frontiers in Oncology

ORIGINAL RESEARCH
published: 04 December 2020
doi: 10.3389/fonc.2020.526577

Patient Stratification of Clear Cell Renal Cell Carcinoma Using the Global Transcription Factor Activity Landscape Derived From RNA-Seq Data

Yanyan Zhu[1], Shundong Cang[1], Bowang Chen[2,3], Yue Gu[4,5,6], Miaomiao Jiang[2,3], Junya Yan[1], Fengmin Shao[4,5,6†] and Xiaoyun Huang[2,3*†‡]*

[1] Department of Oncology, Henan Provincial People's Hospital; Zhengzhou University People's Hospital, Zhengzhou, China, [2] Research and Development, Zhiyu Inc, Shenzhen, China, [3] Zhiyu Center for Systems Biology, Shenzhen, China, [4] Department of Nephrology, Henan Provincial Key Laboratory of Kidney Disease and Immunology, Henan Provincial People's Hospital; Zhengzhou University People's Hospital, Zhengzhou, China, [5] Blood Purification Center, Henan Provincial People's Hospital; Zhengzhou University People's Hospital, Zhengzhou, China, [6] Institute of Nephrology, Henan, China

Clear cell renal cell carcinoma represents the most common type of kidney cancer. Precision medicine approach to ccRCC requires an accurate stratification of patients that can predict prognosis and guide therapeutic decision. Transcription factors are implicated in the initiation and progression of human carcinogenesis. However, no comprehensive analysis of transcription factor activity has been proposed so far to realize patient stratification. Here we propose a novel approach to determine the subtypes of ccRCC patients based on global transcription factor activity landscape. Using the TCGA cohort dataset, we identified different subtypes that have distinct up-regulated biomarkers and altered biological pathways. More important, this subtype information can be used to predict the overall survival of ccRCC patients. Our results suggest that transcription factor activity can be harnessed to perform patient stratification.

Keywords: ccRCC, RNA-Seq, transcription factor, Systems Biology, personalized oncology

OPEN ACCESS

Edited by:
Kanishka Sircar, University of Texas MD Anderson Cancer Center, United States

Reviewed by:
Christoph Oing, University Medical Center Hamburg-Eppendorf, Germany
Luciana Buonerba, Azienda Sanitaria Locale Salerno, Italy

****Correspondence:***
Xiaoyun Huang
x.huang@intelliphecy.com
Fengmin Shao
fmshaoscience@163.com

†These authors share senior authorship

‡Xiaoyun Huang is lead contact

Specialty section:
This article was submitted to Genitourinary Oncology, a section of the journal Frontiers in Oncology

Received: *15 January 2020*
Accepted: *26 October 2020*
Published: *04 December 2020*

Citation:
Zhu Y, Cang S, Chen B, Gu Y, Jiang M, Yan J, Shao F and Huang X (2020) Patient Stratification of Clear Cell Renal Cell Carcinoma Using the Global Transcription Factor Activity Landscape Derived From RNA-Seq Data.10:526577. doi: 10.3389/fonc.2020.526577

INTRODUCTION

Kidney cancer is one major solid cancer, with an estimated 403,262 new cases and 175,098 deaths in 2018 (1). Renal cell carcinoma accounts for around 90% of kidney cancer, with clear cell renal cell carcinoma being one of the most frequent subtypes. Despite recent advances in targeted therapy and immunotherapy, there is still a large gap to be filled for the clinical management of patients diagnosed with clear cell renal cell carcinoma.

Transcription factors are implicated in the initiation and progression of human carcinogenesis. The c-myc oncogene encodes a transcription factor that regulates the transcription of hundreds of genes. It has been well established that transcription factors are situated in the hubs of a complex network shaping the hallmarks of human cancer. Thus, it is intriguing to develop strategies to target transcription factors, especially a combination targeting strategy designed to destroy cancer cell

dependency (2). Small molecules with the capacity to reactivate mutant p53 are now being evaluated in clinical trials (3). Furthermore, transcription factors can be harnessed to develop combination strategy aiming to defeat the intratumor heterogeneity. For example, combination treatments with STAT3 and BCL6 inhibitors were shown to reduce the growth of xenografted tumors (4).

An accurate determination of the subtypes of clear cell renal cell carcinoma is essential to develop personalized therapy for ccRCC patients. ccRCC has been stratified previously with genomic and transcriptomic information (5, 6). However, no studies have focused on transcription factor. Considering the fact that various transcription factors play an important role in metabolic rewiring (7), growth and metastasis (8) of ccRCC cells, it is tempting to approach the molecular stratification of ccRCC using the landscape of transcription factor activity. Here we construct a comprehensive atlas of transcription factor profiles using dataset available in the TCGA cohort comprising of 602 samples.

MATERIALS AND METHODS

Download TCGA Level 3 Data

TCGA level 3 expression data were downloaded from the Pan-Cancer Atlas datasets hosted at Genomic Data Commons (https://gdc.cancer.gov/node/905/). In total, gene expression profiles for 530 ccRCC primary cancer tissues and 72 solid normal tissues were extracted from the Pan-Cancer Atlas.

