

Article

**Metabolic Footprinting of a Clear Cell Renal Cell Carcinoma
in vitro Model for Human Kidney Cancer Detection**María Elena Knott, Malena Manzi, Nicolás Zabalegui, Mario
O. Salazar, Lydia I. Puricelli, and Maria Eugenia Monge*J. Proteome Res.*, **Just Accepted Manuscript** • DOI: 10.1021/acs.jproteome.8b00538 • Publication Date (Web): 27 Sep 2018Downloaded from <http://pubs.acs.org> on October 1, 2018**Just Accepted**

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.



ACS Publications

is published by the American Chemical Society, 1155 Sixteenth Street N.W.,
Washington, DC 20036Published by American Chemical Society. Copyright © American Chemical Society.
However, no copyright claim is made to original U.S. Government works, or works
produced by employees of any Commonwealth realm Crown government in the course
of their duties.

1
2
3 **Metabolic Footprinting of a Clear Cell Renal Cell Carcinoma *in vitro* Model for Human**
4
5 **Kidney Cancer Detection**
6
7

8 María Elena Knott¹[⊥], Malena Manzi¹[⊥], Nicolás Zabalegui¹, Mario O. Salazar², Lydia I.
9 Puricelli³, María Eugenia Monge^{1*}
10
11
12
13
14
15

16
17 ¹Centro de Investigaciones en Bionanociencias (CIBION), Consejo Nacional de Investigaciones
18 Científicas y Técnicas (CONICET), Godoy Cruz 2390, C1425FQD, Ciudad de Buenos Aires,
19 Argentina
20
21
22

23 ²Farmacognosia, Departamento de Química Orgánica, Facultad de Ciencias Bioquímicas y
24 Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531, Rosario S-2002LRK, Santa Fé,
25 Argentina.
26
27
28
29

30 ³Instituto de Oncología Ángel H. Roffo, Facultad de Medicina, Universidad de Buenos Aires,
31 Av. San Martín 5481, C1417DTB, Ciudad de Buenos Aires, Argentina.
32
33
34
35
36

37 [⊥]These authors contributed equally to this work.
38
39
40
41
42
43
44
45
46
47

48 *Corresponding author.
49

50
51 E-mail: maria.monge@cibion.conicet.gov.ar
52

53 Ph: +54 11 4899 5500 (ext. 5614)
54
55
56
57
58
59
60

Abstract

A protocol for harvesting and extracting extracellular metabolites from an *in vitro* model of human renal cell lines was developed to profile the exometabolome by means of a discovery-based metabolomics approach using ultraperformance liquid chromatography coupled to quadrupole-time-of-flight mass spectrometry. Metabolic footprints provided by conditioned media (CM) samples (n=66) of two clear cell Renal Cell Carcinoma (ccRCC) cell lines with different genetic background and a non-tumor renal cell line, were compared with the human serum metabolic profile of a pilot cohort (n=10) comprised of stage IV ccRCC patients and healthy individuals. Using a cross-validated orthogonal projection to latent structures-discriminant analysis model, a panel of 21 discriminant features selected by iterative multivariate classification, allowed differentiating control from tumor cell lines with 100% specificity, sensitivity and accuracy. Isoleucine/leucine, phenylalanine, N-lactoyl-leucine, and N-acetyl-phenylalanine, and cysteineglutathione disulfide (CYSSG) were identified by chemical standards, and hydroxypropyl-valine was identified with MS and MS/MS experiments. A subset of 9 discriminant features, including the identified metabolites except for CYSSG, produced a fingerprint of classification value that enabled discerning ccRCC patients from healthy individuals. To our knowledge, this is the first time that N-lactoyl-leucine is associated to ccRCC. Results from this study provide a proof of concept that CM can be used as a serum proxy to obtain disease-related metabolic signatures.

Keywords: *in vitro* cell culture, conditioned media, metabolomics, ultraperformance liquid chromatography-mass spectrometry, clear cell Renal Cell Carcinoma, metabolic footprinting.

Introduction

Kidney cancer is fundamentally a disease of dysregulated cellular metabolism.¹ Renal cell carcinoma (RCC), originated from the renal epithelium, accounts for >90% of cancers in the kidney,² and is among the 10 most common cancers in both men and women worldwide.^{3, 4} RCC patients are often incidentally diagnosed by imaging procedures. Only 10% of patients exhibit the classic triad of hematuria, low back pain, and flank mass symptoms, and nearly 40% of patients lack all of these and present with systemic symptoms, including weight loss, abdominal pain, anorexia, and fever.⁵ More than 30% of the patients exhibit locally advanced or metastatic RCC at the time of diagnosis.^{6, 7} Although the disease is inherently resistant to chemotherapy⁸ and radiotherapy,⁹ the survival of advanced RCC patients has improved significantly with the advent of tyrosine-kinase inhibitors as standard of care therapy.¹⁰ However, the discovery of early detection biomarkers is an important priority area to give more opportunities for early intervention and improved outcome of ccRCC patients.

Clear cell RCC (ccRCC) is the most common (75%) histological subtype and accounts for most cancer-related deaths.^{2, 11} At molecular level, 50-80% of all ccRCC patients show mutations in *Von Hippel Lindau* (VHL) gene which is involved in hypoxia inducible factor 1 α (HIF 1 α) expression.^{3, 12, 13} Overexpression of HIF 1 α triggers the transcription of genes involved in glucose metabolism.¹⁴ In addition, ccRCC is a lipogenic tumor with abnormal cholesterol metabolism.^{15, 16} In this context, and due to complex pathways contributing to kidney cancer progression, a single molecular marker might not be efficient enough as tumor biomarker, suggesting the need of a multiple biomarker panel to achieve sufficient clinical information.

Mammalian cell metabolomics^{17, 18} has emerged as a promising tool for studying cellular biochemistry and investigate altered metabolic networks that contribute to cell proliferation,

1
2
3 dissemination and survival in RCC. Metabolomics uses a holistic approach to characterize and
4 quantify the metabolome, comprised of all the small molecules (MW<1500) in biological
5 systems.^{19, 20} Different metabolic fingerprinting studies have shown alterations associated with
6
7 RCC. Mass spectrometry (MS)-based urine untargeted metabolomics approaches have suggested
8 alterations in metabolic pathways of tryptophan²¹ and acylcarnitines.²² Weiss and collaborators
9 showed that serum is a more accurate proxy for tissue changes than urine, using a mouse
10 xenograft model of kidney cancer, and suggested that tryptophan degradation is highly
11 represented in RCC.²³ Untargeted serum MS-based metabolomics studies from RCC patients
12 have also suggested disease-related alterations in the phospholipid catabolism, sphingolipid,
13 cholesterol, phenylalanine, and arachidonic acid metabolisms in addition to fatty acid beta-
14 oxidation.^{23, 24} Recently, tumor progression and metastasis have been associated with metabolite
15 increases in glutathione and the cysteine/methionine pathways by means of a metabolomic
16 profiling study on 138 matched ccRCC/normal tissue pairs.²⁵
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

33 *In vitro* cell models are of particular interest for understanding the metabolism of cellular
34 processes, and allow the study of both intracellular (fingerprint) and extracellular (footprint)
35 metabolic profiles,^{17, 18, 26} being the latter a closer proxy of serum. Cell lines can be used as *in*
36 *vitro* models for research including biomarker discovery studies and the evaluation of new drugs
37 for treatment,¹⁷ as recommended by the Food and Drug Administration.²⁷ Recent genomic
38 studies have identified molecular differences between commonly used renal cancer cell lines and
39 human tumor samples.²⁸ However, studies comparing the exometabolome of kidney cell lines
40 with the human serum metabolome to evaluate the feasibility of using *in vitro* models for serum
41 sample classification have not been reported up to date.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 In this study, we have optimized a protocol for harvesting, extracting, lyophilizing and
4 reconstituting conditioned media (CM) metabolites derived from two human ccRCC cell lines
5 786-O (VHL^{-/-}) and Caki-1 (VHL^{+/+}), and the non-tumor human renal cell line HEK-293 (n=22
6 for each cell line); and we have profiled the exometabolome using a discovery-based
7 metabolomics approach by means of ultraperformance liquid chromatography coupled to
8 quadrupole-time-of-flight mass spectrometry (UPLC-QTOF-MS). Metabolic features (Rt, *m/z*
9 pairs) were analyzed using a cross-validated orthogonal projection to latent structures-
10 discriminant analysis model, coupled to a genetic algorithm variable selection method. A panel
11 of 21 discriminant features, obtained from the binary comparison of control and tumor cells,
12 allowed sample classification with 100% specificity, sensitivity and accuracy. In addition, 9 of
13 these compounds were present in human serum samples and enabled discriminating stage IV
14 ccRCC patients from healthy individuals, which could potentially be relevant in kidney cancer
15 diagnosis.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

37 **Materials and Methods**

38 *Chemicals*

39
40
41
42
43 LC-MS grade acetonitrile, methanol, isopropanol, acetic acid and formic acid purchased
44 from Fisher Chemical (NC, USA) and ultrapure water with 18.2 MΩ·cm resistivity (Thermo
45 Scientific Barnstead Micropure UV ultrapure water system, USA) were used to prepare
46 chromatographic mobile phases and solutions. Leucine enkephalin was purchased from Waters
47 Corp. (Milford, MA, USA). Sodium hydroxide was purchased from EMSURE[®] ISO (Merck
48 Millipore, Burlington, MA, USA). Roswell Park Memorial Institute (RPMI) 1640 and
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12) powder culture media
4
5 were purchased from GIBCO™ (Thermo Fisher Scientific, MA USA). Fetal bovine serum (FBS)
6
7 was purchased from Internegocios S.A. (Buenos Aires, Argentina), gentamicin from Laboratorio
8
9 Drewer S.A. (Buenos Aires, Argentina), and L-glutamine from GIBCO™ (Thermo Fisher
10
11 Scientific, MA USA). The analgesic mix comprised of the following standards: acetaminophen,
12
13 2-acetaminophen, acetanilide, acetylsalicylic acid, caffeine, phenacetin and salicylic acid was
14
15 purchased from Waters (Waters Corporation, Manchester, UK). L-cysteine-glutathione disulfide
16
17 (purity $\geq 95\%$) was purchased from Cayman Chemical (MI, USA). L-leucine ($\geq 98\%$) and L-
18
19 phenylalanine ($\geq 98\%$) were purchased from Sigma-Aldrich (St. Louis, MO, USA). L-isoleucine
20
21 ($\geq 98\%$) was purchased from Fluka (Steinheim, Germany). N-acetyl-phenylalanine and N-
22
23 lactoyl-leucine reference compounds were chemically synthesized in our laboratory (Supporting
24
25 Information, Figures S1 and S2).
26
27
28
29

30 31 *In Vitro Model - Cell Culture* 32 33

34
35 The *in vitro* model was comprised of three different kidney cell lines: HEK-293, 786-O
36
37 and Caki-1, which were obtained from the American Type Culture Collection (ATCC). HEK-
38
39 293, a non-tumor human embryonic kidney cell line, was used as control and was compared to
40
41 two ccRCC cell lines with different genetic background, 786-O and Caki-1. The 786-O cell line
42
43 derives from primary ccRCC tumor and has a deletion of a gene encoding the VHL protein
44
45 (VHL^{-/-}). The Caki-1 cell line arises from ccRCC skin metastasis, and expresses wild type VHL
46
47 protein (VHL^{+/+}). HEK-293 and 786-O cell lines were cultured in RPMI culture media and Caki-
48
49 1 cell line was cultured in DMEM/F12 media; all supplemented with 10% FBS, 2 mM L-
50
51 glutamine, and 8 $\mu\text{g mL}^{-1}$ gentamicin, in a humidified atmosphere of 5% CO₂ at 37 °C. Cell
52
53 cultures were routinely checked for mycoplasma contamination by DAPI (2-(4-Amidinophenyl)-
54
55
56
57
58
59
60

1
2
3 1H-indole-6-carboxamide) staining. The protocol designed for cell culture and CM harvesting
4 generated 24 samples for each cell line (Figure 1, Table S1). Briefly, two cryovials with similar
5 cell passage number were thawed for each cell line. Cells from each cryovial were split into two
6 flasks (A and B in Figure 1) and were treated independently during the experiments. After cell
7 counting, two wells were plated for each cell line. Culture media blanks were obtained by
8 incubation of cell culture media with the same protocol used for cell lines (Figure 1). Two CM
9 samples from each cell line were used for analytical method development. Thus, 22 samples per
10 cell line were used for the untargeted metabolic footprinting study.

