Title

Development of a Novel Tracer-based Metabolomics Methodology to Study the Metabolism of Cancer Cells

Author

Misuno, Kaori

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Development of a Novel Tracer-based Metabolomics Methodology to Study the Metabolism of Cancer Cells

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Oral Biology

by

Kaori Misuno

2017

ABSTRACT OF THE DISSERTATION

Development of a Novel Tracer-based Metabolomics Methodology to Study the Metabolism of Cancer Cells

by

Kaori Misuno

Doctor of Philosophy in Oral Biology

University of California, Los Angeles, 2017

Professor Shen Hu, Chair

Tracer-based metabolomics is a special form of targeted metabolomics in which the isotopomer distributions of carbon 13 (^{13}C) from a labeled precursor is used to perform quantitative analysis of metabolic phenotypes in cells. The objective of the study is to develop a novel tracer-based metabolomics methodology to investigate the metabolic alterations between cancer and normal cells as well as between high and low invasive cancer cells by tracing ^{13}C in consumed glucose and its integration into protein synthesis. *Our hypothesis is that the incorporation of ^{13}C from glucose into the backbones of cellular proteins reflects the metabolic phenotypes of cancer cells.* To test the hypothesis, we developed a mass spectrometry (MS) based methodology that traces the incorporation of ^{13}C labels in newly synthesized proteins allowing the understanding of the process of partial protein synthesis through the consumption of glucose. Next, we implemented a set of bioinformatics tools to analyze the tracer-based

metabolomics data and compare the metabolic regulation of protein synthesis between normal pancreas epithelial cells and pancreatic cancer cells. Lastly, we applied this new methodology to investigate the metabolic regulation of protein synthesis between high and low invasive head and neck cancer cells and revealed signature proteins that showed significant difference in the glucose utilization for their protein synthesis. Our studies have demonstrated that this novel tracer-based metabolomics methodology is a powerful tool for studying the metabolic regulation of protein synthesis in cancer cells.

The dissertation of Kaori Misuno is approved.

Robert Chiu

Carl Maida

Yifang Zhu

Shen Hu, Committee Chair

University of California, Los Angeles

2017

I dedicate my dissertation to

My parents, husband and daughter.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank Dr. Shen Hu for all his guidance and support throughout my graduate tenure, as well as my thesis committee Dr. Robert Chiu, Dr. Carl Maida and Dr. Yifang Zhu for all their guidance and constructive comments throughout my dissertation.

A big thanks to my family, my husband for his collaboration, guidance, support, encouragement to follow my dreams: his family, for their support and everlasting encouragement.

To my daughter, that even so young she could understand my passion for studies and gave me the strength to continue.

My graduate studies would not be possible without my parents' support; they instilled in me a great work ethics and for that I am more than grateful. Furthermore, I thank my grandmother Maria Eloisa (in memoriam) for all her wisdom and whom seeded my inspiration to pursue biology as well as breaking the glass ceiling in our family to obtain a PhD.

A special thanks to my friends Diana Rigueur and Grace Ji who helped me throughout my graduate studies; they were always there with words of encouragement and palpable drive to succeed during the difficult times. Finally, I thank all my colleagues, and lab mates for their help throughout this journey. I could not have done it without them.

VITA

Kaori Misuno

EDUCATION

2013 M.S., Oral Biology, University of California Los Angeles

2006 D.D.S., Doctor of Dental Surgery, University of Planalto, Catarinense (UNIPLAC), Brazil.

PUBLICATIONS

Peer-reviewed journals

1. Martha Arellano-Garcia, **Kaori Misuno**, Simon Tran, Shen Hu. Interferon-gamma induces immunoproteasomes in human salivary glands cells. PLoS ONE, 2014; 9(8): e102878

2. Junhua Wang, Terri T. Christison, **Kaori Misuno**, Linda Lopez, Andreas F. Huhmer, Yingying Huang, and Shen Hu. Metabolomic Profiling of Anionic Metabolites in Head and Neck Cancer Cells by Capillary Ion Chromatography with Orbitrap Mass Spectrometry. Analytical Chemistry, 2014, 86(10):5116-24

3. Min Zhang, Yang Chai, Jeffrey Brumbaugh, Xiaojun Liu, Ramin Rabii, Sizhe Feng, **Kaori Misuno**, Diana Messadi, and Shen Hu. Oral cancer cells may rewire alternative metabolic pathways to survive from siRNA silencing of metabolic enzymes. BMC Cancer, 2014. 14(1): p. 223.

4. **Kaori Misuno**, Simon D. Tran, Saeed Khalili, Junwei Huang, Younan Liu, and Shen Hu. Quantitative Analysis of Protein and Gene Expression in Salivary Glands of Sjogren's-Like Disease NOD Mice Treated by Bone Marrow Soup. PLoS ONE, 2014. 9(1): p. e87158.

5. **Kaori Misuno**, Xiaojun Liu, Sizhe Feng, and Shen Hu. Quantitative proteomic analysis of sphere-forming stem-like oral cancer cells. Stem Cell Research & Therapy, 2013. 4(6): p. 156.

6. **Kaori Misuno**, Shen Hu. Molecular-Targeted and Cell-Based Therapies for Sjogrens Syndrome. Immunology, Endocrine, Metabolic Agents - Medicinal Chemistry, 2012, 12 (2), 73-79.

7. Jiye Ai, Sizhe Feng, **Kaori Misuno**, Shen Hu. Integrated Omics Analysis of Sjogren's Syndrome. Journal of Integrated OMICS, 2012, 2(2), 6-10. DOI: 10.5584/jiomics.v2i2.97

Book chapter

8. Xiaojun Liu, Yang Chai, **Kaori Misuno**, Min Zhang, Shen Hu. Tissue, Serum and Saliva Sampling for Proteomic Analysis. In Comprehensive Sampling and Sample Preparation, Editor-in-Chief: Janusz, P., Ed. Academic Press: Oxford, 2012; pp 359-364.

Other publications

9. Shen Hu, Martha Arellano, **Kaori Misuno**. Immunoproteasome is activated in Sjogren's syndrome. International Association for Dental Research (IADR) Annual Meeting (invited talk), San Diego, CA. March 16-19, 2011.

Introduction

Warburg Effect

Cancer is a worldwide disease that involves abnormal cell growth, tissue invasion and potential metastasis to local and distant organs. In 2012, 14.1 million new cases were identified and 8.2 million cancer deaths were reported indicating that cancer is a leading cause of morbidity and mortality in the world (Stewart and Wild 2016). In 2017, 1.688 million new cancer cases and 600,920 cancer deaths are projected to occur in the United States (Siegel, Miller et al. 2017). Despite clinical and treatment advances, the overall 5-year survival rates for many types of cancer have remained low during the past few decades. The high mortality rate of cancer highlights the importance of studying the molecular and cellular mechanisms of the disease progression and metastasis.

Tumor cells require rapid synthesis of macromolecules such as nucleotides, lipids and proteins for growth. Hence, altered cellular metabolic regulation has been intrinsically linked with oncogenic progression (Muñoz-Pinedo, El Mjiyad et al. 2012); these macromolecular processes encompass aerobic glycolysis, fatty acid synthesis, and glutaminolysis. Moreover, cancer cells exhibit reduced oxidative phosphorylation for cell growth and proliferation (Zhao, Butler et al. 2013). One of the first identified biochemical hallmarks of tumor cells was the “Warburg effect,” a shift in the glucose metabolism from oxidative phosphorylation to increased aerobic glycolysis, even in the presence of sufficient oxygen (Kim and Dang 2006). This phenomenon was a unique characteristic of many cancer cells and proliferating cells, and was also coupled with enhanced glucose, glutamine uptake, and lactate production. In addition, as Rasola *et al.* stated, glucose uptake is important for cancer cells to maintain their metabolism and aggressive growth (Rasola, Neckers et al. 2014). Since then, identifying unique aspects of metabolic pathways in cancer cells has received increasing attention (Vaitheesvaran, Xu et al. 2014).

According to Vaitheesvaran *et al.*, Otto Warburg's original observation that cancer cells metabolize glucose via glycolysis, which is a less efficient pathway for generating ATP despite oxygen availability, appears to be a paradoxical phenomenon which has not been satisfactorily resolved (Vaitheesvaran, Xu et al. 2014). Warburg effect is based on the assumptions that lactate is solely produced by glycolysis, and that ATP production by oxidative phosphorylation (OXPHOS) is the most efficient way of producing energy. Current techniques for the regulation of cancer cell metabolism are not capable to validate these assumptions. Moreover, to prove that these assumptions are correct it requires quantitative data encompassing a distribution of glucose carbons among its intermediates and how glucose is used for NADH or ATP production (Ferreira 2010, Levine and Puzio-Kuter 2010). The use of the C13 stable-isotope glucose tracer and mass isotopomer analysis is currently the preferred technique in generating quantitative data for isotopomer distribution (Boros, Cascante et al. 2002, Boren, Lee et al. 2003) and such data can be interpreted using balance of flux analysis metabolomics and phenotypic phase plane analysis (Lee 2006, Orth, Thiele et al. 2010). This method is useful to understand regulation of proteins in cells since the level of protein expression depends on the rates of its synthesis and degradation. It provides an understanding of cellular homeostasis and its changes under the influence of nutrient conditions and pharmaceuticals (Lee, Boros et al. 2012). The application of this methodology is diverse due to the cell dependence of carbon. Although the isotope is incorporated into glucose, the glucose molecules undertaken by cells integrated into all carbon-based products , highlighting the preferred metabolism and thus signature of the disease (Fan, Lorkiewicz et al. 2012). Understanding the metabolomics signatures is a valuable tool for designing approaches for modulating metabolic activity for therapeutics (Cascante, Selivanov et al. 2012).

Glucose Metabolism

Glucose metabolism is involved in a variety of cellular processes, such as nucleotides synthesis, protein synthesis, non-essential amino acids synthesis, and fatty acid synthesis, all of

which are essential for cellular function and cell survival. Cells use the glycolysis pathway to generate pyruvate. This process is conducted by breaking down glucose ($C_6H_{12}O_6$) to pyruvate ($C_3H_3O_3$). In most mammalian cells, glycolysis is an oxygen-independent metabolic pathway; in other words, in the absence of oxygen, pyruvate is converted into lactate and gets secreted out of the cell as waste product. On the contrary, in the presence of oxygen, decarboxylation of pyruvate occurs in the mitochondrial matrix. As mentioned before, conversion of glucose to lactate in the presence of oxygen is known as the "Warburg effect" or aerobic glycolysis and this is observed in cancers (Gatenby and Gillies 2004, DeBerardinis, Sayed et al. 2008). Moreover, pyruvate dehydrogenase converts pyruvate into acetyl-CoA, which then enters the tricarboxylic acid (TCA) cycle where carbon atoms are oxidized to release CO_2, and energy captured by NADH. Thus, intermediates of the TCA cycle are transported out of the mitochondria for synthesis of non-essential amino acids and fatty acids, which are the building blocks of proteins and lipids. An alternative pathway to glycolysis is pentose phosphate pathway (PPP), which is fundamentally an anaerobic process and generates pentose phosphates for ribonucleotide synthesis and nicotinamide adenine dinucleotide phosphate (NADPH). The pentose phosphate pathway entailed two distinctive branches: oxidative branch and non-oxidative branch. The oxidative phase of PPP involves NADPH production for maintenance of cellular reductive power. The non-oxidative phase of PPP involves ribose 5-phosphate (R5P) and erythrose-4-phosphate (E4P) generation for synthesis of nucleotide, nucleic acids (DNA and RNA), and aromatic amino acids (Kruger and von Schaewen 2003). In addition to glucose oxidation, glutamine utilization may also contribute to these metabolic pathways for macromolecule synthesis. Overall these studies were possible because of the development of ^{13}C stable-isotope strategies (Kruger and von Schaewen 2003, Hay 2016).

Metabolomics

The term “metabolome” was first introduced by Oliver *et. al.* in 1998 and it was used to describe the total metabolite content of a biologic sample (Oliver, Winson et al. 1998); subsequently, the term metabonomics was coined in 1999 which describes the analysis of changes in the metabolic status of an organism as a consequence of drug treatment, environmental influences, nutrition, lifestyle, genetic effects, toxic exposure, diseases, and other factors, which was then (Nicholson, Lindon et al. 1999) followed by the introduction of the term metabolomics in 2002: the identification and quantification of every single metabolite in a biological system (Fiehn 2002). Today, the two terms metabolomics and metabonomics are often used interchangeably and the word, metabolomics, is more widely accepted. Therefore, metabolomics is defined as the study of small molecule metabolites such as intermediates, secondary metabolites, hormones and other signaling molecules that are found in cells, tissue and fluids. While mRNA gene expression and proteomic analyses do not tell the whole story of what is happening in a cell, metabolic profiling can give an instantaneous snapshot of the integrated cellular physiology of the cell (Bennett 2005, Jordan, Nordenstam et al. 2009). Figure 1 presents the central dogma of the “omics” that links metabolomics to physiological knowledge from genome-type to phenotype. In metabolomics studies, extracellular metabolite profiles (metabolic footprinting or exometabolome) capture the features of overflow metabolism and offer a noninvasive approach to probe cells’ physiological alterations (nutrient uptake and metabolite secretion) in the appropriate media provoked by environmental or genetic perturbations (Allen, Davey et al. 2003). The human metabolome database (HMDB) consists of more than 40,000 metabolites entries that have already been identified or are likely to be found in the human body, and these include metabolites produced by external and internal compounds which are not only related to our genome but also microfloral genomes and, xenobiotic chemicals (Martin, Collino et

al. 2011, Savorani, Rasmussen et al. 2013, Wishart, Jewison et al. 2013, Wishart, Mandal et al. 2016).

Metabolites can be identified, quantified or classified by using different technologies such as nuclear magnetic resonance spectroscopy (NMR) or MS. MS is a powerful tool for metabolomic analysis with higher sensitivity than NMR (Griffin and Kauppinen 2007, Theodoridis, Gika et al. 2012, Patti, Tautenhahn et al. 2013). Based on MS analysis and database searching, various small molecules such as lipids, sugars, and amino acids can be identified (Patti, Tautenhahn et al. 2013). MS-based metabolomics is best performed when coupled to a separation technique, such as capillary electrophoresis, gas chromatography (GC), or liquid chromatography (LC) (Jellum, Bjørnson et al. 1981, Wysocki, Resing et al. 2005, Hu, Wang et al. 2015). GC-MS was applied to metabolomics studies much earlier than LC-MS and it is a tool limited to the analysis of volatile molecules (such as fossil fuel/oil analysis, doping and illicit drug control, plant metabolomics, screening for inborn errors of metabolism and others) or molecules that can be made to be volatile (Hori, Hasegawa et al. , Brunelli, Bicchi et al. 2006, Brown, Rhodes et al. , Meyer, Peters et al. 2010, Theodoridis, Gika et al. 2012). On the other hand, LC-MS is increasingly used versus GC-MS due to sample preparation that is less time consuming and consequently less prone to errors. In addition, LC-MS is most widely used for metabolomics due to its ability to separate and identify a wide range of molecules. This method allows for both quantitative and structural analyses (Griffin and Shockcor 2004).

In early metabolomics studies, NMR was widely used to study pathological factors and mechanisms (e.g., xenobiotic metabolism related to drug resistance in infectious diseases or cancer chemotherapy) (Griffin 2003). However, to reveal or calculate isotopomers, multi-dimensional NMR is required and the analysis is time consuming. MS has become a common approach to investigate cell metabolism including identification of amino acids, organic acids and sugar phosphate. Modern MS instruments are able to probe many low abundant metabolites, particularly hormones and natural products and may generate comprehensive insight into

genome-wide cell metabolisms. Both are proteome and metabolome context-dependent, that is, the total complement of proteins or metabolites changes according to the physiological, developmental, or pathological state of a cell, tissue, organ, or organism. Moreover, unlike transcriptome analysis of messenger RNA (mRNA) molecules, proteins and metabolites are functional entities within the cell (Raamsdonk, Teusink et al. 2001). Each type of organism may have a unique subset of intracellular metabolites (metabolite fingerprints) that elucidate specific cellular processes (Raamsdonk, Teusink et al. 2001).

Global metabolomic profiling may also be used to link cellular biochemistry to systems biology via tracking alterations in small molecule fluxes that determine cellular dynamics with respect to dysregulation of metabolic flux as a function of time (Patti, Tautenhahn et al. 2011). In a previous study, we have demonstrated a capillary ion chromatography (CapIC/MS) with Q exactive MS for global profiling of metabolites in oral/head and neck cancer cells, including cancer stem-like cells (CSCs) and non-stem cancer cells (NSCCs). Due to the superior resolution and sensitivity of Cap IC, we were able to quantify more than 4000 metabolites in oral/head and neck cancer cells. Our studies also revealed the differential changes of metabolites in CSCs and aided in elucidating the possible role of glycolysis/TCA pathways in oral/head and neck cancer cells (Wang, Christison et al. 2014).

Tracer-based Metabolomics

Metabolic fluxes reflect the quantitative final product between gene expression, protein synthesis, post-translational modifications and thermodynamic constraints; and this represents the cell's final phenotype. Analyzing intracellular metabolic fluxes is essential to investigate the physiological state of a cell and thereby revealing disease specific alterations or dysregulations of metabolic conversion rates and enzyme activities. As a consequence of the complex regulation of metabolic pathways, significant flux changes are sometimes associated only with a modest change in metabolite concentrations (Fell 2005). For instance, target systems (cell culture, tissue,

whole organism) are fed with stable isotopes (^{13}C, ^{15}N) and substrates (glucose, glutamine) labeled with the objective to investigate metabolic pathway activity. The metabolism of glucose is integrated into multiple pathways, which means that the carbon atoms of glucose can potentially be incorporated into a wide-range of metabolites. This provides the ability to simultaneously survey the activity of numerous metabolic pathways using ^{13}C-glucose (Fan, Lorkiewicz et al. 2012).

A stable isotope tracer is a molecule with one or more isotopes with a different mass than the most abundantly occurring mass incorporated somewhere in the molecule. In the case of the most commonly used tracers (e.g., C, H and N), the stable isotope tracers are heavier than the most commonly occurring mass. The fact that the tracer is heavier than the tracee allows for determination of the ratio of tracer to tracee by measuring the mass of the molecule of interest (Kim, Suh et al. 2016).Traditionally, radioactive tracers had been used to determine protein synthesis and turnover because of its high sensitivity. The appearance of radioactivity in the product, or the disappearance of radioactivity in the precursors over time, is used to determine the rate of reaction of enzymes (Kim, Suh et al. 2016). However, it has long been recognized that in metabolic studies in cell cultures or in whole animals, the precursor specific activity is often unknown through dilution by unlabeled precursor or recycling of the tracer through other metabolic intermediates (Hellerstein and Neese 1999). The dilution of the tracer along the path of synthesis greatly complicates the determination of substrate fluxes. Various techniques have been used to overcome such difficulties using specifically (recyclable and non-recyclable) labeled precursor, or performing detailed chemical degradation of the product to identify the product of a specific pathway (Katz and Grunnet 1979). However, with the advance of stable isotope and mass isotopomer analysis, positional isotopomers from isotope exchange and dilution through intersecting pathways can be estimated without the need for chemical degradation. Stable isotope and mass isotopomer analysis has essentially replaced radioactive tracers in complex metabolic studies.

Tracer-based metabolomics is a collective analysis of metabolites on the metabolic pathways that transfer the tracer (e.g., ^{13}C carbon) from precursors to specific end products. Traditional metabolite profiling analysis (measuring the total complement of individual metabolites in a given biological sample) and tracer-based metabolomics are complementary to each other, and the combined data should describe the interaction between the cellular metabolic networks with its nutrient environment (Gill, Thomas et al. 1990, Usenius, Kauppinen et al. 1994, Lee, Wahjudi et al. 2010).

Therefore, the integration of isotope-tracer with metabolomics studies provides a comprehensive understanding of the cellular metabolome and a simultaneous analysis of numerous metabolic pathways. The development of concept and practice of tracer-based metabolomics comes from early radioactive label application in the studies of metabolic network in model cells. With the advance of new MS technology, the use of radioactive tracers for metabolomics analysis have been replaced with stable isotopes (e.g., [U-$^{13}C6$] glucose) and mass isotopomers in complex metabolic studies. The isotopomer [U-$^{13}C6$] of glucose is naturally occurring and stable within cells over time without undergoing spontaneous decay with emission of radiation. It can be used to trace metabolites that target its degradation without the need of purifying targeted products and with its employed utility, we can generate positional mass isotopomers from ^{13}C labeled precursors. Carbon has two stable isotopes, carbon 12 and carbon 13 which natural occurrence of approximately 98.9% and 1.1%, respectively. Because the glucose molecule has six carbons, and each carbon has approximately 1.1% of chance to be ^{13}C, 6.6% of naturally occurring glucose will be glucose with having at least one ^{13}C. When metabolites incorporate stable isotopes such as ^{13}C, labeled and unlabeled metabolites can be separated by MS based on their m/z (mass-to-charge) ratios. Depending on the incorporated labeled atoms, the m/z ratio is shifted by one or more atomic mass units, resulting in different shifted mass spectra and this will generate a mass isotopomer distribution (Kim, Suh et al. 2016).

Tracer-based Metabolomics of Cancer Cells

The ^{13}C glucose has been used as a single tracer in the study of direct/indirect pathways of glycolysis, PPP, gluconeogenesis and TCA cycle metabolism and this has been the main focus for tracer-based metabolomics (Katz, Lee et al. 1989, Katz and Lee 1991, Katz, Wals et al. 1991, Katz, Wals et al. 1993, Lee, Boros et al. 1998). However, the concept and analytical methodology of tracer-based metabolomics can also be applied to studies of many other metabolic systems such as essential amino acids, essential fatty acids and nucleic acids. The methodology for such studies remains to be further developed and validated (Lee, Wahjudi et al. 2010).

Since carbon, nitrogen, and hydrogen atoms are incorporated from their precursor substrates through exchanges during the process of amino acid synthesis, a ^{13}C labeled substrate can be involved in metabolite and protein synthesis (Xiao, Garg et al. 2008). Amino acids that are incorporated with a ^{13}C labeled substrate are heavier than “natural” amino acids because they contain ^{13}C (Xiao, Garg et al. 2008). Stable isotope labeling of non-essential amino acids can be used to determine protein synthesis in cancer cells. ^{13}C labeling (heavy) to replace natural ^{12}C (light) can lead to mass difference in peptides of newly synthesized proteins, and link de novo protein synthesis process to nutrient metabolic pathways to enhance current understanding of nutrient-gene interaction. Cells can breakdown [U-$^{13}C_6$]-glucose for anabolic processes and trap ^{13}C into the carbon backbones of intracellular amino acids during protein synthesis. The mass difference of ^{13}C labeled peptide can then be readily detected by MS as a distribution of mass isotopomers. This pattern distribution reflects respective metabolic fluxes as the amino acid isotopomers carry information regarding the paths they have traversed, and thus the final distribution can reflect the relative contribution from these individual pathways (Xiao, Garg et al. 2008).

By establishing a new methodology that combines traced-based metabolomics with proteomics, we can better understand the metabolic regulation of protein synthesis in cancer cells.

This method is based on developing a bioinformatics tool that analyzes the differences in ^{13}C glucose uptake/conversion to metabolomes and subsequent to amino acids. In addition, we have employed this methodology to study the metabolic phenotypes of pancreatic cancer cells, highly invasive oral cancer cells and CSCs. Therefore, the overall goal of our study is to develop a novel platform to investigate the metabolic phenotypes of oral/pancreatic cancer cells and the metabolic regulation of protein synthesis in those cancer cells, with the following objectives:

This dissertation includes three chapters focused on the development of a novel tracer-based metabolomics technology to understand the protein synthesis in pancreatic and oral/head and neck cancer cells.

Chapter 1: Develop a Bioinformatics Tool for the Data Analysis of Tracer-based Metabolomics

Chapter 2: Investigate the Metabolic Regulation of Protein Synthesis between Normal Pancreas, Pancreatic Cancer Cells and Oral Cancer Stem Cells

Chapter 3: Study the Metabolic Regulation of Protein Synthesis between High and Low Invasive Cancer Cells (head and neck squamous cell carcinoma, HNSCC)

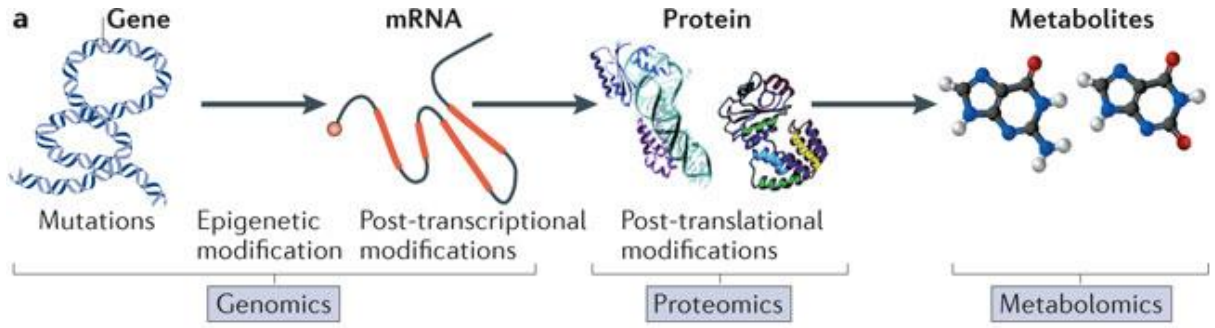


Figure 1. The illustration represents the central dogma of “omics” studies. Whereas genes and proteins are subject to regulatory epigenetic processes and post-translational modifications, respectively, metabolites represent downstream biochemical end products that are closer to the phenotype. The “Omics” cascade comprises complex datasets that as an entity comprehensively describe the response of biological systems to disease, genetic, and environmental perturbations. The most powerful database will integrate data from all omic levels. However, of these databases the metabolome is the most predictive of phenotype. (Adopted by Patti,GJ, Yanes O, Siuzdaz G, Innovation: Metabolomics: the apogee of the omics trilogy, Nature Reviews Molecular Cell Biology 2012).

Chapter 1: Develop a Bioinformatics Tool for the Data Analysis of Tracer-based Metabolomics

Introduction

Metabolome is the ultimate functional expression of gene and protein activities to meet the physiological demands for growth and survival in response to the effects of disease, therapeutic agents, toxicity, or the environment. It is also an integral partner in molecular regulations for homeostasis regulating enzyme activities, as well as protein and gene expression events in a feedback loop (Fan, Lorkiewicz et al. 2012, Gowda and Djukovic 2014). Metabolic reprogramming is a common feature of cancers and other metabolic diseases, which can be caused by genetic mutations. This is exemplified by the long-known aerobic glycolysis or "Warburg effect" (Warburg 1956) in many cancer types and the recent discovery of a defective mitochondrial Krebs cycle associated with a loss of function fumarate hydratase mutation (Gottlieb and Tomlinson 2005) or isocitrate dehydrogenase 1 (IDH1) mutation diverting isocitrate to 2-hydroxyglutarate (Gross, Cairns et al. 2010). Therefore, cancer cells possess altered metabolic pathways that provide a support foundation for the tumorigenicity and malignancy. Cancer research has benefited from the integration of large-scale profiling platforms from genomics, proteomics and metabolomics, which together can provide a global assessment of how enzymes and their parent metabolic network become altered in cancer to fuel tumor growth (Benjamin, Cravatt et al. 2012).

The number of MS-based metabolomics studies has increased exponentially over the last decade, being applied to numerous diseases such as breast cancer (O'Connell 2012), colorectal cancer (Williams, Reeves et al. 2013), prostate cancer (Trock 2011), esophageal and gastric cancer (Abbassi-Ghadi, Kumar et al. 2013), cardiovascular diseases (Rhee and Gerszten 2012, Shah, Kraus et al. 2012) kidney diseases (Weiss and Kim 2012), inborn errors of metabolism

(Mamas, Dunn et al. 2011), effects of toxicology (Robertson, Watkins et al. 2011) and nutrition (Scalbert, Brennan et al. 2009), and metabolic fluxes (Zamboni, Fendt et al. 2009, Reaves and Rabinowitz 2011). Hereafter, identification and quantification analysis of metabolites in humans, animal and cell models related with cancer and other metabolic diseases will open venues in understanding the diseases, diagnosis and creating novel therapeutic options.

Tracer-based metabolomics has emerged as a powerful tool for studying cancer cell metabolism. The approach is based in mass isotopomers of a metabolite which are isomers that differ only in the isotope distributions for example ^{13}C vs ^{12}C. The use of ^{13}C-labeled substrate will incorporate itself either through exchange or by direct synthesis into a wide range of metabolites and proteins (Xiao, Garg et al. 2008). In our study, during protein synthesis, the incorporation of ^{13}C (derived from [U-$^{13}C6$] glucose) generates peptide mass isotopomer distributions that can be detected in MS (Lee, Wahjudi et al. 2010).

The software available for the analysis of labeled MS data include CAMERA, MetExtract, MAVEN, mzMatch-ISO; however, they have some limitations, such as rapid differentiation and relative quantification of isotope patterns is not currently possible with this software. Metabolomic Analysis and Visualization Engine (MAVEN) is a data visualizer and isotopic flux animator which offers automated inference into biological events detected within a study, making it an excellent tool for the analysis of isotope labeling studies in systems biology. Unfortunately, custom algorithm development and data integration that would allow the software's extension require expertise in system level programming language. Thus, for LC-MS study of metabolites, the ideal features for systematic analysis of data derived from heavy isotope labeling experiments would be visualization options for unlabeled and labeled chromatograms for rapid visual comparison, assessment of peak shape and relative intensity across samples, output that facilitates downstream statistical and modeling analyses and easy integration with existing software (Chokkathukalam, Kim et al. 2014).

By developing a methodology to analyze the mass isotopomer distributions of metabolites in conjunction with isotopomer distributions of peptides from newly synthesized proteins, we may understand the metabolic regulation of protein synthesis. As part of this, a new bioinformatics tool based on MATLAB has been developed to analyze the mass isotopomer distributions and to identify the position of ^{13}C label into proteins. This new tool was used as a visualization software to display mass isotopomer pattern differences and to generate global ranking of differences across proteins globally.

The high-level outline of this methodology is:

1. Introduce stable isotope [U-$^{13}C6$] glucose in the culture medium of cell lines.
2. Perform LC-MS/MS analysis of unlabeled and ^{13}C-labeled metabolites and proteins.
3. Identify peptide sequences and proteins using the Proteome Discoverer.
4. Execute comparison analysis of cell lines using our developed bioinformatics tool to identify, visualize and rank differences in ^{13}C incorporation.

The aim of this chapter is to describe our new developed bioinformatics tool. Therefore, this is a new platform combining two technologies proteomics and metabolomics, and creating a novel field defined as metaboproteomics which has a definition of the study of protein synthesis through stable-isotope labeled, tracer-based metabolomics.

