

RESEARCH ARTICLE

Proteotranscriptomic Analysis Reveals Stage Specific Changes in the Molecular Landscape of Clear-Cell Renal Cell Carcinoma

Benjamin A. Neely¹, Christopher E. Wilkins², Laura A. Marlow³, Dariya Malyarenko⁴, Yunee Kim⁵, Alexandr Ignatchenko⁶, Heather Sasinowska⁷, Maciek Sasinowski^{7,8}, Julius O. Nyalwidhe^{2,9}, Thomas Kislinger^{5,6}, John A. Copland³, Richard R. Drake^{1*}

1 Department of Cell and Molecular Pharmacology and Experimental Therapeutics, Medical University of South Carolina, Charleston, South Carolina, United States of America, **2** Department of Microbiology and Molecular Cell Biology, Eastern Virginia Medical School, Norfolk, Virginia, United States of America, **3** Department of Cancer Biology, Mayo Clinic Comprehensive Cancer Center, Jacksonville, Florida, United States of America, **4** Department of Radiology, University of Michigan Medical School, Ann Arbor, Michigan, United States of America, **5** Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada, **6** Princess Margaret Cancer Center, Toronto, Ontario, Canada, **7** INCOGEN, Inc., Williamsburg, Virginia, United States of America, **8** Venebio Group, LLC, Richmond, Virginia, United States of America, **9** Leroy T. Canoles Jr. Cancer Research Center, Eastern Virginia Medical School, Norfolk, Virginia, United States of America

* draker@musc.edu



OPEN ACCESS

Citation: Neely BA, Wilkins CE, Marlow LA, Malyarenko D, Kim Y, Ignatchenko A, et al. (2016) Proteotranscriptomic Analysis Reveals Stage Specific Changes in the Molecular Landscape of Clear-Cell Renal Cell Carcinoma. PLoS ONE 11(4): e0154074. doi:10.1371/journal.pone.0154074

Editor: Lu-Zhe Sun, University of Texas Health Science Center, UNITED STATES

Received: February 29, 2016

Accepted: April 10, 2016

Published: April 29, 2016

Copyright: © 2016 Neely et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data are fully available without restriction. Accession numbers for data deposited to MassIVE and PRIDE are available in the paper.

Funding: This research was supported by the South Carolina SmartState Center of Economic Excellence in Proteomics and the MUSC Proteomics Center (BAN and RRD). TK is supported through the Canadian Research Chair program. This work was funded in part from NIH/NCI grants 2R44CA125807-02 (MS), R01CA104505 (JAC), R01CA136665 (JAC), and R01CA104505-05S1 (JAC); a generous gift from

Abstract

Renal cell carcinoma comprises 2 to 3% of malignancies in adults with the most prevalent subtype being clear-cell RCC (ccRCC). This type of cancer is well characterized at the genomic and transcriptomic level and is associated with a loss of *VHL* that results in stabilization of HIF1. The current study focused on evaluating ccRCC stage dependent changes at the proteome level to provide insight into the molecular pathogenesis of ccRCC progression. To accomplish this, label-free proteomics was used to characterize matched tumor and normal-adjacent tissues from 84 patients with stage I to IV ccRCC. Using pooled samples 1551 proteins were identified, of which 290 were differentially abundant, while 783 proteins were identified using individual samples, with 344 being differentially abundant. These 344 differentially abundant proteins were enriched in metabolic pathways and further examination revealed metabolic dysfunction consistent with the Warburg effect. Additionally, the protein data indicated activation of *ESRRA* and *ESRRG*, and *HIF1A*, as well as inhibition of *FOXA1*, *MAPK1* and *WISP2*. A subset analysis of complementary gene expression array data on 47 pairs of these same tissues indicated similar upstream changes, such as increased *HIF1A* activation with stage, though *ESRRA* and *ESRRG* activation and *FOXA1* inhibition were not predicted from the transcriptomic data. The activation of *ESRRA* and *ESRRG* implied that *HIF2A* may also be activated during later stages of ccRCC, which was confirmed in the transcriptional analysis. This combined analysis highlights the importance of *HIF1A* and *HIF2A* in developing the ccRCC molecular phenotype as well as the potential involvement of *ESRRA* and *ESRRG* in driving these changes. In addition, *cofilin-1*, *profilin-1*, *nicotinamide N-methyltransferase*, and *fructose-bisphosphate aldolase A* were identified

the David & Lois Stulberg Endowed Fund for Kidney Cancer Research (JAC); Mr. and Mrs. Ompal Chauhan Research Fund (JAC); Kidney Cancer Research at Mayo Clinic in Florida (JAC); Scheidel Foundation (JAC). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. INCOGEN, Inc. & Venebio Group, LLC, provided support in the form of salaries for authors HS and MS, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

Competing Interests: The affiliation of authors HS and MS with INCOGEN, Inc. & Venebio Group, LLC, does not alter the authors' adherence to PLOS ONE policies on sharing data and materials.

as candidate markers of late stage ccRCC. Utilization of data collected from heterogeneous biological domains strengthened the findings from each domain, demonstrating the complementary nature of such an analysis. Together these results highlight the importance of the VHL/HIF1A/HIF2A axis and provide a foundation and therapeutic targets for future studies. (Data are available via ProteomeXchange with identifier PXD003271 and MassIVE with identifier MSV000079511.)

Introduction

Renal cell carcinoma accounts for 90% of kidney cancers, and 2 to 3% of malignancies in adults [1]. In 2012, kidney cancer was the ninth and fourteenth most common cancer in men and women worldwide, respectively [2], predominantly affecting men 2:1 [1]. Of these worldwide cases in 2012, there was a 42% mortality rate [2], and in a survey of over 340,000 RCC cases between 2001 and 2010, 25% of cases were stage III or IV [3]. Within renal cell carcinoma, clear-cell renal cell carcinoma (ccRCC) is the most prevalent subtype comprising 70 to 80% of cases [1]. Detection of ccRCC is the result of an incidental diagnosis 50 to 60% of the time, and survival is greatly affected by tumor grade at diagnosis. The 5-year survival rate for grade I and II tumors is 88.9% and grade III and IV tumors is 65.6% and 41.7%, respectively [4]. For this reason it is imperative to develop a clear understanding of the molecular pathogenesis of ccRCC in order to identify new targets related to metastatic ccRCC.

The defining genetic characteristic of ccRCC is the loss of chromosome 3p and/or mutations of the von Hippel-Lindau tumor suppressor gene (*VHL*), especially in sporadic ccRCC where 92% have inactivated *VHL* by mutation, hypermethylation or deletion [5]. In a survey of data from The Cancer Genome Atlas (TCGA) encompassing 12 major cancer types, *VHL* mutation occurred in over 50% of ccRCC cases and was not a factor in other malignancies [6], although some studies have observed *VHL* mutation in over 80% of ccRCC cases [7]. With the loss of *VHL*, hypoxia-inducible factor 1 α (HIF1A) and HIF-2 α (HIF2A; also referred to as EPAS1) are stabilized [8, 9], each regulating at least 350 gene loci [10, 11]. Specifically, stabilization of HIF1A leads to downstream changes resulting in an oncologic shift in glucose and glutamine metabolism consistent with the Warburg effect (first described in the 1920s [12] and expertly reviewed in [13]). This results in increased glucose uptake and increased glycolytic flux, with the top-half of the glycolytic pathway facilitating intermediates for the pentose phosphate pathway and biosynthesis required for tumor growth. The majority of glucose goes to lactate fermentation, thereby uncoupling oxidative metabolism in the tricarboxylic acid (TCA) cycle [14]. In order to facilitate fatty acid synthesis, the TCA cycle instead relies on glutamine flux [15], with primary outputs being citrate and malate. Importantly, HIF1A and HIF2A stabilization have different effects on promoting tumor growth (reviewed in [16]): HIF1A activates glycolytic genes while HIF2A promotes growth and angiogenesis [17, 18]. Moreover, later stages of ccRCC have higher levels of HIF2A and HIF1A and HIF2A are thought to be antagonistic [9, 18].

The seminal TCGA study of ccRCC confirmed many Warburg effect-related changes at the genomic/transcript level and correlated changes with stage and survival [19]. This large-scale characterization study identified increased DNA methylation in higher stage tumors, as well as key transcriptional hubs (HIF1A/ARNT, MYC/MAX, SP1, FOXM1, JUN and FOS). It was also found that down-regulation of genes in the TCA cycle, and up-regulation of genes in the pentose phosphate pathway and fatty acid synthesis correlated with poor survival. Specifically,

reduced transcript levels of AMP-activated kinase (AMPK) and increased levels of acetyl-CoA carboxylase (ACC) corresponded to worse survival, changes that would contribute to increased fatty acid synthesis. In addition to the TCGA work, other transcriptomic studies have been performed using tumor-normal paired samples, identifying almost 6000 differentially expressed genes in ccRCC, and 31 genes required for tumorigenesis have been confirmed *in vitro* [20]. Additionally, genes related to adipogenesis have been identified, likely explaining the hallmark lipoid clear cell morphology of ccRCC [21]. Together these studies have improved our understanding of the molecular changes at the gene and transcript level required for the ccRCC phenotype.

In addition to characterizing genomic and transcriptomic changes in ccRCC, there have been a series of shotgun-proteomic analyses [22–25]. In general these studies have confirmed dysregulated metabolic patterns supporting the Warburg shift predicted at the transcript level [22, 24], including increased abundance of proteins in the pentose phosphate pathway related to tumor aggressiveness [23]. Studying ccRCC across different biological domains (gene, transcript, protein, metabolite) is essential since changes are not linear between these domains. For example, the follow up proteomic analysis by Zhang *et al.* [26] to the TCGA study of colorectal cancer found a mean Spearman's rank correlation coefficient between transcript and protein levels of 0.23. Similarly, a recent metabolomic study of ccRCC compared their results to the TCGA ccRCC study (albeit using different samples) and found little to no correlation between transcript and metabolite levels [27]. In another study, a follow-up metabolomic analyses of stage-dependent proteomics of ccRCC demonstrated protein abundance does not correspond to stage-dependent changes in metabolites [25, 28]. It is important to utilize the complementary nature of different techniques across domains in order to identify changes at each biological hierarchy (abundance, isoforms, post-translational modifications, *etc.*) that are driving phenotypic changes.

In the current study we utilized proteomic analysis to characterize the molecular landscape of ccRCC and interrogated changes in protein abundance and biological pathways with ccRCC stage. Moreover, we used a previously published transcriptomic data set from the same sample cohort to strengthen our conclusions. We also identified cofilin-1 (CFL1), profilin-1 (PFN1), nicotinamide N-methyltransferase (NNMT), and fructose-bisphosphate aldolase A (ALDOA), as candidate markers of late stage ccRCC. This combined proteotranscriptomic analysis not only strengthened our understanding of the underlying metabolic changes that occur in ccRCC, but also highlights probable changes in gene regulation that result in changes to the molecular phenotype of ccRCC. By utilizing data from heterogeneous biological domains, we have improved both analyses and provide the foundation for future studies of therapeutic targets.

Methods

Sample collection and storage

In January, 2000, Mayo Clinic began collecting and storing fresh-frozen patient matched normal renal and tumor tissue samples of individuals undergoing nephrectomy. The tissue samples are linked to the Nephrectomy Registry database and are also available to other investigators conducting renal cancer research. Collections of pathological specimens occurred in an IRB approved manner such that subjects could not be identified. Approximately 300 patients per year diagnosed with renal cell carcinoma undergo nephrectomy at the Mayo Clinic. Patients presenting with local tumors, locally advanced tumors, and patients with metastatic disease undergoing cytoreductive nephrectomy were candidates for this study. The gender breakdown for patients presenting with RCC is approximately 65% male and 35% female,

and the average patient age was 65 years old. Deidentified patient matched normal and ccRCC tissues were collected under a Mayo Clinic Institutional Review Board approved protocol (IRB#1746–03 first approved September 2, 2003; “Cancer Research Investigations Using Discarded Tissues”) to use tissue for molecular analyses for research purposes only (Mayo Clinic OHRP number: FWA00005001). Samples were collected from surgical resections, snap frozen in liquid nitrogen and stored at -80°C . A centralized pathology review on all tumor samples was performed to confirm tumor histologic classification, TNM stage and grade. The following definitions were used: stage I is localized ccRCC with tumor less than 7 cm; stage II is localized ccRCC with tumor >7 cm; stage III is invasive ccRCC into Gerota fascia; stage IV ccRCC is metastatic disease to a distal organ. There were 177 tissue samples used in the current study taken from 84 patients comprised of 34 stage I, 40 stage II, 42 stage III, and 52 stage IV tumor normal-adjacent pairs as well as 9 corresponding metastasis tissues.

