Original Research

Radiomic profiling of clear cell renal cell carcinoma reveals subtypes with distinct prognoses and molecular pathways

Peng Lin [a], Yi-qun Lin [b], Rui-zhi Gao [a], Rong Wen [a], Hui Qin [a], Yun He [a], Hong Yang [a],*

[a] Department of Medical Ultrasound, The First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi Zhuang Autonomous Region 530021, China
[b] Department of Radiology, The Affiliated Dongnan Hospital of Xiamen University, Zhangzhou, Fujian Province 363020, China

ARTICLE INFO

ABSTRACT

Background: To identify radiomic subtypes of clear cell renal cell carcinoma (ccRCC) patients with distinct clinical significance and molecular characteristics reflective of the heterogeneity of ccRCC.

Methods: Quantitative radiomic features of ccRCC were extracted from preoperative CT images of 160 ccRCC patients. Unsupervised consensus cluster analysis was performed to identify robust radiomic subtypes based on these features. The Kaplan–Meier method and chi-square test were used to assess the different clinicopathological characteristics and gene mutations among the radiomic subtypes. Subtype-specific marker genes were identified, and gene set enrichment analyses were performed to reveal the specific molecular characteristics of each subtype. Moreover, a gene expression-based classifier of radiomic subtypes was developed using the random forest algorithm and tested in another independent cohort ($n = 101$).

Results: Radiomic profiling revealed three ccRCC subtypes with distinct clinicopathological features and prognoses. *VHL, MUC16, FBN2*, and *FLG* were found to have different mutation frequencies in these radiomic subtypes. In addition, transcriptome analysis revealed that the dysregulation of cell cycle-related pathways may be responsible for the distinct clinical significance of the obtained subtypes. The prognostic value of the radiomic subtypes was further validated in another independent cohort (log-rank $P = 0.015$).

Conclusion: In the present multi-scale radiogenomic analysis of ccRCC, radiomics played a central role. Radiomic subtypes could help discern genomic alterations and non-invasively stratify ccRCC patients.

Introduction

Clear cell renal cell carcinoma (ccRCC) is the most common subtype of kidney cancer, with over 400,000 cases worldwide annually and nearly 175,000 deaths per year [1,2]. Although several treatment approaches, including radical nephrectomy and ablation, have improved the clinical outcome, the 5-year relative survival rate of ccRCC patients remains poor. With the advancement in high-throughput genomic technology, ccRCC is no longer considered a single uniform entity. Instead, ccRCC is now recognized as an extremely complex cancer with intra-tumor and inter-tumor heterogeneity, which limits individualized treatment and heterogeneous clinical outcomes [3-5]. While the traditional pathological stage is a strong survival indicator, it is insufficient to characterize individual clinical outcomes and molecular characteristics. In

the era of precision medicine, effective risk stratification and the estimation of specific molecular characteristics of an individual could spur personalized treatment and help improve the understanding of the molecular mechanisms underlying ccRCC.

Decades ago, several molecular subtyping approaches for ccRCC based on multi-omics data, including genomic, transcriptomic, and proteomic data, have been proposed to classify the patients' characteristics and predict their prognosis and response to treatment [6-8]. However, these molecular classifiers are limited by the requirement of postoperative pathological or biopsy tissues. Therefore, a non-invasive and reproducible approach is urgently required to evaluate the detailed molecular characteristics and estimate the prognosis of patients. Radiomics is a newly emerging algorithm that analyzes high-throughput imaging features extracted from medical images [9]. The emergence of radiomics

---

* Corresponding author.
E-mail address: yanghong@gxmu.edu.cn (H. Yang).
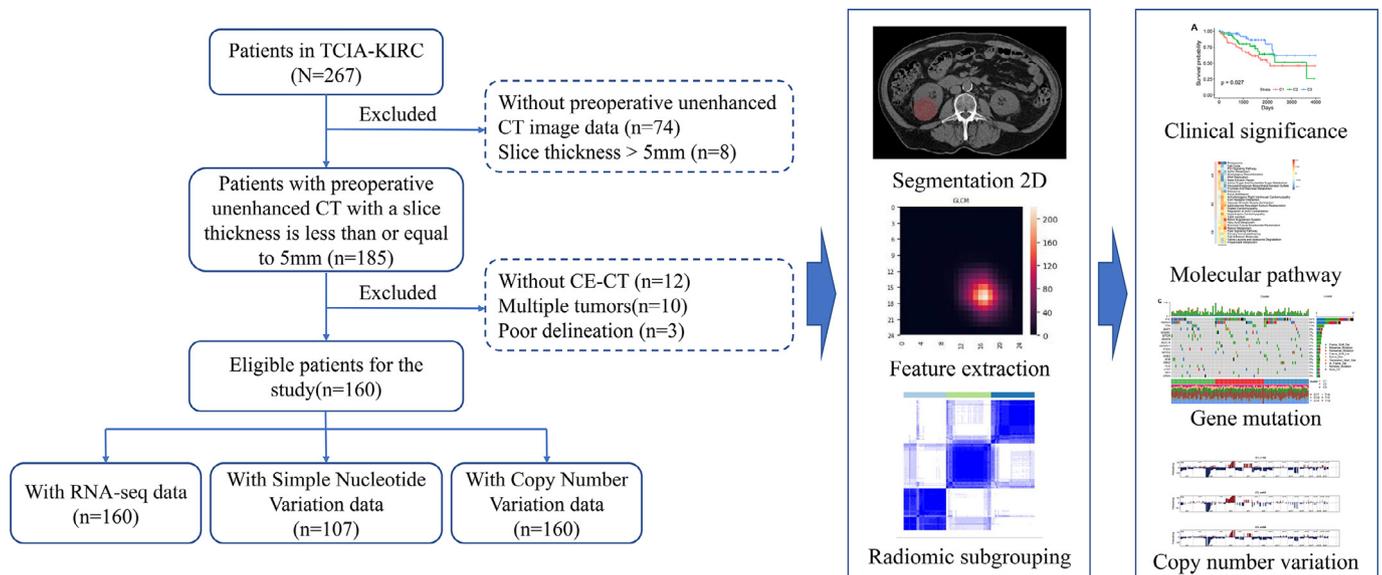
Fig. 1. Flowchart showing the patient selection and radiomic image analysis process
TCGA-KIRC: The Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma, CECT: contrast-enhanced computed tomography, 2D: two-dimensional.

has enabled the non-invasive assessment of molecular characteristics of tumors. Integrated analysis of medical images and multi-omics molecular data helps link our understanding of macroscopic medical images to microscopic molecular counterparts that contribute to patient risk stratification [10]. A previous study identified subtypes of glioblastoma using quantitative perfusion magnetic resonance imaging; patients belonging different subtypes were found to have distinct prognoses and treatment responses [11]. Another study found that quantitative imaging phenotypes could be effective for breast cancer stratification [12]. These findings highlight the importance of radiomics in tumor stratification.

In the present study, we aimed to develop a novel ccRCC classifier using quantitative radiomic profiles. By integrating clinicopathological and multi-omics molecular data, we aimed to elucidate the radiomic subtypes associated with distinct clinical outcomes and molecular characteristics. Our findings could provide a novel approach for determining patients' clinical prognosis and for providing personalized, precision-based clinical care.

