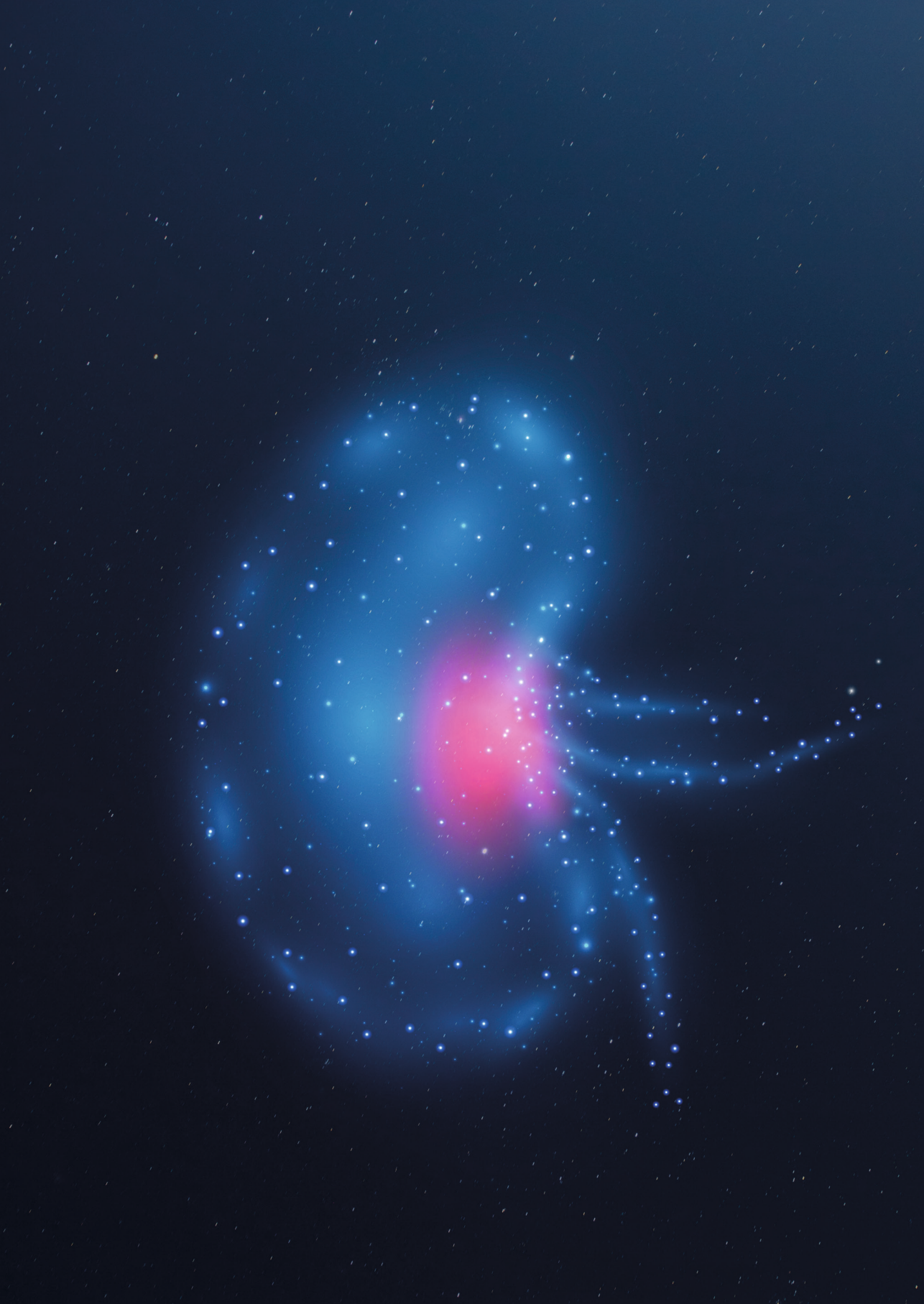
Exploration of renal space: navigating injury and repair through spatial omics

Rietjens, R.G.J.



CHAPTER 5

Investigating the Warburg effect in renal cell carcinoma using spatial dynamic metabolomics

ROSALIE G.J. RIETJENS, GANGQI WANG, BRAM HEIJS

A Practical Guide to Metabolomics Applications in Health and Disease 2023;389-425

1. LEARNING OBJECTIVES

- Spatial lipidomics & metabolomics using mass spectrometry imaging.
- Isotope tracing in biological tissues.
- Visualizing the Warburg effect in cancer tissues.

2. THEORETICAL BACKGROUND

This chapter focuses on the application of stable-isotope tracing in mass spectrometry imaging to unravel changes in the metabolic profile of renal cell carcinoma. Using a pre-recorded dataset, we will illustrate how to use spatial lipidomics data for spatial segmentation of the tissue, and subsequently explain the data analysis strategies for dynamic metabolic measurements. Essential insights into the theory of mass spectrometry (imaging), stable isotope tracing, as well as the human kidney, its metabolism and expected changes upon renal cell carcinoma will be elaborated.

2.1. Spatial metabolomics and lipidomics using mass spectrometry imaging

First, you will find a short introduction into the basic principles underlying spatial metabolomics/lipidomics analysis using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry imaging (MALDI-TOF-MSI). By the end of this section, you will understand the concepts behind this technique and its use in metabolomics/lipidomics, how samples should be treated for MSI, and some of the pitfalls to avoid when designing an experiment.

2.1.1. Mass spectrometry

Many analytical chemistry approaches are based on molecular detection, identification and quantification by mass spectrometry (MS). Although MS encompasses a large variety of technologies, virtually all commercially available MS instruments share the same basic layout, consisting of an ionization source, one or more mass analyzers, a detector, a vacuum system and a computer for instrument setup and data acquisition. There is a large variety in each of these main components, and different combinations will determine important practical and analytical characteristics of specific MS instruments, such as the use of solid *vs.* liquid samples, fragmentation type, resolving power, mass accuracy, and sensitivity. These considerations consequently affect the applications for which a specific mass spectrometer can be used. It is beyond the scope of this introduction to go in-depth into all different MS platforms, but a few essentials will be explained in following sections.

2.1.1.1. Mass spectra

The output of a mass spectrometer is commonly represented as a mass spectrum; the value on the horizontal axis of a mass spectrum reflects the mass-to-charge ratio (*m/z*) of the detected gas-phase analyte ions, and the value on the vertical axis represents

the intensity which is a measure of the abundance of the analyte ion in the analyzed sample. One can make a broad distinction between two types of mass spectra; i) MS1 spectra and ii) fragmentation spectra, or MS*n* spectra, in which *n* represents the number of subsequent fragmentations. MS1 spectra provide a broad overview of all detected compounds in an analyzed sample and their relative quantity compared to each other. An MS1 spectrum can only link an *m/z* feature to a chemical composition (i.e. $C_8H_{15}NO_6$ for *N*-acetylhexosamine). MS*n* spectra provide structural information on selected, isolated, and fragmented compounds, which can be used to identify the molecular structure (e.g. the distinction between the isomeric *N*-acetylgalactosamine and *N*-acetylglucosamine).[1]

2.1.1.2. Mass spectrometer performance

Several important analytical characteristics can be determined from a mass spectrum; **i) mass accuracy:** the difference between the measured mass of a compound and the theoretical mass derived from its chemical formula usually expressed in absolute numbers (10^{-3} u, or mDa), or relative numbers (parts per million, or ppm); **ii) resolving power:** the ability of a mass spectrometer to distinguish two peaks of equal height with a slightly different *m/z*:

$$\boldsymbol{R} = \frac{\boldsymbol{m/z}}{\boldsymbol{\delta m/z}} = \text{resolving power}$$

The smallest peak separation ($\delta m/z$) at which the two peaks can be separated is called resolution, and is defined as the width of a peak, at 50% of its maximum peak height; **iii) sensitivity:** the response of the recorded signal to a change in concentration of the measured analyte. Inherently, mass spectrometers produce and record noise coming from both electrical and chemical interferences. The presence of noise calls for a threshold to distinguish true signals from background noise, the signal-to-noise ratio (S/N).

Since the chapter focuses on the use of MALDI-TOF-MSI, a brief introduction to the MALDI mechanism and TOF mass spectrometry principles will follow. For more details on other ionization methods, analyzers, and fragmentation, see the ***Further reading*** section.

2.1.2. Matrix-assisted laser desorption/ionization

The main function of the ionization source is to convert the from the solid or liquid molecular analytes contained by the sample into gas phase ions; cations in positive ion-mode and anions in negative ion-mode. MSI requires an ionization source that can directly probe and produce ions from a solid sample.[2] One of the most common ionization methods able to directly convert analytes from solid phase molecules to gas phase ions is MALDI. This ionization strategy is based on the illumination of a matrix-doped sample with a pulsed UV laser (Figure 1A). The chemical matrix used is typically a small organic molecule dissolved in organic solvent and has to be a strong absorber of UV light at the wavelength of the laser (Figure 1B). During the evaporation of the organic solvent, the matrix crystallizes, and molecular analytes are embedded and co-crystallized with the matrix. Upon illumination of the matrix with the laser, rapid super heating causes both desorption of single surface

(matrix) ions, as well as an explosive phase transition (ablation) creating larger clusters of neutral and charged matrix and analyte molecules, the combination of which is referred to as the MALDI plume. In the MALDI plume, charges will be transferred mainly through the addition or removal of protons (H^+), or the addition metal ions (Na^+, K^+, Li^+, Ag^+) or halogens (Cl^-) resulting in either cations or anions which are accelerated towards the mass analyzer, which will be discussed in the next section. It is important to note that the majority of ions generated in during the MALDI process only carry single charges. [3]

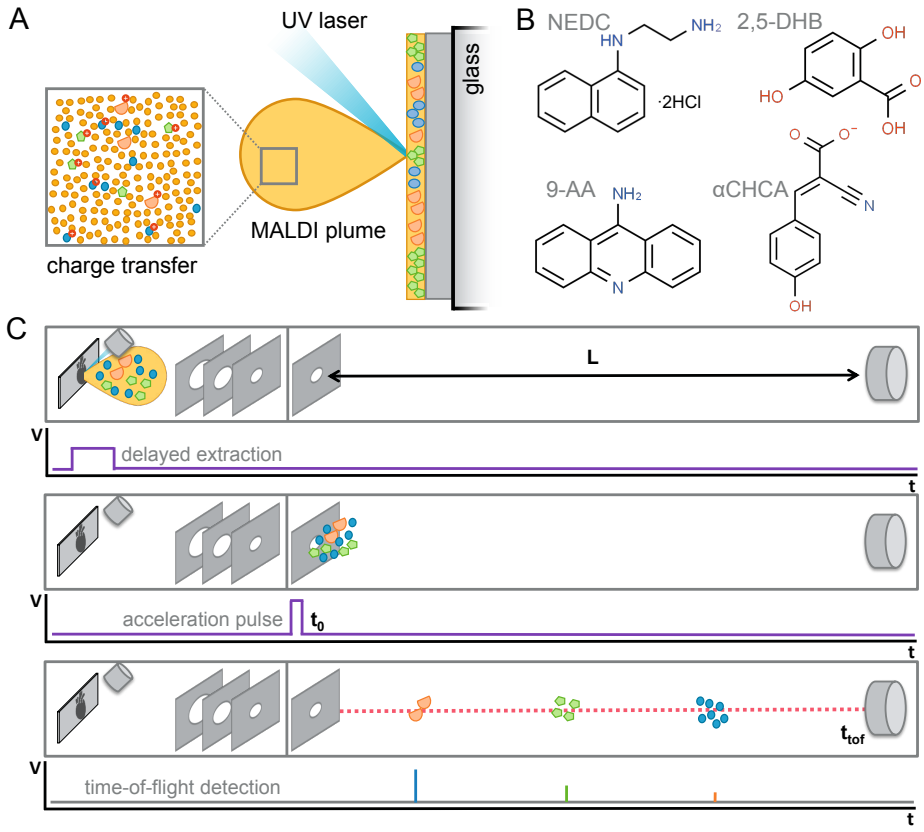


Figure 1. Schematic of the MALDI process. A: In MALDI-MS the analyte molecules are mixed with a chemical matrix and illuminated with a UV laser. In the resulting gas phase ion cloud, the MALDI plume, charge transfer and secondary ionization processes take place, creating gas phase analyte ions. B: Common MALDI matrices used for negative ion-mode metabolomic MALDI-MSI *N*-(1-Naphthyl)ethylenediamine dihydrochloride (NEDC), and 9-aminoacridine (9-AA), and positive ion-mode 2,5-dihydroxybenzoic acid (2,5-DHB), and α-cyano-4-hydroxycinammic acid (αCHCA). C: Schematic of TOF-MS. In a TOF mass analyser, the ions are transferred from the ionization source to a vacuum drift tube. At t_0 the ions are exposed to an electrostatic pulse, accelerating them towards the detector (at distance L). A difference in their resulting velocity separates the ions in space and time. The *m/z* can be calculated for each analyte with a differential t_{tof}.

