Dissecting LncRNA Roles in Renal Cell Carcinoma Metastasis and Characterizing Genomic Heterogeneity by Single-Cell RNA-seq

Xue Li1,2,3, Xianwen Meng1, Cong Wei3, Yincong Zhou1, Hongjun Chen1, He Huang3, and Ming Chen1,2

Abstract

Long noncoding RNAs (IncRNA) have recently emerged as important regulators in cancer cell proliferation and metastasis. However, the role of IncRNAs in metastatic clear cell renal cell carcinoma (ccRCC) remains unclear. Here, single-cell RNA sequencing data were analyzed from primary renal cell carcinoma and paired metastatic renal cell carcinoma specimens, and characterized the expression profiles of over 10,000 genes, including 1,874 IncRNAs. Further analysis revealed that IncRNAs exhibit cancer type– and tissue–specific expression across ccRCC cells. Interestingly, a number of IncRNAs (n = 173) associated with ccRCC metastasis, termed ccRCC metastasis–associated IncRNAs (CMAL). Moreover, functional analysis based on a CMAL-PCG coexpression network revealed that CMALs contribute to cell adhesion, immune response, and cell proliferation. In combination with survival analysis, 12 CMALs were identified that participate in TNF and hypoxia-inducible factor 1 signaling to promote ccRCC metastasis. Further investigation on intratumoral heterogeneity showed that some CMALs are selectively expressed in different subpopulations.

Implications: To explore ccRCC metastasis, the current study performed a global dissection of IncRNAs and a complex genomic analysis of ccRCC tumor heterogeneity. The data shed light on the discovery of potential IncRNA biomarkers and IncRNA therapeutic targets. Mol Cancer Res: 16(12): 1879–88. ©2018 AACR.

Introduction

Kidney cancer, or renal cell carcinoma (RCC), is characterized by comprehensive biological process. Contributing factors include aberrant expression of protein-coding genes (PCG) and noncoding genes (1). The most common type of RCC is clear cell renal cell carcinoma (ccRCC). Patients with ccRCC tend to have worse prognosis than patients with other RCC subtypes, and ccRCC is highly prone to metastasize to the lungs, bones, and lymph nodes (2). As one of the major drivers of mortality in patients with cancer, metastasis is a multistep process that involves dissemination of cancer cells to distant organ sites and subsequent adaptation to foreign tissue microenvironments (3). Previous studies have shown some different molecular and signaling pathway signatures between metastatic renal cell carcinoma (mRCC) and primary renal cell carcinoma (pRCC). For example, the expression of cell-surface protein CD44 was significantly higher in mRCCs compared with pRCCs (4); von Hippel-Lindau tumor suppressor (VHL) gene presents variant genotypes between the primary and metastatic tumor (5); a set of genes involved in the process of epithelial–mesenchymal transition (EMT), their expression levels also vary dramatically in aggressive tumors (6). Currently, some significant progress has been made in the development of targeted therapeutic approaches for mRCC. A vascular endothelial growth factor receptor inhibitor, Sunitinib, is now widely prescribed for the treatment of mRCC (7). However, other kidney cancer metastasis markers are not effectively applied to clinical treatment. Because of concerns with drug resistance and side effects, it would be imperative to find new markers for mRCC prognosis and therapy.

LncRNAs are defined as transcripts longer than 200 nucleotides with no protein-coding potential (8). During the past decade, accumulating evidence demonstrates that LncRNAs play important roles in tumorigenesis and cancer invasiveness. For example, overexpression of the IncRNA HOTAIR promotes breast cancer metastasis (9); oncogenic IncRNA PCAT-1 contributes to prostate cancer cell proliferation via regulating MYC stabilization (10). Thus, LncRNAs as functional transcripts could contribute to improve prognosis in patients with cancer (11). Although thousands of lncRNAs have been discovered, their functions in cancer remain poorly studied, and research on IncRNA signatures for cancer metastasis is still fragmentary. LncRNAs were always lower expressed than PCGs, which leads to a difficulty in identification (8). In recent years, RNA sequencing (RNA-seq) has successfully been applied to assess large-scale IncRNAs. In general, the expression of IncRNA is more cancer type–specific than PCGs (12).

