Radiomic and gEnomic approaches for the enhanced DIagnosis of clear cell REnal Cancer (REDIRECt): a translational pilot methodological study

Francesco Cianflone[1], Dejan Lazarevic[2], Anna Palmisano[3], Giuseppe Fallara[1], Alessandro Larcher[1], Massimo Freschi[4], Giacomo Dell'Antonio[4], Giulia Maria Scotti[2], Marco J. Morelli[2], Anna Maria Ferrara[1], Francesco Trevisani[1], Alessandra Cinque[1], Antonio Esposito[3], Alberto Briganti[1], Carlo Tacchetti[5], Claudio Doglioni[4], Alessandro del Maschio[3], Francesco de Cobelli[3], Roberto Bertini[1], Andrea Salonia[1], Francesco Montorsi[1], Giovanni Tonon[2], Umberto Capitanio[1]

[1]Unit of Urology, Division of Experimental Oncology, Urological Research Institute, IRCCS San Raffaele Scientific Institute, Vita-Salute San Raffaele University, Milan, Italy; [2]Center for Omics Scences, IRCCS San Raffaele Scientific Institute, Milan, Italy; [3]Department of Radiology, Experimental Imaging Centre, IRCCS San Raffaele Scientific Institute, Vita-Salute San Raffaele University, Milan, Italy; [4]Department of Pathology, IRCCS San Raffaele Scientific Institute, Milan, Italy; [5]Unit of Cancer Imaging, Experimental Imaging Centre, IRCCS San Raffaele Scientific Institute, Vita-Salute San Raffaele University, Milan, Italy

Contributions: (I) Conception and design: U Capitanio, G Tonon, D Lazarevic; (II) Administrative support: AM Ferrara; (III) Provision of study materials or patients: A Larcher, D Lazarevic, A Palmisano, M Freschi, G Dell'Antonio, GM Scotti; (IV) Collection and assembly of data: A Larcher, D Lazarevic, A Palmisano, M Freschi, G Dell'Antonio, GM Scotti; (V) Data analysis and interpretation: F Cianflone, A Larcher, D Lazarevic, A Palmisano, M Freschi, G Dell'Antonio, GM Scotti, U Capitanio, G Tonon; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Umberto Capitanio. Division of Oncology/Unit of Urology URI, IRCCS Ospedale San Raffaele, Via Olgettina 60, Milan 20132, MI, Italy. Email: capitanio.umberto@hsr.it.

Background: The combination of radiomic and transcriptomic approaches for patients diagnosed with small clear-cell renal cell carcinoma (ccRCC) might improve decision making. In this pilot and methodological study, we investigate whether imaging features obtained from computed tomography (CT) may correlate with gene expression patterns in ccRCC patients.

Methods: Samples from 6 patients who underwent partial nephrectomy for unilateral non-metastatic ccRCC were included in this pilot cohort. Transcriptomic analysis was conducted through RNA-sequencing on tumor samples, while radiologic features were obtained from pre-operative 4-phase contrast-enhanced CT. To evaluate the heterogeneity of the transcriptome, after a 1,000 re-sampling via bootstrapping, a first Principal Component Analyses (PCA) were fitted with all transcripts and a second ones with transcripts deriving from a list of 369 genes known to be associated with ccRCC from The Cancer Genome Atlas (TCGA). Significant pathways in each Principal Components for the 50 genes with the highest loadings absolute values were assessed with pathways enrichment analysis. In addition, Pearson's correlation coefficients among radiomic features themselves and between radiomic features and transcripts expression values were computed.

Results: The transcriptomes of the analysed samples showed a high grade of heterogeneity. However, we found four radiogenomic patterns, in which the correlation between radiomic features and transcripts were statistically significant.

Conclusions: We showed that radiogenomic approach is feasible, however its clinical meaning should be further investigated.

Keywords: Kidney cancer; renal cancer; genomics; radiomics; transcriptomics

Introduction

Clear-cell renal cell carcinoma (ccRCC) is the most common RCC subtype, accounting for almost 70% of all RCC diagnoses (1). Incidence predominates in men, being male-to-female ratio 1.5:1.0, and peaks at age 60–70 years. RCC is the sixth most common cancer for male and eighth for female in US (2). Estimated five-year relative survival is almost 75%, but it is highly heterogenic depending on stage, grade, histology and several cancer and patient features (2). Employing a combined radiomic and genomic approach might improve decision making, as it is currently happening for breast cancer, lung cancer, head and neck cancer (3-5).

In the past years, the molecular characterization of ccRCC has brought to the discover of new genetic pathways, methylation patterns and transcripts (6-9), suggesting possible targets for new therapies which then became the cornerstone of metastatic ccRCC management (10,11). In addition, gene expression clusters have been found to correlate with patients' long-term oncologic outcomes (12,13). However, genetic analyses are rarely used in everyday practice (14).

Radiomics is a new field based on quantitative analysis of radiological images, either computed tomography (CT), magnetic resonance imaging (MRI) or positron emission tomography (PET) (15,16). It has been shown that radiomic patterns may predict patient's long-term oncologic outcomes in different tumours, such as non-small cell lung cancer and glioblastoma (17,18). However, few evidences exist for radiomics in ccRCC and only a few combined radiomic with genomic data (19,20).

The combination of radiomics and genomics might be useful in several steps of the RCC workup: for example, in diagnosis, to define the nature of small renal masses; in active surveillance, to drive the choice toward surveillance or active treatment; in cancer staging, to better stratify aggressiveness of the tumour; in medical treatment choice, to select optimal candidates for adjuvant/first/second line immune/chemotherapies.

Therefore, the aim of this pilot study is to establish a method for radiogenomic characterisation of small ccRCC masses, focusing on the transcriptomic underpinnings of radiomic features. We present the following article in accordance with the MDAR checklist (available at https://tau.amegroups.com/article/view/10.21037/tau-21-713/rc).

Methods

Study population

Six patients diagnosed with ccRCC were randomly selected from a prospective maintained database, according to the following inclusion criteria: age 18–85 years, single monolateral organ-confined non-metastatic renal mass (pT1a-b stage, 5 cm maximum diameter), clear cell histology, no previous diagnosis of renal cancer, complete clinical history, blood and tumour samples collected at the time of surgery and available in our biobank, and availability of a preoperative contrast-enhanced 4-phase multidetector contrast-enhanced computed tomography (MDCT) performed at our Institution. All patients had been submitted to partial nephrectomy.

The study was approved by the institutional ethical board of San Raffaele Hospital in Milan (protocollo No. URI001-2010 RENE - versione 29/08/2007). Informed consent was collected from each patient. The study was conducted in accordance with the principles outlined in the Declaration of Helsinki (as revised in 2013).

Clinical and pathological evaluation

Complete anamnestic and preoperative data were recorded, including age, patient weight and height, body mass index (BMI) estimated glomerular filtration rate (eGFR) and clinical tumour size, defined as the greatest tumour diameter in centimetres on pre-operative imaging. Tumor was staged according to TNM 8th edition, 2017. eGFR was calculated with the Chronic Kidney Disease Epidemiology Collaboration formula for younger patients (<70 years) and with the Berlin Initiative Study formula for older patients (≥70 years). A dedicated experienced genitourinary pathologist re-examined the surgical specimens to standardize histological evaluation (Appendix 1).

Computed tomography acquisition protocol and images analysis

All patients underwent CT scan with a 64-slice multidetector CT scanner (Philips Brilliance 64, Philips, Best, The Netherlands) with a four-phase protocol: unenhanced (UP), corticomedullary (CMP), nephrographic (NP) and excretory phases (EP), i.e., scans respectively before and 30 s, 90 s and 5 min after contrast injection.

An experienced radiologist selected the axial frame showing the largest tumour diameter for each patient's CT scan. Tumours were manually contoured using a dedicated software (Intellispace portal v.8, Philips, Best, The Netherlands) to obtain the region of interest (ROI). The following quantitative features were measured in the ROI: mean, maximum, minimum and standard deviations of the attenuation [measured in Hounsfield unit (HU)]. The following semiquantitative parameters were measured in the ROI: tumour volume, percentage of exophytic growth, tumour-to-psoas ratio on UP scan, tumour-to-kidney ratio on UP, CMP, NP and EP scan, early and late tumour enhancement (defined as the difference between mean attenuation on CMP and mean attenuation on UP scan and the difference between mean attenuation on NP and mean attenuation on UP scan, respectively). The following qualitative features were obtained in the ROI: presence/absence of calcifications, central scar and pseudocapsule, tumour attenuation (hyperdense, isodense, or hypodense relative to adjacent parenchyma), composition (solid or cystic), necrosis, homogeneity [homogeneous or heterogeneous (uniform or mixed attenuation)] (Figure S1).

RNA extraction

Total RNA was extracted from snap frozen ccRCC tissue specimens using TRIzol reagent (Invitrogen), according to manufacturer's instructions. Purity of recovered RNA were determined with a NanoDrop spectrophotometer (ND-1000, NanoDrop Technologies, ThermoFisher). All 6 RNA samples had ~2.0 A260/A280 ratio and 1.8-2.2 A260/A230 ratio. Integrity and concentration of isolated total RNA was assessed with the RNA 6000 Nano LabChip kit using the Agilent 2100 Bioanalyzer (Agilent Technologies). All 6 RNA samples had RIN values >8.

Transcriptomic data analysis

RNA-sequencing was performed by use of Quant Seq 3'mRNA-Seq library prep kit for Illumina® from the previously collected renal tumour specimen stored in our biobank. The prepared libraries were sequenced on Illumina® Next Seq 500 platform, and reads were generated towards poly(A)-tail. Quality control was performed by use of Multi QC platform, and the Fast QC toolkit (Figures S2-S7). Sequences were then mapped to the human genome using the STAR aligner, v. 2.5.3, and annotated according to Gencode basic annotations, version 25. Data generated in this study were deposited in NCBI's Gene Expression Omnibus accessible through GEO Series (accession number GSE133460, link: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE133460).