21 *In vitro Model - Conditioned Media Collection*

22
23
24
25 In each well of 9.5 cm² area, 5.00×10⁵, 1.00×10⁵, and 1.25×10⁵ cells were plated for
26 HEK-293, 786-O, and Caki-1 cell lines, respectively. Once 80% confluence was reached, cells
27 were gently washed 3 times during 10 minutes with the corresponding culture media, without
28 FBS. Then, cell monolayers were incubated overnight with 800 μL of their respective cell
29 culture media without FBS (starving conditions).²⁹ CM samples and culture media blanks were
30 collected, centrifuged at 300 × g for 5 minutes to remove cell debris, and finally supernatants
31 were collected and immediately frozen at -80 °C.

41 *Serum Samples*

42
43
44
45 Serum samples and associated clinical data were provided by the public oncologic serum
46 biobank “Biobanco Público de Muestras Séricas Oncológicas” (BPMSO) from “Instituto de
47 Oncología A. H. Roffo” (IOAHR), Buenos Aires, Argentina. The patient cohort consisted of 5
48 healthy individuals (age range 40–64, mean (SD) age 56(9) years, 20% male) and 5 ccRCC
49 patients at stage IV (SIV) (age range 53–72, mean age 64(6) years, 100% male). At the 0.05
50
51
52
53
54
55
56
57
58
59
60

1
2
3 level, the means of the age populations were not significantly different with the two-sample t-
4
5 test. Blood samples were drawn from untreated cancer patients. Serum sample collection
6
7 followed the guidelines approved by the IOAHR Institutional Review Board, and samples were
8
9 drawn after signature of the corresponding informed consent. According to the BPMSO standard
10
11 operating procedure, 20 mL of blood were collected in tubes without any anticoagulant and left
12
13 15 minutes at 25 °C to allow the clot formation and centrifuged at $600 \times g$ for 10 minutes. Then,
14
15 serum was split into aliquots and stored at -80 °C. Serum aliquots were used only once after
16
17 thawing.
18
19
20
21

22 *Conditioned Media and Serum Sample Preparation*

23
24

25 All frozen samples were thawed at 0 °C on a water-ice bath. Protein precipitation was
26
27 performed by addition of cold (4 °C) isopropanol to 500 μL of CM or 60 μL of serum samples,
28
29 in a 3:1 solvent:sample volume ratio. Samples were vortex-mixed for 10 seconds and centrifuged
30
31 at $16000 \times g$ for 20 minutes and 4 °C. Supernatants were frozen and lyophilized at -80 °C and 50
32
33 mTorr for 48 hours using a Telstar LYOQuest-85 freeze dryer (Telstar, Madrid, Spain). Sample
34
35 residues were reconstituted in water/methanol (80:20 v/v) with a concentration factor of 7 for
36
37 CM samples, and 1 for serum samples; and analyzed by UPLC-QTOF-MS. Sample preparation
38
39 blanks containing ultrapure water and culture media blanks also went through the same sample
40
41 preparation procedure. CM and serum samples were analyzed in different batches, with quality
42
43 control (QC) and quality assurance (QA) samples to verify the stability of retention times, peak
44
45 shapes and areas during the analysis.³⁰ QCs consisted in randomly pooled CM samples from the
46
47 3 cell lines studied and were processed in an identical approach as samples; i.e., a small aliquot
48
49 of a subset of the CM samples were pooled into a single QC sample, followed by protein
50
51 precipitation and further sample preparation steps. All CM samples were used to prepare QCs.
52
53
54
55
56
57
58
59
60

1
2
3 For each batch a new QC sample was reconstituted and analyzed every 14 sample injections.
4
5 Samples were randomly analyzed by UPLC-QTOF-MS together with QCs, solvent blanks,
6
7 sample preparation blanks and culture media blanks. Additionally, for serum samples, QA was
8
9 verified with the periodical injection of a mix of standards comprised of acetaminophen, 2-
10
11 acetaminophen, acetanilide, acetylsalicylic acid, caffeine, phenacetin and salicylic acid. CM
12
13 samples were analyzed along 5 consecutive days and serum samples were analyzed in one day,
14
15 after conditioning the analytical platform.
16
17
18

19 20 *Ultraperformance Liquid Chromatography-Mass Spectrometry*

21
22

23 Ultraperformance liquid chromatography-mass spectrometry (UPLC-MS) analyses were
24
25 performed using a Waters ACQUITY UPLC I Class system fitted with a Waters ACQUITY
26
27 UPLC BEH C₁₈ column (2.1×100 mm, 1.7 μm particle size, Waters Corporation, Milford, MA,
28
29 USA), coupled to a Xevo G2S QTOF mass spectrometer (Waters Corporation, Manchester, UK)
30
31 with an electrospray ionization (ESI) source. The typical resolving power and mass accuracy of
32
33 the Xevo G2S QTOF mass spectrometer were 32,000 FWHM and 0.2 ppm at m/z 554.2615,
34
35 respectively. Gradient elution was utilized in the chromatographic separation method using water
36
37 with 0.1% acetic acid (mobile phase A) and acetonitrile (mobile phase B), with the following
38
39 program: 0-1 min 10% B; 1-2.5 min 10-15% B; 2.5-4 min 15-22% B; 4-6 min 22-38% B; 6-9
40
41 min 38-65% B; 9-12 min 65-80% B; 12-16 min 80-100% B; 16-19 min 100% B. The flow rate
42
43 was constant at 0.25 mL min⁻¹ for 12 min and was increased to 0.30 mL min⁻¹ between 12 and 19
44
45 min. After each sample injection, the gradient was returned to its initial conditions in 11 min.
46
47
48 The injection volume was 5 μL for CM and 2μL for serum samples. The column and
49
50 autosampler tray temperatures were set at 35 and 5 °C, respectively. The mass spectrometer was
51
52 operated in negative ion mode with a probe capillary voltage of 2.3 kV, and a sampling cone
53
54
55
56
57
58
59
60

1
2
3 voltage of 30.0 V. The source and desolvation gas temperatures were set to 120 and 300 °C,
4
5 respectively. The nitrogen gas desolvation flow rate was 600 L h⁻¹, and the cone desolvation flow
6
7 rate was 10 L h⁻¹. The mass spectrometer was daily calibrated before each batch analysis across
8
9 the range of m/z 50-1200 using a 0.5 mM sodium formate solution prepared in isopropanol/water
10
11 (90:10 v/v). Data were drift corrected during acquisition using a leucine enkephalin (m/z
12
13 554.2615) reference spray infused at 2 $\mu\text{L min}^{-1}$, every 45 seconds. Data were acquired in MS^E
14
15 continuum mode³¹ in the range of m/z 50-1200, and the scan time was set to 0.5 s. Technical
16
17 duplicates were acquired in all cases. For UPLC-MS/MS experiments, the product ion mass
18
19 spectra were acquired with collision cell voltages between 10 and 30 V, depending on the
20
21 analyte. Ultra-high-purity argon ($\geq 99.999\%$) was used as the collision gas. Data acquisition and
22
23 processing were carried out using MassLynx version 4.1 (Waters Corp., Milford, MA, USA).
24
25 The mass spectrometry data have been deposited to the MetaboLights public repository
26
27 (<https://www.ebi.ac.uk/metabolights/index>) with the data set identifier MTBLS737.
28
29
30
31
32
33

34 *Data Analysis*

35
36 Spectral features (retention time (R_t), m/z pairs) were extracted from UPLC-QTOF-MS
37
38 data using Progenesis QI version 2.1 (Nonlinear Dynamics, Waters Corp., Milford, MA, USA).
39
40 The procedure included retention time alignment, peak picking, deisotoping, integration, and
41
42 grouping together adducts derived from the same compound. Subsequently, if a feature had a
43
44 peak area in a CM sample that was 3-fold or less than the mean peak area in the solvent and
45
46 sample preparation blanks of the same feature, then its peak area was set to 0.³² Otherwise, the
47
48 mean peak area in those blanks was subtracted from the feature peak areas in the CM samples.
49
50 After blank subtraction, feature abundances from technical duplicates were averaged, and only
51
52 those that were present in at least 80% of one group class were retained. Since Caki-1 cell lines
53
54
55
56
57
58
59
60

1
2
3 were cultured with a different culture medium, only features from HEK-293 or 786-O were
4 considered to build the feature matrix for multivariate statistical analysis to avoid influence of
5 culture media on the detected exometabolome (cell secretome).³³ Chromatographic peak shape
6 and signal intensity of each feature was further evaluated for data curation. Features with signal
7 intensity $<10^3$ in the continuum mass spectra or with a mass difference larger than 10 mDa along
8 the sample list were discarded. The feature matrix obtained after this procedure was normalized
9 to the number of plated cells for each cell line and to the total peak area for each sample. This
10 matrix (Data Set S1 in the Supporting Information) was then utilized to build unsupervised and
11 supervised multivariate statistical analysis models using MATLAB R2012b (The MathWorks,
12 Natick, MA, USA) with the PLS Toolbox version 8.1 (Eigenvector Research, Inc., Manson, WA,
13 USA). Principal component analysis (PCA) was used to track data quality, reduce the
14 dimensionality, identify and remove outliers in the dataset, as well as to identify sample clusters.
15 Two samples were identified as outliers by PCA (data not shown), one from a HEK-293 CM and
16 another from 786-O CM, and were not further considered for data analysis. Orthogonal
17 projection to latent structures-discriminant analysis (oPLS-DA)^{34, 35} coupled with a genetic
18 algorithm (GA) variable selection method was applied to find a feature panel that maximized
19 classification accuracy for the binary comparison of HEK-293 and 786-O. A panel of 23
20 discriminant features had the lowest root-mean-square error of cross-validation (RMSECV) at
21 the conclusion of the GA variable selection process; however, only 21 features were retained for
22 statistical analysis (see *Metabolite Identification Procedure*). The parameters for genetic
23 algorithm variable selection were as follows: population size: 112, variable window width: 1, %
24 initial terms (variables): 10, target minimum # of variables: 1, target maximum # of variables:
25 30, penalty slope: 0.05, maximum generations: 100, % at convergence: 80, mutation rate: 0.005,
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 crossover: double, regression choice: PLS, # of latent variables: 2, cross-validation: random, # of
4
5 splits: 4, # of iterations: 10, replicate runs: 20. The oPLS-DA model was cross-validated using
6
7 venetian blinds with 10 data splits. Data were preprocessed by autoscaling prior to PCA or
8
9 oPLS-DA analysis. PCA was also performed to inspect data before and after GA variable
10
11 selection (*i.e.*, on the curated spectral feature matrix and on the discriminant feature panel).
12
13 Feature abundances of QC samples for each batch were averaged before PCA.
14
15
16