The existence of heterogeneity between individual tumors has been recognized by experimental cancer research (13). Intratumor heterogeneity represents a major challenge to effective treatment and personalized medicine (14). The recent development of
Materials and Methods

Renal cell carcinoma RNA-seq datasets
Sequences data and expression data were obtained from three previously RNA-seq datasets: (i) 116 single-cell RNA-seq and 3 popular RNA-seq from NCBI-GEO under accession number GSE73122 (ref. 17; Supplementary Table S1), which single cells are derived from parental mRCC (Pt-mRCC, n = 34), patient-derived xenografts mRCC (PDX-mRCC, n = 36), and patient-derived xenografts pRCC (PDX-pRCC, n = 46); (ii) 448 The Cancer Genome Atlas (TCGA) RCC IncRNA expression datasets from TANRIC (http://ibl.mdanderson.org/tanric/_design/basic/download.html; ref. 18); and (iii) 32 expression profiles of normal kidney cortex from GTEx portal (http://www.gtexportal.org/home/#V6p) by converting to transcripts per million (TPM) values.

Detection of differential PCGs and IncRNAs
We applied the single-cell differential expression (SCDE) software package ("single-cell differential expression") to identify altered IncRNAs and PCGs between mRCC and pRCC single cells. To avoid the noise-induced fluctuation in gene expression due to the xenograft and cell culture, we excluded the significant differentially expressed genes (DEG) between PDX-mRCC and Pt-mRCC (21). Note that 173 IncRNAs and 1,445 PCGs were selected as the significant altered genes by the following strategies: Choose the genes with corrected Z-score of expression difference (cZ) between mRCC and pRCC more than 2 and maximum likelihood estimate (mle) more than 1, meanwhile exclude the genes whose cZ between PDX-mRCC and Pt-mRCC is more than 2 and the mle is more than 2. These differentially expressed IncRNAs were identified as CMALs.

TCGA clear cell renal cell samples classify by CMAL
The IncRNA expression and the clinical data were downloaded from TANRIC (18) and TCGA (22), respectively. Forty CMALs that do not exist in the TANRIC dataset were removed. Totally, 448 TCGA RCC samples and 134 IncRNAs were considered in this part. Unsupervised clustering was performed based on the expression level of CMALs with "Euclidean" as the distance measurement. The "cutree" function in R was used to cut the clustering tree. A small amount of patients have shown their own peculiar expression patterns. These patients will no longer be considered in cut tree, and they are referred to as "non-cluster." Then, a Kaplan–Meier survival curve and a Cox proportional hazard regression model were performed to compare survival between the different subclass of TCGA samples.

Functional prediction of differentially expressed IncRNAs
The "guilt-by-association" analysis was used to identify coexpressed IncRNA-PCG pairs, which Pearson correlation coefficients were calculated based on the expression level between CMALs and PCGs. The threshold of Pearson correlation coefficients was set to more than 0.5. Finally, 474 and 375 PCGs were identified to correspond to up- and downregulated CMALs, respectively. All functional enrichment analyses were performed in Database for Annotation, Visualization and Integrated Discovery (23). Redundant gene ontology (GO) terms were reduced by the Revigo web tool (http://revigo.irb.hr; ref. 24). Cytoscape v3.4.0 (25) was used to visualize the function networks.

Quantitative real-time PCR
The human renal proximal tubule epithelial cell line HK-2 and renal cancer lines 786-O and Caki-1 were kind gifts from Dr. Xinqiang Zhu laboratory (Zhejiang University School of Medicine). All experiments performed on cells that were passaged less than 25 times. Mycoplasma testing was performed during the studies, and cell cultures were free of mycoplasma. HK-2 and 786-O were cultured in RPMI-1640 with 10% FBS, and Caki-1 in MEM with 10% FBS. Reverse transcription was carried out with a PrimeScript RT reagent kit with gDNA Eraser kit (RR047A; Takara) after total cellular RNA was extracted by using Trizol reagent (15596-026; Life Technologies). Quantitative PCR was performed by using SYBR Premix Ex Taq II (RR420A; Takara) according to the manufacturer's instructions with a LightCycler system (Roche Diagnostics). Each of the samples was all prepared in triplicate. The expression of GAPDH was detected for normalization. Sequences of primers used for qRT-PCR in this study are shown in Supplementary Table S5.