Statistical analyses

Gene expression for each tumour sample was quantified with logarithmic values of Reads Per Kilobase of transcript per Million mapped reads (RPKM).

To assess the heterogeneity of the transcripts among the six samples, several versions of Principal Component Analysis (PCA) were fitted. First, a PCA with all transcripts was performed. Then, a list of 369 out of 406 expressed genes known to be associated with ccRCC from The Cancer Genome Atlas (TCGA) was obtained after excluding 37 genes with a constant expression in our samples, in order to focus on genes really involved in ccRCC and avoid possible confounding variables (Appendix 1) (6). A second PCA with this restricted list was performed. For PCA analyses zero-centred RPKM values were used. Given the small sample size, to account for PCA instability, a bootstrap with 1,000 re-sampling was performed. For each re-sampling, 50 genes with the highest loadings absolute values were extracted. In order to identify pathways and gene ontologies, Enrichment analysis was performed by use of Enrichr (Ma'ayan Laboratory, Mount Sinai Center for Bioinformatics, New York, NY, USA) (21).

Pearson's correlation coefficients were used to assess correlation among radiomic features and between the radiomic features and transcript expression. Adjusted P values were computed by use of Benjamini-Hochberg false discovery rate. Correlation was considered adequate if $\rho<-0.85$ or $\rho>0.85$.

In order to depict correlations between radiomic features and transcripts, a heatmap was used. Moreover, to assess the relationship between radiomic features and transcripts, a hierarchical clustering dendrogram was plotted. Statistical significance of the radiogenomic correlation patterns

Table 1 Patient's characteristics at surgery

Variable	Median [IQR]/n [%]
Age, years	
Median	68
IQR	55–76
Gender	
Male	5 [83]
Female	1 [17]
BMI	
Median	26.2
IQR	22.1–30.1
CCI	
0	3 [50]
1–2	2 [33]
≥3	1 [17]
Preoperative haemoglobin, mg/dL	
Median	14.5
IQR	13.1–15.6
Preoperative eGFR, mL/min/1.73 m^2	
Median	99.5
IQR	84.6–134.2
Preoperative hypertension	
No	3 [50]
Yes	3 [50]
Smoking status	
No smoking history	3 [50]
Active smoker	0 [0]
Former smoker	3 [50]
Clinical tumor size, cm	
Median	4.1
IQR	3.3–4.4
Year of surgery	
Median	2014
IQR	2012–2015
Surgical approach	
Laparotomic	5 [83]
Laparoscopic	0 [0]

Table 1 (*continued*)

Table 1 (*continued*)

Variable	Median [IQR]/n [%]
Robotic	1 [17]
Affected side	
Left	3 [50]
Right	3 [50]

BMI, body mass index (kg/m^2); CCI, Charlson Comorbidity Index; eGFR, estimated Glomerular Filtration Rates, computed with the Chronic Kidney Disease Epidemiology Collaboration formula for younger patients (<70 years),

was assessed with Mann-Whitney U test, for dichotomic comparison, or with Kruskal-Wallis tests, for multiple comparison.

GraphPad Prism (GraphPad Software La Jolla, CA, USA) and RStudio graphical interface v.0.98 for R software environment v.3.0.2 (http://www.r-project.org) were used for analysis.

Results

Population and tumor characteristics

Median (IQR) age was 68 [55–76] years old (*Table 1*). Five out of 6 patients were male. The median tumour size was 4.1 (3.3–4.4) cm. At histopathological evaluation of the specimens, Fuhrman Grade was 2 in 5 out of 6 specimens while it was 3 in the remaining one; necrosis and lymphoid infiltrates or aggregates were found in half the specimens. Radiologic evidence of intra-tumoral calcifications was found in one case. No central scars were found. The other radiological features are described in Table S1.

RNA sequencing quality control

Quality check for RNA-sequencing was performed. Per base sequence quality of the reads was adequate (Figures S2,S3). The alignment quality was generally good, with almost 80% of reads uniquely mapped on the human genome (Figures S4,S5).

PCA

In the first PCA, using all expressed genes, no clustering was observed between tumour samples (*Figure 1*). When a PCA was ran only against genes significantly associated with
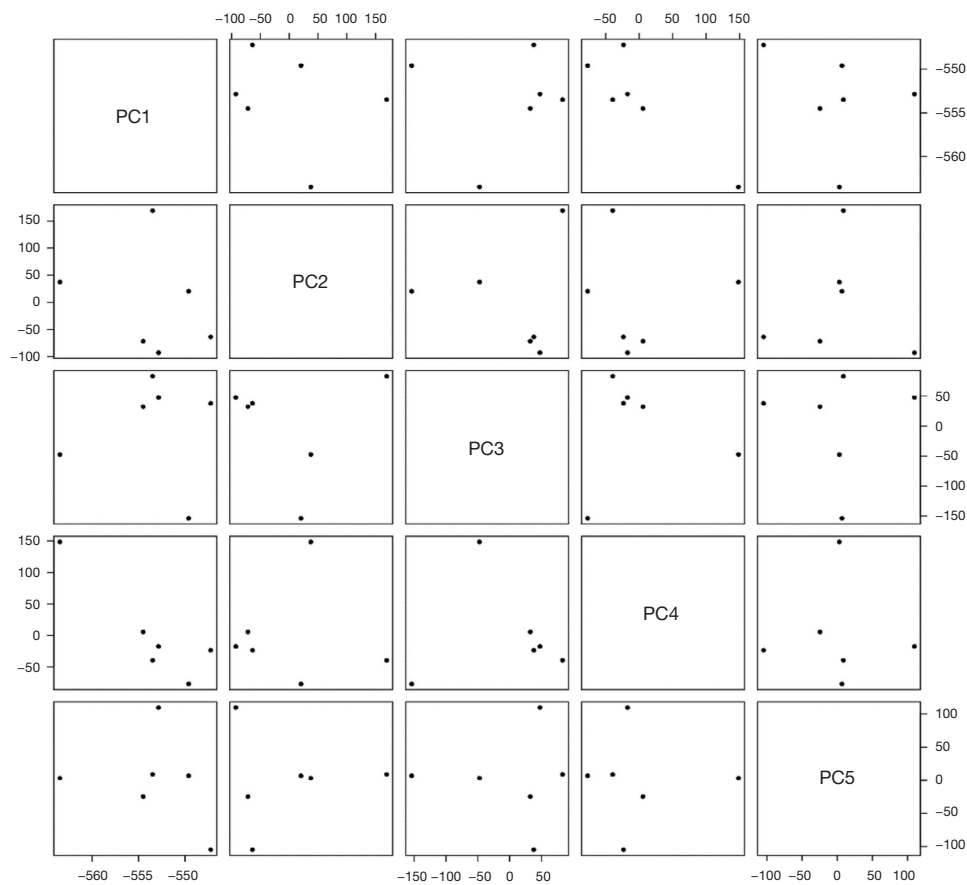
Figure 1 PCA scoreplots for the first five PC, using all expressed genes. Analyses were performed using zero-centred RPKM logarithmic values. The six samples did not cluster in any of the scoreplots. The variances from PC1 to PC5 were 92%, 2.4%, 1.9%, 1.5%, and 1.2%, respectively. PC, principal component; PCA, principal component analysis; RPKM, Reads Per Kilobase of transcript per Million mapped reads.

ccRCC, despite the persistence of high heterogeneity, three tumour samples clustered in PC3 versus PC5 plot, and two in PC1 versus PC5 (*Figure 2*).

Pathways enrichment analysis

Enrichment analysis performed on the top 50 genes with highest loading absolute values in the PCA fitted with transcripts associated with ccRCC showed that the genes driving the clustering in PC3 versus PC5 plot were significantly enriched in pathways related to ECM-receptor interaction pathways, PI3K-Akt signalling pathway, microtubules pathways and calcium homeostasis.

Radiomic features correlation

Among radiomic features, correlation between (I) mean

tumour attenuation in CMP and early enhancement (Pearson's correlation ρ=0.99, adjusted P=0.003), (II) mean tumour attenuation in NP and late enhancement (ρ=0.99, P=0.003), and (III) lowest tumour attenuation in CMP and NP (ρ=0.98, P=0.03) was found to be statistically significant.