The HNSCC, CSC, NSCC and pancreatic cell lines were cultured in two different groups: Group #1 without glucose supplement and Group #2 with ^{13}C glucose. After LC-MS/MS analysis, the RAW data files were used for the informatics analysis, as detailed below:

Data Processing Phase

All the RAW files were reformatted using different converters in order to get specific data formats compatible with the MATLAB. First, RAW files were converted to mzXML which is XML (eXtensible Markup Language) based common file format for MS-based proteomics data (Pedrioli, Eng et al. 2004, Lin, Zhu et al. 2005). This format was developed at the Seattle

Proteome Center/Institute for Systems Biology. The .mzXML files were processed with msconvert and MM File Converter **(**Error! Reference source not found.**)**. Msconvert is a command line tool c onverting to/from various MS data formats (Yonezawa, Nishiumi et al. 2013) and it was used to obtain scan number, precursor number, charge, detailed m/z (mass/charge) and intensity. We also use MM File Converter, which converts common input formats: RAW (Thermo) to mzXML, MGF, and this was used to acquire scan number, m/z and coarse m/z and intensity.

The Proteome Discoverer software supports multiple dissociation techniques, quantitation technologies, and database search algorithms for comprehensive protein identification, characterization, and quantification. Thus, this software was used to perform a MASCOT database search and to obtain outputs (.xls) such as peptide sequence, accession number, protein name, charge, peptide spectrum matches (PSM) ambiguity and MS2 scan number **(**Error! R eference source not found.**)**. MASCOT is a software search engine that uses MS data to identify proteins from peptide sequence databases (Cottrell and London 1999, Koenig, Menze et al. 2008). Moreover, MASCOT is widely used by research facilities around the world. In addition, it uses a probabilistic scoring algorithm for protein identification that was adapted from the MOWSE algorithm (Koenig, Menze et al. 2008), and identifies proteins by interpreting MS data (Elias, Haas et al. 2005). The prevailing experimental method for protein identification is a bottom-up approach, where a protein sample is typically digested with trypsin to form smaller peptides. While most proteins are too big, peptides usually fall within the limited mass range that a typical mass spectrometer can measure. Mass spectrometers measure the molecular weights of peptides in a sample. MASCOT then compares these molecular weights against a database of known peptides. The program cleaves every protein in the specified search database in silico according to specific rules depending on the cleavage enzyme used for digestion and calculates the theoretical mass for each peptide. Mascot then computes a score based on the probability that the peptides from a sample match those in the selected protein database. The more peptides

Mascot identifies from a particular protein, the higher the Mascot score for that protein (Pappin, Hojrup et al. 1993).

Data Analysis Phase

All the data were imported into the MATLAB by using our JDImportData script (Error! R eference source not found.). MATLAB (matrix laboratory) is a multi-paradigm numerical computing environment and it is used for engineers and scientists worldwide. In addition, MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages, including C, C++, C#, Java, Fortran and Python.

The imported spectra data was stored in a variable called “ms”, and the imported mascot search data was stored in the “mascot” variable. The data were then analyzed by our JDFindEnvelopes script which performs the following steps (Error! Reference source not found.).

Identify MS2 Scans from Same Isotopomer Distribution Scan:

The first step of ^{13}C label analysis algorithm was grouping MS2 scans—contained in the groupMS script—by using 6 criteria: (1) same Sequence, (2) same Accessions, (3) same Protein Descriptions, (4) same precursor (MS1) scan, (5) same charge and (6) within a reasonable m/z distance (skip at most 3 peaks). Each group gets assigned a GroupID.

The first 3 criteria come from the MASCOT search results, whereas the last 3 come from the raw MS data itself.

Pseudocode for the groupMS script

Group MS2s initially by Sequence, Accessions, Protein Descriptions, same precursor (MS1) scan #, and same charge (z)

For each group:

Sort MS2 scans by their m/z value (ascending)

Breakdown into groups of peaks spaced by $4 * {}^{1}/_{z} \pm \varepsilon$ at most

*where **z** is the ionization charge and **ε** is the tolerable margin of error*

Assign each group a GroupID

Based on sample testing, the tolerance value was experimentally optimized to 0.02.

Identify Range of Each Isotopomer Distribution Scan:

For each isotopomer distribution scan, we looked up the monoisotopic m/z corresponding to the sequence included in the mascot results; this corresponds to the first peak of the distribution. Then, we find isotopomer peptides by looking at 1/z separations for all peaks selected for MS2. Thereafter, keep selecting peaks with the same 1/z separation if the peak's intensity is at least 10% of the basepeak's intensity; basepeak is defined as the peak in the distribution with highest intensity which tends to be the monoisotopic peak. The stopping criteria is the second occurrence that does not meet the intensity criterion, which allows for the case where 1 peak is missing but the next peak is present.

Pseudocode to Identify range of each Isotopomer Distribution Scan

For each peptide sequence

Let M be the m/z of the monoisotopic peak

Let $i = 0$ be the peak's index (from 0 as the monoisotopic peak)

Let n be the index of the last peak selected for MS2 scanning

Let $misses = 0$ be the stop search criterion

While $misses < 2$

Set peak P_i as the peak with highest intensity and m/z within $\left(M + \frac{i}{z}\right) \pm \varepsilon$

If $i > n \; and \; intensity(P_i) \; not \; within \; 10\% \; and \; 100\% \; of \; intensity(P_0)$

Set $misses = misses + 1$

End If

Set $i = i + 1$

End While

End For

*where **ε** is the tolerable margin of error*

Based on sample testing, the tolerance value was experimentally optimized to 0.015.

Combine Isotopomer Distributions of each Unique Peptide Sequence:

There may be multiple isotopomer distributions scans for the same peptide sequence which may have been ionized by different charge (z) values by the MS machine. Therefore prior to combining these, the m/z values need to be converted to non-ionized mass values (M) using the following equation:

$M_i = ([m/z]_i * z) - z$ where $i \geq 0$ is the peak's index within the distribution

Thus M_0 corresponds to the theoretical mass of the monoisotopic peptide sequence.

A list of unique peptide sequences is built—defining "unique" as a permutation of Sequence, Accessions, and Protein Descriptions. There may be multiple isotopomer distribution scans corresponding to the same unique peptide sequence so the natural next step is to combine them into the average isotopomer distribution observed across these samples. This process is in the "AggregatePeaks" script (Error! Reference source not found.).

The process begins by ensuring that the monoisotopic mass (M_0) of the samples to be combined are similar. Based on sample testing, the tolerance value was experimentally optimized to within 0.1 of each other; this means samples whose monoisotopic masses (M_0) differ by more than 0.1 are excluded.

Combining the masses of the isotopomer distributions for each unique peptide sequence results in defining $\bar{M}_i$ as the average non-ionized mass of the i[th] peak:

$\bar{M}_i = average(M_i^x)$ where $x > 1$ is the sample isotopomer distribution index

Similarly, the intensities of the different sample isotopomer distributions are combined as:

$$I_i^{sum} = \sum_x I_i^x$$

Finally, there are 2 versions of relative abundance defined:

$$relative\ abundance\ A_i^{Nsum} = \frac{I_i^{sum}}{\sum I_i^{sum}}\ normalized\ by\ the\ sum\ of\ the\ peaks$$

$$relative\ abundance\ A_i^{Nmax} = \frac{I_i^{sum}}{\max_i\left(I_i^{sum}\right)}\ normalized\ by\ the\ basepeak's\ intensity$$

The property of A_i^{Nsum} is that the relative abundances of the peaks in the isotopomer distribution aggregates to 100%, whereas the property of A_i^{Nmax} is that the relative abundance of its highest peak is set to 100%. Thus, the best metric to use depends on which property is favored for the given analysis framework.

Compare Isotopomer Distributions for the Same Peptide Sequence between Two Cell Lines:

The comparison between two cell lines starts with performing a union between the two lists of unique peptide sequences (**Error! Reference source not found.**). Some unique peptide s equences appear in only one of the 2 cell lines whereas others—the intersection between the two sets - appear on both.

For those peptide sequences that appear on both, the Chi-square distance metric was used to measure the difference between the isotopomer distributions. Normalized Chi-square distance is defined as:

$$d = \frac{1}{n}\sum_{i}^{n} \frac{(Y1_i - Y2_i)^2}{\frac{Y1_i + Y2_i}{2}} \; where \; 0 \leq d \leq \infty$$

Chi-square distance is extensively applied in correspondence analysis (Chardy, Glemarec et al. 1976) for ecological research, as well as more recently in computer vision. It aims to quantify the differences between distributions/histograms by summing up the pair-wise differences scaled by their individual averages. It is similar to Euclidean distance, except the differences are squared and thus it does not assume a linear relationship between sample dissimilarity.

In the above formula, the distance has been normalized by the number of peaks n in the distributions since in our analysis the number of peaks in the isotopomer distributions across sequences is variable.

For this analysis, the relative abundance A^{max} was selected because it is more suitable for a peak by peak distribution comparison. For example, if the two distributions - UM1 and UM5 - are identical except the isotopomer distribution in UM1 has one extra peak at the end, the Chi-square distance d would simply equate to the A_i^{max} of the extra peak.

Results Visualization:

For ease of analysis the results were exported from the MATLAB to macro-enabled Microsoft Excel spreadsheets that possess the ability to provide a mirror display for the comparison of two different cell lines (**Error! Reference source not found.**).

Enrichment Ratio Analysis:

The observed spectrum of a peptide (both pre-existing and newly synthesized) is then a linear combination of isotopomers of natural (unlabeled) and the expected labeled peptides, and the ratio of labeled to the total isotopomers provides a molar fraction of the newly synthesized protein. The normalization converts the individual intensity to relative intensity with the sum of intensities of all peaks equal to 1 (Zhao, Lee et al. 2009).The isotopomer distribution of peptides after labeling can be calculated by the enrichment ratio which is the calculation of the ratio of labeled or new (N) and unlabeled, old (O) and was determined by the formula:

$$\frac{N}{(O+N)}$$

Given that N is unknown since the enrichment process was stopped before its full completion, the formula was reworked using the following:

$$O + N = {}^{13}C$$

$$N = (O+N) - O = {}^{13}C - {}^{12}C$$

$$\frac{N}{(O+N)} = \frac{{}^{13}C - {}^{12}C}{{}^{13}C}$$

For this metric, the relative abundance A^{sum} was selected because it is more suitable for trapezoidal area comparison analysis. This approach is also aligned with earlier descriptions and uses of this metric in literature (Cargile, Bundy et al. 2004, Zhao, Lee et al. 2009).

Discussion

In principle, measuring protein turnover through stable isotope tracers (^{13}C) relies on the enrichment of ^{13}C in the labeling agent ([U-$^{13}C_6$]-glucose 99%) and the enrichment of ^{13}C in the protein (Xiao, Garg et al. 2008). Enrichment of ^{13}C into amino acids of proteins will produce a heavier mass than the natural ^{12}C. In the case of the amino acid highly enriched with ^{13}C, it will generate a larger isotopomer distribution that will have a slow decay compared with ^{12}C unlabeled amino acid (Xiao, Garg et al. 2008). Therefore, the spectrum will be smaller in unlabeled ^{12}C (containing approximately 3 peaks in the spectrum). The peak intensities and mass differences between ^{13}C and ^{12}C can be calculated by the enrichment ratio that represent the relative protein turnover dynamics in the cell (Cargile, Bundy et al. 2004, Xiao, Garg et al. 2008, Zhao, Lee et al. 2009). In our project, we considered an isotopomer distribution (^{13}C) to be more than four peaks where carbon ^{13}C starts to get incorporated otherwise is a ^{12}C natural distribution.

Cargile *et al.* described a method to determine protein synthesis without the effort of determining precursor enrichment which they introduced [U-$^{13}C_6$]-glucose into bacteria. The amino acids synthesis in bacteria achieved >50% enrichment of ^{13}C labeling newly synthesized amino acids. In other words, in organisms, which can synthesize essential and nonessential amino acids from glucose and nitrogen, [U$^{13}C_6$]-glucose effectively replaces ^{12}C by ^{13}C in protein creating a heavy protein which can be separated by MS. The intensities of the labeled and unlabeled peaks were quantified to calculate synthesis/degradation ratio to reflect the relative dynamics of protein turnover (Cargile, Bundy et al. 2004). This method was a valid approach for studying simple model organisms such as bacteria, because they can synthesize both essential and non-essential amino acids. In multicellular organisms that can only synthesize non-essential amino acids, this approach has limited application (Cargile, Bundy et al. 2004). However, Vogt *et al.* developed a technique with ability to quantify the relative synthesis rate in complex organisms based on changes in average mass of the protein (Vogt, Schroer et al. 2003, Vogt, Hunzinger et

al. 2005). The average mass shift from the precursor amino acid (acyl-tRNA bound) can then be linked to the average mass shift of the labeled protein peptides and their fractional synthesis rate (Vogt, Hunzinger et al. 2005). This interrelation in mass shift provided an approach to assess precursor labeling via plasma apolipoproteins or secretory protein (close to 100% FSR) comparison with protein peptide of interest to assess the fractional synthesis of their parent protein (Vogt, Hunzinger et al. 2005).

Mass isotopomer distribution analysis (MIDA) is an approach that can be combined with the method by Vogt *et al.* for measuring precursor enrichment (Hellerstein and Neese 1999, Papageorgopoulos, Caldwell et al. 1999). This technique involves quantifying relative abundance, difference in mass shift and isotopomer distribution of molecular polymers (peptide) via a combinatorial probability model to compare measured abundances (^{13}C-99% labeled peptide) to theoretical (^{12}C unlabeled or natural ^{13}C abundance) distributions predicted from the binomial or multinomial expansion (Papageorgopoulos, Caldwell et al. 1999). MIDA was previously attempted by Papageorgopoulos *et al.* to determine protein synthesis *in vivo* by infusing rat with [5,5,5-$^{2}H_3$] leucine (99%) via the jugular catheter for 24h at a rate of 50 mg·kg^{-1}·h^{-1} using a mini-pump (Papageorgopoulos, Caldwell et al. 1999, Xiao, Garg et al. 2008). ESI-MS analysis determined incorporation of [5,5,5-$^{2}H_3$] leucine into protein peptides, which resulted in a mass shift of +3, +6 Da (Papageorgopoulos, Caldwell et al. 1999, Xiao, Garg et al. 2008)25-26]. However, low protein turnover rate and low amount of the +3 or +6 isotopomers found, described MIDA to be non-practical for in-*vivo* and low enrichment studies (Hellerstein and Neese 1999, Papageorgopoulos, Caldwell et al. 1999, Papageorgopoulos, Caldwell et al. 2002, Xiao, Garg et al. 2008); for combinatorial probabilities to be applicable, a labeled precursor must combine with itself in the form of two or more repeating subunits (Papageorgopoulos, Caldwell et al. 1999, Xiao, Garg et al. 2008).

In addition, Xiao et. al. proposed that the ratio of ^{13}C protein enrichment to ^{13}C enrichment of the labeling agent can determine the newly synthesized fraction, which can then be used to

calculate protein synthesis rate (PSR= [protein concentration] x [newly synthesized fraction]/ (unit time)) (Xiao, Garg et al. 2008).

Determining protein synthesis requires labeling on the precursor amino acids (acyl-tRNA bound) and the fraction of the specific proteins that was de novo synthesized during the labeling period (Zhao, Lee et al. 2009). However, ^{13}C labeling via [U-$^{13}C_6$]-glucose can generate complex labeling patterns that are difficult to link labeled pattern to its constituent amino acids (Zhao, Lee et al. 2009).

Furthermore, in the present study we present the mirror display data by using mass versus relative abundance since this was the best approach to average between the two samples. Hence, the same peptide sequence has an identical pattern in isotopically labelled mass distributions no matter the charge was. Therefore, we opt to combine the spectra independent of the charge and obtain as many isotopically labelled mass distribution.

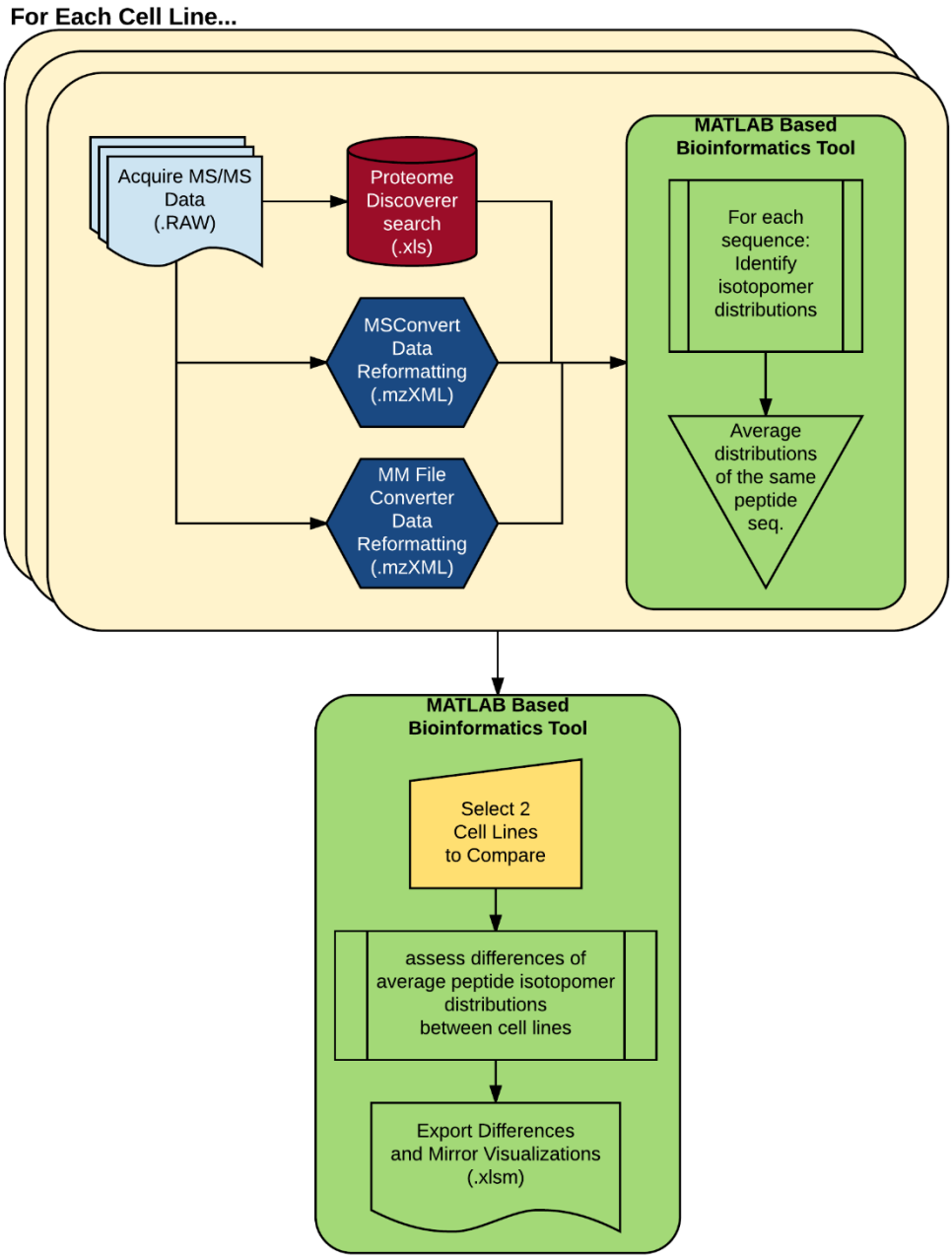
In summary, we have developed a bioinformatics tool to analyze the isotopomer distributions. The new methodology allows to:

- Generate mirror display of mass isotopomer patterns to demonstrate the different isotope distributions,
- Calculate chi-square distance demonstrating the difference in glucose utilization,
- Calculate of enrichment ratios which serve to quantify the glucose incorporation,

The tool was developed to process and analyze in large-scale metabolomics. Being an open-source software and developed in a popular analysis platform, this tool can be extended for further analysis of metabolites or other tracer-based metabolites.

Conclusion

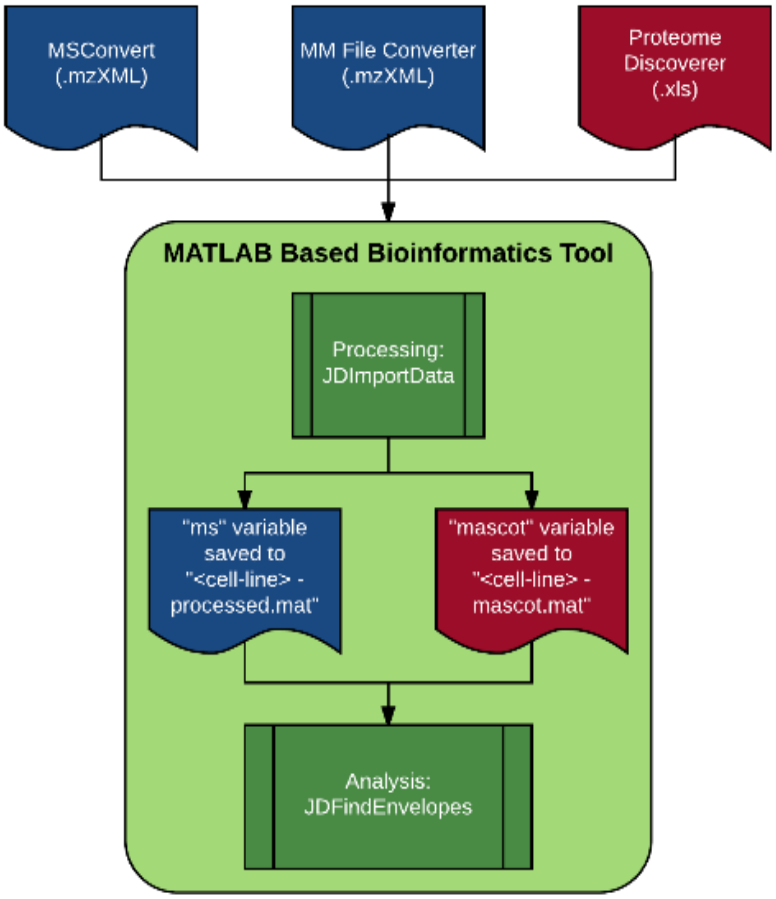
We have developed a bioinformatics tool based on MATLAB software to analyze tracer-based metabolomics data produced by LC-MS/MS and to study protein synthesis through ^{13}C stable-isotope labeling. Essentially, this new software tool can display mass isotopomer pattern differences and generate ranking of differences across proteins globally.



1

Figure 2: Overview of the data flow through the methodology

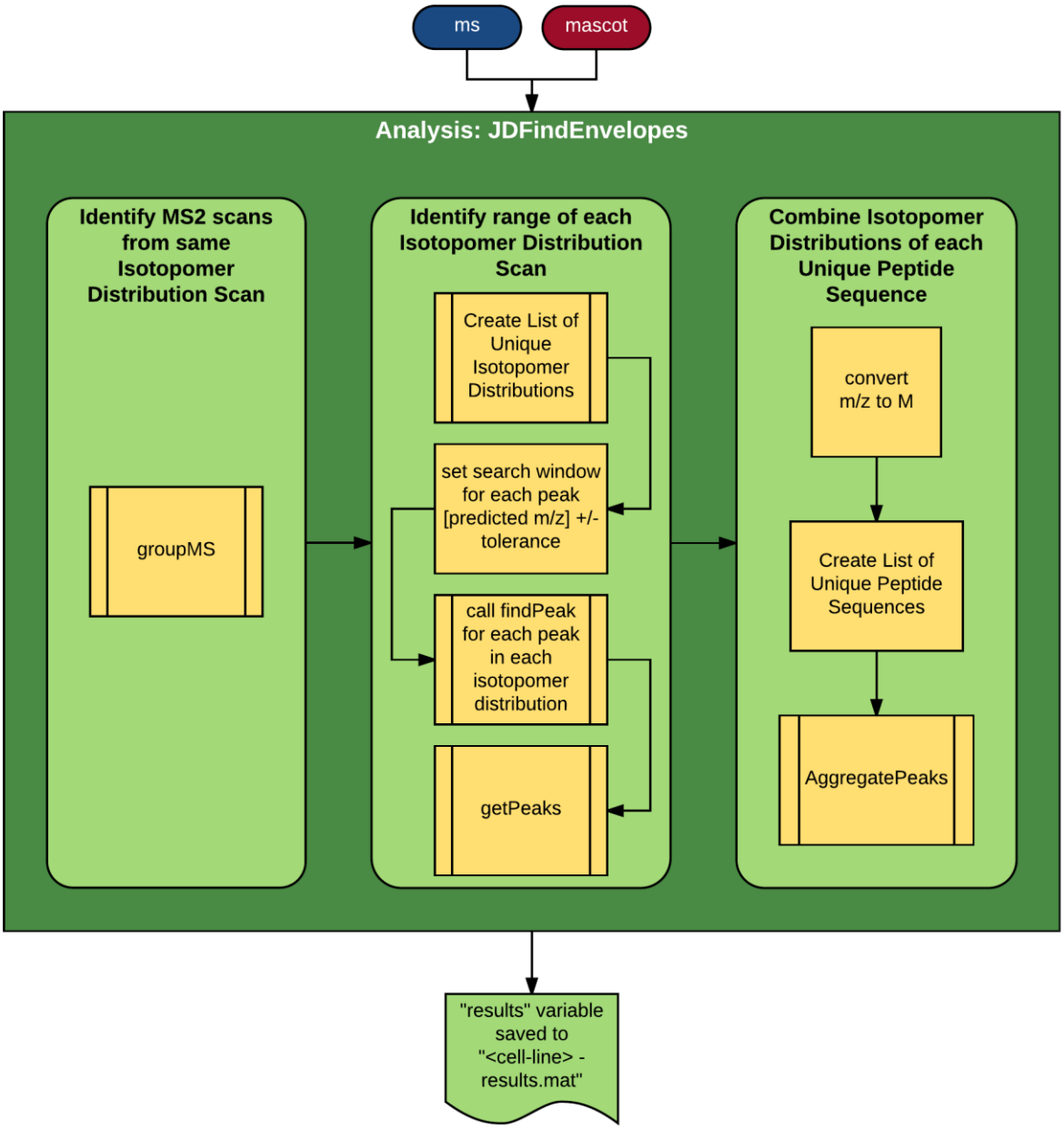
2



3

Figure 3: Detailed overview of the Pre-Analysis and Analysis Phases of the data analysis process

4



5

Figure 4: Detailed visualization of the tasks called by the JDFindEnvelopes script.

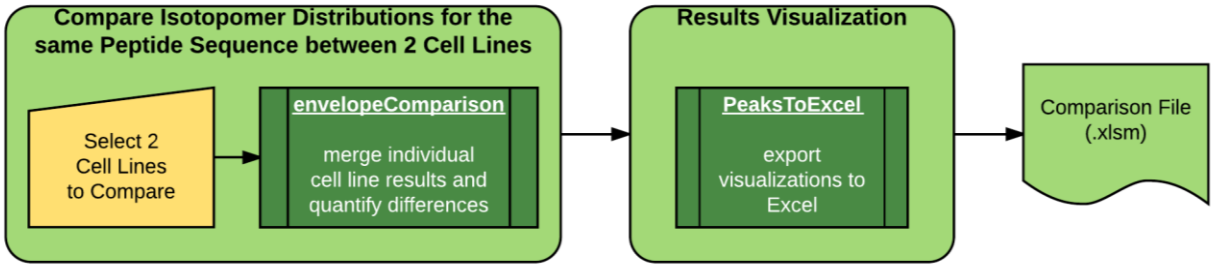


Figure 5: Flowchart of the results comparison, visualization and export module.

Chapter 2: Investigate the Metabolic Regulation of Protein Synthesis between Normal Pancreas Cells, Pancreatic Cancer Cells and Oral Cancer Stem Cells

Introduction

Cancer Stem Cells (CSCs)

Oral squamous cell carcinoma (OSCC) is a solid tumor that was reported to have the presence of CSCs. CSCs have the ability to undergo self-renewal and differentiation and these properties are also attributed to normal stem cells (Wicha, Liu et al. 2006). These properties are very important for stem cells, which allow them to regenerate and develop into different types of cells. However, in the context of CSCs, these properties are aberrant and uncontrolled. Thus, CSCs are predicted to show resistance to DNA damage, evasion of apoptosis, high tumorigenicity, low turnover and high invasion/migration resulting in resistance to drugs and radiation and recurrence. It is therefore believed that CSCs might play an important role in the metastatic process of OSCC and also in the resistance to chemotherapy and radiotherapy treatments (Prince, Sivanandan et al. 2007).

In our previous study, we compared the expression of two metabolic enzymes phosphoglycerate kinase 1 (PGK1) and hypoxia-inducible factor 1- alpha (HIF-1α) between CSCs and NSCCs. HIF-1α and PGK1 were found to be down-regulated in CSCs cancer cells compared to NSCCs. The protein plays an important role in cancer cell metabolism by activating the genes for glucose transporters and glycolytic enzymes resulting in an increased glycolytic rate (Vaupel 2004). PGK1 is an ATP-generating enzyme in the glycolytic pathway regulated by HIF-1α. Recent studies suggest that prostate cancer cell-derived PGK-1 induces osteoblastic differentiation of bone marrow stromal cells, affecting bone formation at the metastatic site. PGK-1 may also induce the differentiation of gastric cancer stem cells (Jung, Shiozawa et al. 2009, Zieker, Bühler et al.

2013). Since HIF-1α positively regulates glycolytic enzymes, including PGK1, it is not surprising that PGK-1 was down-regulated, the same as HIF-1α, in CSC-like UM1 cells.

Pancreatic Cancer

Pancreatic cancer (PC) is a malignant neoplasm of the pancreas. Each year in the US, more than 40,000 individuals are diagnosed with PC, and approximately 35,000 die from the disease (Boring, Squires et al. 1994, Jemal, Bray et al. 2011). Patients diagnosed with PC typically have poor prognosis mainly due to lack of symptoms early on, resulting in locally advanced or metastatic disease at time of diagnosis (Priante, Carvalho et al. 2010). The median survival from diagnosis is around 3 to 6 months, and the 5-year survival is less than 5% (Boring, Squires et al. 1994). PC is one of the few cancers for which survival rate has not improved substantially over the past 30 years (Lippman and Hong 2001, Park, Chiosea et al. 2010). In terms of its etiology, pathogenesis, detection, and treatment, PCs lags far behind the other forms of cancer. Therefore, the development of a new methodology to diagnosis and treatment is urgently needed.

Recent studies have identified oncogenes and tumor suppressor genes that play a role in the regulation of metabolism through their regulation of signaling pathways and enzyme expression (Feng and Levine , Levine and Puzio-Kuter 2010, Zheng 2012, Faubert, Boily et al. , Vaitheesvaran, Xu et al. 2014). A study done in pancreatic cancer MIA PaCa-2 cells revealed that inhibition of glycogenolysis by glycogen phosphorylase inhibitor resulted in significant reduction in both oxidative and non-oxidative pentose cycle activities, decreased pyruvate dehydrogenase activity and de novo lipogenesis (Lee, Guo et al. 2004). Similarly, it was observed in other studies that metabolic changes in pentose cycle and de novo lipogenesis by inhibiting oxamate, a lactic acid dehydrogenase (LDHA) or inhibition of Bcr-ABL tyrosine kinase in human leukemia cells can have an effect in hexokinase, G6PDH, and transketolase (TKT) metabolic pathways, which are dominant in Warburg effect (Boren, Cascante et al. 2001).