17
18 Fold changes were calculated as the ratio of median peak areas between CM from 786-O
19
20 and HEK-293, Caki-1 and HEK-293, and Caki-1 and 786-O samples. Mann-Whitney U tests
21
22 were used to calculate statistical significance, and *p* values were corrected using the Benjamini-
23
24 Hochberg³⁶ procedure for multiple comparisons with a FDR of 0.1.
25
26

27 *Metabolite Identification Procedure*

28
29

30
31 Metabolite identification was attempted for the 23 discriminant features resulting from
32
33 the GA variable selection process. Mass spectral adduct ions and elemental formula generated
34
35 based on accurate masses and isotopic patterns, were searched against the Human Metabolome
36
37 Database (HMDB).³⁷ Based on the list of tentative candidates, metabolite identification was
38
39 performed considering the accurate mass, isotopic pattern, and the fragmentation pattern
40
41 obtained from tandem QTOF-MS experiments. Tandem MS spectra were compared to the Metlin
42
43 database,³⁸ and MassBank,³⁹ and for cases where MS/MS spectra were not available in
44
45 databases, fragmentation patterns were manually interpreted for metabolite annotation. Putative
46
47 identifications were validated with chemical standards. Though Progenesis performed adduct
48
49 grouping, the software did not account for NaCl adduct ions, which were evidenced by isotopic
50
51 pattern analysis of the selected features. Out of the 23 features, 2 features were identified as
52
53 different NaCl adduct ions of the same compound; *i.e.* different $[M+x\text{NaCl-H}]^-$ ionic species, and
54
55
56
57
58
59
60

1
2
3 one of the features was identified as paraben, which is a non-endogenous metabolite. Therefore,
4
5 only 21 features were considered for further statistical analysis. Chemical standards were
6
7 prepared in ultrapure water, and were analyzed under identical conditions as CM and serum
8
9 samples to validate putative metabolite identities by chromatographic retention time and MS/MS
10
11 fragmentation pattern matching. Spiking experiments were also conducted with the standards on
12
13 CM and serum samples as well as on cultured media blanks, to address retention time differences
14
15 caused by matrix effects.
16
17
18
19
20
21

22 **Results and Discussion**

23 *Metabolic Footprint & CM Sample Classification*

24
25 A protocol for harvesting and extracting extracellular metabolites from an *in vitro* model
26
27 of 3 different human renal cell lines was developed (Figure 1, Table S1) and implemented to
28
29 profile the cell secretome³³ with a UPLC-QTOF-MS-based method. In addition, serum samples
30
31 from 5 healthy individuals and 5 SIV ccRCC patients were analyzed with the same optimized
32
33 analytical method to evaluate the use of CM as a serum proxy to obtain a disease-related
34
35 metabolic signature from the detected metabolomes.
36
37
38
39
40
41

42 A total of 6002 features (Rt, *m/z* pairs) were extracted by Progenesis software from the
43
44 UPLC-MS negative ion mode data from the three studied cell lines. The number of features was
45
46 reduced to 5030 after deconvolution, to group together adducts derived from the same
47
48 compound. Following solvent blank deduction, feature abundances from technical replicates
49
50 were averaged, and only features that were present in at least 80% of one group class were
51
52 retained. Since Caki-1 cell line was cultured with a different culture media, features that were
53
54
55
56
57
58
59
60

1
2
3 only present in Caki-1 CM were excluded for further analysis to avoid bias in sample
4 classification by the media composition.³³ This data set, comprised of 2358 features, was further
5 manually filtered to keep only those features with signal intensity, peak shape and mass variance
6 within the established thresholds that would allow accurate identification, leaving 755 spectral
7 features. This matrix was normalized (Data Set S1 in the Supporting Information) and utilized to
8 build a PCA model. The 2D score plot illustrated in Figure 2A shows distinguishable separation
9 between the three classes, mainly achieved by the contribution of the first PC. In addition, QC
10 data points clustering around the origin of the PCA score plot indicates reproducibility in the
11 sample preparation method, high data quality and adequate performance of the analytical
12 platform. The PCA model consisted of 5 PCs with 70.67% total captured variance, with the first
13 two PCs accumulating 47.44% of the total variance. Figure 2B shows the score plot for CM
14 samples from the non-tumor control cell line HEK-293 and the tumor cell line 786-O, which
15 were cultured and incubated with the same culture media. The model, which consisted of 2 PCs
16 with 53.38% total captured variance, provided a clear degree of class separation, mainly
17 achieved by the contribution of the first PC. The loadings plot associated to PC1 showed similar
18 contribution from all spectral features to sample separation in the score plot (figure not shown).
19 Thus, sample discrimination was further analyzed by means of oPLS-DA coupled to a GA
20 variable selection method to find a reduced metabolic feature set that would allow sample
21 classification and class membership prediction. A panel of 23 metabolic features with the lowest
22 RMSECV was selected through the GA process; although 2 of them were removed for further
23 statistical analysis (see *Metabolite Identification Procedure*). Figure 2C shows the cross-
24 validated prediction plot using the 21 discriminant metabolic features. The model consisted of 1
25 latent variable that interpreted 50.55% and 99.47% variance from the X- (feature peak areas) and
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Y- (class membership) blocks, respectively. This oPLS-DA model resulted in 100% cross-
4 validated accuracy, sensitivity, and specificity; therefore, no CM samples were misclassified.
5
6
7

8 *Comparison of in vitro Exometabolome with Serum Metabolome*

9

10
11 Figure S3 shows the different base peak intensity (BPI) chromatograms obtained for CM
12 samples of each cell line, and for serum samples of a healthy individual and a SIV ccRCC
13 patient. Differences observed were probably associated to metabolite concentration levels, and
14 matrix effects. Additionally, human serum metabolic profiles are influenced by the presence of
15 the disease, lifestyle, diet, environmental exposures, i.e. the exposome,^{40, 41} and molecules
16 derived from the interactions with associated microorganisms (the microbiome),⁴² demonstrating
17 the need for multivariate statistical analysis to extract a disease-related metabolic signature from
18 the detected metabolomes.
19
20
21
22
23
24
25
26
27
28
29

30 To evaluate the ccRCC *in vitro* exometabolome as a good proxy to study metabolic
31 changes in serum, the 755 features of the CM matrix were searched in the feature matrix
32 extracted from human serum sample analysis. A total of 163 features were found to be common
33 to both CM and sera. In addition, out of the 21 metabolic features that perfectly classified CM
34 samples of the renal *in vitro* model, 9 were present in serum samples. These feature sets were
35 further utilized to build unsupervised models to explore sample clustering in both types of
36 biological samples, i.e. CM and sera. Figures 3A and 3B show the 2D PCA score plots of serum
37 samples using the 163 common feature set and the smaller subset of 9 common features from the
38 CM discriminant panel, respectively. Using the former set, the PCA model composed of 2 PCs
39 with 57.97% total captured variance did not show sample separation (Figure 3A). However, the
40 PCA model based on the subset of 9 common CM discriminant features utilized 2 PCs that
41 captured a larger percentage of total variance (70.87%) after dimension reduction, and was able
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 to provide a good degree of sample separation between SIV ccRCC patients and healthy
4 individuals (Figure 3B). The overlapping of two samples from both classes in the score plot was
5
6 not associated to gender, even if the latter was not balanced between sample classes. Therefore,
7
8 the metabolic footprint of the ccRCC *in vitro* model assisted in differentiating human individuals
9
10 based on the presence of disease.
11
12
13
14

15 The PCA model that used the 163 common features provided similar class separation in
16
17 CM samples (Figure 3C) as the one based on 755 features (Figure 2A). The first PC of the scores
18
19 plot allowed sample separation between tumor (786-O and Caki-1) and non-tumor (HEK-293)
20
21 cell lines, whereas the second PC provided separation between 786-O and Caki-1 cell lines.
22
23 Interestingly, an improved CM class separation with larger captured variance (65.02%) was
24
25 observed in the score plot of the PCA model using the subset of 9 discriminant features (Figure
26
27 3D) compared with 755 and 163 features (Figure 2A and 3C, respectively). Actually, this
28
29 reduced panel provided an improved class separation for the binary comparison of HEK-293 vs.
30
31 786-O cell lines (Figure S4A) than with 755 features (Figure 2B) and also a clear separation
32
33 between tumor cells with different genetic background (Figure S4B). Considering that these 9
34
35 features provide a good simplified ccRCC metabolic signature, and allow serum sample
36
37 classification, their analysis could potentially offer useful information in studies of diagnosis
38
39 purposes, and drug discovery for cancer treatment.
40
41
42
43
44
45

46 *Discriminant Metabolite Identification*

47
48

49 Metabolite identification was attempted for the 21 features of the discriminant panel, as
50
51 illustrated in Figure 4 for N-lactoyl-leucine (N-Lac-Leu). First, extracted ion chromatograms
52
53 (EICs) and the corresponding mass spectra were obtained for each feature in CM and serum
54
55 samples. Mass spectral detected adduct ions and elemental formula of possible candidates were
56
57
58
59
60

1
2
3 searched in the HMDB³⁷ database. Subsequently, fragmentation patterns obtained from tandem
4 MS experiments were compared to MS/MS spectra in the HMDB³⁷ or Metlin database,³⁸ or
5 manually interpreted. Finally, the tentatively identified metabolites were confirmed by matching
6 retention times and fragmentation patterns in CM and serum samples with chemical standards,
7 whenever possible. Chemical standards also aided to discard non matches with putatively
8 identified compounds, leaving several features with no ID (Table 1). The identity of some
9 candidate molecules could not be confirmed due to i) co-elution with similar molecular weight
10 compounds that interfered in the quadrupole selection process of the precursor ion, providing
11 product ion overlap with the target feature in the tandem MS spectra, ii) insufficient precursor
12 ion intensity for MS/MS experiments, or iii) limitations associated to metabolite databases.
13
14
15
16
17
18
19
20
21
22
23
24
25
26