Radiogenomic correlation

Pearson correlation coefficients were assessed between radiomic features and RPKM values of the ccRCC-associated gene transcripts (data not shown). Using heatmap (*Figure 3*) and dendrogram (*Figure 4*), it was possible to create 4 radiogenomic correlation patterns between expressed genes and radiomic features. In the first pattern, named Radio Genomic Pattern 1 (RGP1), global tumour density and tumour vascular behaviour (i.e., mean tumour attenuation in UP, CMP, NP and early and late tumour
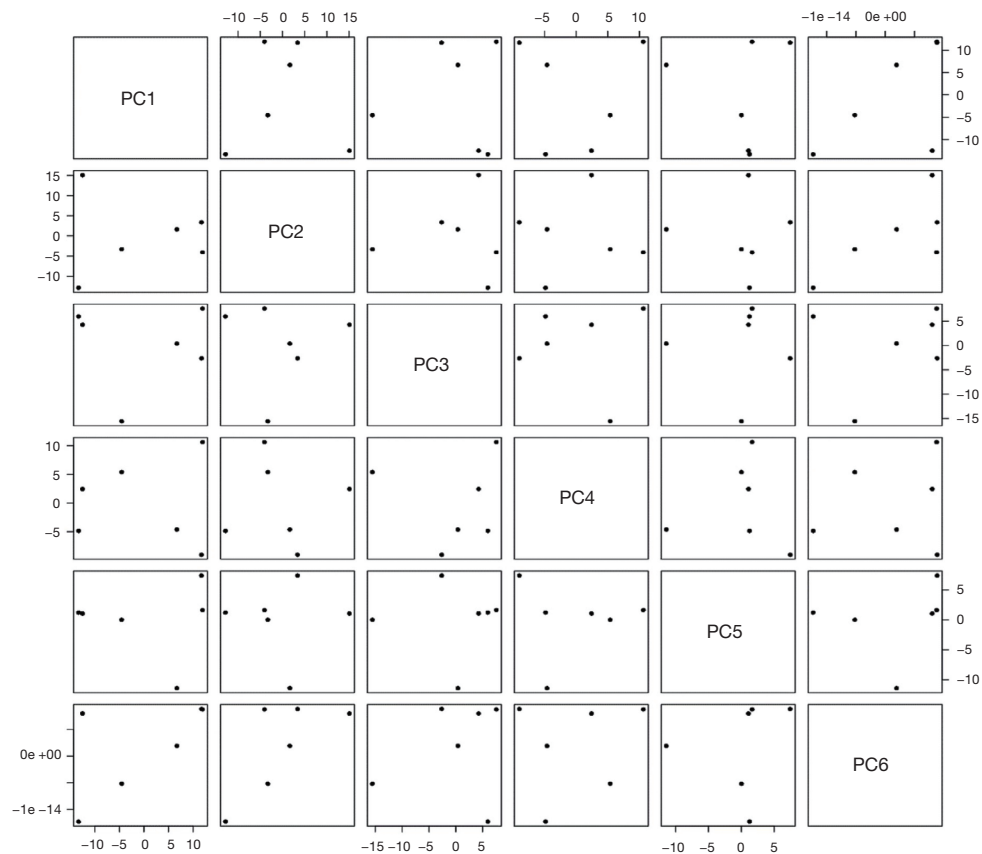
Figure 2 PCA scoreplots for the first six PC restricting the analysis to a list of 369 genes associated with clear cell renal cell carcinoma from TCGA RNAseq and transcriptomic analysis. The list of 369 genes used in the Appendix 1. Analyses were perfomed using zero-centred RPKM logarithmic values. Three samples clustered in PC3 versus PC5 plot, and two tumour samples clustered in PC1 versus PC5 plot. The variances from PC1 to PC6 were 35.1%, 22.4%, 18.6%, 14.1%, 9.9%, and <0.1%, respectively. PC, principal component; PCA, principal component analysis; TCGA, The Cancer Genome Atlas; RPKM, Reads Per Kilobase of transcript per Million mapped reads.

enhancement) were correlated positively with the expression of *VHL*, *COL5A3* and *ANKRD50*, and negatively with *PCDH7*, *PAPOLG* and *PTEN* genes. In the second pattern, named Radio Genomic Pattern 2 (RGP2), hypovascular fat and necrotic components (i.e., lowest tumour attenuation in CMP and NP) were correlated positively with the expression of *PLEC*, *TNR*, *CUL9*, and *UGGT1*, and negatively with *RFC1*, *ESPL1*, *SACS*, *ACVR1B*, *DNAH5*, *KIAA0368*, *ANK3*, *VWA8*, *PHF20*, *USF3*, *DNAH7*, *ARHGAP5*, *SMC3*, *ZMYM1* and *CMYA5* genes. In the third pattern, named Radio Genomic Pattern 3 (RGP3), mean vascularization of kidney and tumour mass (i.e., mean attenuation of the renal cortex in the NP and maximum tumour attenuation in the NP) were correlated positively with the expression *ATM*, and negatively with *PAPOLG* gene. In the fourth pattern, named Radio Genomic

Pattern 4 (RGP4), maximum tumour hypervascular areas (i.e., maximum tumour attenuation in the CMP) and kidney-to-tumour attenuation ratio in the CMP were correlated positively with the expression of *CHD9*, *ALMS1*, *SMARCA4*, *GPATCH8*, *ITSN2*, *CDH8*, *KIF21B*, *NCOR1*, *KMT2C* and *SPAG17*, and negatively with *SVEP1* and *LAMA2* genes.

Discussion

Our pilot methodological study was designed to assess the feasibility of a combined radiogenomic approach in the work-up of organ-confined non-metastatic monolateral ccRCC, based on RNA sequencing and radiomic features.

The application of radiomics in RCC is still at a very early stage, with few reports published and a huge
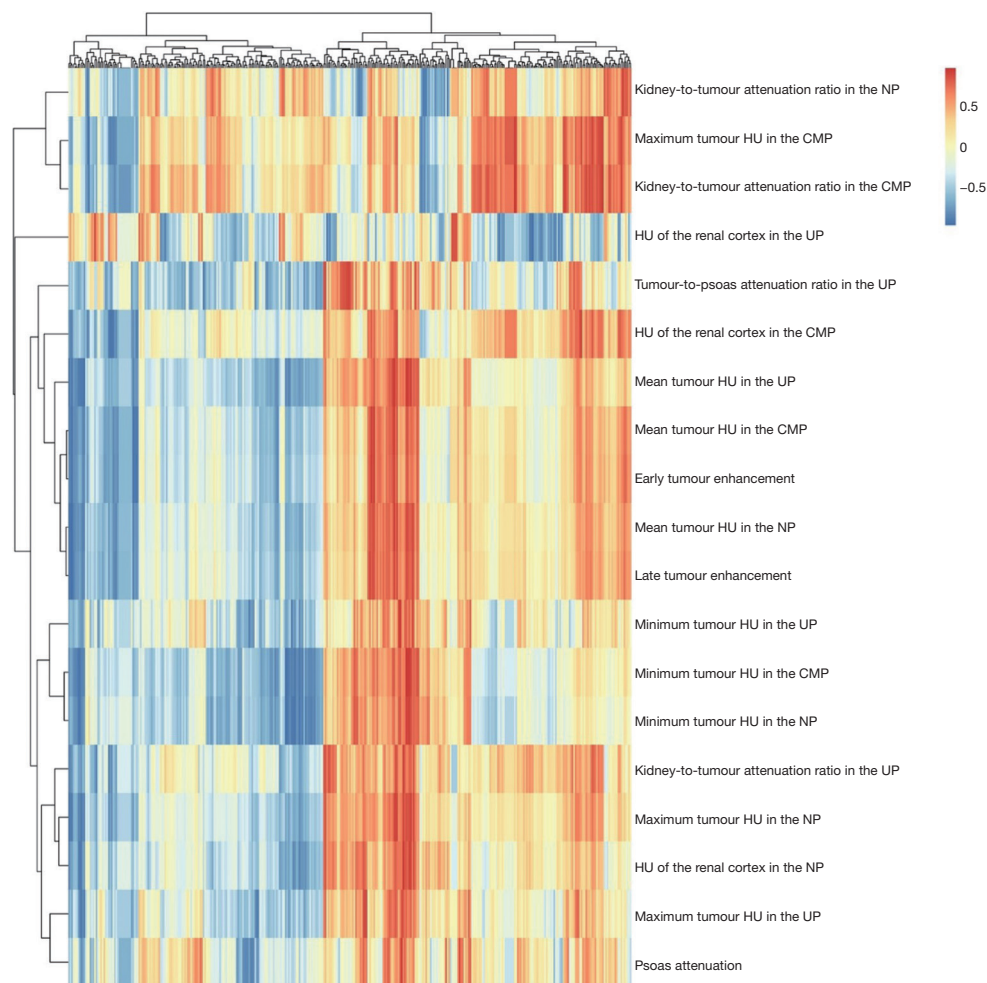
Figure 3 Heatmap showing the correlation between expressed genes and 19 radiomic features. Correlation strength is shown from red to blue, ranging from complete positive to complete negative correlation on the base of Pearson's correlation coefficients. HU, Hounsfield unit; UP, unenhanced phase; CMP, corticomedullary phase; NP, nephrographic phase.

heterogeneity of methodologies applied (19,20). A potential role of radiomics in prediction of RCC subtype, grade and long-term oncologic outcomes has been suggested. Concerns have been raised on the reproducibility of this approach because of use of unstandardized CT-scan protocols in different centres. The main criticisms to radiomics is the lack of a straightforward interpretation of the association between imaging and biological features (15,16). However, studies focusing on the relationship between genomic and radiomic features in different oncological setting are gaining momentum, and recently a review on RCC radiogenomics has been published (22).

In the current pilot study, we investigated the association between RCC transcriptomic and qualitative, semiquantitative and quantitative radiomic features. Initial PCA showed high transcriptomes heterogeneity, as no samples clustered in any of the PCA matrixes. A second PCA was run using a restricted list of 369 ccRCC-associated genes (Appendix 1). In this PCA, heterogeneity was still high, but three sample clustered in the PC3 versus PC5 plot. PC3 and PC5 explained a not negligible portion of the transcriptomic variance (18.6% and 9.9%, respectively).