Materials and methods

Data source and inclusion criteria

All data were from public datasets and study were approved by our institutional review board. Informed consent was not required in the present study because the analyses were based on public and freely available datasets for scientific purposes. Two publicly accessible ccRCC cohorts were retrospectively analyzed in the present study. In cohort 1, The Cancer Genome Atlas (TCGA) Kidney Renal Clear Cell Carcinoma collection consisting of the imaging data of 267 ccRCC patients was downloaded from The Cancer Imaging Archive database [13,14]. Corresponding clinical follow-up data and clinicopathological characteristics were acquired from the Pan-Cancer analysis project by TCGA [15]. In cohort 2, microarray-based gene expression data and clinical data of 101 ccRCC patients were downloaded from the ArrayExpress database (https://www.ebi.ac.uk/arrayexpress/; accession ID: E-MTAB-1980) [16].

The inclusion criteria of the patients in cohort 1 are summarized in Fig. 1. The inclusion criteria were as follows: (a) patients with pathologically diagnosed ccRCC, (b) those who preoperatively underwent unenhanced CT with slice thickness ≤ 5 mm (c), those who preoperatively

Table 1
Characteristics of the Study Population.

Characteristic	Cohort 1($n = 160$)	Cohort 2($n = 101$)
Age (years)		
<60	79	41
≥60	81	60
Gender		
Male	105	77
Female	55	24
Histological grade		
G1	1	13
G2	62	59
G3	70	22
G4	27	5
NA	0	2
Pathological stage		
Stage I	85	66
Stage II	14	10
Stage III	40	13
Stage IV	21	12

underwent unenhanced CT and contrast-enhanced CT (CECT), (d) those with solitary tumors, and (e) those with well-preserved imaging data.

Tumor segmentation and radiomic feature extraction

The largest tumor slice in the cross-section map was subjected to tumor segmentation. The region of interest (ROI) was manually segmented using ITK-SNAP software (version 3.8) [17]. During tumor segmentation, available CECT images were also used to recognize and determine the tumor boundaries in all the patients. The manual segmentation of ROI was performed by a single radiologist.

Radiomic features were extracted using Intelligence Foundry (GE Healthcare, version 1.3). In total, 122 radiomic features were extracted, including 18 first-order features, 23 gray-level co-occurrence matrix features, 16 gray-level run-length matrix features, 16 gray-level size-zone matrix features, five neighboring gray-tone difference matrix features, 13 gray-level dependence matrix features, 18 shape features, and 13 textural phenotype features. All 122 radiomic features fully respecting standard Image Biomarker Standardization Initiative.
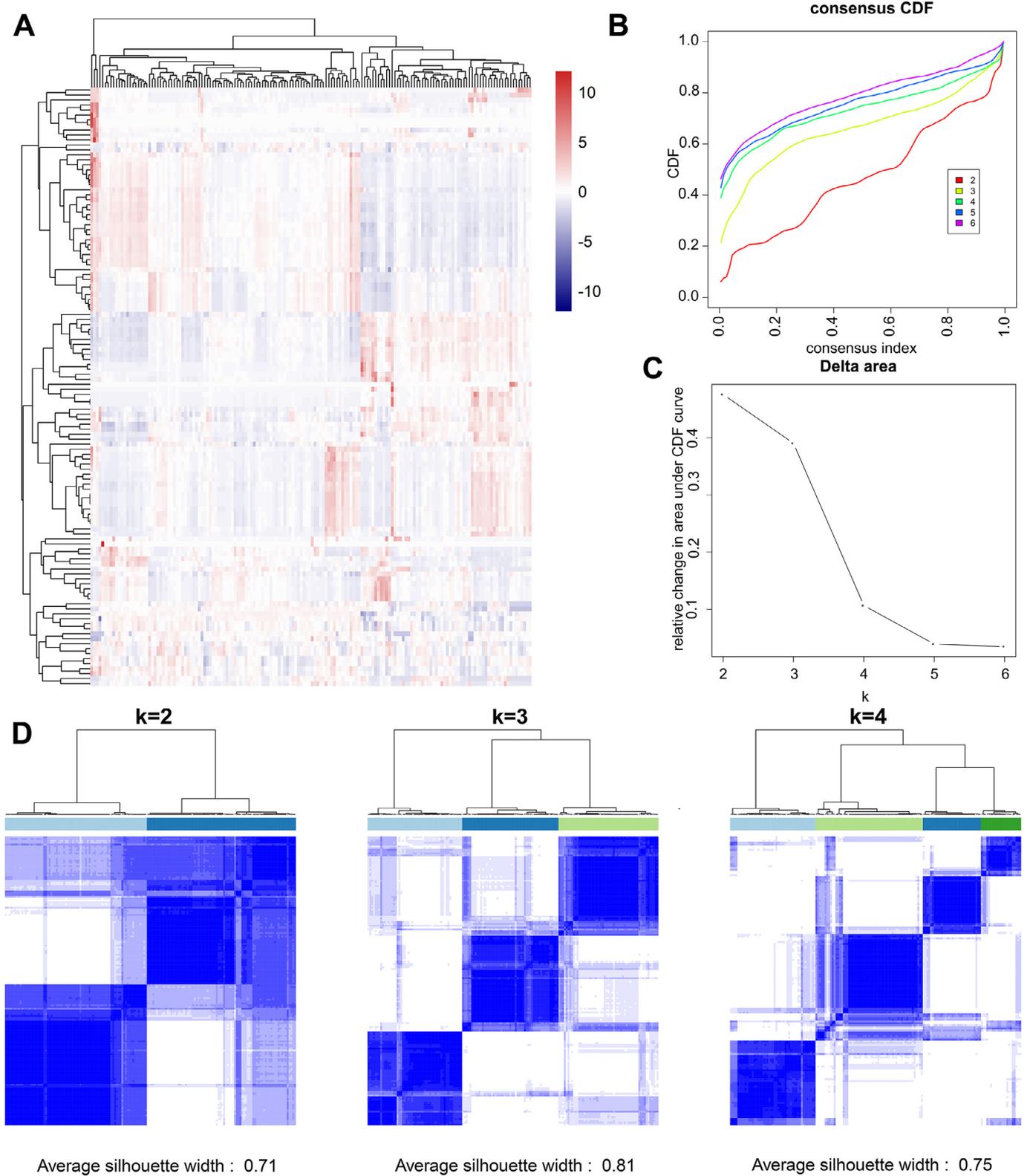
Fig. 2. Consensus clustering for radiomic stratification of ccRCC patients

(A) Heatmap of extracted radiomic features reveals a certain heterogeneity of medical images of ccRCC patients.

(B) The cumulative distribution function (CDF) plot shows the probability distribution for each k.

(C) The delta plot shows the relative change between k and $k-1$ under the CDF curve.