27 The isotopic pattern analysis of the discriminant features suggested the presence of
28 several ionic species being the product of non-specific binding caused by the matrix
29 composition, which occurs when chemical species that are trapped in the same ionization
30 droplets start to interact during solvent evaporation in the ESI mechanism.^{43, 44} Out of the 21
31 features, 6 were identified as $[M+xNaCl-H]^-$ ionic species, with x between 1 and 7, for different
32 compounds (M), which were not accounted by Progenesis software. These type of adduct ions
33 were detected for N-acetyl-phenylalanine (N-Ac-Phe); phenylalanine (Phe); isoleucine/leucine
34 (Ile/Leu), and for hydroxypropyl-valine (OHPro-Val) (Table 1, Figures S5-S8). Since the
35 chemical composition of the culture media utilized to incubate the cell lines favored this type of
36 nonspecific ion pairing, spiking experiments were conducted with the standards on CM and
37 serum samples, as well as on cultured media blanks, to validate metabolite identity, and address
38 retention time shifts caused by matrix effects (Figure 4, Figures S5-S7). The salt content of CM
39 samples due to the culture media formulation and the concentration factor optimized for
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 metabolite detection (7 for CM samples vs. 1 for serum samples), was actually translated, in
4
5 some cases, into 3-5 seconds shorter retention times compared to metabolites detected in serum
6
7 samples or in standard solutions (Figure 4 and Figures S5-S7).
8
9

10 Matrix composition mainly affected Ile/Leu identification, since the chromatographic
11
12 gradient implemented for sample analysis was not able to separate these isomers, but instead
13
14 provided partial resolution for a Ile+Leu standard aqueous solution (Figure S7A, S7B). A
15
16 different chromatographic initial gradient was utilized in combination with different spiking
17
18 experiments to illustrate the presence of both Ile and Leu in CM samples, as expected from the
19
20 culture media composition (Figure S7B). Since neither the chromatographic separation nor
21
22 tandem MS experiments allowed identifying the discriminant amino acid (Figure S7D), for
23
24 further discussion both isomers will be considered together as Ile+Leu.
25
26
27
28
29

30 Both $[M-H]^-$ and $[M+NaCl-H]^-$ adduct ions were detected for $M = \text{OHPro-Val}$. To
31
32 distinguish between the two possible isomers of this dipeptide, i.e., OHPro-Val or valyl-
33
34 hydroxyproline (Val-OHPro), tandem MS experiments were conducted both in negative and
35
36 positive ESI modes, and the product ions detected in the corresponding mass spectra suggested
37
38 OHPro-Val being the discriminant metabolite (Figure S8). Since no chemical standard was
39
40 analyzed for this compound, its identification was given a different confidence level (Table 1).
41
42 For cysteine glutathione disulfide (CYSSG), only the $[M-H]^-$ ion was detected, and both a spiked
43
44 CM sample and a CYSSG aqueous standard solution provided identical retention times (Figure
45
46
47
48
49
50 S9).

51 Overall, 6 of the 21 CM discriminant metabolic features were successfully identified by
52
53 MS and MS/MS experiments, while 5 were further chromatographically confirmed by chemical
54
55
56
57
58
59
60

standards (Table 1). In addition, 5 out of 6 identified features were detected in human serum samples (Table 1).

Biological significance of Identified Discriminant Metabolites in ccRCC

Metabolic patterns are powerful tools for sample classification, though understanding their biological significance can become more challenging.⁴⁵ Changes in gene expression do not necessarily correlate with changes in metabolites of a given pathway during ccRCC tumor progression.²⁵ Table 1 summarizes the identification of the 6 discriminant endogenous metabolites, discussed below based on their level change in CM and serum samples and the affected pathways, many of which have been reported to be involved in RCC progression. Fold changes were calculated as the ratio of median peak areas between CM samples, and *p* values were corrected using the Benjamini-Hochberg³⁶ procedure for multiple comparisons with a FDR of 0.1. Though no significant fold changes were obtained for serum samples, probably due to the small cohort, trends between classes are shown in Table 1.

Isoleucine and leucine (Ile+Leu), Phe, N-Lac-Leu and N-Ac-Phe exhibited significant higher levels in CM of 786-0 cells (CM-7) than HEK-293 cells (CM-H) with fold changes between 1.6 and 7.3 (Table 1, Figure 5). In contrast, CYSSG showed decreased levels in CM-7 compared to CM-H with the highest significant fold change equal to 18. Regarding the binary comparison between CM from non-tumor cells (CM-H) and Caki-1 cells (CM-C), CYSSG and OHPro-Val presented significant decreased levels in CM-C, with fold changes of 270 and 10, respectively; whereas N-Lac-Leu levels were 1.7 higher (Table 1, Figure 5). Significantly decreased levels of Ile+Leu, Phe, N-Ac-Phe and OHPro-Val were detected in CM-C compared to CM-7 with fold changes between 1.7 and 200, while CYSSG levels were 1.8-fold higher in CM-C (Table 1, Figure 5). Ile and Leu concentration was equivalent in both culture media

1
2
3 formulations, but Phe concentration was doubled in DMEM/F12 compared to RPMI. Therefore,
4
5 comparisons between CM-C and CM-H or/and CM-7 should be cautiously interpreted for Phe.
6
7

8 *Cysteineglutathione disulfide (CYSSG) – Redox state*

9

10
11 CYSSG can be endogenously produced *via* a thiol-disulfide exchange reaction between γ -
12 glutamyl-cysteinyl-glycine (GSH) and L-cystine or glutathione disulfide (GSSG) and Cys.⁴⁶
13
14 The balance between GSH and GSSG is crucial for regulating the redox potential of the cell.^{47, 48}
15
16 GSH counteracts the increased ROS production, thus minimizing oxidative damage to tissues
17 and cells.^{47, 49} GSH is synthesized from amino acid precursors (Cys, Glu, and Gly) in the cell
18 cytosol, and it is primarily regulated by GCS (glutathione-S-transferase), Cys availability, and
19 GSH feedback inhibition.^{50, 51} The intracellular GSH concentration depends on a dynamic
20 balance between its synthesis and consumption inside the cell, and its efflux.⁵² Since oxidative
21 stress occurs in kidney cancer,⁵³ as well as in RCC cell lines including 786-O and Caki-1 cells,⁵⁴
22 it is not surprising to find that CYSSG levels in CM-7 and CM-C were lower than those in CM-
23 H. Even if we do not have information about the ratio GSH/GSSC, our results suggest an
24 alteration in the cell redox state. Along these lines, higher GSH levels were detected in tumor
25 tissue of ccRCC patients compared to normal tissue,⁵⁵ while levels of GSH, GSSG, Cys-Gly and
26 α -hydroxybutyrate, all involved in the GSH biosynthetic pathway, were increased in late-stage
27 tumors compared to early stages.²⁵ As well, higher levels of CYSSG were detected in tumor
28 tissue of a mouse xenograft model characterized by implantation of Caki-1 cells in the kidney,
29 compared to control, presenting the highest fold change among all measured metabolites.²³ We
30 were not able to detect CYSSG in the patient cohort, probably due to the relatively lower GSH
31 levels reported in human plasma, compared to intracellular GSH content.⁴⁷
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 *Isoleucine, Leucine, Phenylalanine, N-Lactoyl-Leucine, N-Acetyl-Phenylalanine and*
4
5 *Hydroxyproline-Valine –Metabolism of Amino acids and Proteins*
6
7

8 The identified amino acids, Ile+Leu and Phe, and the amino acid derivatives N-Lac-Leu
9 and N-Ac-Phe, exhibited significantly increased levels in CM-7 compared to CM-H, while
10 similar levels of Ile+Leu, Phe and N-Ac-Phe were detected between CM-H and CM-C (Table 1,
11 Figure 5). However, as mentioned above, the biological interpretation of Phe fold changes in
12 CM-C could be biased by the different culture media composition. A different trend was
13 obtained for OHPro-Val, which showed decreased levels in CM-7 and CM-C, with significant
14 differences between CM-H and CM-C (Table 1, Figure 5). These compounds are all involved in
15 the metabolism of amino acids and proteins. In this regard, Hakimi *et al.* identified that most of
16 the pathways significantly decreased in ccRCC tumor tissue were involved in amino acid
17 metabolism, including Ile, Leu and Phe metabolism.²⁵
18
19
20
21
22
23
24
25
26
27
28
29
30
31

32 Previous evidence showed that a large number of dipeptides were detected in higher
33 levels in tumor tissue of patients with SIV ccRCC compared to early stages.²⁵ Our *in vitro* model
34 analysis showed lower OHPro-Val levels in CM from tumor cell lines compared to the non-
35 tumor cell line (with significant decrease for CM-C vs. CM-H), and serum samples exhibiting
36 the same trend (Table 1, Figure 5). A possible explanation for this result is that Caki-1 cells may
37 reuse dipeptides as a source of intracellular amino acids so that, even if they are exported to the
38 extracellular milieu they could be internalized faster than in normal cells. In this sense, it is well
39 known that peptides can penetrate into the cells by endocytosis.⁵⁶
40
41
42
43
44
45
46
47
48
49
50

51 Lower serum amino acid levels can be expected if biosynthetic requirements are
52 enhanced in tumor cells. Our results showed no change in Phe levels and reduced levels of
53 Ile+Leu in serum samples from SIV ccRCC patients (Table 1, Figure 5). Nevertheless, there are
54
55
56
57
58
59
60

1
2
3 some contradictory results in the literature. Non-significant changes in the levels of Ile, Leu, and
4 Phe were reported in a serum sample cohort with 65% ccRCC (out of RCC) American patients
5 by means of ion exchange chromatography with post derivatization⁵⁷ whereas increased levels of
6 Ile, Leu and Phe were detected in a cohort with 100% ccRCC Chinese patients using an LC-
7 MS/MS-based method, demonstrating the possible influence of ethnicity, histology, and/or to the
8 techniques used in the results obtained.⁵⁸
9

10
11
12
13
14
15
16
17
18 Regarding N-Ac-Phe, its levels were increased in CM from 786-O cells compared to CM-
19 H, and a similar trend was observed for SIV ccRCC serum samples in relation to healthy
20 individuals (Table 1, Figure 5). Elevated N-Ac-Phe in ccRCC serum has already been reported
21 coupled to a tyrosine decline,²⁴ associating this alteration either to the inhibition of the
22 phenylalanine hydroxylase, the enzyme necessary to metabolize phenylalanine to tyrosine,⁵⁹ or to
23 impaired glycine N-acyltransferase activity.⁶⁰ Interestingly, the loss of VHL in ccRCC, among
24 other changes, leads to elevated levels of citrate and cytosolic AcetylCoA,^{25, 61} that may favor
25 Phe acetylation. Higher level of acetylated proteins and/or amino acids could lead to higher
26 levels of N-Ac-Phe, among others, in the extracellular compartment. A similar process may
27 explain the experimentally observed higher levels of N-Ac-Phe in CM of 786-O cells, which lack
28 VHL.
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

There is evidence showing that N-lactoyl-amino acids are synthesized from lactic acid and amino acids, catalyzed by cytosolic non-specific dipeptidase 2, and then exported by ATP-binding cassette subfamily C member 5 to the extracellular compartment.⁶² High levels of lactate and certain amino acids have been shown to correlate with high levels of N-lactoyl-amino acids in plasma.⁶² This has been observed in healthy individuals subjected to exercise, and patients with phenylketonuria, with high levels of Phe and N-lactoyl-phenylalanine. ccRCC exhibits high