Enrichment analysis performed on the top 50 genes with highest loading absolute values in the second PCA revealed different pathways and gene ontologies: the three samples clustering in PC3 versus PC5 plot had similar expression of the genes regulating the ECM-receptor interaction pathways, PI3K-Akt signalling pathway, microtubules
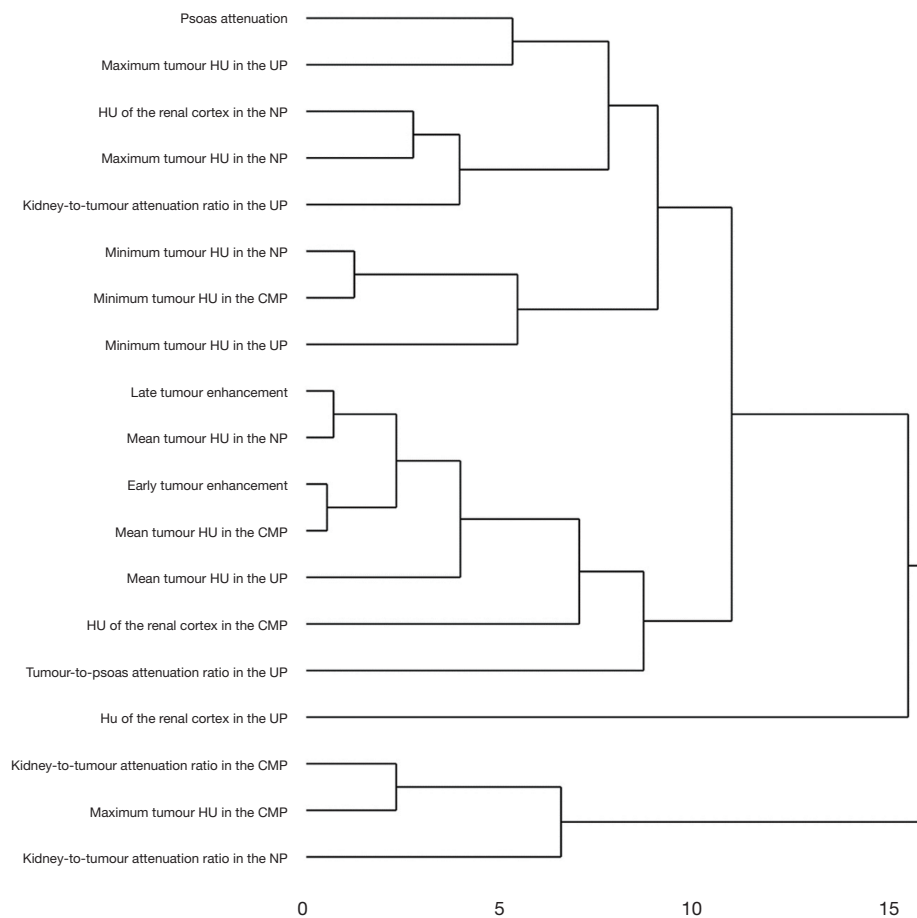
Transl Androl Urol 2022;11(2):149-158 | https://dx.doi.org/10.21037/tau-21-713

Figure 4 Dendrogram showing hierarchical clustering. Clustering was performed on the base of correlation coefficients between gene expression and radiomic feature. HU, Hounsfield unit; UP, unenhanced phase; CMP, corticomedullary phase; NP, nephrographic phase.

function and calcium homeostasis.

In addition, despite the low numerosity of our study population, we were able to finds correlations between expressed genes and some radiologic features. The radiologic features included in RGP1 showed a strong positive correlation with the expression of *VHL* gene, a key gene in RCC oncogenesis, and *ANKRD50* gene, involved in endosome-to-plasma membrane trafficking and recycling of SNX27-retromer-dependent cargo proteins (23). In RPG1 radiomic features were also positively associated with the expression of *COL5A3* gene, an ubiquitous fibrillar collagen. On the contrary, a strong negative correlation was found between radiomic features and the expression of *PCDH7* gene, whose product is an integral membrane protein involved in cell-cell recognition and adhesion (24), with the expression of *PAPOLG* gene, which encodes a member of

the poly(A) polymerase family (25), and with the expression of *PTEN* gene, a tumor suppressor gene whose product antagonizes the PI3K-AKT/PKB signalling pathway and modulates cell cycle progression and cell survival. Of note, RGP1 radiomics features included lesion attenuation and early and late tumour enhancement, which are all radiologic vascular features. VHL and PTEN are frequently mutated in ccRCC, and are usually associated with aggressive tumoral features and neoangiogenesis (26). Expressed genes in RGP2 included cytoskeleton regulatory genes involved in microtubules and ECM function and architecture, protein misfolding pathways and chromosome clustering during mitosis (27-29). Of note, RGP2 radiomic features included hypovascular areas, as for fat, cystic or necrosis areas, where some of the aforementioned genes might play a role (27). In RGP3 the mean contrast enhancement of renal cortex and

of tumor clustered together, suggesting a relation between renal cortical and tumoral vascular features, which might be driven by similar expressed genes. Finally, in RPG4 tumor hypervascular features were related to transcripts for nucleic acid binding, chromatin organization, microtubule motor activity and calcium ion binding (30,31).

This study is not devoid of limitation. A major limitation is the small size of patient's cohort and their strict inclusion criteria. In addition, we were not able to investigate the impact of radiogenomic patterns on oncologic outcomes. Likely, further radiogenomic analyses might yield new evidences on long-term oncologic outcomes according to the different radiogenomic patterns, and the inclusion of radiogenomic data in predictive-prognostic tools might be of key clinical value.

However, with the current study, we demonstrated that radiogenomic approach in ccRCC is methodological feasible.

Conclusions

In this pilot methodological study, in patients with small low-grade organ-confined non-metastatic ccRCC, the transcriptomes of tumor samples were characterized by high heterogeneity. Gene expression was associated with different radiologic features and four main patterns were found, each including different radiomic and transcripts correlates, likely linked by biologic underpinnings. However, further investigations on radiogenomic approaches in ccRCC work-up are needed.

Acknowledgments

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at https://tau.amegroups.com/article/view/10.21037/tau-21-713/rc

Data Sharing Statement: Available at https://tau.amegroups.com/article/view/10.21037/tau-21-713/dss

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tau.amegroups.com/article/view/10.21037/tau-21-713/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was approved by the institutional ethical board of San Raffaele Hospital in Milan (protocol No. URI001-2010 RENE - version 29/08/2007). Informed consent was collected from each patient. The study was conducted in accordance with the principles outlined in the Declaration of Helsinki (as revised in 2013).

References

1. Capitanio U, Montorsi F. Renal cancer. Lancet 2016;387:894-906.
2. Siegel RL, Miller KD, Fuchs HE, et al. Cancer Statistics, 2021. CA Cancer J Clin 2021;71:7-33.
3. Zwirner K, Hilke FJ, Demidov G, et al. Radiogenomics in head and neck cancer: correlation of radiomic heterogeneity and somatic mutations in TP53, FAT1 and KMT2D. Strahlenther Onkol 2019;195:771-9.
4. Zanfardino M, Pane K, Mirabelli P, et al. TCGA-TCIA Impact on Radiogenomics Cancer Research: A Systematic Review. Int J Mol Sci 2019;20:6033.
5. Zhu Z, Albadawy E, Saha A, et al. Deep learning for identifying radiogenomic associations in breast cancer. Comput Biol Med 2019;109:85-90.
6. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature 2013;499:43-9.
7. Guo G, Gui Y, Gao S, et al. Frequent mutations of genes encoding ubiquitin-mediated proteolysis pathway components in clear cell renal cell carcinoma. Nat Genet 2011;44:17-9.
8. Dalgliesh GL, Furge K, Greenman C, et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. Nature 2010;463:360-3.
9. Brugarolas J. Molecular genetics of clear-cell renal cell

carcinoma. J Clin Oncol 2014;32:1968-76.

10. Motzer RJ, Tannir NM, McDermott DF, et al. Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma. N Engl J Med 2018;378:1277-90.

11. Hutson TE, Escudier B, Esteban E, et al. Randomized phase III trial of temsirolimus versus sorafenib as second-line therapy after sunitinib in patients with metastatic renal cell carcinoma. J Clin Oncol 2014;32:760-7.

12. Wei JH, Haddad A, Wu KJ, et al. A CpG-methylation-based assay to predict survival in clear cell renal cell carcinoma. Nat Commun 2015;6:8699.

13. Kapur P, Peña-Llopis S, Christie A, et al. Effects on survival of BAP1 and PBRM1 mutations in sporadic clear-cell renal-cell carcinoma: a retrospective analysis with independent validation. Lancet Oncol 2013;14:159-67.

14. Ljungberg B, Albiges L, Abu-Ghanem Y, et al. European Association of Urology Guidelines on Renal Cell Carcinoma: The 2019 Update. Eur Urol 2019;75:799-810.

15. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749-62.

16. Avanzo M, Stancanello J, El Naqa I. Beyond imaging: The promise of radiomics. Phys Med 2017;38:122-39.

17. Kickingereder P, Neuberger U, Bonekamp D, et al. Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma. Neuro Oncol 2018;20:848-57.

18. Zhang Y, Oikonomou A, Wong A, et al. Radiomics-based Prognosis Analysis for Non-Small Cell Lung Cancer. Sci Rep 2017;7:46349.

19. Ding J, Xing Z, Jiang Z, et al. CT-based radiomic model predicts high grade of clear cell renal cell carcinoma. Eur J Radiol 2018;103:51-6.

20. Yu H, Scalera J, Khalid M, et al. Texture analysis as a radiomic marker for differentiating renal tumors. Abdom Radiol (NY) 2017;42:2470-8.

21. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016;44:W90-7.

22. Alessandrino F, Shinagare AB, Bossé D, et al. Radiogenomics in renal cell carcinoma. Abdom Radiol (NY) 2019;44:1990-8.

23. Kvainickas A, Orgaz AJ, Nägele H, et al. Retromer- and WASH-dependent sorting of nutrient transporters requires a multivalent interaction network with ANKRD50. J Cell Sci 2017;130:382-95.