Transcription Factor Scoring

For each transcription factor, the target genes with known regulation modes were extracted from TTRUST database (9), resulting in a list of genes activated by the transcription factor and a list of genes repressed by the transcription factor. The ratio between the median expression level of activated target gene and the median expression level of repressed target gene was calculated for each transcription factor and log2 transformed to obtain a final transcription factor score.

k-Means Clustering

k-means clustering was performed to unbiasedly analyze the structures in the dataset. Only those transcription factors which have no NAs were included in the clustering analysis. In total, 238 transcription factors were analyzed. Subpopulations of patients were estimated using only the transcription factor score matrix. The optimal number of k was determined by gap statistic using clusGap function in package "cluster". For each number of clusters k, it compares log(W(k)) with E*[log(W(k))] where the latter is defined *via* bootstrapping. The parameter for bootstrapping is 100.

Differentially Expressed Genes

The unique biomarkers characterizing one cluster were determined by statistical test comparing samples within that particular cluster and the rest of all other samples. Two-sided t test was implemented with t.test function in package "stats" using default parameters. "Bonferroni" correction was employed to reduce false hits in multiple comparisons. Adjusted p-value less than 0.05 was considered as statistically significant. Genes with fold change greater than 1.5 were selected as candidate genes for downstream analysis.

Gene List Analysis

Gene list analysis was performed using metascape as previously described (10). Multiple gene lists were analyzed by the online tool metascape (https://metascape.org/) using default parameters to obtain pathway enrichment result. Four major steps were included in the pipeline: ID conversion, Annotation, Membership determination and Enrichment analysis. Differentially expressed gene list was input into the online tool and analyzed using default parameters. Statistically enriched terms were identified and filtered. The remaining significant terms were then hierarchically clustered into a tree based on Kappa-statistical similarities among their gene memberships. A subset of representative terms was selected from the full cluster and converted into a network layout. More specifically, each term is represented by a circle node, where its size is proportional to the number of input genes that fall into that term, and its color represent its cluster identity. Terms with a similarity score > 0.3 are linked by an edge (the thickness of the edge represents the similarity score). The network is visualized with Cytoscape (v3.1.2) with "force-directed" layout and with edge bundled for clarity. One term from each cluster is selected to have its term description shown as label.

Protein–Protein Interaction

All protein–protein interactions (PPI) among each input gene list were extracted from PPI databases and formed a PPI network. GO enrichment analysis was applied to the original PPI network and its MCODE network components to assign biological "meanings", where top three best p-value terms were retained. All input gene lists were also merged into one list and resulted in a PPI network. MCODE components were identified from the merged network. Each MCODE network is assigned a unique color. Network nodes are also displayed as pies. Color code for pie sector represents a gene list.

Survival Analysis

Kaplan–Meier analysis was used for survival curve analysis. It measures the percent of patients surviving with time. The analysis is implemented with R package "survival". Patients were stratified using cluster assignment from k-means clustering. Log-rank test was performed to measure statistical difference between survival curves. p-value < 0.05 was considered as statistically different.

Statistical Method

All statistical analyses were performed using R (version 3.5.1), including basic packages "stats," "graphics," "grDevices," "utils,"

"datasets," "methods," "base," as well as additional packages "factoextra" (version 1.0.7), "ggplot2" (version 3.2.1), "gplots" (3.0.1.1), "survival" (3.1.11). p-value < 0.05 was considered statistically significant. All plots were generated with R.

RESULTS

Establishment of an Analytical Framework to Quantify the Global Transcription Factor Activity

One typical transcription factor regulates the expression of hundreds of downstream gene either positively or negatively. We established an analytical framework to quantify the global transcription factor activity by transforming the gene expression matrix into a transcription factor activity matrix (**Supplementary Figure 1**). In a nutshell, the ratio between the median expression level of activated target gene and the median expression level of repressed target gene was calculated for each transcription factor and log2 transformed to obtain a final transcription factor score, enabling the construction of a comprehensive landscape of transcription factor activity (**Figure 1A**). Globally, all transcription factors were grouped into "activated" and "repressed" states. To further gain insights into the patterns in the transcription factor landscape, the correlation matrix was visualized with a heatmap. It was shown that all samples fell into two apparent groups, and the larger group harbored finer structures (**Figure 1B**).