2.1.4. Time-of-flight mass spectrometry

One of the simpler mass analyzers to comprehend is the axial time-of-flight, or TOF, mass spectrometer (Figure 1C). TOF mass analyzers are pulsed systems, and therefore perfectly compatible with MALDI-based ion generation. The ions produced in the MALDI process are transferred from the sample target in the ionization source to the ion optics by means of a strong electric field between the sample target and the first counter electrode in the optics. The ions are accelerated into a drift tube, which they enter all having the same kinetic energy. The time-of-flight (t_{tof}) can be defined as the time interval between the MALDI laser pulse, and the impact of the ion on the detector. The *m/z* for each ion can be calculated using the following formula:

$$\frac{m}{z} = \left(\frac{L^2}{2eV}\right) t_{tof}^2$$

In which the constants L, the length of the ion flight path, and eV, the electrostatic potential of the accelerating pulse are within brackets, and the t_{tof} is the measured time-of-flight. It can easily be deduced that the higher the molecular weight of the analyte, the longer its time-of-flight. [4]

2.1.5. Mass spectrometry imaging

MSI is based on the acquisition of spatially correlated mass spectra from discrete positions in a Cartesian coordinate system virtually projected on a sample surface. Each recorded spectrum is barcoded with an XY-coordinate and placed in a virtual data cube, in which the X and Y axes represent the X and Y coordinates, and the Z-axis represents the *m/z* axis of the mass spectrum. Each individual voxel, or 3D pixel, in this data cube contains the intensity of a single *m/z* feature at the given XY-coordinate. Through the selection of a specific *m/z* feature, representing an analyte, one can visualize the intensity distribution of the analyte over the sample surface (Figure 2). [5]

2.1.6. Sample preparation for in situ metabolomics using MALDI-MSI

Direct molecular imaging by MALDI-MSI is one of the most common tools for *in situ* metabolomics. While there are many applications beyond metabolomics (i.e. proteomics and glycomics) and clinical research (i.e. food, insect and plant biochemistry), the majority of applications focuses on the analysis of mammalian tissues. Metabolomics by MALDI-MSI is commonly applied to thin sections obtained from fresh frozen tissue material, although analysis of metabolites from formalin-fixed and paraffin-embedded material also has been reported[6], and this immediately poses the two main challenges in the studying the metabolome in its spatial context; i) **post-mortem degradation**, and ii) **molecular delocalization** through lateral diffusion. [7,8]

2.1.6.1. Post-mortem degradation

The metabolome is extremely dynamic and alters rapidly upon changes in the environment. Resecting a tissue specimen (i.e. an organ or tumor) from a human or animal body inevitably

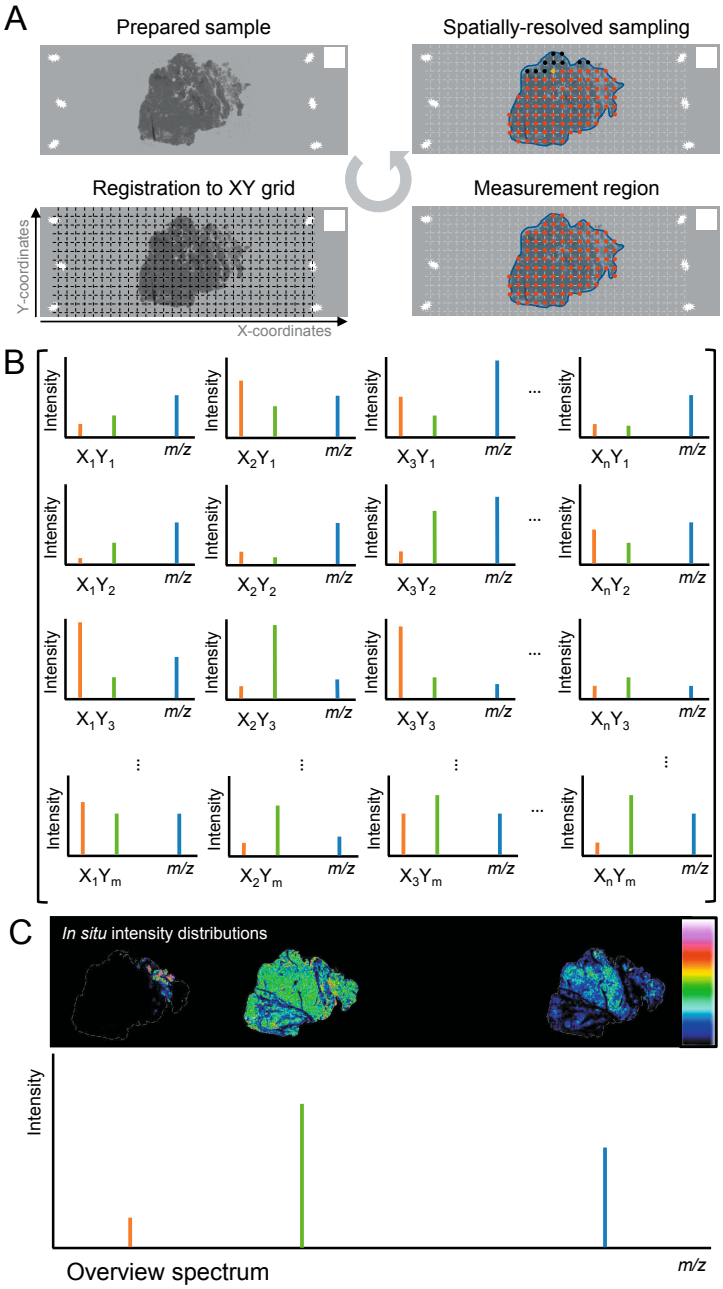


Figure 2. Schematic representation of the mass spectrometry imaging (MSI) principle. A: Registration of the prepared sample to a Cartesian grid with predefined raster width. Followed by definition of the measurement area and spatially resolved sampling. B: Construction of a data cube from the collection of single mass spectra recorded for every pixel coordinate in the defined measurement area**.** C: Calculation of a representative overview spectrum, and visualization of intensity distributions for *m/z* features of interest. The colour scale represents the relative intensity differences of the selected *m/z* feature between the measured pixels.

requires the disconnection of that tissue from the blood flow, causing nutrient and oxygen depletion which immediately start to affect the metabolome.[9] Especially the intracorporeal ischemia time, during which the tissue is still inside the body but disconnected from blood flow, is problematic due to the tissue being present at the optimal working temperature of all endogenous enzymes. Naturally, for a representative metabolomics study, this time should be minimized upon sample collection. Once resected, a tissue should be handled swiftly and is typically flash frozen in seconds using liquid nitrogen-cooled isopentane.

2.1.6.2. Molecular delocalization

MSI is used to analyze molecules in the spatial context of the tissue, thus the localization of the metabolites is of utmost importance. Most metabolites are polar molecules and easily dissolve in water, which makes avoiding condensation one of the primary objectives during the entire MALDI-MSI sample preparation.[8] Tissues, stored at -80°C, should be transferred on dry ice at all times, and equilibrated to room temperature using a vacuum freeze-drier prior to MALDI matrix application. Once at room temperature, the slide-mounted sections should be handled swiftly. After taking a pre-MSI optical scan of the glass slide, required for setting up the virtual Cartesian coordinate system the spatially correlated analysis will be based on, the sample preparation for MALDI-MSI typically only involves applying the MALDI matrix.

2.1.6.3. MALDI matrix application

The application of the MALDI matrix is typically done in one of two ways: spray-based (Figure 3A), or sublimation-based (Figure 3B) matrix application. For the spray-based matrix application, the MALDI matrix should be dissolved. For metabolomics approaches this is typically done in a high-organic solvent, minimizing the amount of water to limit delocalization. The preparation of the high-organic solvent is a balancing act, since there should still be some water present to extract the polar metabolites from the tissue. For example, for the measurements described below, the *N*-(1-naphthyl) ethylenediamine dihydrochloride (NEDC) matrix was dissolved in methanol:acetonitrile:water (70:25:5 % v/v/v). The dissolved matrix is then homogeneously sprayed over the tissue using a robot. The nebulization of the matrix solution into a fine spray is achieved pneumatically, using ultrasound or electrospray. The fine matrix droplets that land on the tissue allow analytes to extract from the tissue, and upon drying, incorporate into the matrix crystal. It is important here that the resulting matrix crystals should not exceed the dimensions of the desired pixel size, as it is impossible to determine the exact location of origin from an analyte within the confines of a single matrix crystal.[8] The use of solvents makes that the spray-based approach comes with the risk of slightly delocalizing the analyte molecules.

In sublimation systems, the solid MALDI matrix and sample are brought into a vacuum chamber. The matrix is heated under vacuum, causing it to sublimate. The sample, placed above the matrix, is cooled, causing the gas phase matrix to condensate onto the sample surface, allowing surface molecules to co-crystalize with the matrix. The sublimation

method has some advantages over the spray-based method; i) it is a solvent-free approach, limiting analyte delocalization, and ii) during the procedure the sample is cooled and stored in vacuum, limiting post-mortem degradation. Naturally, sublimation-based matrix application also has its limitations. The most obvious one is the limited extraction of analytes into the matrix which affects the measurement sensitivity of certain analytes.

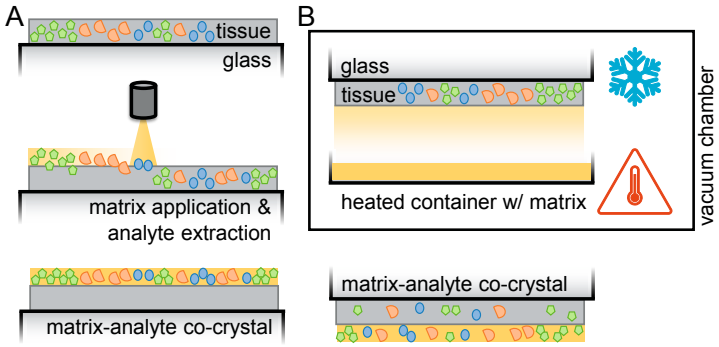


Figure 3. Spray-based matrix application of MALDI matrix is commonly performed using a pneumatic or ultrasound nebulizer mounted on a X-Y-Z-stage robot. A: The dissolved matrix is sprayed homogeneously onto the tissue section, allowing for maximum extraction of analytes, although at the risk of delocalization. B: Sublimation-based application of MALDI matrix is a solvent-free matrix application approach performed under vacuum conditions. Through heating the matrix, it sublimates into the vacuum chamber. Upon touching the cooled sample, the matrix condensates, and forms a fine crystal layer on top of the tissue. Although the cooling and vacuum conditions aid in minimizing delocalization and post-mortem degradation, molecular extraction using sublimation-based matrix application is compromised leading to lower sensitivity for some analytes.

2.1.7. MALDI-MSI measurement setup

Once the MALDI matrix has been applied, the tissue will be transferred to the MALDI-MS system to set up MSI the measurement. Depending on the instrument vendor, the slide with sections is usually mounted in a target carrier that is positioned in a XY robotic MALDI stage. The laser is focused on a fixed position, and between different pixels the target carrier with the sample is moved in the X and Y plane. Upon loading the target in the MS system, the first thing to do is to register the pre-MSI optical scan to the acquisition software of the MS system, essentially linking the MALDI stage XY motor positions to specific pixels in the pre-MSI optical scan image. This should be done as accurately as possible.[10] Once the registration of the pre-MSI optical image to the acquisition software is performed, the MS method can be optimized. This consists of five steps. i) Setting up the desired *m/z* range; for metabolites and lipids this is typically *m/z* 80-1500 Th. ii) Setting up the desired ablation field size and matching laser focus. Note that the ablation field size should not exceed the desired spatial resolution of the MSI analysis, as it leads to undesired oversampling. iii) Determining the optimal laser energy; too low of a laser energy results in insufficient ionization, and consequently produces poor spectra that will translate into "dead pixels". Too high laser energies result in extensive matrix cluster formation and analyte fragmentation, as well as an increase of the effective laser spot diameter which

might compromise spatial resolution. iv) Determine the optimal number of laser shots per pixel; based on the tradeoff between sensitivity (high number of shots per pixel) and throughput (low number of shots per pixel) one can select the optimal number of laser shots per pixel for the experiment. v) External mass calibration; using a known compound mixture calibrate the instrument mass response.