1
2
3 rates of glycolysis and activation of lactate dehydrogenase,⁶¹ leading to high cytosolic lactate
4 content. Based on this evidence, we hypothesized that elevated N-Lac-Leu levels in the CM from
5 tumor cells could result from the intracellular reaction of lactate and Leu, followed by N-Lac-
6 Leu release. N-lactoyl-amino acids might represent useful extracellular biomarkers for
7 intracellular amino acid concentration because they are only formed inside cells.⁶² Interestingly,
8 N-Lac-Leu exhibited the same increasing trend in CM-7 and CM-C as well as in SIV ccRCC
9 serum samples compared to CM-H and healthy individuals, respectively (Table 1, Figure 5).
10 Thus, it could be a more robust potential biomarker to be evaluated in a larger human serum
11 cohort including different stages of disease.
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26

27 **Conclusions**

28
29 In the present study, we developed a protocol for harvesting and extracting extracellular
30 metabolites from an *in vitro* model of human renal cell lines. The exometabolome was profiled
31 using a discovery-based metabolomics approach *via* UPLC-QTOF-MS. The metabolic footprints
32 of ccRCC cell lines and a non-tumor renal cell line were compared with the human serum
33 metabolic profile of SIV ccCRR and healthy individuals. A panel of 21 discriminant features
34 obtained from the binary comparison of control HEK-293 and tumor 786-O cells allowed
35 differentiating kidney cell lines with 100% specificity, sensitivity and accuracy. A subset of 9
36 discriminant features from CM samples was detected in human serum, and produced a
37 fingerprint that enabled discerning stage IV ccRCC patients from healthy individuals. Identified
38 discriminant metabolites suggest alterations in amino acid metabolism, and the redox status of
39 cells. To our knowledge, this is the first time that N-lactoyl-leucine is associated to ccRCC. Even
40 if results in serum need to be validated in a larger cohort, our study highlights the utility of RCC
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 *in vitro* models for finding disease-related discriminant metabolites capable of human serum
4 classification.
5
6

7 8 **Acknowledgements** 9

10
11 We acknowledge CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas,
12 Argentina), and the National Agency of Scientific and Technological Promotion (PRH-PICT-
13 2015-0022 project) for providing the funding; and the “Biobanco Público de Muestras Séricas
14 Oncológicas” (BPMSO) from Instituto de Oncología A. H. Roffo for providing the samples.
15
16 M.E.M., M.O.S. and L.I.P. are Research Staff members from CONICET. M.O.S. acknowledges
17 the National Mass Spectrometry System (SNEM) for the travel grant received for his research
18 stance at CIBION. We also would like to thank Dr. Ana Bellomo and Dr. Mariela Bollini for
19 providing useful guidance regarding the synthesis of chemical standards.
20
21
22
23
24
25
26
27
28
29
30
31
32

33 **Figure Captions** 34 35 36

37 **Figure 1.** Scheme of conditioned media (CM) incubation and collection for the different cell
38 lines studied. O/N: overnight.
39
40
41
42
43

44 **Figure 2.** (A) Principal Component Analysis (PCA) scores plot of conditioned media (CM)
45 samples using the set of 755 spectral features for HEK-293 non-tumor control cell line (blue
46 squares), 786-O cell line (red triangles), Caki-1 cell line (black circles); and quality controls
47 (green diamonds), with a model that consisted of 5 PCs with 70.67% total captured variance. (B)
48 PCA scores plot for HEK-293 (blue squares) and 786-O cell line (red triangles), with a model
49 that consisted of 2 PCs with 53.38% total captured variance. (C) Cross-validated prediction plot
50
51
52
53
54
55
56
57
58
59
60

1
2
3 from the orthogonal projection to latent structures-discriminant analysis model of CM samples
4 from a tumor cell line 786-O (red triangle) versus a non-tumor control cell line HEK 293 (blue
5 squares) using the 21 discriminant metabolic feature panel obtained from genetic algorithm
6
7
8
9
10 variable selection.

11
12
13
14 **Figure 3.** Principal Component Analysis (PCA) scores plot of serum (top panel) and CM (low
15 panel) samples using the set of 163 common spectral features (left) and the set of 9 discriminant
16 features of CM samples (right) that were also present in serum samples. The model consisted of
17 (A) 2 PCs with 57.97% total captured variance; (B) 2 PCs with 70.87% total captured variance;
18 (C) 4 PCs with 69.29% total captured variance, and (D) 2 PC with 65.02% total captured
19 variance. Serum samples from healthy individuals (blue squares) and stage IV ccRCC patients
20 (magenta diamonds); CM samples from the non-tumor control cell line HEK-293 (blue squares);
21 and tumor cell lines 786-O (red triangles), and Caki-1 (black circles).
22
23
24
25
26
27
28
29
30
31
32
33
34
35

36 **Figure 4.** (A) Extracted ion chromatograms for [N-lactoyl-leucine-H]⁻ ion at m/z 202.1079 ±
37 0.0500 generated from conditioned media (CM, dotted line) and serum (solid line), non-spiked
38 (blue), and 150 μM spiked (red) samples; and a 15 μM N-lactoyl-leucine (N-Lac-Leu) standard
39 solution (red dashed-dotted line). (B) Mass spectrum for [N-Lac-Leu-H]⁻ ion with m/z 202.1079
40 in CM sample (red), and its simulated isotopic pattern (black). (C) Product ion mass spectra of
41 [N-Lac-Leu-H]⁻ precursor ion for a CM sample (red), and for a 15 μM N-Lac-Leu standard
42 (green) using a collision cell voltage of 20 V.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **Figure 5.** Box plots for discriminant metabolites in CM and serum samples. Comparisons are
4 shown for CM of HEK-293 (n= 21), 786-O (n= 21) and Caki-1 (n=22) cell lines; and between
5 stage IV ccRCC patients and healthy individuals. Mean values are represented by a filled circle
6 in the box; median values are represented by a line in the box; the upper and lower edges of the
7 box are the 25th and 75th percentiles; the whisker extends to the most extreme values in data, not
8 including outliers defined as 1.5 of the interquartile range and represented by colored crosses.
9
10 Fold changes are calculated as the ratio of median peak areas between CM from 786-O and
11 HEK-293 (CM-7/CM-H), Caki-1 and HEK-293 (CM-C/CM-H), and Caki-1 and 786-O (CM-
12 C/CM-7) samples. Mann-Whitney U tests were used to calculate statistical significance, and *p*
13 values were corrected using the Benjamini-Hochberg procedure for multiple comparisons with a
14 FDR of 0.1. Statistically significant differences after correction for multiple comparisons
15 between classes are indicated on top of the boxes with i) * for CM-7 or CM-C vs CM-H; and ii)
16 # for CM-C vs CM-7. Metabolite identities are indicated for each case.
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 1.

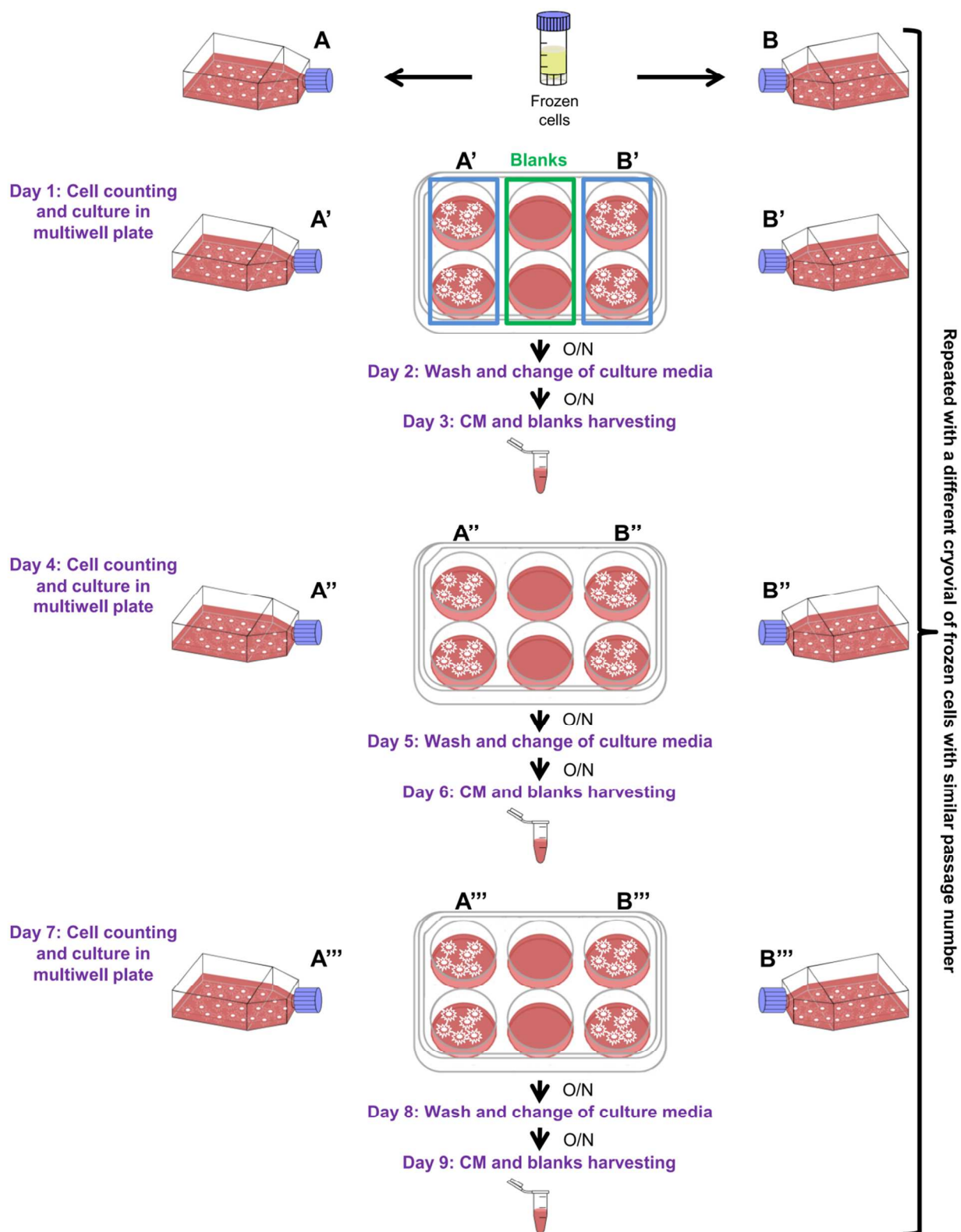


Figure 2.

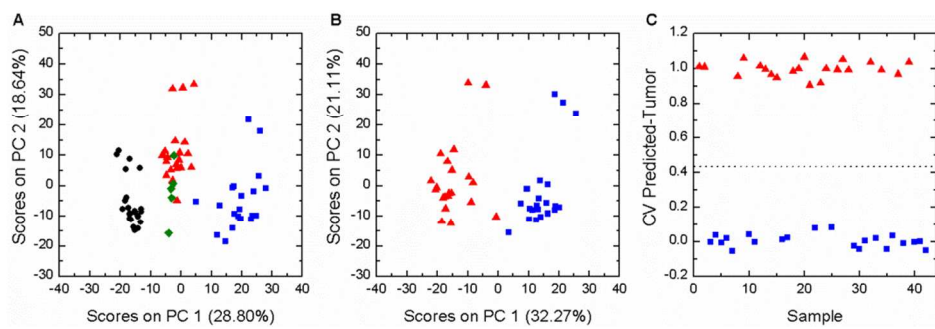


Figure 3.