Clustering single cells and identifying specific genes
A total of 10,508 abundant genes were used in the principal component analysis (PCA). We ranked all genes based on their highest absolute values of gene loading scores across the first four principal components in PCA. Using the top 1,000 genes ranked by PCA loading scores, hierarchical clustering was performed for mRCC single cells. The dendrogram was divided into four clusters. Two scores were used to measure the specificity of genes according to the principle previously described (26). Gene specificity score
of each comprising cells was computed as the OR for each gene, within a cluster compared with outside a cluster, as shown in formula: $\text{OR}_{ij} = \log \left( \frac{p_{ij}(1-q_{ij})}{p_{ij}q_{ij}} \right)$, where, for a given threshold, $p_{ij}$ is the fraction of cells expressing gene $i$ above this threshold in cluster $j$, and $q_{ij}$ is the fraction of cells expressing gene $i$ above this threshold that are not in cluster $j$. Here, we used each gene’s own 75 percentiles for expression thresholds. Then gene specificity score was used to rank elements within each cluster, and we chose the top 15 PCGs and IncRNAs in each cluster as specific genes. Gene expression enrichment score for each gene in each cluster is calculated by: $e_i = \frac{u_i}{v_i}$, where $u_i$ is the mean expression level of gene $i$ within cluster $j$, and $v_i$ is the mean expression level of gene $i$ outside of cluster $j$.

**Results**
Composition of expressed genes and IncRNAs in single cells
We reanalyzed a previously reported scRNA-seq dataset that profiled 119 single cells and 3 population samples isolated from human mRCC, PDX-mRCC, and paired PDX-pRCC tumor tissues (Supplementary Table S1; ref. 17). In total, 8,634 PCGs and 1,874 IncRNAs were detected after filtering cells and genes with low coverage, which is comparable with previous results of IncRNA detection in KIRC tissue samples (18). We applied GENCODE IncRNAs categorization to our expressed IncRNAs, resulting in the following distribution: antisense (958), lincRNA (714), sense intronic (162), and sense overlapping (40). Among them, 450 (42%) IncRNAs and 4,947 (74%) PCGs were commonly abundant expressed in all three tissues (Supplementary Fig. S1A).

Meanwhile, we calculated the SD of IncRNA’s and PCG’s expression across all single cells. Although the variation was not distinguishable between four IncRNA subtypes, total IncRNAs displayed significantly higher variation than PCGs (Fig. 1A). This result indicated that the heterogeneity of IncRNA was consistently higher than PCGs whether between different tissues or cells in same tissue. Moreover, to investigate whether these IncRNAs could distinguish metastasis tumor tissue from primary tumor tissue, PCA was performed on IncRNAs (Fig. 1B) and PCGs (Supplementary Fig. S1B). PCA results showed that mRCC and pRCC cells represented distinct populations. In addition, the cluster of PDX-mRCC cells overlapped with which of Pt-mRCC cells, suggesting that PDX-mRCC cells can approximately represent parental mRCC.

![Figure 1](image_url)

**Characterization of IncRNA expression profiles in single cells**
We compared the expressed IncRNAs and PCGs in each single cell. Genes with expression levels greater than zero were regarded as expressed genes. On average, 26.6% IncRNAs and 75.1% PCGs were expressed in single cells. We also found that the percentages of IncRNAs are positively correlated with PCGs across single cells and bulk tissues, and there are more expressed genes in three bulk tissue than single cells (Supplementary Fig. S2A). Consistent with the observation, we found that IncRNAs and PCGs that were lower abundant in bulk tissues were generally detected in fewer single cells (Fig. 1C). This result suggests that gene expression level in bulk tissue may reflect the expressed proportion in single cells.
Furthermore, lncRNAs were detected at levels 1.14-fold lower than PCGs on average (P < 2.2e−16, Mann–Whitney U) in our single-cell datasets, and it showed no difference between the four categories of lncRNAs (P > 0.01, Mann–Whitney U, Fig. 1D). This is consistent with former studies (26). Previous studies also suggested that lncRNAs can be expressed at high levels in individual cells (26), so we calculated the expression level ratio of lncRNA:PCG in bulk tissues, single cells, and normal kidney cortex. Our result demonstrated that the expression level of lncRNAs was much lower than that of PCGs (lncRNA:PCG < 1) for most single cells, but there was still few single cells with lncRNA:PCG > 1 (Supplementary Fig. S2B). All of these may serve as evidences that the expression of lncRNAs tends to have lower level and higher difference than PCGs, whereas a small-cell subpopulation tends to express lncRNAs at high levels.