Immunohistochemistry

A tissue microarray was created using paired normal-adjacent and tumor tissue cores from 54 patients with stage I ccRCC. The resulting formalin-fixed, paraffin-embedded block was cut into $5\ \mu\text{m}$ sections, deparaffinized, hydrated, antigen retrieved and blocked with diluent that contained background reducing components (Dako; Carpinteria, CA). The section was probed for 6-phosphofructokinase (PFKP antibody at 1:1500; LifeSpan BioSciences, Inc.; Seattle, WA) and pyruvate kinase (PKM2 antibody at 1:2500; Cell Signaling Technology; Danvers, MA). The specificity of the antibodies was confirmed using positive and negative controls from breast tumor tissue for PFKP and lung normal and tumor tissue for PKM2. Positive controls were performed and carried out to high titers to demonstrate decreased staining intensity (S1 Fig). Negative controls were performed with the absence of primary antibody (S1 Fig). For detection, the Envision Dual Labeled Polymer kit (Dako) was used according to the manufacturer’s instructions and then lightly counterstained with Gill I hematoxylin (Sigma-Aldrich) before dehydration and mounting. Images were obtained at 20X using an Aperio AT2 Scanscope (Leica Biosystems, Buffalo Grove, IL). An average IHC staining intensity score for each core was generated with triplicate measurement using an in-house Imagescope algorithm. The number of cores read varied from 36 to 48, and the values were compared using a two-sample *t*-test (Excel).

Sample preparation

Individual tissues (10 to 15 mg) were processed by trifluoroethanol (TFE) solubilization and sonication as adapted from Wang *et al.*[29]. Briefly, tissue in 50% TFE was disrupted sequentially by repeated probe sonication and heat/vortexing (1 hr, 60°C), with at least two repeats. At this step, $10\ \mu\text{g}$ of an internal protein standard, maltose-binding periplasmic protein from *Escherichia coli*, was added to the lysate. Reducing agent [tris(2-carboxyethyl)phosphine] and alkylating agent (iodoacetamide; IAA) were added, followed by overnight trypsin digestion (1:50 based on protein concentration) at 37°C . The resulting peptides were desalted with a C-18 column (1cm x 1cm), dried down by SpeedVac and reconstituted with Mobile Phase A [5% ACN, 0.1% formic acid (FA), 0.005% heptafluorobutyric acid (HFBA)]. Pooled samples were composites of five samples, and these were tumor normal-adjacent matched between the pooled samples.

Data acquisition

For the pooled samples, a fully automated five-cycle two-dimensional high-performance liquid chromatography sequence was set up as previously described [30]. Peptides were loaded onto a

7-cm Kasil fritted pre-column (150 μm inner diameter) packed with 3.5 cm of 5 μm Magic C-18 100 \AA reversed-phase material (Michrom Bioresources Inc., Auburn, CA) followed by 3.5 cm of Luna 5 μm SCX 100- \AA strong cation exchange resin (Phenomenex, Torrance, CA). Samples were loaded automatically from a 96-well microplate autosampler at 3 $\mu\text{L}/\text{min}$ using the EASY-nLC system (Thermo Scientific). The pre-column was connected to a fused silica analytical column (8 cm long, 75 μm inner diameter) via a microsplitter tee (Thermo Scientific) to which a distal 2.0 kV spray voltage was applied. The analytical column was pulled to a fine electrospray emitter using a laser puller. For peptide separation on the analytical column, a water-ACN gradient, controlled by the EASY-nLC (Thermo Scientific), was applied at an effective flow rate of 400 nL/min. Ammonium acetate salt bumps (8 μL) at concentrations of 100, 150, 200, and 500 mM were sequentially loaded, and peptides were eluted by a water-ACN gradient as described previously [30]. Sample analysis was performed on an LTQ Orbitrap XL (Thermo Scientific) using previously described instrument parameters [30]. For individual tissue sample analysis, tumor-normal pairs were analyzed by tandem mass spectrometry temporally close together. Digests were resuspended in 20 μL Mobile Phase A and a 15 μL aliquot of this peptide solution was separated on a 12 cm x 0.075 mm fused silica capillary column packed with 5 μm diameter C-18 beads (The Nest Group, Southborough, MA) across a 90 min linear gradient from 5% ACN, 0.1% FA, 0.005% HFBA to 95% ACN, 0.1% FA, 0.005% HFBA at 300 nL/min. The LC was interfaced by electrospray ionization with an LTQ (ThermoFinnigan, San Jose, CA). Data-dependent analysis was used to perform MS/MS on the five most intense ions between $m/z = 400$ and 2000 in each MS spectra with a minimum signal of 1000 cps. Dynamic exclusion was used with a repeat count of two and an exclusion duration of 180 s.

Data processing

Raw data from pooled sample tandem mass spectrometry runs were converted to mzXML files using ReAdW (v1.1) and searched against the Human UniProtKB SwissProt database (2011_3 release; 20,227 sequences plus the addition of *E. coli* maltose-binding periplasmic protein, malE, P0AEX9) using X!Tandem (CYCLONE v2011.12.01.1), OMSSA (v2.1.8), and MyriMatch (v2.1.97) search algorithms. The search was conducted with a fragment ion mass tolerance of 0.40 Da and a parent ion tolerance of 10 ppm. Complete tryptic digestion was assumed with one allowed missed cleavage site. Methionine oxidation was specified as a variable modification and alkylation of cysteine with IAA as a static modification. For protein inference minimization, an in-house grouping scheme was applied, reporting only proteins with substantial peptide information [31]. Target/decoy searches were performed to experimentally estimate the protein false discovery rate, which was determined to be <1%. Protein identifications with at least two unique tryptic peptides were considered [31]. The mass spectrometry data have been deposited to MassIVE (MSV000079511).

Raw data from individual sample tandem mass spectrometry runs were converted to peak list (mgf format) using MSConvert (ProteoWizard 3.0.4243; Jan 3, 2013). The default parameters were used, including 'Prefer Vendor for Peak Picking'. These mgfs were searched with Mascot (v2.4.1; Matrix Sciences) using the following parameters: trypsin as the enzyme with a maximum of two mis-cleavages; 1+, 2+, and 3+ charged peptides; carbamidomethyl (C) as a fixed modification, and protein N-term acetylation, deamidation (NQ) and oxidation (M) as variable modifications; instrument type was ESI-TRAP; a precursor tolerance of 2 Da and fragment ion tolerance 0.5 Da. These thresholds were chosen based on a test analysis which resulted in a <0.1% local FDR (above identity threshold) while maintaining the highest number of protein hits. The database used was a *Homo sapiens* database (taxon ID: 9606) retrieved from the 2013_04 release of the UniProtKB SwissProt database along with the SwissProt

varsplc database, a cRAP database (common Repository of Adventitious Proteins, v 2012.01.01; The Global Proteome Machine) and the entry *E. coli* malE (P0AEX9), resulting in 38,480 sequences. A separate search was also performed against a reversed decoy version of this database in order to calculate global FDR. Resulting .dat files were loaded into ProteoIQ (v2.3.08; NuSep, Inc; Bogart, GA) and protein results were filtered by requiring a 5% FDR and minimum of two peptides per protein identification. Following this step, there were 350,630 identified peptides (10,590 unique) belonging to 786 proteins. Three of these proteins were contaminants and malE, which were removed from the analysis resulting in 783 proteins identified experiment wide. Spectral counts were normalized according to total spectral counts per sample and exported from ProteoIQ for downstream analysis. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium [32] via the PRIDE partner repository with the dataset identifier PXD003271 and 10.6019/PXD003271.

Affymetrix Human Genome U133 Plus 2.0 array data (.CEL files) were retrieved from the NCBI Gene Expression Omnibus (GSE 53757), which contains data for the 144 arrays used by von Roemeling *et al.* [20]. Prior to analysis, a subset was created of just the 94 arrays (paired tumor normal-adjacent samples) that correspond directly to samples used in the current proteomic study. The data were processed with the *affy* R-package using RMA normalization. The array was annotated with the *hgu133plus2.db*, *annotate*, and *R2HTML* R-packages. A key linking sample IDs and mzIdentML files between this proteomic analysis and the transcriptomic analysis can be found in [S1 Table](#).

Data analysis

Proteomic data from the pooled samples were evaluated by calculating adjusted spectral counts (asp), as recently described [33]. The asp values for each protein were used to perform a Wilcoxon rank sum test using the exact method (Matlab, v8.5.0.197613; MathWorks) followed by a Benjamini-Hochberg (BH) procedure to correct for multiple hypothesis testing. Proteins were considered differentially abundant between tumor and normal-adjacent samples at BH adjusted $p < 0.05$. Normalized protein spectral count data from patient matched tumor and normal-adjacent samples were used to calculate a tumor to normal-adjacent ratio for each protein, which was \log_2 transformed. These \log_2 transformed ratios were evaluated with a moderated *t*-test (*limma* package [34]; R v3.2.1) followed by a BH procedure to correct for multiple hypothesis testing. This comparison was performed within tumor stages or by using data from all stages. Proteins were considered differentially abundant at a BH adjusted $p < 0.05$. These results were analyzed through QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN, Redwood City) by specifying human as a species, experimentally observed confidence, and not limiting the search space to a tissue or cell line. All other parameters were defaults within IPA. Pathway enrichment was performed within IPA using a Fisher's exact test (right-tailed), and pathways were considered significant at BH adjusted $p < 0.05$, and at least two proteins per pathway. Results were evaluated within stages and across all stages. Upstream regulator analysis was performed in IPA which uses a z-score algorithm, such that an activation z-score ≥ 2 is considered activated and ≤ -2 is considered inhibited. Regulator effects were predicted using the regulator effects algorithm in IPA that links upstream regulator analysis (upstream regulators) and downstream effects analysis (disease or functions) and merges networks with overlapping targets. Array data were also evaluated in IPA by using RMA normalized values and a moderated *t*-test. Probes with an absolute \log_2 fold-change ≥ 2 and BH adjusted $p < 0.001$ were used for IPA analysis ($n = 1003$) using the same parameters as the proteomic IPA analysis. These transcriptomic results were compared to identify overlapping patterns of pathway enrichment and/or upstream regulator effects. These same parameters were used to evaluate

the following comparable published proteomic data sets: Perroud *et al.*[25] Supp. Table 3 (180 proteins), Masui *et al.*[35] Supp. Table 5 (29 proteins), White *et al.*[22] Supp. Table 7 (55 proteins), and Zhao *et al.*[24] Online resource 1 (213 proteins).

Results

Molecular phenotype of ccRCC

In order to characterize proteomic changes related to ccRCC, two proteomic approaches were employed. The first approach was to catalog proteins in ccRCC tissue samples by using in-line multi-dimensional separation techniques prior to high resolution bottom-up shotgun proteomic analysis [30]. This approach was used on pooled samples from tumor or normal-adjacent tissues: four stage I tumor, two stage I normal-adjacent, two stage II tumor, two stage II normal-adjacent, two stage III tumor, two stage III normal-adjacent, two stage IV tumor and two stage IV normal-adjacent. The average Pearson's linear correlation coefficient (r) within replicates was 0.8835 demonstrating low intra-condition variability (S2 Fig). From these pooled samples, 1551 proteins were identified (S2 Table), and by comparing label-free quantification between tumor and normal-adjacent tissues, 290 were detected as differentially abundant at BH adjusted $p < 0.05$ (Wilcoxon rank sum test). Of these, 249 (85.9%) were decreased and 41 (14.1%) were increased in tumor samples. To further characterize the molecular changes involved in the progression of ccRCC, samples from tumor and normal-adjacent tissues from 84 individuals with stage I (17 pairs), II (20 pairs), III (21 pairs), and IV (26 pairs) ccRCC were evaluated individually using a shorter separation gradient than the pooled analysis. Nine individuals with stage IV ccRCC had samples from metastasized tissue in addition to tumor and normal-adjacent tissues. Using this approach, 783 proteins were identified (global protein FDR < 5%) and normalized spectral counts were calculated (S3 Table). In order to normalize data across the individual analysis data set, each patient's normal-adjacent tissue was used as a reference to calculate fold-change values for each protein in the tumor tissue (or metastasis tissue). These values were log-transformed and 344 proteins were identified as being differentially abundant (moderated t -test, BH adjusted $p < 0.05$; Fig 1A). Similar to the pooled analysis, 245 (71.2%) were decreased, and 99 (28.2%) were increased in tumor samples.