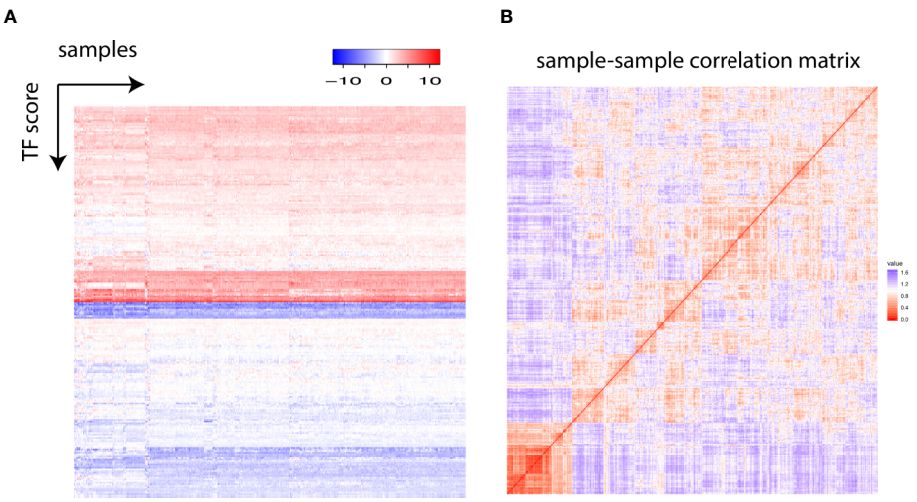

FIGURE 1 | **(A)** Each row represents one transcription factor, and each column stands for one individual. The transcription factor score is plotted. **(B)** The sample–sample correlation matrix is visualized as a heatmap.

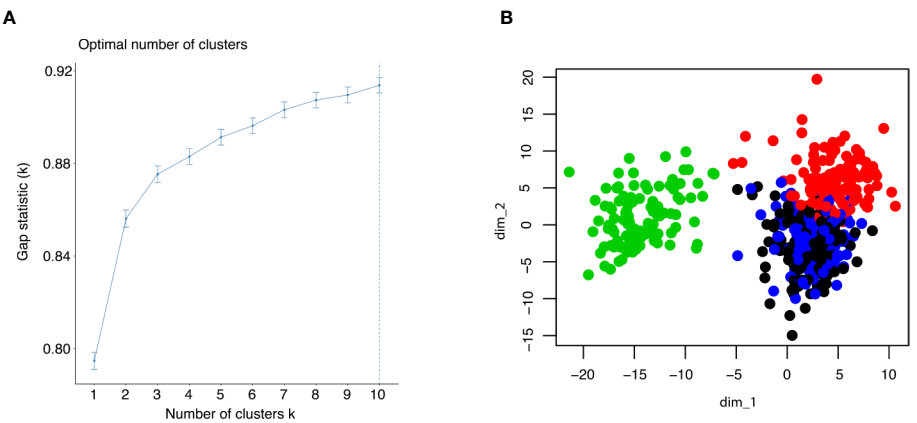

FIGURE 2 | **(A)** Gap statistic was shown up to 10 clusters to determine the optimal k for k means clustering. **(B)** The result of k means clustering was plotted for k = 4, using all individuals.

Patient Stratification Based on Global Transcription Factor Activity

To estimate the number of clusters in the samples, gap statistic was employed. It was found that four clusters could contain more than 88% of the information in the TCGA dataset (**Figure 2A**). Thus, we decided to use k = 4 in subsequent k-means clustering arbitrarily considering the tradeoff between sufficiently few clusters and information retained. To identify the clusters, k-means clustering was performed using all the samples in the TCGA dataset, which included 530 cancer samples and 72 normal tissues (**Figure 2B**). Of note, all normal samples were grouped together as Subtype_3. No normal samples were in the other clusters, while there were also some cancer samples in Subtype_3. Those cancer samples were normal like in the transcription factor activity.

Next, differential gene expression analysis was performed using statistical test. Each subtype identified previously was compared with the rest of all subtypes, resulting a list of genes whose expression is characteristic for that subtype. The top 10 up-regulated genes for each subtype were visualized with heatmap (**Figure 3A**). The up-regulated genes with the Subtype_3 included genes essential for the normal functionality of kidney cells, while the other three subtypes over-expressed genes with known roles in carcinogenesis. Alteration of angiogenesis is common in ccRCC, and this typically involves the VEGF (**Figure 3B**). Not surprisingly, cancer clusters expressed significantly higher level of VEGF. Besides an up-regulated angiogenesis, expression of PD-1 on cancer cells was correlated with the ability of cancer cells to evade the immune systems. Consistent with previous observations, cancer clusters up-regulated the expression of PD-1(**Figure 3B**).

Different Patient Groups Have Distinct Pathway Alterations

To investigate what pathways or biological processes were uniquely altered for the different subtypes, GO analysis was performed. Interestingly, the Subtype_3 stood out while the other three subtypes were grouped together after hierarchical clustering of the enriched GO terms (**Figure 4A**). GO terms uniquely enriched in the Subtype_3 included lipid biosynthetic process, TCA cycle, monovalent inorganic cation homeostasis, mitochondrion organization, metabolism of vitamins and cofactors, suggesting the downregulation of the normal kidney functions in the other subtypes. The fact that only a few cancer samples were in Subtype_3 indicates that most ccRCC cancer cells tune down normal processes exerted by normal kidney cells during carcinogenesis. Subtype_1 and Subtype_2 have altered T cell signaling pathways, such as adaptive immune response,