Altered genes expression during RCC metastasis

As shown in Supplementary Fig. S3A, a hierarchical clustering of single cells was performed on a set of previously published metastatic tumor–specific genes (27, 28). The mRCC and pRCC cells were segregated into distinct clusters, but Pt-mRCC (green) and PDX-mRCC (yellow) cells cannot be distinguished by clustering, indicating their high proximity. These results provided preliminary evidence on the alteration of the expression profile during RCC metastasis. Although these metastasis-specific genes were identified from different cancer types and individuals, some of them have also exhibited differential expression in our ccRCC single-cell samples. For example, TGFBR2 can induce EMT in cancer progression and metastatic dissemination (29). In previous study, metastatic tumor–specific genes were found to be involved in the DNA damage response, chromatin modification, differentiation, apoptosis, and the cell cycle (28). However, metastasis-related genes for ccRCC are still lacking, especially for lncRNAs. Therefore, we aim to find more ccRCC metastasis–related genes and lncRNAs.

We identified DEGs using a recently developed Bayesian approach called SCDE (ref. 30; Supplementary Table S2). Note that 1,445 differentially expressed PCGs were identified from metastatic ccRCC (Supplementary Fig. S4). Nine of these genes are included in the previously published metastatic tumor–specific genes (Supplementary Fig. S3A, red bold; Supplementary Fig. S3B: CD24, PEN1). In addition to these 9 PCGs, several other significantly differentially expressed PCGs have also been proved to be associated with tumor metastasis in previous studies. For example, SMAD3, decreased expressed in mRCC, is a target of Notch signaling, regulating autoimmunity (ref. 31; Supplementary Fig. S3B). SERPINE1, increased expressed in mRCC, is implicated in invasion and metastasis (ref. 32; Supplementary Fig. S3B). This may serve as a candidate set for finding the genes involved in the ccRCC metastasis. Further, we investigated the biological function of those differential expressed PCGs by functional enrichment analysis. We found both upregulated genes and downregulated genes were enriched for regulating cell death, apoptosis, and cell adhesion (Fig. 2A). To explore whether there is any functional difference between upregulated PCGs and downregulated PCGs, we separately performed biological functional analysis on each PCG set. Strikingly, upregulated PCGs were significantly associated with mRNA synthesis (Fig. 2A, light pink background), whereas downregulated PCGs were associated with membrane process (Fig. 2A, light blue background). Cell–cell adhesion is mediated by membrane process, and diminished intercellular adhesion is commonly believed as an important sign for the initiation of metastasis. Similarly, as a basic functional process, the increase of mRNA synthesis is always required for tumor cell invasion. Together, these evidences provide a possible interpretation for the metastasis mechanisms. The downregulation of the cell adhesion–related genes would facilitate the generation of free cancer cells, whereas the upregulation of the mRNA synthesis–related genes would promote the proliferation of the cancer cells. Similar results were obtained using Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis (Fig. 2B), and the consistency suggests the reliability of our results.

We identified 173 differentially expressed lncRNAs from metastatic ccRCC (Supplementary Fig. S4). These lncRNAs may be involved in ccRCC metastasis, so we called them CMALs. From our result, there is no differentially expressed lncRNAs on chromosome 13. According to the fold changes of expression level between mRCC and pRCC, we counted the CMALs and altered PCGs for each fold-change condition. We found that the values of lncRNAs’ fold changes detected within single cells were higher than PCGs (Supplementary Fig. S4, inner circle). Intriguingly, some cancer-related lncRNAs, such as HIF1A-AS1, MALAT1, and PTEN (33), are not present in our differentially expressed lncRNA set. This observation may suggest that these cancer-related lncRNAs are not metastasis-related. But, MEG3 (34), previously known as kidney cancer–related lncRNA, also showed low expression in metastatic RCC cells.

