-specific death and recurrence).

Groups Dichotomized by the Expression of DDX1, TMEM38B and PRUNE2	Cancer-Specific Death			Multivariate OR (95% CI)	<i>p</i> -Value <sup>b</sup>
	RCC with Cancer-Specific Death (n = 8)	RCC without Cancer-Specific Death (n = 16)	<i>p</i> -Value <sup>a</sup>		
DDX11+	7/8 (87.5%)	5/16 (31.3%)	0.027	-	0.065
TMEM38B-	8/8 (100.0%)	4/16 (25.0%)	0.001	-	0.998
PRUNE2-	7/8 (87.5%)	5/16 (31.3%)	0.027	-	0.065
DDX11+ and TMEM38B-	7/8 (87.5%)	2/16 (12.5%)	0.001	49.000 (3.765–637.794)	0.003
DDX11+ and PRUNE2-	6/8 (75.0%)	2/16 (12.5%)	0.005	21.000 (2.372–185.930)	0.006
Recurrence					
	RCC with Recurrence (n = 7)	RCC without Recurrence (n = 17)	<i>p</i> -Value <sup>a</sup>	Multivariate OR (95% CI)	<i>p</i> -Value <sup>b</sup>
DDX11+	7/7 (100.0%)	5/17 (29.4%)	0.005	-	0.999
TMEM38B-	6/7 (85.7%)	6/11 (54.5%)	0.069	11.000 (1.061–114.086)	0.045
PRUNE2-	7/7 (100.0%)	5/17 (29.4%)	0.005	-	0.999
DDX11+ and TMEM38B-	6/7 (85.7%)	3/17 (17.6%)	0.004	28.000 (2.399–326.735)	0.008
DDX11+ and PRUNE2-	7/7 (100.0%)	1/17 (5.9%)	<0.001	-	0.998

Data are shown as the number of patients (%); <sup>a</sup> *p*-value calculated using the chi-square test or Fisher's exact test; <sup>b</sup> *p*-value calculated using logistic regression for multivariate analysis; DDX11+, FPKM (fragments per kilobase of exon per million fragments mapped) > 20.0; TMEM38B-, FPKM < 5.0; PRUNE2-, FPKM < 32.0.

Patients with *DDX11*-positive and *TMEM38B*-negative tumors had a median CSS of 42.5 months (95% CI 28.3–56.7 months), which was substantially shorter than that of patients who had *DDX11*-negative or *TMEM38B*-positive tumors, whose median CSS was 86.4 months (95% CI 76.0–96.8 months). The differences in CSS corresponded to a hazard ratio (HR) of 16.057 (95% CI 1.866–138.169,  $p = 0.011$ ; Figure 4). For RFS, patients with *DDX11*-positive and *TMEM38B*-negative tumors had a median RFS of 27.3 months (95% CI 15.4–39.2 months), which was substantially shorter than that of patients with *DDX11*-negative or *TMEM38B*-positive tumors, who presented a median RFS of 84.9 months (95% CI 71.9–97.8 months). The differences in RFS corresponded to an HR of 23.052 (95% CI 2.686–197.844,  $p = 0.004$ ; Figure 4).



**Figure 4.** Kaplan–Meier curves of cancer-specific survival and recurrence-free survival according to the expression of *DDX11* and *TMEM38B*. (A) Cancer-specific survival (B) recurrence-free survival. CI, confidence interval; HR, hazard ratio.

### 2.5. Survival Analysis of 16 Genes

Using the Kaplan–Meier plotter database, the prognostic value of the 16 genes was evaluated in 530 patients with ccRCC. We found that 13 of the 16 genes (*MOCOS*, *BAIAP2L1*, *DDX11*, *CLIP4*, *SLC16A9*, *FRAS1*, *NPR3*, *AQP1*, *PRUNE2*, *TMEM38B*, *PBRM1*, *BAP1*, and *SETD2*) were related to overall survival (OS) in patients with ccRCC (Figure S6). High expression of *MOCOS*, *BAIAP2L1*, *DDX11*, and *CLIP4*, and low expression of *SLC16A9*, *FRAS1*, *NPR3*, *AQP1*, *PRUNE2*, *TMEM38B*, *PBRM1*, *BAP1*, and *SETD2*, was significantly associated with poor OS ( $p < 0.05$  for both). Expression of *RGPD8*, *KDM5C*, and *FOXC2* was not significantly associated with poor OS in patients with ccRCC.