24. Zhou X, Updegraff BL, Guo Y, et al. PROTOCADHERIN 7 Acts through SET and PP2A to Potentiate MAPK Signaling by EGFR and KRAS during Lung Tumorigenesis. Cancer Res 2017;77:187-97.

25. Topalian SL, Kaneko S, Gonzales MI, et al. Identification and functional characterization of neo-poly(A) polymerase, an RNA processing enzyme overexpressed in human tumors. Mol Cell Biol 2001;21:5614-23.

26. Jonasch E, Walker CL, Rathmell WK. Clear cell renal cell carcinoma ontogeny and mechanisms of lethality. Nat Rev Nephrol 2021;17:245-61.

27. Anlar B, Gunel-Ozcan A. Tenascin-R: role in the central nervous system. Int J Biochem Cell Biol 2012;44:1385-9.

28. Lopez J, Tait SW. Killing the Killer: PARC/CUL9 promotes cell survival by destroying cytochrome C. Sci Signal 2014;7:pe17.

29. Daikoku S, Seko A, Son SH, et al. The relationship between glycan structures and expression levels of an endoplasmic reticulum-resident glycoprotein, UDP-glucose: Glycoprotein glucosyltransferase 1. Biochem Biophys Res Commun 2015;462:58-63.

30. van Riel WE, Rai A, Bianchi S, et al. Kinesin-4 KIF21B is a potent microtubule pausing factor. Elife 2017;6:24746.

31. Martínez-Iglesias O, Alonso-Merino E, Aranda A. Tumor suppressive actions of the nuclear receptor corepressor 1. Pharmacol Res 2016;108:75-9.

Supplementary methods

Patients' characteristics

Complete anamnestic data were recorded for all the patients, including age, patient weight and height, body mass index (BMI), previous abdominal operations, smoking status, Eastern Cooperative Oncology Group (ECOG) performance status, American Society of Anaesthesiologists (ASA) score, Charlson Comorbidity Index (CCI), history of hypertension and its treatment, history of acute myocardial infarction and its treatment, history of coronary artery disease and its treatment, history of peripheric vasculopathies, history of other comorbidities, presence of symptoms or sings at diagnosis, clinical tumor size, tumour location in the kidney, nephrometric scores (P.A.D.U.A. and R.E.N.A.L.) and TNM stage (according to 8th edition, 2017). Preoperative data were also recorded for all the patients, regarding preoperative imaging, cardiological evaluation, serum haemoglobin and creatinine levels, and complete blood count. Estimated glomerular filtration rate (EGFR) was calculated with the Chronic Kidney Disease Epidemiology Collaboration formula for younger patients (<70 years), and with the Berlin Initiative Study formula for older patients (≥70 years).

Histologic assessment

The following histological parameters were re-evaluated by an experienced and dedicated pathologist: tumour diameter, WHO/ISUP grade, histologic heterogeneity, presence of necrosis, of cystic component, of regions with a different WHO/ISUP grade and of lymphoid infiltration or aggregates.

Computed tomography acquisition protocol and image analysis

A 7 years experienced radiologist measured qualitative, semi-quantitative and quantitative radiological features from CT scan images obtained with a 64-slice multidetector CT scanner (Philips Brilliance 64, Philips, Best, The Netherlands). The CT protocol included scanning acquisition in four phases: unenhanced phase (UP), corticomedullary phase (CMP, at 30 seconds delay after contrast injection), nephrographic phase (NP, at 90 seconds delay after contrast injection) and excretory phases (EP, at 5 minutes delay after contrast injection).

The region of interest (ROI) was defined as the tumour area delimited in axal scan by use of a dedicated software (Intellispace portal v.8, Philips, Best, The Netherlands).

Central scar was defined as a central stellate hypoattenuating area in corticomedullary phase with a surface area lower than 5% than the scan area, with or without progressive enhancement in nephrographic phase. Pseudocapsule was defined as a high- or low-attenuation rim surrounding the tumor. Heterogeneity of the lesion was defined as the presence of different radiologic appearance. Calcification presence was also assessed.

Tumour enhancement was defined high if similar to renal cortex enhancement, moderate if similar to soft tissue enhancement but lower than renal cortex, low if slightly higher than water attenuation and with measurable contrast enhancement.

With regards to their attenuating pattern, both in the unenhanced and nephrographic phase, the lesions were defined as hypoattenuating, isoattenuating, hyperattenuating or mixed compared to adjacent parenchyma. Tumour composition (solid or cystic), necrosis and homogeneity [homogeneous (uniform in attenuation) or heterogeneous (mixed areas of attenuation)] were also assessed.

The following volumetric features were analysed: length of tumour in the short axis (measured in millimetres), length of tumour in the long axis (measured in millimetres); total tumour volume (measured in cubic centimetres); exophytic tumour volume (measured in cubic centimetres), and percentage of exophytic tumour volume.

The following features were analysed in UP images: mean, standard deviation, minimum and maximum tumour attenuation (measured in HU), attenuation of the psoas (measured in HU), the tumour-to-psoas attenuation ratio, calculated as the ratio between the maximum attenuating region of the tumour and the psoas attenuation; the tumour-to-kidney attenuation ratio, calculated as the ratio between the maximum attenuating region of the tumour and the renal cortex attenuation.

The following features were analysed on CMP images: mean, standard deviation, minimum and maximum tumour attenuation (measured in HU), the attenuation of the renal cortex (measured in HU), and tumour-to-kidney attenuation ratio.

The following features were analysed on nephrographic phase (NP) images: mean, standard deviation, minimum and maximum tumour attenuation (measured in HU), the attenuation of the renal cortex (measured in HU), and tumour-to-kidney attenuation ratio.

The following enhancement features were also examined: (I) the early tumour enhancement (measured in HU), calculated as difference between the mean tumour attenuation in the CMP and the mean tumour attenuation in the UP, and (II) the late tumour enhancement (measured in HU), calculated as between the mean tumour attenuation in the NP and the mean tumour attenuation in the UP.

The following 19 features were selected for correlation analysis with transcriptomic signature:
(I)   The mean attenuation of the tumour in unenhanced phase,
(II)   The maximum attenuation of the tumour in the unenhanced phase,
(III)   The minimum attenuation of the tumour in the unenhanced phase,
(IV)   The attenuation of the renal cortex in the unenhanced phase,
(V)   The tumour-to-psoas attenuation ratio in the unenhanced phase,
(VI)   The tumour-to-kidney attenuation ratio in the unenhanced phase,
(VII)   The mean attenuation of the tumour in the corticomedullary phase,
(VIII) The maximum attenuation of the tumour in the corticomedullary phase,
(IX)   The minimum attenuation of the tumour in the corticomedullary phase,
(X)   The attenuation of the renal cortex in the corticomedullary phase,
(XI)   The tumour-to-kidney attenuation ratio in the corticomedullary phase,
(XII)   The mean attenuation of the tumour in the nephrographic phase,
(XIII) The maximum attenuation of the tumour in the nephrographic phase,
(XIV)   The minimum attenuation of the tumour in the nephrographic phase,
(XV)   The attenuation of the renal cortex in the nephrographic phase,
(XVI)   The tumour-to-kidney attenuation ratio in the nephrographic phase,
(XVII) The early tumour enhancement
(XVIII) The late tumour enhancement
(XIX)   The attenuation of the psoas.
An example of 4 phase CT imaging is the showed in *Figure S1*.

Transcriptomics and RNA sequencing quality control

The tumoral molecular landscape was assessed through transcriptomic signature analysis. RNA sequencing was perfumed using quantseq 3' mrna-Seq library prep kit and the prepared libraries were sequenced on the Illumina nextseq 500 platform. Reads were generated towards a poly(A) tail.

Rnaseq quality control was performed by use of the multiqc v1.0. Dev0 (a4e3db2) platform (https://github.com/ewels/multiqc, developed by Phil Ewels et al., Science for Life Laboratory, Stockholm University, Stockholm 106 91, Sweden) (1) implementing the Bamtools toolkit to manage BAM files (http://github.com/pezmaster31/bamtools, developed by Derek Barnett et al., Marth Lab, Biology Dept., Boston College, Boston, USA) (2).

After calculation of Reads Per Kilobase of transcript per Million mapped reads (RPKM) values, the sequencing depth and the quality of alignment of the 6 samples were examined. Sequencing depth was found not to be high, with the reads being roughly between 3 and 5 million per sample. The majority of aligned genes are protein-coding genes. However, fractions of mitochondrial-ribosomal RNA (Mt_rrna), processed pseudogene and long intergenic noncoding RNAs (lincrna) were present (*Figure S2*).

The alignment quality, evaluated with the STAR ultrafast universal RNA-seq aligner plot (https://github.com/alexdobin/STAR developed by Alex Dobin et al., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA) (3), was good, with a high percentage of reads (around 80%) univocally mapped on the human genome (*Figure S3*).

The mean quality value across each base position in the read, measured by fastqc toolkit (developed by Simon Andrews et al., Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom), was satisfactory as the Phred scored above 30 from base 1 to base 75 for each sample (*Figure S4*). The per sequence quality scores, measured through the fastqc toolkit,

assessed the number of reads with average quality scores and suggested that no subset of reads had poor quality (*Figure S5*). Sequence duplication levels, measured by fastqc toolkit, by analysing the level of duplication found in each sequence, suggested that duplication levels were low: from 65.5% to 85% of the libraries in the different samples had no duplication (*Figure S6*). In addition, the total amount of overrepresented sequences found in each library, measured by fastqc toolkit, showed only 0.15% - 0.79% over-represented sequences in the different samples (*Figure S7*).

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE133460 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?Acc=GSE133460).