Once the MS method is optimized and calibrated, indicate the measurement areas on the tissues, and start the spatially correlated data acquisition.

2.1.8. MSI data pre-processing and feature extraction

After the MSI data acquisition the data needs pre-processing prior to data analysis. Typical steps in pre-processing are:

i) **Baseline subtraction**. Setting the noise level of the recorded single spectra to zero to ultimately enhance the signal-to-noise ratio (S/N). [11]
ii) **Normalization.** Multiplying mass spectra with an intensity-scaling factor to correct for and minimize the effect of systematic errors introduced during the MALDI-MSI analysis. [12]
iii) **Feature selection**, or peak picking. Defining true *m/z* features from noise in the representative (i.e. average, sum, or base peak) spectrum using a pre-defined S/N cutoff value. [13]
iv) **Feature extraction**. Extracting the intensity information for each *m/z* feature defined in the feature selection from each of the single pixel spectra. [14]

The result of the pre-processing is a workable peak matrix with the per-pixel intensity information for all selected *m/z* features. This pre-processed peak matrix is the starting point for the data analysis procedure described in the Exercises below.

2.2. Stable isotope tracing in (pre)clinical tissue specimens

Cell metabolism is a dynamic process characterized by parameters such as cellular metabolite levels, metabolic flux and nutrient contributions to different metabolic pathways. [15] Mass spectrometry-based metabolomics has a central role in measuring metabolite levels in both physiological and pathological conditions. Changes in metabolite levels indicate altered cellular metabolic states and are related to processes such as biosynthesis, energy metabolism and catabolism. However, metabolite levels per se do not directly reflect the metabolic rates, or fluxes, of the pathways, nor do they reflect the origin sources of the measured metabolites. Think of it as a bank account, both a rich and a poor person can have the same amount of money on the bank. Despite having the same balance, the rich person likely has a much higher in- and outflow of money and can thus afford a different lifestyle. Stable isotope tracing is a common tool to get insight in the fluxes of metabolism, as well as nutrient partitioning. Here, so-called "heavy" nutrients (e.g. $^{13}C_6$-glucose or $^{13}C_5$-, $^{15}N_2$-glutamine) are introduced into a biological system, and the incorporation and enrichment of the stable isotopes (^{13}C or ^{15}N) into downstream metabolic compounds is assessed. [16] Measuring both metabolite levels and metabolic fluxes at a single timepoint usually requires a metabolic steady or pseudo-steady state. This is characterized by constant or minimal changes in metabolite levels or metabolic fluxes during the time

course of the isotope tracing experiment. Sometimes steady state cannot be achieved in a natural biological system, which emphasizes the need for time course experiments, and dynamic labeling calculations, as well as non-stationary flux analysis, such as acute signaling events or nutrient modulations. [15] The duration of this time course may vary depending on the research question, e.g. glycolysis reaches metabolic steady state within approximately 10 minutes, whereas for the TCA cycle this often takes several hours.

Different stable isotope-labeled nutrients can be used to targeted different metabolic pathways. For example, uniformly labeled glucose (U-$^{13}C_6$-glucose, also noted as M+6 glucose) is the most commonly used nutrient to trace glycolysis (Figure 4A), the TCA cycle (Figure 4B), as well as other metabolic pathways related to glucose metabolism. However, to measure flux of the oxidative pentose phosphate pathway (PPP), 1,2-$^{13}C_2$-glucose is more common. [16] Uniformly labeled glutamine (U-$^{13}C_5$-glutamine; M+5 glutamine) is often used for TCA cycle flux estimation, as it results in highly abundant labeling of TCA cycle intermediates. Its conversion, through a-ketoglutarate via reductive carboxylation, results in the production of M+5 labeled citrate, which means U-$^{13}C_5$-glutamine can also be used to elucidate the contribution of glutamine to lipogenesis via the reductive carboxylation pathway – which is the reversed direction of the TCA cycle. Alternatively, when labeled glutamine enters the oxidative TCA cycle it will result in M+4 labeled succinate, and malate, as well as M+3 labeled a-ketoglutarate for the second cycle through the TCA cycle. [17] To measure the direct contribution of different nutrients to metabolic pathways, it is necessary to conduct tracer experiments with all circulating nutrients of interest, which can be determined by a straightforward matrix calculation. [18v]

Stable isotope tracing lends itself perfectly for *in vitro* studies, however *in vivo* experiments have been performed and applied in cancer patients to study tumor cell metabolism via either bolus injection or constant infusion. Nutrient partitioning has proven important for tumor cell survival and the function of immune cells in the tumor microenvironment. [19] Unfortunately, to study nutrient partitioning of tumor cells directly in patients multi-tracer analyses would be required, and these are not feasible using either both bolus injection or constant infusion. *Ex vivo* culturing of human tissue, following vibratome slicing, has provided a promising strategy that allows multi-tracer experiments using single tissue samples. [20] This approach can be combined with the parallelized introduction of various stable isotope labeled nutrients to the incubation medium, which allows for an efficient and biochemically meaningful labeling of metabolically active cells. [21] Given that the metabolic labeling takes place *in situ*, makes incubation of vibratome sectioned tissue slices with stable isotope labeled tracers perfectly compatible with spatial metabolomics tools such as MALDI-MSI. [22]

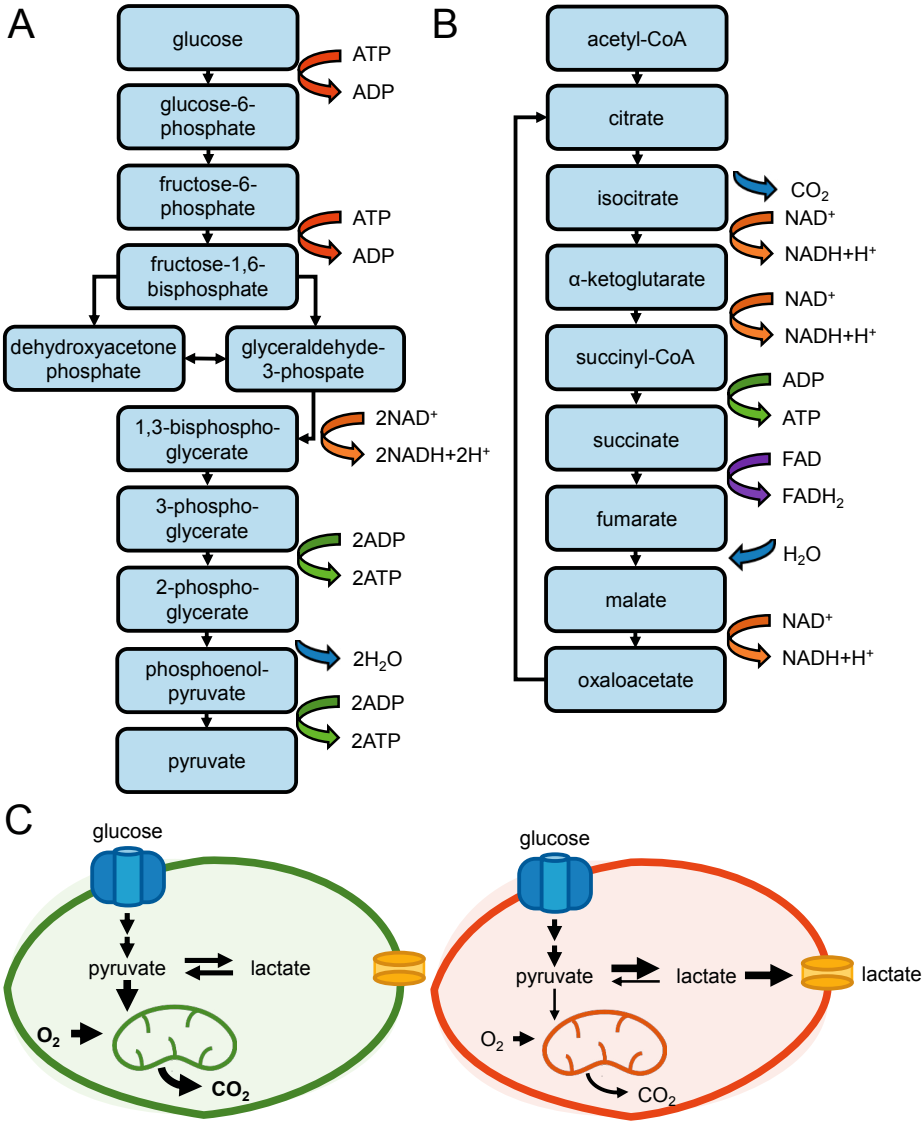


Figure 4. Metabolism and the Warburg effect. A: Schematic overview of glycolysis. B: Schematic overview of the tricarboxylic acid (TCA) cycle. C: Schematic representation of the Warburg effect.

2.3. The kidney and renal cell carcinoma

The human kidney is a highly complex organ, with up to twenty known cell types contributing to its main function of filtering our blood. Maintaining the molecular integrity of all these cell types is a complex process, requiring strict control of the transcriptome, proteome and metabolome. Disruption of these processes can emerge in a variety of diseases, ranging from chronic kidney disease to renal cancer. The most common type of kidney cancer in adults is renal cell carcinoma (RCC), a type of cancer which originates in the proximal tubule cells, which transport the primary urine after filtration of the blood. [23] The 5-year survival rate of RCC patients is around 50 – 70%, however when the cancer metastasizes the prognosis is substantially worse, with a median survival time of 13 months and a 5-year survival rate under 10%.[23] Giving its severe nature, RCC is heavily studied to better understand the disease pathogenesis and progression, as well as how effective treatment can be provided. One common finding amongst these studies is the role of metabolism; RCC cells display a grade-dependent metabolic reprogramming. [24] In this chapter, we will have a closer look into the metabolism of the healthy human kidney (Figure 5A) as well as RCC and its surrounding tissue (Figure 5B) using the *in situ* stable isotope tracing method and spatial metabolomics by MALDI-MSI described above.

2.3.1. Metabolism of the kidney and RCC

Different renal segments contribute to the role of the kidney as filter of our blood. The primary functional unit of the kidney, the nephron, consists of a glomerulus and Bowman's capsule, connected serially to a proximal tubule, loop of Henle and distal convoluted tubule. The various tubules play an important role in the reabsorption of water and salts from the filtrate originating from the glomerulus. These reabsorption processes are mostly mediated by active ion transport channels, making the kidney one of the most energy demanding organs of our body. This makes that the human kidney is highly metabolically active, with an estimated metabolic rate of >400 kcal/kg tissue/day. [25,26] To meet this energy demand, the kidney mostly uses the TCA cycle. To this end, the kidney is able to directly take up citrate, one of the TCA cycle intermediates, from the blood to fuel the TCA cycle. [27] Besides citrate, also lactate, uric acid and glutamine are reabsorbed in high levels to fuel the TCA cycle through side branches of the central carbon metabolism. Besides its metabolite burning character, the kidney also portrays significant gluconeogenetic capabilities. This results in the organ net-oxidizing lactate into pyruvate, thereby contributing to maintaining the circulating redox homeostasis.

RCC originates in the high energy demanding proximal tubule cells. As described above, the healthy kidney relies on the TCA cycle for its energy demand. However, upon the manifestation of RCC this drastically changes; the TCA cycle and subsequent oxidative phosphorylation (OxPhos) are downregulated, whereas anaerobic glycolysis and the PPP are activity are increased. The metabolic switch from OxPhos to anaerobic glycolysis a distinct feature of cancer cells, and is also known as the Warburg effect (Figure 4C). [28] Even though there is enough oxygen available for OxPhos, cancer cells preferentially use glycolysis for

energy production. To still provide sufficient ATP, the cell has to drastically increase its glycolytic flux since the ATP yield of anaerobic glycolysis (2 mol ATP/mol glucose) is much lower compared to OxPhos (~38 mol ATP/mol glucose). This results in a net increase of lactate production, which subsequently can be used for biomass incorporation and cell proliferation; highly beneficial for the fast-dividing cancer cell. Another phenomenon that can be attributed to the Warburg effect is the decreased glucose contribution to the TCA cycle. Of course, these two phenomena go hand in hand, and are both indicative of the metabolic shift resulting from the Warburg effect.