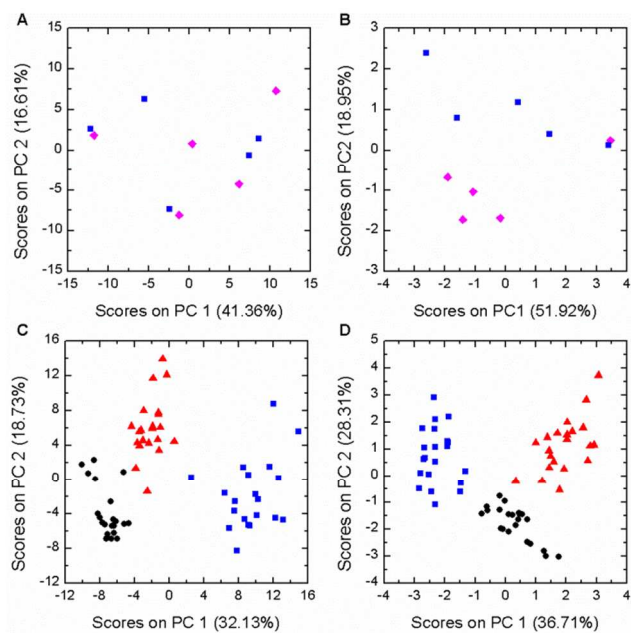


Figure 4.

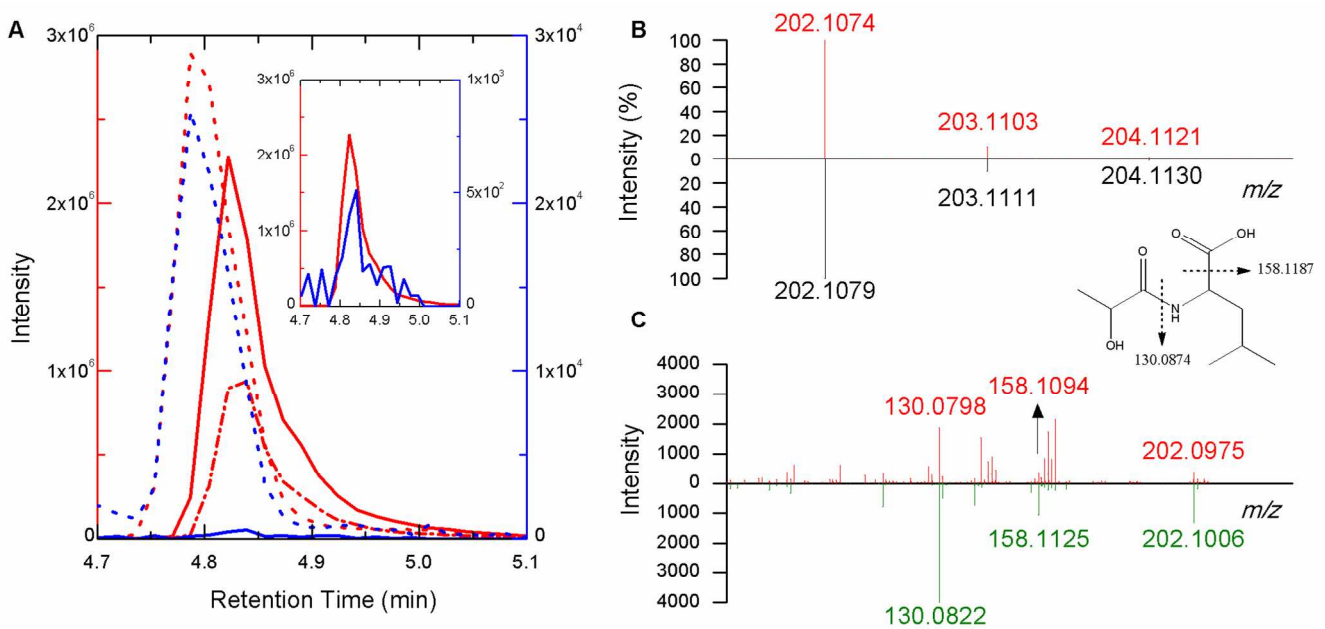


Figure 5.

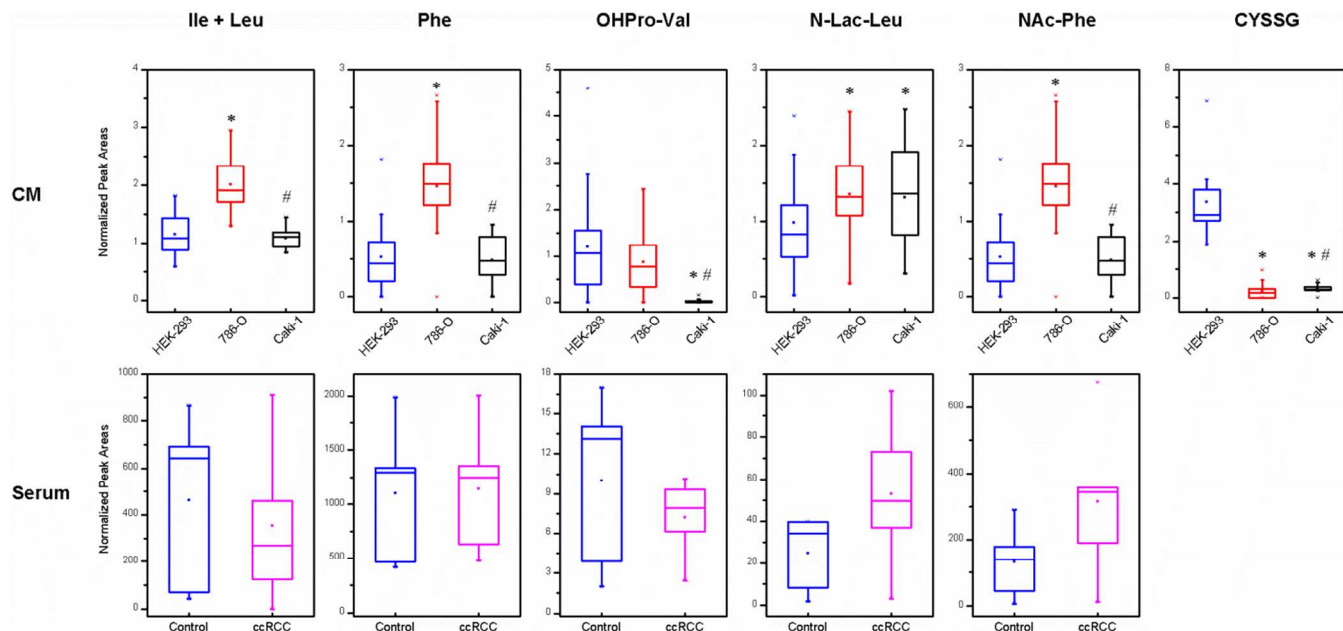


Table 1. Identification of discriminant feature panel, based on accurate mass (a), isotopic pattern (b), MS/MS experiments (c), and retention time match with standards (d). Compounds and features highlighted in bold are detected in serum and metabolites chromatographically confirmed by chemical standards are indicated in italics. Δm is calculated as the mass difference between the m/z value obtained from Progenesis and the theoretical mass. Fold changes are calculated as the ratio of median peak areas between CM from 786-O and HEK-293 (CM-7/CM-H), Caki-1 and HEK-293 (CM-C/CM-H), and Caki-1 and 786-O (CM-C/CM-7) samples; except for cases with median value equal to zero. p values were calculated using Mann-Whitney U tests. NS: non-significant differences after correction with the Benjamini-Hochberg procedure for multiple comparisons with a FDR of 0.1. Trends in binary comparisons are indicated with arrows, (\uparrow): increased levels, and (\downarrow): decreased levels. For CM samples, fold changes of 1-9.9, 10-99.99 and >100 are indicated with one, two or three arrows, respectively. Non-statistically significant trends are indicated for serum samples in the comparison of SIV ccRCC patients and healthy individuals (HI).

Rt (min)	m/z	Ion Type	Elemental Formula	Δm (mDa)	CM-7/CM-H			CM-C/CM-H			CM-C/CM-7			Serum ccRCC/HI	Metabolite Identification	ID Confirmation
					p	Trend	Fold Change	p	Trend	Fold Change	p	Trend	Fold Change			
0.92	425.0825	[M-H] ⁻	C ₁₃ H ₂₂ N ₄ O ₈ S ₂	0.2	2.5x10 ⁻⁸	↓↓	18	2.1x10 ⁻⁸	↓	9.9	6.6x10 ⁻²	↑	1.8	-	<i>Cysteineglutathione-disulfide</i>	a, b, c, d
1.09	225.9854	[M+NaCl-H] ⁻	-	-	2.2x10 ⁻⁸	-	-	-	-	-	1.7x10 ⁻⁸	↓	3.8	-	NO ID	-
1.30	303.9641	[M+3NaCl-H] ⁻	C ₆ H ₁₃ NO ₂	1.4	1.4x10 ⁻⁶	↑	1.8	NS	-	1.0	3.7x10 ⁻⁸	↓	1.7	↓	<i>Isoleucine+Leucine</i>	a, b, c, d
1.67	569.7810	[M+7NaCl-H] ⁻	C ₉ H ₁₁ NO ₂	-0.5	3.5x10 ⁻⁴	↑	3.4	NS	↑	1.2	3.4x10 ⁻⁵	↓↓	2.8	-	<i>Phenylalanine</i>	a, b, c, d
2.07	287.0872	[M+NaCl-H] ⁻	C ₁₀ H ₁₈ N ₂ O ₄	9.7	NS	↓	1.4	1.3x10 ⁻⁶	↓↓↓	270	5.4x10 ⁻⁷	↓↓↓	200	↓	<i>Hydroxyprotyl-valine</i>	a, b, c
2.52	142.0687	[M-H] ⁻	-	-	2.0x10 ⁻⁶	↑	3.1	2.1x10 ⁻⁸	↑	5.7	1.0x10 ⁻⁵	↑	1.8	↓	NO ID	-
3.19	558.9097	[M+4NaCl-H] ⁻	-	-	2.1x10 ⁻⁸	-	-	2.5x10 ⁻⁷	-	-	1.2x10 ⁻³	↓	1.9	-	NO ID	-
3.38	230.9961	[M-H] ⁻	-	-	7.9x10 ⁻⁸	-	-	1.3x10 ⁻⁵	↓	3.6	1.0x10 ⁻⁶	-	-	-	NO ID	-
3.97	180.0137	[M-H] ⁻	-	-	NS	↑	1.6	8.7x10 ⁻⁸	-	-	2.6x10 ⁻⁶	-	-	-	NO ID	-
4.38	283.0891	[M-H] ⁻	-	-	2.0x10 ⁻⁸	-	-	1.0x10 ⁻⁸	-	-	4.3x10 ⁻³	↑	1.3	↑	NO ID	-
4.91	202.1104	[M-H] ⁻	C ₉ H ₁₇ NO ₄	2.5	3.9x10 ⁻²	↑	1.6	6.7x10 ⁻²	↑	1.7	NS	-	1.0	↑	<i>N-lactoyl-leucine</i>	a, b, c, d
5.03	264.0427	[M+NaCl-H] ⁻	C ₁₁ H ₁₃ NO ₃	2.4	3.6x10 ⁻⁸	↑	7.3	NS	↓	1.1	2.8x10 ⁻⁸	↓	7.8	↑	<i>N-acetyl-phenylalanine</i>	a, b, c, d
5.40	319.0772	[M+Cl] ⁻	-	-	3.0x10 ⁻⁸	↑↑	39	9.4x10 ⁻⁸	↑↑	14	2.8x10 ⁻⁵	↓	2.8	-	NO ID	-
6.34	206.0815	[M-H] ⁻	-	-	2.1x10 ⁻⁸	-	-	1.9x10 ⁻⁸	↓↓↓	200	-	-	-	-	NO ID	-
6.58	138.0199	[M-H] ⁻	-	-	3.1x10 ⁻⁸	↓	5.4	2.1x10 ⁻⁸	↓	6.3	7.4x10 ⁻²	↓	1.2	↓	NO ID	-
6.81	213.1167	[M-H] ⁻	-	-	4.1x10 ⁻⁵	↓	2.2	1.9x10 ⁻⁷	↓	5.1	1.9x10 ⁻²	↓	2.4	-	NO ID	-
7.90	420.0948	[M-H] ⁻	-	-	9.6x10 ⁻³	↑	1.6	1.8x10 ⁻⁹	-	-	1.8x10 ⁻⁹	-	-	-	NO ID	-
8.36	333.1066	[M-H] ⁻	-	-	4.8x10 ⁻³	↑	2.4	NS	↓	1.6	3.1x10 ⁻⁴	↓	3.9	-	NO ID	-
8.41	211.1355	[M-H] ⁻	-	-	7.2x10 ⁻⁷	↑	7.6	2.6x10 ⁻⁴	↑	3.0	1.1x10 ⁻⁴	↓	2.5	↓	NO ID	-
9.44	177.0970	[M+Cl] ⁻	-	-	4.1x10 ⁻⁵	↓	1.5	8.7x10 ⁻⁷	↓	2.1	7.8x10 ⁻²	↓	1.4	-	NO ID	-
11.29	291.1608	[M-H] ⁻	-	-	7.2x10 ⁻⁷	↓	4.3	2.3x10 ⁻⁶	↓	3.4	NS	↓	1.3	-	NO ID	-