Initially the pooled and individual analysis approaches were compared by evaluating overlap of identified proteins. To directly compare the two datasets, seven isoforms were removed from the individual analysis to eliminate ambiguity in assignments, leaving 776 proteins. Of the proteins identified in the individual analysis, 663 (85.4%) were identified in the pooled analysis (S3A Fig). Next, the fold-change of each of these shared proteins in each analysis was visualized using a scatter plot (S3C Fig), and 540 (81.5%) of the proteins exhibited conserved directionality change, showing good agreement between the analyses. Then the 290 proteins that were differentially abundant in the pooled analysis were compared to the 342 proteins that were differentially abundant in the individual analysis (S3B Fig). There was considerable overlap between the analyses (169 proteins), with 58.3% and 49.4% of the differentially abundant proteins in the pooled analysis or individual analysis being shared, respectively. Overall this showed good agreement between analyses, and the individual analysis was used for further bioinformatic interrogation.

Since a subset of these tissues had previously been analyzed using gene expression arrays [20], we were interested in identifying parsimony between the analytical domains. Previously 144 samples were analyzed, and 94 samples (47 individuals) of these were directly related to the 84 individuals used in the current proteomic study. Therefore a subset analysis comparing tumor and normal-adjacent tissues was performed with just data from these 94 arrays. Of the 54,677 probes present, 37,334 had a BH adjusted $p < 0.05$ (moderated t -test; S4 Fig). Similar to

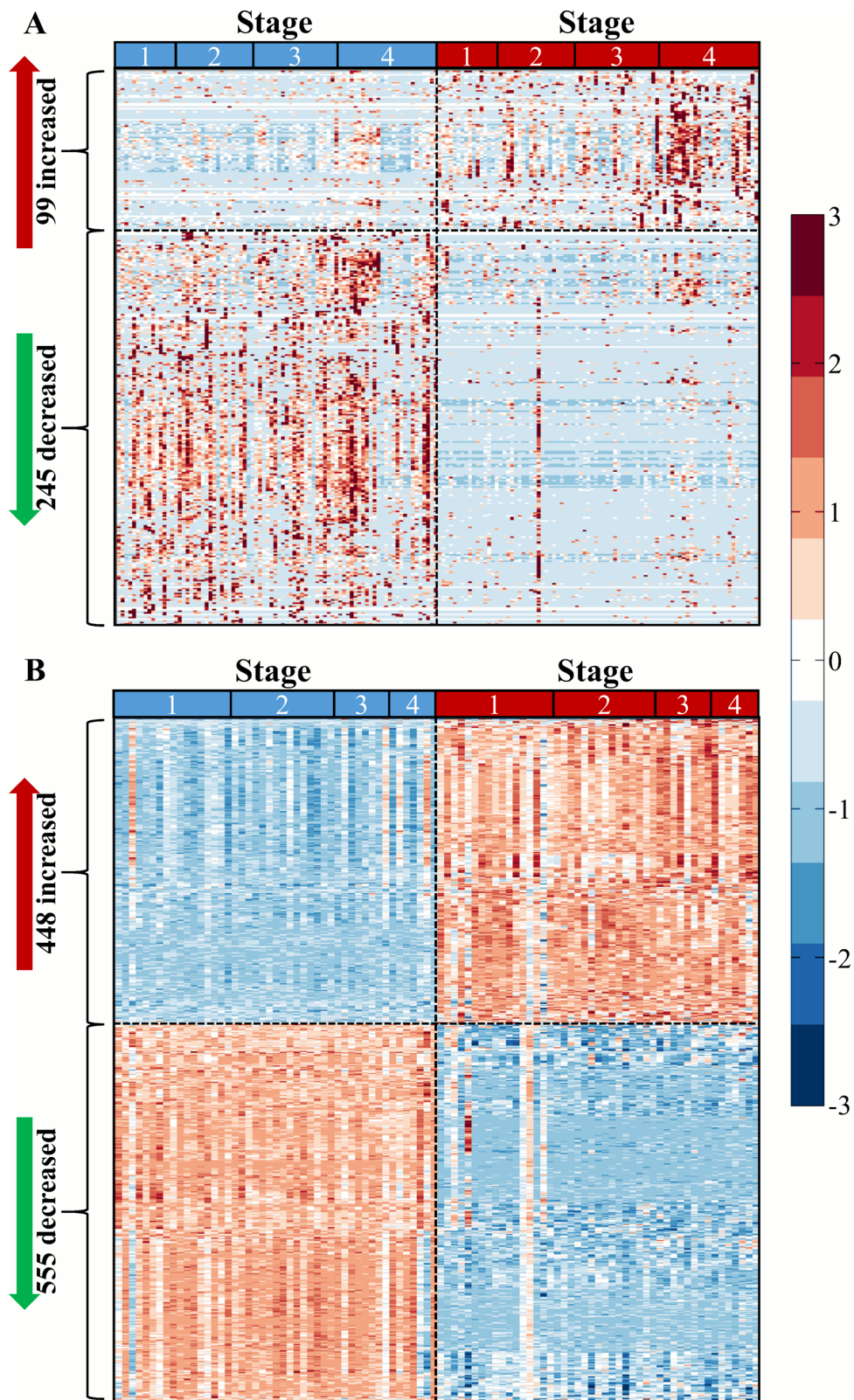


Fig 1. Differential protein abundance and gene expression between tumor and normal-adjacent ccRCC tissues. (A) Heatmap of 344 proteins with differential abundance between tumor and normal-adjacent samples (moderated *t*-test, Benjamini-Hockberg adjusted *p*-value < 0.05). (B) Heatmap of 1003

genes with differential expression between 94 tumor and normal-adjacent samples (moderated t -test BH adjusted $p < 0.001$ and absolute fold-change ≥ 4). Scale bar is standard deviation units around the mean of each protein abundance or gene expression level.

doi:10.1371/journal.pone.0154074.g001

von Roemeling *et al.* [20], we applied an additional fold-change cutoff to identify a subset of highly differentially expressed genes. Whereas von Roemeling applied a cutoff of \log_2 fold-change of 1 (resulting in 5937 genes), in the current comparison a \log_2 fold-change cutoff of 2 was used to identify 1003 differentially expressed genes between tumor and normal-adjacent tissues (Fig 1B). Genes with increased and decreased expression were approximately evenly distributed (S4 Fig).

After performing a subset analysis on the gene expression data to identify differentially expressed genes, we evaluated the correlation between protein abundance and gene transcript levels. Of the 783 proteins identified, there were 770 unique HGNC gene symbols, of which 725 overlapped with 1764 gene probes. The majority of these were positively correlated (1285; 72.8%), and the average Pearson's linear correlation coefficient (r) was 0.157 (Fig 2A). When only proteins that were differentially abundant between tumor and normal-adjacent tissue were interrogated (344, of which 318 had complementary expression data), the majority were positively correlated (301; 94.7%) with only 17 negatively correlated, and an average r of 0.347 (Fig 2B). Interestingly, 287 of the 318 probes that corresponded to differentially abundant proteins were also differentially expressed (BH adjusted $p < 0.05$). This provided evidence that important changes at the protein level were largely a reflection of signal at the transcript level, although the relatively low average correlation indicates this relationship was not linear.

Pathway enrichment analysis

Differentially abundant proteins identified in the analysis of individual samples were used to identify 88 canonical pathways that were enriched in ccRCC. The top enriched pathways were those related to metabolism (Fig 3). There was decreased abundance of proteins involved in ketolysis, the TCA cycle, ketogenesis, fatty acid β -oxidation, oxidative phosphorylation and degradation of isoleucine, valine, glutaryl-CoA, and ethanol, while glycolysis and gluconeogenesis had increased protein abundance. Pyruvate fermentation to lactate via LDH was more complicated with LDHA and LDHAL6B being increased in tumor tissues (2.55 and 0.69 \log_2 fold-change, respectively), while LDHB and LDHAL6A were decreased (-2.11 and -0.24 \log_2 fold-change, respectively). Likewise, sucrose degradation had increased levels of ALDOA, ALDOC, and TPI1 (0.90, 0.27, and 0.61 \log_2 fold-change, respectively), and decreased levels of ALDOB, KHK, and TKFC (-3.36, -0.74, and -0.90 \log_2 fold-change, respectively; TKFC was previously listed as DAK in the UniProtKB entry Q3LXA3 in S3 and S4 Tables, and was modified to TKFC in August 2015 after this analysis was completed). Also, although all the TCA cycle proteins identified were significantly decreased, the degree of change varied. Specifically, proteins responsible for transforming citrate and malate to *cis*-aconitate and oxaloacetate, respectively, were the most decreased proteins in the TCA cycle (ACO2-1.9 \log_2 fold-change and MDH2-1.6 \log_2 fold-change, respectively), supporting accumulation of these two substrates for fatty acid biosynthesis (similar to [22]).

To better understand the metabolic dysregulation in ccRCC, we evaluated protein abundance changes in glycolytic proteins (Fig 4). Contrary to previous metabolomic studies indicating the top-half of glycolysis being asynchronously increased in ccRCC [27], at the protein level there were widespread increases in glycolytic proteins. Many of these proteins were increased by relatively the same fold-change at all stages (similar to [25, 28]), although some proteins were highest at stage IV (GPI, ALDOA, ALDOC, TPI, GAPDH, PGK1, PGAM,

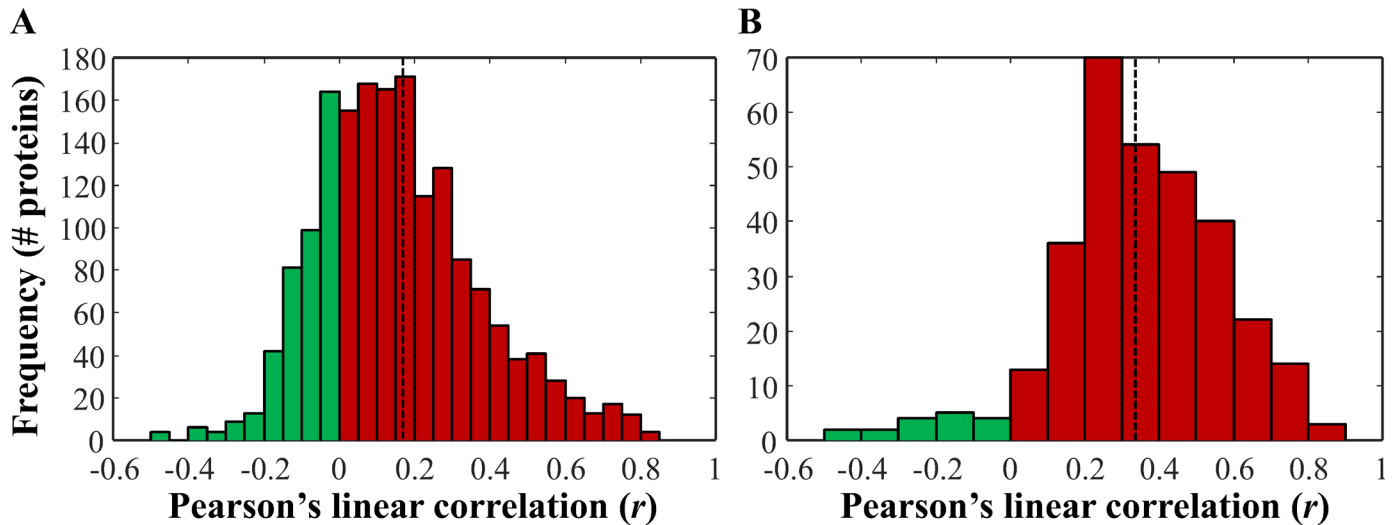


Fig 2. Correlation of differentially abundant proteins and respective gene expression levels in matched samples. (A) There were 725 proteins identified which had 1764 corresponding probes in the corresponding transcriptomic data. Pearson's linear correlation coefficient was used to correlate normalized spectral count levels and RMA normalized microarray data in matched samples (94 samples). The average Pearson's linear correlation coefficient (r) was 0.157 (dotted line). (B) Distribution of r for just the 344 differentially abundant proteins, of which 318 were also measured in the corresponding transcriptomic study. The average r was 0.347 (dotted line).