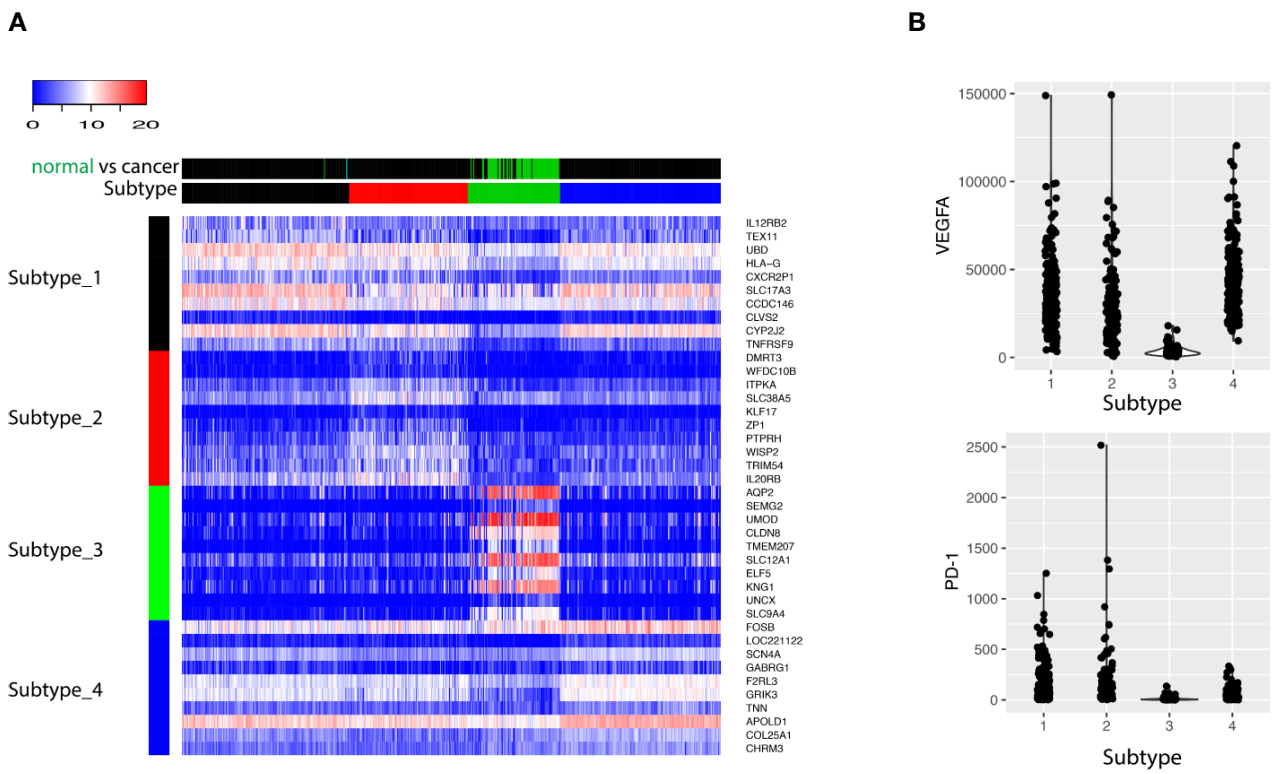

FIGURE 3 | **(A)** The top 10 biomarkers up-regulated in each subtype were shown as heatmap. Each row represents one gene, and each column stands for one sample. The column side color bar labels the information about the sample. In Subtype bar, Subtype_1, Subtype_2, Subtype_3 and Subtype_4 are colored as black, red, green, and blue, respectively. In normal *vs* cancer bar, cancer tissue is labeled as black and normal tissue is labeled as green. **(B)** The expression of VEGFA and PD-1 was shown for the four subtypes: Subtype_1, Subtype_2, Subtype_3, and Subtype_4.