Statistical analysis

To assess the heterogeneity of the transcriptomes in the samples, Principal Component Analysis (PCA) was fitted. First, a PCA with all transcripts was performed. Then, a list of 369 out of 406 genes known to be associated with ccRCC from the 2013 Cancer Genome Atlas (TCGA) rnaseq and transcriptomic analysis was used (4), after excluding those genes with a constant expression in our samples in order to focus on transcripts really involved in ccRCC and avoid possible confounding (the list of the involved genes is below). A second PCA with this restricted list was performed, considering zero-centred RPKM values. While elaborating the ccRCC-associated gene list, we noted that 26 genes listed in the TGCA analysis were not referenced with the GENCODE basic annotation but with an alias, so a translation to the GENCODE basic annotation was made. Since PCA is unstable when the number of features is greater than the number of samples, we ran 1,000 re-sampling via bootstrapping. For each re-sampling the corresponding PCA was calculated and 50 genes with the highest loadings were obtained. Significant pathways in each Principal Components for these top 50 genes were assessed with Enrichr (Ma'ayan Laboratory, Computational Systems Biology, Mount Sinai Center for Bioinformatics, One Gustave L. Levy Place, Box 1603, New York, NY, USA), in relation to the KEGG 2016 database, while significant gene ontologies were assessed in relation to GO Cellular Component 2018, GO Biological Process 2018, and GO Molecular Function 2018, to evaluate respectively which pathways, cellular components, biological process, and molecular function impacted on the variance of the PC. Statistically significance was set at adjusted P value <0.05.

Pearson's correlation coefficients were used to assess correlation among 19 selected radiomic features themselves and between the radiomic features and RPKM values. Adjusted p-values were computed through Benjamini-Hochberg false discovery rate. Correlation was considered adequate if $\rho < -0.85$ or $\rho > 0.85$.

To interpret the correlation data, a heatmap showing correlation between genes expression and radiomic features was drawn. A dendogram showed the hierarchical relationship between radiomic features and transcripts. Statistical significance of the radiogenomic correlation patterns was assessed with Mann-Whitney U test, for dichotomic comparison, or with Kruskal-Wallis tests, for multiple comparison. In order to obtain the graphics graphpad Prism (graphpad Software La Jolla, CA, USA) was used.

References

1. Ewels P, Magnusson M, Lundin S, et al. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 2016;32:3047-8.
2. Barnett DW, Garrison EK, Quinlan AR, et al. BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics 2011;27:1691-2.
3. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15-21.
4. Creighton CJ, Morgan M, Gunaratne PH, et al. Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature 2013;499:43-9.

List of the 369 genes obtained from The Cancer Genome Atlas (TCGA) Rnaseq and transcriptomic analysis used for PCA

	Genes
1	*AGRN*
2	*CHD5*
3	*KIF1B*
4	*MTOR*
5	*CLCN6*
6	*VPS13D*
7	*SPEN*
8	*UBR4*
9	*ARID1A*
10	*CSMD2*
11	*ZMYM1*
12	*MACF1*
13	*USP24*
14	*PATJ*
15	*DOCK7*
16	*AGL*
17	*COL11A1*
18	*SLC16A4*
19	*LRIG2*
20	*ATP1A1*
21	*SPAG17*
22	*NOTCH2*
23	*PDE4DIP*
24	*NUP210L*
25	*ASH1L*
26	*GON4L*
27	*SPTA1*
28	*TNR*
29	*LAMC2*
30	*HMCN1*
31	*TPR*
32	*CFH*
33	*ASPM*
34	*CRB1*
35	*KIF14*

continued

	Genes
36	*KIF21B*
37	*NFASC*
38	*CR1*
39	*DNAH14*
40	*OBSCN*
41	*SIPA1L2*
42	*TARBP1*
43	*LYST*
44	*RYR2*
45	*PXDN*
46	*MYT1L*
47	*KIDINS220*
48	*GREB1*
49	*APOB*
50	*ITSN2*
51	*EMILIN1*
52	*CAD*
53	*BIRC6*
54	*PLEKHH2*
55	*LRPPRC*
56	*NRXN1*
57	*PAPOLG*
58	*USP34*
59	*XPO1*
60	*ALMS1*
61	*AFF3*
62	*GCC2*
63	*RANBP2*
64	*UGGT1*
65	*THSD7B*
66	*LRP1B*
67	*NEB*
68	*BAZ2B*
69	*SLC4A10*
70	*SCN1A*
71	*LRP2*

continued

	Genes
72	*NFE2L2*
73	*TTN*
74	*DNAH7*
75	*NDUFS1*
76	*ZDBF2*
77	*CPS1*
78	*ABCA12*
79	*FN1*
80	*ZFAND2B*
81	*DOCK10*
82	*SPHKAP*
83	*COL6A3*
84	*SETD5*
85	*VHL*
86	*SETD2*
87	*DOCK3*
88	*BAP1*
89	*PBRM1*
90	*USF3*
91	*ZBTB38*
92	*MED12L*
93	*ZBBX*
94	*PIK3CA*
95	*ATP13A4*
96	*BOD1L1*
97	*PCDH7*
98	*RFC1*
99	*KDR*
100	*ADGRL3*
101	*ANKRD17*
102	*FRAS1*
103	*WDFY3*
104	*PTPN13*
105	*TET2*
106	*NPNT*
107	*ANK2*

	Genes
108	*ANKRD50*
109	*FAT4*
110	*PCDH10*
111	*MAML3*
112	*DCHS2*
113	*FNIP2*
114	*RAPGEF2*
115	*TRAPPC11*
116	*FAT1*
117	*DNAH5*
118	*CDH18*
119	*PDZD2*
120	*ADAMTS12*
121	*NIPBL*
122	*MAST4*
123	*BDP1*
124	*CMYA5*
125	*VCAN*
126	*ADGRV1*
127	*DMXL1*
128	*SLC12A2*
129	*FBN2*
130	*RAPGEF6*
131	*RAD50*
132	*PCDHA12*
133	*PCDHB11*
134	*NSD1*
135	*FLT4*
136	*RREB1*
137	*HIVEP1*
138	*KIF13A*
139	*DNAH8*
140	*CUL9*
141	*ZNF318*
142	*XPO5*
143	*PKHD1*

	Genes
144	*DST*
145	*MDN1*
146	*LAMA2*
147	*UTRN*
148	*PPIL4*
149	*SYNE1*
150	*IGF2R*
151	*PLG*
152	*SDK1*
153	*ABCB5*
154	*DNAH11*
155	*HECW1*
156	*ABCA13*
157	*PCLO*
158	*AKAP9*
159	*COL1A2*
160	*RELN*
161	*CPED1*
162	*ZNF800*
163	*RBM28*
164	*KMT2C*
165	*CSMD1*
166	*CDCA2*
167	*NSD3*
168	*RP1*
169	*CHD7*
170	*ZFHX4*
171	*LRRCC1*
172	*VPS13B*
173	*ZFPM2*
174	*TG*
175	*PLEC*
176	*FREM1*
177	*GBA2*
178	*PRUNE2*
179	*COL15A1*

	Genes
180	*ABCA1*
181	*SVEP1*
182	*KIAA0368*
183	*RGS3*
184	*TNC*
185	*ODF2*
186	*LAMC3*
187	*CAMSAP1*
188	*DIP2C*
189	*FAM208B*
190	*FBXO18*
191	*UPF2*
192	*CUBN*
193	*MYO3A*
194	*ANK3*
195	*JMJD1C*
196	*DDX50*
197	*GRID1*
198	*PTEN*
199	*KIF20B*
200	*BTAF1*
201	*RRP12*
202	*GBF1*
203	*SMC3*
204	*SFXN4*
205	*MKI67*
206	*NAV2*
207	*CCDC73*
208	*KIAA1549L*
209	*TNKS1BP1*
210	*AHNAK*
211	*LRP5*
212	*PPFIA1*
213	*TENM4*
214	*SYTL2*
215	*FAT3*

	Genes
216	*BIRC2*
217	*DYNC2H1*
218	*ATM*
219	*EXPH5*
220	*CEP164*
221	*KMT2A*
222	*HSPA8*
223	*WNK1*
224	*CACNA1C*
225	*VWF*
226	*CHD4*
227	*PZP*
228	*GRIN2B*
229	*ABCC9*
230	*LRRK2*
231	*KMT2D*
232	*ACVR1B*
233	*ESPL1*
234	*ERBB3*
235	*LRP1*
236	*LRIG3*
237	*NAV3*
238	*NT5DC3*
239	*SART3*
240	*NOS1*
241	*GCN1*
242	*DNAH10*
243	*RIMBP2*
244	*ZMYM2*
245	*LATS2*
246	*SACS*
247	*PARP4*
248	*ATP8A2*
249	*MTUS2*
250	*BRCA2*
251	*NBEA*

	Genes
252	*FREM2*
253	*VWA8*
254	*MYCBP2*
255	*SLITRK6*
256	*NALCN*
257	*NYNRIN*
258	*ARHGAP5*
259	*RALGAPA1*
260	*TOGARAM1*
261	*SYNE2*
262	*PCNX1*
263	*YLPM1*
264	*FLRT2*
265	*AHNAK2*
266	*HERC2*
267	*RYR3*
268	*AQR*
269	*STARD9*
270	*FBN1*
271	*SECISBP2L*
272	*DMXL2*
273	*PRTG*
274	*VPS13C*
275	*HERC1*
276	*ITGA11*
277	*IL16*
278	*AKAP13*
279	*ACAN*
280	*ANPEP*
281	*IQGAP1*
282	*LRRK1*
283	*KIAA0430*
284	*MYH11*
285	*RBBP6*
286	*ZNF423*
287	*SALL1*