doi:10.1371/journal.pone.0154074.g002

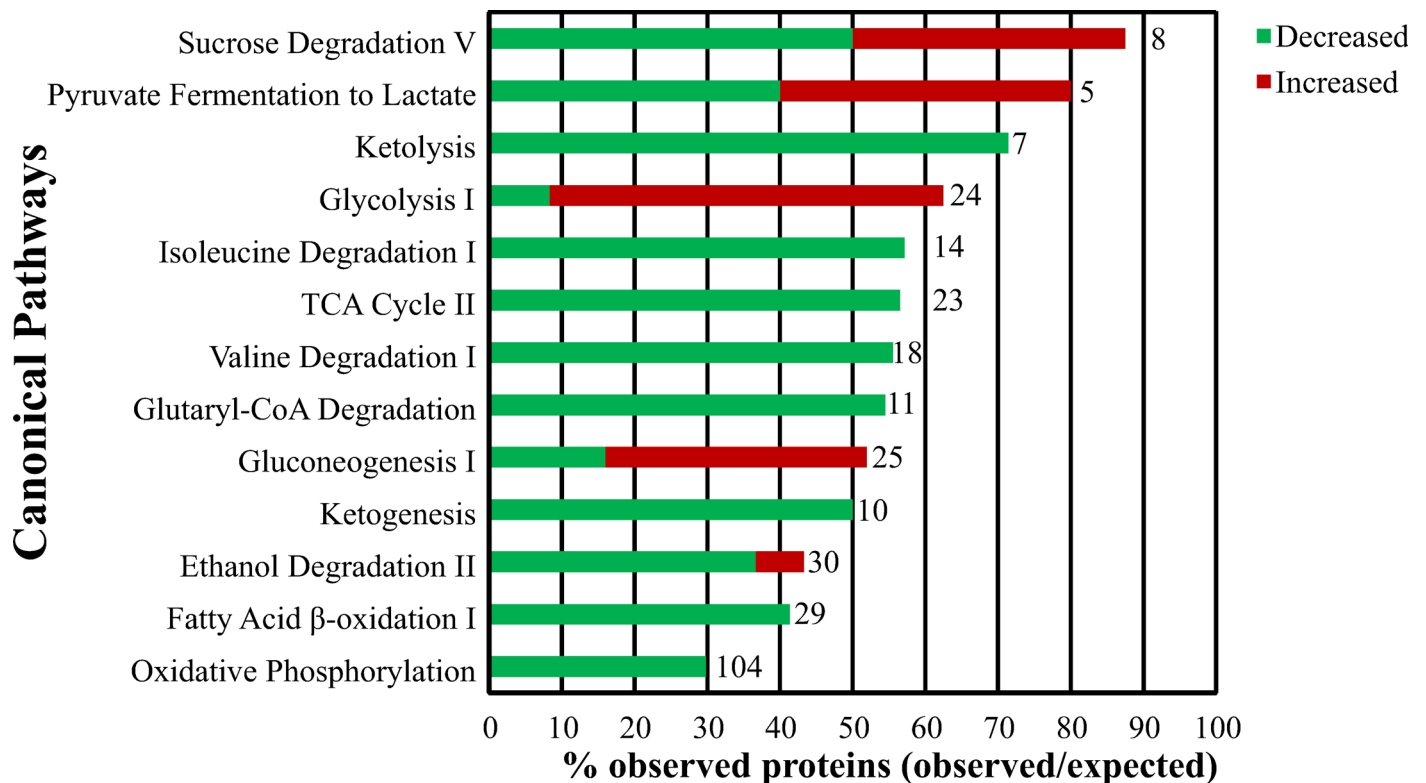


Fig 3. Enriched pathways related to metabolic dysfunction in ccRCC. Pathway enrichment analysis using IPA was performed using the 344 proteins with differential abundance between normal and tumor samples. Of the 88 pathways identified at an FDR < 5% and containing more than one protein, the following top 13 pathways are shown based on their relationship to Warburg effect related changes. The ratio of enrichment (or % observed) is further divided into those proteins increased or decreased in tumor samples. The total number of possible (or expected) proteins in each pathway is given to the right of the bar.

doi:10.1371/journal.pone.0154074.g003

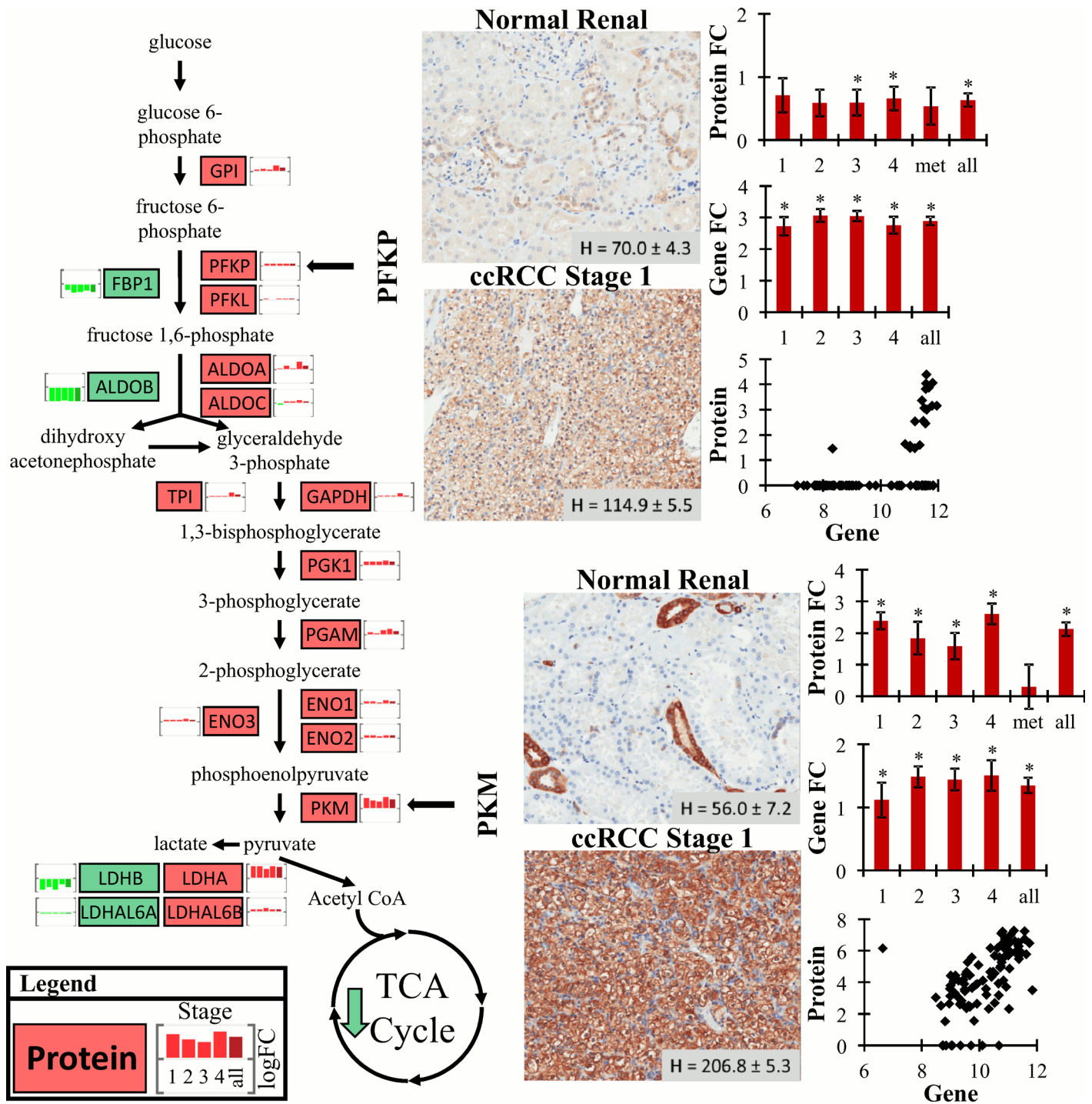


Fig 4. Protein abundance changes with stage in the glycolysis pathway. Using only proteins that were differentially abundant at each stage, the glycolysis pathway was interrogated. Directionality of change is indicated by red (increased in tumor) or green (decreased in tumor) and the small bar graph next to each protein symbol in the pathway is protein fold-change (tumor/normal) at stage I, II, III, IV and then all stages (from left to right). Levels of two proteins, PFKP and PKM, were confirmed using IHC staining of stage I ccRCC tissue (other stages were not evaluated by IHC). A representative IHC image is shown for PFKP and PKM along with average staining intensity (H value) \pm standard error of a stage I ccRCC TMA. For both PFKP and PKM the average \log_2 fold-change (FC) levels \pm standard error for protein abundance (stage I-IV and metastasis tissue) and gene expression (stage I-IV) are displayed as bar graphs (** indicates significance, BH adjusted $p < 0.05$). Below each pair of bar graphs is a scatter plot of \log_2 normalized spectral counts (protein) versus RMA normalized array gene expression data from the 94 tissues with proteomic and transcriptomic data.

doi:10.1371/journal.pone.0154074.g004

ENO1, ENO3 and PKM). Using IHC, increased levels of PFKP and PKM in stage I ccRCC were confirmed. Using a TMA of stage I ccRCC tumor and normal-adjacent tissue, PFKP was increased 1.64-fold in tumor tissue (*t*-test, *p* < 0.001), and increased 3.63-fold in tumor (*t*-test, *p* < 0.001). Expectedly, the protein immunohistochemistry levels aligned with protein abundance changes, but not transcript expression changes, though both transcript and protein levels were mildly correlated (*r* of 0.503 and 0.576 for PFKP and PKM, respectively). These results indicate increased glycolytic flux in ccRCC consistent with the Warburg effect.

Upstream target activation and inhibition

Using fold-change data for the 344 differentially abundant proteins, upstream regulator analysis was performed with IPA using data for each stage, as well as all 84 pairs, to predict activation or inhibition of regulators based on changes in protein abundance. There were six upstream targets predicted to be activated or inhibited in tumor versus normal-adjacent samples (Fig 5A). Of these, three were likely activated (activation z-score ≥ 2), estrogen-related receptor- α (ESRRA), ESRR- γ (ESRRG), and HIF1A, and three were likely inhibited (activation z-score ≤ -2), WNT1 inducible signaling pathway protein 2 (WISP2; also referred to as CCN5), FOXA1, and MAPK1. Except for ESRRG and MAPK1, the activation score was only significant when the complete data set was used (*i.e.*, “all”). To further evaluate the predicted upstream changes, the downstream stage specific protein abundance changes associated with HIF1A, ESRRA, ESRRG, WISP2, FOXA1, and MAPK1 were plotted (S5 Fig). The activated networks (HIF1A, ESRRA, ESRRG) had many proteins that were increased in tumor tissues that overlapped between the three targets, such as ENO1/ENO2, ALDOA, and LDHA.