## 2.6. GO Analysis and KEGG Analysis of DEGs

DEGs were functionally classified into biological process (BP), cellular component (CC), and molecular function (MF) categories (Figure S7). In the BP category, the top three most enriched terms were “single-organism process,” “cellular process,” and “single-organism cellular process” (Figure S7A). In the CC category, the top three most enriched terms were “cell part,” “cell,” and “organelle” (Figure S7B). In the MF category, the top three most enriched terms were “binding,” “protein binding,” and “ion binding” (Figure S7C). Moreover, the top three most enriched terms in the KEGG analysis were “metabolic pathways,” “cell cycle,” and “complement and coagulation cascades” (Figure S7D).

## 2.7. Validation of Target Genes Using Frozen Tissue PCR

Of the ten genes newly identified through RNA-seq analysis, *AQP1*, *DDX11*, and *BAIAP2L1* were previously reported to be potent indicators of outcome in patients with ccRCC. We analyzed the expression of these three genes along with that of the six genes previously identified by the qRT-PCR analysis of frozen tissue samples. Our results showed that *DDX11* expression and tumor size were significantly greater in patients with high Fuhrman grade (3 and 4), both in the univariate and multivariate analyses (Table S5).

## 3. Discussion

This is the first study to identify 251 DEGs in early-stage aggressive ccRCCs measuring  $\leq 7$  cm using RNA-seq. Among the potential target candidate genes, *BAP1*, *KDM5C*, *FOXC2*, and *CLIP4* mutations were enriched in patients with aggressive ccRCC. Among the DEGs identified in this study, *MOCOS*, *RGPD8*, *BAIAP2L1*, and *DDX11* were significantly upregulated; in contrast, *SLC16A9*, *FRAS1*, *NRP3*, *AQP1*, *TMEM38B*, and *PRUNE2* were significantly downregulated.

A panel of six genes (*PBRM1*, *BAP1*, *SETD2*, *KDM5C*, *FOXC2*, and *CLIP4*) was chosen for mutational frequency analysis based on the clinical relevance of significantly mutated genes across all ccRCC cohorts [12,13]. However, these six genes were not differentially expressed, perhaps because our patient cohort included unique cases for which cancers exhibited aggressive characteristics, such as synchronous metastasis, recurrence, or cancer-specific death, even though the clinical stage was low. Moreover, small sample sizes would have affected this result. As in other large NGS studies, *PBRM1* mutation was identified in 45.8% patients. However, *BAP1* mutation was observed in 25% of patients, which is slightly higher than the value obtained in other studies, perhaps because our cohort comprised patients with more aggressive ccRCC; we intentionally collected aggressive ccRCC to pair with non-aggressive ccRCC. *BAP1*, a critical gatekeeper for disease progression, has a mutually exclusive interaction with *PBRM1*, whereas the *BAP1* mutation is associated with worse prognosis and a higher Fuhrman grade than the *PBRM1* mutation [7,14]. This finding was further supported by the MSKCC cohort study where *BAP1* mutations were associated with poor prognostic factors (higher T stage, higher nuclear grade, large size, more necrosis), and the presence of metastatic disease at presentation) [15].

*SETD2* mutation was observed in all cases of ccRCC in our study, which diverges from other NGS studies in which the frequency of *SETD2* mutation was substantially lower (approximately 11%) [12,13]. However, Liu et al. reported a multicenter study in China that had evaluated radical nephrectomy cases, wherein the *SETD2* protein deficiency rate was 34.1%, although there may have been a discrepancy between protein deficiency and gene mutation rates [16]. *KDM5C* mutations were found in less than 40% of the study population, compared to 3.8–6.8% reported in other NGS studies [12,13]. We believe that this difference in *KDM5C* mutation frequency could result from selection bias because of the unique group of patients who were defined as aggressive. However, owing to the intratumoral heterogeneity (ITH), there could also have been a sampling bias. Moreover, gene mutation status was based on

multiple types of sequencing that only reflect the gene status of very circumscribed cancer lesions, and thus could underestimate mutational rates [16].