In Chapter 2, our developed methodology will be used to distinguish the mass isotopomer patterns between pancreatic cancer cells and normal pancreas epithelial cells and between CSCs and NSCCs. Therefore, human pancreatic duct epithelial (HPDE) cells, MIA-PaCa-2, CSCs and NSCCs will be used for comparative analysis. HPDE cell line is an immortalized epithelial cell line derived from normal pancreatic duct epithelial cells, MIA-PaCa-2 is derived from pancreas adenocarcinoma of a 65-year-old man and involved the body and tail of the pancreas (Ouyang, Mou et al. 2000, Deer, Gonzalez-Hernandez et al. 2010) CSCs and NSCCs are subset of oral cancer cells, as reported in our previous studies (Misuno, Liu et al. 2013). The objective is to reveal the differences in mass isotopomer distributions for peptides (derived from newly synthesized proteins) between HPDE and MIA-PaCa-2 cells and between CSCs and NSCCs.

Materials and Methods

Cell Culture and Sphere Formation

The cell cultures were divided into two groups:

Group #1 (unlabeled ^{12}C): UM1 cells were cultured in Dulbecco's modified eagle medium (DMEM with 10% fetal bovine serum and 1% penicillin/streptomycin (Invitrogen, Carlsbad, CA, USA) and maintained at 37°C in a humidified atmosphere of 5% CO_2. When the UM1 cell line reached confluence of approximately 90%, the culture medium was changed to serum-free medium (DMEM) containing 1% penicillin/streptomycin and growth factors bFGF (10 ng/mL) and EGF (10 ng/mL) (Geminni Bio, Sacramento, CA, USA). The cells was maintained at 37°C in a humidified atmosphere of 5% CO_2 for about 3 weeks. The spherogenic and non-spherogenic UM1 cancer cells were harvested through trypsinization, as described earlier in our published study (Misuno, Liu et al. 2013).

MIA PaCa-2 were obtained from the American Tissue Culture Collection (ATTCC, Rockville, MD) and cultured in glucose free DMEM (GIBCO), 10% fetal bovine serum, 100 units/mL penicillin and 100µg/mL streptomycin. Immortalized HPDE cell line were cultured with glucose free, keratinocyte serum-free media, epidermal growth factor, bovine pituitary extract (Life Technologies, Inc., Grand Island, NY), 100units/mL penicillin and 100µg/mL streptomycin. All cells were maintained at 5% CO_2-95% O_2 and 37°C and washed with PBS prior to cell lysis with Radio-Immunoprecipitation Assay (RIPA) buffer containing 1% Nonidet P-40, 0.1% SDS, 50mM Tris-HCL (pH 7.4), 150mM NaCl, 0.5% Sodium Deoxycholate and 1mM EDTA.

Cell lysates containing 25ug of total proteins were treated with 100mM dithiothreitol (DTT) for 1hr to reduce protein disulfide bonds, and then treated with 150 mM iodoacetic acid (IAA) alkylating agent in 200mM NH_4HCO_3 for 1hr in the dark to prevent re-formation of disulfide bonds. Protein samples were digested with 10 ng/µL enzyme-grade trypsin (Promega, Madison, WI) in 200 mM NH_4HCO_3 at 37°C incubation for 24 hours. After proteolytic cleavage, samples were spun

down at 14000g. The supernatant was transferred and dried. The peptide samples were dissolved in 0.1% formic acid and transferred to sample injection vials. Each experiment was in triplicate.

Group #2 (labeled ^{13}C): UM1s were cultured in Dulbecco's modified eagle medium (DMEM with 10% fetal bovine serum and 1% penicillin/streptomycin (Invitrogen, Carlsbad, CA, USA) and maintained at 37°C in a humidified atmosphere of 5% CO_2. When UM1 cell line reached confluence of approximately 90%, the culture medium was changed to serum-free medium (DMEM) containing 1% penicillin/streptomycin and growth factors bFGF (10 ng/mL) and EGF (10 ng/mL) (Geminni Bio, Sacramento, CA, USA). After ~ 3 weeks, when UM1 form sphere structures the culture medium was changed to serum-free medium supplemented with 4.5g/L [U-$^{13}C_6$, 99%] D-glucose (Cambridge Isotope Laboratories, Inc.) and maintained at 37°C in a humidified atmosphere of 5% CO_2. The cells were treated with ^{13}C medium for 3 days. The spherogenic and non-spherogenic UM1 cancer cells were then harvested using trypsinization.

MIA PaCa-2 were cultured in glucose free DMEM (GIBCO) supplemented with 4.5g/L [U-$^{13}C_6$, 99%] D-glucose (Cambridge Isotope Laboratories, Inc.), 10% fetal bovine serum, 100 units/mL penicillin and 100µg/mL streptomycin. Immortalized HPDE cell line were cultured with glucose free, keratinocyte serum-free media supplemented with [U$^{13}C_6$]-glucose, epidermal growth factor, bovine pituitary extract (Life Technologies, Inc., Grand Island, NY), 100units/mL penicillin and 100µg/mL streptomycin. All cells were maintained at 5% CO_2-95% O_2 and 37°C and washed with PBS prior to cell lysis with RIPA buffer containing 1% Nonidet P-40, 0.1% SDS, 50mM Tris-HCL (pH 7.4), 150mM NaCl, 0.5% Sodium Deoxycholate and 1mM EDTA.

Cell lysates containing 25ug of total proteins were treated with 100mM dithiothreitol (DTT) for 1hr to reduce protein disulfide bonds, and then treated with 150 mM iodoacetic acid (IAA) alkylating agent in 200mM NH_4HCO_3 for 1hr in the dark to prevent re-formation of disulfide bonds. Protein samples were digested with 10 ng/µL enzyme-grade trypsin (Promega, Madison, WI) in 200 mM NH_4HCO_3 at 37°C incubation for 24 hours. After proteolytic cleavage, samples were spun down at 14000g. The supernatant was transferred and dried. The peptide samples were dissolved

in 0.1% formic acid and transferred to sample injection vials. Each experiment was conducted in triplicate.

Liquid Chromatography with Tandem Mass Spectroscopy.

The samples were fractioned and each fraction was desalted using the PepClean C18 spin columns (Pierce, Rockford, IL, USA) and further analyzed by RPLC-MS/MS on the Orbitrap Velos instrument. The peptide samples were loaded onto a C18 trap column at a flow rate of 30 mL/min. Then, the peptides were separated on a C18 nano column (75 mm 150 mm, 3 mm, 100 A˚ , C18; Dionex, Sunnyvale, CA, USA) at a flow rate of 250 nL/min on an UltiMate 3000 LC system (Dionex) using 3 h gradient. The MS/MS spectra were acquired on the Orbitrap Velos mass spectrometer using a data-dependent analysis mode, which the top 10 most abundant ions in each MS scan (m/z 350–2000) were selected for MS/MS analysis. The capillary temperature was set to 275C and the spray voltage set to 2 kV. The resolution for MS scan is 60,000 full width at half maximum (FWHM) and 7500 FWHM for MS/MS scan. The lock mass was used for accurate mass measurement. The MS/MS spectra were searched against the UniRef100 human database (120,982 entries) using the Mascot Daemon (Version 2.3). The mass tolerance was 10 ppm for MS and 0.1 Da for MS/MS. Database search was performed using Mascot database and Proteome Discoverer search engine against UniProt. We re-acquired MS/MS spectra of CSC, NSCC, HPDE and MIA PaCa-2 cell lines with “Exclude isotopes” turned off in the MS method on Q Exactive instrument.

Database Searching

Database search was performed using the Mascot (Matrix Science, London, UK; version 2.4.1) and Proteome Discoverer search engine against the UniProt database. Mascot was searched with a fragment ion mass tolerance of 0.100 Da and a parent ion tolerance of 10.0 PPM. Oxidation of methionine, acetyl of the n-terminus and carbamidomethyl of cysteine were specified in the Mascot database search as variable modifications.

Criteria for protein identification -- Scaffold (version Scaffold_4.4.8, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability by the Peptide Prophet algorithm (Keller, Nesvizhskii et al. 2002) with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii, Keller et al. 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

Data Processing

The RAW files were converted to the mzXML format and processed with the Proteome Discoverer. The mzXML files were generated by using both msconvert and MM File Converter. The two converters were used to obtain the MS2 scan number, precursor (MS1) scan number, charge (z), precursor (MS1) m/z as well as the MS1 and MS2 spectra. The Proteome Discoverer was used to perform the Mascot database search to acquire peptide sequence, accession number, protein name, charge, peptide spectrum matches (PSM) ambiguity and MS2 scan number. Next, these parameters were all imported into the MATLAB and processed with our developed bioinformatics tool to identify the average isotopomer distribution for each unique peptide sequence observed. See chapter 1 for further details on this methodology. Overall, the tool first identifies individual isotopomer distribution scans and then it proceeds to combine the isotopomer distributions belonging to the same peptide sequence into an average profile.

Results

Identification and Analysis of Isotopomer Distributions in CSC, NSCC, HPDE and MIAPaCa-2 cells

To investigate the utilization of glucose for protein synthesis, we introduced stable isotope-labeled glucose, [U-$^{13}C_6$] glucose, in the culture medium of CSCs, NSCCs, HPDE and MIA PaCa-2 cells and compare with unlabeled (without ^{13}C glucose) CSCs, NSCCs, HPDE and MIA PaCa-2 cells. Cellular proteins were harvested and analyzed by LC-MS/MS. Peptide sequence and protein identification were conducted by searching the obtained MS data against the UniProt database with the Proteome Discoverer searching engine. In total, we identified 7,517 unique peptides in CSCs, 6,555 unique peptides in NSCCs, 5,198 unique peptides in HPDE and 4,324 unique peptides in MIA-PaCa-2 for labelled cell lines, whereas for the unlabeled cell lines, we identified 12,218 unique peptides in CSCs, 13,338 unique peptides in NSCCs, 9,355 unique peptides in HPDE cells and 10,720 unique peptides in MIA PaCa-2 cells (**Figure 6**). The reduction number observed in the identification of unique peptides in ^{13}C labeled could be due to the complexity of ^{13}C labeled isotopic peak patterns since ^{13}C labeling adds 1-Da mass *versus* the ^{12}C (no labeling) and this may cause a confusion in the MS instrument for the selection of monoisotopic peak for fragmentation and protein identification.

Identification and analysis of isotopomer distribution pattern from MS1 were processed by our newly developed bioinformatics tool. By matching the MS2 scans to the monoisotopic peak (M0, ^{12}C) and the 2nd isotopic peak of the labeled peptide (displaying mass shift of +1 Da; M0, M1, M2, M3…), we identified the isotopomer distributions. The number of peaks per isotopomer distribution in CSC, NSCC, HPDE and MIA PaCa-2 cells is shown in **Figure 7**. The histograms show that there is an increase in glucose incorporation since we observed isotopomer distributions with more than 4 peaks. The total number of isotopomer distribution in CSCs is 9,403 of which 84.7% are isotopomer distributions with more than 4 peaks and the total number in NSCCs is 8,273 of which 86.2% are isotopomer distributions with more than 4 peaks. In HPDE cells, the total number is 6,867 with 90.3% > 4 peaks and in MIA PaCa-2 cells, the total number is 5,521 with 89.4% >4 peaks.

In order to analyze the differences between the samples, we created a table containing peptide sequence, accession number, protein descriptions, number of peaks, abundance difference and chi-square distance calculations to quantify the differences for comparison between each cell phenotypes and the chi-square distances were normalized by the number of peaks enclosing in an isotopomer distribution. **Table 1** and **Table 2** show the top 5% chi-square distance normalized. In addition, a mirror display was built from the excel table and it shows a plot with comparison of the isotopomer distributions of two cellular phenotypes. The following figures are the mirror display with description of normalized chi-square distance and sum abundance difference between CSC and NSCC or between HPDE cells and MIA PaCa-2 cells.

In our study, we observed different expression of ^{13}C labeled proteins including but not limited to integrin alpha and beta, tumor rejection antigen (Gp96), serpin B5 and phosphoglycerate kinase 2 between CSCs and NSCCs. Note: chi-square distance is a metric calculation of the difference between the isotopomer distributions; therefore, if the number is closest to “zero” they are similar and the more different the more it is closest to “one”. In the sum abundance difference negative numbers are with the intent to differentiate between the two samples since they are mirror displays for example negative for NSCC and MIA PaCa-2 cell lines. **Figure 8** integrin α-6 isoform shows a normalized chi-square distance of 0.488 and the sum abundance difference of -2.383. **Figure 9** demonstrates an isoform β- 4D of integrin β-4 which shows a normalized chi-square 0.115 and the abundance sum difference -1.154. **Figure 10** is a tumor rejection antigen (Gp96) with normalized chi-square distance 0.302 and abundance sum difference -3.987. In **Figure 11** serpin B5 with normalized chi-square distance 0.128 and abundance sum difference -1.798. Therefore, we observed different expression of glucose ^{13}C labeled in integrin α-6 isoform, isoform β- 4D of integrin β-4, Gp96 and serpin B5 suggesting that NSCCs utilized more glucose in order to synthesize for integrin α-6 isoform, isoform β- 4D of integrin β-4, Gp96 and serpin B5 in NSCCs compared with CSCs.

Figure 12-Figure *17* exhibit a mirror display of peptide isotopomers from several proteins between HPDE and MIAPaCA-2 cells. In **Figure 12,** desmoplakin shows normalized chi-square 0.571 and abundance sum difference 4.892, whereas, in **Figure 13,** hexokinase-1 (isoform 4) shows normalized chi-square 0.145 and abundance sum difference 1.897. **Figure 14** indicates a normalized chi-square distance 0.108 and abundance sum difference 0.913 for transketolase. This suggests that HPDE cells utilized more glucose in order to synthesize desmoplakin, hexokinase-1 and transketolase compare with MIA PaCa-2 cells.

In addition, the results for heat shock protein (HSP) 90, peptidyl-prolyl cis-trans-isomerase and glutamate dehydrogenase (GLDH) seem to indicate that there is more utilization of glucose in MIAPaCa-2 cells compare to HPDE cells. As shown in **Figure 15,** HSP 90-beta protein shows normalized chi-square 0.472 and abundance sum difference -0.066. Meanwhile, peptidyl-prolyl cis-trans- isomerase was found to have normalized chi-square 0.228 and abundance sum difference -3.192 (**Figure 16**). **Figure 17** shows GLDH with normalized chi-square 0.279 and abundance sum difference -3.927.

In NSCCs, we observed elevated expression of ^{13}C in Keratin 1, type I cytoskeletal 9, 14, 16, 17 and 18 and type II cytoskeletal 5 (**Table 4**). We also observed high expression ^{13}C in CSC Keratin 1, type I cytoskeletal 9, 10, 14, 16, 17, and 18 and type II cytoskeletal 2 epidermal (**Table 3**). In HPDE type I cytoskeletal 9, 14, 16 and 17 and type II cytoskeletal 5, 8, 6A, 6B and 3 (**Table 5**). In MIA PaCa-2 Keratin 1, type I cytoskeletal 1, 9, 10, 13, 17 and 19, type II 5, 8 and cytoskeletal 2 epidermal 6A (**Table 6**). Therefore, CSC, NSCC, HPDE cells and MIA PaCa-2 cells utilized glucose to synthesize different types of keratins.

Analysis of Enrichment Ratios in CSC/NSCC and HPDE/MIAPaCa-2 cells

The enrichment of ^{13}C in a sample was measured by calculating the ratio of partial incorporation of ^{13}C and the natural ^{12}C. In total, the number of enrichment ratios were 4,533 in CSCs, 3,896 in NSCCs, 1,917 in HPDE cells and 1,637 in MIA PaCa-2 cells (**Figure 18**). The

negative values of $^{13}C/^{12}C$ represent that the ^{12}C peak area is greater than the ^{13}C peak area which could indicate no ^{13}C incorporation. We also quantify the number of peaks with ^{12}C and ^{13}C and all cell lines—CSC, NSCC, HPDE and MIA PaCa-2—show a rapid decay in approximately 9 peaks for unlabeled peptides and for labeled peptides they show a slow decay around 20 or more peaks. This implies that extensive decay has an incorporation of ^{13}C glucose (**Figure 19*Figure 20***). **Figure 21** presents the total number of peptide sequences in CSCs (15,140), NSCCs (20,067), HPDE cells (16,357), MIA PaCa-2 cells (17,590) versus the total number of peptide sequences with enrichment ratio in CSCs with 4,533, NSCCs 3,896, HPDE cells 1,917 and MIA PaCa-2 cells 1,637.

The isotopomer distribution of $^{12}C/^{13}C$ for peptide sequence LIATFPDTLTYSAYR derived from integrin-alpha in CSCs demonstrates that there is no glucose enrichment since there is 4 peaks in unlabeled and labeled ^{13}C, the enrichment ratio is low 0.044 and abundance sum difference is negative 0.264. However, for the same sequence in NSCCs, the unlabeled peptides contain 5 peaks and the labeled peptides show 9 peaks with an enrichment ratio of 0.113 and abundance sum difference of 1.920 (**Figure 22 A & B**). **Figure 22 C & D** show an isotopomer pattern of a peptide TQDYPSVPTLVR (derived from integrin beta) in CSCs with rapid decay of 3 peaks in unlabeled peptide and labeled peptide with slow decay of 8 peaks, the enrichment ratio was 0.175 and abundance sum difference was 1.156. Similarly, in NSCCs, unlabeled peptides display rapid decay of 3 peaks and labeled display slow decay of 9 peaks with enrichment ratio of 0.140 and abundance sum difference of 2.620. This implies that synthesis of integrin-alpha and -beta have more glucose enrichment in NSCCs than CSCs.

Figure 22E illustrates the isotopomers pattern for a peptide ELISNASDALDKIR derived from tumor rejection antigen (GP96) in CSCs. The unlabeled peptide displays rapid decay of 5 peaks and labeled peptide displays extensive decay of 15 peaks with an enrichment ratio 0.138 and abundance sum difference 2.480. While in NSCCs (**Figure 22F**), the unlabeled peptides show slow decay of 6 peaks and labeled peptides display extensive decay of 15 peaks with an

enrichment ratio of 0.136 and abundance sum difference of 5.974 indicating more glucose enrichment for synthesis of the protein in NSCCs than CSCs.

In **Figure 23** HSP90 peptide sequence DQVANSAFVER in MIAPaCa-2 cells demonstrates a rapid decay in unlabeled peptides of 4 peaks and slow decay for labeled 10 peaks, enrichment ratio 0.195 and abundance sum difference 5.858. Moreover, isotopomer pattern is shown in **Figure 24**, NAD-dependent malic enzyme peptide sequence HISDSVFLEAAK in MIA PaCa-2 cells with unlabeled peptides presenting 2 peaks and a slow decay of 17 peaks in labeled peptides, enrichment ratio 0.458 and abundance sum difference 5.799. Proteins HSP90 and NAD-dependent malic enzyme were not found in HPDE cells which indicates that there is more glucose enrichment for synthesis of protein in MIA PaCa-2 cells than HPDE cells.

In **Figure 25** peptidyl-prolyl cis-trans isomerase peptide sequence EGmNIVEAmER in HPDE cells displays a rapid decay rapid decay in unlabeled peptide of 3 peaks while labeled ^{13}C presents a slow decay 9 peaks, enrichment ratio 0.202 and abundance sum difference 1.912. No enrichment ratio and abundance sum difference for peptidyl-prolyl cis-trans in MIA PaCa-2 cells which indicates more glucose enrichment in HPDE cells compare with MIA PaCa-2 cells.

Discussion

The key to understanding the regulation of protein expression in cells, which is determined by protein synthesis and degradation, is to study the dynamics of protein turnover; in other words, dynamics of protein turnover reflects the regulation of protein expression in cell division and cell function. With tracers, we can now investigate how protein synthesis depends on the precursor/product relationship (Zhao, Lee et al. 2009).

Recently, it has been reported metabolic enzymes play an important role in carcinogenesis and this leading to the design of potential drugs targeting glycolysis and other metabolic processes (Boros, Cascante et al. 2002, Cairns, Harris et al. 2011). In order to understand how nutrition (e.g., glucose) is used for biosynthesis of metabolic enzymes, stable isotopes can be used to label the non-essential amino acids and study the protein synthesis in cancer cells. The mass difference of ^{13}C labeled peptide can then be readily detected by MS as a distribution of mass isotopomers. This distribution patterns can reflect respective metabolic fluxes, because the amino acid isotopomers carry information regarding the paths they have traversed, and the final distribution can reflect the relative contribution from these individual pathways (Lee, Wahjudi et al. 2010). Quantification of the isotopic peak intensities from the labeled and the unlabeled peaks can generate a synthesis-to-degradation ratio to represent the relative dynamics in protein turnover (Xiao, Garg et al. 2008). In addition, comparing the ^{13}C labeled and ^{12}C unlabeled ion fragments can identify peptide amino acid sequences to depict carbon flow from glucose, and to possibly map out protein synthesis pathways and alterations in nutrient-gene interactions between cancer and normal cells. Therefore, we developed a novel tool using MATLAB to assist us in the identification of the isotopomer distributions differences, and to create a visual tool to envision the enrichment ratio between CSCs/NSCCs and HPDE/MIA PaCa-2 cells.

We observed in our study that the synthesis of integrin α6 and integrin β4, PG96 and serpin B5 requires more glucose (^{13}C uptake) in NSCCs compared to CSCs. Integrins are one of

the major families of heterodimeric cell adhesion receptors containing an α and β subunit, which are important in regulating cellular responses to control cell adhesion, migration, proliferation, survival, and differentiation (Humphries, Byron et al. 2006). A study by Dyce *et.al.* comparing the invasiveness of UM-SCC-1 and JHU-022-SCC HNSCC cell lines advocates that differential integrin expression may be related with invasion behavior of the tumor. Moreover, the study also advocates that high expression of α6 and β4 subunits and presumably, the α6β4 receptor was associated with a less invasive squamous cell carcinoma phenotype (Dyce, Ziober et al. 2002). In our study, we found that the synthesis of integrin α6 and integrin β4 requires high glucose uptake in NSCCs compared to CSCs. Thus, integrin α6 and β4 could be related with the invasion behavior between CSCs and NSCCs.

Serpin B5 (Maspin) belongs to the serpin family of protease inhibitors and it has been suggested to have a tumor-suppressor function. In OSCC cell lines, Yoshizawa *et al.* demonstrated that positive maspin expression had a better prognosis and lower expression stronger invasive potential than other grades of OSCC since it acts as a tumor- suppressor. The authors suggested that Maspin may be a useful marker to identify the potential for progression in OSCC (Yoshizawa, Nozaki et al. 2009). In our study, the synthesis of serpin B5 requires more glucose utilization in NSCCs than CSCs, suggesting that serpin B5 may act as tumor-suppressor function in NSCCs.

The synthesis of desmoplakin, hexokinase and transketolase were found to require more glucose utilized in HPDE cells than MIAPaCa-2 cells. On the other hand, HSP90, peptidyl-prolyl-cis-trans and GLDH were found to require more ^{13}C uptake in MIAPaCa-2 cells than HPDE cells. Desmoplakin is a component of desmosome structures in cardiac muscle and epidermal cells, which function to maintain the structural integrity at adjacent cell contacts. Decreased desmosomal protein expression was found in breast cancer (Oshiro, Kim et al. 2005), oropharyngeal squamous cell carcinoma (Papagerakis, Shabana et al. 2009), cervical carcinoma (Alazawi, Morris et al. 2003), colorectal cancer (Cui, Chen et al. 2011) and pancreatic cancer

(Hamidov, Altendorf-Hofmann et al. 2011, Yang, Chen et al. 2012). Hexokinase is an enzyme that phosphorylates hexoses (six-carbon sugars), forming hexose phosphate. In most organisms, glucose is the most important substrate of hexokinases, and glucose-6-phosphate is the most important product (Kowalczyk and Green 2013). Transketolase is an enzyme that plays an important role in pentose phosphate pathway (Wang, Zhang et al. 2013). Studies have shown that inhibition of transketolase will suppress the pentose phosphate pathway and interrupt the synthesis of important coenzymes ATP, CoA, NAD+, FAD and genetic material, RNA and DNA in cancer cells (Rodríguez-Enríquez, Vital-González et al. 2006).

Peptidyl-prolyl isomerases (PPIs), also known as cyclophilin A (CypA), belong to a group of proteins that possess peptidyl prolyl isomerase activity and catalyse the cis-trans conversion of proline peptide bonds. Members of the cyclophilin family are overexpressed in a range of human malignancies including hepatocellular cancer, pancreatic cancer, non-small cell lung cancer, gastric cancer, colorectal cancer and glioblastoma multiforme (Li, Wang et al. , Paul and Margaret 2016). Li *et al.* demonstrated that in both pancreatic cancer cell lines and tissues the expression of CypA and CD147 was significantly increased. They also claimed that exogenous CypA promotes pancreatic cancer cell growth, which may be mediated through the interaction with CD147 and the activation of ERK1/2 and p38 MAPKs (Li, Zhai et al. 2006). Thus, this is congruent with our finding that synthesis of cyclophilin A requires high glucose uptake in MIA PaCa-2 compare with HPDE cells.

A protein HSP 90-beta was found to have high utilized glucose ^{13}C in MIA PaCa-2 cells. Hsp90 is a chaperon protein that assists other proteins to fold properly, stabilizes proteins against heat stress, and aids in protein degradation. In addition, it appears to be utilized in carcinogenesis in order for cells to escape the pathways of tumor suppression, to promote progression in more advanced stage, to become treatment-resistant, and to facilitate metastasis (Calderwood 2010). In breast neoplasia, it interacts with a variety of proteins; including estrogen receptors (ER), tumor suppressor p53 protein, angiogenesis transcription factor HIF-1 alpha,

antiapoptotic kinase Akt, Raf-1 MAP kinase and a variety of receptor tyrosine kinases (Beliakoff and Whitesell 2004). Therefore, HSP-90 may play an important role in promoting tumor progression in MIA PaCa-2 cells.

Our study demonstrates that GLDH has a high glucose utilization in MIA PaCa-2 cells, compares with HPDE cells. GLDH is an enzyme that converts glutamate to α-ketoglutarate (α-KG). Thus, conversion of glutamate to α-ketoglutarate occurs either through oxidative deamination by GLDH in the mitochondrion or by transamination to produce nonessential amino acids in either the cytosol or the mitochondrion. During glucose metabolism, the transamination pathway predominates (Yang, Sudderth et al. 2009). When glucose is scarce, GDH becomes a major pathway to supply glutamine carbon to the TCA cycle, and is required for cell survival (Yang, Sudderth et al. 2009, Choo, Kim et al. 2010). Methods to image or otherwise quantify glutamine metabolism in vivo would be useful in this regard (Rajagopalan and DeBerardinis 2011). Infusions of pre-surgical subjects with isotopically labeled glutamine, followed by extraction of metabolites from the tumor and analysis of ^{13}C enrichment, can be used to detect both glutamine uptake and catabolism (Marin-Valencia, Yang et al. 2012, Yuneva, Fan et al. 2012).

Metastatic disease is the major cause of cancer mortality and there is a need to unravel the mechanisms that govern cancer cell spread from a primary tumor. Cytoskeleton, a structure essential for motility, is an example of metastatic behavior involving a vascular invasion, during which cancer cells bind to, separate, and then migrate through intact endothelial layers (Byers and Bhawan 1998, Hicks and Flaitz 2000, Safadi, Bader et al.). Therefore, in cancers, keratins have been recognized to regulate many cellular properties and functions, including apico-basal polarization, motility, cell size, protein synthesis, membrane trafficking, and membrane signaling. In addition, several studies have provided evidence for active keratin involvement in cancer cell invasion and metastasis, as well as in treatment responsiveness, and have set the foundation for further exploration of the role of keratins as multifunctional regulators of epithelial tumorigenesis (Moll, Divo et al. 2008, Karantza 2011, Kanojia, Sawant et al. 2012, Berens, Sharif et al. 2016,

Misiorek, Lähdeniemi et al. 2016, Safadi, Bader et al.). In a study, expression of the stratified epithelial keratins K5, K14 and K17 and the hyperproliferative keratinocyte-type keratins K6 and K16 are found in squamous cell carcinomas, independently of their site of origin (Wetzels, Schaafsma et al. 1992, Takahashi, Shikata et al. 1995, Luo, Kong et al. 2003, Cohen-Kerem, Rahat et al. 2004, Kitamura, Toyoshima et al. 2012). Moreover, keratins are being used as diagnostic markers in tumor pathology, such as in cases where is unclear on the basis of clinical presentation and conventional histopathology, including carcinomas that are poorly differentiated or metastasized over several organs and metastases of unknown primary tumor site, keratin typing is especially valuable for correct tumor type identification and subsequent selection of the most appropriate treatment plan. In additional, there is evidence using keratins as prognostic markers more interestingly, as active regulators of epithelial tumorigenesis and treatment responsiveness. Thus, we need further investigation of the role of keratins in cancer and the metabolic regulation of keratin synthesis and this will lead to possible therapeutic targets and markers for early cancer detection (Crowe, Milo et al. 1999, Karantza 2011). In our study, we observed high glucose utilization for the synthesis of keratins between CSCs/NSCCs and between HPDE/MIA PaCa-2 cells which may be use as prognostic marker.

Conclusion

Based on the tracer-based metabolomic analysis, we could identify a number of proteins that show significant differences in utilization of glucose for their biosynthesis between CSCs and NSCCs and between HPDE and MIA PaCa-2 cells. To the best of our knowledge, this is the first study of cancer cells with tracer-based metabolomics at the amino acid level. By analyzing the isotopomer patterns it allows us to calculate the differences in protein synthesis using chi-square distance and enrichment ratio for the quantification of ^{13}C glucose utilization for protein synthesis. We also identified unique ^{13}C labeling signatures distinctive to cancer cells. Our next step is to apply the novel methodology to find distinct signatures between high and low invasive oral cancer cells.