(D) Consensus matrix silhouette plots ($k = 2$–4) are displayed.
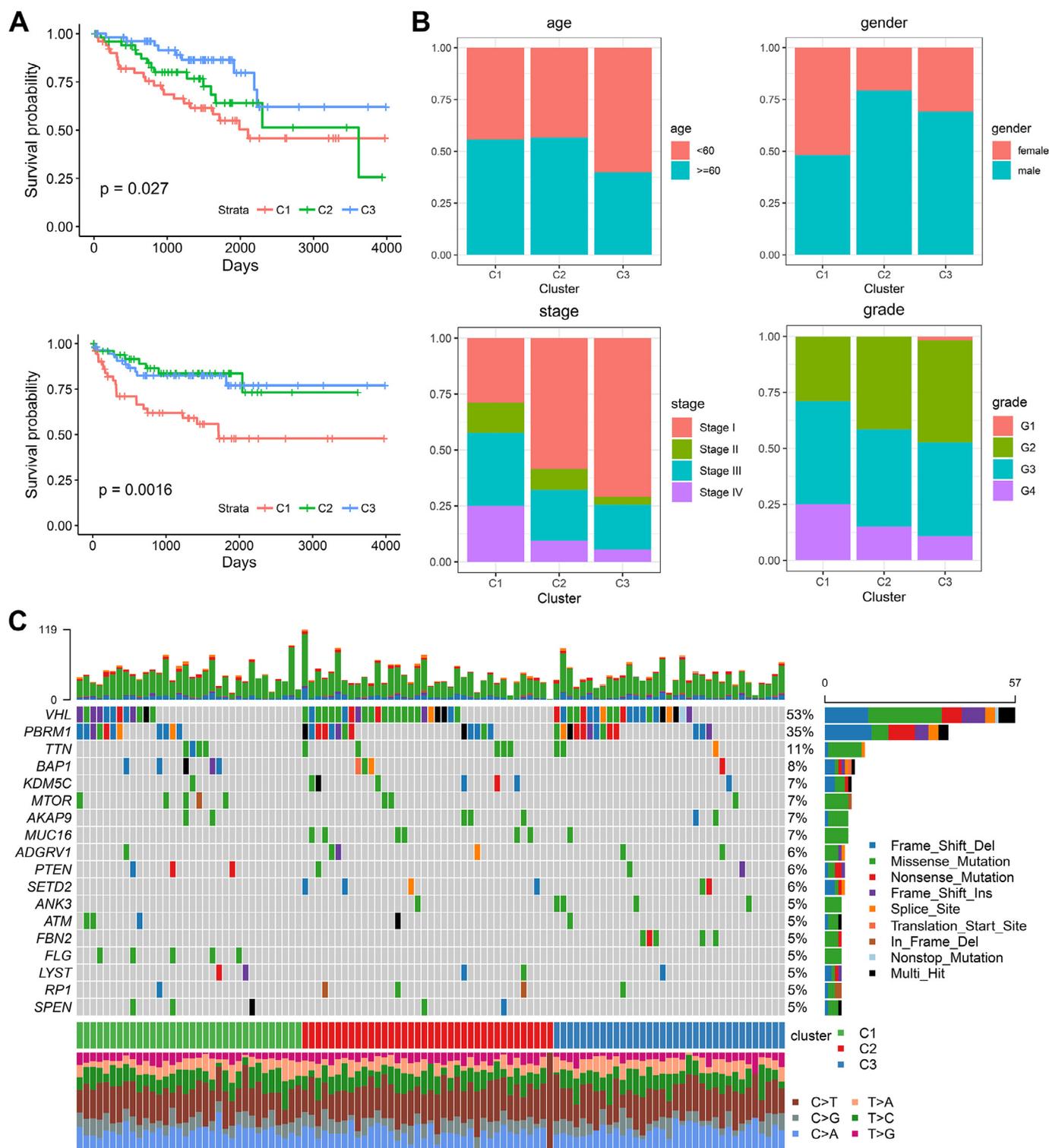
Fig. 3. Radiomic stratification of ccRCC and the clinicopathological and genomic correlations
(A) Kaplan–Meier curves for overall survival and progression-free interval based on radiomic stratification (log-rank test).
(B) The distribution of age, gender, stage, and grade in tumors with different radiomic stratifications.
(C) Genes with mutation rates greater than 5% are ranked on the left side in the waterfall plot, where the mutational burden for each sample is calculated above the legend.

Radiomic subgrouping analysis

Consensus clustering, an unsupervised subtype discovery method, was performed to identify the intrinsic radiomic subtypes of the ccRCC patients using the "CancerSubtypes" R package [18]. In particular, parti-tioning around medoid (a type of consensus clustering) was performed with the Euclidean distance metric to generate the subtypes. In total, 80% of the patients were randomly sampled each time to obtain the randomly sampled data set. The process was repeated 1000 times to ob-tain the conservative matrix result. Subtype numbers from two to six

Table 2
Mutation frequency in different radiomic subtypes.

Gene	C1 (mutation/wild)	C2 (mutation/wild)	C3 (mutation/wild)	P value
VHL	12/22	24/14	21/14	0.038[a]
PBRM1	11/23	13/25	13/22	0.915[a]
TTN	4/30	5/33	3/32	0.864[b]
BAP1	5/29	3/35	1/34	0.207[b]
KDM5C	1/33	6/32	1/34	0.103[b]
MTOR	5/29	2/36	1/34	0.166[b]
AKAP9	2/32	3/35	2/33	>0.99[b]
MUC16	0/34	6/32	1/34	0.017[b]
ADGRV1	1/33	3/35	2/33	0.870[b]
PTEN	3/31	1/37	2/33	0.439[b]
SETD2	0/34	4/34	2/33	0.185[b]
ANK3	0/34	1/37	4/31	0.078[b]
ATM	3/31	1/37	1/34	0.443[b]
FBN2	0/34	0/38	5/30	0.006[b]
FLG	5/29	0/38	0/35	0.003[b]
LYST	2/32	2/36	1/34	0.866[b]
RP1	1/33	3/35	1/34	0.617[b]
SPEN	3/31	2/36	0/35	0.233[b]

[a] : Chi-square test;.
[b] : Fisher test.

were analyzed. The optimal cluster number was determined according to the following factors: the relative change between k and k - 1 under the cumulative distribution function (CDF) curve, significance of separation of the consensus clusters in the heatmap, and average silhouette distance of the consensus clusters.

Molecular characteristic identification

Masked somatic mutation data, which were processed with VarScan2 software, were downloaded using "TCGAbiolinks" package in R software [19]. Mutation Annotation Format (MAF) files were then processed using the "maftools" R package, which provides analysis and visualization models for MAF files from large-scale sequencing studies [20]. Similarly, masked copy number alteration data were downloaded using "TCGAbiolinks" package in R software [19]. GISTIC_2.0 was used to detect significant amplification or deletion regions in the genomes [21]. Following this, the burden of copy number gain or loss was calculated on the basis of the total number of genes with a change in copy number at the focal and arm levels.

Genome-wide RNA sequencing data generated using the Illumina HiSeq 2000 RNA Sequencing platform were also acquired from the Pan-Cancer project by TCGA [22]. Based on RSEM-normalized gene-level data, the differential expression of one subtype versus all other subtypes was identified using the "limma" package in R software. Following this, upregulated genes in each radiomic subtype were used to estimate specific activated molecular pathways using gene set enrichment analysis (GSEA). GSEA was performed using the GSEA software (version 3.0) [23] using the GSEA pre-ranked function. The goal of GSEA is to determine whether differentially expressed genes are randomly distributed or primarily found at the top or bottom of specific pathway. Based on the results of differentially expression analysis, enrich score for each pathway was calculated, significantly enriched pathways were ranked based on enrich scores. Top 10 enriched pathways with largest enrich scores were identified as the pathway alterations characteristics for each subtype, respectively. Annotated pathways were based on Kyoto Encyclopedia of Genes and Genomes (version 7.1) annotated pathways downloaded from the Molecular Signatures Database [23]. The top 10 significantly enriched pathways were used for heatmap visualization.