Validation of CMALs in the ccRCC metastasis

In order to prove that CMALs are indeed associated with ccRCC metastasis, we derived a set of 448 TCGA RCC tumors (22) according to CMALs’ expression levels. The most widely used features for predicting tumor aggressiveness are the cancer stage, grade, and degree of metastasis. Besides this clinical information, mRNA clusters and miRNA clusters obtained from the previous work (22) were also used for measurement patient status. Next, unsupervised consensus clustering was used to subdivide the patient samples into six robust molecular subpopulations, and a nonclustered group was detected (Fig. 3A). We found that CMAL-based classification was highly concordant with the TCGA clinical features by χ² tests. Although CMAL-cluster2 and CMAL-cluster4 were classified as two small subclusters, they tend to include higher stage, higher grade, higher metastasis rate, and more frequent lymph node involvement (P < 10⁻⁵, χ² test; Supplementary Table S3). This indicated that CMAL-cluster2, 4 tend to be more readily metastasize to other organ and more aggressive. CMAL-clusters were also significantly associated with previously studied cluster results in mRNA and miRNA. Especially, CMAL-cluster2 and CMAL-cluster4 were significantly enriched in the mRNA-cluster2, 3 and miRNA-cluster2 (P value < 10⁻⁵, χ² test) which showed high mortality in previous study (22). As aggressive clusters, the proportion of CMAL-cluster2 and CMAL-cluster4 increases with the increase of malignancy grade (Supplementary Fig. S5). Next, we performed Kaplan–Meier survival analyses to compare survival rate of CMAL-clusters (Fig. 3B). Patients with CMAL-cluster2 and CMAL-cluster4 showed lower probability of overall survival. This is consistent with the above conclusion, indicating that the detected CMALs are significantly correlated with ccRCC metastasis and do play roles in predicting patient outcome. Interestingly, patients within the nonclustered group showed relatively the most aggressive and worst overall survival rate (P < 10⁻⁶).
Activated IncRNAs during metastasis

To identify the putative roles of the activated IncRNAs in ccRCC metastasis, we performed a 'guilt-by-association' analysis to assess the close correlation between PCGs and CMALs, and constructed a coexpressed network between the CMALs and all PCGs in the single cells (Fig. 4A). Subsequently, the functional analysis of these correlated PCGs indicated that downregulated CMALs are clearly functionally related to cell adhesion, immune response, and cell proliferation, which are all known to be associated with cancer metastasis (Supplementary Fig. S6; refs. 35, 36). The similar result was found in downregulated PCGs.

The result of pathway analysis showed that CMALs were activated in cell adhesion, p53, and AMPK signaling pathway that were associated with cancer development. Both up- and downregulated CMALs were activated in the TNF signaling pathway (Fig. 4B). This pathway can induce MAPK and PI3K–AKT signaling pathway, thus regulating the process of apoptosis and cell adhesion (37). Previous work has indicated that TNFα induces EMT and promotes tumorigenicity of RCC by repressing E-cadherin (38). In the EMT process, epithelial cells could break free from the primary tumor site and metastasize at distant sites by attaining a mesenchymal phenotype (29). We extracted 27 IncRNAs involved in TNF signaling pathway (Supplementary Table S4). Overall
survival analysis was performed to further investigate the prognostic potential of these TNF-related CMALs. We found that low expression of MEG3 and CTA-392C11 in the metastasis ccRCC cells was correlated with poor survival. Other 7 upregulated CMALs, including CTB-78F1, RP11-66B24, CTD-3131K8, DGUOK-AS1, RP11-323H4, RP11-563N4, and CTD-2020R17, have also indicated poor survival. Expression profiles of the 9 TNF-related CMAL were displayed in Fig. 4C.

The hypoxia-inducible factor (HIF)-1 signaling pathway activation by overexpression of HIF-1a is a significant pathway that promotes metastasis and EMT (39). We obtained nine HIF-1-related CMALs, including five upregulated CMALs (RP11-167N5, CTB-78F1, RP11-527B8, RP11-563N4, and RP11-66B24) and four downregulated CMALs (LINC01611, CTA-392C11, MEG3, and RP11-380J14; Fig. 4C; Supplementary Table S4). In addition, except for RP11-527B8, LINC01611, and RP11-380J14, other CMALs are also TNF-related. Previous work showed that TNF1a can induce hin expression, suggesting a complex system of crosstalk between hypoxia and inflammation (40). We assume that these overlapping CMALs may play a regulatory role between these two pathways.

To validate the results of single-cell RNA-seq, we performed qRT-PCR to examine the aberrant expression of CMALs in the metastasis ccRCC cell line (CAKI-1) compared with the primary ccRCC cell line (786-O) and the normal kidney cell line (HK-2). Because of the restriction of PCR primer design and the low expression of IncRNAs, we examined two PCGs (SMAD3 and SERPINE1) mentioned above. Indeed, IncRNA DGUOK-AS1 and protein gene SERPINE1 were confirmed to be highly expressed in the CARI-1 cell line, IncRNA LINC01611, and RP11-380J14, and protein gene SMAD3 showed lower expressed in the 786-O cell line, whereas IncRNA RP11-323H4 showed no significant difference in expression (Fig. 4D).