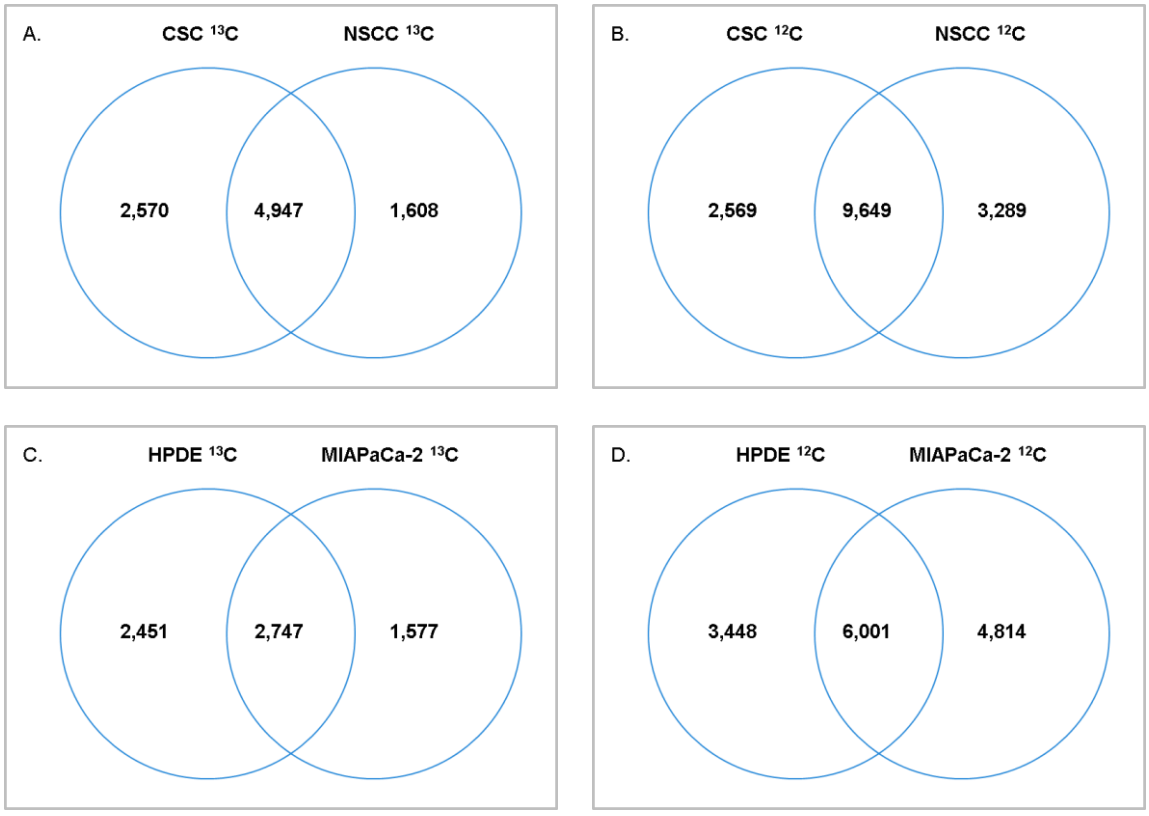
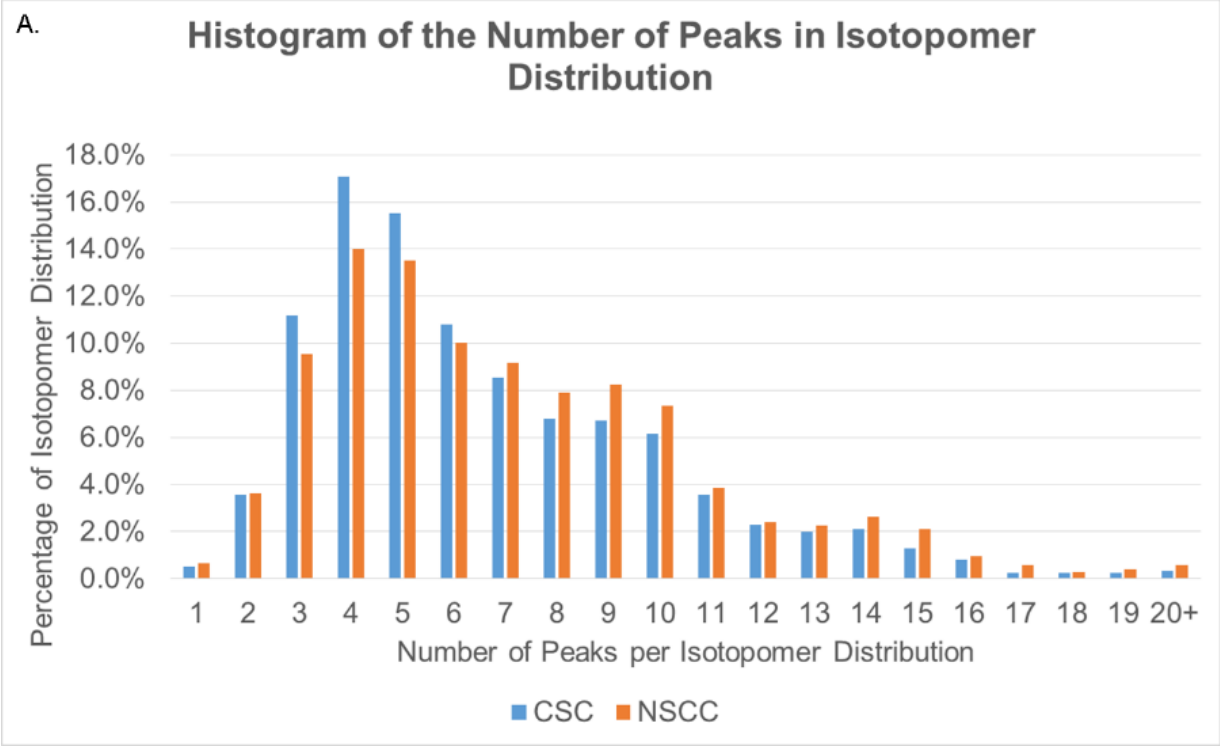


Figure 6A-D: The Venn diagrams show the overlap of total unique peptides found in ^{13}C and ^{12}C on CSCs/NSCCs and HPDE/MIAPaCa-2 cells provided by scaffold software.



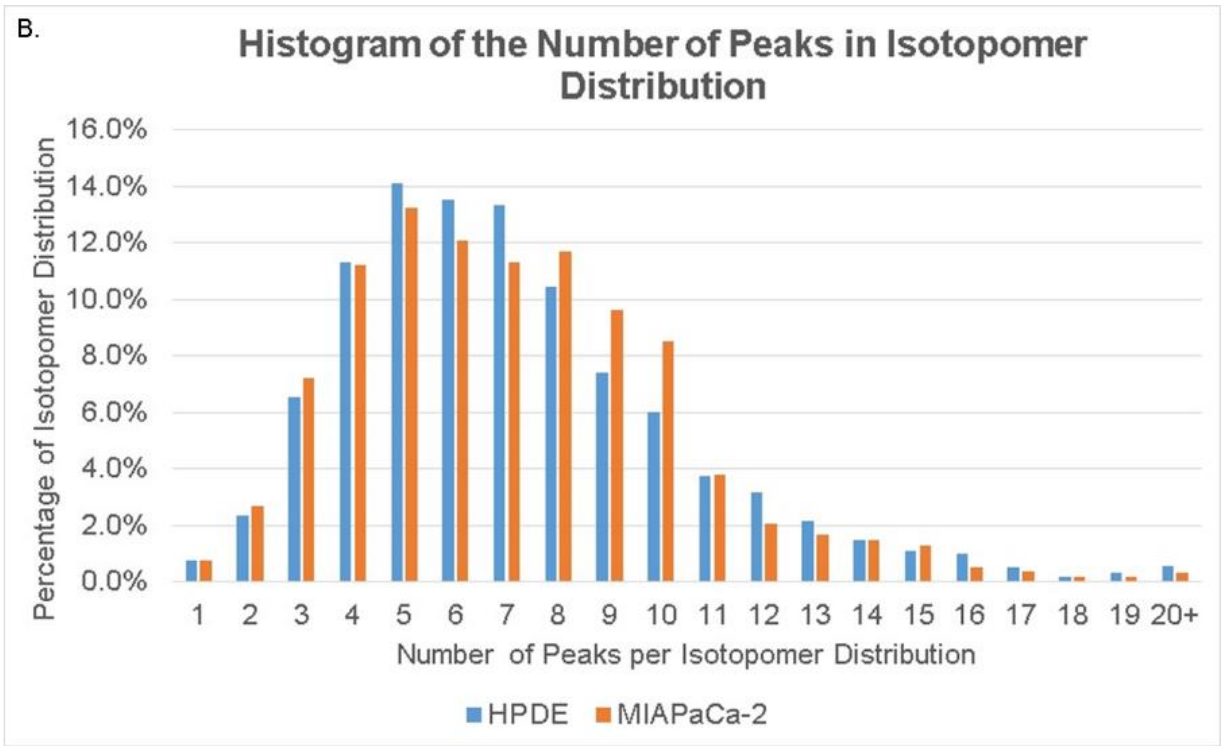


Figure 7A & B: Total number of peaks per isotopomer distribution between CSCs/NSCCs and HPDE/MIA PaCa-2 cells.

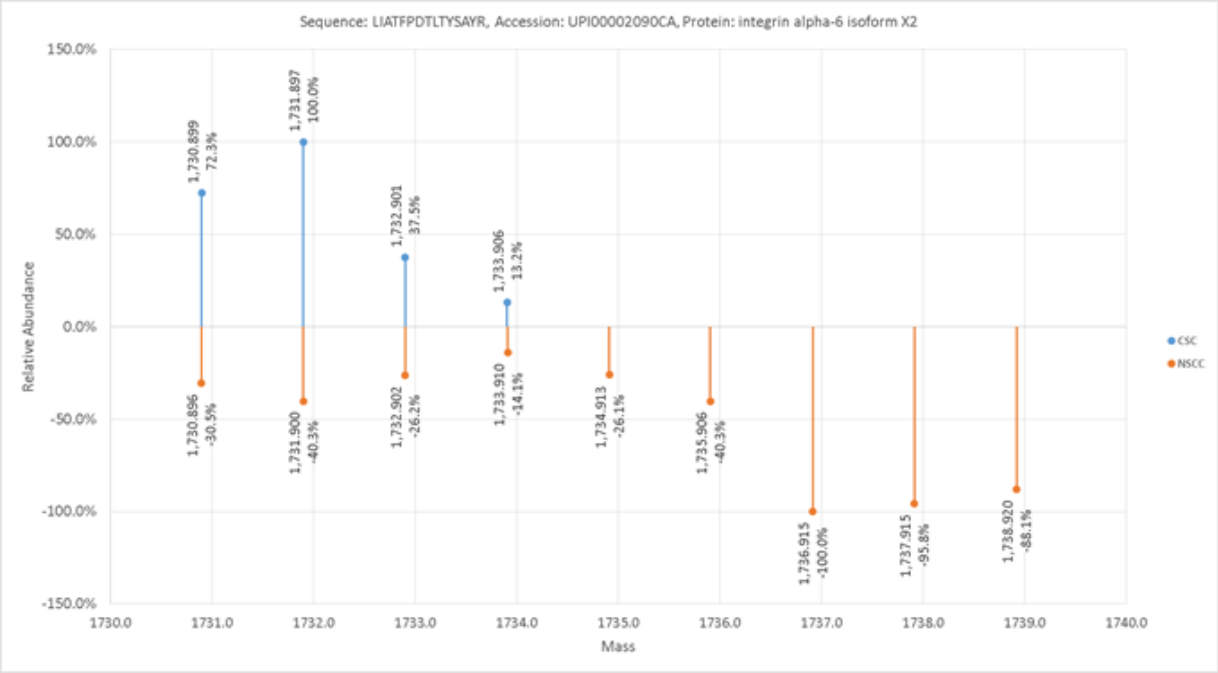


Figure 8: Comparison of the isotopomer distribution in α-6 isoform X2 for CSCs (in blue) and NSCCs (in orange). The mirror display shows high glucose utilized for the synthesis of protein α-6 isoform X2 in NSCCs.

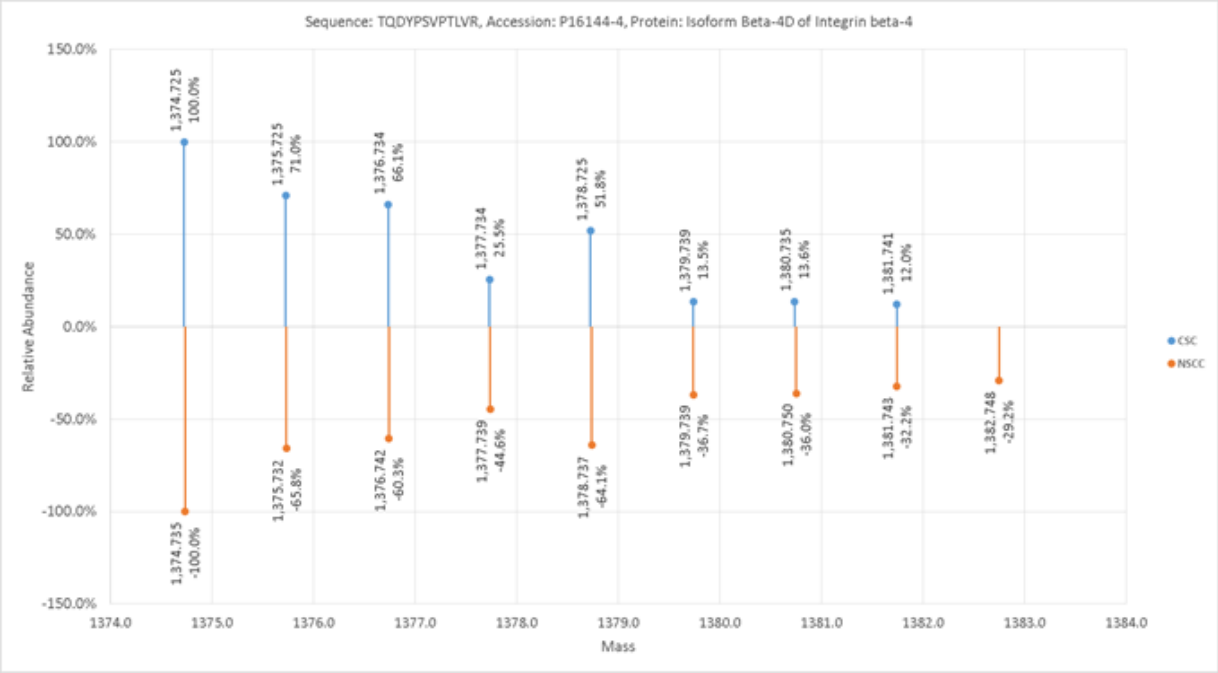


Figure 9: Comparison of the isotopomer distribution in isoform β-4D of integrin β-4 for CSCs (in blue) and NSCCs (in orange). The mirror display shows high glucose utilized for the synthesis of protein isoform β-4D of integrin β-4 in NSCCs.

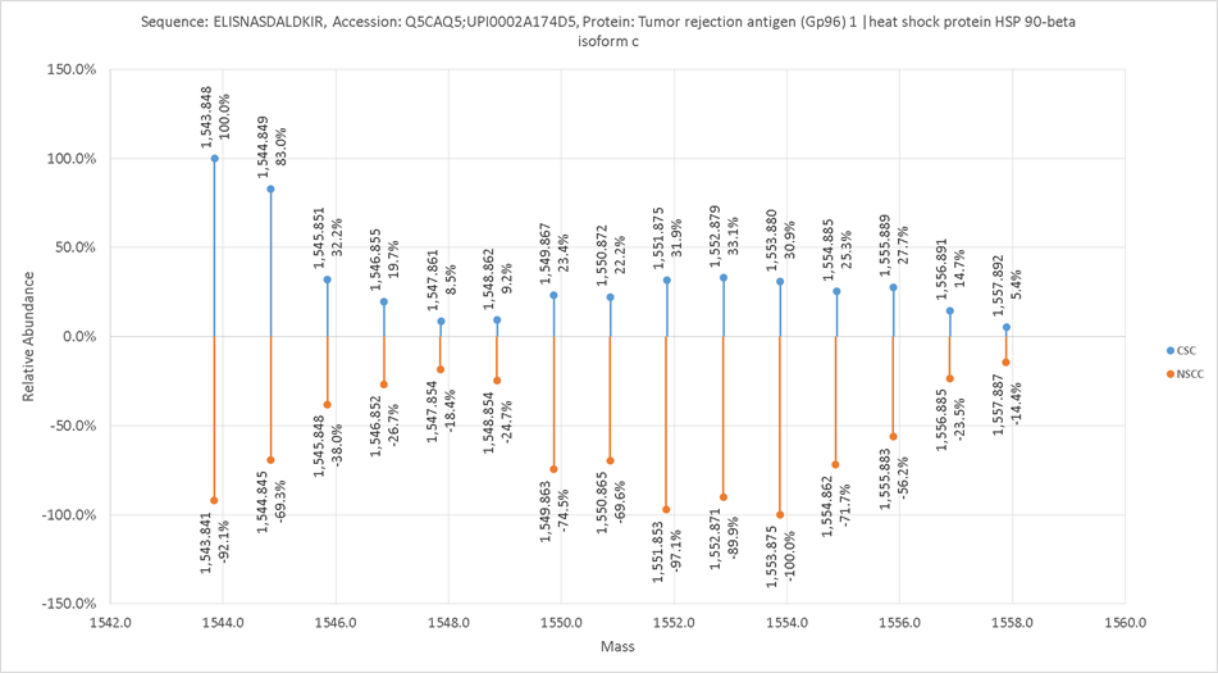


Figure 10: Comparison of the isotopomer distribution in Gp96 for CSCs (in blue) and NSCCs (in orange). The mirror display shows high glucose utilized for the synthesis of protein Gp96 in NSCCs.

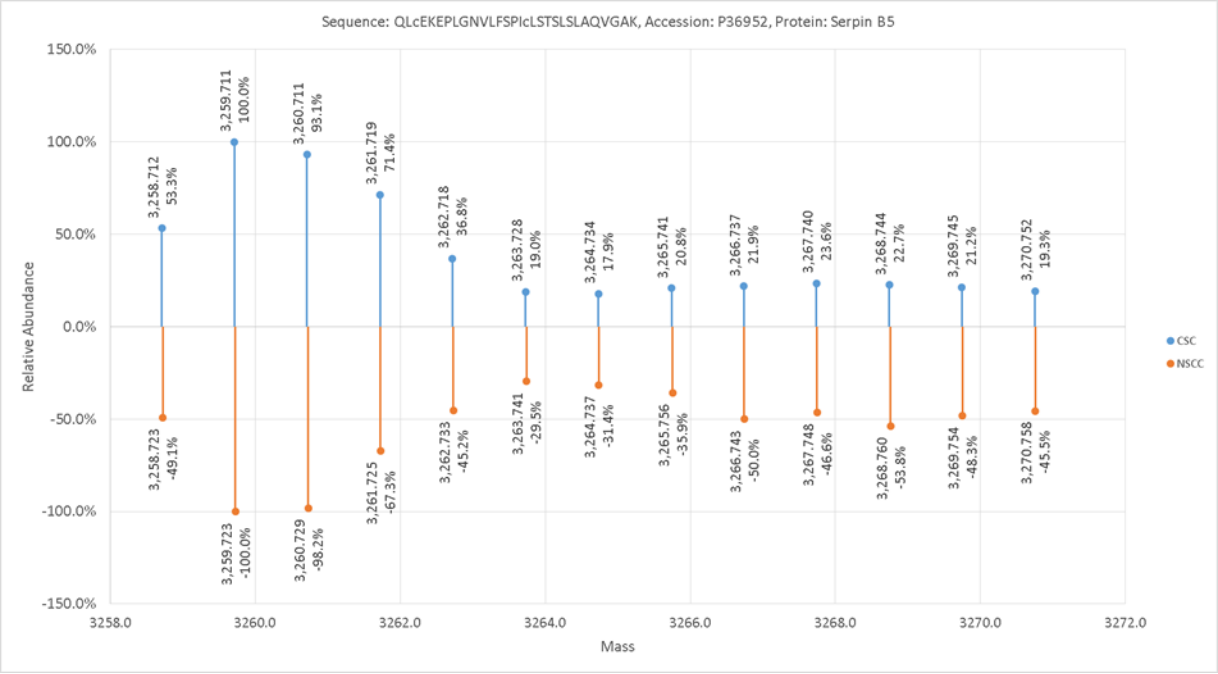


Figure 11: Comparison of the isotopomer distribution in serpin B5 for CSCs (in blue) and NSCCs (in orange). The mirror display shows high glucose utilized for the synthesis of protein serpin B5 in NSCCs.

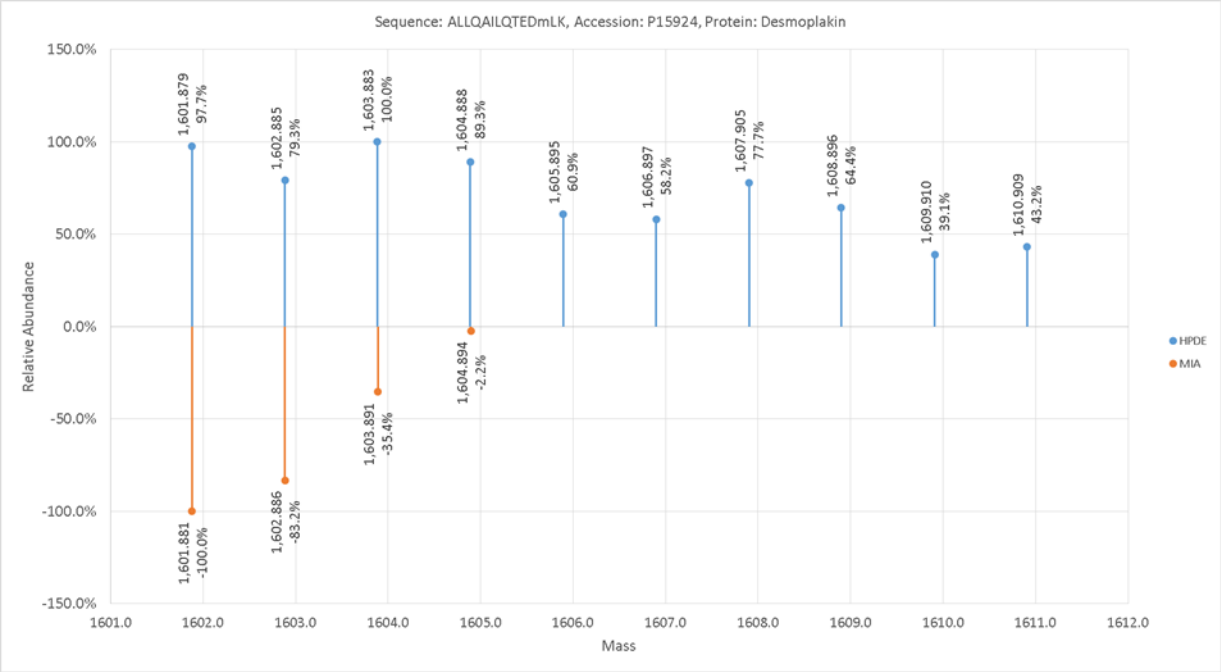


Figure 12: Comparison of the isotopomer distribution in desmoplakin for HPDE cells (in blue) and MIA PaCa-2 cells (in orange). The mirror display shows high glucose utilized for the synthesis of protein desmoplakin in HPDE cells.

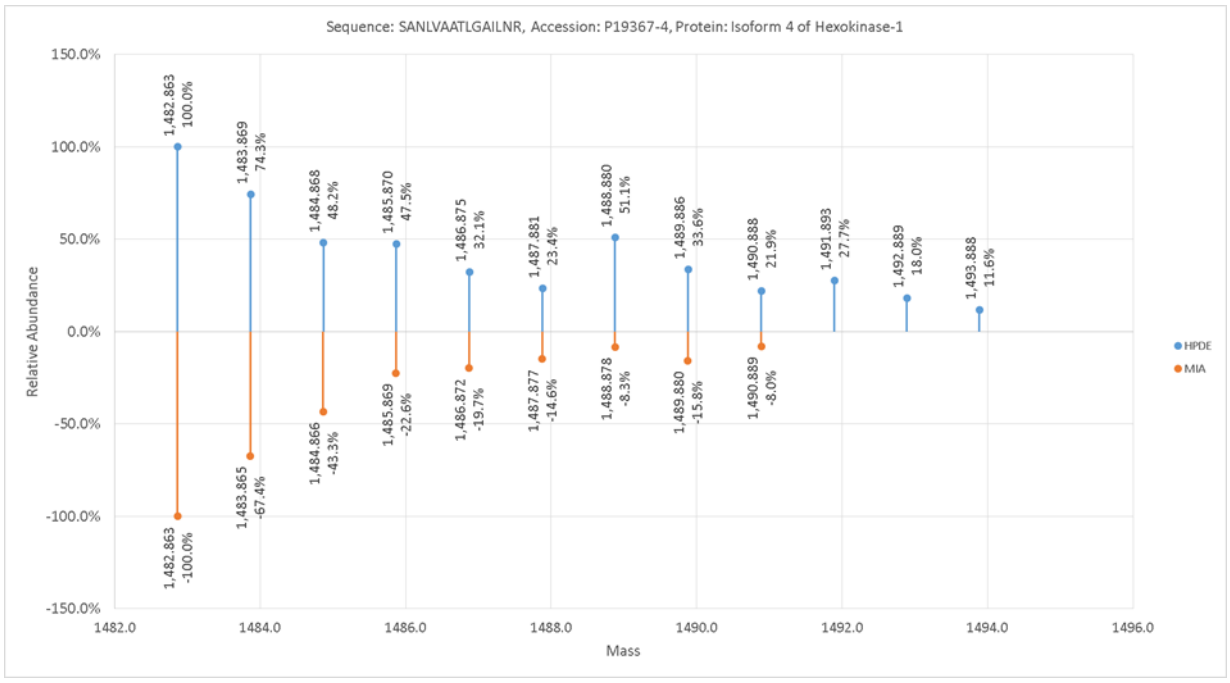


Figure 13: Comparison of the isotopomer distribution in isoform 4 of hexokinase for HPDE cells (in blue) and MIA PaCa-2 cells (in orange). The mirror display shows high glucose utilized for the synthesis of protein isoform 4 of hexokinase in HPDE cells.

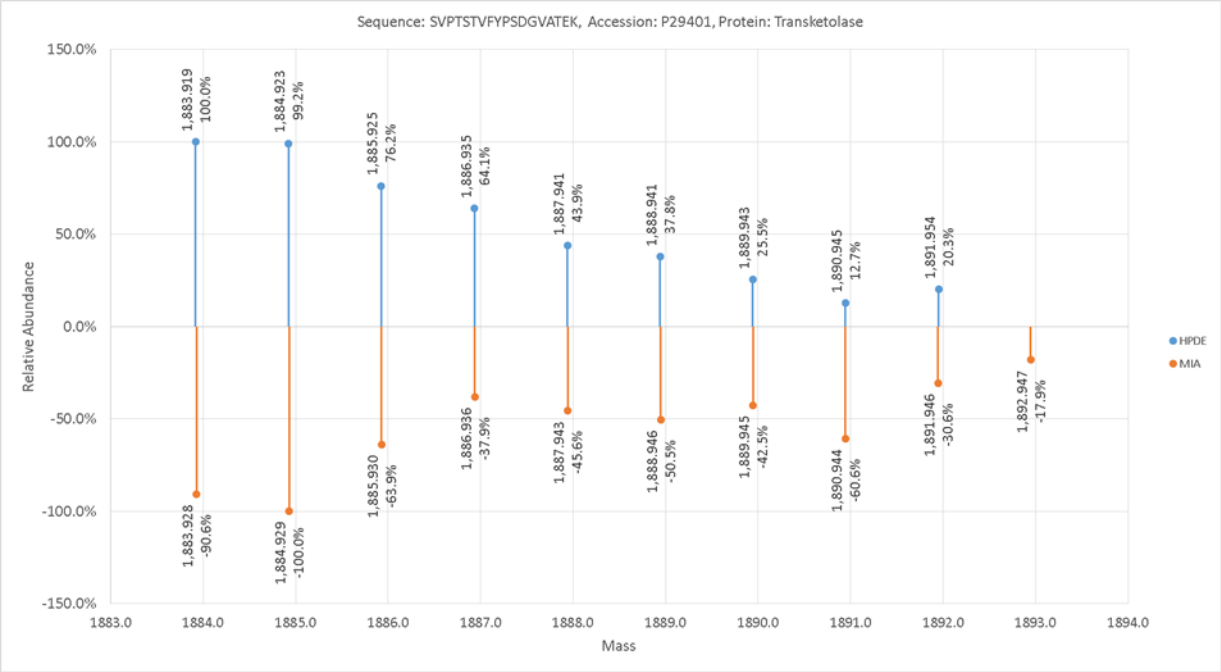


Figure 14: Comparison of the isotopomer distribution in transketolase for HPDE cells (in blue) and MIA PaCa-2 cells (in orange). The mirror display shows high glucose utilized for the synthesis of protein transketolase in HPDE cells.

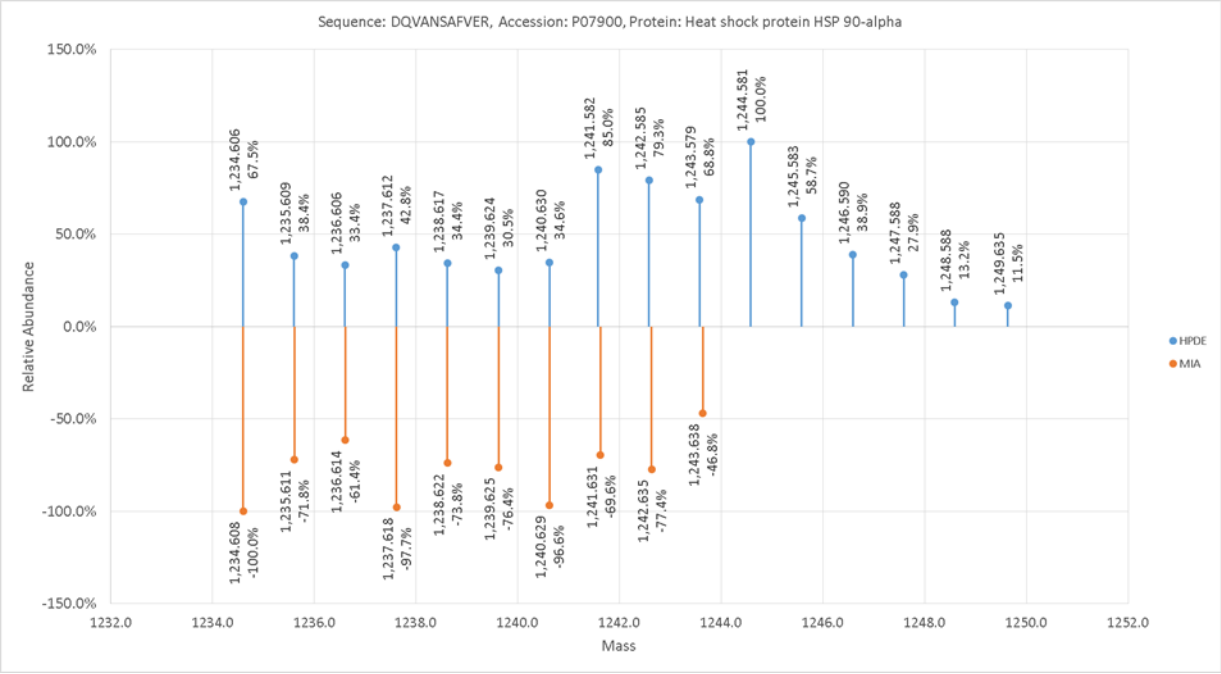


Figure 15: Comparison of the isotopomer distribution in HSP90 for HPDE cells (in blue) and MIA PaCa-2 cells (in orange). The mirror display shows high glucose utilized for the synthesis of protein HSP90 in MIA PaCa-2 cells.