Radiomic subtype validation

In order to examine the risk stratification performance of our radiomic subtypes in ccRCC patients with longer follow-up times, cohort 2 was used to validate the clinical significance of radiomic subtypes.

First, the marker genes for radiomic subtypes were analyzed using differentially expressed genes in cohort 1. Based on these marker genes, the least absolute shrinkage and selection operator (LASSO) algorithm was further used to select the core genes that could be used to identify the subtypes. LASSO algorithm selection was performed using the "glmnet" R package (version 2.0–18) [24]. Finally, a classifier was developed using the random forest algorithm and was used to predict the radiomic subtypes in the validation cohort. When developing the classifier model, five-fold cross validation was adopted to get the best parameters. The random forest algorithm was used to develop and validate the classifier using the "rpart" R package (version 4.1–15).

Statistical analysis

The overall survival (OS) and progression-free interval (PFI) of ccRCC patients belonging to distinct radiomic subtypes were compared. The endpoint of OS was death from any cause, while that of PFI was the occurrence of a new tumor event, including disease progression, local recurrence, distant metastasis, or a new primary tumor event, or the occurrence of death due to cancer without a new tumor event[15]. Survival differences among the patients were estimated by Kaplan–Meier (K–M) curves using the log-rank test. The associations between subtypes and clinical factors were assessed using the chi-squared test. The chi-squared test or Fisher's exact test was performed to identify significant differences in the distribution of each mutation among the subtypes. In addition, univariate Cox analysis was performed to assess the survival of specific upregulated genes in each subtype. A P value < 0.05 was considered statistically significant. All statistical analyses were performed using R software (version 3.6.0).

Results

Demographic and clinical characteristics

In total, 160 and 101 ccRCC patients were included in cohorts 1 and 2, respectively. Demographic and clinical characteristics of the patients are provided in Table 1.

Radiomic subtypes in the discovery cohort

A total of 122 radiomic features were extracted from ROIs (Fig. 2A). Based on consensus clustering of the radiomic features, a three-cluster solution was identified as the optimal solution: (1) The consensus CDF and delta plot exhibited little relative change between $k = 3$ and $k = 4$
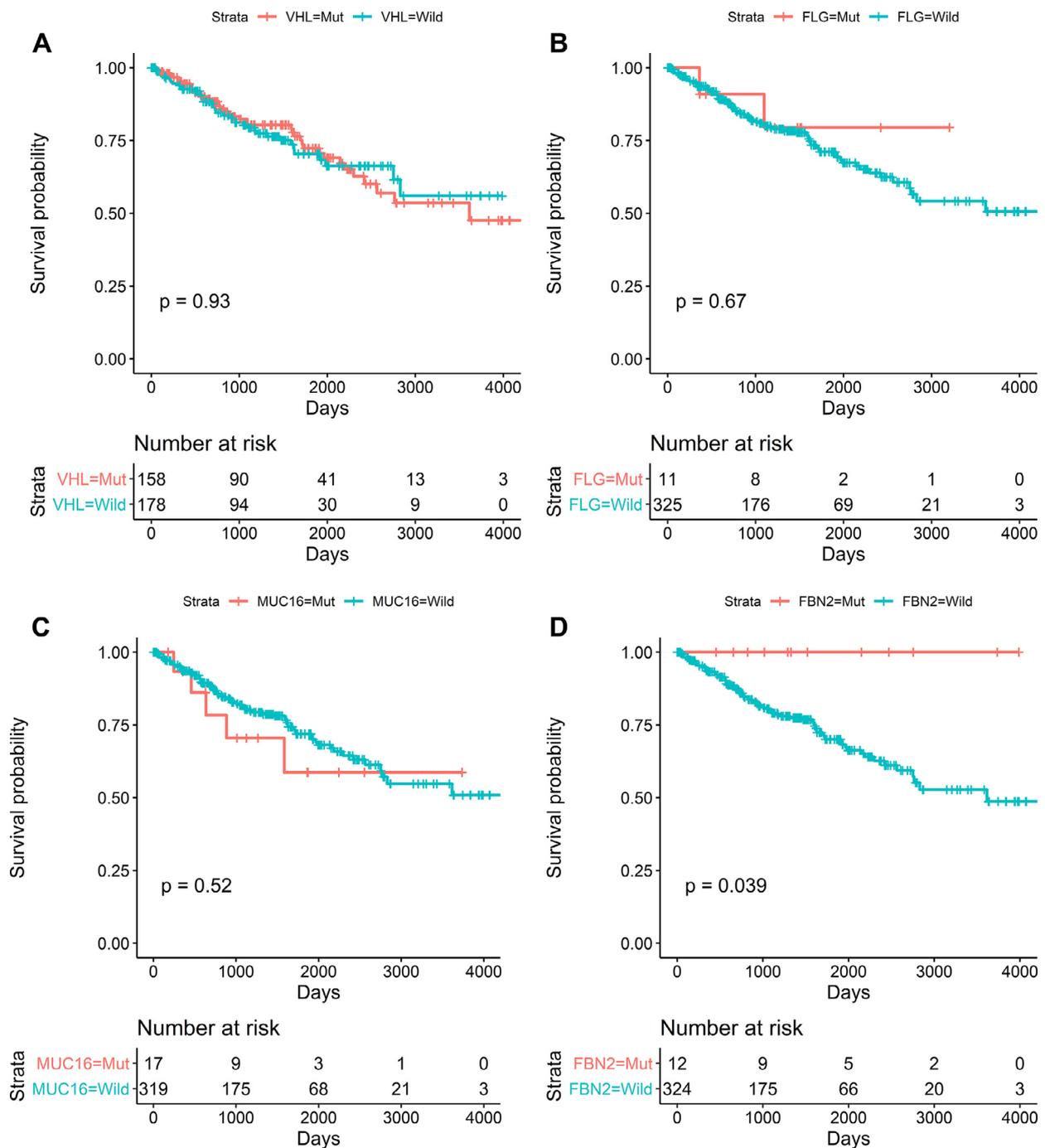
Fig. 4. Relationships between gene mutation and patients' OS
(A) VHL; (B) FLG; (C) MUC16; (D) FBN2.

(Fig. 2B and 2C). (2) The consensus matrix heatmap revealed three clusters with significant interconnectivity (Fig. 2D). (3) The average silhouette distance for $k = 3$ (0.81) was greater than that for $k = 2$ (0.71) or $k = 4$ (0.75) and did not have significant negative values. Accordingly, the 160 ccRCC patients could be divided into three subtypes: 52 patients (32.5%) in cluster 1 (C1), 53 patients (33.1%) in cluster 2 (C2), and 55 patients (34.4%) in cluster 3 (C3).

Distinctive clinical characteristics of radiomic subtypes

Differences in clinical characteristics and clinical outcomes among the three subtypes were then explored. K–M curves indicated that the C1 subtype had significantly lower OS ($P = 0.027$) and PFI ($P = 0.002$)

than the C2 and C3 subtypes (Fig. 3A). Furthermore, the distribution of different clinical factors among the subtypes was not random. Significant differences were observed among the subtypes with regard to gender (chi-square value = 11.75, $P = 0.003$) and pathological stage (chi-square value = 13.01, $P = 0.001$). No association was noted between the subtypes and age (chi-square value = 4.03, $P = 0.133$) and histological grade (chi-square value = 3.953, $P = 0.139$) (Fig. 3B).