Supporting Information

The following Supporting Information is available free of charge at ACS website

<http://pubs.acs.org>.

Data Set S1. Metabolic feature matrix for conditioned media samples.

Materials and Methods. Organic Synthesis of N-acetyl-phenylalanine and N-lactoyl-leucine.

Figure S1. ^1H NMR and ^{13}C NMR spectra (500 MHz, D_2O) of N-acetyl-phenylalanine.

Figure S2. ^1H NMR and ^{13}C NMR spectra (500 MHz, D_2O) of N-lactoyl-leucine.

Figure S3. Representative base peak intensity chromatograms obtained for conditioned media and serum samples.

Figure S4. Principal Component Analysis scores plot of conditioned media (CM) samples using the set of 9 metabolic features from the discriminant CM panel that were present in serum samples.

Figure S5. Identification of N-acetyl-phenylalanine.

Figure S6. Identification of phenylalanine.

Figure S7. Identification of isoleucine/leucine.

Figure S8. Identification of hydroxypropyl-valine.

Figure S9. Identification of cysteineglutathione disulfide.

Table S1. Conditioned media samples collected at different time points based on the cell culture protocol.

References

- (1) Linehan, W. M.; Srinivasan, R.; Schmidt, L. S., The genetic basis of kidney cancer: a metabolic disease. *Nat. Rev. Urol.* **2010**, *7*, 277-85.
- (2) Hsieh, J. J.; Purdue, M. P.; Signoretti, S.; Swanton, C.; Albiges, L.; Schmidinger, M.; Heng, D. Y.; Larkin, J.; Ficarra, V., Renal cell carcinoma. *Nat. Rev. Dis. Primers* **2017**, *3*, 17009.
- (3) Linehan, W. M.; Bratslavsky, G.; Pinto, P. A.; Schmidt, L. S.; Neckers, L.; Bottaro, D. P.; Srinivasan, R., Molecular Diagnosis and Therapy of Kidney Cancer. *Annu. Rev. Med.* **2010**, *61*, 329-343.

- 1
2
3 (4) IARC, Tumors of the kidney. In *World Health Organization Classification of Tumors. Pathology and Genetics of Tumours of the Urinary System and Male Genital Organs*, Eble, J. N.; Sauter, G.; Epstein, J. I.; Sesterhenn, I. A., Eds. IARC press: Lyon, 2004; pp 9-43.
- 4
5 (5) Gibbons, R. P.; Monte, J. E.; Correa, R. J., Jr.; Mason, J. T., Manifestations of renal cell carcinoma. *Urology* **1976**, *8*, 201-6.
- 6
7 (6) Hu, B.; Lara, P. N., Jr.; Evans, C. P., Defining an Individualized Treatment Strategy for Metastatic Renal Cancer. *Urol. Clin. North Am.* **2012**, *39*, 233-249.
- 8
9 (7) Graves, A.; Hessamodini, H.; Wong, G.; Lim, W. H., Metastatic renal cell carcinoma: update on epidemiology, genetics, and therapeutic modalities. *Immunotargets Ther.* **2013**, *2*, 73-90.
- 10
11 (8) Diamond, E.; Molina, A. M.; Carbonaro, M.; Akhtar, N. H.; Giannakakou, P.; Tagawa, S. T.; Nanus, D. M., Cytotoxic chemotherapy in the treatment of advanced renal cell carcinoma in the era of targeted therapy. *Crit. Rev. Oncol. Hematol.* **2015**, *96*, 518-526.
- 12
13 (9) De Meerleer, G.; Khoo, V.; Escudier, B.; Joniau, S.; Bossi, A.; Ost, P.; Briganti, A.; Fonteyne, V.; Van Vulpen, M.; Lumen, N.; Spahn, M.; Mareel, M., Radiotherapy for renal-cell carcinoma. *Lancet Oncol.* **2014**, *15*, e170-e177.
- 14
15 (10) Rodriguez-Vida, A.; Hutson, T. E.; Bellmunt, J.; Strijbos, M. H., New treatment options for metastatic renal cell carcinoma. *ESMO Open* **2017**, *2*, e000185.
- 16
17 (11) Hakimi, A. A.; Pham, C. G.; Hsieh, J. J., A clear picture of renal cell carcinoma. *Nat. Genet.* **2013**, *45*, 849-850.
- 18
19 (12) Calzada, M. J.; Esteban, M. A.; Feijoo-Cuaresma, M.; Castellanos, M. C.; Naranjo-Suarez, S.; Temes, E.; Mendez, F.; Yanez-Mo, M.; Ohh, M.; Landazuri, M. O., von Hippel-Lindau tumor suppressor protein regulates the assembly of intercellular junctions in renal cancer cells through hypoxia-inducible factor-independent mechanisms. *Cancer Res.* **2006**, *66*, 1553-60.
- 20
21 (13) Banks, R. E.; Tirukonda, P.; Taylor, C.; Hornigold, N.; Astuti, D.; Cohen, D.; Maher, E. R.; Stanley, A. J.; Harnden, P.; Joyce, A.; Knowles, M.; Selby, P. J., Genetic and epigenetic analysis of von Hippel-Lindau (VHL) gene alterations and relationship with clinical variables in sporadic renal cancer. *Cancer Res.* **2006**, *66*, 2000-11.
- 22
23 (14) Godinot, C.; de Laplanche, E.; Hervouet, E.; Simonnet, H., Actuality of Warburg's views in our understanding of renal cancer metabolism. *J. Bioenerg. Biomembr.* **2007**, *39*, 235-41.
- 24
25 (15) Pinthus, J. H.; Whelan, K. F.; Gallino, D.; Lu, J.-P.; Rothschild, N., Metabolic features of clear-cell renal cell carcinoma: mechanisms and clinical implications. *Can. Urol. Assoc. J.* **2011**, *5*, 274-282.
- 26
27 (16) Gebhard, R. L.; Clayman, R. V.; Prigge, W. F.; Figenshau, R.; Staley, N. A.; Reese, C.; Bear, A., Abnormal cholesterol metabolism in renal clear cell carcinoma. *J. Lipid Res.* **1987**, *28*, 1177-84.
- 28
29 (17) Cuperlovic-Culf, M.; Barnett, D. A.; Culf, A. S.; Chute, I., Cell culture metabolomics: applications and future directions. *Drug Discov. Today* **2010**, *15*, 610-21.
- 30
31 (18) León, Z.; García-Cañaveras, J. C.; Donato, M. T.; Lahoz, A., Mammalian cell metabolomics: Experimental design and sample preparation. *Electrophoresis* **2013**, *34*, 2762-2775.
- 32
33 (19) Fiehn, O.; Kopka, J.; Dormann, P.; Altmann, T.; Trethewey, R. N.; Willmitzer, L., Metabolite profiling for plant functional genomics. *Nat. Biotechnol.* **2000**, *18*, 1157-61.
- 34
35 (20) Nicholson, J. K.; Lindon, J. C., Systems biology: Metabonomics. *Nature* **2008**, *455*, 1054-6.
- 36
37 (21) Kim, K.; Taylor, S. L.; Ganti, S.; Guo, L. N.; Osier, M. V.; Weiss, R. H., Urine Metabolomic Analysis Identifies Potential Biomarkers and Pathogenic Pathways in Kidney Cancer. *OMICS* **2011**, *15*, 293-303.
- 38
39 (22) Ganti, S.; Taylor, S. L.; Kim, K.; Hoppel, C. L.; Guo, L. N.; Yang, J.; Evans, C.; Weiss, R. H., Urinary acylcarnitines are altered in human kidney cancer. *Int. J. Cancer* **2012**, *130*, 2791-2800.
- 40
41 (23) Ganti, S.; Taylor, S. L.; Abu Aboud, O.; Yang, J.; Evans, C.; Osier, M. V.; Alexander, D. C.; Kim, K.; Weiss, R. H., Kidney Tumor Biomarkers Revealed by Simultaneous Multiple Matrix Metabolomics Analysis. *Cancer Res.* **2012**, *72*, 3471-3479.
- 42
43 (24) Lin, L.; Huang, Z. Z.; Gao, Y.; Yan, X. M.; Xing, J. C.; Hang, W., LC-MS based serum metabonomic analysis for renal cell carcinoma diagnosis, staging, and biomarker discovery. *J. Proteome Res.* **2011**, *10*, 1396-1405.
- 44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- (25) Hakimi, A. A.; Reznik, E.; Lee, C.-H.; Creighton, Chad J.; Brannon, A. R.; Luna, A.; Aksoy, B. A.; Liu, Eric M.; Shen, R.; Lee, W.; Chen, Y.; Stirdivant, Steve M.; Russo, P.; Chen, Y.-B.; Tickoo, Satish K.; Reuter, Victor E.; Cheng, Emily H.; Sander, C.; Hsieh, James J., An Integrated Metabolic Atlas of Clear Cell Renal Cell Carcinoma. *Cancer Cell* **2016**, *29*, 104-116.
- (26) Kell, D. B.; Brown, M.; Davey, H. M.; Dunn, W. B.; Spasic, I.; Oliver, S. G., Metabolic footprinting and systems biology: the medium is the message. *Nat. Rev. Microbiol.* **2005**, *3*, 557-565.
- (27) Halama, A., Metabolomics in cell culture--a strategy to study crucial metabolic pathways in cancer development and the response to treatment. *Arch. Biochem. Biophys.* **2014**, *564*, 100-9.
- (28) Sinha, R.; Winer, A. G.; Chevinsky, M.; Jakubowski, C.; Chen, Y. B.; Dong, Y.; Tickoo, S. K.; Reuter, V. E.; Russo, P.; Coleman, J. A.; Sander, C.; Hsieh, J. J.; Hakimi, A. A., Analysis of renal cancer cell lines from two major resources enables genomics-guided cell line selection. *Nat Commun.* **2017**, *8*, 15165.
- (29) Zheng, X.; Baker, H.; Hancock, W. S.; Fawaz, F.; McCaman, M.; Pungor, E., Jr., Proteomic analysis for the assessment of different lots of fetal bovine serum as a raw material for cell culture. Part IV. Application of proteomics to the manufacture of biological drugs. *Biotechnol. Progr.* **2006**, *22*, 1294-300.
- (30) Broadhurst, D.; Goodacre, R.; Reinke, S. N.; Kuligowski, J.; Wilson, I. D.; Lewis, M. R.; Dunn, W. B., Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics* **2018**, *14*, 72.
- (31) Bateman, K. P.; Castro-Perez, J.; Wrona, M.; Shockcor, J. P.; Yu, K.; Oballa, R.; Nicoll-Griffith, D. A., MSE with mass defect filtering for in vitro and in vivo metabolite identification. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 1485-96.
- (32) Kirwan, J. A.; Weber, R. J. M.; Broadhurst, D. I.; Viant, M. R., Direct infusion mass spectrometry metabolomics dataset: a benchmark for data processing and quality control. *Scientific Data* **2014**, *1*, 140012.
- (33) Daskalaki, E.; Pillon, N. J.; Krook, A.; Wheelock, C. E.; Checa, A., The influence of culture media upon observed cell secretome metabolite profiles: The balance between cell viability and data interpretability. *Anal. Chim. Acta* **2018**, DOI: 10.1016/j.aca.2018.04.034.
- (34) Trygg, J.; Wold, S., Orthogonal projections to latent structures (O-PLS). *J. Chemometr.* **2002**, *16*, 119-128.
- (35) Bylesjö, M.; Rantalainen, M.; Cloarec, O.; Nicholson, J. K.; Holmes, E.; Trygg, J., OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J. Chemometr.* **2006**, *20*, 341-351.
- (36) Benjamini, Y.; Hochberg, Y., Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Series B, Stat. Methodol.* **1995**, *57*, 289-300.
- (37) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; Fung, C.; Nikolai, L.; Lewis, M.; Coutouly, M. A.; Forsythe, I.; Tang, P.; Shrivastava, S.; Jeroncic, K.; Stothard, P.; Amegbey, G.; Block, D.; Hau, D. D.; Wagner, J.; Miniaci, J.; Clements, M.; Gebremedhin, M.; Guo, N.; Zhang, Y.; Duggan, G. E.; Macinnis, G. D.; Weljie, A. M.; Dowlatabadi, R.; Bamforth, F.; Clive, D.; Greiner, R.; Li, L.; Marrie, T.; Sykes, B. D.; Vogel, H. J.; Querengesser, L., HMDB: the Human Metabolome Database. *Nucleic Acids Res.* **2007**, *35*, D521-6.
- (38) Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G., METLIN: a metabolite mass spectral database. *Ther. Drug Monit.* **2005**, *27*, 747-51.
- (39) Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.; Sawada, Y.; Hirai, M. Y.; Nakanishi, H.; Ikeda, K.; Akimoto, N.; Maoka, T.; Takahashi, H.; Ara, T.; Sakurai, N.; Suzuki, H.; Shibata, D.; Neumann, S.; Iida, T.; Tanaka, K.; Funatsu, K.; Matsuura, F.; Soga, T.; Taguchi, R.; Saito, K.; Nishioka, T., MassBank: a public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.* **2010**, *45*, 703-714.