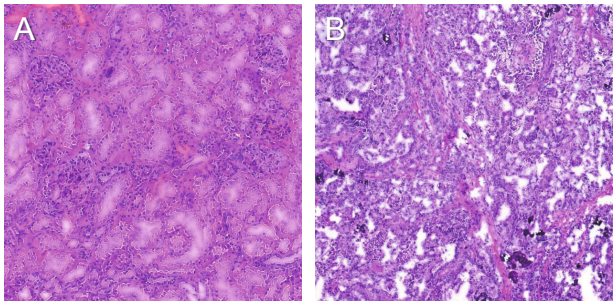


Figure 5. Kidney and renal cell carcinoma histology. A: Representative image of H&E stained normal human kidney. B: Representative image of H&E stained human RCC.

3. RESEARCH AIM

Establish an experiment that allows the metabolic differentiation between healthy proximal tubular cells and RCC cells, by visualizing dynamic differences in glucose metabolism within the tissue.

4. HYPOTHESIS AND EXPERIMENTAL SETUP

As the healthy human kidney relies predominantly on the TCA cycle for energy production and RCC relies on glycolysis, an *in situ* dynamic metabolic tracing experiment of glycolysis activity will allow the distinction between healthy proximal tubule cells and RCC tumor cells, and visualization of the Warburg effect.

4.1. Experimental setup

4.1.1. Tissue preparation, in situ isotope incubation, MALDI-MSI, and staining

A patient with RCC underwent surgical resection to remove the cancer. A biopsy from the RCC tissue was taken and preserved for metabolomics purposes. Besides the RCC tissue, the surrounding healthy tissue was sampled serving as control. Tissues were sliced using

a vibratome (Figure 6A). Since for this particular *ex vivo* experiment we are interested in visualizing the dynamics of the Warburg effect, U-$^{13}C_6$-glucose was used as metabolic tracer of the glycolysis. Tissue slices were incubated for 2 hours after which they were quenched using liquid nitrogen. Different tissue slices underwent 0 (control), 15, 30, 60, and 120 minutes of incubation with labeled glucose (Figure 6B). After snap freezing, the tissues were further prepared for MALDI-MSI analysis. First, 10 µm thick tissue sections were sectioned using a cryotome, and thaw-mounted on indium-tin-oxide (ITO)-coated glass slides. *N*-(1-napthylyl)-ethylenediamine dihydrochloride (NEDC) matrix was dissolved at 7 mg/mL in a mixture of solvents (70:25:5 methanol:acetonitrile:deionized water (% v/v/v)) and applied to the tissue section using a pneumatic sprayer. Then, negative ion-mode MALDI-TOF-MSI analysis of the sections was performed using a Bruker Daltonics rapifleX system at a 5 × 5 µm^2 spatial resolution. During these analyses, anions within a *m/z* range of 60-1000 Th were recorded. After MALDI-MSI data acquisition, the remaining MALDI matrix was removed from the MSI-analyzed tissue by various organic solvent washing steps. The remaining tissue was then stained with several immunofluorescence markers (LTL for proximal tubular cells, ECAD for distal tubular cells and collecting duct, and NPHS1 for podocytes), which allowed us to identify the different epithelial cell types in the tissue.

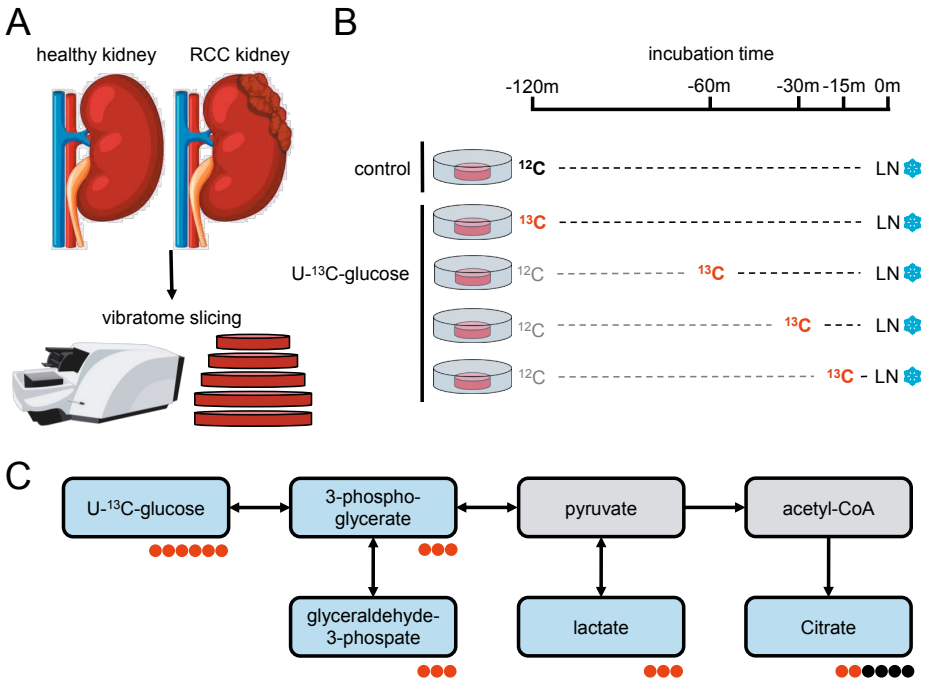


Figure 6. Stable isotope tracing in tissue culture. A: Schematic overview of the vibratome slicing procedure. B: Overview of the stable isotope tracing tissue culture time course experiment. At the indicated time points the label with regular glucose was exchanged with medium containing U-^{13}C-glucose. At time point 0m the tissues were quenched using liquid nitrogen. C: Detected metabolites with the number of incorporated stable isotope labelled carbon atoms in red.

4.2. MSI data pre-processing

Following the MALDI-MSI data acquisition, the data for every incubation time point was loaded into a proprietary software package provided by the instrument vendor (SCiLS Lab PRO, v2023a) with baseline correction using a convolution algorithm. The dataset was normalized to the total ion count (TIC). Spectral recalibration and a two-step peak picking on the average spectrum were performed in mMass[29]; i) untargeted peak picking (S/N > 3) was performed on the *m/z* range between 450-1000 Th), and ii) targeted *m/z* feature selection was performed on specific metabolites and isotopologues expected to derive from the stable isotope tracing experiment and based on the theoretical *m/z* values (Figure 6C). The peak list was imported into SCiLS Lab, which was used for per-pixel feature extraction and data exporting.

Table 1. Overview with datasets containing the per-pixel intensity information for all selected *m/z* features.

Timepoint	**Dataset**
Control	Kidney_RCC_lipids.csv
15 min	Kidney_RCC_13C_15min.csv
30 min	Kidney_RCC_13C_30min.csv
60 min	Kidney_RCC_13C_60min.csv
120 min	Kidney_RCC_13C_120min.csv
Additional files	
Pixel ID and coordinates	Kidney_RCC_coordinate.csv

Table 2. Files needed for the isotope correction package IsoCorrectoR.

File	**Contains**
ElementFile	Information on the elements important for the isotope correction process
MoleculeFile	Information on the molecules to be corrected for natural isotope abundance/tracer purity
MeasurementFile	The measured data that needs to be corrected

Table 3. Overview of datasets containing the isotope corrected values for enrichment visualization.

Timepoint	**Dataset**
15 min	CorrectedFractions15.csv
30 min	CorrectedFractions30.csv
60 min	CorrectedFractions60.csv
120 min	CorrectedFractions120.csv

4.3. Necessary software and exemplary dataset

Since the SCiLS Lab software is not freely available, we provide the pre-processed datasets (Table 1). The associated datasets, IsoCorrectoR file templates and the R scripts (in Rmarkdown) which are referred to throughout the chapter, are available for downloading from OSF.io.

The programming language R was used for most data analysis steps, e.g. data transformation, spatial segmentation, data integration, metabolite intensity imputation and visualization. A list of required R packages is provided below (Table 4).

Table 4. Overview of R packages required for the data analysis strategies described below.

Package	Purpose
ggplot2	Visualization
ggrepel	Visualization; repel overlaying labels in ggplot2
ggcorrplot	Visualization; easy correlation matrix
pheatmap	Visualization; drawing clustered heatmaps
patchwork	Visualization; combining plots
viridis	Visualization; colors suited for black-white and color
IsoCorrectoR	Isotope abundance correction
Seurat	Clustering, data integration, metabolite imputation
dplyr	Data manipulation
Tidyverse	Data manipulation
reshape2	Data manipulation

4.4. Research questions and exercises

Using your newly acquired knowledge from the introduction, as well as the provided code and the example datasets, you can train yourself to perform the data transformations and steps to perform *in situ* metabolic dynamics analysis. In this chapter we aim to answer the following central research questions:

1. Can we differentiate RCC from healthy kidney tissue on the basis of their metabolic histology? In other words, can we use unsupervised multivariate statistical approaches to isolate pixels obtained from a cancerous tissue from those obtained from a healthy kidney tissue?
2. Can we differentiate RCC from healthy kidney by visualizing the Warburg effect using *in situ* dynamic metabolomics? In other words, can we find differences in the contribution of U-$^{13}C_6$-glucose to glycolysis and TCA cycle between cancerous tissue and healthy kidney?

5. EXERCISES

Throughout the remainder of the chapter, you will be guided through the workflow outlined in Figure 7. This will be a good starting point for any *in situ* dynamic metabolomics study. Obviously, the options to expand on this analysis pipeline are endless and will not be within the scope of this chapter.

5.1. Dimensionality reduction and spatial segmentation of MSI data

The analysis of a tissue section by MSI generates very high dimensional data. At a 5×5 μm^2 spatial resolution, a measurement area of 1 mm^2 contains 40,000 pixels. For each pixel, a mass spectrum is recorded each consisting of typically 250.000 datapoints, resulting in a total of 10×10^9 datapoints per square millimeter analyzed. In order to create interpretable data, one needs to reduce the data complexity. Following complexity reduction that is achieved through peak picking and feature selection (section 2.1.8), the next step in dimensionality reduction is achieved through spatial segmentation of the data. Here, multivariate statistical tools (i.e. principal component analysis (PCA), t-distributed stochastic neighbor embedding (tSNE), uniform manifold approximation and projection (UMAP) etc.) are used to calculate groups of pixels that are highly similar, which in this case means they have mass spectra with comparable peak intensity profiles. By color coding the clusters, and plotting the clusters using the pixel XY coordinates, one can construct an image that shows the spatial distribution of metabolically similar pixels. [30]

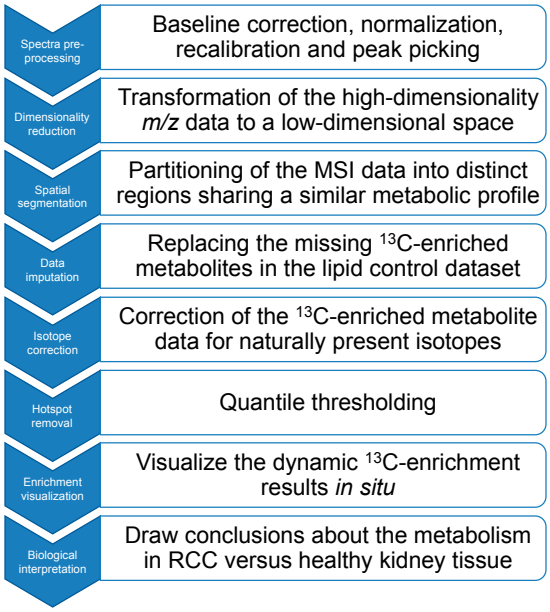


Figure 7. Schematic of the data analysis workflow used for the in situ spatial metabolomics analysis.

During the first exercise in our dynamic metabolism data analysis pipeline, you will perform a metabolome-driven spatial segmentation of the healthy kidney control and RCC tissue to identify groups of pixels with similar metabolic profiles. The clustering algorithm chosen for the dimensionality reduction is UMAP. [31] In the final UMAP plot, pixels that have a similar metabolic profile will end up in close together and consequently will be assigned to the same cluster. Subsequently you will reconstruct the cluster image, resulting in a chemically

segmented visualization of the tissue. This image will be referred to as the "*metabolic histology image*" of the tissue.

5.1.1. R code and explanations

The first goal of this exercise is to perform the UMAP-based dimensionality reduction, you can use the R Script "**Reduction_Segmentation**" with corresponding .csv files "Kidney_RCC_lipids.csv" and "Kidney_RCC_coordinate.csv" for this.