Molecular characteristics of each subtype

In addition to assessing the differences in clinical characteristics among the three subtypes, the differences in molecular characteristics among them were assessed in the present study. All the patients had
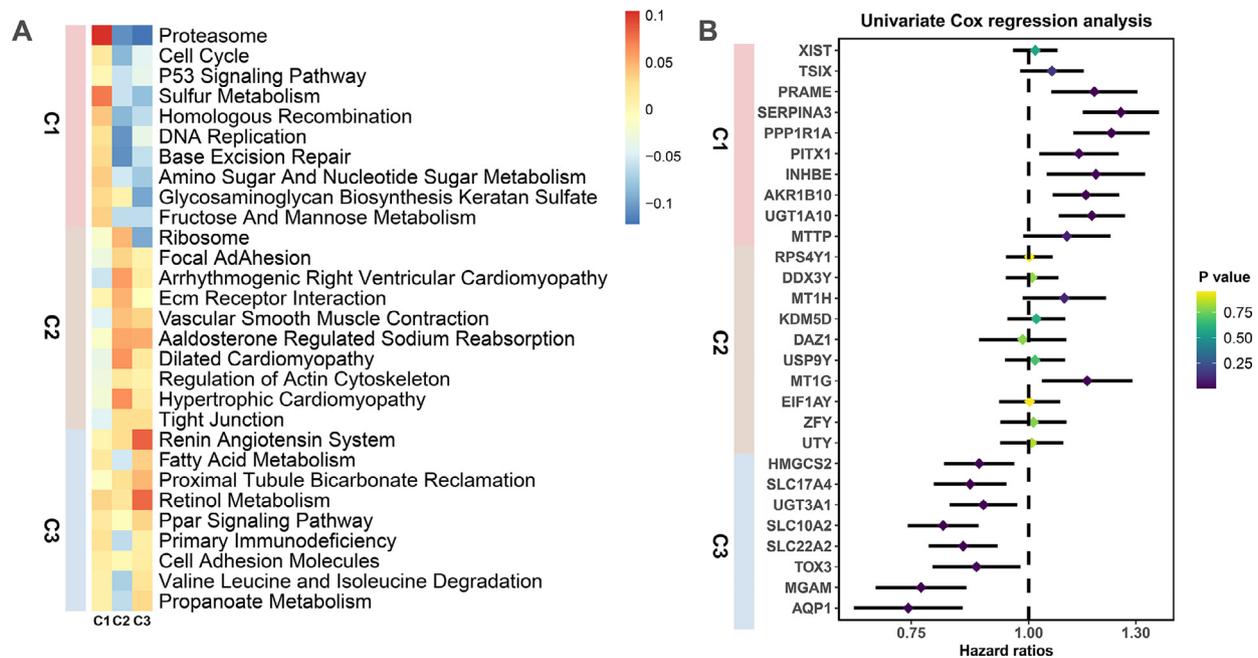
Fig. 5. Distinct enriched pathways and genes between subtypes
(A) GSEA reveals distinct enriched gene sets between the radiomic subtypes. In the heatmap, rows represent the enriched gene sets, while columns represent the enrichment scores for each subtype.
(B) The top 10 upregulated genes in each subtype and their survival significance estimated by univariate Cox analysis.

RNA-seq and masked copy number alteration data, while 107 out of 160 patients had gene mutation data. The association of the clusters with the gene mutation data is shown in Fig. 3C; 18 genes had mutation rates over 5%. The tumors classified into different subtypes had specific gene mutation events. *VHL* was the most frequently mutated gene; it was observed in 53% of the patients (57/107). Next, the frequency of each gene mutation was compared among the subtypes. While *VHL* had a high mutation rate in all the subtypes, *VHL* mutations were less frequently observed in the C1 subtype than in the C2 and C3 subtypes (Table 2). In addition, *FLG, MUC16*, and *FBN2* mutations were specifically observed in the C1, C2, and C3 subtypes, respectively (Table 2). Based on TCGA-KIRC cohort, K-M analysis also revealed that VHL, FLG and MUC16 mutation status did not show survival differences (Fig. 4A-C) while patients with FBN2 mutation have superior OS (Fig. 4D).

A total of 20,531 genes were used to differentially expression analysis. By further using GSEA, which revealed that the tumors belonging to the C1 subtype were characterized by significantly activated proteasome, cell cycle, and p53 signaling pathway genes (Fig. 5A). Univariate Cox analysis of the top 10 upregulated genes also suggested that most upregulated genes in the C1 subtype were risk factors, while most upregulated genes in the C3 subtype were protective factors. However, most of the top 10 upregulated genes in the C2 subtype had no significant survival significance (Fig. 5B).

The copy number burden of each sample was also calculated. Fig. 6A shows the copy number frequency distribution across all chromosomes in the three radiomic subtypes. A higher burden of copy number loss was noted in the C1 subtype than in the C2 and C3 subtypes at the focal and arm levels, respectively (Fig. 6B). Next, the deletions of chromosomal regions in the C1 subtype were observed (Fig. 6C).

Radiomic subtype validation

Based on the criteria of |log2FC| > 1 and *P* < 0.05, 25 upregulated and 37 downregulated genes were identified in the C1 subtype, 23 upregulated and three downregulated genes were identified in the C2 subtype, and eight upregulated and 22 downregulated genes were identi-

fied in the C3 subtype. After removing duplicated genes, 89 genes were finally included in cohort 1 as cluster markers for further classifier development (Fig. 7A). By LASSO analysis, 21 core genes were further selected for model construction using the random forest algorithm. Based on the developed classifier, the patients in cohort 2 were also classified into C1–C3 radiomic subtypes. In cohort 1, the patients belonging to the C1 subtype were found to have a significantly lower survival rate and a higher risk of postoperative death than those belonging to the C2 and C3 subtypes (Fig. 2A). This result was further validated in cohort 2, an external clinically annotated ccRCC cohort (Fig. 7B). More importantly, the distribution of different clinicopathological features in radiomic subtypes was found to be similar in training and validation cohort (Fig. 7C).

Discussion

Molecular subtypes based on a multi-omics view of ccRCC have broadened the understanding of the heterogeneity of ccRCC [25,26]. However, because of the difficulties involved in performing a biopsy and sequencing assay for each patient, non-invasive approaches are needed to classify patients with distinct clinical outcomes and molecular characteristics. The global radiomic information obtained in the present study provides novel insights into the clinicopathological and molecular characteristics of ccRCC. Integrated radiomic characterization and genomic data revealed clinically relevant subtypes and non-invasively acquired genomic data for specific subtypes. Our integrated analysis revealed that radiomic profiles are heterogenous, having the potential for risk stratification. By consensus clustering, the patients were classified into three distinct subtypes. Interestingly, the patients belonging to the three subtypes had distinct clinical outcomes.