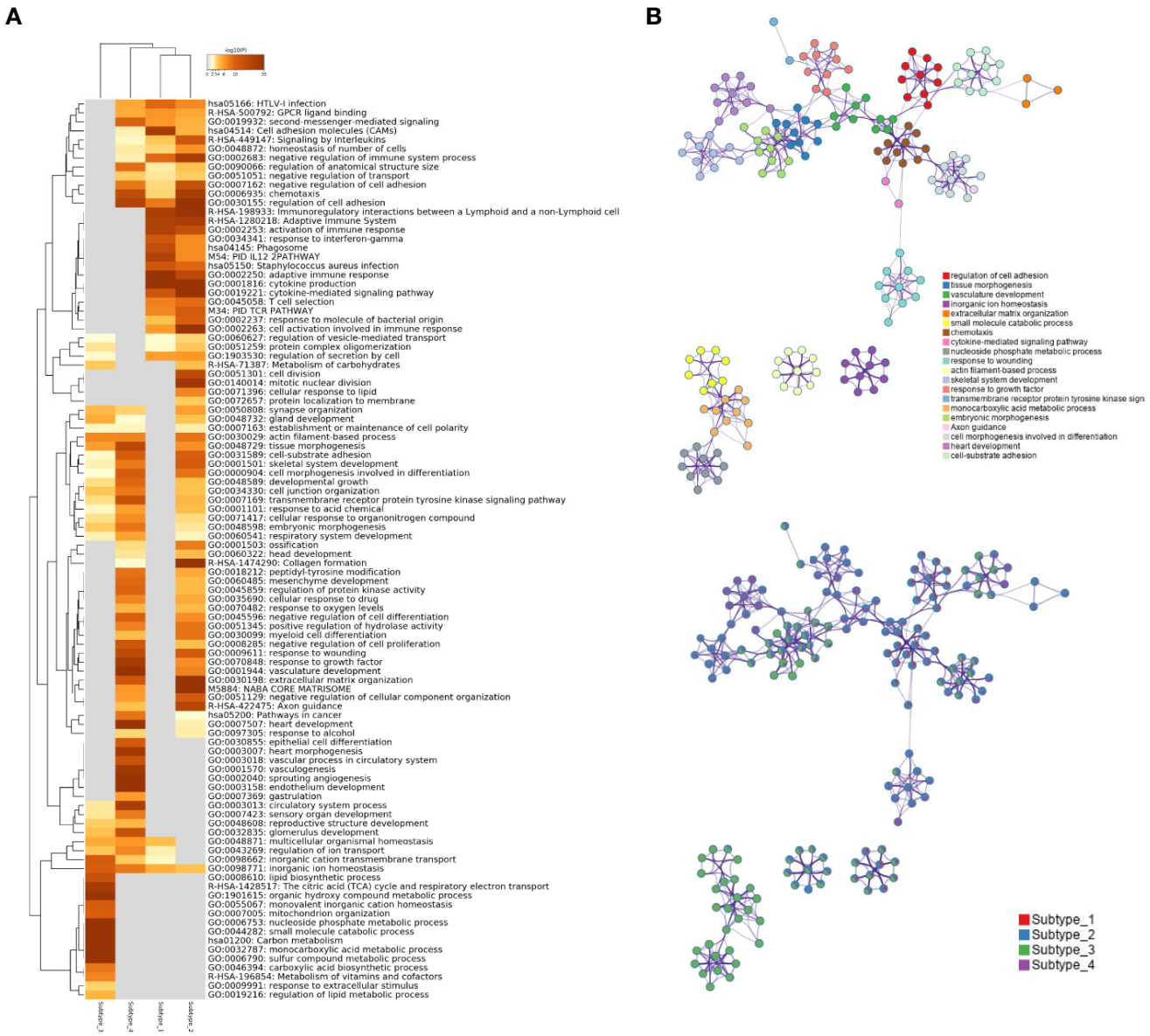

FIGURE 4 | **(A)** Statistically enriched terms were identified and then hierarchically clustered into a tree based on Kappa-statistical similarities among their gene memberships. The heatmap cells are colored by their p-values; white cells indicate the lack of enrichment for that term in the corresponding gene list. **(B)** A subset of representative terms from the full cluster was selected and converted into a network layout. Each term is represented by a circle node, where its size is proportional to the number of genes satisfying that term and its color represents its cluster identity. Terms with a similarity score > 0.3 are linked by an edge where the thickness of the edge indicates the magnitude of similarity. The same enrichment network has its nodes displayed as pies. Each pie sector is proportional to the number of hits originated from a gene list.

cytokine production, cell activation in the immune system and T cell selection. It was consistent with some previous reports showing that cancer rewired the tumor microenvironment to become immune suppressive environment. Subtype_4 preserved some of the functionality of normal kidney cells, such as regulation of ion transport. Of note, Subtype_2 has significantly altered cell division, mitotic nuclear division and cellular response to lipids, unlike any other subtypes. Globally, regulation of cell adhesion was altered in Subtype_1 and Subtype_2 (**Figure 4B**).

Immunity-Related Networks Are Up-Regulated in Subtype_1 and Subtype_2

Using public databases with protein–protein interaction information, MCODE motifs in the regulated genes were extracted (**Figure 5A**). Subtype_1 and Subtype_2 shared

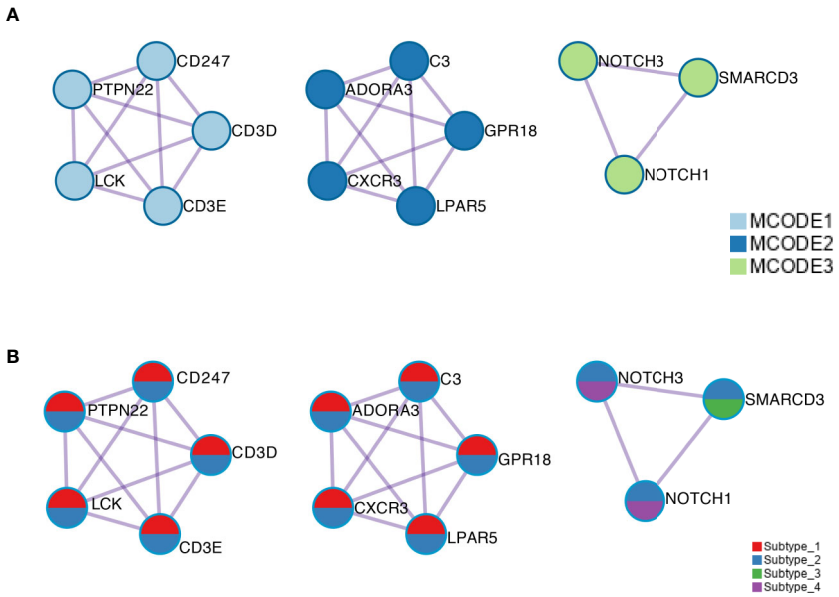

FIGURE 5 | **(A)** The MCODE components were identified from the merged protein–protein interaction network. Each MCODE network is assigned a unique color. **(B)** The same MCODE networks were displayed as in **(A)**, where network nodes shown as pies. Color within pie sector indicates the subtype origin.