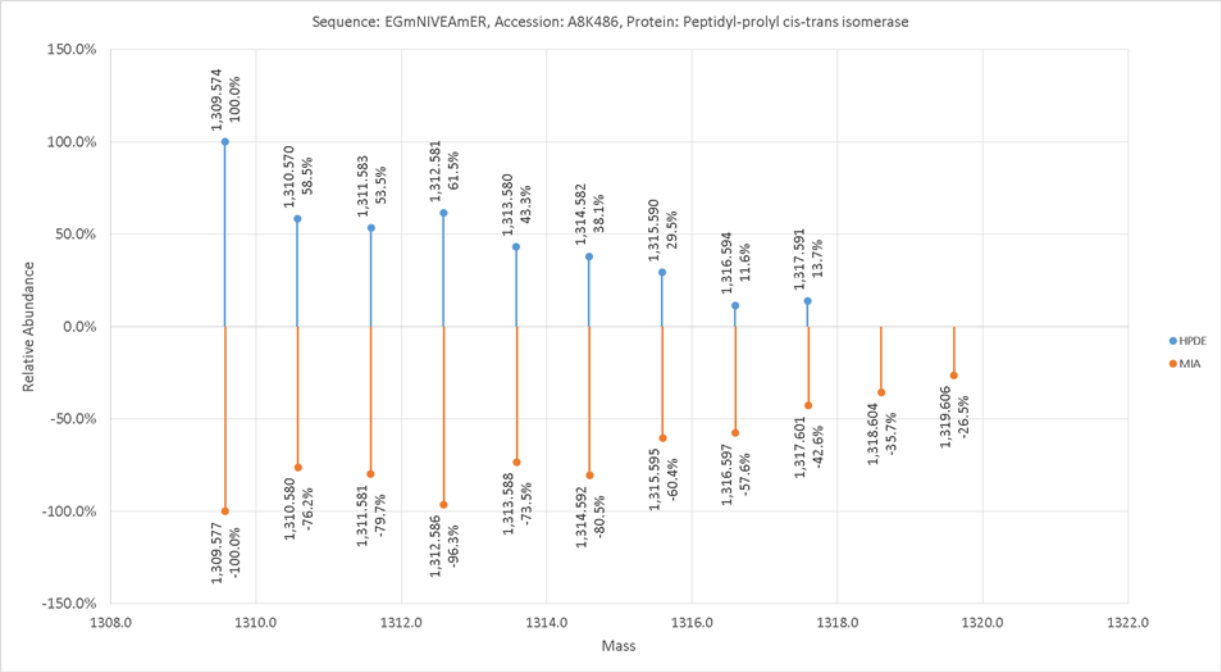


Figure 16: Comparison of the isotopomer distribution in peptidyl-prolyl cis-trans isomerase for HPDE cells (in blue) and MIA PaCa-2 cells (in orange). The mirror display shows high glucose utilized for the synthesis of protein peptidyl-prolyl cis-trans isomerase in MIA PaCa-2 cells.

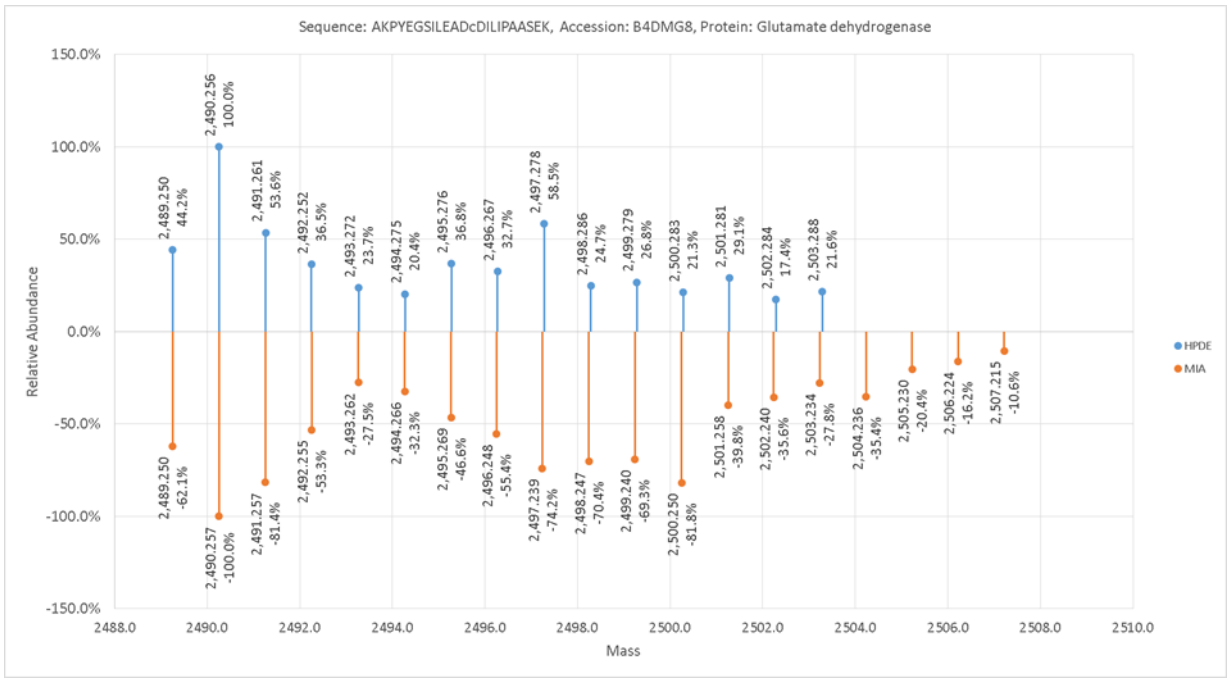


Figure 17: Comparison of the isotopomer distribution in GLDH for HPDE cells (in blue) and MIA PaCa-2 cells (in orange). The mirror display shows high glucose utilized for the synthesis of protein GLDH in MIA PaCa-2 cells.

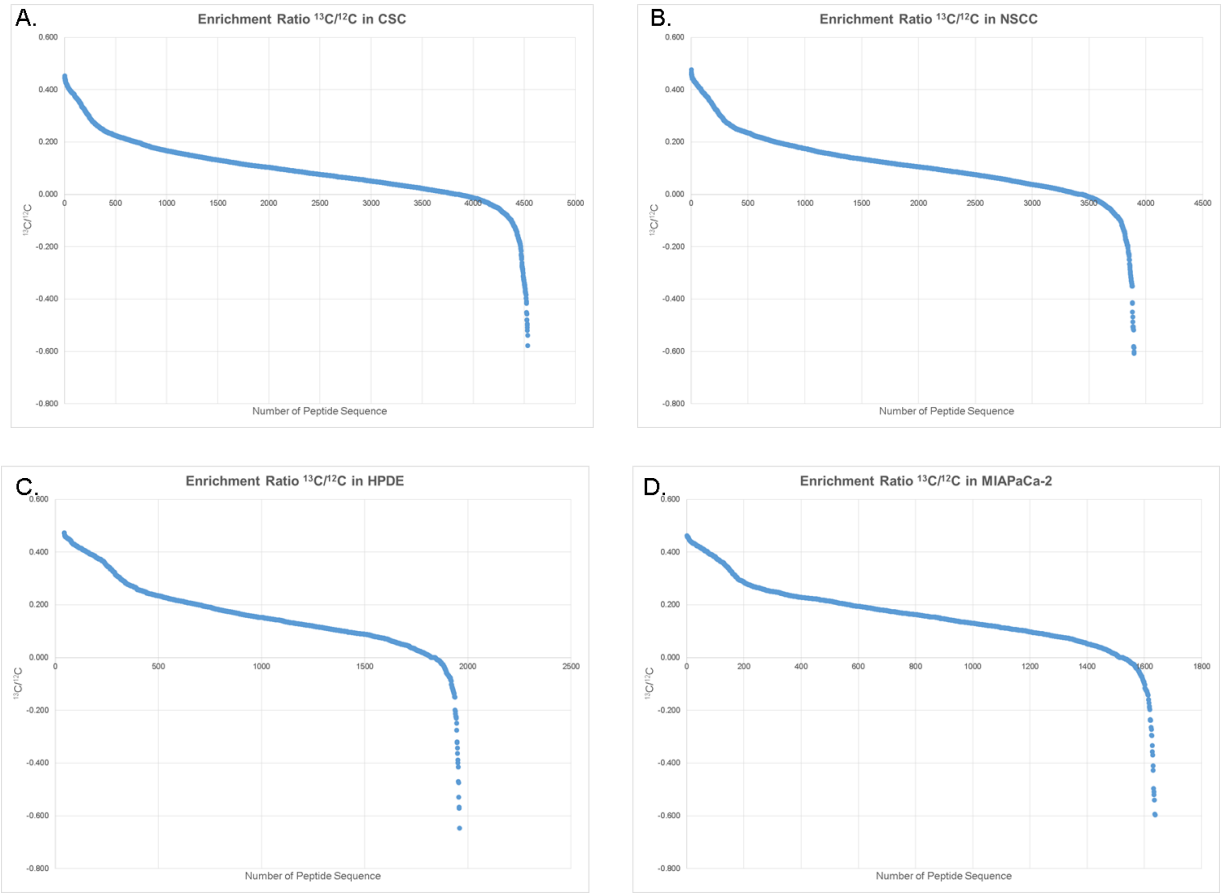
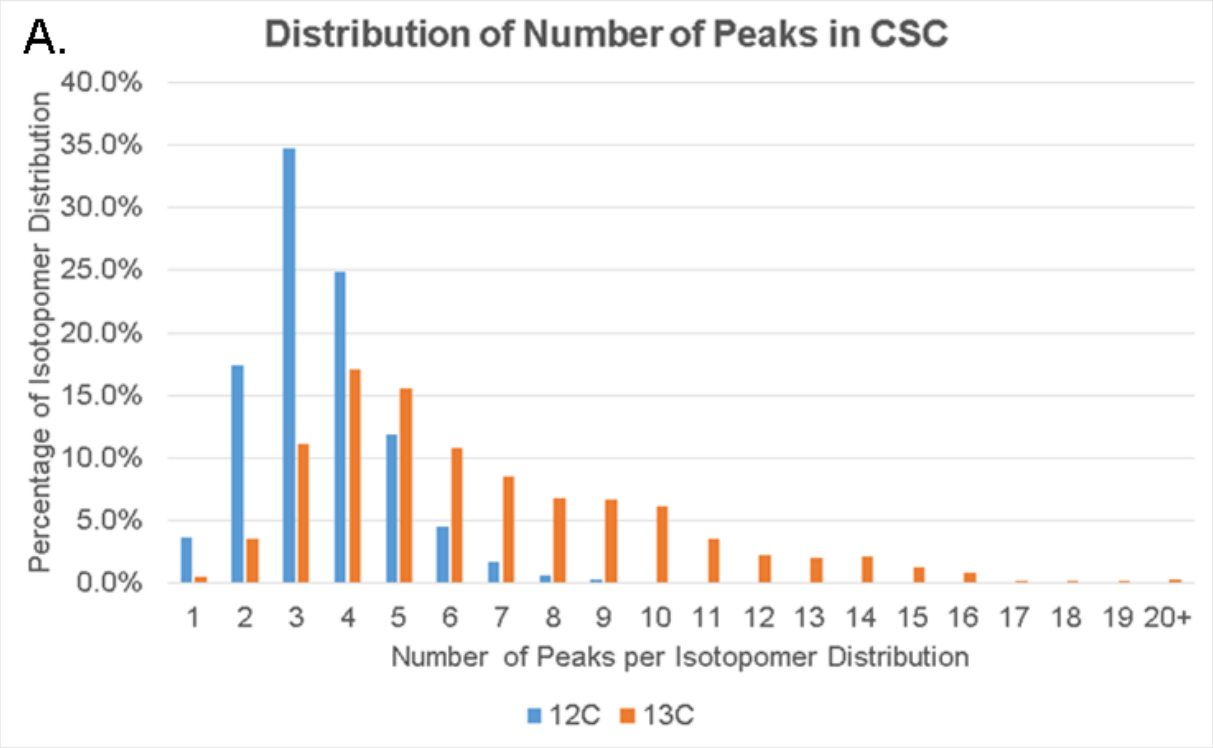


Figure 18A-D: Quantification of enrichment ratio between the cells lines. This diagram represents the number of enrichment ratio of label and unlabeled ($^{13}C/^{12}C$) peptide sequences by percentage of glucose uptake.



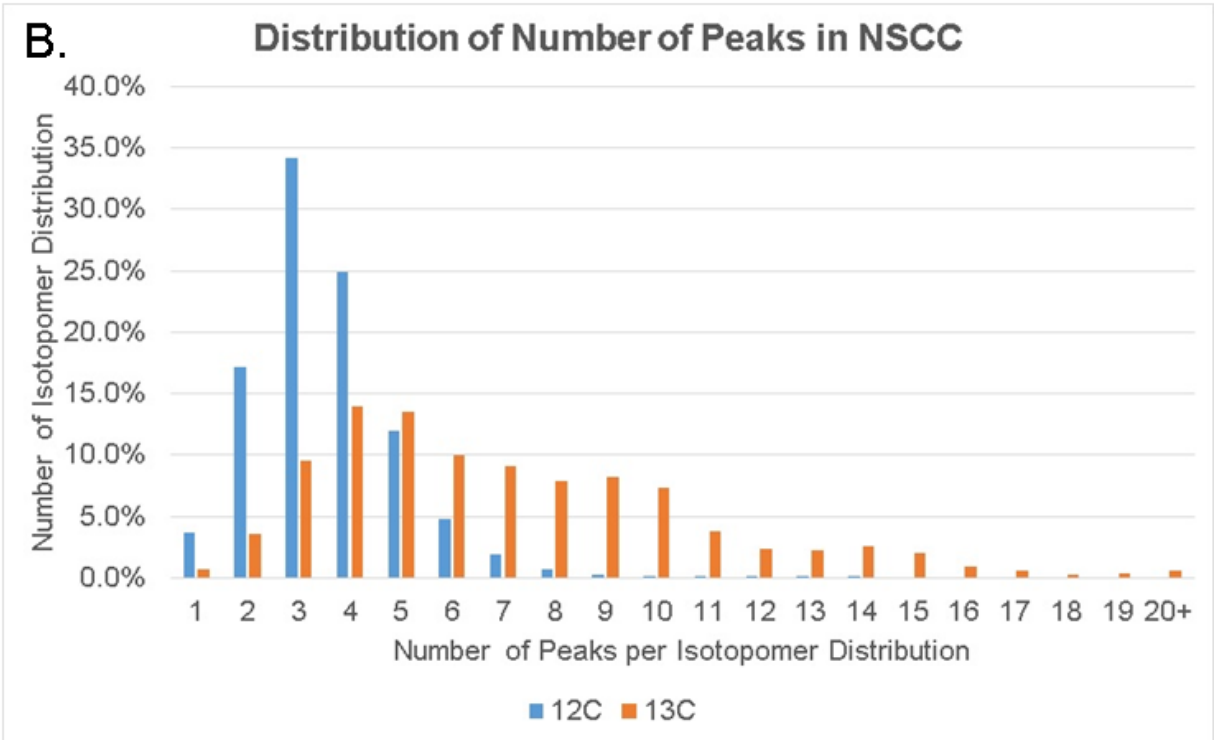
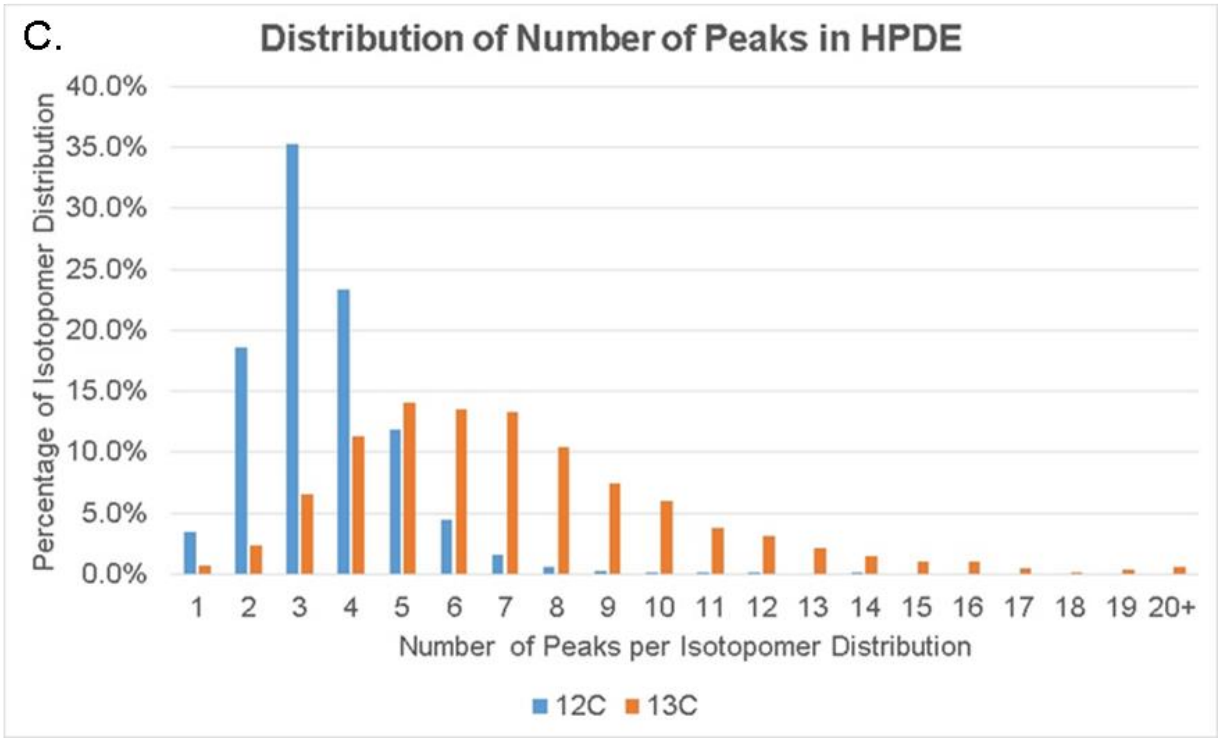


Figure 19A-B: The histograms show the distribution of ^{12}C (blue) and ^{13}C (red) in CSCs and NSCCs.



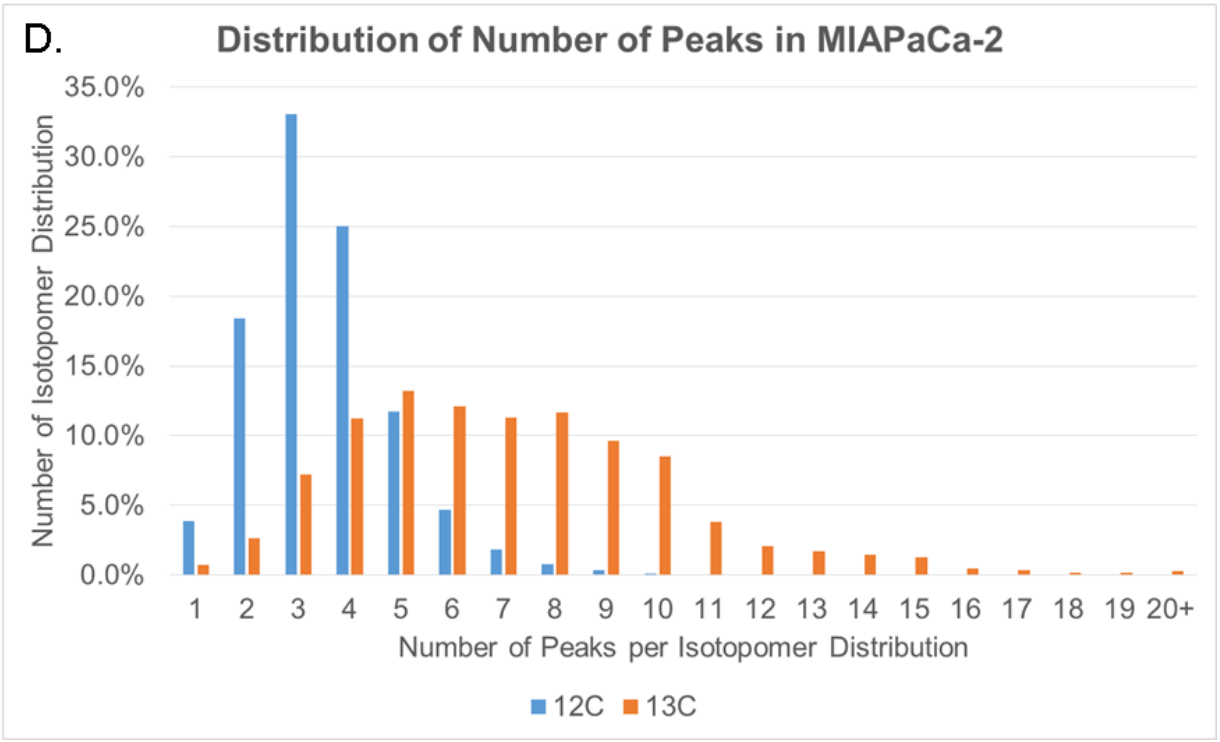


Figure 20C-D: The histograms show the distribution of ^{12}C (blue) and ^{13}C (red) in HPDE cells and MIAPaCa-2 cells.

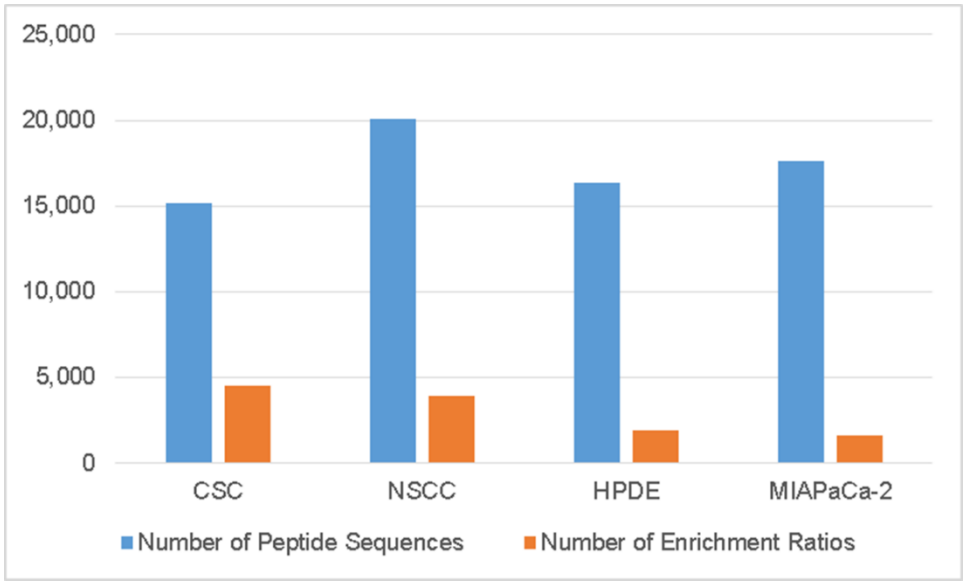


Figure 21: Global number of enrichment ratio between cell lines. The histogram quantifies the total number of peptide sequences and the number of peptide sequences that contains enrichment ratio in each cell line.

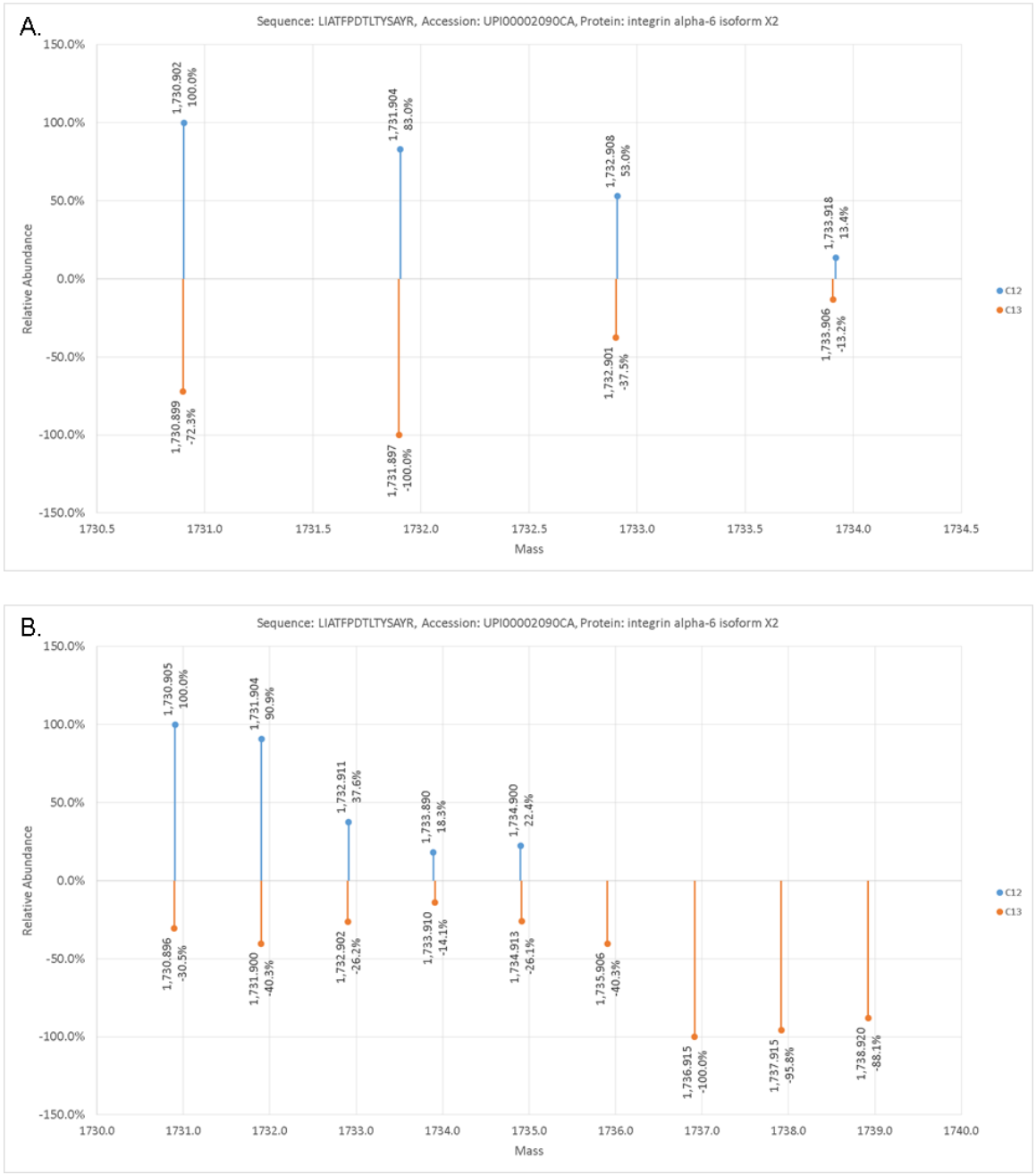


Figure 22: (A) Comparison of ^{12}C unlabeled and ^{13}C labeled peptide isotopomer distribution of integrin α-6 isoform X2 in CSCs. The isotopomer distribution shows that there is no enrichment ratio in protein integrin α-6 isoform X2 in CSCs compare with NSCCs. **(B)** Comparison of ^{12}C unlabeled and ^{13}C labeled peptide isotopomer distribution of protein integrin α-6 isoform X2 in NSCCs. The isotopomer distribution shows an enrichment ratio in NSCCs.

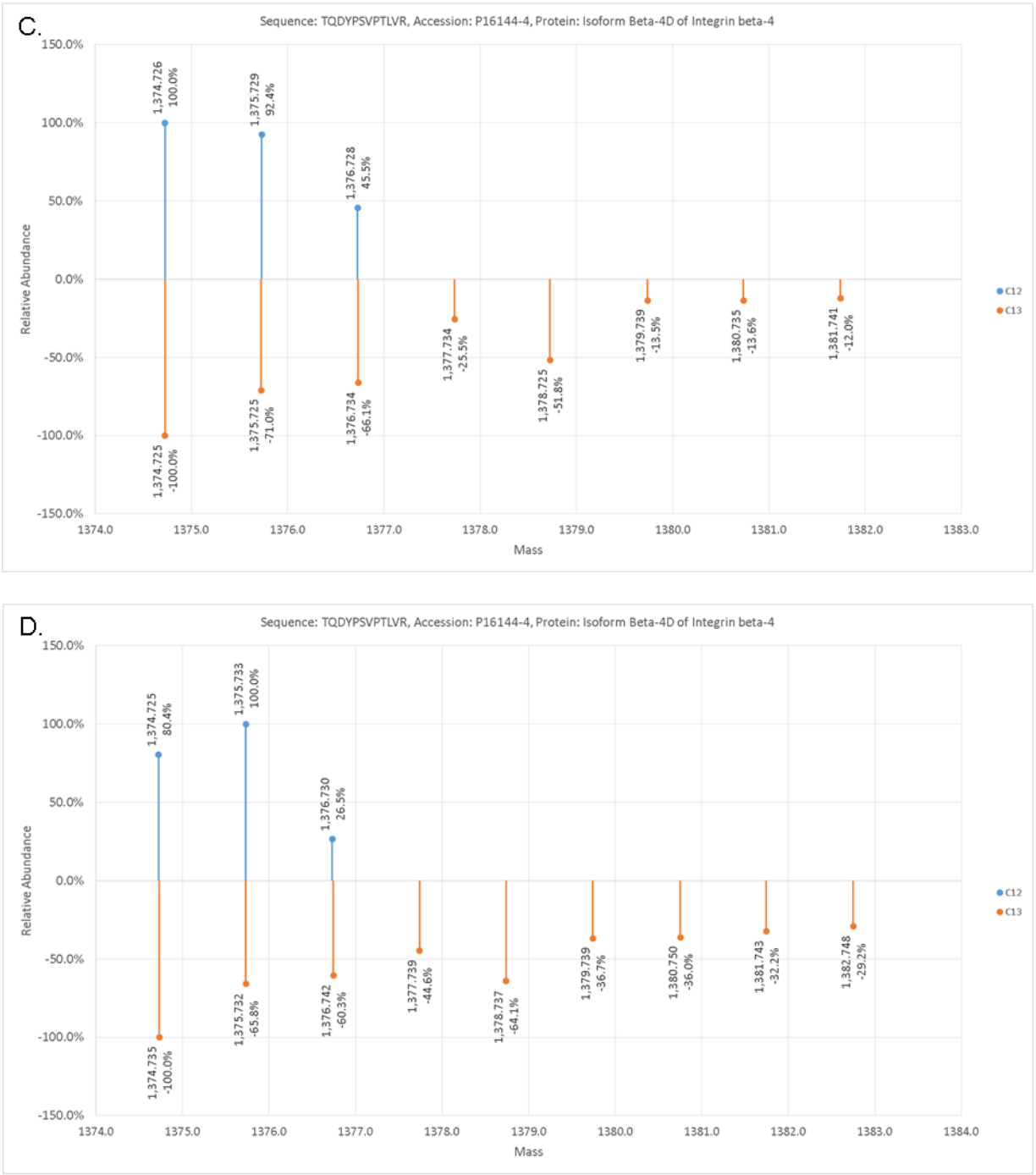


Figure 22: (C) Comparison of ^{12}C unlabeled and ^{13}C labeled peptide isotopomer distribution of protein isoform β-4D of integrin β-4 in CSCs. **(D)** Comparison of ^{12}C unlabeled and ^{13}C labeled peptide isotopomer distribution of protein isoform β-4D of integrin β-4 in NSCCs.