Radiogenomics serves as a bridge that connects non-invasive radiomic phenotypes with genomic alteration data [27]. Feng Z et al. reported that the use of CT features is a potential and feasible approach for predicting *BAP1* mutations in ccRCC patients[28]. Moreover, a study involving 233 ccRCC patients identified associations between some CT features, such as well-defined tumor margins and nodular tumor enhance-
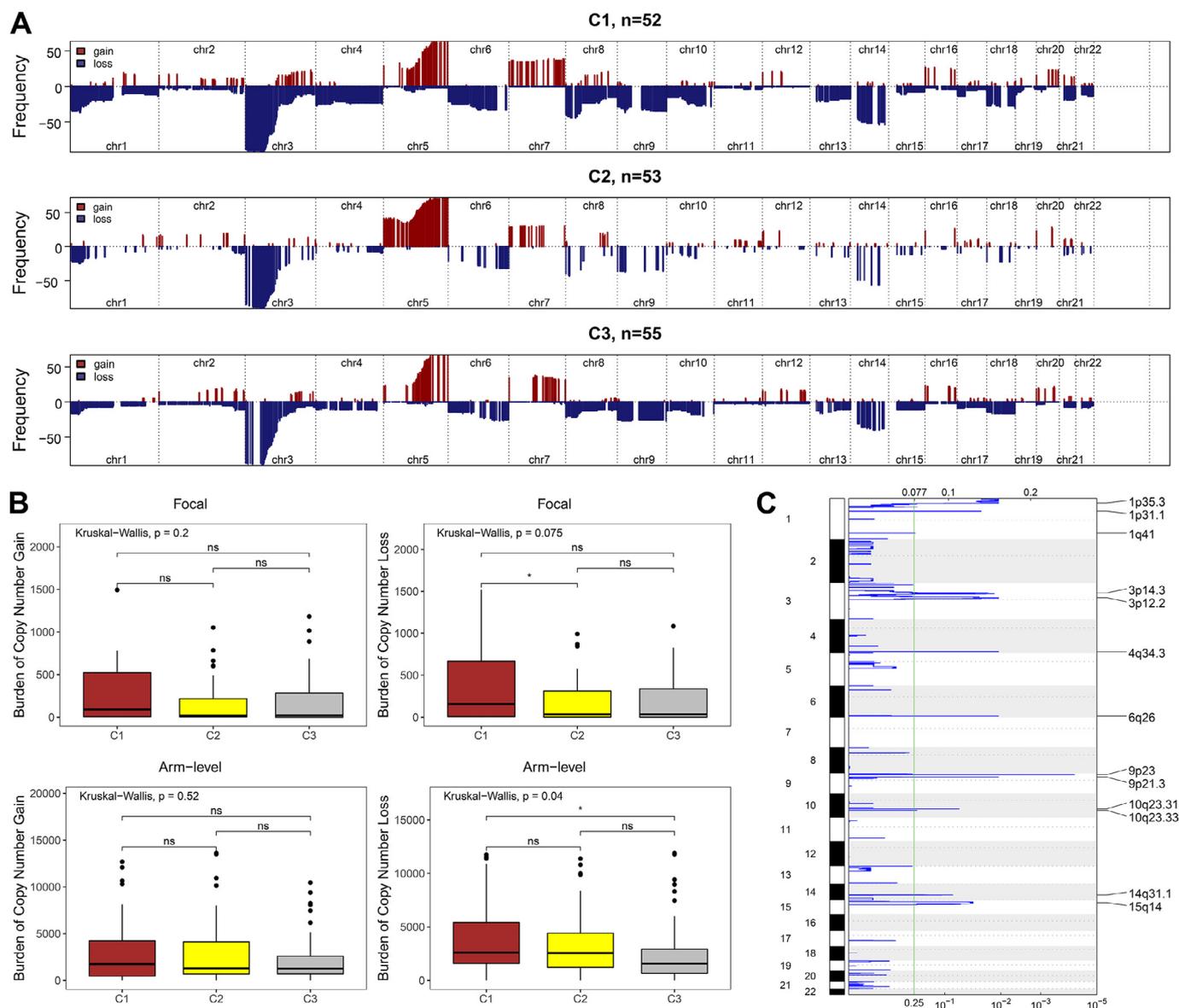
Fig. 6. Copy number alterations across radiomic subtypes

(A) Copy number profiles for the C1, C2, and C3 subtypes. Dark red represents gain, while midnight blue represents loss. Gene segments are placed according to their location on chromosomes, ranging from chromosome 1 to chromosome 22.

(B) The distribution of focal and broad copy number alterations among the radiomic subtypes.

Ns represents not significant ($P > 0.05$). * represents $P \leq 0.05$.

(C) A detailed cytoband with focal deletion (right) in the C1 subtype generated using GISTIC_2.0 software.

ment, and the presence of mutations in several genes (*VHL, PBRM1, SETD2, KDM5C,* and *BAP1*) [29]. These studies have revealed the ability of radiomics to access internal tumor data. With the rapid advances in next-generation sequencing and increased public multi-omics data, it is now possible to perform integrated analysis of radiomic phenotypes and genomic data. The present findings indicate that medical imaging phenotypes may help infer dysregulated gene mutation statuses or molecular pathways.

Genetically, ccRCC is characterized by a high proportion of *VHL* mutations, which play a central role in the current understanding of the fundamental principles of ccRCC development [30]. *VHL* loss often leads to the unregulated accumulation of hypoxia-inducible factor 1, resulting in activated metabolism and angiogenesis [31]. Accordingly, it is a reasonable phenomenon that the features of medical images could, in a certain way, reflect *VHL* mutations. A previous study revealed that *VHL* mutations are significantly associated with ill-defined tumor margins,