MCODE networks in T cell signaling pathways, suggesting an important role played by the immune systems. To consider the contribution of genes for the identified MCODE, it was found that Subtype_2 was the only cluster that contributes to all identified MCODE (**Figure 5B**). Besides two MCODE related to immunity, there was one MCODE consisting of three players: NOTCH1, NOTCH3, and SMARCD3. SMARCD3 is an essential component of the SWI/SNF chromatin-remodeling complex, which epigenetically modulates the expression of downstream genes. Based on the altered pathways, a potential list of drugs approved or under clinical development was assembled for the distinct subtypes identified based on the global transcription factor activity (**Table 1**).

Prognostic Value of the Proposed Transcription Factor Stratification

The transcription factor landscape is sufficient to divide the cancer patients into subtypes, with distinct biomarker and altered pathways. We then ask whether different subtypes have different overall survival. The survival information was downloaded to devise the survival curves for different subtypes (**Figure 6**). Here we used only the cancer samples in the dataset. For the three cancer only subtypes, Subtype_2 has the worse OS as compared with all the other subtypes. To confirm whether a patient who belongs to Subtype_2 is suggestive of prognosis, we undertook a multivariable cox regression analysis (**Supplementary Figure 2**). As shown, the artificial variable "Subtype_2" had a hazard ratio of 1.67 (95% CI: 1.06–2.6), and stood out as an independent prognostic indicator as known clinical variables, including nodal status (HR:1.84, 95% CI: 0.92–3.7), metastasis status (HR:2.07, 95% CI: 0.97–4.4) and stage (HR: 1.34, 95% CI: 0.82–2.2).

TABLE 1 | Potential drugs approved or under clinical development.

Subtype	Target	Drug
Subtype_1	VEGFA	Sorafenib
		Sunitinib
		Pazopanib
		Bevacizumab
	PD-1	Pembrolizumab
		Nivolumab
	CXCR3	AMG487
Subtype_2	VEGFA	Sorafenib
		Sunitinib
		Pazopanib
		Bevacizumab
	PD-1	Pembrolizumab
		Nivolumab
	NOTCH signaling	Brontictuzumab
		Tarextumab
	CDK4/CDK6	Abemaciclib
		Palbociclib
		Ribociclib
	CXCR3	AMG487
Subtype_3	NOTCH signaling	Brontictuzumab
		Tarextumab
Subtype_4	VEGFA	Sorafenib
		Sunitinib
		Pazopanib
		Bevacizumab
	NOTCH signaling	Brontictuzumab
		Tarextumab

Subtype_4 has the longest OS in the three cancer only subtypes. Interestingly, the normal like Subtype_3 did not have significantly better OS compared with Subtype_1 and Subtype_4.

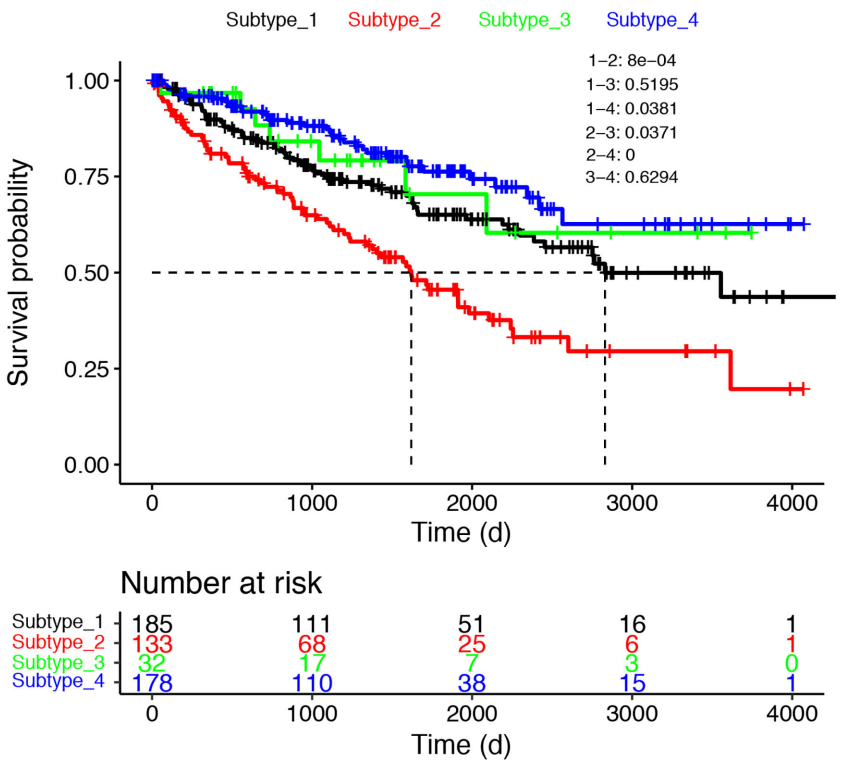

FIGURE 6 | The survival curves were shown for the four subtypes. Only cancer samples were used for the analysis.