Heterogeneous expression of IncRNAs

Cancer cells in a solid tumor behave as communities. To obtain an overview of single-cell heterogeneity of IncRNAs and PCGs, we calculated the Spearman correlation coefficients between single cells and bulk tissues based on the expression data of IncRNAs and PCGs, respectively. We found that the correlation of cells in the same tissue was higher than that between different tissues, but there were still differences between the cells within the same tissue (Supplementary Fig. S7). We extracted the correlation value between individual tissues. Correlation coefficients derived from IncRNAs profile data (0.494) are significantly lower than those derived from PCGs profile data (0.612; Fig. 5A). From the result, we found that the intratumoral heterogeneity of IncRNAs is higher than PCGs. This finding indicated that IncRNAs contribute more to intratumoral heterogeneity than PCGs.

To investigate whether intratumoral heterogeneity exist in our ccRCC metastasis tumor samples, we performed hierarchical clustering of mRCC (PDX-mRCC and Pt-mRCC) single cells with genes having high loading scores calculated from the PCA. Our algorithm identified four clusters within mRCC single cells (Fig. 5B). To identify the signatures of these subtypes of mRCC cells, we determined the specific expressed IncRNAs and PCGs in each cluster (Fig. 5C). The set of Cluster-1-specific genes contained the greatest number of specific genes compared with other subtypes. Of the top 45 specific PCGs (15 in each cluster), 6 were not annotated in GO biological function. Several ribosomal proteins, such as RPL12A, RPL27A, and RPS25, had high expression in Cluster3, and GO biological function enrichment showed that they participate in SRP-dependent cotranslational protein targeting to membrane. HEG1, HRSPI2, and ARID5B from Cluster4 function in kidney development or lung development.

For each cluster, specific IncRNAs and PCGs can also be found in the DEG set we identified previously (Fig. 5C, red bold). This result suggests that these ccRCC metastasis-associated genes did not express in all mRCC cells, but specifically expressed in different subpopulations. This finding may indicate that cancer cell subtypes may be able to achieve the same or similar function by expressing different specific genes and activating complementary pathways to maintain cancer cell proliferation and differentiation. Specifically, except Cluster3, PRO6A, KLF4, and DAA1 were highly expressed in other three clusters, respectively, and they are
all involved in the Wnt signaling pathway, which can mediate tumor metastasis (41). SEMA6B and FBXL11 (42) in Cluster1, RSL1 in Cluster2, and DNAJB1 and RPL29 in Cluster3 have biological functions in cell–cell adhesion and cell migration. KLF4 and MED110 in Cluster2, TPT1 in Cluster3, and ANXA1 in Cluster4 take part in stem cell population maintenance and cell differentiation. Among them, lncRNA CCAT1 have been identified to have functional roles in cancer invasion and be deregulated in many cancer types (43). We suggest that clusters of tumor cell-specific genes do exist and might relate to different pathways and through different mechanisms to promote tumorigenesis and neoplastic progression.

**Discussion**

LncRNAs are emerging as key players in tumorigenesis. Plenty of previous works demonstrated that these molecules have functional roles in tumor metastasis via associating with apoptosis,
RNA-seq data to prove reanalyzed the public paired pRCC and mRCC single-cell cycle, and cellular differentiation (9, 44). In this study, we reanalyzed the public paired pRCC and mRCC single-cell RNA-seq data to profile lncRNA expressions and investigated for the differences between them. We demonstrated that the expression of lncRNAs was heterogeneous across all ccRCC cells both intertumor and intratumor. To our knowledge, this is the first study systematically analyzing the differentially expressed lncRNAs and PCGs in human ccRCC metastasis. The altered expression of specific lncRNAs between metastasis and primary tumors may work as independent biomarkers for diagnosis and prognosis (45).

We identified 173 CMALs, most of which were not previously reported to be implicated in cancer invasion and metastasis. CMAL clustering with TCGA expression data was significantly associated with clinical manifestations. Notably, the nonclustered patients with unique aberrant expression patterns had the worst overall survival rate. Compared with clustered patients, the CMAL expression patterns of these nonclustered patients are more complex and diverse, making it difficult to classify and analyze. Thus, the treatment and diagnosis of these patients may require more precise personalized care.