	Genes
288	*CHD9*
289	*CNOT1*
290	*CDH8*
291	*CDH11*
292	*NFATC3*
293	*ZFHX3*
294	*ADAMTS18*
295	*ZC3H18*
296	*PRPF8*
297	*ZZEF1*
298	*POLR2A*
299	*TP53*
300	*DNAH2*
301	*CHD3*
302	*MYH13*
303	*NCOR1*
304	*MYO15A*
305	*SSH2*
306	*ATAD5*
307	*NF1*
308	*C17orf75*
309	*HEATR9*
310	*CDK12*
311	*NBR1*
312	*GPATCH8*
313	*KANSL1*
314	*MED13*
315	*SDK2*
316	*TTYH2*
317	*DNAH17*
318	*RNF213*
319	*LAMA1*
320	*ASXL3*
321	*SETBP1*
322	*LOXHD1*
323	*MYO5B*

	Genes
324	*ALPK2*
325	*CDH19*
326	*ZNF407*
327	*C3*
328	*FBN3*
329	*MUC16*
330	*COL5A3*
331	*SMARCA4*
332	*CACNA1A*
333	*ADGRE3*
334	*CPAMD8*
335	*TSHZ3*
336	*RYR1*
337	*SIGLEC8*
338	*PEG3*
339	*MACROD2*
340	*PHF20*
341	*PTPRT*
342	*NCOA3*
343	*PREX1*
344	*ARFGEF2*
345	*ZNFX1*
346	*SCAF4*
347	*SYNJ1*
348	*SON*
349	*BRWD1*
350	*PCNT*
351	*MICAL3*
352	*PI4KA*
353	*PRR14L*
354	*TRIOBP*
355	*EP300*
356	*TCF20*
357	*FBLN1*
358	*CELSR1*
359	*TUBGCP6*

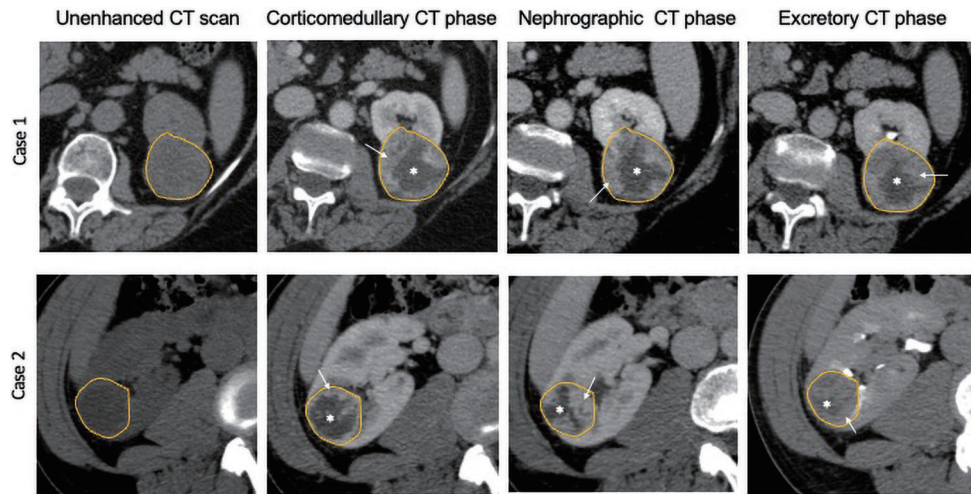|	Genes
360	*SBF1*
361	*DMD*
362	*HDAC6*
363	*KDM5C*
364	*HUWE1*
365	*TAF1*
366	*ATRX*
367	*COL4A5*
368	*STAG2*
369	*TENM1*

Figure S1 An example of 4 phase CT imaging is the following. Four-phase CT protocol (unenhanced, corticomedullary, nephrographic and excretory phases) in two cases of renal cell carcinoma (yellow ROIs in all images). Case 1 is a ccRCC of 4 cm diameter involving the upper third of the left kidney, Case 2 is a ccRCC of 3 cm involving the upper third of the right kidney. Both tumours showed similar features at qualitative assessment of each CT scan, being characterized by round shape, an hypervascular peripheral region (white arrows) and an irregular hypovascular central region (asterisks). At radiomic evaluation case 1 showed slightly lower HU in CMP and NP compared to case 2 (63 and 64 *vs.* 69 and 84 HU) but with higher maximum HU in both phases (168 and 139 HU *vs.* 120 and 106 HU) as for more necrotic lesion with higher degree of vascularization in the periphery. HU, Hounsfield Unit; CMP, corticomedullary phase; NP, nephrographic phase; ROI, Region Of Interest; ccRCC, clear cell Renal Cell Carcinoma; CT, Computed Tomography.

Figure S2 The mean quality value across each base position in the read was good, as the Phred scored remaining above 30 from base 1 to base 75 for each sample.
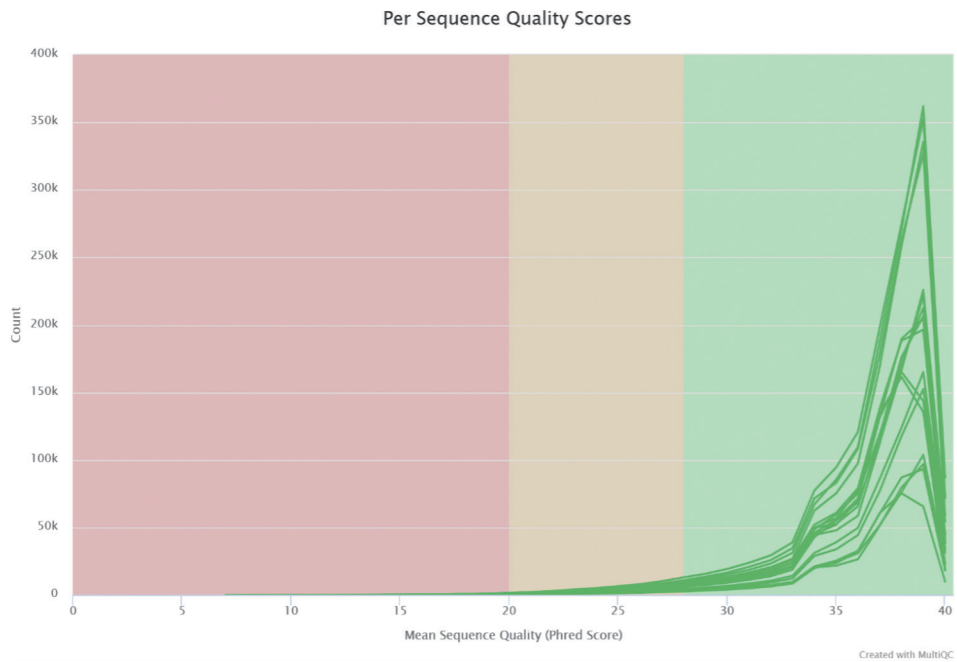

Figure S3 The per sequence quality scores, which evaluates the number of reads with average quality, suggested that no subset of reads has poor quality.
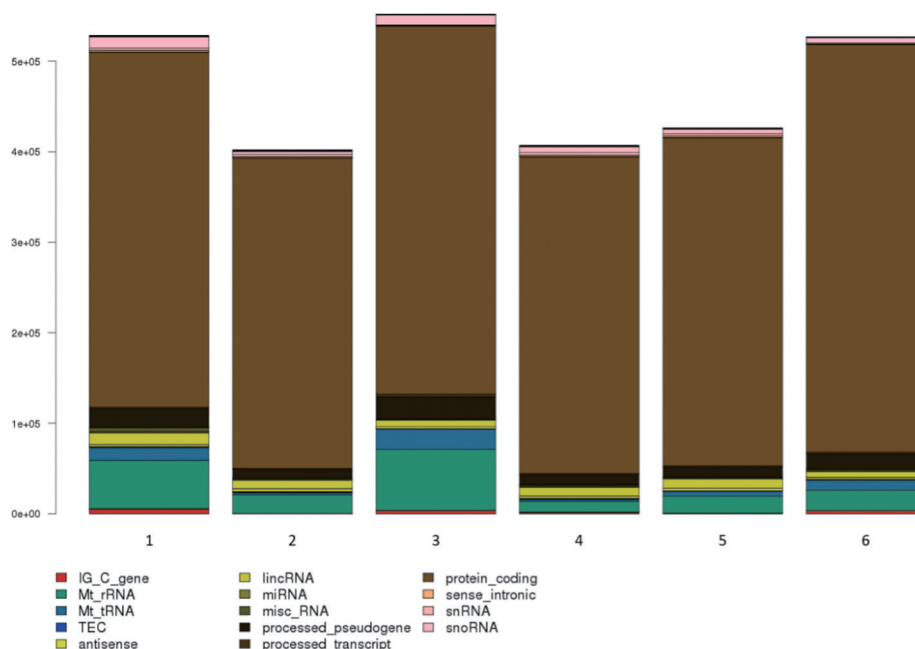
https://dx.doi.org/10.21037/tau-21-713