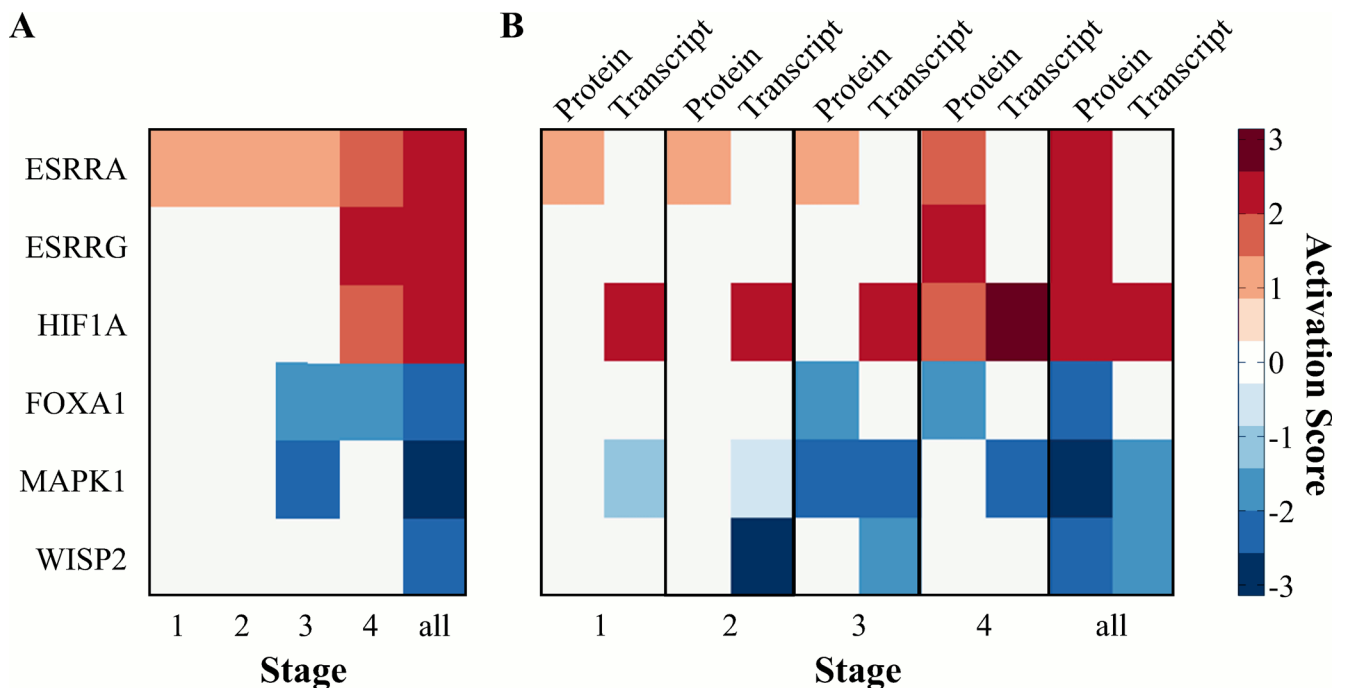


Fig 5. Upstream regulation with stage. (A) Heatmap of six upstream targets predicted to be activated/inhibited in tumor versus normal at each ccRCC stage and using all stage data together. Scale is activation z-score, with ≥ 2 being likely activation and ≤ -2 being likely inhibition. Using all stage data together, activated or inhibited upstream targets (ESRRA, ESRRG, HIF1A, FOXA1, MAPK1, and WISP2) are shown. (B) Heatmap of six upstream targets predicted to be activated/inhibited in tumor versus normal-adjacent tissues using the proteomic data set. Corresponding activation z-scores from transcriptomic data analysis are included to demonstrate conserved trends at each ccRCC stage and using all stage data together.

doi:10.1371/journal.pone.0154074.g005

Next, the transcriptomic results were subjected to the same analysis in order to evaluate whether the proteomic based activation/inhibition predictions were supported at the transcript level. Only data for the six upstream targets identified with the proteomic dataset were compared to the activation z-scores calculated using the proteomic data (Fig 5B). Of the six upstream targets found to be activated/inhibited based on the proteomic data, only HIF1A, WISP2, and MAPK1 were supported by the analysis based on transcriptomic data (the other three upstream targets were absent from the transcriptomic based upstream regulator analysis results). Interestingly, for all three of these targets, activation/inhibition is predicted by the transcript data at an earlier ccRCC stage when using transcriptomic data. Specifically, HIF1A is predicted to be activated at all stages using transcript data, while using protein data, HIF1A is only predicted to be activated when all the data is used. Similarly, WISP2 is predicted to be inhibited at stage II and III when using the transcriptomic data, while only the full protein data set predicted it to be inhibited. On the other hand, MAPK1 is predicted to be inhibited at stage III and IV based on the transcript data, which aligns well with predictions based on the protein data. The agreement between the two analyses provided increased confidence in the proteomic results and highlights potential therapeutic targets of ccRCC.

Candidate markers of aggressive ccRCC

In addition to evaluating the proteomic and transcriptomic data to better understand the systematic molecular changes in ccRCC, we were also interested in whether there were proteins that could be used to discriminate stage IV ccRCC specifically. Of the 344 differentially abundant proteins, 50 were significantly different at stage IV only. To evaluate the discriminatory power of these 50 proteins, receiver operator characteristic (ROC) curves were constructed and area under the curve was estimated (AuROC). Of the 50 proteins, four had an AuROC > 0.7: cofilin-1 (CFL1), profilin-1 (PFN1), nicotinamide N-methyltransferase (NNMT), and fructose-bisphosphate aldolase A (ALDOA). All four proteins were increased at stage IV (Fig 6), yet CFL1, PFN1, and ALDOA were decreased at early stages relative to normal tissue. In the case of NNMT and ALDOA these trends were also reflected in metastasis tissues, while CFL1 and PFN1 levels were only slightly increased in metastasis tissues. These results highlight the dynamic nature of the molecular progression of ccRCC in addition to identifying candidate markers of ccRCC aggressiveness.

Discussion

We present here the results of a comprehensive proteomic analysis of ccRCC tissues grouped by histopathology stages along with proteotranscriptomic analysis using previously reported gene expression array data from the same tissues. The goal of this analysis was not only to catalog proteins responsible for the molecular phenotype of ccRCC, but also evaluate stage dependent changes that reflect the molecular pathogenesis of ccRCC progression. These results confirm previous studies characterizing metabolic dysfunction in ccRCC [19, 21, 22, 24, 27, 28, 35], while also highlighting upstream gene targets that are predicted to be activated/inhibited using multi-domain analysis. Utilizing data across heterogeneous biological domains of the same samples not only strengthened these conclusions but also demonstrated the complementary nature of such an approach. Additionally, general ccRCC candidate biomarkers were observed, as well as stage specific markers related to high mortality metastatic ccRCC. These results highlight the benefits of a combined proteotranscriptomic approach and lay the foundation for future studies to confirm candidate therapeutic targets.

A hallmark feature of ccRCC is that cells undergo a metabolic shift consistent with the Warburg effect. These changes have been confirmed at the gene transcript level in ccRCC [19], and

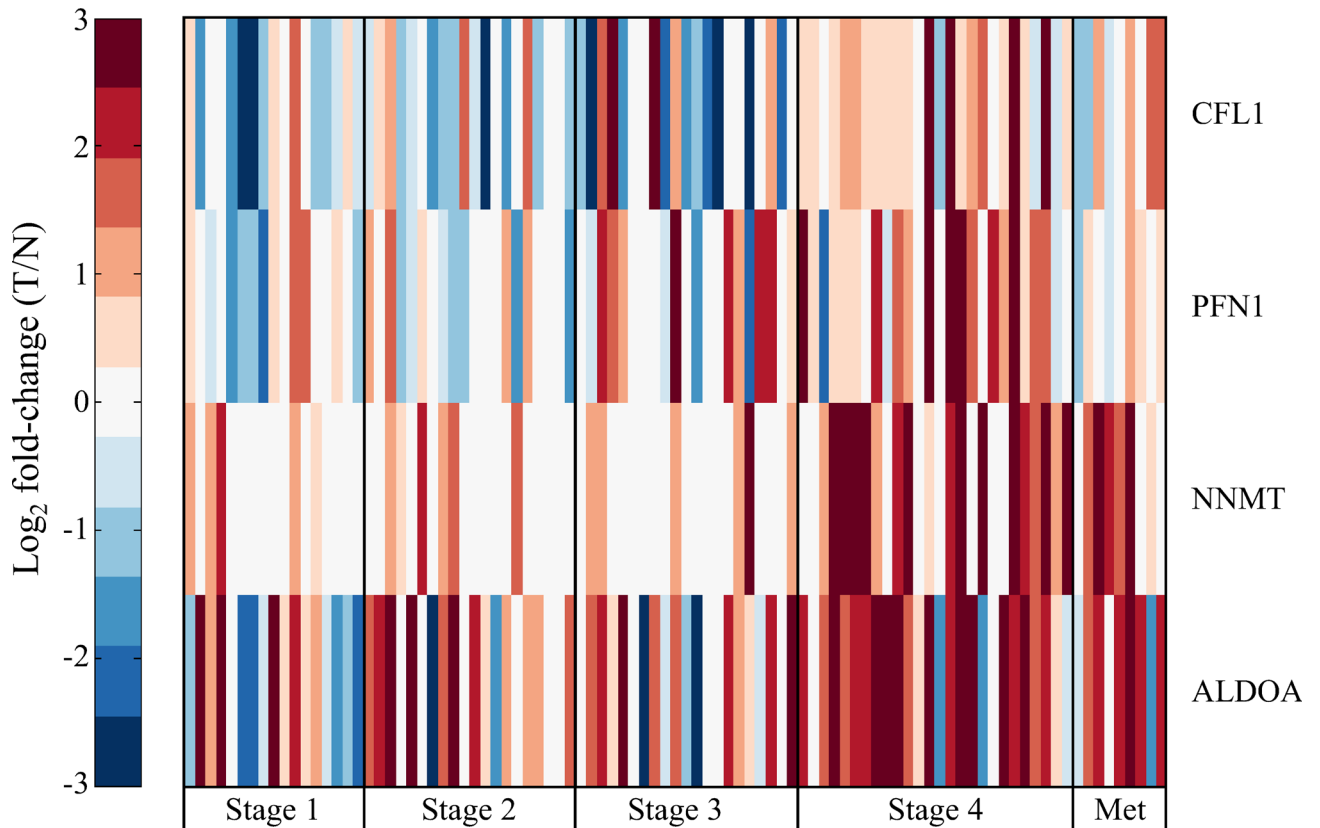


Fig 6. Candidate markers of advanced stage ccRCC. Heatmap of log₂ fold-change in protein abundance of four candidate markers of late stage ccRCC, cofilin-1 (CFL1), profilin-1 (PFN1), nicotinamide N-methyltransferase (NNMT), and fructose-bisphosphate aldolase A (ALDOA), in paired tumor and normal-adjacent tissues from 84 individuals, as well as 9 pairs that also included metastasis tissue.

doi:10.1371/journal.pone.0154074.g006

recently a series of proteomic analyses have confirmed similar widespread effects at the protein level [22–25, 35]. These studies identified between 770 and 1900 proteins with 30 to 350 differentially abundant between normal and tumor ccRCC tissues. Similarly, in the current study using a pooled sample approach we identified 1551 proteins, 290 of which were differentially abundant, while using individual samples we identified 783 proteins, 344 of which were differentially abundant. Using pathway enrichment analysis metabolic dysfunction was evident in ccRCC at all stages, as well as pathways involved in biosynthesis. Using data from recent proteomic studies of ccRCC [22, 24, 25, 35], similar pathway enrichment patterns were observed (S5 Table). In the current study we found that the key glycolytic enzyme PFKP was increased in all stages of ccRCC by both proteomic analyses and IHC. The companion transcriptomic analysis and prior proteomic studies [24, 25] also identified PFKP as being increased in ccRCC. This protein regulates an important control point in glycolysis and is an essential enzyme to drive glycolysis despite downstream feedback (as reviewed by [14]). Additionally, we observed increased PKM at all stages by both proteomic analyses and IHC, similar to other studies [22, 24, 25, 28, 35], which is another key change driving pyruvate production, highlighting potential roles PKM can play in tumor progression (as reviewed by [36]). In the current study we also observed decreased abundance of TCA cycle proteins, similar to results at the transcript level [19], specifically MDH2 and ACO2, which is similar to other proteomic studies [22, 24, 25, 35]. A recent metabolomic study found increased levels of citrate, *cis*-aconitate and succinate in ccRCC [27]. This could be partly explained by significantly decreased ACO2 protein

abundance, but highlights that discrepancies between metabolite and proteotranscriptomic or proteomic and transcriptomic results should be investigated by acquiring data across domains in parallel on the same sample set. Overall, these results provide detailed empirical support for metabolic dysregulation and the Warburg effect in ccRCC and will help to improve our understanding of the underlying metabolic changes associated with ccRCC.