The frequency of *PBRM1* mutation was not different in ccRCC with and without aggressive characteristics. *BAP1*, *KDM5C*, *FOXC2*, and *CLIP4* mutations were enriched in aggressive ccRCC, which was not statistically significant because of small sample sizes. This finding is similar to the results of other studies that have reported an association between *BAP1* and *KDM5C* mutations and aggressiveness of tumors [9,10]. *FOXC2*, a gene known to be associated with tumor aggressiveness and synchronous metastasis in ccRCC, based on our previous studies [2,11], was found to be associated with aggressive characteristics of tumors in early-stage ccRCC, in the present study. The frequency of *PBRM1* mutation was similar between patients with aggressive and non-aggressive ccRCC, perhaps because it is an early, essential event in tumorigenesis that does not impact clinical outcome, but instead plays a principal role in tumor initiation. *BAP1* mutation was more frequently observed in aggressive ccRCC; *BAP1* is believed to be associated with disease progression and a worse prognosis [17].

With regard to metastasis, the lung and bone were the two most common sites, with up to 60% and 40% of patients having lung and bone metastasis, respectively [18]. A recent study suggests that patients with bone metastases have an unfavorable prognosis; however, another study found no prognostic differences between metastatic sites [19], although this remains controversial. According to this study, synchronous bone metastatic patients were grouped with non-aggressive ccRCC patients, while synchronous lung metastatic patients were grouped with the aggressive ccRCC group. These results show that the early bone metastasis group genetically clusters with the less aggressive ccRCC group and suggest that the prognosis of early bone metastasis ccRCC patients might not be bad, although future studies are needed to verify this finding. The exact role of *MOCOS* and *BAIAP2L1* in lung metastasis needs to be studied further, as these genes may help us understand differences in metastatic potential in ccRCC and the prognostic value of different metastatic sites.

We included *AQP1*, *DDX11*, and *BAIAP2L1* from the newly identified gene set because these were previously associated with ccRCC or have a known relation to other tumors [20–22]. A prospective study examining frozen tissue qRT-PCR of *AQP1*, *DDX11*, *BAIAP2L1*, and six previously identified genes (*PBRM1*, *BAP1*, *SETD2*, *KDM5C*, *FOXC2*, and *CLIP4*) showed that expression of *DDX11* was a significant factor for predicting tumor aggressiveness based on Fuhrman grade. Bhattacharya et al. demonstrated that inhibiting *DDX11* in melanoma cells decreased proliferation and rapidly increased apoptosis [21]. In the future, we will perform an in vitro and in vivo study using *DDX11* since ccRCC has similar tumor characteristics as melanoma.

We only included *DDX11* and *TMEM38B* in our final analysis of CSS and RFS since although using three genes increased the statistical significance (Table S4), the combination of *DDX11* and *PRUNE2* was not associated with recurrence in the multivariate analysis (Table 4). Therefore, although combining three genes would of course increase the statistical significance compared to two genes in clinical settings, it is also important to reduce the number of genes used. Therefore, we tried to select the most important gene that is both related to CSF and RFS.

We propose that the factors involved in tumor aggressiveness and metastasis or recurrence are different. Previous studies have focused on using Fuhrman grade as the golden standard and have compared gene expression or mutation based on Fuhrman grade. However, Fuhrman grade only reflects cell differentiation and does not reflect metastatic and recurrence potential. Recent studies have reported clonal and subclonal mutations in ccRCC, whereas clonal mutations are present in all tumor regions, and subclonal mutations are present in some, but not all, tumors [23]. Clonal cancer evolution is also known to be involved in ccRCC through Darwinian selection [23]. TRACERx Renal classified tumors into seven distinct evolutionary subtypes [6,7], and primary tumors with low ITH, in addition to a low fraction of the tumor genome affected by somatic copy-number alterations (SCNAs), had an overall low metastatic potential. In contrast, primary tumors with low ITH but elevated SCNAs were associated with rapid progression at multiple sites [7]. Thus, our knowledge regarding potential ccRCC biomarkers for guiding intervention and surveillance has recently been dramatically expanded. Recent

article by Wach et al. reported the RNA sequencing of collecting duct carcinoma, a rare renal cell carcinoma subtype with a very poor prognosis and solute carrier (SLC) gene family (*SLC3A1*, *SLC9A3*, *SLC26A7*, and *SLC47A1*) was included in the most downregulated genes [24]. According to this study trend, our newly identified biomarkers may further aid in unraveling the evolutionary trajectories of ccRCC tumor growth.