DISCUSSION

Patient stratification is key to personalized therapy of cancer patients. One milestone paper (11) divided ccRCC into two major subtypes: clear cell type A (ccA) and clear cell type B (ccB). Subtype ccB was correlated with poor survival compared with subtype ccA. Pathway analysis suggested that ccA group over-expressed genes associated with hypoxia, angiogenesis, and metabolism, while ccB group over-expressed genes involved in EMT, cell cycle and wound healing. Unsupervised clustering of the TCGA ccRCC cohort revealed four stable subtypes using RNA expression data designated m1-m4 (6). The subtype m1 corresponded to the ccA group mentioned earlier, while the subtype m2 and m3 corresponded to the ccB group. The remaining subtype m4 represented a previously uncharacterized subtype with a frequency of around 15%. The m1 subtype was characterized by up-regulation of genes involved in chromatin remodeling.

Transcription factors regulate the transfer of genetic information from DNA to RNA. Traditionally, transcription factors were considered as undruggable. However, a shift of paradigm is realized as multiple approaches to modulate the activity of transcription factors have been demonstrated both preclinically and clinically (12). Targeting transcription factors can be an effective strategy against RCC. It has been demonstrated that STAT3 inhibitor WP1066 inhibited the growth of renal cancer cell lines or xenografted renal cancer cells (13).

Checkpoint inhibition is emerging as a promising therapy for ccRCC patients, with a subset of patients responding to anti-PD-1 monotherapy extremely well. Gene expression analysis uncovered an altered transcriptional output in Janus kinase-signal transducers and activators of transcription (JAK-STAT) signaling (14). It was shown that alteration of transcription factor activity can influence the response to checkpoint inhibitors.

Large cancer genome programs such as TCGA have generated an enormous amount of data that is publicly available to the cancer research community. Mining those datasets has enabled novel insights into cancer biology (15–19). Our analysis suggests that predictive models can be derived from those massive data and used to assign patients to distinctive subtypes. Currently, lots of signature models based on RNA-seq data have been generated but very few made their way into routine clinical use due to lack of reproducibility. As sequencing cost continues to drop, we believe our transcription factor omic (TFO) approach based patient stratification can significantly increase the robustness of signature modeling and transform precision cancer medicine.

Our study can be extended by proof-of-concept experiments to establish the feasibility of predicting drug sensitivity based on the proposed patient stratification. For example, experiments can be carefully designed to investigate the effect of CDK4/CDK6

inhibitors in treating Subtype_2 cancer patients using *in vitro* cancer cell line model, *ex vivo* or *in vivo* cancer models. One popular model to test is the patient derived organoids (20). Subtyping can be performed with transcriptomic analysis using RNA-seq, and the organoids can be perturbed with drug treatments to show proof-of-concept that optimal treatment strategy can be predicted with transcription factor activity based ccRCC subtyping.

In conclusion, we have revisited the molecular stratification of clear cell renal cancer by integrative analysis of multiple publicly available datasets with a focus in the landscape of transcription factor activity. We demonstrated how a deep understanding of ccRCC subtypes can be obtained by a comprehensive and integrative reanalysis of publicly available datasets.

AUTHORS SUMMARY

This study takes advantage of the public available cancer transcriptome data in TCGA to model patient-to-patient variability. A novel computational framework is established to quantify the global transcription factor activity from RNA-seq data. Patient stratification was achieved by unsupervised analysis of the global transcription factor activity landscape. As proof-of-concept, this study highlights the value of publicly available datasets in patient stratification based precision medicine and underscores the importance of transcriptomic profiling in precision medicine for cancer patients.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These can be found in The Cancer Genome Atlas (https://portal.gdc.cancer.gov/).

AUTHOR CONTRIBUTION

YZ, BC, MJ, and XH performed the analysis. YZ, SC, YG, JY, and XH interpreted the results. XH and FS conceived and supervised the study. All authors contributed to the article and approved the submitted version.

FUNDING

This study is partially supported by the Henan Provincial Medical Science and Technology Advancement Program (201702161) and National Science Foundation (Grant ID: 81903106).

ACKNOWLEDGMENTS

This manuscript has been released as a pre-print at biorxiv (21) (Yanyan Zhu et al.).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.526577/full#supplementary-material

SUPPLEMENTARY FIGURE 1 | Workflow illustrating the analysis pipeline undertaken. Gene expression matrix was projected to transcription factor activity matrix, which was used as input for unsupervised clustering. Cluster labels were used for patient stratification. Different clusters were annotated using the differentially expressed genes as compared to the rest of clusters. Survival analysis was performed to obtain an overview of the OS of distinct patient populations.