Relative to other cancers, VHL inactivation is relatively specific to ccRCC and is both critical to ccRCC evolution [37] and common across ccRCC [6, 7]. Inactivation of VHL leads to the stabilization of HIF1A and HIF2A (also referred to as EPAS1) [8, 9], with the former being responsible for many downstream changes related to the Warburg effect [13]. In the current study, HIF1A activation was evident by proteomic analysis of all tissue samples, but not within stages. In contrast to this, the companion transcriptomic data indicated increasing HIF1A activation with ccRCC stage, though HIF1A mRNA levels were significantly decreased 1.25-fold in tumor tissue. Alternatively, in the transcriptomic data HIF2A expression was significantly increased 1.4-fold and was predicted to be activated in stage IV ccRCC samples (activation scores increased with stage from 1.3 at stage I to 2.2 at stage IV). This apparent HIF2A activation is supported by the analysis of the proteomic data that indicated activation of ESRRA and ESRRG. Interestingly, although there is relatively low overlap between differentially abundant proteins detected in the current study and previous proteomic analyses [22, 24, 25, 35] (7 to 38%; S6 Fig), upstream analysis of these studies indicates ESRRA and ESRRG activation (S6 Table). Both of these nuclear receptors have increased expression in breast cancer [38] and ovarian cancer [39], while in neuroblastoma ESRRA expression has been shown to increase HIF2A expression and correlates with poor survival [40]. Given the additional involvement of ESRRA and ESRRG predicted by the proteomic data, HIF2A likely plays a key role in ccRCC, especially in later stages. A survey of 160 ccRCC tumors found that VHL-deficient tumors expressed either HIF1A and HIF2A or just HIF2A, and that tumors expressing only HIF2A had higher c-Myc activity and higher stage [9], correlating with the known proliferative effects of c-Myc in ccRCC [41]. Clearly the VHL/HIF1A/HIF2A axis is important in ccRCC development and proliferation, and these results highlight the potential involvement of ESRRA and ESRRG in driving these changes.

In addition to identifying upstream targets related to HIF1A and HIF2A, the proteomic data analysis predicted WISP2 (also referred to as CCN5) inhibition in ccRCC tissues. Using the transcriptomic data, WISP2 was predicted to be activated at stage II, and was significantly increased at the transcript level at stage II and III (1.5- and 2.7-fold, respectively). At the protein level, WISP2 activation was supported by increased vimentin and fibronectin levels, which have been previously reported in ccRCC [21, 25, 42, 43] (S6 Table), and decreased keratin 8 levels. In gastric [44], breast [45] and pancreatic cancer [46], WISP2 is a known tumor suppressor and likely regulates invasion and motility through MMPs [47]. The involvement of WISP-2 in ccRCC has not been reported, and additional targeted experiments should be performed to confirm its possible role in ccRCC progression. Overall, these results highlight the complementary nature of proteomic and transcriptomic analyses when used in conjunction to identify and confirm molecular changes.

To date there have been numerous genomic, transcriptomic, and proteomic analyses of ccRCC tissue, but there has not been a study that interrogates multiple biological domains of the same samples. In a prior study, similar ccRCC tissues were analyzed by cDNA arrays and 2D-GE, identifying 119 and 334 significantly different genes and proteins, respectively, with an overlap of only 12 genes/proteins [48]. Companion proteomic studies to the seminal TCGA studies are being published, such as the follow up proteomic analysis of colorectal cancer by Zhang *et al.* [26]. Similar to the analysis presented in the current study, they evaluated the correlation of mRNA levels and protein abundance and found the mean Spearman's rank

correlation coefficient of 0.23, which is similar to our results (average $r = 0.157$) and other similar study designs [49, 50]. This low level of correlation is not entirely surprising since it is understood that the transcriptome and proteome are not linearly related [49]. For example, in the Zhang *et al.* proteomic study only 60% of a small group genes of interest identified at the mRNA level were also significant at the protein level. In the current study we utilized the transcriptomic data to support changes seen in the proteomic data, but, as is evident with the ESRRRA and ESRRG findings (*i.e.*, activation is not predicted using transcriptome data), changes at the protein level may not be directly mirrored at the transcriptome level. Protein changes are likely more indicative of broader phenotypic changes, and also likely represent a non-linear composite effect of many upstream message level changes. Using data from both biological domains of the same samples has improved the confidence of our findings and may serve to limit false discoveries in future confirmation experiments.

Early diagnosis of ccRCC significantly improves patient outcomes, and for this reason there have been numerous studies looking for transcriptomic, peptidomic, proteomic, miRNA, and metabolomic signatures of late stage ccRCC in tissue, serum, and urine (reviewed by [51–53]). To date there are no confirmed biomarkers that can be used for screening late-stage ccRCC, though there are many candidate markers of ccRCC relative to normal tissue. Many of these markers were detected in the current study, such as MnSOD [24, 54] and vimentin [55] (to name a few), but our focus was to investigate candidate markers specific to late-stage ccRCC. After eliminating non-stage specific markers of ccRCC, four proteins were identified that were specific to stage IV. Three of these proteins have been confirmed by western blotting and/or IHC in ccRCC. Cofilin (CFL1) has been confirmed to be increased in ccRCC [54] but has not been confirmed in late-stage ccRCC, though it is known to be associated with metastasis in many solid tumors [56–58]. Profilin (PFN1) has been shown to be increased in metastatic ccRCC by IHC [35] and is also a candidate marker of bladder cancer metastasis [59], though it is also down-regulated in numerous other cancers (as discussed by [35]). Nicotinamide N-methyltransferase (NNMT) has recently been shown to be an interesting candidate marker of aggressive ccRCC by two recent studies: Lebdai *et al.* demonstrated NNMT overexpression by western blotting in ccRCC tissues with high SSIGN scores [23], while Zaravinos *et al.* identified NNMT following a large meta-analysis of five published transcriptomic data sets and confirmed overexpression by IHC in ccRCC tissues [60]. Our data also indicate fructose-bisphosphate aldolase A (ALDOA) is a candidate marker of late-stage ccRCC, which has been observed in other ccRCC studies, but not specifically as a marker of late-stage ccRCC. ALDOA has been identified as a marker of late-stage colorectal cancer [61] and lung squamous cell carcinoma [62], and is involved in osteosarcoma metastasis [63]. Further studies are required to confirm that increased abundance of these four proteins correlate specifically to late-stage ccRCC.

One of the most defining traits of ccRCC is late-stage tumor diversity [37] along with intra-tumor heterogeneity [64]. Developing molecular tools to subtype ccRCC beyond TNM staging are key, but also developing ways to accurately evaluate a heterogeneous genomic, and therefore molecular, landscape are key. Based on the results herein, it is evident that there are many shared traits among ccRCC stages related to the Warburg effect, but it seems likely that there are major differences between tumor stage related to the VHL/HIF1A/HIF2A axis that likely affect tumor aggressiveness and treatment. Developing study designs that focus on stratified samples (by VHL and/or HIF1A/HIF2A status) along with multiple samples from tumors and a multi-domain approach (such as proteotranscriptomic) is necessary to more clearly define the molecular pathogenesis of ccRCC and identify new therapeutic targets that are required for metastatic ccRCC.

Supporting Information

S1 Fig. Positive and negative controls of immunohistochemical staining. (A-C) Controls for 6-phosphofructokinase (PFKP antibody): (A) Negative control on breast cancer tissue. (B) Positive control at 1:1000 and (C) 1:2000 on breast cancer tissue. (D-F) Controls for and pyruvate kinase (PKM2 antibody): (D) Negative control on normal lung tissue. (E) Positive control at 1:2500 and (F) 1:3000 on lung cancer tissue. Scale bar is 200 μm .
(TIF)

S2 Fig. Replicate variability within pooled analysis. Variability between biological replicates in the pooled analysis was visualized using scatter plots of adjusted spectral counts for the 1551 proteins identified, and the Pearson's linear correlation coefficient (r) was calculated for each pair. There from four stage I tumor (t1_#), two stage I normal-adjacent (n1_#), two stage II tumor (t2_#), two stage II normal-adjacent (n2_#), two stage III tumor (t3_#), two stage III normal-adjacent (n3_#), two stage IV tumor (t4_#) and two stage IV normal-adjacent (n4_#) pooled samples. The axis identifiers are the same used in [S2 Table](#). The average r within replicates was 0.8835.
(TIF)

S3 Fig. Comparison of proteomic results from pooled and individual samples. (A) Overlap between the proteins identified in the pooled and individual sample analysis. (B) Comparison of directionality agreement between the 663 overlapping proteins using \log_2 fold-change (tumor/normal-adjacent). There were 123 proteins changing in different directions but these differences were not dramatic. (C) Overlap of differentially abundant proteins identified using each approach.
(TIF)

S4 Fig. Distribution of mRNA expression levels across all stages. Data from 94 arrays with 54,677 probes was RMA normalized and evaluated using a moderated t -test comparing all tumor normal-adjacent pairs. RMA normalized expression values were plotted against Benjamini-Hockberg (BH) adjusted p -values. Emphasized in red are probes with BH adjusted $p < 0.001$ and absolute fold-change ≥ 4 (1003 genes).
(TIF)

S5 Fig. Predicted upstream effects. Using protein abundance data from all four stages, HIF1A, ESRRG, and ESRRR were predicted to be activated and WISP2, FOXA1, and MAPK1 were predicted to be inhibited. Bar graphs are four stage and all stage comparisons. Image generated with QIAGEN's Ingenuity® Pathway Analysis.
(TIF)

S6 Fig. Similarity of other shotgun proteomics studies with the current proteotranscriptomic study. Differentially abundant proteins and genes from this study (pooled, individual and 94 array) were compared to differentially abundant proteins identified in four other studies (granular proteomic results from Lebdaï *et al.*[23] were not available): Perroud *et al.*[25] Supp. Table 3 (180 proteins), Masui *et al.*[35] Supp. Table 5 (29 proteins), White *et al.*[22] Supp. Table 7 (55 proteins), and Zhao *et al.*[24] Online resource 1 (213 proteins). Two separate list of unique gene symbols for proteins or genes that increased ($n = 458$) or decreased ($n = 719$) in ccRCC tissue across data sets were created, and a binary matrix of presence/absence was constructed for these across experiments. A dendrogram for each list was created using the unweighted pair group method with arithmetic mean method (euclidian distance; Matlab). Specifically, of the increasing proteins the overlap of each study with the individual analysis presented herein was 28.4%, 10.5%, 12.6%, and 21.1% for Perroud, Masui, White, and Zhao

respectively, and of the decreasing proteins the overlap of each study with the individual analysis was 35.7%, 7.0%, 8.2%, and 37.7% respectively. The overlap between the individual and pooled analysis described in [S3 Fig](#) is due to the 59% overlap in decreasing proteins resulting in their proximity on the dendrogram.

(TIF)

S1 Table. Data organization key. The tissues analyzed in this study were also part of two previous expression array studies. The proteomic data has been deposited to the ProteomeX-change Consortium via the PRIDE partner repository with the dataset identifier PXD003271 and 10.6019/PXD003271. The PRIDE repository assigns assay numbers, but maintains mzIdentML filenames (.mzid) for each run. These correspond directly to the array data that was deposited on the NCBI Gene Expression Omnibus (GSE 53757). These.cel files have been linked to the respective proteomic mzIdentML files.

(XLSX)

S2 Table. Proteomic analysis of pooled samples. Proteomic data from the pooled samples were evaluated by using the \log_2 values for each protein to perform a Wilcoxon rank sum test using the exact method followed by a Benjamini-Hochberg (BH) procedure to correct for multiple hypothesis testing. Whether the sample is from pooled tumor tissues or normal-adjacent is indicated by a 't' or 'n', respectively, followed immediately by a number which refers to stage. Fold-change (FC) values were determined by taking the \log_2 of each spectral count (in case of zero, 0.5 was used), and \log_2 fold-change was the difference in averages.

(XLSX)

S3 Table. Proteomic analysis of individual samples. Normalized spectral count values for the 783 proteins identified across 177 samples is given along with sample information and protein information.