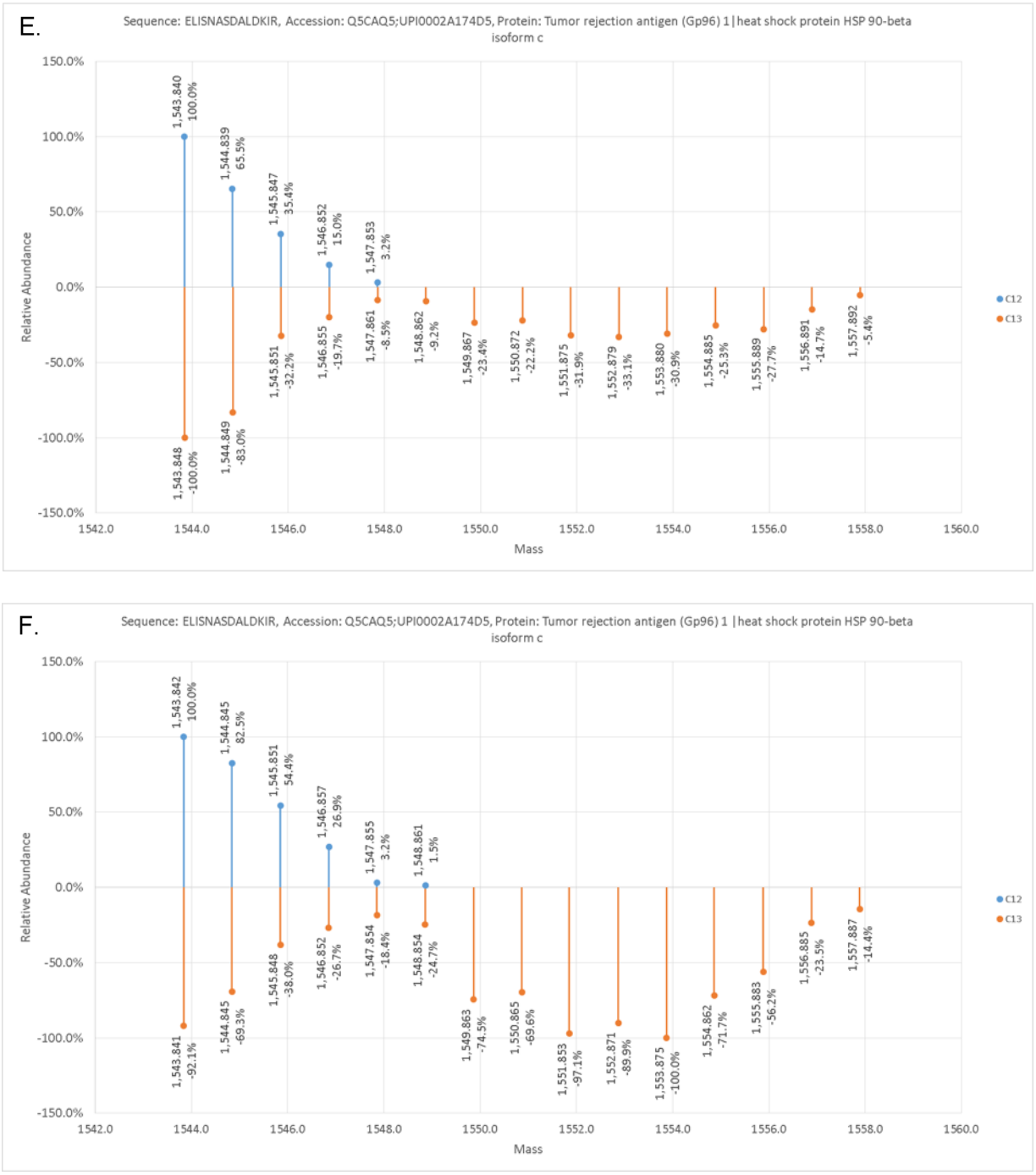


Figure 22: (E) Comparison of ^{12}C unlabeled and ^{13}C labeled peptide isotopomer distribution of protein Gp96 in CSCs. **(F)** Comparison of ^{12}C unlabeled and ^{13}C labeled peptide isotopomer distribution of protein GP96 in NSCCs.

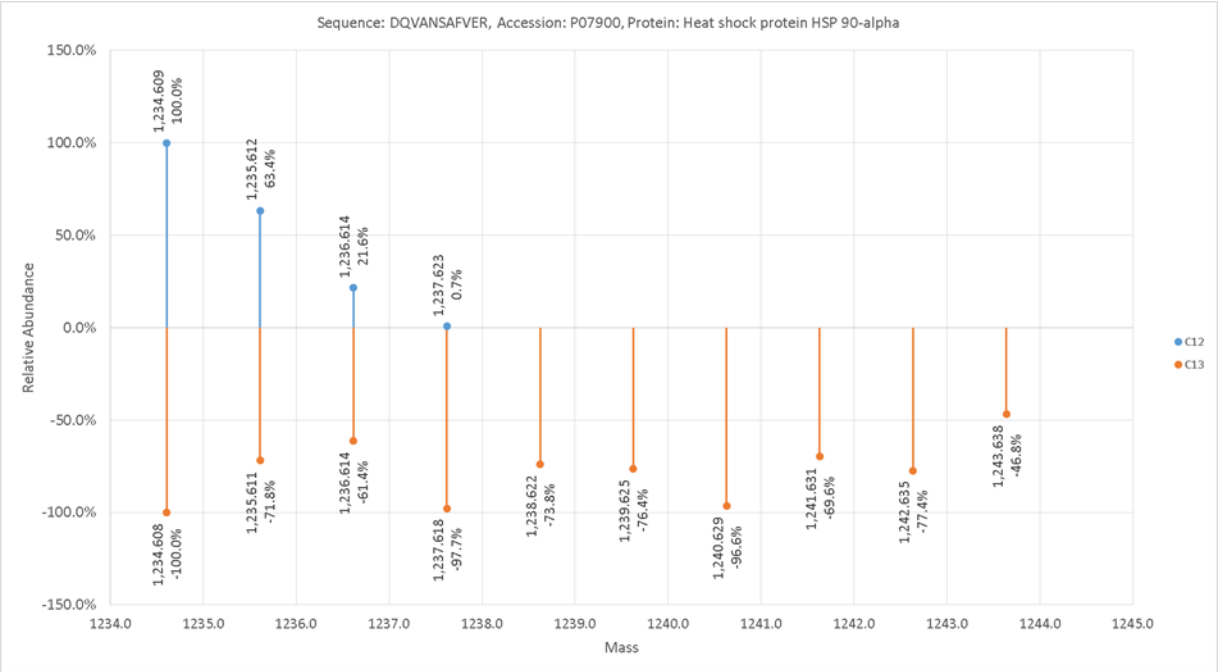


Figure 23: Comparison of ^{12}C unlabeled and ^{13}C labeled peptide isotopomer distribution of protein HSP90 in MIA PaCa-2 cells.

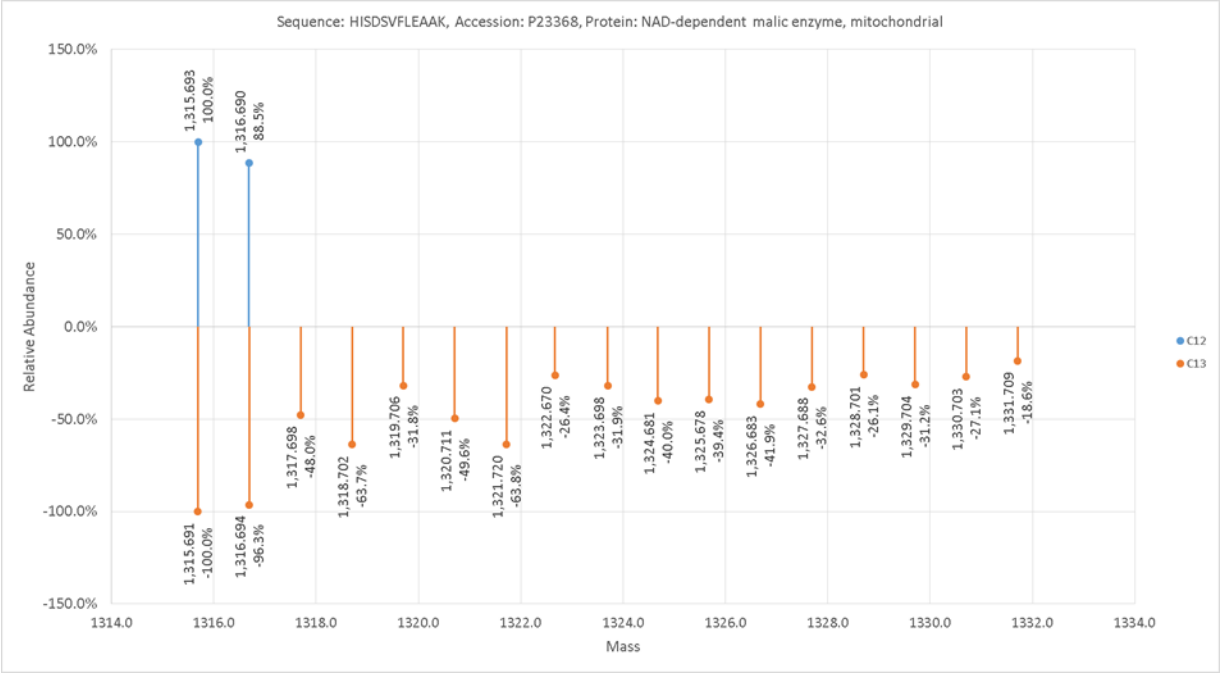


Figure 24: Comparison of 12C unlabeled and 13C labeled peptide isotopomer distribution of protein NAD-dependent malic enzyme in MIA PaCa-2 cells.

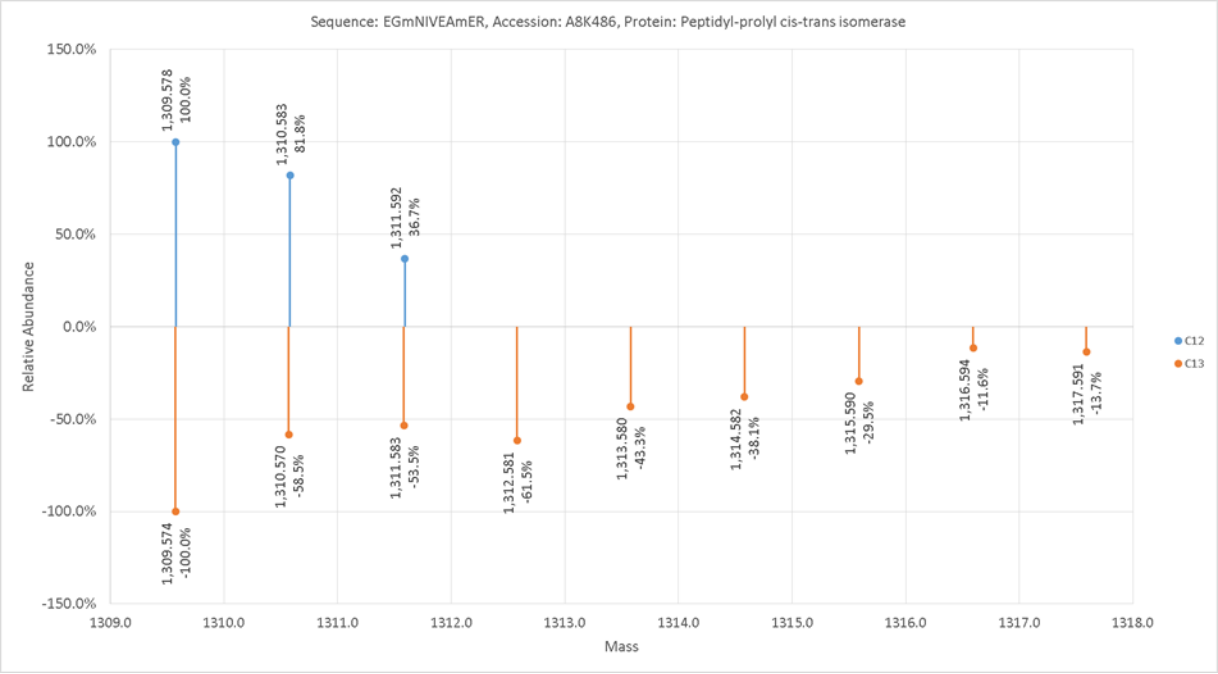


Figure 25: Comparison of ^{12}C unlabeled and ^{13}C labeled peptide isotopomer distribution of protein name peptidyl-prolyl cis-trans isomerase in HPDE cells.

Table 1: ^{13}C labeling of proteins in CSCs/NSCCs. Note: Abundance sum is the sum of relative abundance in an isotopomer distribution. Abundance sum difference is the difference in the abundance sum between two cell lines.

Sequence	Accessions	Protein Descriptions	# Peaks CSC	# Peaks NSCC	# Peaks Difference	Abundance Sum CSC	Abundance Sum NSCC	Abundance Sum Difference	Chi-Squared Normalized
LNEDMAcSVAGITSDANVL TNELR	B2RDG0;H0YLC2	Proteasome subunit alpha type n=1	5	4	1	3.168	3.354	-0.186	0.694
DGFIDKEDLHDMLASLGK	J3QRS3	Myosin regulatory light chain 12A	10	13	-3	3.544	9.466	-5.922	0.621
GAGTGGLGLTVEGPcEAK	O75369-2	Isoform 2 of Filamin-B	11	11	0	3.331	7.142	-3.811	0.608
LPEEWSQWLGGSSWPGYV RPLPPAAIESPAVAAPAYSR	F8WE04	Heat shock protein beta-1	7	6	1	3.858	3.762	0.096	0.586
MLVVLLQANRDPDAGIDEA QVEQDAQALFQAGELK	D6RBL5	Annexin	8	7	1	4.186	3.977	0.209	0.573
LADDVDLEQVANETHGHV GADLAALcSEAALQAIR	Q9HAP0	Valosin-containing protein (Fragment)	8	6	2	3.745	4.088	-0.343	0.560
HVLVEYPMTLSLAAAQELW ELAEQK	P53004	Biliverdin reductase A	5	5	0	2.929	3.432	-0.503	0.558
SQSSNDTFPTAMHIAAAIE VHEVLLPGLQK	P07954	Fumarate hydratase, mitochondrial	6	8	-2	3.669	4.111	-0.442	0.556
VHSPSGALEEcYVTEIDQDK YAVR	P21333	Filamin-A	5	6	-1	3.366	4.080	-0.714	0.528
ALEEANTELEVK	Q04695	Keratin, type I cytoskeletal 17	10	10	0	2.159	4.924	-2.765	0.526
GPIVESYIGFIESYR	Q9NY33	Dipeptidyl peptidase 3	3	7	-4	2.180	3.358	-1.178	0.493
LIATFPDTLTYSAYR	UPI00002090CA	integrin alpha-6 isoform X2	4	9	-5	2.229	4.613	-2.383	0.488
TMVGFGPEDDHFVAELTYN YGVGDYK	B7Z403	Glyoxalase domain-containing protein 4	5	5	0	3.339	3.219	0.120	0.469
TIFPLFMK	Q8WYA6-2	Isoform 2 of Beta-catenin-like protein 1	3	3	0	1.691	2.069	-0.378	0.468
LTTDFNVIVEALSK	B5BUB5	Autoantigen La (Fragment)	9	10	-1	2.819	5.714	-2.894	0.467
DLGEELEALK	P35579	Myosin-9	2	9	-7	1.729	3.006	-1.277	0.454
LFVYDPNNPPSSEVLR	Q15165-3	Isoform 3 of Serum paraoxonase/arylesterase 2	3	12	-9	2.089	7.466	-5.377	0.451
RAEFTVETR	P21333	Filamin-A	15	8	7	5.766	5.121	0.645	0.450
VDcTAHSDVcSAQGVR	Q86UY0	TXNDC5 protein	12	14	-2	3.802	8.067	-4.265	0.449
LSLNIDPDAK	Q5CAQ5	Tumor rejection antigen (Gp96) 1	9	2	7	2.574	1.763	0.811	0.446
FGAVFLWDSGSSVGEITGH NK	O75083	WD repeat-containing protein 1	5	11	-6	3.131	8.449	-5.318	0.441
TSIFWNDVKDPVSIEER	B4DXN6	Eukaryotic translation initiation factor 3 subunit B	4	10	-6	2.642	5.910	-3.268	0.441
GYGFVHFETQEAAER	A0A024R9E2	Poly(A) binding protein, cytoplasmic 1, isoform CRA_c	5	14	-9	2.747	8.923	-6.176	0.438
RPYGVGLLIAGYDDMGPHIF QTcPSANYFDcR	V9HW55	Epididymis secretory protein Li 275	8	10	-2	3.920	3.942	-0.021	0.437
GVQYLNEIKDSVVAGFQW ATK	B4DPU3	cDNA FLJ56548, highly similar to Elongation factor 2	15	16	-1	5.744	9.304	-3.560	0.434
DPETLVGYSMVGcQR	P49327	Fatty acid synthase	7	1	6	4.021	1.000	3.021	0.432

Table 2: ^{13}C labeling of proteins in HPDE/MIA PaCa-2 cells.

Sequence	Accessions	Protein Description	# Peaks HPDE	# Peaks MIAPaCa-2	# Peaks Difference	Abundance Sum HPDE	Abundance Sum MIAPaCa-2	Abundance Sum Difference	Chi-squared Normalized
ALLQAILQTEDmLK	P15924	Desmoplakin	10	4	6	7.100	2.208	4.892	0.571
NcVVELNFGQK	Q1KMD3	Heterogeneous nuclear ribonucleoprotein U-like protein 2	1	4	-3	1.000	3.197	-2.197	0.549
ASLENSLEETK	P02533;P08779	Keratin, type I cytoskeletal 14|Keratin, type I cytoskeletal 16	10	6	4	5.388	1.948	3.440	0.499
YVEcSALTQK	B4E1U9	Cell division control protein 42 homolog n=1	6	7	-1	2.240	4.206	-1.966	0.494
SDFFELLSNHHLDSQSR	UPI000387B8CC	transcription elongation regulator 1 isoform X3	4	9	-5	2.443	5.708	-3.266	0.489
FGDPEcQVILPLLK	P22102	Trifunctional purine biosynthetic protein adenosine-3	3	1	2	2.002	1.000	1.002	0.479
VTEAEIVPmGK	O75369-2	Isoform 2 of Filamin-B	1	19	-18	1.000	8.675	-7.675	0.478
ALELFLESLLK	E9PNC7	Dr1-associated corepressor (Fragment)	7	9	-2	2.479	6.242	-3.763	0.477
DQVANSAFVER	P07900	Heat shock protein HSP 90-alpha	16	10	6	7.649	7.715	-0.066	0.472
FFGDSAASmAIK	B7Z2N4	150 kDa oxygen-regulated protein (Orp150)	12	5	7	6.194	3.323	2.871	0.451
TAAAVAAQSGILDR	Q9UHS8	PRO1975	14	3	11	7.935	2.121	5.814	0.445
LAMQEFmILPVGAANFR	UPI0003EAF40A	alpha-enolase isoform X1	4	8	-4	2.384	6.017	-3.633	0.440
VSLAGAcGVGGYGSR	P13647	Keratin, type II cytoskeletal 5	6	5	1	3.724	2.198	1.526	0.421
FAEFLLPLLIEK	Q5T454	MMS19 nucleotide excision repair protein homolog	7	2	5	4.724	1.876	2.849	0.416
NPPGFAFVEFEDPRDAEDAVR	B4DEK2	Splicing factor, arginine/serine-rich 7	4	13	-9	2.341	7.924	-5.583	0.412
LASVPGSQTVVVK	Q96N83	podocalyxin-like protein mRNA	6	7	-1	2.843	5.427	-2.584	0.402
EALDVLGAVLK	B4DT94	Mitochondrial 28S ribosomal protein S27	4	10	-6	3.160	6.785	-3.626	0.399
GITAFIVEK	H0YKV0	Isovaleryl-CoA dehydrogenase, mitochondrial	1	9	-8	1.000	4.591	-3.591	0.399
LNEQASEEILK	Q01105-3	Isoform 3 of Protein SET	10	10	0	4.122	6.394	-2.272	0.398
mLmmmALLVPLVYTIK	P08574	Cytochrome c1, heme protein, mitochondrial	4	7	-3	2.941	5.445	-2.505	0.398
SAIYQLEEEYENLLK	P15924	Desmoplakin	11	5	6	7.287	3.037	4.250	0.397
KPAAGLSAAPVPTAPAAGAPLmDFGNDFVPPAPR	Q6IPN0	RTN4 protein	26	5	21	13.789	3.072	10.717	0.396
AGNNmLLVGVHGPR	Q5HY54	Filamin-A	6	1	5	3.340	1.000	2.340	0.390
AELIVQPELK	P15924	Desmoplakin	8	4	4	4.883	2.134	2.749	0.388
VAQVAEITYGQK	E5RJ09	Erlin-2	15	2	13	6.522	1.688	4.834	0.387
HGGVIHIYVDK	H0Y4X3	RNA-binding protein 39	1	5	-4	1.000	2.933	-1.933	0.387
EIDLVNRDPK	Q7Z4F3	Caveolin	7	12	-5	4.008	7.736	-3.729	0.382
IFFAGDTIPK	O75369-2	Isoform 2 of Filamin-B	12	15	-3	4.260	7.250	-2.990	0.376
VIFLENYR	P06737-2;Q59GM9	Isoform 2 of Glycogen phosphorylase|Phosphorylase	7	5	2	4.135	2.561	1.574	0.376
VDIQTEDLEDGTcK	O75369-2	Isoform 2 of Filamin-B	6	3	3	4.582	2.395	2.187	0.376
SLVASLAEPDFVVTDFAK	P22314	Ubiquitin-like modifier-activating enzyme 1	5	9	-4	3.227	6.921	-3.693	0.374
IAYSLLGLK	B4DL04	Ephrin type-A receptor 2	2	6	-4	1.772	3.335	-1.563	0.373
VALAGLLGFGLGK	Q56VL3	OCIA domain-containing protein 2	10	12	-2	4.191	6.804	-2.613	0.372
GLAVFISDIR	M0R2D9	AP-2 complex subunit alpha-1	4	6	-2	2.026	4.250	-2.224	0.372
NVGLDIEAEVPAVK	Q53HU0	Chaperonin containing TCP1, subunit 8 (Theta)	5	17	-12	3.351	8.919	-5.568	0.367
QcANLQNAIADAEQR	P13647	Keratin, type II cytoskeletal 5|Keratin, type II cytoskeletal 5	17	11	6	7.767	2.758	5.009	0.362
LYSILGTTLK	O75083	WD repeat-containing protein 1	2	1	1	1.722	1.000	0.722	0.361
NLFILAYNYK	B3KN05	Long-chain-fatty-acid--CoA ligase 3	5	12	-7	3.287	5.387	-2.100	0.360
LTAELIEQAAQYTNAVR	Q53G21	Small nuclear ribonucleoprotein polypeptide A' variant	16	3	13	7.651	2.780	4.871	0.356

Table 3: ^{13}C labeling of keratins in CSCs.

Sequence	Accessions	ProteinDescriptions	# Peaks CSC	# Peaks NSCC	# Peaks Difference	Abundance Sum CSC	Abundance Sum NSCC	Abundance Sum Difference	Chi-squared Normalized
LLEGEDAHLSSQQASGQSYSSR	P08779	Keratin, type I cytoskeletal 16	15	5	10	8.293	2.945	5.348	0.349
LLxGEDAHLSSQQASGQSYSSR	P08779	Keratin, type I cytoskeletal 16	15	5	10	8.293	2.945	5.348	0.349
WELLQQMNVGTRPINLEPI	P35908	Keratin, type II cytoskeletal 2 epidermal	7	6	1	4.044	3.668	0.376	0.346
TEELNKEVASNSELVQSSR	P08779	Keratin, type I cytoskeletal 16	15	4	11	5.914	3.219	2.695	0.204
LLEGEDAHLSSSQFSSGSQS	P02533	Keratin, type I cytoskeletal 14	15	4	11	5.454	2.976	2.478	0.166
EVATNSELVQSGK	P02533;Q04695	Keratin, type I cytoskeletal 14|Keratin, type I cytoskeletal	13	3	10	3.806	2.261	1.545	0.149
QSVENDIHGLR	F8VZY9	Keratin, type I cytoskeletal 18	8	7	1	4.038	3.871	0.167	0.142
TDLEmQIEGLKEELAYLR	P08779	Keratin, type I cytoskeletal 16	19	13	6	4.726	3.170	1.556	0.122
VLYDAEISQIHQSVTDTNVI	P35908	Keratin, type II cytoskeletal 2 epidermal	11	6	5	4.632	3.707	0.925	0.115
LNDLEEALQQAKEDLAR	P35908	Keratin, type II cytoskeletal 2 epidermal	4	4	0	2.895	2.821	0.074	0.086
NKLNDLEEALQQAK	P35908	Keratin, type II cytoskeletal 2 epidermal	4	3	1	2.515	2.120	0.395	0.085
DAETWFLSK	P08779	Keratin, type I cytoskeletal 16	10	8	2	3.795	2.763	1.031	0.084
ADLEmQIENLKEELAYLK	Q04695	Keratin, type I cytoskeletal 17|Junction plakoglobin	14	7	7	3.705	2.643	1.062	0.082
TDLEMQIEGLKEELAYLR	P08779	Keratin, type I cytoskeletal 16	14	5	9	5.640	3.942	1.698	0.074
DVDNAYmIK	P35908	Keratin, type II cytoskeletal 2 epidermal	3	3	0	1.902	1.565	0.337	0.066
GSSSGGGYSSGSSSYGSGGR	P35908	Keratin, type II cytoskeletal 2 epidermal	4	3	1	2.473	2.329	0.144	0.065
NVQDAIADAEQRGEHALK	P35908	Keratin, type II cytoskeletal 2 epidermal	6	4	2	3.078	2.700	0.377	0.064
TAAENDFVTLKK	P35908	Keratin, type II cytoskeletal 2 epidermal	8	4	4	2.473	2.048	0.424	0.058
VQALEEANNDLENKIQDWYDK	P35527	Keratin, type I cytoskeletal 9	5	5	0	3.018	2.905	0.113	0.055
TSQNSELNNmQDLVEDYKK	P35908	Keratin, type II cytoskeletal 2 epidermal	6	5	1	3.053	2.961	0.092	0.054
EVASNSELVQSSR	P08779	Keratin, type I cytoskeletal 16	10	11	-1	4.029	3.373	0.655	0.051
GQTGGDVNVEMDAAPGVDLSR	P08779	Keratin, type I cytoskeletal 16	16	5	11	3.676	2.832	0.845	0.049
SDLEMQYETLQEELmALK	P35527	Keratin, type I cytoskeletal 9	5	6	-1	3.032	2.712	0.320	0.048
QRPSEIKDYSPYFK	P08779	Keratin, type I cytoskeletal 16	13	9	4	4.744	4.159	0.585	0.046
ILNEmRDQYEK	P02533;Q04695	Keratin, type I cytoskeletal 14|Keratin, type I cytoskeletal 17|Junction plakoglobin	9	8	1	3.633	3.373	0.259	0.044
ADLEMQIESLTEELAYLKK	P13645	Keratin, type I cytoskeletal 10	6	6	0	3.102	2.669	0.433	0.043
LLRDYQELMNTK	H6VRF8	Keratin 1	4	3	1	2.577	2.450	0.127	0.039

Table 4: ^{13}C labeling of keratins in NSCCs.

Sequence	Accessions	ProteinDescriptions	# Peaks CSC	# Peaks NSCC	# Peaks Difference	Abundance Sum CSC	Abundance Sum NSCC	Abundance Sum Difference	Chi-squared Normalized
ALEEANTELEVK	Q04695	Keratin, type I cytoskeletal 17	10	10	0	2.159	4.924	-2.765	0.526
ASLEGNLAETENR	Q04695	Keratin, type I cytoskeletal 17|Junction plakoglobin	14	10	4	2.153	5.212	-3.059	0.422
ADLEMQIENLKEELAYLK	Q04695	Keratin, type I cytoskeletal 17 |Junction plakoglobin	15	15	0	4.341	8.155	-3.814	0.304
ADLEMQIESLKEELAYLK	P02533	Keratin, type I cytoskeletal 14	14	14	0	3.607	6.599	-2.991	0.304
DAEDWFFSK	Q04695	Keratin, type I cytoskeletal 17|Junction plakoglobin	11	10	1	3.774	6.768	-2.993	0.258
NHEEEMNALR	P02533;Q04695	Keratin, type I cytoskeletal 14|Keratin, type I cytoskeletal 17|Junction plakoglobin	15	15	0	3.684	7.069	-3.385	0.215
APSTYGGGLSVSSR	P08779	Keratin, type I cytoskeletal 16	2	7	-5	1.936	2.872	-0.936	0.181
DAEEWFFTK	P02533	Keratin, type I cytoskeletal 14	10	10	0	3.087	4.997	-1.911	0.169
ADLEmQIESLKEELAYLK	P02533	Keratin, type I cytoskeletal 14	5	16	-11	2.613	5.355	-2.741	0.159
YcmQLAQIQEMIGSVEEQL AQLR	P02533	Keratin, type I cytoskeletal 14	16	15	1	5.756	7.839	-2.083	0.150
YcMQLAQIQEMIGSVEEQL AQLR	P02533	Keratin, type I cytoskeletal 14	15			5.919	8.339	-2.419	0.133
ILNEMRDQYEK	P02533;Q04695	Keratin, type I cytoskeletal 14|Keratin, type I cytoskeletal 17|Junction plakoglobin	10	14	-4	3.751	5.735	-1.984	0.105
QVLDNLTmEK	P35527	Keratin, type I cytoskeletal 9	7	3	4	2.319	3.505	-1.186	0.102
FEMEQNLR	P35527	Keratin, type I cytoskeletal 9	4	6	-2	2.439	2.678	-0.238	0.100
QEIEcQNQEYSLLLSIK	P35527	Keratin, type I cytoskeletal 9	3	5	-2	2.665	2.818	-0.153	0.099
TLNDmRQEYEQLIAK	P35527	Keratin, type I cytoskeletal 9	5	5	0	2.699	2.750	-0.051	0.089
NMQDLVEDFK	P13647	Keratin, type II cytoskeletal 5	10	10	0	2.641	3.711	-1.070	0.086
DVDNAYMIK	P35908	Keratin, type II cytoskeletal 2 epidermal	4	8	-4	2.022	2.730	-0.709	0.083
NHKEEMSQLTGQNSGDVN VEINVAPGK	P35527	Keratin, type I cytoskeletal 9	5	5	0	3.099	3.515	-0.417	0.081
ASLENSLEETKGR	P02533;P08779	Keratin, type I cytoskeletal 14|Keratin, type I cytoskeletal 16	14	13	1	4.101	5.554	-1.453	0.076
IKFEmEQNLR	P35527	Keratin, type I cytoskeletal 9	10	11	-1	1.950	2.789	-0.839	0.075
cEMEQQNQEYK	P02533;Q04695	Keratin, type I cytoskeletal 14|Keratin, type I cytoskeletal 17|Junction plakoglobin	12	9	3	3.986	4.874	-0.888	0.069
VSLAGAcGVGGYGSR	P13647	Keratin, type II cytoskeletal 5	10	10	0	3.829	4.542	-0.713	0.068
SDLEmQYETLQEELMALKK	P35527	Keratin, type I cytoskeletal 9	7	7	0	2.548	3.108	-0.560	0.067
YETELNLR	P02533	Keratin, type I cytoskeletal 14	6	8	-2	2.528	3.377	-0.849	0.067
ASLENSLEETK	P02533;P08779	Keratin, type I cytoskeletal 14|Keratin, type I cytoskeletal 16	11	9	2	3.383	4.156	-0.773	0.066
SEISELR	P02533;Q04695	Keratin, type I cytoskeletal 14|Keratin, type I cytoskeletal 17 |Junction plakoglobin	7	8	-1	2.118	2.707	-0.590	0.062
TIVEEVQDGK	Q04695	Keratin, type I cytoskeletal 17|Junction plakoglobin	10	10	0	3.465	4.732	-1.267	0.061
TVQSLEIDLDSMR	F8VZY9	Keratin, type I cytoskeletal 18	10	9	1	4.800	5.060	-0.260	0.055
TKYETELNLR	P02533	Keratin, type I cytoskeletal 14	8	9	-1	3.221	4.021	-0.799	0.054
KQISNLQQSISDAEQR	H6VRF8	Keratin 1	4	10	-6	2.442	3.059	-0.617	0.049

Table 5: ^{13}C labeling of keratins in HPDE cells.