This study has a few limitations. First, this study intentionally collected aggressive cases of ccRCC; thus, selection bias may have resulted in statistical differences from other NGS studies. However, this collection method has benefits, including the examination of unique cases of interest to clinicians. Second, because of the retrospective nature of this study, the study sample size was small; this limited our ability to further investigate the ITH of primary tumors, although the ITH in ccRCC tumors, including even small renal masses, may be substantial [1]. In future studies, we plan to overcome this ITH limitation by analyzing the circulating tumor DNA in plasma. Third, RNA-seq was based only on surgical FFPE samples and we could not provide validation results from a different independent dataset. To overcome these limitations and validate our findings, a prospective clinical trial is now underway (ClinicalTrials.gov Identifier: NCT03694912) that will include an examination of the mutation profiles of the selected genes (*PBRM1*, *BAP1*, *SETD2*, *KDM5C*, *FOXC2*, *CLIP4*, *MOCOS*, *BAIAP2L1*, *DDX11*, and *AQP1*) for aggressive ccRCC. However, our preliminary data obtained using frozen tissue PCR show promising results for *DDX11*, and in vivo and in vitro studies on each gene are ongoing. The demonstration of immunohistochemistry in selected genes would have provided the association between RNA and protein. However, we have focused on the use of RNA expression profiles since it is easier and faster to use in the clinical settings compared to immunohistochemistry. Nevertheless, the protein expression profiles would provide more specific information regarding the role of selected genes. Therefore, we are planning to include protein expression analysis in the future study.

## 4. Materials and Methods

### 4.1. Patients and Tissues

For this retrospective study, data were gathered from 1132 patients with ccRCC ( $\leq 7$  cm) who underwent radical and partial nephrectomy between January 2008 and December 2014. An aggressive tumor was defined as a tumor exhibiting synchronous metastasis, recurrence, or cancer-specific death; synchronous metastasis was defined as metastasis detected at or within three months of the primary RCC diagnosis [25]. Twelve patients with aggressive ccRCC and 12 matched patients in terms of gender, BMI, percentage of radical surgeries performed, clinical tumor sizes, and Fuhrman grades with non-aggressive ccRCC were examined in order to adjust the confounding factors other than aggressiveness of ccRCC. This study was approved by the Institutional Review Board of the Yonsei University Health System (project no: 4-2013-0742).

We included patients with ccRCC ( $\leq 7$  cm) treated with nephrectomy alone and for whom formalin-fixed paraffin-embedded (FFPE) tumor tissue samples were available; we also included patients who did not exhibit the typical characteristics of ccRCC. Specimens that were inappropriate for molecular analysis were excluded, and clinicopathologic features were recorded for each patient (Supplementary Method 1). Diameters of the primary tumors were measured using imaging modalities.

For validation analyses, we used a prospective cohort (ClinicalTrials.gov Identifier: NCT03694912) sampled between November 2018 and July 2019 and included cases with small ccRCCs ( $\leq 4$  cm). Data on clinicopathologic features were included, as previously described in our retrospective study group. The levels of six previously identified genes (*PBRM1*, *BAP1*, *SETD2*, *KDM5C*, *FOXC2*, and *CLIP4*), plus three genes (*AQP1*, *DDX11*, and *BAIAP2L1*) among the ten newly identified genes, were considered potent indicators of outcome in ccRCC patients according to a literature search. Gene expression was measured by reverse-transcription polymerase chain reaction (qRT-PCR).

#### 4.2. Tissue Preparation

Formalin-fixed, paraffin-embedded (FFPE) sections from patients with ccRCC were obtained from the archives of the Department of Pathology at Yonsei University College of Medicine (Seoul, Korea). All cases were reviewed and classified by a urologic pathologist (N.H.C.). Non-tumor elements were identified based on hematoxylin and eosin-stained slides from each sample by a urologic pathologist (N.H.C.). The samples were subsequently cut into 20- $\mu$ m sections before being transferred to an extraction tube.

#### 4.3. RNA Extraction and Sequencing

Briefly, for FFPE tissues, RNA was extracted from 20  $\mu$ m sections using an FFPE RNeasy kit (Qiagen, Gaithersburg, MD, USA), and 100 ng of RNA was used to construct a sequencing library using a TruSeq RNA Access library prep kit (Illumina, San Diego, CA, USA) according to the manufacturer's protocols (Supplementary Method 2).

The sequencing libraries were quantified by the Kapa Biosystems Library Quantification Kit for Illumina Sequencing platforms according to the qPCR Quantification Protocol Guide (KapaBiosystems, Wilmington, MA, USA Cat #KK4854) and qualified by the TapeStation D1000 ScreenTape (Agilent Technologies, Palo Alto, CA, USA Cat # 5067-5582). Indexed libraries were then submitted to Illumina NovaSeq 6000 (Illumina), and paired-end (2  $\times$  101 bp) sequencing was performed by MacroGen Incorporated.