Figure S4 Biotypes distribution among samples, using RPKM values. 15935 expressed genes have been considered. IG_C_gene are Immunoglobulin (Ig) variable chain and T-cell receptor (tcr) genes imported or annotated according to IMGT (http://www.imgt.org/). Mt_rrna, Mt_trna, mirna, misc_RNA, snrna, snorna are non-coding RNA predicted using sequences from Rfam and mirbase (http://rfam.xfam.org/ & http://www.mirbase.org/). TEC are rnas that need to be experimentally validated, used for non-spliced EST clusters that have polya features; this category has been specifically created for the ENCODE project for regions that could indicate the presence of protein coding genes that require experimental validation. Antisense are transcripts that overlap the genomic span (i.e. Exon or introns) of a protein-coding locus on the opposite strand. LincRNA are long, intervening noncoding (linc) RNA that can be found in evolutionarily conserved, intergenic regions. Processed_pseudogene are pseudogenes that lack introns and are thought to arise from reverse transcription of mrna followed by reinsertion of DNA into the genome. Processed_transcript are transcripts not containing an open reading frame (ORF). Protein_coding are transcripts containing an ORF, thus thought to be protein coding. Sense_intronic are long non-coding transcript in introns of a coding gene that does not overlap any exons. RPKM, Reads Per Kilobase of transcript per Million mapped reads; Ig, Immunogolobuline; RNA, RiboNucleic Acid.
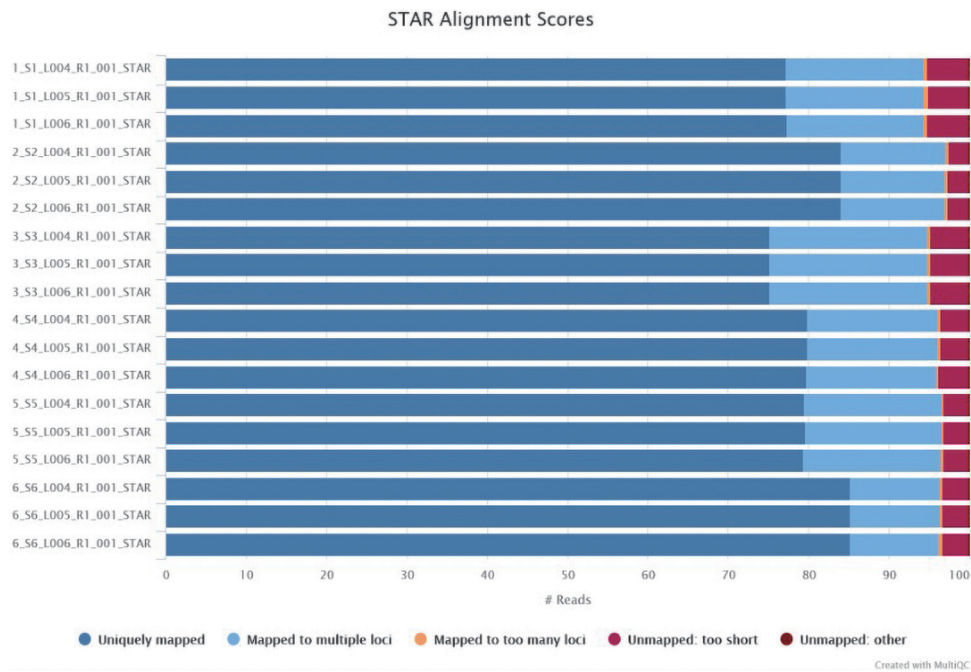
Figure S5 STAR alignment scores plot. RNA uniquely mapped ranges from 75.1% to 85.2%. RNA mapped to multiple loci ranged from 11% to 19.7%. RNA unmapped ranged from 2.6% to 5%.
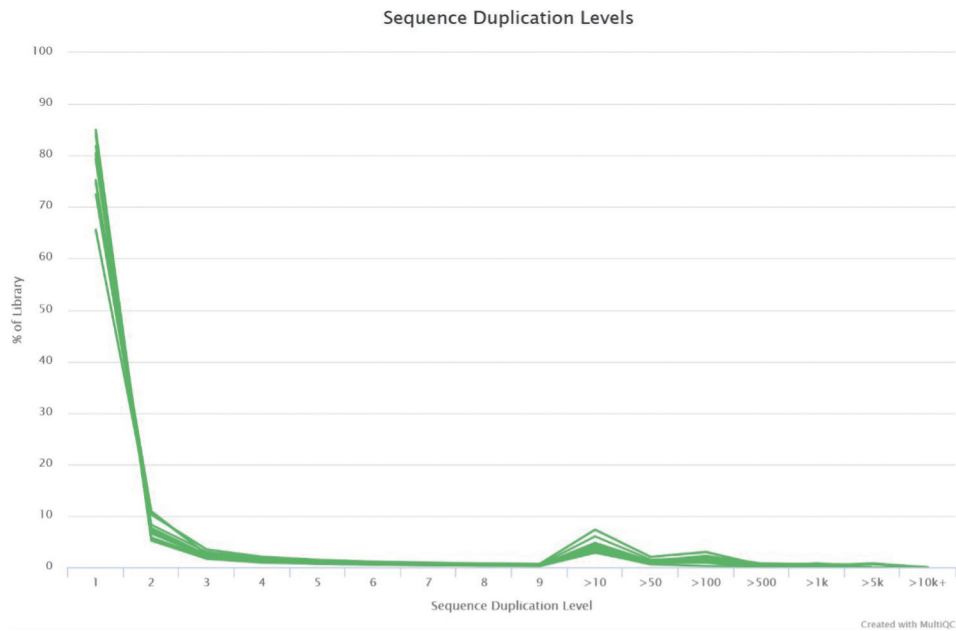

Figure S6 Sequence duplication levels, that show the relative level of duplication found in each sequence, were extremely low: from 65.5% to 85% of the libraries in the samples had no duplication.
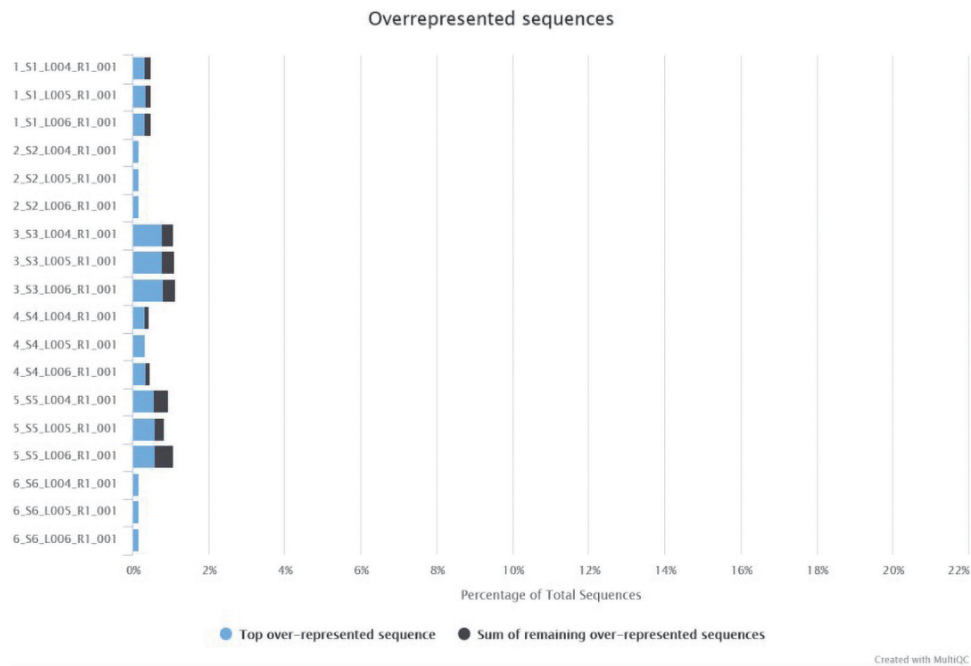
Figure S7 The graph shows the total amount of overrepresented sequences found in each library, with the top over-represented sequences ranging from 0.15% to 0.79% in the different samples.

Table S1 Radiomics characteristics at computed tomography imaging

Variable	
Volumetric data	
Long axis, mm	41 (32.5-46)
Short axis, mm	32 (30.5-41.5)
Total volume, cc	19.5 (16.2-38.4)
Exophytic volume, cc	12 (5-28.1)
Exophytic growth percentage, %	14.5 (13.1-15.6)
Unenhanced phase data	
Mean tumor attenuation, HU	25.5 (21-33.4)
Standard deviations tumor attenuation, HU	14 (11.8-19.3)
Maximum tumor attenuation, HU	35.6 (31.4-41.1)
Minimum tumor attenuation, HU	13 (9.8-27.8)
Psoas attenuation, HU	53.6 (50.6-59.6)
Tumor-to-Psoas ratio	0.67 (0.60-0.7)
Renal cortex attenuation, HU	31.3 (27.3-35.6)
Tumor-to-Kidney Ratio	1.22 (0.94-1.38)
Corticomedullary phase data	
Mean tumor attenuation, HU	133.9 (71.2-164.3)
Standard deviations tumor attenuation, HU	47.6 (32.2-67.7)
Maximum tumor attenuation, HU	168.3 (129.9-232.5)
Minimum tumor attenuation, HU	23.7 (14.7-65.4)
Renal cortex attenuation, HU	173 (155.5-204.8)
Tumor-to-Kidney Ratio	0.92 (0.81-1.19)
Nephrographic phase data	
Mean tumor attenuation, HU	97.9 (74.6-141.2)
Standard deviations tumor attenuation, HU	35.9 (27.5-44.3)
Maximum tumor attenuation, HU	139.4 (114.3-169.5)
Minimum tumor attenuation, HU	32.2 (17.2-74)
Renal cortex attenuation, HU	171 (146.7-241)
Tumor-to-Kidney Ratio	0.75 (0.69-0.81)
Enhancement data	
Early tumor enhancement	106.5 (50.2-131.9)
Late tumor enhancement	70.5 (53.6-108.7)

Four phases at CT scan were evaluated: unenhanced, i.e., before contrast injection; corticomedullary, i.e., 30 seconds after contrast injection; nephrographic, i.e., 90 seconds after contrast injection; excretory, i.e., 5 minutes after contrast injection. No data from the excretory phase were used for the analysis. Data are presented as median and interquartile range (IQR). HU, Hounsfield Units.