Sequence	Accessions	Protein Description	# Peaks HPDE	# Peaks MIAPaCa-2	# Peaks Differences	Abundance Sum HPDE	Abundance Sum MIAPaCa-2	Abundance Sum Difference	Chi-squared Normalized
ASLENSLEETK	P02533;P08779	Keratin, type I cytoskeletal 14|Keratin, type I cytoskeletal 16	10	6	4	5.388	1.948	3.440	0.499
VSLAGAcGVGGYGSR	P13647	Keratin, type II cytoskeletal 5	6	5	1	3.724	2.198	1.526	0.421
QcANLQNAIADAEQR	P13647	Keratin, type II cytoskeletal 5|Keratin, type II cytoskeletal 5	17	11	6	7.767	2.758	5.009	0.362
VDALmDEINFmK	P13647	Keratin, type II cytoskeletal 5| Keratin, type II cytoskeletal 5		2		4.854	2.447	2.406	0.334
xDALmDEINFmK	P13647	Keratin, type II cytoskeletal 5		2		4.854	2.447	2.406	0.334
LAELEEALQK	P13647	Keratin, type II cytoskeletal 5	11	5	6	4.587	1.904	2.682	0.296
ADLEmQIESLK	P02533	Keratin, type I cytoskeletal 14	6	2	4	3.452	1.573	1.879	0.295
QVLDNLTMEK	P35527	Keratin, type I cytoskeletal 9	10	4	6	4.525	2.118	2.408	0.281
ISSVLAGGScR	P02533;P08779	Keratin, type I cytoskeletal 14 |Keratin, type I cytoskeletal 16	6	4	2	2.906	1.741	1.165	0.270
NMQDLVEDFK	P13647	Keratin, type II cytoskeletal 5	8	4	4	3.573	1.934	1.640	0.244
ASLEGNLAETENR	Q04695	Keratin, type I cytoskeletal 17 |Junction plakoglobin	9	5	4	4.164	2.262	1.902	0.234
NmQDLVEDFK	P13647	Keratin, type II cytoskeletal 5	8	5	3	3.483	1.924	1.559	0.231
NmQDLVEDFKNK	P13647	Keratin, type II cytoskeletal 5	8	4	4	3.953	2.016	1.937	0.220
EVASNSELVQSSR	P08779	Keratin, type I cytoskeletal 16	9	4	5	3.794	2.225	1.569	0.220
ALEEANTELEVK	Q04695	Keratin, type I cytoskeletal 17	5	4	1	2.988	2.135	0.853	0.219
NKLAELEEALQK	P13647	Keratin, type II cytoskeletal 5	9	3	6	3.866	2.075	1.790	0.208
GmQDLVEDFK	B4DRR0	Keratin, type II cytoskeletal 6A	5			3.345	2.024	1.321	0.206
LSVEADINGLR	Q04695	Keratin, type I cytoskeletal 17 |Junction plakoglobin n=1	9	6	3	4.788	3.283	1.505	0.200
TEmENEFVLIKK	P05787	Keratin, type II cytoskeletal 8	12	3	9	4.934	2.817	2.118	0.197
TTAENEFVmLKK	P13647	Keratin, type II cytoskeletal 5	9			5.057	2.453	2.604	0.195
ATGGGLSSVGGGSSTIK	B4DKV4	Keratin, type II cytoskeletal 6B	7	4	3	3.268	2.097	1.171	0.190
DLNLDGIIAEVK	Q9NSB2	Keratin, type II cuticular Hb4	7	7	0	3.853	2.836	1.016	0.186
GmQDLVEDFKNK	B4DRR0	Keratin, type II cytoskeletal 6A	7	3	4	3.286	2.078	1.208	0.184
AQYEEIANR	P13647	Keratin, type II cytoskeletal 5	15	5	10	4.400	1.810	2.590	0.182
TKYETELNLR	P02533	Keratin, type I cytoskeletal 14	8	4	4	3.500	2.039	1.461	0.182
FETEQALR	Q04695;P08727	Keratin, type I cytoskeletal 17|Keratin, type I cytoskeletal 19 |Junction plakoglobin	12	2	10	3.693	1.398	2.295	0.181
TTAENEFVmLK	P13647	Keratin, type II cytoskeletal 5	10			4.382	2.059	2.324	0.175
mSVEADINGLR	P02533;P08727	Keratin, type I cytoskeletal 14|Keratin, type I cytoskeletal 19| Junction plakoglobin	8			3.327	2.529	0.798	0.173
TmQNLEIELQSQLSmK	P02533	Keratin, type I cytoskeletal 14	8	7	1	4.509	3.042	1.467	0.170
YEELQQTAGR	P13647	Keratin, type II cytoskeletal 5	12	4	8	3.693	1.901	1.792	0.168
NmEDLVEDFK	UPI0003EAE85F	keratin, type II cytoskeletal 3 isoform X3	4			3.269	1.973	1.296	0.165
SEISELR	P02533;Q04695	Keratin, type I cytoskeletal 14|Keratin, type I cytoskeletal 17|Junction plakoglobin	5	4	1	2.295	1.588	0.707	0.162
mSVEADINGLRR	P02533;P08727	Keratin, type I cytoskeletal 14 |Keratin, type I cytoskeletal 19| Junction plakoglobin	11			3.805	2.525	1.280	0.146
APSTYGGGLSVSSR	P08779	Keratin, type I cytoskeletal 16	9	5	4	3.221	2.123	1.097	0.139
SFSTASAITPSVSR	P13647	Keratin, type II cytoskeletal 5	16	10	6	4.176	2.618	1.559	0.123

Table 6: ^{13}C labeling of keratins in MIA PaCa-2 cells

Sequence	Accessions	Protein Description	# Peaks HPDE	# Peaks MIAPaCa-2	# Peaks Difference	Abundance Sum HPDE	Abundance Sum MIAPaCa-2	Abundance Sum Difference	Chi-squared Normalized
AALEDTLAETEAR	P08727	Keratin, type I cytoskeletal 19 |Junction plakoglobin	17	10	7	4.291	8.218	-3.927	0.304
ALEAANGELEVK	P08727	Keratin, type I cytoskeletal 19	13	11	2	3.946	7.500	-3.554	0.292
ILGATIENSR	P08727	Keratin, type I cytoskeletal 19| Junction plakoglobin	7	8	-1	2.708	4.644	-1.936	0.234
NHEEEISTLR	P08727	Keratin, type I cytoskeletal 19 |Junction plakoglobin	8	9	-1	3.035	4.795	-1.760	0.166
FGPGVAFR	P08727	Keratin, type I cytoskeletal 19	7	8	-1	2.417	3.692	-1.275	0.160
QLETLGQEK	P05787	Keratin, type II cytoskeletal 8	8	11	-3	2.473	3.552	-1.079	0.085
LEGLTDEINFLR	P05787	Keratin, type II cytoskeletal 8	8	10	-2	3.537	4.507	-0.970	0.077
LNDLEEALQQAKEDLAR	P35908	Keratin, type II cytoskeletal 2 epidermal	3	5	-2	2.181	2.772	-0.591	0.071
LNDLEDALQQAKEDLAR	H6VRF8	Keratin 1	10	12	-2	4.303	5.577	-1.275	0.066
FSSSGGGGGGGRFSSSSGYGGGSSR	P35527	Keratin, type I cytoskeletal 9	4	5	-1	2.511	2.890	-0.379	0.060
SLDMDSIIAEVK	P05787	Keratin, type II cytoskeletal 8	10	10	0	5.456	6.339	-0.883	0.058
DYQELMNTK	H6VRF8	Keratin 1	3	4	-1	1.637	1.998	-0.362	0.053
LVSESSDVLPK	P05787	Keratin, type II cytoskeletal 8	8	9	-1	3.132	3.682	-0.551	0.050
QLYEEEIR	P05787	Keratin, type II cytoskeletal 8	7	7	0	3.068	3.448	-0.380	0.047
TLQGLEIELQSQLSmK	P13646-3;P08727	Isoform 3 of Keratin, type I cytoskeletal 13 |Keratin, type I cytoskeletal 19 |Junction plakoglobin	6			3.827	4.474	-0.647	0.040
SLVNLGGSKSISISVAR	H6VRF8	Keratin 1	3	5	-2	2.120	2.506	-0.386	0.039
QISNLQQSISDAEQR	H6VRF8	Keratin 1	6	9	-3	2.444	2.922	-0.478	0.035
NYSPYYNTIDDLK	P35527	Keratin, type I cytoskeletal 9	3	4	-1	2.482	2.487	-0.004	0.035
FSSSSGYGGGSSR	P35527	Keratin, type I cytoskeletal 9	2	2	0	1.771	1.944	-0.173	0.031
SISISVAGGGGGFGAAGGFGGR	P35908	Keratin, type II cytoskeletal 2 epidermal	4	5	-1	2.474	2.666	-0.192	0.030
NMQDmVEDYR	H6VRF8	Keratin 1 n		3		1.989	2.071	-0.082	0.028
TSQNSELNNMQDLVEDYKK	P35908	Keratin, type II cytoskeletal 2 epidermal	4	5	-1	3.108	3.166	-0.058	0.027
SGFSSVSVSR	B4DRR0	Keratin, type II cytoskeletal 6A	10	5	5	2.128	2.410	-0.282	0.024
TSFTSVSR	P13647	Keratin, type II cytoskeletal 5	5	3	2	1.688	1.695	-0.007	0.022
KDIENQYETQITQIEHEVSSSGQEVQSSAK	P35527	Keratin, type I cytoskeletal 9	6	8	-2	3.420	3.883	-0.463	0.019
TKFETEQALR	Q04695;P08727	Keratin, type I cytoskeletal 17 |Keratin, type I cytoskeletal 19 | Junction plakoglobin |Junction plakoglobin	16	10	6	4.063	4.253	-0.191	0.019
HGGGGGGFGGGGFGSR	P35908	Keratin, type II cytoskeletal 2 epidermal	3	5	-2	2.001	2.017	-0.016	0.018
LESGmQNmSIHTK	P05787	Keratin, type II cytoskeletal 8	6			3.420	3.749	-0.329	0.016
QEIEcQNQEYSLLLSIK	P35527	Keratin, type I cytoskeletal 9	4	6	-2	2.906	2.932	-0.027	0.016
RVLDELTLTK	P13645	Keratin, type I cytoskeletal 10	3	4	-1	1.799	1.954	-0.155	0.015
HGVQELEIELQSQLSKK	P35527	Keratin, type I cytoskeletal 9	4	5	-1	2.489	2.734	-0.245	0.015
SQYEQLAEQNRK	P13645	Keratin, type I cytoskeletal 10	6	13	-7	2.181	2.446	-0.266	0.013
NHKEEmSQLTGQNSGDVNVEINVAPGK	P35527	Keratin, type I cytoskeletal 9		6		3.309	3.365	-0.057	0.011
QGVDADINGLR	P35527	Keratin, type I cytoskeletal 9	7	8	-1	3.318	3.485	-0.167	0.011

Chapter 3: Study the Metabolic Regulation of Protein Synthesis between High and Low Invasive Head and Neck Cancer Cells

Introduction

Head and neck squamous cell carcinoma (HNSCC) arises from the upper aerodigestive tract (oral cavity, oropharynx, hypopharynx and larynx) and is the sixth most common type of cancer worldwide (Chung, Parker et al. 2004, Kamangar, Dores et al. 2006). In 2017, it is estimated that 49,670 new cases cancer of the oral cavity and pharynx in the United States, with approximately 9,700 deaths from the disease (Siegel, Miller et al. 2017). Despite the advances in diagnostics and treatment, 50% of patients with HNSCC have detectable node involvement due to local cancer metastasis, and less than 40% of these patients survive five years (Chung, Parker et al. 2004, Arellano-Garcia, Li et al. 2010). Therefore, early detection is important to improve the clinical outcome of the disease.

With the introduction of the “omics” technologies (genomics, proteomics and metabolomics), now we have global profiling methodologies to identify potential biomarkers for early detection of HNSCC (Yonezawa, Nishiumi et al. 2013). Metabolomics is the study of the chemical processes related to metabolites in a biological system. The identification of biomarkers by metabolomics can be performed by using two different technologies, nuclear magnetic resonance spectroscopy (NMR) (Zhang, Liu et al. 2011, Wang, Chen et al. 2013, Yang, Wang et al. 2013) and mass spectrometry (MS) (Wu, Xue et al. 2009, Wei, Xie et al. 2011, Xu, Chen et al. 2013). In a recent study, the metabolite profiles were generated from the plasma samples of patients with esophageal squamous cell carcinoma (ESCC) before, during, and after chemoradiotherapy (CRT) as well as healthy controls by LC-MS in order to identify potentially applicable markers for the disease diagnosis and monitoring treatment effects. The comparative analysis between pre- and post-treatment patients generated 11 metabolites as potential therapeutic biomarkers and 18 altered metabolites in ESCC patients compared to healthy

controls. These biomarkers may be useful in diagnosis, as well as in monitoring therapeutic responses and predicting outcomes of the ESCC (Xu, Chen et al. 2013). In 2008, Yan et al., demonstrated a metabolomics-based diagnostic approach for OSCC and its precancerous lesions, including oral lichen planus (OLP) and oral leukoplakia (OLK). Their results indicate that metabolomics profiling of saliva samples may properly represent the pathologic characteristics of OSCC, OLP and OLK, suggesting that metabolomics might be a potential approach for detection of cancers and precancers in the oral cavity (Yan, Wei et al. 2008). Similar LC-MS approach was also demonstrated, by Wei et al., for saliva metabolite profiling of 37 OSCC patients, 32 oral leukoplakia (OLK) patients and 34 healthy subjects. The OSCC, OLK and healthy control groups demonstrate characteristic metabolic signatures in their saliva samples. A panel of five salivary metabolites including γ-aminobutyric acid, phenylalanine, valine, *n*-eicosanoic acid and lactic acid were selected and their predictive power was evaluated for OSCC detection. Valine, lactic acid and phenylalanine in combination yielded satisfactory accuracy (0.89, 0.97), sensitivity (86.5% and 94.6%), specificity (82.4% and 84.4%) and positive predictive value (81.6% and 87.5%) in distinguishing OSCC from the controls or OLK, respectively. These results suggest that metabolomics approach may be complementary to the clinical detection of OSCC and may lead to an improved disease diagnosis (Wei, Xie et al. 2011). In a recent study by Sandulache et al., a panel of 15 HNSCC cell lines was assayed for glucose and glutamine dependence and sensitivity to metabolic inhibitors. In addition, LC-MS based metabolomic analysis was combined with individual measurements of reducing potential, adenosine triphosphate, and lactate production to characterize cellular metabolic phenotypes. Their studies revealed that HNSCC tumor cells are dependent on glucose, but not glutamine, for energy production and survival. This provides a rationale for treatment strategies of head and neck cancer that target glucose catabolism (Sandulache, Ow et al. 2011).

In this chapter, we propose to use the developed methodology to study the metabolic regulation of protein synthesis between highly (UM1 and UM-SCC5) and low invasive (UM2 and

UM-SCC6) cells. Both UM1 and UM2 cell lines were established from the same tongue cancer patient (Nakayama, Sasaki et al. 1998), but UM1 cells are significantly more invasive than UM2 cells. UM-SCC6 is derived from squamous carcinoma anterior tongue and is considered less metastatic than UM-SCC5 (Bonner, Maihle et al. 1994, Brenner, Graham et al. 2010). Our studies have demonstrated significant differences in metabolic phenotypes, particularly regarding protein synthesis, between high and low invasive HNSCC cells lines.

Materials and Methods

Cell Culture

The cell cultures were divide into two groups:

Group #1 (unlabeled ^{12}C): Oral and head/neck cancer cell lines UM1, UM2, UM-SCC5, and UM-SCC6 were cultured in Dulbecco's modified eagle medium (DMEM) with 10% fetal bovine serum and 1% penicillin/streptomycin (Invitrogen, Carlsbad, CA, USA). All cells were maintained at conditions of 5% CO_2-95% O_2 and 37°C and washed with PBS prior to cell lysis with Radio-Immunoprecipitation Assay (RIPA) buffer containing 1% Nonidet P-40, 0.1% SDS, 50mM Tris-HCL (pH 7.4), 150mM NaCl, 0.5% Sodium Deoxycholate and 1mM EDTA. Cell lysates containing 25ug of total proteins were treated with 100mM dithiothreitol (DTT) for 1hr to reduce protein disulfide bonds, and then treated with 150 mM iodoacetic acid (IAA) alkylating agent in 200mM NH_4HCO_3 for 1hr in the dark to prevent re-formation of disulfide bonds. Protein samples were digested with 10 ng/μL enzyme-grade trypsin (Promega, Madison, WI) in 200 mM NH_4HCO_3 at 37°C incubation for 24 hours. After proteolytic cleavage, samples were spun down at 14000g. The supernatant was transferred and dried. The peptide samples were dissolved in 0.1% formic acid and transferred to sample injection vials. Each experiment was in triplicate.

Group #2 (labeled ^{13}C): UM1, UM2, UM-SCC5, and UM-SCC6 cell lines were cultured in glucose-free DMEM (GIBCO) supplemented with 4.5g/L [U-$^{13}C_{6,}$ 99%] D-glucose (Cambridge Isotope Laboratories, Inc.), 10% fetal bovine serum and 1% penicillin/streptomycin (Invitrogen, Carlsbad, CA, USA) and maintained at 37°C in a humidified atmosphere of 5% CO2-95% O_2. All cell lines were treated with supplemented medium of [U-13C6, 99%] D-glucose for 3 days and washed with PBS prior to cell lysis with Radio-Immunoprecipitation Assay (RIPA) buffer containing 1% Nonidet P-40, 0.1% SDS, 50mM Tris-HCL (pH 7.4), 150mM NaCl, 0.5% Sodium Deoxycholate and 1mM EDTA. Cell lysates containing 25ug of total proteins were treated with 100mM dithiothreitol (DTT) for 1hr to reduce protein disulfide bonds, and then treated with 150

mM iodoacetic acid (IAA) alkylating agent in 200mM NH_4HCO_3 for 1hr in the dark to prevent re-formation of disulfide bonds. Protein samples were digested with 10 ng/μL enzyme-grade trypsin (Promega, Madison, WI) in 200 mM NH_4HCO_3 at 37°C incubation for 24 hours. After proteolytic cleavage, samples were spun down at 14000g. The supernatant was transferred and dried. The peptide samples were dissolved in 0.1% formic acid and transferred to sample injection vials. Each experiment was in triplicate.

Liquid Chromatography with Tandem Mass Spectrometry

The samples were fractioned and each fraction was desalted using the PepClean C18 spin columns (Pierce, Rockford, IL, USA) and further analyzed by RPLC-MS/MS on the Orbitrap Velos instrument. The peptide samples were loaded onto a C18 trap column at a flow rate of 30 ml/min. Then, the peptides were separated on a C18 nano column (75 mm 150 mm, 3 mm, 100 A˚ , C18; Dionex, Sunnyvale, CA, USA) at a flow rate of 250 nl/min on an UltiMate 3000 LC system (Dionex) using 3 h gradient. The MS/MS spectra were acquired on the Orbitrap Velos mass spectrometer using a data-dependent analysis mode, which the top 10 most abundant ions in each MS scan (m/z 350–2000) were selected for MS/MS analysis. The capillary temperature was set to 275C and the spray voltage set to 2 kV. The resolution for MS scan is 60,000 full width at half maximum (FWHM) and 7500 FWHM for MS/MS scan. The lock mass was used for accurate mass measurement. The MS/MS spectra were searched against the UniRef100 human database (120,982 entries) using the Mascot Daemon (Version 2.3). The mass tolerance was 10 ppm for MS and 0.1 Da for MS/MS. Database search was performed using Mascot database and Proteome Discoverer search engine against UniProt.Database search was performed using Mascot database and Proteome Discoverer search engine against UniProt. We re-acquired MS/MS spectra of UM1, UM2, UM-SCC5, and UM-SCC6 cell lines with “Exclude isotopes” turned off in the MS method on Q Exactive instrument.

Database Searching

Database search was performed using Mascot database (Matrix Science, London, UK; version 2.4.1) and Proteome Discoverer search engine against UniProt. We re-acquired MS/MS spectra of CSCs, NCSCs, MIA PaCa-2 and HPDE cell lines with “Exclude isotopes” turned off in the MS method on Q Exactive instrument. Mascot was searched with a fragment ion mass tolerance of 0.100 Da and a parent ion tolerance of 10.0 PPM. Oxidation of methionine, acetyl of the n-terminus and carbamidomethyl of cysteine were specified in Mascot as variable modifications.

Criteria for protein identification -- Scaffold (version Scaffold_4.4.8, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability by the Peptide Prophet algorithm (Keller, Nesvizhskii et al. 2002) with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii, Keller et al. 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

Data Processing

The RAW files were converted to mzXML and also processed with the Proteome Discoverer. The mzXML files were then processed by using both msconvert and MM File Converter. The two converters were used to obtain the MS2 scan number, precursor (MS1) scan number, charge (z), precursor (MS1) m/z as well as the MS1 and MS2 spectra. The Proteome Discoverer with the Mascot database search engine was used to acquire peptide sequence, accession number, protein name, charge, peptide spectrum matches (PSM) ambiguity and MS2 scan number. Next, these parameters were all imported into the MATLAB and processed with our

developed bioinformatics tool to identify the average isotopomer distribution for each unique peptide sequence observed. See chapter 1 for further details on this methodology. In essence, the tool first identifies individual isotopomer distribution scans and then proceeds to combine the isotopomer distributions belonging to the same peptide sequence into an average profile.

Results

Identification and Analysis of Isotopomer Distributions in UM1, UM2, UM-SCC5 and UM-SCC6 cells

To compare the metabolic phenotypes of high and low invasive cells, we divided the cell cultures in two groups: for Group #1, we introduced stable isotope-labeled glucose, [U-$^{13}C_6$] glucose, in the culture media for UM1, UM2, UM-SCC5 and UM-SCC6 cells, and, in Group #2, we used regular glucose for the culture of UM1, UM2, UM-SCC5 and UM-SCC6 cells. Then, the cellular proteins were harvested and analyzed by LC-MS/MS. Peptide sequence and protein identification were conducted by searching the MS data against the UniProt database with the Proteome Discoverer searching engine. The database search with the Proteome Discoverer in ^{13}C labeled group indicated that 2,906 unique peptides were identified in UM1 cells, 3,724 unique peptides in UM2 cells, 2,627 unique peptides in UM-SCC5 cells and 3,106 unique peptides in UM-SCC6 cells. For unlabeled group, the total unique peptides found in UM1 was 1,517, in UM2 was 5,686, in UM-SCC5 was 3,170 and in UM-SCC6 was 5,757 (**Figure 26**).

Identification and analysis of isotopomer distribution was generated by our developed bioinformatics tool described earlier. The number of peaks in an isotopomer distribution is shown in the **Figure 27**. The histograms suggest that there are an increase in ^{13}C uptake since we observe extensive isotopomer distributions containing more than 4 peaks. The total number of isotopomer distribution for UM1 is 4,101 (>4 peaks, 91.7%), for UM2 the total is 5,214 (>4 peaks, 65.7%), for UM-SCC5, the total is 3,780 (>4 peaks, 85.1%) and, for UM-SCC6, the total is 4,588 (>4 peaks, 76%).

An excel table was created based on our bioinformatics tool in MATLAB which contains peptide sequence, accession number, protein descriptions, MS scan, number of peaks, abundance difference and chi-square distance calculations to quantify the differences between two cell phenotypes as well as the chi-square distances, which were normalized by the number

of peaks in the isotopomer distribution. The table also contains a display function to generate a mirror graphics showing the visual isotopomer distribution between two cell lines. **Table 7** and **Table 8** show the proteins with the top 5% chi-square distance normalized as well as the protein identification and peptide sequences.

Our next step was to analyze the differences between UM1/UM2 and UM-SCC5/UM-SCC6. We generated a mirror display for cellular proteins between UM1 and UM2 and between UM-SCC5 and UM-SCC6. Heat shock protein 90 (HSP90) beta was found to have normalized chi-square distance of 0.449 in UM1/UM2 with abundance sum difference of 4.489 and 0.654 in UM-SCC5 and UM-SCC6 with the abundance sum difference of 5.932 for UM-SCC5 (**Figure 28).** Moreover, in **Figure 29** is a mirror display of heat shock 70kDa with normalized chi-square distance is 0.130 (UM1/UM2) and 0.672 (UM-SCC5/UM-SCC6) and abundance sum difference of 0.781 (UM1/UM2) and 5.011 (UM-SCC5/UM-SCC6). This suggests that the cells use more glucose for synthesis of protein HSP90/HSP70 in UM1 than UM2 and UM-SCC5 than UM-SCC6.

Figure 30 shows phosphoglycerate kinase 1 (PGK-1) with normalized chi-square distance of 0.175 and sum abundance difference of 0.196 in UM1/UM2. For UM-SCC5/UM-SCC6 the same sequence we did not we observed a normalized chi-square distance. Therefore, there was a glucose ^{13}C utilization for synthesis of PGK-1 in UM1 cells. In **Figure 31** shows abundance sum difference of 1.101 in UM1/UM2 for alpha- enolase and normalized chi-square distance is 0.202 (UM1/UM2). For UM-SCC5/UM-SCC6 group did not show normalized chi-square distance. This illustrates that PGK-1 and α-enolase has more glucose utilized in UM1 than UM2 and UM-SCC5 than UM-SCC6. **Figure 32** vimentin in UM1/UM2 the abundance difference is 1.784 and normalized chi-square is 0.162 suggesting that there is more glucose utilized in UM1 than UM2.

Analysis of Enrichment Ratio of High vs Low Invasive Cells

The enrichment ratio was calculated by quantifying the intensity of ^{12}C and ^{13}C peaks of high and low invasive cell lines. The total number of enrichment ratios in UM1 was 719, in UM2

was 2,196, in UM-SCC5 was 861 and in UM-SCC6 was 2,440 (**Figure 33**). The negative values indicate that ^{12}C has greater area than ^{13}C which may suggest no incorporation of ^{13}C. **Figure 34** and **Figure *35*** present the number of peaks for isotopomer distributions between cells cultured with ^{12}C glucose and ^{13}C glucose. UM1, UM2, UM-SCC5 and UM-SCC6 cells cultured in ^{12}C glucose show approximately a rapid decay at 8 peaks compared with UM1, UM2, UM-SCC5 and UM-SCC6 cells cultured in ^{13}C-glucose that have extensive decay containing more than 20 peaks. This suggests an incorporation of ^{13}C into the backbone of cellular proteins through labeled glucose. **Figure 36** presents the total number of peptide sequences in UM1 (5,661), UM2 (10,537), UM-SCC5 (7,261) and UM-SCC6 (9,984) versus the total number of peptide sequences with enrichment ratio in UM1 with 719, UM2 with 2,195, UM-SCC5 with 860 and UM-SCC6 with 2,440.

The protein synthesis rates were calculated based on the enrichment ratios of the same peptide from cells grown with or without ^{13}C glucose and they were comparable to each other. **Figure 37A** shows the isotopomer patterns of a natural unlabeled (^{12}C) and labeled (^{13}C) peptide for HSP90. Unlabeled peptide displayed rapid isotopic decay pattern within 3 peaks (M0,M1,M2). Similarly, labeled ^{13}C of the same peptide display a slow decay of 12 isotopic peaks which is typically seen when there is ^{13}C incorporation into peptides, the enrichment ratio is 0.267 and the abundance sum difference is 6.127. This demonstrates the incorporation of ^{13}C glucose in the peptide sequence YESLTDPSKLDSGK. The same peptide sequence can be found in UM2 and UM-SCC6; however, the enrichment ratio was not calculated since we just had one of the abundance sum. In addition, the same protein can be found in different cell lines but with different peptide sequences. This shows that ^{13}C incorporation can be different among different peptide sequences of the same protein (HSP90) (**Table 2**).

Figure 37B shows isotopomer distribution of heat shock 70 kDa of unlabeled peptide display 3 peaks. Similarly, labeled ^{13}C display a slow decay of 14 isotopic peaks, the enrichment ratio of 0.228 and abundance sum difference of 8.674 in UM1. This demonstrates the

incorporation of ^{13}C glucose in the peptide sequence ATAGDTHLGGEDFDNR in UM1. The same peptide sequence is not found in UM2 and UM-SCC5 and UM-SCC6. This may suggests no incorporation of ^{13}C in UM2. UM-SCC5 and UM-SCC6.