#### 4.4. Analysis of RNA Sequencing—Differentially Expressed Gene (DEG) Selection

Paired-end sequencing reads from cDNA libraries (101 bp) were generated and the trimmed reads were aligned to the reference human genome ([University of California Santa Cruz] UCSC hg19). Transcriptome assembly of known transcripts, novel transcripts, and alternative splicing transcripts was processed.

Based on the results of this processing, the abundance of transcript and gene expression was calculated as read count (the number of reads mapping to a gene) or FPKM value (fragments per kilobase of exon per million fragments mapped) per sample. To identify DEGs from aggressive and non-aggressive ccRCC groups, genes with more than one sample of zero read count were pre-filtered. We then measured the expression levels of 27,685 RefSeq genes and 12,636 genes (Supplementary Method 3).

For DEGs, unsupervised hierarchical clustering analysis was performed with the complete linkage method and Euclidean distance as a measure of similarity. Principal component analysis (PCA) was performed to reduce the dimensionality of the data set by transforming it into a new set of variables to summarize the data features process (Supplementary Method 3). All data analysis and visualization of DEGs were conducted using R.3.5.1 ([www.r-project.org](http://www.r-project.org)).

#### 4.5. Variant Calling

For variant calling of RNA-seq data, trimmed reads were aligned to the human genome (UCSC hg19). The alignment of the reads used in this analysis was created through Split 'N' Trim, mapping quality reassignment, indel realignment, and base recalibration process (Supplementary Method 4). For our target gene list (*PBRM1*, *BAP1*, *SETD2*, *KDM5C*, *FOXC2*, and *CLIP4*), the frequency of alterations in all samples was represented on the heatmap. Individual genes are represented as rows, and individual patients are represented as columns.

#### 4.6. qRT-PCR

Total RNA was extracted from frozen tissue samples using TRIzol<sup>®</sup> reagent (Ambion, Life Technologies, Carlsbad, CA, USA), and 1  $\mu$ g of total RNA was reverse-transcribed into first-strand

cDNA using an iNtRon Maxime RT PreMix (Intronbio, Sungnam, Korea Cat No. 25081) according to the manufacturer's protocol (Supplementary Method 5).

#### 4.7. GO Analysis and KEGG Analysis

Gene Ontology (GO) analysis per biological process (BP), cellular component (CC), and molecular function (MF) for DEGs were performed based on Gene Ontology (<http://geneontology.org/>). KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis for DEGs was also performed based on the KEGG pathway (<https://www.genome.jp/kegg/>) database (Supplementary Method 6).

#### 4.8. Validation of Gene Expression, Survival Analyses, and Statistical Analyses

UALCAN (<http://ualcan.path.uab.edu>) was used to validate the expression of ten newly identified genes and six target genes according to the stage of the cancer. The Kaplan–Meier plotter database (<http://kmplot.com>) was used to analyze the associations between 16 genes (10 newly identified and six target genes) and overall patient survival. Additional details for statistical analyses are provided in Supplementary Method 7.