(XLSX)

S4 Table. Statistical analysis of fold-change protein differences between tumor and normal-adjacent pairs. For each of the 783 proteins identified in the individual analysis \log_2 fold-change is given as well as p -value (moderated t -test) and BH adjusted p -value for each stage and using all samples.

(XLSX)

S5 Table. Pathway enrichment analysis of previously published proteomic analyses of ccRCC. Differentially abundant proteins from this study (pooled and individual analyses) were compared to differentially abundant proteins identified in four other studies (granular proteomic results from Lebdai *et al.* [23] were not available): Perroud *et al.* [25] Supp. Table 3 (180 proteins), Masui *et al.* [35] Supp. Table 5 (29 proteins), White *et al.* [22] Supp. Table 7 (55 proteins), and Zhao *et al.* [24] Online resource 1 (213 proteins). Pathway enrichment analysis was performed using IPA as described in the current study and the top 15 pathways with more than one molecule and $p < 0.05$ are given (ranked by p -value).

(XLSX)

S6 Table. Upstream regulator analysis of previously published proteomic analyses of ccRCC. Differentially abundant proteins from this study (pooled and individual analyses) were compared to differentially abundant proteins identified in four other studies (granular proteomic results from Lebdai *et al.* [23] were not available): Perroud *et al.* [25] Supp. Table 3 (180 proteins), Masui *et al.* [35] Supp. Table 5 (29 proteins), White *et al.* [22] Supp. Table 7 (55 proteins), and Zhao *et al.* [24] Online resource 1 (213 proteins). Upstream regulator analysis was performed using IPA as described in the current study and the top 15 targets are given (ranked

by p -value). The activation z-score is given and an activation z-score ≥ 2 is considered activated and ≤ -2 is considered inhibited. The six targets identified in the current study as being activated or inhibited (ESRRA, ESRRG, HIF1A, FOXA1, MAPK1, and WISP2; Fig 5) are highlighted in bold or listed in bolded italics below the top 15 targets if not included in the top 15 (ranked by p -value). Also, EPAS1 (also referred to as HIF2A) is listed if present since its activation is implied by ESRRA/ESRRG activation and the transcriptomic analysis in the current study.
(XLSX)

Author Contributions

Conceived and designed the experiments: DM HS MS JON TK JAC RRD. Performed the experiments: CEW LAM YK AI JON TK JAC RRD. Analyzed the data: BAN AI TK DM LAM HS MS JAC RRD. Contributed reagents/materials/analysis tools: BAN AI DM HS MS JON TK JAC RRD. Wrote the paper: BAN RRD JAC AI TK LAM JON DM HS MS.

References

- Rini BI, Campbell SC, Escudier B. Renal cell carcinoma. *Lancet* (London, England). 2009; 373(9669):1119–32.
- Moch H. Kidney Cancer. In: Stewart BW, Wild CP, editors. *World Cancer Report 2014*. Lyon, France: International Agency for Research on Cancer; 2014. p. 436–43.
- King SC, Pollack LA, Li J, King JB, Master VA. Continued increase in incidence of renal cell carcinoma, especially in young patients and high grade disease: United States 2001 to 2010. *J Urol*. 2014; 191(6):1665–70. doi: [10.1016/j.juro.2013.12.046](https://doi.org/10.1016/j.juro.2013.12.046) PMID: [24423441](https://pubmed.ncbi.nlm.nih.gov/24423441/)
- Sun M, Lughezzani G, Jeldres C, Isbarn H, Shariat SF, Arjane P, et al. A Proposal for Reclassification of the Fuhrman Grading System in Patients with Clear Cell Renal Cell Carcinoma. *European Urology*. 2009; 56(5):775–81. doi: [10.1016/j.eururo.2009.06.008](https://doi.org/10.1016/j.eururo.2009.06.008) PMID: [19573980](https://pubmed.ncbi.nlm.nih.gov/19573980/)
- Sato Y, Yoshizato T, Shiraiishi Y, Maekawa S, Okuno Y, Kamura T, et al. Integrated molecular analysis of clear-cell renal cell carcinoma. *Nature genetics*. 2013; 45(8):860–7. doi: [10.1038/ng.2699](https://doi.org/10.1038/ng.2699) PMID: [23797736](https://pubmed.ncbi.nlm.nih.gov/23797736/)
- Kandath C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. *Nature*. 2013; 502(7471):333–9. doi: [10.1038/nature12634](https://doi.org/10.1038/nature12634) PMID: [24132290](https://pubmed.ncbi.nlm.nih.gov/24132290/)
- Nickerson ML, Jaeger E, Shi Y, Durocher JA, Mahurkar S, Zaridze D, et al. Improved identification of von Hippel-Lindau gene alterations in clear cell renal tumors. *Clin Cancer Res*. 2008; 14(15):4726–34. doi: [10.1158/1078-0432.CCR-07-4921](https://doi.org/10.1158/1078-0432.CCR-07-4921) PMID: [18676741](https://pubmed.ncbi.nlm.nih.gov/18676741/)
- Gossage L, Eisen T, Maher ER. VHL, the story of a tumour suppressor gene. *Nature reviews Cancer*. 2015; 15(1):55–64. doi: [10.1038/nrc3844](https://doi.org/10.1038/nrc3844) PMID: [25533676](https://pubmed.ncbi.nlm.nih.gov/25533676/)
- Gordan JD, Lal P, Dondeti VR, Letrero R, Parekh KN, Oquendo CE, et al. HIF- α effects on c-Myc distinguish two subtypes of sporadic VHL-deficient clear cell renal carcinoma. *Cancer Cell*. 2008; 14(6):435–46. doi: [10.1016/j.ccr.2008.10.016](https://doi.org/10.1016/j.ccr.2008.10.016) PMID: [19061835](https://pubmed.ncbi.nlm.nih.gov/19061835/)
- Mole DR, Blancher C, Copley RR, Pollard PJ, Gleadow JM, Ragoussis J, et al. Genome-wide association of hypoxia-inducible factor (HIF)-1 α and HIF-2 α DNA binding with expression profiling of hypoxia-inducible transcripts. *The Journal of biological chemistry*. 2009; 284(25):16767–75. doi: [10.1074/jbc.M901790200](https://doi.org/10.1074/jbc.M901790200) PMID: [19386601](https://pubmed.ncbi.nlm.nih.gov/19386601/)
- Schodel J, Oikonomopoulos S, Ragoussis J, Pugh CW, Ratcliffe PJ, Mole DR. High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq. *Blood*. 2011; 117(23):e207–17. doi: [10.1182/blood-2010-10-314427](https://doi.org/10.1182/blood-2010-10-314427) PMID: [21447827](https://pubmed.ncbi.nlm.nih.gov/21447827/)
- Koppenol WH, Bounds PL, Dang CV. Otto Warburg's contributions to current concepts of cancer metabolism. *Nature reviews Cancer*. 2011; 11(5):325–37. doi: [10.1038/nrc3038](https://doi.org/10.1038/nrc3038) PMID: [21508971](https://pubmed.ncbi.nlm.nih.gov/21508971/)
- Semenza GL. HIF-1 mediates the Warburg effect in clear cell renal carcinoma. *Journal of bioenergetics and biomembranes*. 2007; 39(3):231–4. PMID: [17551816](https://pubmed.ncbi.nlm.nih.gov/17551816/)
- Lu J, Tan M, Cai Q. The Warburg effect in tumor progression: mitochondrial oxidative metabolism as an anti-metastasis mechanism. *Cancer letters*. 2015; 356(2 Pt A):156–64. doi: [10.1016/j.canlet.2014.04.001](https://doi.org/10.1016/j.canlet.2014.04.001) PMID: [24732809](https://pubmed.ncbi.nlm.nih.gov/24732809/)

15. DeBerardinis RJ, Mancuso A, Daikhin E, Nissim I, Yudkoff M, Wehrli S, et al. Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104(49):19345–50. PMID: [18032601](#)
16. Frew IJ, Moch H. A clearer view of the molecular complexity of clear cell renal cell carcinoma. *Annual review of pathology*. 2015; 10:263–89. doi: [10.1146/annurev-pathol-012414-040306](#) PMID: [25387056](#)
17. Hu C-J, Wang L-Y, Chodosh LA, Keith B, Simon MC. Differential Roles of Hypoxia-Inducible Factor 1 α (HIF-1 α) and HIF-2 α in Hypoxic Gene Regulation. *Molecular and Cellular Biology*. 2003; 23(24):9361–74. PMID: [14645546](#)
18. Raval RR, Lau KW, Tran MGB, Sowter HM, Mandriota SJ, Li J-L, et al. Contrasting Properties of Hypoxia-Inducible Factor 1 (HIF-1) and HIF-2 in von Hippel-Lindau-Associated Renal Cell Carcinoma. *Molecular and Cellular Biology*. 2005; 25(13):5675–86. PMID: [15964822](#)
19. The Cancer Genome Atlas Research N. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature*. 2013; 499(7456):43–9. doi: [10.1038/nature12222](#) PMID: [23792563](#)
20. von Roemeling CA, Marlow LA, Radisky DC, Rohl A, Larsen HE, Wei J, et al. Functional genomics identifies novel genes essential for clear cell renal cell carcinoma tumor cell proliferation and migration. *Oncotarget*. 2014; 5(14):5320–34. PMID: [24979721](#)
21. Tun HW, Marlow LA, von Roemeling CA, Cooper SJ, Kreinest P, Wu K, et al. Pathway Signature and Cellular Differentiation in Clear Cell Renal Cell Carcinoma. *PloS one*. 2010; 5(5):e10696. doi: [10.1371/journal.pone.0010696](#) PMID: [20502531](#)
22. White NM, Masui O, Desouza LV, Krakovska O, Metias S, Romaschin AD, et al. Quantitative proteomic analysis reveals potential diagnostic markers and pathways involved in pathogenesis of renal cell carcinoma. *Oncotarget*. 2014; 5(2):506–18. PMID: [24504108](#)
23. Lebdaï S, Verhoest G, Parikh H, Jacquet SF, Bensalah K, Chautard D, et al. Identification and validation of TGFBI as a promising prognosis marker of clear cell renal cell carcinoma. *Urologic oncology*. 2015; 33(2):69.e11–8.
24. Zhao Z, Wu F, Ding S, Sun L, Liu Z, Ding K, et al. Label-free quantitative proteomic analysis reveals potential biomarkers and pathways in renal cell carcinoma. *Tumour biology: the journal of the International Society for Oncodevelopmental Biology and Medicine*. 2015; 36(2):939–51.
25. Perroud B, Ishimaru T, Borowsky AD, Weiss RH. Grade-dependent Proteomics Characterization of Kidney Cancer. *Molecular & cellular proteomics: MCP*. 2009; 8(5):971–85.
26. Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z, et al. Proteogenomic characterization of human colon and rectal cancer. *Nature*. 2014; 513(7518):382–7. doi: [10.1038/nature13438](#) PMID: [25043054](#)
27. Hakimi AA, Reznik E, Lee CH, Creighton CJ, Brannon AR, Luna A, et al. An Integrated Metabolic Atlas of Clear Cell Renal Cell Carcinoma. *Cancer Cell*. 2016; 29(1):104–16. doi: [10.1016/j.ccell.2015.12.004](#) PMID: [26766592](#)
28. Wettersten HI, Hakimi AA, Morin D, Bianchi C, Johnstone ME, Donohoe DR, et al. Grade-Dependent Metabolic Reprogramming in Kidney Cancer Revealed by Combined Proteomics and Metabolomics Analysis. *Cancer research*. 2015; 75(12):2541–52. doi: [10.1158/0008-5472.CAN-14-1703](#) PMID: [25952651](#)
29. Wang H, Qian WJ, Mottaz HM, Clauss TR, Anderson DJ, Moore RJ, et al. Development and evaluation of a micro- and nanoscale proteomic sample preparation method. *Journal of proteome research*. 2005; 4(6):2397–403. PMID: [16335993](#)
30. Taylor P, Nielsen PA, Trelle MB, Horning OB, Andersen MB, Vorm O, et al. Automated 2D peptide separation on a 1D nano-LC-MS system. *Journal of proteome research*. 2009; 8(3):1610–6. doi: [10.1021/pr800986c](#) PMID: [19178303](#)
31. Gortzak-Uzan L, Ignatchenko A, Evangelou AI, Agochiya M, Brown KA, St Onge P, et al. A proteome resource of ovarian cancer ascites: integrated proteomic and bioinformatic analyses to identify putative biomarkers. *Journal of proteome research*. 2008; 7(1):339–51. PMID: [18076136](#)
32. Vizcaïno JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Ríos D, et al. ProteomeXchange provides globally co-ordinated proteomics data submission and dissemination. *Nat Biotechnol*. 2014; 32(3):223–6. doi: [10.1038/nbt.2839](#) PMID: [24727771](#)
33. Sepiashvili L, Waggott D, Hui A, Shi W, Su S, Ignatchenko A, et al. Integrated omic analysis of oropharyngeal carcinomas reveals human papillomavirus (HPV)-dependent regulation of the activator protein 1 (AP-1) pathway. *Molecular & cellular proteomics: MCP*. 2014; 13(12):3572–84.
34. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology*. 2004; 3:Article3.