We start by loading the exemplary lipid MSI dataset into our R working environment. Since the lipid profile is highly cell-type specific and stable throughout the isotope-labelling experiment, we can use these features for spatial segmentation, cell-type identification, and anchor-based data integration.[22] The file we have provided you with is the combined data of both the healthy kidney and the RCC tissue. Throughout the exercise it will appear that these two tissues are metabolically indeed very distinct from one another.

```
# Load in the lipidome dataset
 MSIref <- read.csv(file = 'Kidney_RCC_lipids.csv', row.names = 1, header   = TRUE, sep = "
,")

# Transform the countmatrix into dataframe suitable for Seurat
 MSIref <- MSIref * 100 %>%
     round(digits = 0)
 MSIref <-  as.data.frame(t(MSIref))
```

After data transformation, the data is now in a suitable format to load it into the *Seurat* package using the following code:

```
MSIdata <- CreateSeuratObject(counts = MSIref, project = "RCC")
```

To put the dataset to a common scale, without distorting the relative differences in ranges of intensity values, the data needs to be normalized and scaled. After data transformation, a PCA will be performed which determines the neighbors of each pixel. The results will later be used by the UMAP algorithm.

```
MSIdata <- SCTransform(MSIdata, verbose = F)
MSIdata <- RunPCA(MSIdata, assay = "SCT", verbose = FALSE)

# Perform Elbowplot to assess suitable number of PCs for subsequent
analysis
ElbowPlot(MSIdata)
```

The elbow plot (Figure 8) shows the standard deviation for each of the calculated principal components (PCs), and is a useful tool to determine how many PCs should be selected to represent the majority of the variation held within the dataset. Based on visual inspection of the plot, you determine the PC at which the change in standard deviation starts to taper off. Despite being subjective, it is a quick and efficient method for choosing the number of PCs to use.

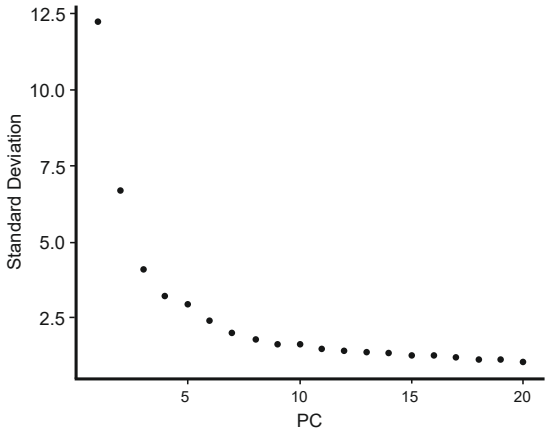


Figure 8. Elbow plot displaying the standard deviation for each calculated principal component.

Based on the elbow plot above, we chose to work with the first eight PCs to run the UMAP algorithm. There are two important, so-called, hyper parameters that have a significant effect on the results: `resolution` and `min.dist`. You can play around yourself with these parameters to see how they affect the resulting UMAP embedding.

```
# Look for pixels that overlap in the PCA space
MSIdata <- FindNeighbors(MSIdata, dims = 1:8)
# Iteratively groups cells together to a certain optimal point
MSIdata <- FindClusters(MSIdata, resolution = 0.5)

# The RunUMAP function learns the underlying manifold of the data in order to place similar
cells together in a low-dimensional space
MSIdata <- RunUMAP(object = MSIdata, dims = 1:8, n.neighbors = 15L,
min.dist = 0.05, check_duplicates = FALSE)

DimPlot(object = MSIdata, reduction = 'umap', label = TRUE, pt.size = 1,
label.size = 5)

DimPlot(object = MSIdata, reduction = 'umap', label = TRUE, pt.size = 1,
label.size = 5, group.by = "orig.ident")
```

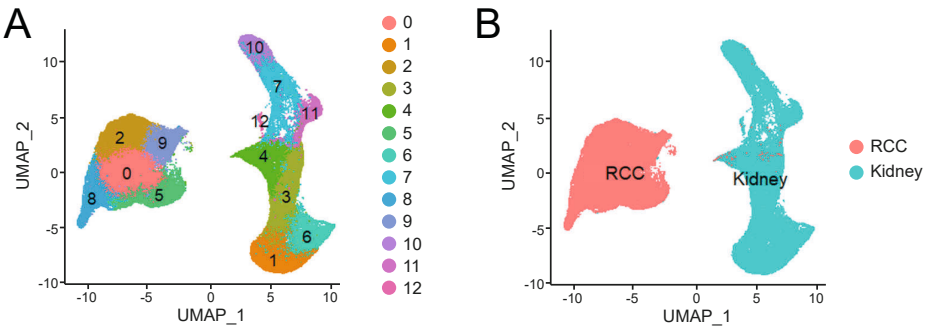


Figure 9. Two-dimensional scatterplot visualization of the UMAP embedding. A: UMAP representation with color-coding based on cluster identities. B: UMAP representation with color-coding based on original pixel identities.

After the dimensionality reduction, the clusters are named using meaningless integers which do not give any information about the biological meaning of these clusters (Figure 9A). Since in our experimental setup we expect to find a mixture of cells, including a variety of healthy kidney cells as well as cancer cells, we are interested in assigning the different cell types to the different clusters. To achieve this, we need to compare the spatial representation of the dimensionality reduction – the "metabolic histology" – with the celltype information obtained from the immunofluorescence microscopy images. This allows us to determine which clusters are positive for which cell-type markers.

The second part of the exercise is to reconstruct the segmentation cluster distributions and generate the metabolic histology image. To recreate the cluster images, we first need to load the XY coordinates for each of the analyzed MSI pixels in R.

```
# Create a data frame with the xy coordinates from the imaging run
xycoord <- read.csv(file = 'Kidney_RCC_coordinate.csv', row.names = 1,
                    header = TRUE, sep = ",")

xycoord$y1 <- xycoord$y * -1
```

The next step is to associate the XY coordinates to the clustered pixels in the Seurat object and visualize the individual cluster images. An example for cluster 1 is shown below (Figure 10).

```
# Extract the cluster information from the dimensionality reduction
cluster <- as.data.frame(as.matrix(MSIdata@active.ident))

# Select which cluster you want to visualize, by setting this to 1 and all others to 0
cluster_int <- "1"
cluster$V1 <- replace(cluster$V1, cluster$V1 != cluster_int, 0)
cluster$V1 <- replace(cluster$V1, cluster$V1 == cluster_int, 1)

# Merge the cluster information with the xy coordinate system
dataframe <- merge(cluster, xycoord, by = 'row.names')

# Transform data for pheatmap
dataframe$V1 <- as.numeric(dataframe$V1)
```

Figure 10. Visualization of the spatial distribution of cluster 1. The left panel is the healthy kidney tissue, the right panel represents the RCC tissue.

```
a = dcast(dataframe, y~x, value.var = "V1")
row.names(a) <- a[,1]
data <- as.matrix(subset(a, select = -c(y) ))

# Visualize the spatial distribution of cluster of interest
pheatmap(data, scale = "none", cellwidth = 0.6, cellheight = 0.6,
cluster_rows = FALSE, cluster_cols = FALSE, legend = T, show_rownames = F,show_colnames = F
, border_color = FALSE, fontsize = 10,
color = viridis(250), na_col = "WHITE", breaks = NA, main = cluster_int)
```

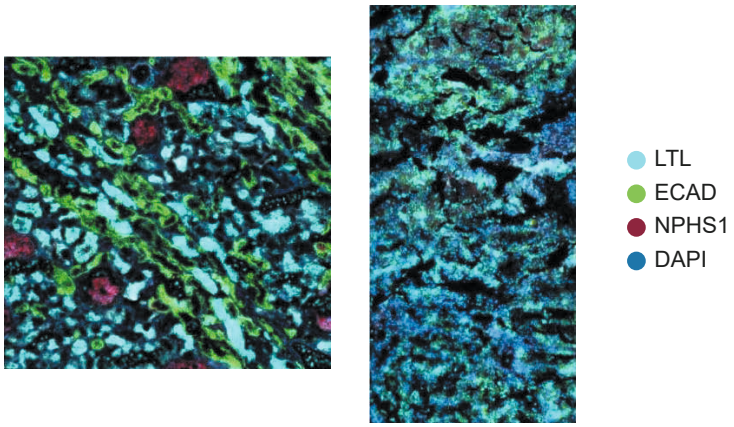


Figure 11. Immunofluorescence staining. The left panel represents the healthy kidney tissue, the right panel represents the RCC tissue – LTL (turquoise) for proximal tubular cells, ECAD (green) for distal tubular cells and collecting duct, NPHS1 (red) for podocytes, and DAPI (blue) for cell nuclei.

To assign cell type information to the clusters we compare the visualizations of each of the clusters with the IF images we have of the post-MSI analyzed tissue (Figure 11). From these IF images it becomes apparent immediately that major histological transformations have occurred in the RCC sample compared to the healthy kidney tissue. The glomerular and tubular structures have mostly disappeared, leaving a dedifferentiated and unstructured tissue which mainly consists of cancer cells and stroma tissue. Although the RCC tissue is quite heterogeneous, both on the histological level evidenced by the IF staining, as well as the lipidomic level evidenced by the presence of several RCC clusters in the UMAP embedding, it is beyond the scope of this chapter to go into the details of intratumor heterogeneity. Therefore, in further processing steps the RCC tissue as a whole will be regarded as a single group.

For this chapter, we have provided you with a vector with the cell types of interest (`new.cluster.ids`) which you can use to assign the cluster identities.

```
# Change the cluster names to the newly identified cell type names
new.cluster.ids <- c("RCC", "LTL_1", "RCC", "Unidentified_1","Glom/Vessel"
, "RCC", "LTL_2", "Unidentified_2", "RCC", "RCC",
                     "ECAD_1", "ECAD_2", "Unidentified_3")
names(new.cluster.ids) <- levels(MSIdata)
MSIdata <- RenameIdents(MSIdata, new.cluster.ids)

DimPlot(object = MSIdata, reduction = "umap", label = TRUE, pt.size = 1,
          label.size = 5)
```

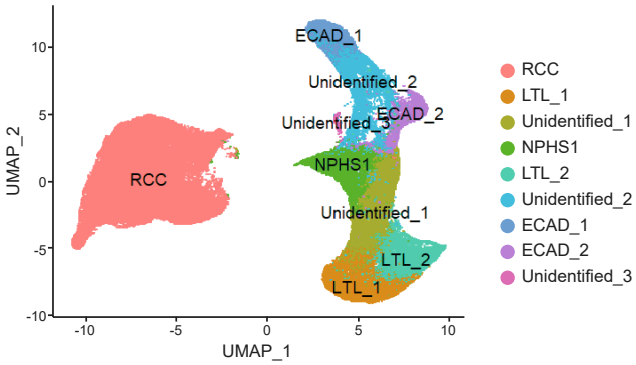


Figure 12. Two-dimensional scatterplot visualization of the UMAP embedding with annotated cluster identities.

As final part of this exercise, you can now visualize the metabolic histology of the tissue. Once again, a data transformation step is required to enable the spatial visualization of the metabolic histology using `ggplot`.

```
# Data extraction out of the Seurat object
embeddings <- as.data.frame(MSIdata@reductions[["umap"]]@cell.embeddings)
ident <- as.data.frame(MSIdata@active.ident)
# Data transformation
vector <- row.names(xycoord)
```

```
xycoord$pixID <- vector

vector <- row.names(embeddings)
embeddings$pixID <- vector

vector <- row.names(ident)
ident$pixID <- vector
names(ident)[1] <- "Ident"

# Merging everything into 1 dataset
spat_UMAP_Kidney <- merge(xycoord, embeddings, by = 'pixID') %>%
  merge(ident, by = 'pixID')


# Using ggplot to visualize the metabolic histology
ggplot(spat_UMAP_Kidney, aes(x = x, y = y))+
  geom_tile(aes(fill = Ident))+
  coord_fixed()+
  theme_void()
```

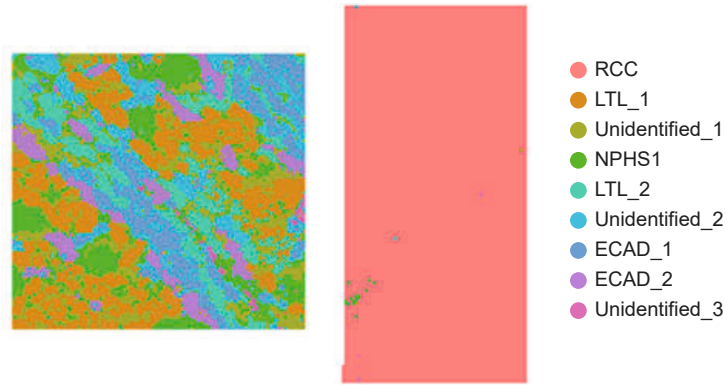


Figure 13. Metabolic histology visualization of the UMAP embeddings. The color-coding represents the different cluster identities in the spatial context of the tissue.