Heterogeneous tumors are composed of differentiated cancer cells, stromal cells, cancer stem cells (CSC), and so on (46). CSC is a subset of cancer cells that possess the driving force of tumorigenesis and metastasis (47). Several molecular markers, such as CD44 (48), CXCR4 (48), and ALDH1A1 (49), have been applied to identify CSCs in renal tumor. However, in our study, the SDs of these three stem-cell markers were less than 1, suggesting four subtypes do not tend to be “stem-like.” One possible explanation for this finding is the small number of cell samples in our research, and the possibility of exiting CSCs in the 80 mRCC cells is slim. Another explanation is that CSCs did not exist in our ccRCC tumor sample.

Furthermore, we examined intratumor heterogeneity properties of mRCC in the context of specific genes in each subtype. Previous studies showed that tumor heterogeneity manifested in many aspects, such as genetic mutation, plastic gene expression, and protein modification (50). Our study suggests that due to tumor heterogeneity in gene expression profile, there are different tumor cell populations addicted to different oncogene pathways. These subpopulations may be developed in the differentiation of tumor stem cells or primitive cells, and have formed a complete

Li et al.
system to resist the resistance of human immunology and drug therapy. Compared with single genes, common pathways with subtype-specific genes may represent more robust prognostic signatures and deserve further investigation.

Our study reveals a prevalent and complex lncRNA-mediated cRCC metastasis. Those cMAls might have oncogenic or tumor-suppressive functions, and the perturbation of the IncRNA-mediated tumor metastasis might be exploited for cancer prognosis and therapy.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: X. Li, C. Wei, H. Huang, M. Chen
Development of methodology: X. Li, C. Wei, H. Huang
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): H. Huang
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): X. Li, X. Meng, Y. Zhou

Writing, review, and/or revision of the manuscript: X. Li, H. Chen
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): H. Chen, H. Huang
Study supervision: H. Huang, M. Chen

Acknowledgments
This work was supported by the National Key Research and Development Program of China (2016YFA0501704), National Natural Sciences Foundation of China (Nos. 31571366 and 31771477), and the Fundamental Research Funds for the Central Universities.

The authors would like to thank Yongjing Liu and Dr Ruxiu Tie for helpful discussions. We thank Dr. Lijuan Wu for sharing the 789-O, Caki-1, and HK-2 cell lines, and Drs. Lijun Zhu and Huilin Zheng for helping with PCR probe design. We also thank the authors who generously shared their data.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 19, 2017; revised March 9, 2018; accepted July 26, 2018; published first August 6, 2018.