35. Masui O, White NM, DeSouza LV, Krakovska O, Matta A, Metias S, et al. Quantitative proteomic analysis in metastatic renal cell carcinoma reveals a unique set of proteins with potential prognostic significance. *Molecular & cellular proteomics: MCP*. 2013; 12(1):132–44.
36. Iqbal MA, Gupta V, Gopinath P, Mazurek S, Bamezai RN. Pyruvate kinase M2 and cancer: an updated assessment. *FEBS letters*. 2014; 588(16):2685–92. doi: [10.1016/j.febslet.2014.04.011](https://doi.org/10.1016/j.febslet.2014.04.011) PMID: [24747424](https://pubmed.ncbi.nlm.nih.gov/24747424/)
37. Gerlinger M, Horswell S, Larkin J, Rowan AJ, Salm MP, Varela I, et al. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nature genetics*. 2014; 46(3):225–33. doi: [10.1038/ng.2891](https://doi.org/10.1038/ng.2891) PMID: [24487277](https://pubmed.ncbi.nlm.nih.gov/24487277/)
38. Ariazi EA, Clark GM, Mertz JE. Estrogen-related receptor alpha and estrogen-related receptor gamma associate with unfavorable and favorable biomarkers, respectively, in human breast cancer. *Cancer research*. 2002; 62(22):6510–8. PMID: [12438245](https://pubmed.ncbi.nlm.nih.gov/12438245/)
39. Sun P, Sehoul J, Denkert C, Mustea A, Konsgen D, Koch I, et al. Expression of estrogen receptor-related receptors, a subfamily of orphan nuclear receptors, as new tumor biomarkers in ovarian cancer cells. *Journal of molecular medicine (Berlin, Germany)*. 2005; 83(6):457–67.
40. Hamidian A, von Stedingk K, Munksgaard Thorén M, Mohlin S, Pålman S. Differential regulation of HIF-1 α and HIF-2 α in neuroblastoma: Estrogen-related receptor alpha (ERR α) regulates HIF2A transcription and correlates to poor outcome. *Biochemical and Biophysical Research Communications*. 2015; 461(3):560–7. doi: [10.1016/j.bbrc.2015.04.083](https://doi.org/10.1016/j.bbrc.2015.04.083) PMID: [25912138](https://pubmed.ncbi.nlm.nih.gov/25912138/)
41. Tang SW, Chang WH, Su YC, Chen YC, Lai YH, Wu PT, et al. MYC pathway is activated in clear cell renal cell carcinoma and essential for proliferation of clear cell renal cell carcinoma cells. *Cancer letters*. 2009; 273(1):35–43. doi: [10.1016/j.canlet.2008.07.038](https://doi.org/10.1016/j.canlet.2008.07.038) PMID: [18809243](https://pubmed.ncbi.nlm.nih.gov/18809243/)
42. Kim HL, Seligson D, Liu X, Janzen N, Bui MHT, Yu H, et al. Using Protein Expressions to Predict Survival in Clear Cell Renal Carcinoma. *Clinical Cancer Research*. 2004; 10(16):5464–71. PMID: [15328185](https://pubmed.ncbi.nlm.nih.gov/15328185/)
43. Yokomizo A, Takakura M, Kanai Y, Sakuma T, Matsubara J, Honda K, et al. Use of quantitative shotgun proteomics to identify fibronectin 1 as a potential plasma biomarker for clear cell carcinoma of the kidney. *Cancer biomarkers: section A of Disease markers*. 2011; 10(3–4):175–83. doi: [10.3233/CBM-2012-0243](https://doi.org/10.3233/CBM-2012-0243) PMID: [22674303](https://pubmed.ncbi.nlm.nih.gov/22674303/)
44. Ji J, Jia S, Jia Y, Ji K, Hargest R, Jiang WG. WISP-2 in human gastric cancer and its potential metastatic suppressor role in gastric cancer cells mediated by JNK and PLC-gamma pathways. *British journal of cancer*. 2015; 113(6):921–33. doi: [10.1038/bjc.2015.285](https://doi.org/10.1038/bjc.2015.285) PMID: [26291058](https://pubmed.ncbi.nlm.nih.gov/26291058/)
45. Haque I, Banerjee S, De A, Maity G, Sarkar S, Majumdar M, et al. CCN5/WISP-2 promotes growth arrest of triple-negative breast cancer cells through accumulation and trafficking of p27(Kip1) via Skp2 and FOXO3a regulation. *Oncogene*. 2015; 34(24):3152–63. doi: [10.1038/onc.2014.250](https://doi.org/10.1038/onc.2014.250) PMID: [25132260](https://pubmed.ncbi.nlm.nih.gov/25132260/)
46. Dhar G, Mehta S, Banerjee S, Gardner A, McCarty BM, Mathur SC, et al. Loss of WISP-2/CCN5 signaling in human pancreatic cancer: a potential mechanism for epithelial-mesenchymal-transition. *Cancer letters*. 2007; 254(1):63–70. PMID: [17383817](https://pubmed.ncbi.nlm.nih.gov/17383817/)
47. Frewer KA, Sanders AJ, Owen S, Frewer NC, Hargest R, Jiang WG. A role for WISP2 in colorectal cancer cell invasion and motility. *Cancer genomics & proteomics*. 2013; 10(4):187–96.
48. Seliger B, Dressler SP, Wang E, Kellner R, Recktenwald CV, Lottspeich F, et al. Combined analysis of transcriptome and proteome data as a tool for the identification of candidate biomarkers in renal cell carcinoma. *Proteomics*. 2009; 9(6):1567–81. doi: [10.1002/pmic.200700288](https://doi.org/10.1002/pmic.200700288) PMID: [19235166](https://pubmed.ncbi.nlm.nih.gov/19235166/)
49. Foss EJ, Radulovic D, Shaffer SA, Ruderfer DM, Bedalov A, Goodlett DR, et al. Genetic basis of proteome variation in yeast. *Nature genetics*. 2007; 39(11):1369–75. PMID: [17952072](https://pubmed.ncbi.nlm.nih.gov/17952072/)
50. Gygi SP, Rochon Y, Franz BR, Aebersold R. Correlation between Protein and mRNA Abundance in Yeast. *Molecular and Cellular Biology*. 1999; 19(3):1720–30. PMID: [10022859](https://pubmed.ncbi.nlm.nih.gov/10022859/)
51. Rose T, Rathmell WK. Biomarkers for Renal Cell Carcinoma. In: Lara PN, Jonasch E, editors. *Kidney Cancer*: Springer International Publishing; 2015. p. 59–82.
52. Cooper SJ, Tun HW, Roper SM, Kim Y, Kislinger T, Drake RR, et al. Current Status of Biomarker Discovery in Human clear Cell Renal Cell Carcinoma. *J Mol Biomark Diagn*. 2012.
53. Moch H, Srigley J, Delahunt B, Montironi R, Egevad L, Tan PH. Biomarkers in renal cancer. *Virchows Archiv: an international journal of pathology*. 2014; 464(3):359–65.
54. Unwin RD, Craven RA, Harnden P, Hanrahan S, Totty N, Knowles M, et al. Proteomic changes in renal cancer and co-ordinate demonstration of both the glycolytic and mitochondrial aspects of the Warburg effect. *Proteomics*. 2003; 3(8):1620–32. PMID: [12923786](https://pubmed.ncbi.nlm.nih.gov/12923786/)
55. Morgan TM, Seeley EH, Fadare O, Caprioli RM, Clark PE. Imaging the clear cell renal cell carcinoma proteome. *J Urol*. 2013; 189(3):1097–103. doi: [10.1016/j.juro.2012.09.074](https://doi.org/10.1016/j.juro.2012.09.074) PMID: [23009866](https://pubmed.ncbi.nlm.nih.gov/23009866/)

56. Yang ZL, Miao X, Xiong L, Zou Q, Yuan Y, Li J, et al. CFL1 and Arp3 are biomarkers for metastasis and poor prognosis of squamous cell/adenosquamous carcinomas and adenocarcinomas of gallbladder. *Cancer investigation*. 2013; 31(2):132–9. doi: [10.3109/07357907.2012.756113](https://doi.org/10.3109/07357907.2012.756113) PMID: [23320827](https://pubmed.ncbi.nlm.nih.gov/23320827/)
57. Wang W, Mouneimne G, Sidani M, Wyckoff J, Chen X, Makris A, et al. The activity status of cofilin is directly related to invasion, intravasation, and metastasis of mammary tumors. *The Journal of cell biology*. 2006; 173(3):395–404. PMID: [16651380](https://pubmed.ncbi.nlm.nih.gov/16651380/)
58. Zhu B, Fukada K, Zhu H, Kyprianou N. Prohibitin and cofilin are intracellular effectors of transforming growth factor beta signaling in human prostate cancer cells. *Cancer research*. 2006; 66(17):8640–7. PMID: [16951178](https://pubmed.ncbi.nlm.nih.gov/16951178/)
59. Zoidakis J, Makridakis M, Zerefos PG, Bitsika V, Esteban S, Frantzi M, et al. Profilin 1 is a potential biomarker for bladder cancer aggressiveness. *Molecular & cellular proteomics: MCP*. 2012; 11(4): M111.009449.
60. Zaravinos A, Pieri M, Mourmouras N, Anastasiadou N, Zouvani I, Delakas D, et al. Altered metabolic pathways in clear cell renal cell carcinoma: A meta-analysis and validation study focused on the deregulated genes and their associated networks. *Oncoscience*. 2014; 1(2):117–31. PMID: [25594006](https://pubmed.ncbi.nlm.nih.gov/25594006/)
61. Peng Y, Li X, Wu M, Yang J, Liu M, Zhang W, et al. New prognosis biomarkers identified by dynamic proteomic analysis of colorectal cancer. *Molecular bioSystems*. 2012; 8(11):3077–88. doi: [10.1039/c2mb25286d](https://doi.org/10.1039/c2mb25286d) PMID: [22996014](https://pubmed.ncbi.nlm.nih.gov/22996014/)
62. Du S, Guan Z, Hao L, Song Y, Wang L, Gong L, et al. Fructose-bisphosphate aldolase a is a potential metastasis-associated marker of lung squamous cell carcinoma and promotes lung cell tumorigenesis and migration. *PloS one*. 2014; 9(1):e85804. doi: [10.1371/journal.pone.0085804](https://doi.org/10.1371/journal.pone.0085804) PMID: [24465716](https://pubmed.ncbi.nlm.nih.gov/24465716/)
63. Long F, Cai X, Luo W, Chen L, Li K. Role of aldolase A in osteosarcoma progression and metastasis: in vitro and in vivo evidence. *Oncology reports*. 2014; 32(5):2031–7. doi: [10.3892/or.2014.3473](https://doi.org/10.3892/or.2014.3473) PMID: [25215901](https://pubmed.ncbi.nlm.nih.gov/25215901/)
64. Tomaszewski JJ, Uzzo RG, Smaldone MC. Heterogeneity and renal mass biopsy: a review of its role and reliability. *Cancer Biol Med*. 2014; 11(3):162–72. doi: [10.7497/j.issn.2095-3941.2014.03.002](https://doi.org/10.7497/j.issn.2095-3941.2014.03.002) PMID: [25364577](https://pubmed.ncbi.nlm.nih.gov/25364577/)