5.1.2. Description of the results

Based on the selected parameters for the dimensionality reduction using the UMAP algorithm, we find a coarse split of the data into two large clusters, each subdivided in several smaller subclusters (Figure 9A). Based on the visualization of the origin of the pixel (`orig.ident`), it appears that the largest contributor to variation within the tissue is whether the pixel comes from a healthy or a RCC kidney (Figure 9B). The sub clustering of the healthy kidney, and comparison of the cluster distributions to the IF images, additionally shows that the individual cell types within the healthy kidney each have their own metabolic profile. For example, we were able to assign various clusters to cells which were positive for LTL, indicating various proximal tubular cells in the healthy kidney tissue (Figure 11, 12 & 13).

5.2. Metabolite abundance imputation

To perform the dynamic metabolic measurements, we have performed a time course of *in situ* isotope tracing experiments with U-^{13}C-glucose in both the RCC tissue and its surrounding healthy kidney tissue. Since the different timepoints are represented by different tissue slices taken from the same tissue following MALDI measurements were performed on physically different pieces of tissue. This effectively means we now have five metabolic snapshots of the healthy kidney and the RCC tissue. The control tissues (timepoint 0 min) will not contain any ^{13}C-enriched isotopologues, while the other timepoints (15, 30, 60 and 120 min) will contain different levels of the ^{13}C-enriched downstream metabolic intermediates. To efficiently and properly evaluate the dynamics of the cellular metabolism in the different cell types, we need to impute the ^{13}C-enriched metabolic snapshots in our control dataset. This allows the direct comparison of the different timepoints using the same tissue, and consequently the direct comparison of the metabolic dynamics of specific cell types. For this data imputation, we need to get the intensities of the ^{13}C-enriched metabolite isotopologues into the control dataset which was not incubated with U-^{13}C-glucose. To achieve this, we make use of a data enhancing strategy. This strategy takes a query dataset (lipid control dataset) which lacks intensity information on features of interest

(the ^{13}C-enriched metabolite isotopologues and downstream metabolites), and a reference dataset, that does contain the intensity information of the features of interest (the ^{13}C-datasets of different timepoints). Based on the intensity profiles of so-called common features, shared between the query and reference datasets (the lipid *m/z* features), a k-nearest neighbor (kNN) analysis can determine the most similar pixels between the query and reference datasets. It then takes the pixel-specific intensity of the features of interest from the reference dataset and imputes them in the most similar pixel in the query dataset. By doing this for each time point, we end up with five datasets in which we can calculate the ^{13}C-enrichment in various cell types for different metabolic intermediates. The differences in metabolic flux of the different cell types will influence the dynamics of ^{13}C-label incorporation, which we can now visualize over time, and in context of the morphology of the analyzed tissues.

The exercises below will take you through the process of imputing the different ^{13}C-labeling timepoints to the control dataset. Following data imputation, an isotope correction step is performed. This is necessary since isotopes are not only introduced with the labeling experiment, but these isotopes are also naturally abundant. This natural abundance of isotopes leads to convoluted signals in the MSI dataset, which could lead to distorted biological findings. To correct for the natural abundance of isotopes, we therefore perform an isotope correction step using the IsoCorrectoR package. Since it takes a significant amount of time to run this package (for this dataset the isotope correction took over 48 hours) and since it requires manual data transformation in Excel, we will provide the code for this step of the process as well as the resulting data frame with isotope corrected data required for the next exercise so you are not required to perform this step yourself.

5.2.1. R code and explanations

The goal of this exercise is to perform data imputation, for which you can use the R Script "**Data_imputation**" with corresponding .csv files "Kidney_RCC_13C_15min.csv" and "Kidney_RCC_lipids.csv" for this. In the chapter, only the example for the 15-minute timepoint will be shown, you can perform the other timepoints in a similar way yourself.

Start by loading in the ^{13}C-enriched processed MALDI-MSI dataset, which will be used to impute ^{13}C isotopologue data to the lipid dataset.

```
# Load reference dataset
MSIref <- read.csv(file = 'Kidney_RCC_13C_15min.csv', row.names = 1,
header = TRUE, sep = ",")

# Reshape data into suitable format for subsequent analysis
MSIref <- MSIref * 100 %>%
  round(digits = 0)
MSIref <- as.data.frame(t(MSIref))

# Create Seurat objects
MSIdata_ref <- CreateSeuratObject(counts = MSIref, project = "RCC")
# Data normalization - ref
MSIdata_ref <- SCTransform(MSIdata_ref, verbose = FALSE)
MSIdata_ref <- RunPCA(MSIdata_ref, assay = "SCT", verbose = FALSE)
ElbowPlot(MSIdata_ref)
```

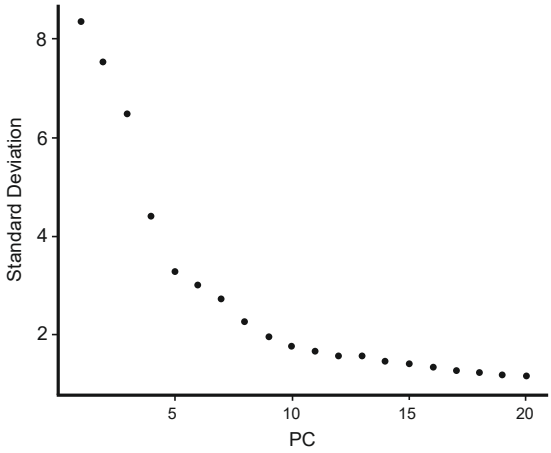


Figure 14. Elbow plot displaying the standard deviation for each calculated principal component.

Based on the elbow plot (Figure 14) the first eight PCs were used to run the UMAP algorithm.

```
# UMAP analysis - ref
MSIdata_ref <- RunUMAP(MSIdata_ref, dims = 1:8)
```

Then, the lipid control dataset can be loaded into the R environment and processed in a similar way as the ^{13}C isotoplogue data.

```
# Load query dataset and reshape accordingly
MSIque <- read.csv(file = 'Kidney_RCC_lipids.csv', row.names = 1,
           header = TRUE, sep=",")
MSIque <- MSIque * 100 %>%
  round(digits = 0)
MSIque <- as.data.frame(t(MSIque))

# Create Seurat object
MSIdata_que <- CreateSeuratObject(counts = MSIque, project = "RCC")

# Data normalization - que
MSIdata_que <- SCTransform(MSIdata_que, verbose = FALSE)
MSIdata_que <- RunPCA(MSIdata_que, assay = "SCT", verbose = FALSE)
```

After processing of both the reference and query datasets, we can continue with the data imputation.

```
# Find the common features between reference and query dataset
anchors <- FindTransferAnchors(reference = MSIdata_ref, query =
              MSIdata_que, normalization.method = "SCT")


# Filling in the labelling data in the control dataset based on KNN
predictions.assay <- TransferData(anchorset = anchors, refdata =
      GetAssayData(MSIdata_ref[['RNA']]), prediction.assay = T,
      weight.reduction = MSIdata_que[["pca"]], dims = 1:8)


# Write out csv files for furhter processing in Excel
data_to_write_out <- as.data.frame(as.matrix(predictions.assay@data))
data_to_write_out <- as.data.frame(t(data_to_write_out))

# Select only the 13C enriched metabolites for visualization
data_to_write_out <- data_to_write_out[ ,1:15]
write.csv(x = data_to_write_out, row.names = T, file = "15min.csv")
```

The resulting .csv file has the following format:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1		X89.05	X92.04	X146.07	X147.06	X148.07	X149.09	X185.01	X188.02	X191.04	X192.04	X193.05	X215.07	X221.06	X265	X268.01
2	Kidney_1	164.2049	78.60585	3170.251	493.208	349.3059	90.98812	999.3672	688.9032	2601.084	398.3326	614.0775	3958.361	3196.907	517.793	571.1652
3	Kidney_2	131.0715	65.20852	3102.507	518.2193	330.5496	154.996	824.6879	745.2096	2528.768	579.7447	775.9839	3999.402	3337.128	566.1404	670.5507
4	Kidney_3	208.5753	189.01	6768.506	809	428.2251	258.9036	1142.644	1040.014	3750.223	689.4207	946.8252	8188.514	5513.889	648.4032	490.0168
5	Kidney_4	143.8233	97.82871	3847.535	561.4539	272.4467	199.8724	979.172	1048.885	3133.234	635.8555	839.0305	5057.1	3818.871	743.6843	598.5983
6	Kidney_5	269.3624	127.382	5101.758	956.4896	567.3086	178.4838	1153.071	1115.954	4130.564	801.2893	1136.845	6062.964	5591.37	697.8259	628.7008
7	Kidney_6	149.8921	133.5138	3759.849	565.6191	310.5893	155.7036	805.2148	872.2529	2883.201	542.5479	750.4825	4174.72	3880.181	667.6901	813.5342
8	Kidney_7	245.1582	111.4794	4346.197	490.9912	333.5009	140.0484	921.652	1091.595	3160.725	545.0753	875.3529	5181.107	4460.052	410.5579	649.548
9	Kidney_8	241.6892	92.76664	4327.685	572.2231	330.8739	201.0429	1006.773	1125.495	3187.372	708.5589	815.7928	4865.497	4251.21	662.7086	557.124
10	Kidney_9	176.681	82.74212	4045.127	524.0093	392.7269	128.7123	1062.987	914.3264	3304.156	505.4223	764.0083	5159.441	4274.472	813.3812	608.3343
11	Kidney_10	192.2751	86.38057	4397.66	529.3585	465.6634	287.0805	964.8196	930.5935	3581.457	710.9646	814.2171	5439.056	4445.235	810.1312	901.0591

Each row represents a pixel from the lipid control tissue and each of the columns contains imputed intensities for the ^{13}C isotopologues represented here by their *m/z* value. In this experiment we only introduced U-$^{13}C_6$-glucose, therefore the number of ^{13}C isotoplogues from the downstream metabolites was limited to the ones represented in the example format.

The next step of the procedure is the isotope correction using the IsoCorrectoR package.[32] The package requires the input of three files which are necessary for proper isotope correction: the ElementFile, MeasurementFile and MoleculeFile. These three files all have a set layout that should be used when importing your own data. For more information about the package and these files, see the reference listed in **Further reading**. A short reminder of the fact that running the isotope correction for these datasets took over 48 hours. The output files of the isotope correction have been provided for a smooth continuation of the workflow.

```
# Load query dataset and reshape accordingly
MeasurementFile <- read.csv(file = '15min.csv', row.names = 1, sep = ";")
a = MeasurementFile
a = data.frame(t(a))
b = data.frame(row.names(a))
# Rename file to “Measurements/Samples”
fix(b)
```

```
c = cbind(b,a)
write.csv(c, quote = F, row.names = F, file = "File_path.csv")

# Get path of IsoCorrectoR files
path.molecule <- system.file("data", "MoleculeFile.csv",
                             package = "IsoCorrectoR", mustWork = TRUE);
path.element <- system.file("data", "ElementFile.csv",
                             package = "IsoCorrectoR", mustWork = TRUE);
path.measurement <- system.file("data", "MeasurementFile15.csv",
                             package = "IsoCorrectoR", mustWork = TRUE);

# Run correction algorithm and save results in new variable
correctionResults <- IsoCorrection(MeasurementFile = path.measurement,
                                   ElementFile = path.element,
                                   MoleculeFile = path.molecule)
```

For the enrichment visualization, we use the “IsoCorrectoR_result_CorrectedFractions” file. This file contains the corrected measurement data as fractions of the total abundance of a specific metabolite. The output format is shown in the example below:

	A	B	C	D	E	F	G	H	I	J	K
1		Kidney_1	Kidney_2	Kidney_3	Kidney_4	Kidney_5	Kidney_6	Kidney_7	Kidney_8	Kidney_9	Kidney_10
2	Lactate_0	0.683064801	0.67466836	0.532385099	0.602666919	0.685699512	0.536667229	0.694084977	0.728848771	0.687797054	0.696648258
3	Lactate_1	0	0	0	0	0	0	0	0	0	0
4	Lactate_2	0	0	0	0	0	0	0	0	0	0
5	Lactate_3	0.316935199	0.32533164	0.467614901	0.397333081	0.314300488	0.463332771	0.305915023	0.271151229	0.312202946	0.303351742
6	Glutamate_0	0.814854513	0.796687867	0.863668963	0.831067397	0.790829455	0.827455296	0.86313783	0.840111403	0.838113844	0.816200256
7	Glutamate_1	0.081818866	0.089026013	0.055918456	0.075514291	0.104370744	0.078878797	0.050287717	0.064948208	0.062569222	0.053530309
8	Glutamate_2	0.083436971	0.078352393	0.050097036	0.053412666	0.080695045	0.062575415	0.061681887	0.05912233	0.076000162	0.081361744
9	Glutamate_3	0.01988965	0.035933726	0.030315545	0.040005645	0.024104756	0.031090492	0.024892566	0.03581806	0.023316772	0.048907691
10	Glutamate_4	0	0	0	0	0	0	0	0	0	0
11	Glutamate_5	0	0	0	0	0	0	0	0	0	0
12	X3PG_0	0.599466781	0.533091921	0.531292374	0.4906127	0.515979805	0.487815273	0.465553126	0.479947787	0.545343138	0.51682877
13	X3PG_1	0	0	0	0	0	0	0	0	0	0
14	X3PG_2	0	0	0	0	0	0	0	0	0	0
15	X3PG_3	0.400533219	0.466908079	0.468707626	0.5093873	0.484020195	0.512184727	0.534446874	0.520052213	0.454656862	0.48317123
16	Citrate_0	0.755579175	0.682700839	0.730686784	0.713450956	0.714060671	0.724461814	0.72373226	0.71002588	0.758396476	0.736427914
17	Citrate_1	0.074052538	0.1183213	0.09380059	0.10507245	0.098840788	0.096112782	0.084758849	0.118168885	0.074196989	0.105230839
18	Citrate_2	0.170368287	0.198977861	0.175512626	0.181476595	0.187098541	0.179425404	0.191508891	0.171805235	0.167406535	0.158341248
19	Citrate_3	0	0	0	0	0	0	0	0	0	0
20	Citrate_4	0	0	0	0	0	0	0	0	0	0
21	Citrate_5	0	0	0	0	0	0	0	0	0	0
22	Glucose_0	0.56807141	0.560052411	0.61201798	0.584476828	0.535270067	0.533326963	0.552357064	0.548671807	0.561809231	0.565155709
23	Glucose_1	0	0	0	0	0	0	0	0	0	0
24	Glucose_2	0	0	0	0	0	0	0	0	0	0
25	Glucose_3	0	0	0	0	0	0	0	0	0	0
26	Glucose_4	0	0	0	0	0	0	0	0	0	0
27	Glucose_5	0	0	0	0	0	0	0	0	0	0
28	Glucose_6	0.43192859	0.439947589	0.38798202	0.415523172	0.464729933	0.466673037	0.447642936	0.451328193	0.438190769	0.434844291
29	X13BPG_0	0.483285929	0.465545798	0.577201429	0.561745202	0.533832269	0.458510079	0.394714295	0.551014282	0.579737899	0.481220583
30	X13BPG_1	0	0	0	0	0	0	0	0	0	0
31	X13BPG_2	0	0	0	0	0	0	0	0	0	0
32	X13BPG_3	0.516714071	0.534454202	0.422798571	0.438254798	0.466167731	0.541489921	0.605285705	0.448985718	0.420262101	0.518779417

5.3. Metabolic dynamics calculations and visualization

As the title of this chapter states, the aim is to perform *in situ* analysis of the dynamic metabolism. The datasets resulting from the exercise in section 5.2 contain information on the ^{13}C-enrichment for each of the pixels in the lipid control dataset. This means that similar to what was done before for constructing the metabolic histology, it is now possible to create pseudo-images containing ^{13}C-enrichment of isotopologues of various metabolites over the 120 min time course experiment. A first step in creating these pseudo-images, is to perform hotspot removal on the images. This is merely an image processing step where for each image the intensity hotspots are removed, which results in higher contrast images. The hotspot removal is based on a quantile thresholding approach in which for each image, the intensities of the pixels in the highest quantile (1%) are set to the 99th quantile. The data after hotspot removal can be used to reconstruct the tissue morphology using the XY coordinates, and, since the coordinates were previously also linked to the UMAP clusters, data can be extracted to directly compare various cell types with one another. The exercises below will allow you to perform enrichment visualization after hotspot removal. Note that the graphs for the ^{13}C-enrichment over time were not prepared using R, and as such there will be no code for these figures. The results with corresponding biological interpretation will however be discussed later in this section.

5.3.1. R code and explanations

This final exercise is focused on the visualization of the ^{13}C-enrichment results. These visualizations show the metabolic dynamics in the context of the tissue. To continue with this exercise, you can use the provided output .csv files from the IsotopeCorrectoR package "CorrectedFractions15.csv", "CorrectedFractions30.csv", "CorrectedFractions60.csv" and "CorrectedFractions120.csv" which contain the isotope corrected data as described in section 5.2. The script you can use for the exercise is "**Hotspot_Enrichment**". This exercise will focus on the 15 minutes timepoint, but the same R code can be used for the other timepoints.

```
# Loading the results from IsoCorrectoR
df <- read.csv("CorrectedFractions15.csv", row.names = 1)
```

Data transformation is necessary to be able to perform visualization of ^{13}C-enrichment. This is similar to what you have seen in a previous exercise.

```
# Hotspot removal, automatically working through all columns
for(i in 1:ncol(df)){
  q <- quantile(df[,i], c(.25, .50, .75, .90, .99))
  max <- q[5]
  df[,i][which(df[,i]>max)] <- max
}
```

After hotspot removal, we can bring in the XY coordinates again to make a spatial representation of the metabolite-specific ^{13}C enrichment onto our tissues of interest. We

first have to combine the data frame containing the ^{13}C enrichment information with the XY coordinates.

```
# Combine the processed isotope corrected data with XY coordinates
xy <- read.csv(file = 'Kidney_RCC_coordinate.csv', row.names = 1,
          header = T, sep = ",")
merged_df <- merge(xy, df, by = 0)
rownames(merged_df) <- merged_df$Row.names
merged_df <- merged_df[,-1 ]
```

Since we know which column contains which metabolite, we can generate heatmaps for all our metabolites of interest.

```
# Extract the metabolites of interest
lactate3   <- merged_df[,c(1,2,6)]
glutamate2 <- merged_df[,c(1,2,9)]
x3pg3      <- merged_df[,c(1,2,16)]
citrate2   <- merged_df[,c(1,2,19)]

# Make a list with all metabolites of interest
metabolites <- list(lactate3, glutamate2, x3pg3, citrate2)

# Write a pdf file in a desired file path containing all enrichment
  visualizations
pdf("File_path.pdf")
for (n in metabolites) {
  a = dcast(n, y~x)
  data <- as.matrix(a)
  data <- data[,-1]
  print(pheatmap(data, scale= "none", cellwidth = 0.6, cellheight = 0.6,
             cluster_rows = F, cluster_cols = F, legend = T,
             show_rownames = F, show_colnames = F, border_color= F,
             color=viridis(250),  na_col = "WHITE"))
}
```

In the end, these heatmaps can be put together to get a first insight into the changes in metabolic dynamics which occur between the healthy kidney and RCC tissues (Figure 15).

5.3.2. Description of the results

From the ^{13}C-enrichment visualizations in Figure 15 and the graph representations of the results in Figure 16, it becomes apparent that there is a substantial decrease of the TCA cycle-derived ^{13}C isotopomer enrichments when comparing the RCC tissue to the healthy kidney. The decrease of the citrate M+2 and glutamate M+2 fractions indicate a reduced contribution of U-$^{13}C_6$-glucose to the synthesis of these metabolites, and therefore allude to the finding that the TCA cycle activity is lower in the RCC tissue compared to the healthy kidney. Additionally, looking at the fractions of 3PG M+3 and lactate M+3, there does not seem to be much difference between the RCC tissue and the proximal tubules in the healthy kidney.

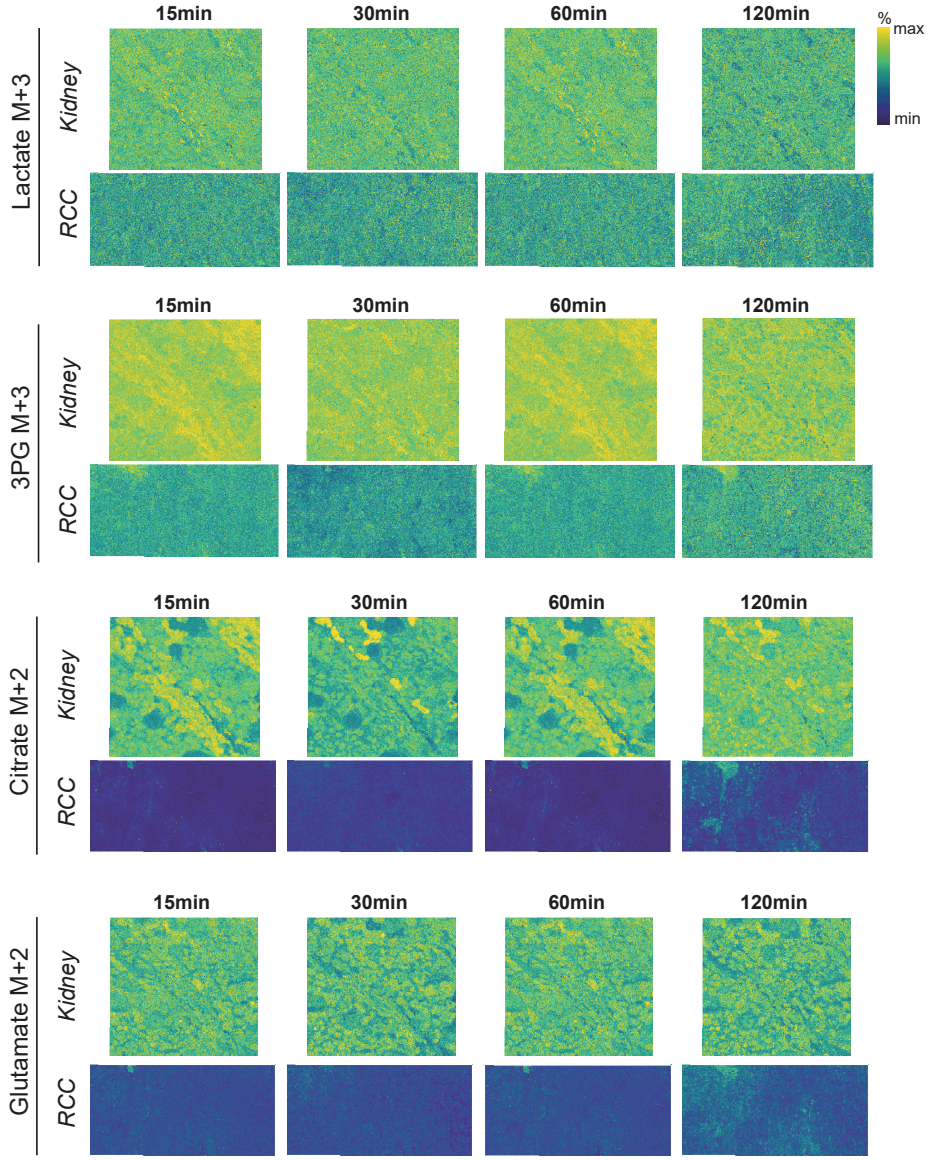


Figure 15. Pseudo-images showing ^{13}C-enrichment of various glycolysis and TCA cycle intermediates over a 120-minute incubation with U-$^{13}C_6$-glucose. The color scale represents the relative ^{13}C-enrichment.

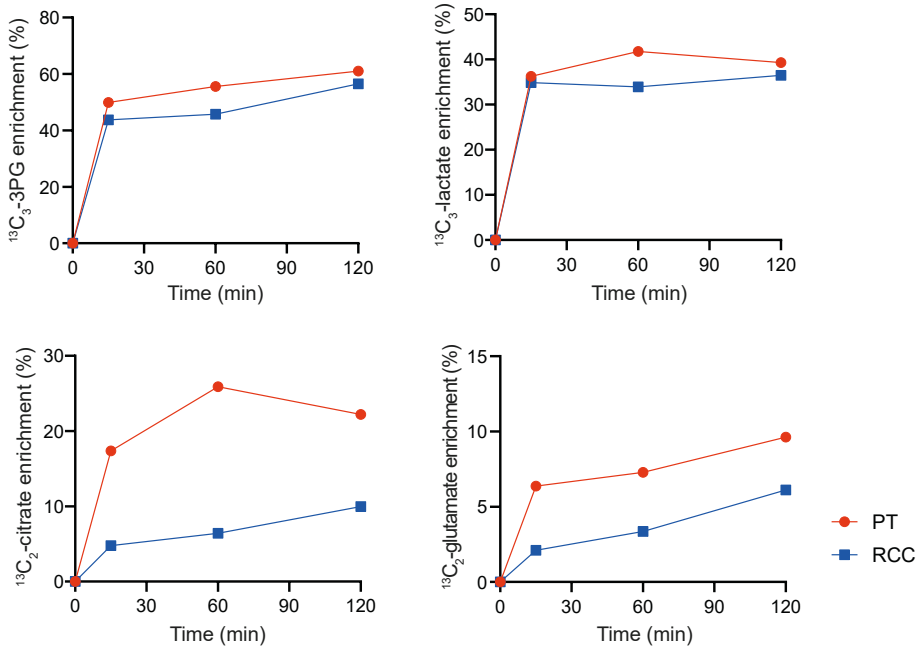


Figure 16. Metabolite-specific ^{13}C enrichment over a 120-minute incubation with U-$^{13}C_6$-glucose.