References
1. Linehan WM, Srinivasan R, Schmidt LS. The genetic basis of kidney cancer: a metabolic disease. Nat Rev Urol 2010;7:277–85.
2. Cheville JC, Lohse CM, Zincke H, Weaver AL, Blute ML. Comparisons of outcome and prognostic features among histologic subtypes of renal cell carcinoma. Am J Surg Pathol 2003;27:612–24.
3. Steeg PS. Tumor metastasis: mechanistic insights and clinical challenges. Nat Med 2006;12:895–904.
4. Lim SD, Young AN, Paner GP, Amin MB. Prognostic role of CD44 cell adhesion molecule expression in primary and metastatic renal cell carcinoma: a clinicopathologic study of 125 cases. Virchows Arch 2008;452:49–55.
5. Vaziri SA, Tavares EI, Golshayan AR, Rini BI, Aydin H, Zhou M, et al. Differing von hippel lindau genotype in paired primary and metastatic tumors in patients with clear cell renal cell carcinoma. Front Oncol 2012;2:51.
6. Harada K, Miyake H, Kusuda Y, Fujisawa M. Expression of epithelial-mesenchymal transition markers in renal cell carcinoma: impact on prognostic outcome in patients undergoing radical nephrectomy. BJU Int 2012;110:E131–7.
7. Kollmannsberger C, Soulieres D, Wong R, Scalera A, Gaspo R, Rijarna S. Sunitinib therapy for metastatic renal cell carcinoma: recommendations for management of side effects. Can Urol Assoc J 2007;1:S41–54.
8. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of transcriptome data with or without a reference genome. BMC Bioinformatics 2011;12:323.
9. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498–504.
10. Prensner JR, Chen W, Han S, Iyer MK, Cao Q, Kothari V, et al. The long non-coding RNA PCAT-1 promotes prostate cancer cell proliferation through c-Myc. Neoplasia 2014;16:900–8.
11. Li J, Chen Z, Tian L, Zhou C, He MY, Gao Y, et al. LncRNA profile study reveals a three-lncRNA signature associated with the survival of patients with oesophageal squamous cell carcinoma. Gut 2014;63:1700–10.
12. Huarte M. The emerging role of lncRNAs in cancer. Nat Med 2015; 21:1253–61.
13. Dexter DL, Kowalski HM, Blazer BA, Fligiel Z, Vogel R, Heppner GH. Heterogeneity of tumor cells from a single mouse mammary tumor. Cancer Res 1978;38:1174–81.
14. Almenndor F, Matsumyk A, Polyak K. Cellular heterogeneity and molecular evolution in cancer. Ann Rev Pathol 2013;8:277–302.
15. Tang F, Lao K, Surani MA. Development and applications of single-cell transcriptome analysis. Nat Methods 2011;8:S6–11.
32. Klein RM, Bernstein D, Higgins CE, Higgins PJ. SERPINE1 expression discriminates site-specific metastasis in human melanoma. Exp Dermatol 2012;21:551–4.
33. Chen G, Wang Z, Wang D, Qiu C, Liu M, Chen X, et al. LncRNA Disease: a database for long non-coding RNA-associated diseases. Nucleic Acids Res 2013;41:D983–6.
34. Martens-Uzunova ES, Böttcher R, Croce CM, Jenster G, Visakorpi T, Calin GA. Long noncoding RNA in prostate, bladder, and kidney cancer. Eur Urol 2014;65:1140–51.
35. Desgrosellier JS, Cheresh DA. Integrins in cancer: biological implications and therapeutic opportunities. Nat Rev Cancer 2010;10:9–22.
36. Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. Cell 2010;140:883–99.
37. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 2017;45:D353–D61.
38. Ho MY, Tang SJ, Chuang MJ, Cha TI, Li JY, Sun GH, et al. TNF-alpha induces epithelial-mesenchymal transition of renal cell carcinoma cells via a GSK3beta-dependent mechanism. Mol Cancer Res 2012;10:1109–19.
39. Giannoni E, Bianchini F, Calortina L, Chiarugi P. Cancer associated fibroblasts exploit reactive oxygen species through a proinflammatory signature leading to epithelial mesenchymal transition and stemness. Antioxid Redox Signal 2011;14:2361–71.
40. Palazon A, Goldrath AW, Nietz V, Johnson RS. HIF transcription factors, inflammation, and immunity. Immunity 2014;41:518–28.
41. Xi Y, Chen Y. Wnt signaling pathway: implications for therapy in lung cancer and bone metastasis. Cancer Lett 2014;353:8–16.
42. Ihlyehanda SS, Das M, Ma YQ, Ding K, Wang X, Gupta S, et al. Migflin, a molecular switch in regulation of integrin activation. J Biol Chem 2009;284:4713–22.
43. Yang F, Xue X, Bi J, Zheng L, Zhi K, Gu Y, et al. Long noncoding RNA CCAT1, which could be activated by c-Myc, promotes the progression of gastric carcinoma. J Cancer Res Clin Oncol 2013;139:437–45.
44. Serviss IT, Johnson P, Grander D. An emerging role for long non-coding RNAs in cancer metastasis. Front Genet 2014;5:234.
45. Zhang XQ, Sun S, Lam KF, Kiang KM, Pu JK, Ho AS, et al. A long non-coding RNA signature in glioblastoma multiforme predicts survival. Neurobiol Dis 2013;58:123–31.
46. Marusyk A, Almendro V, Polyak K. Intra-tumour heterogeneity: a looking glass for cancer? Nat Rev Cancer 2012;12:323–34.
47. Medema JP. Cancer stem cells: the challenges ahead. Nat Cell Biol 2013;15:338–44.
48. Peired AJ, Sisti A, Romagnani P. Renal cancer stem cells: characterization and targeted therapies. Stem Cells Internat 2016;2016: 8342625.
49. Mataki D, Szymbaski L, Szczylk C, Sledziewski R, Lian F, Bartnik E, et al. Biology of renal tumour cancer stem cells applied in medicine. Contemporary Oncol 2015;19:A44–51.
50. Sun XX, Yu Q. Intra-tumor heterogeneity of cancer cells and its implications for cancer treatment. Acta Pharmacol Sin 2015;36:1219–27.