- 1
2
3 (40) Wild, C. P., Complementing the Genome with an “Exposome”: The Outstanding Challenge of
4 Environmental Exposure Measurement in Molecular Epidemiology. *Cancer Epidemiol. Biomarkers*
5 *Prev.* **2005**, *14*, 1847-1850.
- 6 (41) Rappaport, S. M.; Smith, M. T., Environment and Disease Risks. *Science* **2010**, *330*, 460-461.
- 7 (42) Nicholson, J. K.; Holmes, E.; Kinross, J.; Burcelin, R.; Gibson, G.; Jia, W.; Pettersson, S., Host-gut
8 microbiota metabolic interactions. *Science* **2012**, *336*, 1262-7.
- 9 (43) Kebarle, P.; Verkerk Udo, H., Electrospray: From ions in solution to ions in the gas phase, what
10 we know now. *Mass Spectrom. Rev.* **2009**, *28*, 898-917.
- 11 (44) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S., Unraveling the Mechanism of Electrospray
12 Ionization. *Anal. Chem.* **2013**, *85*, 2-9.
- 13 (45) Johnson, C. H.; Ivanisevic, J.; Siuzdak, G., Metabolomics: beyond biomarkers and towards
14 mechanisms. *Nat. Rev. Mol. Cell Biol.* **2016**, *17*, 451-459.
- 15 (46) Eriksson, B.; Eriksson, S. A., Synthesis and characterization of the L-cysteine-glutathione mixed
16 disulfide. *Acta Chem. Scand.* **1967**, *21*, 1304-1312.
- 17 (47) Giustarini, D.; Colombo, G.; Garavaglia, M. L.; Astori, E.; Portinaro, N. M.; Reggiani, F.;
18 Badalamenti, S.; Aloisi, A. M.; Santucci, A.; Rossi, R.; Milzani, A.; Dalle-Donne, I., Assessment of
19 glutathione/glutathione disulphide ratio and S-glutathionylated proteins in human blood, solid
20 tissues, and cultured cells. *Free Radical Biol. Med.* **2017**, *112*, 360-375.
- 21 (48) Chakravarthi, S.; Jessop, C. E.; Bulleid, N. J., The role of glutathione in disulphide bond
22 formation and endoplasmic-reticulum-generated oxidative stress. *EMBO Rep.* **2006**, *7*, 271-5.
- 23 (49) Meister, A.; Anderson, M. E., Glutathione. *Annu. Rev. Biochem.* **1983**, *52*, 711-60.
- 24 (50) Estrela, J. M.; Ortega, A.; Obrador, E., Glutathione in cancer biology and therapy. *Crit. Rev. Clin.*
25 *Lab. Sci.* **2006**, *43*, 143-81.
- 26 (51) Lu, S. C., Glutathione synthesis. *Biochim. Biophys. Acta* **2013**, *1830*, 3143-53.
- 27 (52) Ballatori, N.; Krance, S. M.; Marchan, R.; Hammond, C. L., Plasma membrane glutathione
28 transporters and their roles in cell physiology and pathophysiology. *Mol. Aspects Med.* **2009**, *30*, 13-
29 28.
- 30 (53) Ganesamoni, R.; Bhattacharyya, S.; Kumar, S.; Chauhan, A.; Mete, U. K.; Agarwal, M. M.;
31 Mavuduru, R.; Kaushik, G.; Mandal, A. K.; Singh, S. K., Status of oxidative stress in patients with renal
32 cell carcinoma. *J. Urol.* **2012**, *187*, 1172-6.
- 33 (54) Tew, K. D.; Monks, A.; Barone, L.; Rosser, D.; Akerman, G.; Montali, J. A.; Wheatley, J. B.;
34 Schmidt, D. E., Jr., Glutathione-associated enzymes in the human cell lines of the National Cancer
35 Institute Drug Screening Program. *Mol. Pharmacol.* **1996**, *50*, 149-59.
- 36 (55) Li, B.; Qiu, B.; Lee, D. S. M.; Walton, Z. E.; Ochocki, J. D.; Mathew, L. K.; Mancuso, A.; Gade, T. P.
37 F.; Keith, B.; Nissim, I.; Simon, M. C., Fructose-1,6-bisphosphatase opposes renal carcinoma
38 progression. *Nature* **2014**, *513*, 251.
- 39 (56) Goldstein, J. L.; Anderson, R. G.; Brown, M. S., Coated pits, coated vesicles, and receptor-
40 mediated endocytosis. *Nature* **1979**, *279*, 679-85.
- 41 (57) Mustafa, A.; Gupta, S.; Hudes, G. R.; Egleston, B. L.; Uzzo, R. G.; Kruger, W. D., Serum amino acid
42 levels as a biomarker for renal cell carcinoma. *J. Urol.* **2011**, *186*, 1206-12.
- 43 (58) Zhang, F.; Ma, X.; Li, H.; Guo, G.; Li, P.; Gu, L.; Li, X.; Chen, L.; Zhang, X., The predictive and
44 prognostic values of serum amino acid levels for clear cell renal cell carcinoma. *Urol. Oncol.* **2017**, *35*,
45 392-400.
- 46 (59) Gerlo, E.; Van Coster, R.; Lissens, W.; Winckelmans, G.; De Meirleir, L.; Wevers, R., Gas
47 chromatographic-mass spectrometric analysis of N-acetylated amino acids: the first case of
48 aminoacylase I deficiency. *Anal. Chim. Acta* **2006**, *571*, 191-9.
- 49 (60) Kelley, M.; Vessey Donald, A., Isolation and characterization of mitochondrial acyl - coa:
50 Glycine n - acyltransferases from kidney. *J. Biochem. Mol. Toxicol.* **1993**, *8*, 63-69.
- 51 (61) Wettersten, H. I.; Aboud, O. A.; Lara, P. N., Jr.; Weiss, R. H., Metabolic reprogramming in clear
52 cell renal cell carcinoma. *Nat. Rev. Nephrol.* **2017**, *13*, 410-419.
- 53
54
55
56
57
58
59
60

1
2
3 (62) Jansen, R. S.; Addie, R.; Merkx, R.; Fish, A.; Mahakena, S.; Bleijerveld, O. B.; Altelaar, M.; Ijlst, L.;
4 Wanders, R. J.; Borst, P.; van de Wetering, K., N-lactoyl-amino acids are ubiquitous metabolites that
5 originate from CNDP2-mediated reverse proteolysis of lactate and amino acids. *Proc. Natl. Acad. Sci.*
6 *U. S. A.* **2015**, *112*, 6601.
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For TOC only