## 5. Conclusions

In summary, we identified DEGs in aggressive early-stage ccRCC, defined as those with synchronous metastasis, recurrence, or cancer-specific death, and found that *MOCOS*, *RGPD8*, *BAIAP2L1*, and *DDX11* showed significantly higher expression levels, whereas *SLC16A9*, *FRAS1*, *AQP1*, *TMEM38B*, and *PRUNE2* showed significantly lower expression levels compared to those in non-aggressive ccRCC. Moreover, we verified previously known genes that are associated with tumor aggressiveness by analyzing mutational frequency. Further research on these molecular biomarkers may help stratify patients with early-stage ccRCC by molecular subtyping and aid in making clinical decisions accordingly.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2072-6694/12/1/222/s1>. Table S1: PCR primer sequences of target genes. Table S2: Patient baseline characteristics. Table S3: Mutational frequencies of the candidate aggressiveness-associated genes. Table S4. Expression of combination of *DDX11*, *TMEM38B*, and *PRUNE2* by oncological outcomes (cancer-specific death and recurrence) (in addition to Table 4). Table S5. Comparison of expression levels of target genes according to Fuhrman grade in clear cell renal cell carcinoma ( $\leq 4$  cm). Figure S1. Volcano plots showing a comparison of gene expression levels in aggressive ccRCC and non-aggressive ccRCC. Figure S2. Box plots showing a comparison of expression levels of ten selected genes in aggressive and non-aggressive ccRCC. Figure S3. Supervised hierarchical clustering analysis (red, high relative expression; green, low relative expression) of aggressive ccRCC patients ( $n = 12$ , red) versus matched non-aggressive ccRCC patients ( $n = 12$ , blue) based on the ten selected genes. Figure S4. Expression levels of *BAIAP2L1* and *MOCOS* according to the synchronous metastatic sites (lung and bone) shown as box plots. Figure S5. Validation of the expression of 16 genes in the TCGA database according to stage. Figure S6. Association of the 16 genes with overall survival (OS) of clear cell renal cell carcinoma (ccRCC) patients, as analyzed by Kaplan–Meier survival plots. Figure S7. GO annotation and pathway enrichment analysis. Supplementary Method 1: Patients and tissues. Supplementary Method 2: RNA extraction and sequencing. Supplementary Method 3: Analysis of RNA sequencing. Supplementary Method 4: Variant calling. Supplementary Method 5: qRT-PCR. Supplementary Method 6: GO and KEGG analysis. Supplementary Method 7: Validation of gene expression, survival analyses, and statistical analyses. References [26–36] are cited in the Supplementary Materials.