SUPPLEMENTARY FIGURE 2 | Multivariable cox regression.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492
2. Bhagwat AS, Vakoc CR. Targeting Transcription Factors in Cancer. *Trends Cancer* (2015) 1(1):53–65. doi: 10.1016/j.trecan.2015.07.001
3. Oren M, Tal P, Rotter V. Targeting mutant p53 for cancer therapy. *Aging (Albany NY)* (2016) 8(6):1159–60. doi: 10.18632/aging.100992
4. Deb D, Rajaram S, Larsen JE, Dospoy PD, Marullo R, Li LS, et al. Combination Therapy Targeting BCL6 and Phospho-STAT3 Defeats Intratumor Heterogeneity in a Subset of Non-Small Cell Lung Cancers. *Cancer Res* (2017) 77(11):3070–81. doi: 10.1158/0008-5472.CAN-15-3052
5. Beuselinck B, Job S, Becht E, Karadimou A, Verkarre V, Couchy G, et al. Molecular subtypes of clear cell renal cell carcinoma are associated with sunitinib response in the metastatic setting. *Clin Cancer Res* (2015) 21 (6):1329–39. doi: 10.1158/1078-0432.CCR-14-1128
6. Cancer Genome Atlas Research, N. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* (2013) 499(7456):43–9. doi: 10.1038/nature12222
7. Bleu M, Gaulis S, Lopes R, Sprouffske K, Apfel V, Holwerda S, et al. PAX8 activates metabolic genes via enhancer elements in Renal Cell Carcinoma. *Nat Commun* (2019) 10(1):3739. doi: 10.1038/s41467-019-11672-1
8. Lin ZZ, Ming DS, Chen YB, Zhang JM, Chen HH, Jiang JJ, et al. KMT5A promotes metastasis of clear cell renal cell carcinoma through reducing cadherin-1 expression. *Oncol Lett* (2019) 17(6):4907–13. doi: 10.3892/ol.2019.10163
9. Han H, Cho JW, Lee S, Yun A, Kim H, Bae D, et al. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res* (2018) 46(D1):D380–6. doi: 10.1093/nar/gkx1013
10. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* (2019) 10(1):1523. doi: 10.1038/s41467-019-09234-6
11. Brannon AR, Reddy A, Seiler M, Arreola A, Moore DT, Pruthi RS, et al. Molecular Stratification of Clear Cell Renal Cell Carcinoma by Consensus Clustering Reveals Distinct Subtypes and Survival Patterns. *Genes Cancer* (2010) 1(2):152–63. doi: 10.1177/1947601909359929
12. Bushweller JH. Targeting transcription factors in cancer - from undruggable to reality. *Nat Rev Cancer* (2019) 19(11):611–24. doi: 10.1038/s41568-019-0196-7
13. Horiguchi A, Asano T, Kuroda K, Sato A, Asakuma J, Ito K, et al. STAT3 inhibitor WP1066 as a novel therapeutic agent for renal cell carcinoma. *Br J Cancer* (2010) 102(11):1592–9. doi: 10.1038/sj.bjc.6605691
14. Miao D, Margolis CA, Gao W, Voss MH, Li W, Martini DJ, et al. Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. *Science* (2018) 359(6377):801–6. doi: 10.1126/science.aan5951
15. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The Immune Landscape of Cancer. *Immunity* (2019) 51(2):411–2. doi: 10.1016/j.immuni.2018.03.023
16. Ricketts CJ, De Cubas AA, Fan H, Smith CC, Lang M, Reznik E, et al. The Cancer Genome Atlas Comprehensive Molecular Characterization of Renal Cell Carcinoma. *Cell Rep* (2018) 23(12):3698. doi: 10.1016/j.celrep.2018.03.075

17. Kahles A, Lehmann KV, Toussaint NC, Huser M, Stark SG, Sachsenberg T, et al. Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. *Cancer Cell* (2018) 34(2):211–24 e6. doi: 10.1016/j.ccell.2018.07.001
18. Hoadley KA, Yau C, Hinoue T, Wolf DM, Lazar AJ, Drill E, et al. Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* (2018) 173(2):291–304 e6. doi: 10.1016/j.cell.2018.03.022
19. Ellrott K, Bailey MH, Saksena G, Covington KR, Kandoth C, Stewart C, et al. Scalable Open Science Approach for Mutation Calling of Tumor Exomes Using Multiple Genomic Pipelines. *Cell Syst* (2018) 6(3):271–81 e7. doi: 10.1016/j.cels.2018.03.002
20. Gao D, Vela I, Sboner A, Iaquinta PJ, Karthaus WR, Gopalan A, et al. Organoid cultures derived from patients with advanced prostate cancer. *Cell* (2014) 159(1):176–87. doi: 10.1016/j.cell.2014.08.016
21. Zhu Y, Cang S, Chen B, Gu Y, Jiang M, Yan J, et al. Patient stratification of clear cell renal cell carcinoma using the global transcription factor activity landscape derived from RNA-seq data. *bioRxiv* [Preprint] (2019) 829796. doi: 10.1101/829796

Conflict of Interest: Authors BC, MJ and XH were employed by the company Zhiyu Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.