6. DISCUSSION

Besides providing the reader with an opportunity to learn the ins and outs of spatial dynamic metabolomics, we wanted to test the hypothesis that the combination of *in situ* stable isotope tracing and spatial metabolomics using MALDI-MSI was able to visualize the metabolic anomalies introduced by Warburg effect in RCC.

Since RCC is a tumor that derives from the proximal tubuli in the healthy kidney, we wanted to perform the metabolic comparison between tumor cells and proximal tubule cells. Using a combination of the metabolic histology, based on cell-type specific lipid profiles resulting from the MALDI-MSI analysis (Figure 13), and post-MSI IF staining of several kidney-specific cell types (Figure 11), we were able to isolate the pixels associated to the proximal tubules (PT; pixels positive for LTL staining). Here we performed the IF-based cluster identity assignment based on visual inspection, which is effective, but not the most accurate approach. Beyond the scope of this book chapter, but to achieve the highest level of integration between IF images and MALDI-MSI data one would need to resort to image co-registration strategies. [10]

The Warburg effect encompasses a metabolic shift from aerobic glycolysis and OxPhos to anaerobic glycolysis and loss of OxPhos activity, even in the presence of sufficient oxygen. [28]

Based on this, we expected to find a reduced contribution of stable ^{13}C-isotope labeled glucose to the intermediates of the TCA cycle, as well as increased levels of glycolysis intermediates in the RCC tissues, with an ultimate increase in lactate production. We were indeed able to show the reduced contribution of U-$^{13}C_6$-glucose to citrate M+2, which enters the TCA cycle through the production of pyruvate M+3. Additionally, we see the reduced contribution of labeled glucose to glutamate M+2, which is the glutamate isotopomer specific to the first pass through the TCA cycle, originating from pyruvate M+3, without the interference of other metabolic pathways. [15] These findings point towards the expected decrease of TCA cycle activity in the RCC tissue in comparison to the PTs in the healthy kidney tissue. Despite this finding, however, we could not find the expected changes in dynamics of the glycolysis. Whereas we expected an RCC-specific increase in lactate M+3 enrichment, this remained nearly identical to the PTs of the healthy kidney. This could have several reasons. First, one of the inevitable effects of excessive lactate production upon the Warburg effect is lactic acidosis, which has shown to result in lower glucose consumption, and a reduction in lactate production in favor of a more oxidative, and sustainable non-glycolytic phenotype. [33] To get a conclusive answer to this, we would need to quantify the total lactate pool and show a comparison of the total amount of lactate produced in the healthy kidney compared to RCC. The absolute quantification of metabolites using MALDIMSI is in theory possible but would require a completely different experimental setup than the one presented here and was outside of the scope of this chapter. [34] Alternatively, a tissue homogenate-based technique such as nuclear magnetic resonance (NMR) spectroscopy could be used, which is capable of measuring absolute quantities of lactate. [35] Another contribution to the observation could be related to the post-translational modification of proteins, like for example histone lactylation. When cells are exposed to high levels of lactate, for example as a consequence of the Warburg effect, they will start to utilize lactate as substrate for the post-translational modification of histones, and as such use it as an epigenetic modulator of gene transcription. [36] This lactate-driven gene regulation has recently shown important in the regulation of immune cells and is associated to the modulation of disease-specific immunity status. [37] To exclude this contributor would require an additional quantitative epi-proteomics study to quantify the differential lactylation in histones originating from RCC compared to healthy kidney, this was also deemed beyond the scope of this book chapter.

Ultimately, we have been able to show that the metabolic histology can be used to differentiate RCC from healthy kidney, and that it is able to reveal intratumor heterogeneity, as well as the cellular heterogeneity in the healthy kidney. Furthermore, using our spatial dynamic metabolomics platform we have been able to establish that RCC tissue indeed has reduced TCA cycle activity, and that these tumor cells predominantly rely on glycolysis for energy production. In the end, mass spectrometry imaging in combination with stable isotope tracing provides a powerful platform to show the *in situ* metabolic dynamics of tissues.

7. TAKE-HOME MESSAGES

- Mass spectrometry imaging can be used to gain molecular insight into complex tissue architecture.
- Mass spectrometry imaging is able to provide additional and comprehensive morphological information to conventional histopathological staining.
- The combination of spatial metabolomics and stable isotope tracing provides a powerful platform to show cell-type-specific differences in the dynamics of metabolism.

8. FURTHER READING

Mass Spectrometry: Principles and Applications
Edmond de Hoffmann, Vincent Stroobant
2007, John Wiley & Sons, Ltd, West Sussex England

Correcting for natural isotope abundance and tracer impurity in MS-, MS/MS- and high-resolution-multiple-tracer-data from stable isotope labeling experiments with IsoCorrectoR.
Paul Heinrich, Christian Kohler, Lisa Ellmann, Paul Kuerner, Rainer Spang, Peter Oefner, Katja Dettmer.
2018, Scientific Reports 8, 17910

Getting started with Seurat
Satija Lab 2022.

R for Data Science
Hadley Wickham, Garrett Grolemund 2017.

9. REFERENCES

1 Halim, A. *et al.* Assignment of saccharide identities through analysis of oxonium ion fragmentation profiles in LC-MS/MS of glycopeptides. *J Proteome Res* 13, 6024-6032 (2014).

2 Ma, X. & Fernandez, F. M. Advances in mass spectrometry imaging for spatial cancer metabolomics. *Mass Spectrom Rev*, e21804 (2022).

3 Dreisewerd, K. The Desorption Process in MALDI. *Chemical Reviews* 103, 395-426 (2003).

4 Boesl, U. Time-of-flight mass spectrometry: Introduction to the basics. *Mass Spectrometry Reviews* 36, 86-109 (2017).

5 McDonnell, L. A. & Heeren, R. M. Imaging mass spectrometry. *Mass Spectrom Rev* 26, 606-643 (2007).

6 Ly, A. *et al.* High-mass-resolution MALDI mass spectrometry imaging of metabolites from formalin-fixed paraffin-embedded tissue. *Nat Protoc* 11, 1428-1443 (2016).

7 Goodwin, R. J. A. Sample preparation for mass spectrometry imaging: Small mistakes can lead to big consequences. *Journal of Proteomics* 75, 4893-4911 (2012).

8 Ščupáková, K. *et al.* Cellular resolution in clinical MALDI mass spectrometry imaging: the latest advancements and current challenges. *Clinical Chemistry and Laboratory Medicine (CCLM)* 58, 914-929 (2020).

9 Hattori, K. *et al.* Paradoxical ATP Elevation in Ischemic Penumbra Revealed by Quantitative Imaging Mass Spectrometry. *Antioxidants & Redox Signaling* 13, 1157-1167 (2010).

10 Balluff, B., Heeren, R. M. A. & Race, A. M. An overview of image registration for aligning mass spectrometry imaging with clinically relevant imaging modalities. *Journal of Mass Spectrometry and Advances in the Clinical Lab* 23, 26-38 (2022).

11 Williams, B. *et al.* (ACM).

12 Deininger, S.-O. *et al.* Normalization in MALDI-TOF imaging datasets of proteins: practical considerations. *Analytical and Bioanalytical Chemistry* 401, 167-181 (2011).

13 Abdelmoula, W. M. *et al.* Peak learning of mass spectrometry imaging data using artificial neural networks. *Nature Communications* 12 (2021).

14 Ràfols, P. *et al.* rMSIproc: an R package for mass spectrometry imaging data processing. *Bioinformatics* 36, 3618-3619 (2020).

15 Buescher, J. M. *et al.* A roadmap for interpreting (13)C metabolite labeling patterns from cells. *Curr Opin Biotechnol* 34, 189-201 (2015).

16 Jang, C., Chen, L. & Rabinowitz, J. D. Metabolomics and Isotope Tracing. *Cell* 173, 822-837 (2018).

17 Antoniewicz, M. R. A guide to 13C metabolic flux analysis for the cancer biologist. *Experimental & Molecular Medicine* 50, 1-13 (2018).

18 Hui, S. *et al.* Quantitative Fluxomics of Circulating Metabolites. *Cell Metab* 32, 676-688 e674 (2020).

19 Reinfeld, B. I. *et al.* Cell-programmed nutrient partitioning in the tumour microenvironment. *Nature* 593, 282-288 (2021).

20 Roelants, C. *et al.* Ex-Vivo Treatment of Tumor Tissue Slices as a Predictive Preclinical Method to Evaluate Targeted Therapies for Patients with Renal Carcinoma. *Cancers* 12, 232 (2020).

21 Fan, T. W., Lane, A. N. & Higashi, R. M. Stable Isotope Resolved Metabolomics Studies in Ex Vivo TIssue Slices. *Bio Protoc* 6 (2016).

22 Wang, G. *et al.* Analyzing cell-type-specific dynamics of metabolism in kidney repair. *Nat Metab* 4, 1109-1118 (2022).

23 Padala, S. A. & Kallam, A. (StatPearls, Treasure Island, 2023).

24 Bianchi, C. *et al.* The glucose and lipid metabolism reprogramming is grade-dependent in clear cell renal cell carcinoma primary cultures and is targetable to modulate cell viability and proliferation. *Oncotarget* 8, 113502-113515 (2017).

25 Tian, Z. & Liang, M. Renal metabolism and hypertension. *Nature Communications* 12 (2021).

26 Elia, M. in *Energy Metabolism: Tissue Determinants and Cellular Corollaries* (eds J.M. Kinney & H.N. Tucker) 61-79 (Raven Press, 1992).

27 Jang, C. *et al.* Metabolite Exchange between Mammalian Organs Quantified in Pigs. *Cell Metabolism* 30, 594-606.e593 (2019).

28 Liberti, M. V. & Locasale, J. W. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends in Biochemical Sciences* 41, 211-218 (2016).

29 Strohalm, M., Kavan, D., Novák, P., Volný, M. & Havlíček, V. mMass: A Cross-Platform Software Environment for Precise Analysis of Mass Spectrometric Data. *Analytical Chemistry* 82, 4648-4651 (2010).

30 Alexandrov, T. *et al.* Spatial Segmentation of Imaging Mass Spectrometry Data with Edge-Preserving Image Denoising and Clustering. *Journal of Proteome Research* 9, 6535-6546 (2010).

31 Smets, T. *et al.* Evaluation of Distance Metrics and Spatial Autocorrelation in Uniform Manifold Approximation and Projection Applied to Mass Spectrometry Imaging Data. *Analytical Chemistry* 91, 5706-5714 (2019).

32 Heinrich, P. *et al.* Correcting for natural isotope abundance and tracer impurity in MS-, MS/MS- and high-resolution-multiple-tracer-data from stable isotope labeling experiments with IsoCorrectoR. *Scientific Reports* 8 (2018). https://doi.org:10.1038/s41598-018-36293-4

33 Xie, J. *et al.* Beyond Warburg effect – dual metabolic nature of cancer cells. *Scientific Reports* 4 (2014).

34 Tobias, F. & Hummon, A. B. Considerations for MALDI-Based Quantitative Mass Spectrometry Imaging Studies. *Journal of Proteome Research* 19, 3620-3630 (2020).

35 Kostidis, S., Addie, R. D., Morreau, H., Mayboroda, O. A. & Giera, M. Quantitative NMR analysis of intra- and extracellular metabolism of mammalian cells: A tutorial. *Analytica Chimica Acta* 980, 1-24 (2017).

36 Zhang, D. *et al.* Metabolic regulation of gene expression by histone lactylation. *Nature* 574, 575-580 (2019).

37 Chen, A.-N. *et al.* Lactylation, a Novel Metabolic Reprogramming Code: Current Status and Prospects. *Frontiers in Immunology* 12 (2021).