**Author Contributions:** Research design, data acquisition and analysis, and manuscript—writing and editing, J.S.P.; data analysis, manuscript editing, P.M.P.; data analysis, manuscript editing, J.H.L.; methodology, H.J.L.; methodology, Y.S.L.; data acquisition, W.S.J.; data acquisition, J.K.; data acquisition, S.H.L.; data acquisition, K.H.R.; data—acquisition and analysis, N.H.C.; research design, data—acquisition and analysis, manuscript editing, W.S.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by a grant from the Korea Health Technology R&D Project of the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI17C1095).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Manley, B.J.; Reznik, E.; Ghanaat, M.; Kashan, M.; Becerra, M.F.; Casuscelli, J.; Tennenbaum, D.; Redzematovic, A.; Carlo, M.I.; Sato, Y.; et al. Characterizing recurrent and lethal small renal masses in clear cell renal cell carcinoma using recurrent somatic mutations. *Urol. Oncol.* **2019**, *37*, 12. [[CrossRef](#)] [[PubMed](#)]
2. Rini, B.I.; Campbell, S.C.; Escudier, B. Renal cell carcinoma. *Lancet* **2009**, *373*, 1119. [[CrossRef](#)]
3. Ahn, J.; Han, K.S.; Heo, J.H.; Bang, D.; Kang, Y.H.; Jin, H.A.; Hong, S.J.; Lee, J.H.; Ham, W.S. FOXC2 and CLIP4: A potential biomarker for synchronous metastasis of  $\leq 7$ -cm clear cell renal cell carcinomas. *Oncotarget* **2016**, *7*, 51423. [[CrossRef](#)] [[PubMed](#)]
4. Rini, B.; Goddard, A.; Knezevic, D.; Maddala, T.; Zhou, M.; Aydin, H.; Campbell, S.; Elson, P.; Koscielny, S.; Lopatin, M.; et al. A 16-gene assay to predict recurrence after surgery in localised renal cell carcinoma: Development and validation studies. *Lancet Oncol.* **2015**, *16*, 676. [[CrossRef](#)]
5. Ricketts, C.J.; Linehan, W.M. The origin, evolution and route to metastasis of clear cell RCC. *Nat. Rev. Nephrol.* **2018**, *14*, 538. [[CrossRef](#)]
6. Mitchell, T.J.; Turajlic, S.; Rowan, A.; Nicol, D.; Farmery, J.H.; O'Brien, T.; Martincorena, I.; Tarpey, P.; Angelopoulos, N.; Yates, L.R.; et al. Timing the landmark events in the evolution of clear cell renal cell cancer: TRACERx Renal. *Cell* **2018**, *173*, 611. [[CrossRef](#)]
7. Turajlic, S.; Xu, H.; Litchfield, K.; Rowan, A.; Chambers, T.; Lopez, J.I.; Rowan, A.; Chambers, T.; Lopez, J.I.; Nicol, D.; et al. Tracking cancer evolution reveals constrained routes to metastases: TRACERx Renal. *Cell* **2018**, *173*, 581. [[CrossRef](#)]
8. Turajlic, S.; Xu, H.; Litchfield, K.; Rowan, A.; Horswell, S.; Chambers, T.; O'Brien, T.; Lopez, J.I.; Watkins, T.B.K.; Nicol, D.; et al. Deterministic evolutionary trajectories influence primary tumor growth: TRACERx Renal. *Cell* **2018**, *173*, 595. [[CrossRef](#)]
9. Peña-Llopis, S.; Vega-Rubin-de-Celis, S.; Liao, A.; Leng, N.; Pavía-Jiménez, A.; Wang, S.; Yamasaki, T.; Zhrebker, L.; Sivanand, S.; Spence, P.; et al. BAP1 loss defines a new class of renal cell carcinoma. *Nat. Genet.* **2012**, *44*, 751. [[CrossRef](#)]
10. Dalglish, G.L.; Furge, K.; Greenman, C.; Chen, L.; Bignell, G.; Butler, A.; Davies, H.; Edkins, S.; Hardy, C.; Latimer, C.; et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **2010**, *463*, 360. [[CrossRef](#)]
11. Park, J.S.; Lee, H.J.; Cho, N.H.; Kim, J.; Jang, W.S.; Heo, J.E.; Ham, W.S. Risk prediction tool for aggressive tumors in clinical T1 stage clear cell renal cell carcinoma using molecular biomarkers. *Comput. Struct. Biotechnol. J.* **2019**, *17*, 371. [[CrossRef](#)]
12. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* **2013**, *499*, 43. [[CrossRef](#)] [[PubMed](#)]
13. Sato, Y.; Yoshizato, T.; Shiraishi, Y.; Wu, C.P.; Yang, K.; Wang, S.P.; Wu, S. Integrated molecular analysis of clear-cell renal cell carcinoma. *Nat. Genet.* **2013**, *45*, 860. [[CrossRef](#)]
14. Kapur, P.; Pena-Llopis, S.; Christie, A.; Zhrebker, L.; Pavía-Jiménez, A.; Rathmell, W.K.; Xie, X.J.; Brugarolas, J. Effects on survival of BAP1 and PBRM1 mutations in sporadic clear-cell renal-cell carcinoma: A retrospective analysis with independent validation. *Lancet Oncol.* **2013**, *14*, 159. [[CrossRef](#)]
15. Hakimi, A.A.; Ostrovnaya, I.; Reva, B.; Schultz, N.; Chen, Y.B.; Gonen, M.; Liu, H.; Takeda, S.; Voss, M.H.; Tickoo, S.K.; et al. ccRCC Cancer Genome Atlas (KIRC TCGA) Research Network investigators. Adverse outcomes in clear cell renal cell carcinoma with mutations of 3p21 epigenetic regulators BAP1 and SETD2: A report by MSKCC and the KIRC TCGA research network. *Clin. Cancer Res.* **2013**, *19*, 3259. [[CrossRef](#)] [[PubMed](#)]
16. Liu, L.; Guo, R.; Zhang, X.; Liang, Y.; Kong, F.; Wang, J.; Xu, Z. Loss of SETD2, but not H3K36me3, correlates with aggressive clinicopathological features of clear cell renal cell carcinoma patients. *Biosci. Trends* **2017**, *11*, 214. [[CrossRef](#)]
17. Joseph, R.W.; Kapur, P.; Serie, D.J.; Parasramka, M.; Ho, T.H.; Cheville, J.C.; Frenkel, E.; Parker, A.S.; Brugarolas, J. Clear cell renal cell carcinoma subtypes identified by BAP1 and PBRM1 expression. *J. Urol.* **2016**, *195*, 180. [[CrossRef](#)]
18. Motzer, R.J.; Bander, N.H.; Nanus, D.M. Renal-cell carcinoma. *N. Engl. J. Med.* **1996**, *335*, 865. [[CrossRef](#)]