multicystic tumors, nodular tumor enhancement, and the intratumoral vasculature [29]. In the present study, the proportion of *VHL* mutations was found to differ in different subtypes; in particular, it was low in the C1 subtype. However, a recent meta-analysis suggested that *VHL* mutations are not significantly associated with the survival of ccRCC patients [32] In addition, *FLG, MUC16*, and *FBN2* mutations specifically occurred in the C1, C2, and C3 subtypes, respectively. FLG have been reported that significantly mutated or amplified in some types of cancer [33,34]. However, the molecular function of FLG in ccRCC did not be fully studied. Future in vivo and/or in vitro analysis were required to validate its molecular characteristics in ccRCC. MUC16, also known as the CA125 antigen, is the third most frequently mutated gene in TCGA database [35]. Several previous studies have revealed that MUC16 mutation is closely related to prognosis of many types of cancer [36,37]. Similarly, MUC16 is frequently mutated in papillary renal cell carcinoma [38]. FBN2 is key components of human microfibrils and could regulate TGF-
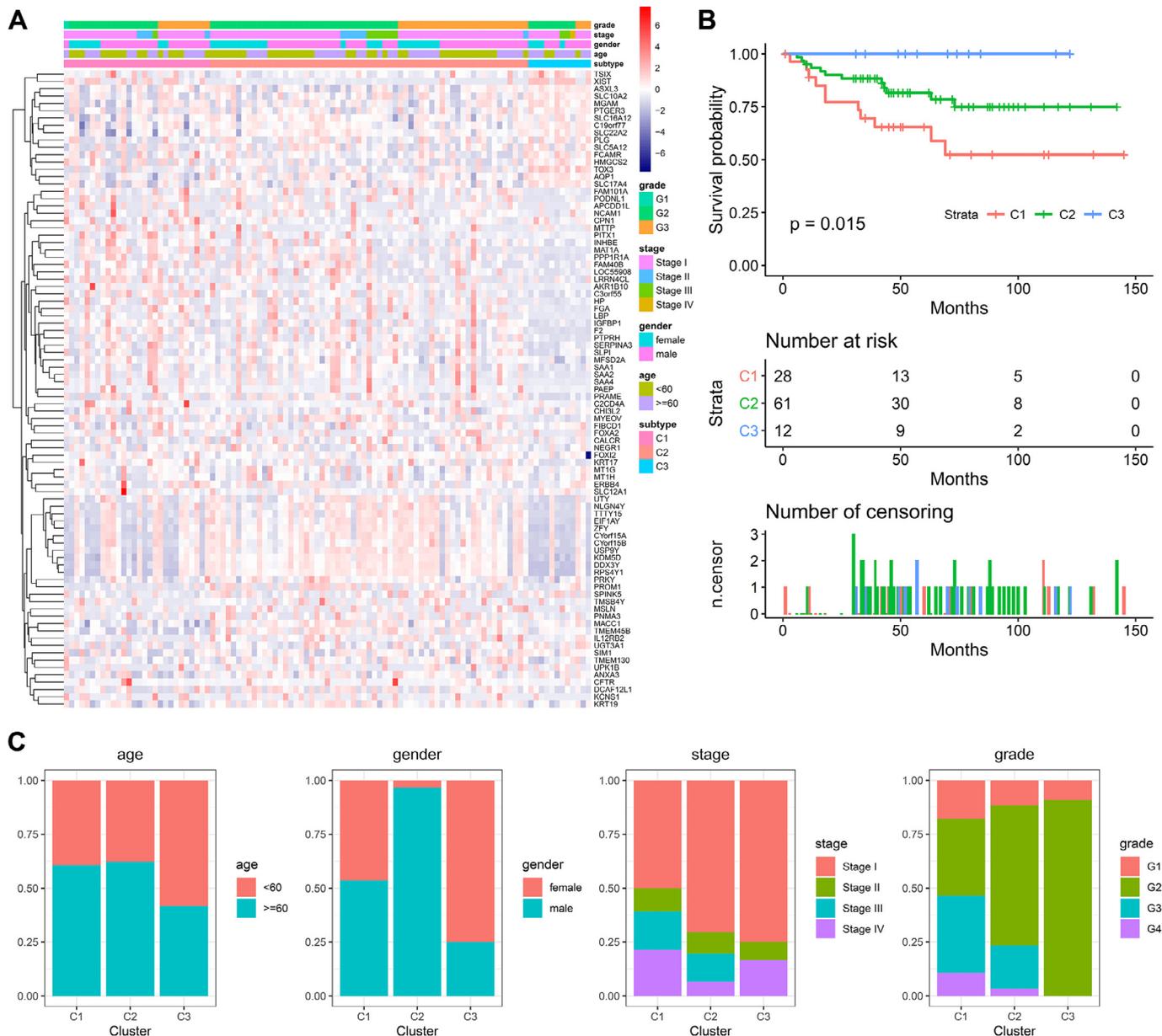
Fig. 7. Radiomic subtype validation using another independent cohort
(A) The heatmap of marker genes in the validation cohort displays a similar expression pattern to that in the training cohort.
(B) Kaplan–Meier curves for radiomic subtype classification based on gene expression data.
(C) The distribution of age, gender, stage, and grade in tumors belonging to different radiomic subtypes.

b signaling. FBN2 was very frequently hypermethylated in ccRCC and the gene demonstrated a relatively high somatic mutation rate [39]. Hypermethylation of FBN2 has been previously reported in renal cell carcinoma and knockdown for FBN2 also lead to an anchorage-independent growth advantage[40]. We found that FBN2 mutation is related to the prognosis of ccRCC. FBN2 mutation also related with patients' OS, it may present as potential biomarker for ccRCC. However, mRNA expression levels of these genes are relatively low in ccRCC, molecular mechanisms are need further exploration.

In addition to gene mutation, transcriptome data analysis revealed some interesting findings. Surprisingly, in the C1 subtype, upregulated genes were mainly detected in the proteasome, cell cycle, and p53 signaling pathways, indicating that sustained proliferation was a response to poor survival in this subgroup. Notably, C1-specific upregulated genes and C3-specific upregulated genes were found to have a distinct prognostic value. Previous studies have reported that some genes are acti-

vated and involved in the development and progression of ccRCC. For example, TOX3 inhibits ccRCC cell migration and invasion via the transcriptional regulation of *SNAI1* [41]. Moreover, *SLC10A2* expression is an independent prognostic indicator of OS of ccRCC patients and is downregulated in tyrosine kinase inhibitor-resistant samples [42]. Based on our findings and previous findings, these marker genes may have moderate potential for acting as biomarkers for monitoring ccRCC prognosis.

The limitations of the present study include its retrospective nature. However, we have used two independent cohorts to validate the results. Future prospective research will be needed in a larger cohort of ccRCC patients in order to assess the generalizability of our stratification. Another limitation is the use of two-dimensional tumor segmentation for radiomic feature extraction. Three-dimensional segmentation may be more suitable for tumor texture extraction. To overcome this limitation, we sincerely hope that automated segmentation will be the gold stan-

dard instead of manual segmentation for ccRCC in the future. Moreover, we did not analyze CECT data because TCGA cohort was collected from multiple hospitals and the CECT clinical protocols were inconsistent.

In conclusion, the present comprehensive radiomic characterization of ccRCC provides comprehensive and integrated insights and is more suitable for clinical application. Our findings revealed that radiomic data could be a potential non-invasive tool for the risk stratification of ccRCC patients. Radiogenomic analysis suggested that the activation of cell cycle-related pathways is the molecular mechanism responsible for the poor clinical outcome in the C1 subtype. The classifier model was validated using another cohort, which could be useful in clinical decision making. These findings may serve as the basis for further in-depth studies to assess the radiomic subtypes in a reliable manner for risk stratification and precision medicine.

Contributors

All authors read and approved the final version of the manuscript.
Research conception, design and manuscript revision:
Peng Lin, Yi-qun Lin, Rui-zhi Gao, Rong Wen, Hui-Qin, Yun He and Hong Yang
Data processing, statistical analysis, drafting of the manuscript:
Peng Lin, Yi-qun Lin and Rui-zhi Gao
Detection of multi-omics subtypes: Rong Wen, Hui-Qin, Yun He
Delineation region of interest (ROI) of the tumor contours: Yi-qun Lin

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 68 (6) (2018) 394–424.

[2] J.J. Hsieh, M.P. Purdue, S. Signoretti, et al., Renal cell carcinoma, Nat. Rev. Dis. Prim. 3 (2017) 17009.

[3] M. Gerlinger, A.J. Rowan, S. Horswell, et al., Intratumor heterogeneity and branched evolution revealed by multiregion sequencing, N. Engl. J. Med. 366 (10) (2012) 883–892.

[4] A. Sankin, A.A. Hakimi, N. Mikkilineni, et al., The impact of genetic heterogeneity on biomarker development in kidney cancer assessed by multiregional sampling, Cancer Med. 3 (6) (2014) 1485–1492.

[5] J.J. Hsieh, D. Chen, P.I. Wang, et al., Genomic biomarkers of a randomized trial comparing first-line everolimus and sunitinib in patients with metastatic renal cell carcinoma, Eur. Urol. 71 (3) (2017) 405–414.

[6] Cancer genome atlas research N. Comprehensive molecular characterization of clear cell renal cell carcinoma, Nature 499 (7456) (2013) 43–49.

[7] P. Msaouel, G.G. Malouf, X. Su, et al., Comprehensive molecular characterization identifies distinct genomic and immune hallmarks of renal medullary carcinoma, Cancer Cell 37 (5) (2020) 720–734 e13.

[8] A. Reustle, M. Di Marco, C. Meyerhoff, et al., Integrative -omics and HLA-ligan-domics analysis to identify novel drug targets for ccRCC immunotherapy, Genome Med. 12 (1) (2020) 32.