19. Mani, S.; Todd, M.B.; Katz, K.; Poo, W.J. Prognostic factors for survival in patients with metastatic renal cancer treated with biological response modifiers. *J. Urol.* **1995**, *154*, 35. [[CrossRef](#)]
20. Huang, Y.; Murakami, T.; Sano, F.; Kondo, K.; Nakaigawa, N.; Kishida, T.; Kubota, Y.; Nagashima, Y.; Yao, M. Expression of aquaporin 1 in primary renal tumors: A prognostic indicator for clear-cell renal cell carcinoma. *Eur. Urol.* **2009**, *56*, 690. [[CrossRef](#)]
21. Bhattacharya, C.; Wang, X.; Becker, D. The DEAD/DEAH box helicase, DDX11, is essential for the survival of advanced melanomas. *Mol. Cancer* **2012**, *11*, 82. [[CrossRef](#)] [[PubMed](#)]
22. Chao, A.; Tsai, C.L.; Jung, S.M.; Chuang, W.C.; Kao, C.; Hsu, A.; Chen, S.H.; Lin, C.Y.; Lee, Y.C.; Lee, Y.S.; et al. BAI1-associated protein 2-like 1 (BAIAP2L1) is a potential biomarker in ovarian cancer. *PLoS ONE* **2015**, *10*, e0133081. [[CrossRef](#)]
23. Soultati, A.; Stares, M.; Swanton, C.; Larkin, J.; Turajilic, S. How should clinicians address intratumour heterogeneity in clear renal cell carcinoma? *Curr. Opin. Urol.* **2015**, *25*, 358. [[CrossRef](#)]
24. Wach, S.; Taubert, H.; Weigelt, K.; Hase, N.; Köhn, M.; Misiak, D.; Hüttelmaier, S.; Stöhr, C.G.; Kahlmeyer, A.; Haller, F.; et al. RNA Sequencing of Collecting Duct Renal Cell Carcinoma Suggests an Interaction between miRNA and Target Genes and a Predominance of Deregulated Solute Carrier Genes. *Cancers* **2019**, *12*, 64. [[CrossRef](#)]
25. Ingimarsson, J.P.; Sigurdsson, M.I.; Hardarson, S.; Petursdottir, V.; Jonsson, E.; Einarsson, G.V.; Gudbjartsson, T. The impact of tumour size on the probability of synchronous metastasis and survival in renal cell carcinoma patients: A population-based study. *BMC Urol.* **2014**, *14*, 72. [[CrossRef](#)]
26. Eble, J.N.; Sauter, G.; Epstein, J.; Sesterhenn, I. *World Health Organization Classification of Tumors. Pathology and Genetics of Tumours of the Urinary System and Male Genital Organs*; IARC Press: Lyon, France, 2004.
27. Edge, S.B.; Compton, C.C. The American Joint Committee on Cancer: The 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann. Surg. Oncol.* **2010**, *17*, 1471. [[CrossRef](#)] [[PubMed](#)]
28. Fuhrman, S.A.; Lasky, L.C.; Limas, C. Prognostic significance of morphologic parameters in renal cell carcinoma. *Am. J. Surg. Pathol.* **1982**, *6*, 655. [[CrossRef](#)] [[PubMed](#)]
29. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114. [[CrossRef](#)]
30. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast-spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357. [[CrossRef](#)]
31. Pertea, M.; Pertea, G.M.; Antonescu, C.M.; Chang, T.C.; Mendell, J.T.; Salzberg, S.L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **2015**, *33*, 290. [[CrossRef](#)]
32. Pertea, M.; Kim, D.; Pertea, G.M.; Leek, J.T.; Salzberg, S.L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **2016**, *11*, 1650. [[CrossRef](#)] [[PubMed](#)]
33. Van der Auwera, G.A.; Carneiro, M.O.; Hartl, C.; Poplin, R.; Del Angel, G.; Levy-Moonshine, A.; Jordan, T.; Shakir, K.; Roazen, D.; Thibault, J.; et al. From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinform.* **2013**, *43*, 10–11.
34. Chandrashekar, D.S.; Bashel, B.; Balasubramanya, S.A.H.; Creighton, C.J.; Ponce-Rodriguez, I.; Chakravarthi, B.V.S.K.; Varambally, S. UALCAN: A portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia* **2018**, *19*, 649. [[CrossRef](#)]
35. Lanczky, A.; Nagy, A.; Bottai, G.; Munkácsy, G.; Szabó, A.; Santarpia, L.; Győrffy, B. miRpower: A web-tool to validate survival-associated miRNAs utilizing expression data from 2178 breast cancer patients. *Breast Cancer Res. Treat.* **2016**, *160*, 439. [[CrossRef](#)]
36. DePristo, M.A.; Banks, E.; Poplin, R.; Garimella, K.V.; Maguire, J.R.; Hartl, C.; Philippakis, A.A.; del Angel, G.; Rivas, M.A.; Hanna, M.; et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **2011**, *43*, 491. [[CrossRef](#)] [[PubMed](#)]