[9] P. Lambin, E. Rios-Velazquez, R. Leijenaar, et al., Radiomics: extracting more information from medical images using advanced feature analysis, Eur. J. Cancer 48 (4) (2012) 441–446.

[10] C.N. Andreassen, L.M. Schack, L.V. Laursen, J. Alsner, Radiogenomics - current status, challenges and future directions, Cancer Lett. 382 (1) (2016) 127–136.

[11] T.T. Liu, A.S. Achrol, L.A. Mitchell, et al., Magnetic resonance perfusion image features uncover an angiogenic subgroup of glioblastoma patients with poor survival and better response to antiangiogenic treatment, Neuro. Oncol. 19 (7) (2017) 997–1007.

[12] J. Wu, Y. Cui, X. Sun, et al., Unsupervised clustering of quantitative image phenotypes reveals breast cancer subtypes with distinct prognoses and molecular pathways, Clin. Cancer Res. 23 (13) (2017) 3334–3342.

[13] K. Clark, B. Vendt, K. Smith, et al., The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository, J. Digit. Imaging 26 (6) (2013) 1045–1057.

[14] O. Akin, P. Elnajjar, M. Heller, R. Jarosz, B.J. Erickson, S. Kirk, J. Filippini, Radiology data from the cancer genome atlas kidney renal clear cell carcinoma [TCGA-KIRC] collection, Cancer Imaging Arch. (2016).

[15] J. Liu, T. Lichtenberg, K.A. Hoadley, et al., An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics, Cell 173 (2) (2018) 400–416 e11.

[16] Y. Sato, T. Yoshizato, Y. Shiraishi, et al., Integrated molecular analysis of clear-cell renal cell carcinoma, Nat. Genet. 45 (8) (2013) 860–867.

[17] P.A. Yushkevich, J. Piven, H.C. Hazlett, et al., User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability, Neuroimage 31 (3) (2006) 1116–1128.

[18] T. Xu, T.D. Le, L. Liu, et al., CancerSubtypes: an R/Bioconductor package for molecular cancer subtype identification, validation and visualization, Bioinformatics 33 (19) (2017) 3131–3133.

[19] A. Colaprico, T.C. Silva, C. Olsen, et al., TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data, Nucleic. Acids. Res. 44 (8) (2016) e71.

[20] A. Mayakonda, D.C. Lin, Y. Assenov, C. Plass, H.P. Koeffler, Maftools: efficient and comprehensive analysis of somatic variants in cancer, Genome Res. 28 (11) (2018) 1747–1756.

[21] C.H. Mermel, S.E. Schumacher, B. Hill, M.L. Meyerson, R. Beroukhim, G. Getz, GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers, Genome Biol. 12 (4) (2011) R41.

[22] K.A. Hoadley, C. Yau, T. Hinoue, et al., Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer, CellCell 173 (2) (2018) 291–304 e6.

[23] A. Subramanian, P. Tamayo, V.K. Mootha, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. U. S. A. 102 (43) (2005) 15545–15550.

[24] J. Friedman, T. Hastie, R. Tibshirani, Regularization paths for generalized linear models via coordinate descent, J. Stat. Softw. 33 (1) (2010) 1–22.

[25] L. Wang, S.R. Williamson, M. Wang, et al., Molecular subtyping of metastatic renal cell carcinoma: implications for targeted therapy, Mol. Cancer 13 (2014) 39.

[26] D.J. Clark, S.M. Dhanasekaran, F. Petralia, et al., Integrated proteogenomic characterization of clear cell renal cell carcinoma, Cell 179 (4) (2019) 964–983 e31.

[27] F. Alessandrino, A.B. Shinagare, D. Bosse, T.K. Choueiri, K.M. Krajewski, Radiogenomics in renal cell carcinoma, Abdom. Radiol. (NY) 44 (6) (2019) 1990–1998.

[28] Z. Feng, L. Zhang, Z. Qi, Q. Shen, Z. Hu, F. Chen, Identifying BAP1 mutations in clear-cell renal cell carcinoma by CT radiomics: preliminary findings, Front. Oncol. 10 (2020) 279.

[29] C.A. Karlo, P.L. Di Paolo, J. Chaim, et al., Radiogenomics of clear cell renal cell carcinoma: associations between CT imaging features and mutations, Radiology 270 (2) (2014) 464–471.

[30] R. Elias, Q. Zhang, J. Brugarolas, The von Hippel-Lindau tumor suppressor gene: implications and therapeutic opportunities, Cancer J. 26 (5) (2020) 390–398.

[31] P.E. Clark, The role of VHL in clear-cell renal cell carcinoma and its relation to targeted therapy, Kidney Int. 76 (9) (2009) 939–945.

[32] H.S. Kim, J.H. Kim, H.J Jang, VHLClinicopathologic Significance of Gene Alteration in Clear-Cell Renal Cell Carcinoma: An Updated Meta-Analysis and Review, Int J Mol Sci 19 (9) (2018) 180–186.

[33] O.F. Femi, Genetic alterations and PIK3CA gene mutations and amplifications analysis in cervical cancer by racial groups in the United States, Int. J. Health Sci. (Qassim) 12 (1) (2018) 28–32.

[34] Y. Xu, H. Luo, Q. Hu, H. Zhu, Identification of potential driver genes based on multi-genomic data in cervical cancer, Front. Genet. 12 (2021) 598304.

[35] L. Zhang, X. Han, Y. Shi, Association of MUC16 mutation with response to immune checkpoint inhibitors in solid tumors, JAMA Netw. Open 3 (8) (2020) e2013201.

[36] N. Jonckheere, I. Van Seuningen, Integrative analysis of the cancer genome atlas and cancer cell lines encyclopedia large-scale genomic databases: MUC4/MUC16/MUC20 signature is associated with poor survival in human carcinomas, J. Transl. Med. 16 (1) (2018) 259.

[37] X. Li, B. Pasche, W. Zhang, K. Chen, Association of MUC16 mutation with tumor mutation load and outcomes in patients with gastric cancer, JAMA Oncol. 4 (12) (2018) 1691–1698.

[38] C. Zhang, Y. Zheng, X. Li, X. Hu, F. Qi, J. Luo, Genome-wide mutation profiling and related risk signature for prognosis of papillary renal cell carcinoma, Ann. Transl. Med. 7 (18) (2019) 427.

[39] C.J. Ricketts, V.K. Hill, W.M. Linehan, Tumor-specific hypermethylation of epigenetic biomarkers, including SFRP1, predicts for poorer survival in patients from the TCGA kidney renal clear cell carcinoma (KIRC) project, PLoS ONE 9 (1) (2014) e85621.

[40] M.R. Morris, C.J. Ricketts, D. Gentle, et al., Genome-wide methylation analysis identifies epigenetically inactivated candidate tumour suppressor genes in renal cell carcinoma, Oncogene 30 (12) (2011) 1390–1401.

[41] B. Jiang, W. Chen, H. Qin, et al., TOX3 inhibits cancer cell migration and invasion via transcriptional regulation of SNAI1 and SNAI2 in clear cell renal cell carcinoma, Cancer Lett. 449 (2019) 76–86.

[42] W. Fangning, M. Chunguang, Z. Hailiang, et al., Identification and validation of soluble carrier family expression signature for predicting poor outcome of renal cell carcinoma, J. Cancer 8 (11) (2017) 2010–2017.