Figure 38C shows isotopomer distribution of PGK1 of unlabeled peptide sequence display 6 peaks. Similarly, labeled ^{13}C demonstrates a slow decay of 16 peaks with enrichment ratio of 0.121 and abundance sum difference 6.268 in UM2. This demonstrates incorporation of 13C glucose in the peptide sequence ALESPERPFLAILGGAK in UM2. The same peptide sequence was not found in UM1. The same peptide sequence was found in UM-SCC5 to display 4 unlabeled peaks and a slow decay of 14 peaks in labeled peptides with enrichment ratio of 0.153 and abundance sum difference of 5.806 (**Figure 38D**); the same peptide sequence was not found in UM-SCC6. **Figure 39** shows isotopomer distribution of α- enolase of unlabeled display 3 peaks and labeled display slow decay of 6 peaks, enrichment ratio 0.128 and abundance sum difference 1.449. This demonstrates the incorporation of ^{13}C glucose in the peptide sequence LNVTEQEKIDK in UM2. In the cell lines UM1, UM-SCC5 and UM-SCC6 the peptide sequence was not found. This suggests that there is a glucose enrichment in UM2 than UM1/UM-SCC5 and UM-SCC6.

Discussion

Cancer cells have a high rate of glucose uptake and glycolytic metabolism. In recent years, this metabolic alteration of malignant cells has been observed in multiple types of cancer cells, and there has been a growing interest in developing anticancer drugs that inhibit glycolysis and other relevant metabolic processes. The purpose of this study is to investigate the differences in the utilization of glucose for protein synthesis between high and low invasive oral cancer cell lines.

HSPs are found in all cell types usually expressed at low levels. Under normal physiological conditions, many members of the HSP family are involved in protein synthesis. Four main members of the HSP family are known: HSP65, HSP70, HSP90, and HSP100. During carcinogenesis, HSPs have been reported to alter their expression levels, showing either an increase or a decrease (Ciocca, Adams et al. 1983). For instance, HSP70 expression in colorectal carcinoma and breast carcinoma has been significantly correlated with low differentiation and poor prognosis (Lazaris, Theodoropoulos et al. 1995, Lazaris, Chatzigianni et al. 1997), whereas in renal cell carcinoma it has been reported to be associated with good prognosis (Santarosa, Favaro et al. , Tavassol, Starke et al. 2011). In a study, HSP90 was found to be overexpressed in human HNSCC and in UMSCC cell lines (Friedman, Wise et al.). Furthermore, previous studies provide evidence that HSP90 activation by interferon may contribute to EGF-mediated protection against the apoptotic effects of interferon in HNSCC cells (Caraglia, Abbruzzese et al. 1999, Ahsan, Ramanand et al. 2012). In our study, HSP90 and HSP70 were found to have high glucose enrichment in UM1 and UM-SCC5, high invasive cell lines. Based on this finding, it seems to suggest that highly invasive HNSCC cells utilize more glucose for synthesis of HSP90, in order to facilitate progression, resistance and metastasis of high invasive cancer cells.

PGK-1 is a major enzyme in the glycolysis pathway. In carcinogenesis, its overexpression has been linked to the regulation of the Warburg effect, which describes increased aerobic glycolysis in tumor cells despite the presence of oxygen (Lu, Gao et al. 2015). PGK1 has been

shown to play multiple cellular roles related to tumor biology, although such functions have not been fully elucidated. PGK1 has disulfide reductase activity during tumor angiogenesis (Lu, Gao et al. 2015). Besides as a glycolytic pathway enzyme, PGK1 is also believed to be involved angiogenesis, DNA replication and repair and cancer metastasis. In addition, PGK1 is involved in the onset of malignancies, such as pancreatic cancer, liver cancer and gastric cancer (Sun, Liang et al. 2015). In our study, PGK1 was found to have high glucose enrichment in UM1 and UM-SCC5 cells, which indicates that PGK1 may have a role in the high invasive oral cancer cells.

In our study, α- enolase was found to have high glucose enrichment in highly invasive UM1 and UM-SCC5 cells. Alpha- enolase is one of the key enzymes that can catalyze the conversion of glycerol phosphate to phosphoenolpyruvate in the glycolytic process, and plays an important role in energy metabolism. α-Enolase gene expression correlates with tumor size and shorter disease-free interval in breast cancer and its higher expression is associated with poorer clinical outcomes in head and neck and non-small-cell lung cancer patients (De Paepe 2012). Furthermore, it was also found to express on the cell surface promoting metastasis and migration of tumor cells (Liu, Huang et al. 2015). Overexpression of α-enolase is associated with tumor development and represents a potential diagnostic and prognostic marker. Thus, our study suggests that highly invasive HNSCC cells utilize more glucose for synthesis of α-enolase which can promote their invasiveness.

Vimentin, a mesenchymal cell marker, associates with components of the cytoskeleton and membrane adhesions. Studies of human epithelial carcinomas, such as breast cancer, hepatocellular carcinoma, colon carcinoma, and prostatic adenocarcinoma, have shown that vimentin expression can be correlated with tumor invasion and a poor prognosis. Several previous publications have detected the expression of vimentin in oral squamous cell carcinoma patients or cell lines. In oral squamous cell carcinoma patients, tumors lacking or expressing a low level of vimentin were correlated with a better prognosis than tumors with high vimentin expression. Moreover, strong vimentin expression was found in the invading tumor cells at the invasive front

(Liu, Jiang et al. 2009). Vimentin was found to have higher glucose enrichment in UM1 than UM2 from our studies.

Conclusion

Cancer cells rewire their metabolism and energy production networks to enable rapid proliferation, invasion, metastasis, and resistance to cancer treatment. Our study has demonstrated that there is an important difference in glucose enrichment between high and low invasive head and neck cancer cells for protein synthesis. By ranking the glucose enrichment for protein synthesis, we could reveal the proteins in that cancer cells preferentially utilize nutritious resources (e.g., glucose) for their synthesis, and these proteins may be involved in cancer invasion and metastasis, such as HSP90 and 70, PGK-1, α-enolase and vimentin.

Limitations, Clinical Applications and Future Directions

One of the limitations of this study is that the accuracy of the searching algorithms may be affected by using carbon 13 glucose since the peptide fingerprint for a labeled protein is different from unlabeled protein. Therefore, a large number of proteins were not identified based on mass spectrometry and database searching. Future studies will need to improve the database searching engine so that more proteins can be identified for data analysis using our developed bioinformatics tool. Another potential limitation is the use of single tracer for tracer-based metabolomics in our study. In the future, multiple tracers can be used for tracer-based metabolomics analysis that will lead to a more in-depth understanding of the metabolic phenotypes of cancer cells.

By using this novel methodology, we will be able to rank proteins and metabolites in cancer cells according to their enrichment ratios and chi square distances. Those signature proteins may be used as biomarkers for diagnostic/prognostic applications or targets for therapeutic interventions. In addition, all the studies conducted in this dissertation are built on the use of *in vitro* cell line models. In the future, we plan to apply our novel methodology *in vivo* by infusion of ^{13}C-labeled glucose into mouse xenografts and extract the tumor tissues to perform tracer-based metabolomic analysis.

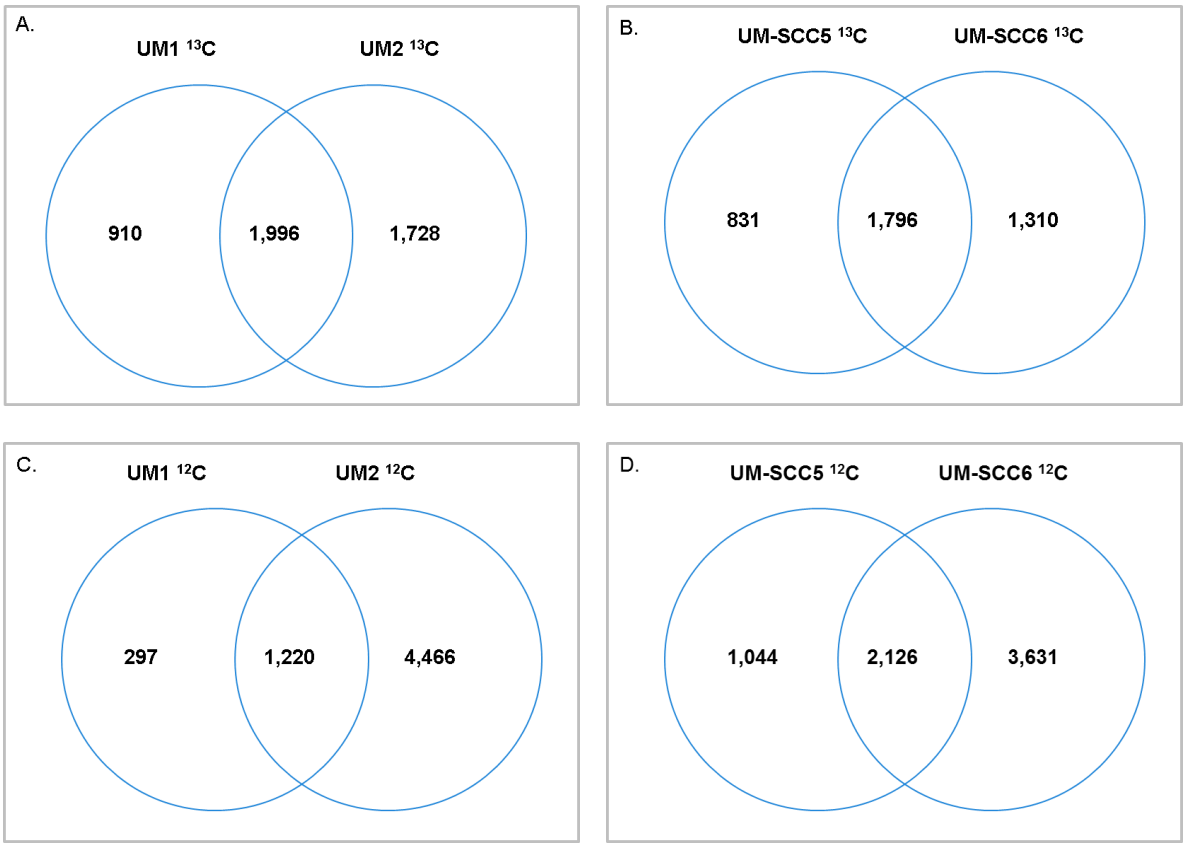
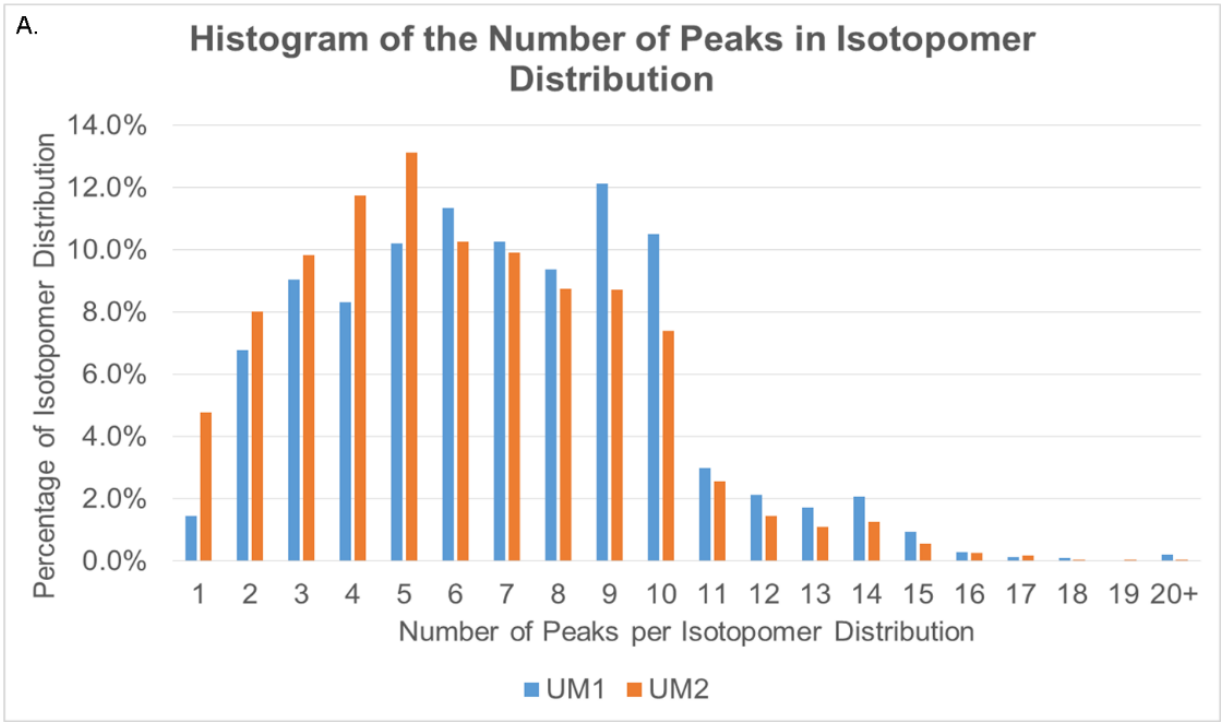


Figure 26: The Venn diagrams show the overlap of total unique peptide numbers found in UM1/UM2 and UM-SCC5/UM-SCC6 cells obtained by using scaffold software.



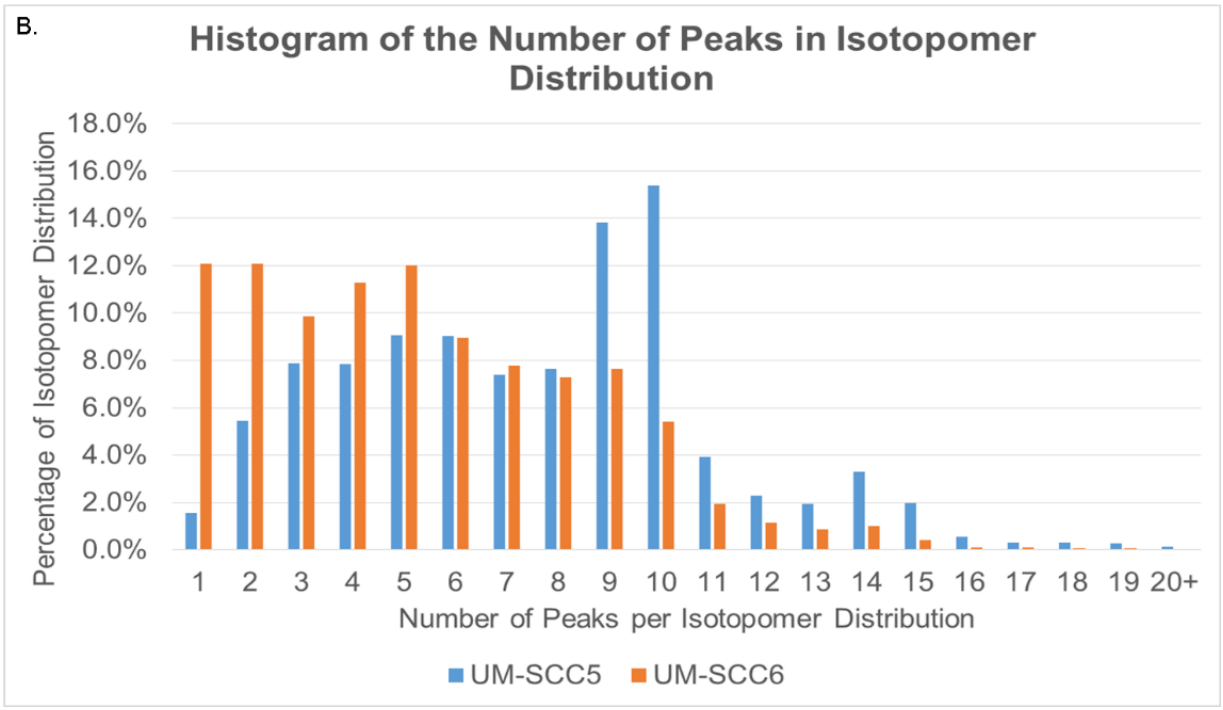


Figure 27: Histogram shows the total number of peaks per isotopomer distribution between UM1/UM2 and UM-SCC5/UM-SCC6 cells.

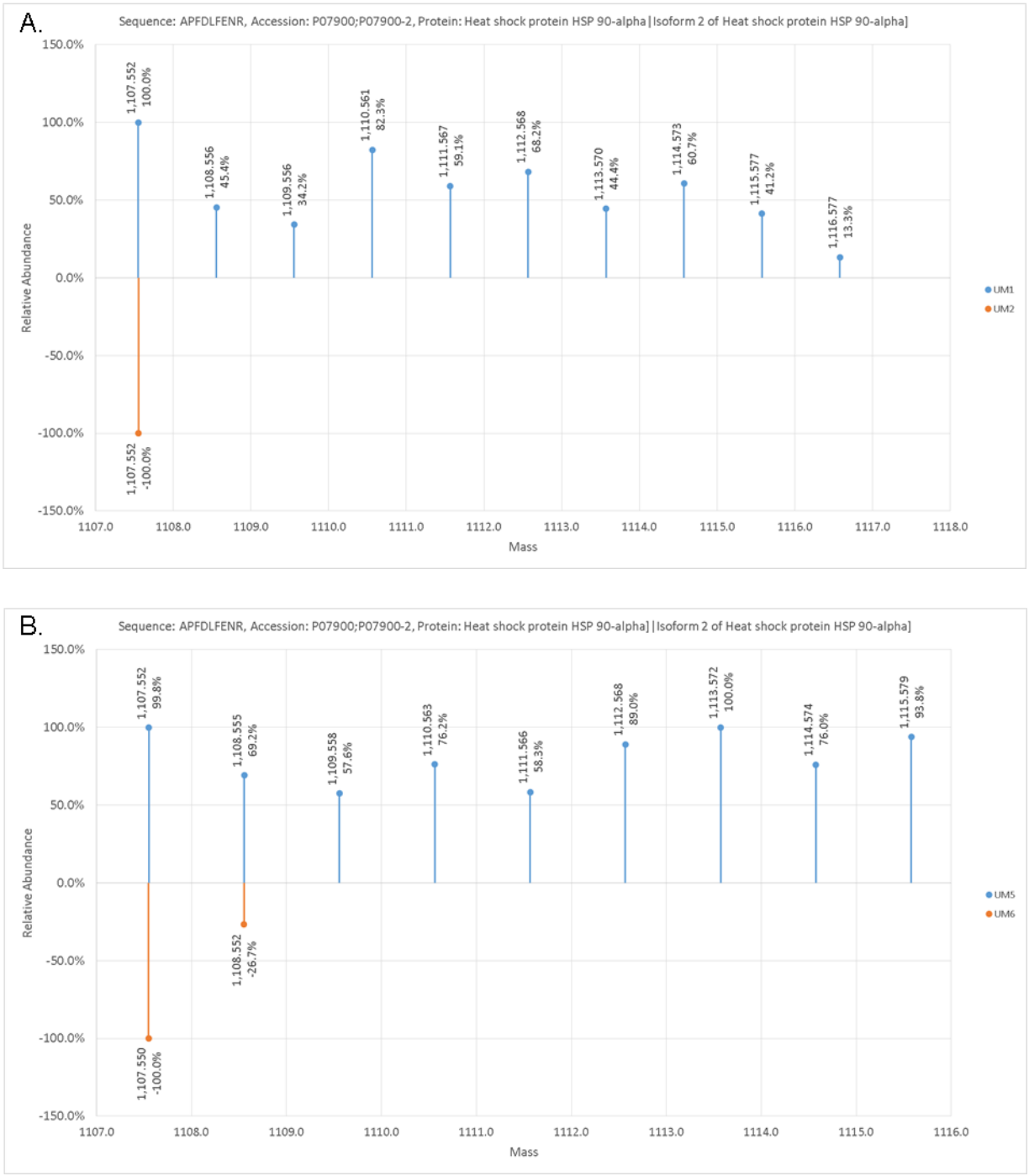
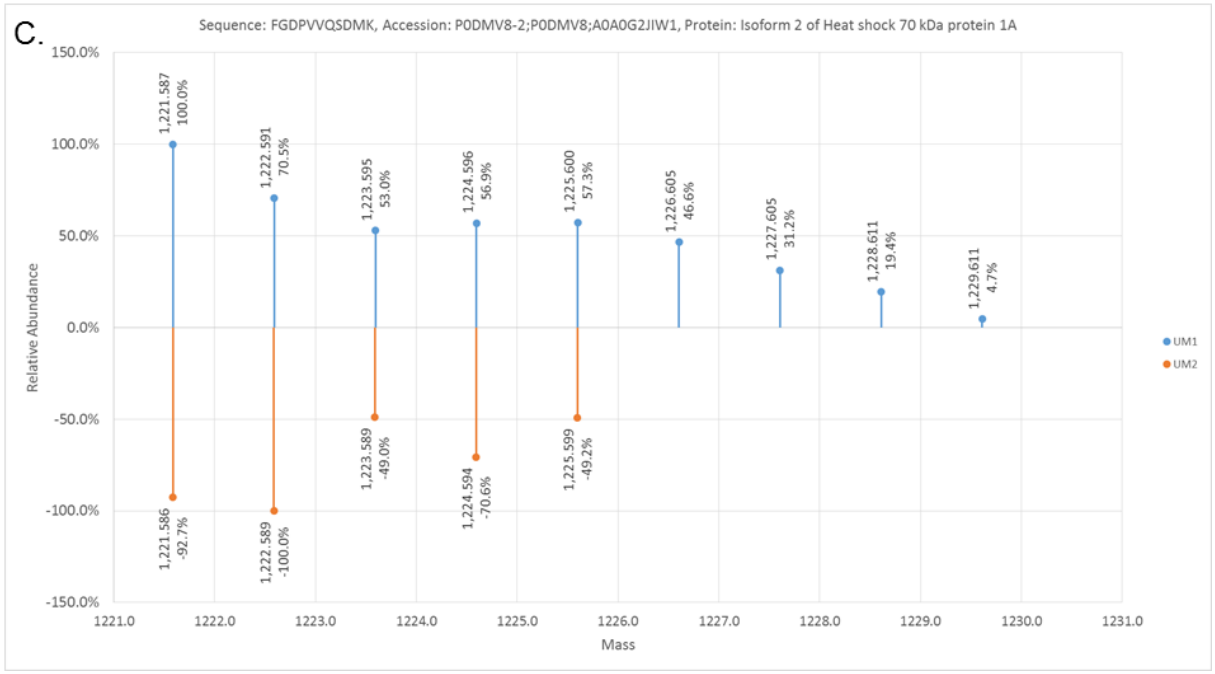


Figure 28: Comparison of the isotopomer distribution in HSP90 for UM1, UM-SCC5 cells (in blue) and UM2, UM-SCC6 cells (in orange). The mirror display shows high glucose utilization for the synthesis of protein HSP90 in UM1/UM-SCC5 cells compare with UM2/UM-SCC6 cells.



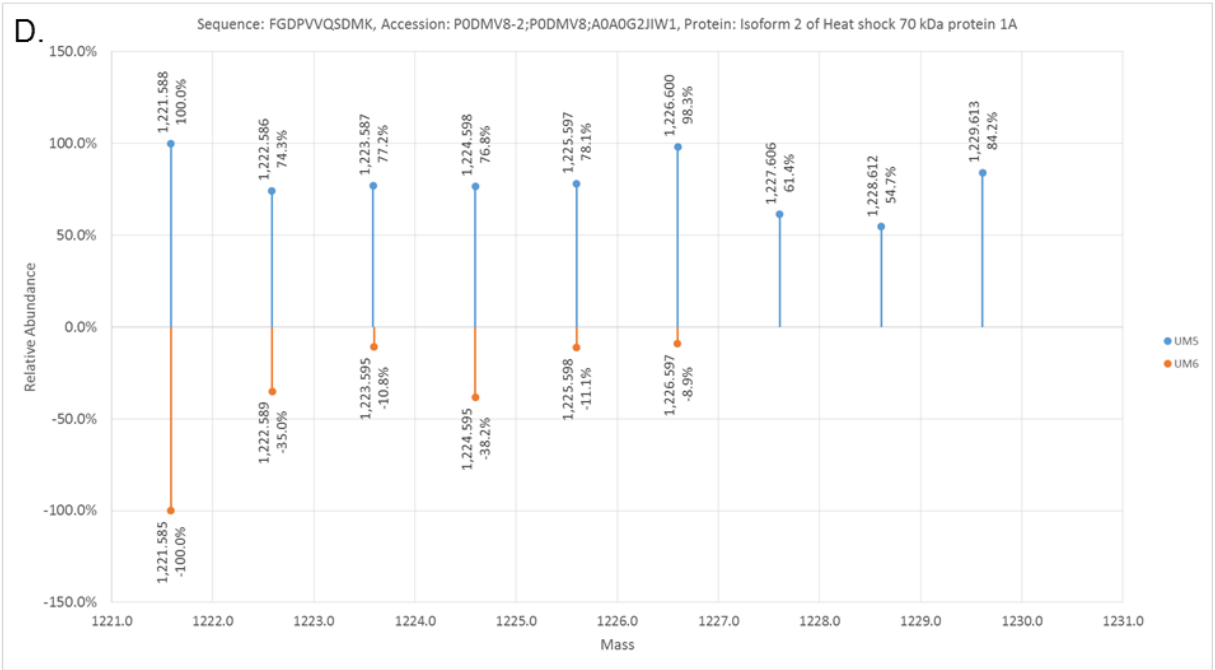


Figure 29: Comparison of the isotopomer distribution in HSP70 for UM1, UM-SCC5 cells (blue) and UM2, UM-SCC6 cells (orange). The mirror display shows high glucose utilization for the synthesis of protein HSP70 in UM1/UM-SCC5 cells compare with UM2/UM-SCC6 cells.

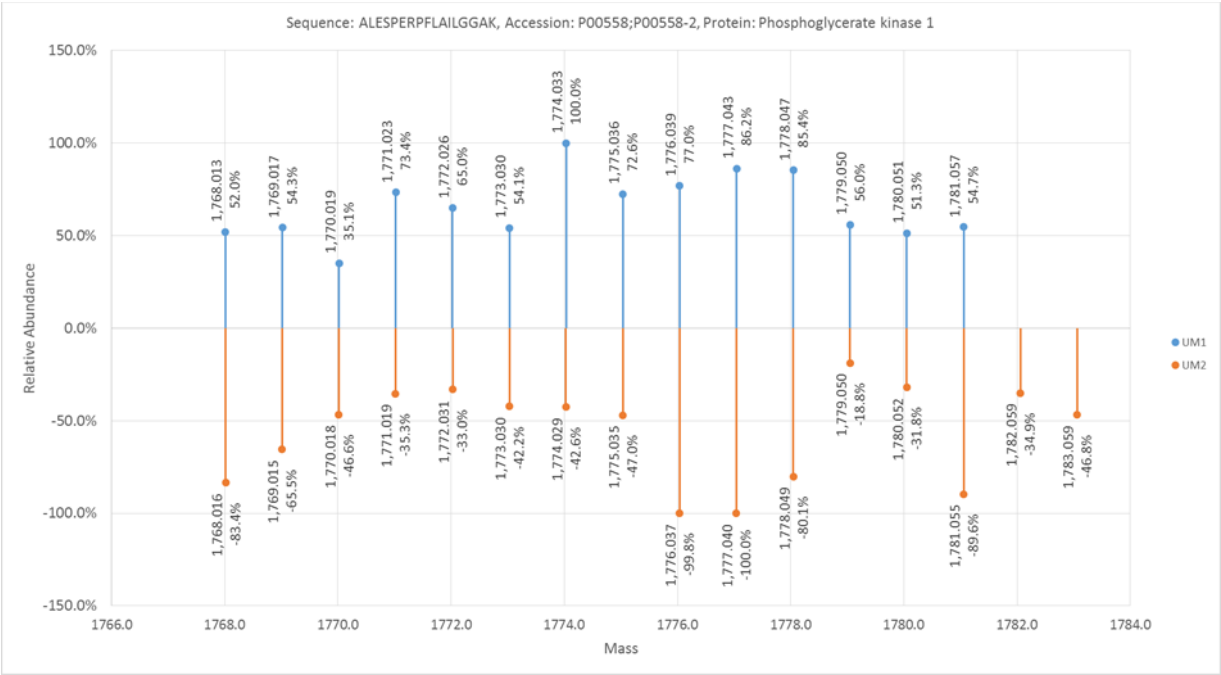


Figure 30: Comparison of the isotopomer distribution in PGK-1 for UM1 cells (blue) and UM2 cells (orange). The mirror display shows high glucose utilization for the synthesis of protein PGK-1 in UM1 cells.

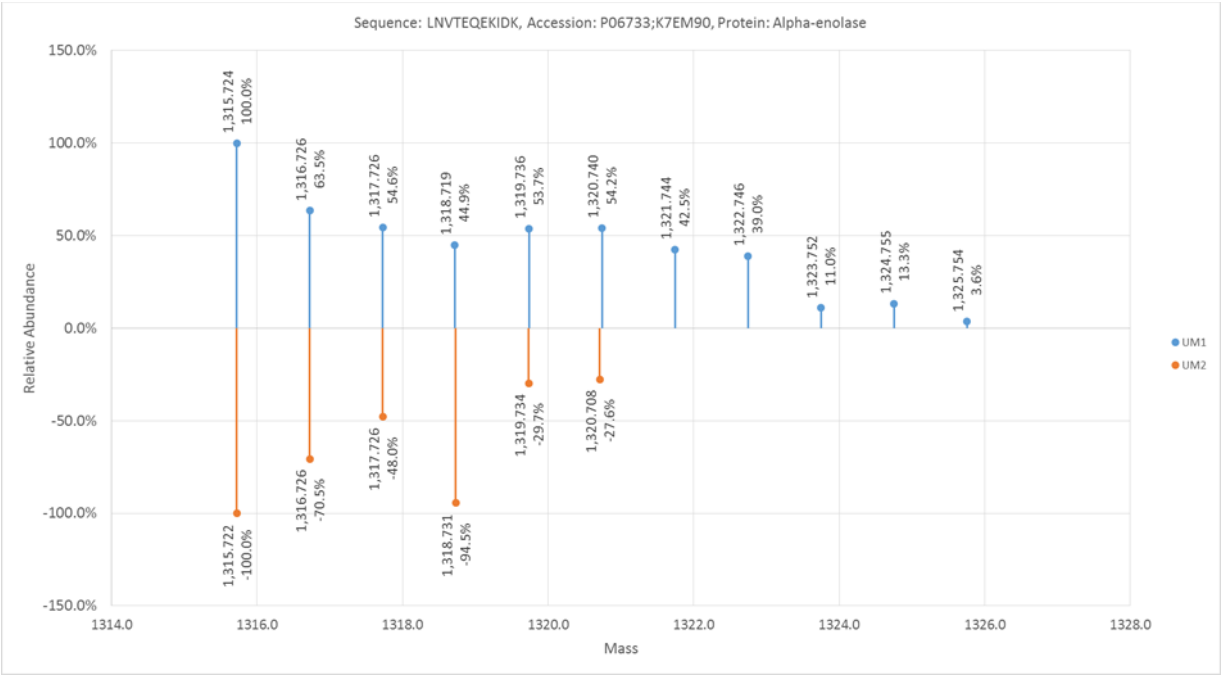


Figure 31: Comparison of the isotopomer distribution in α-enolase for UM1 cells (blue) and UM2 cells (orange). The mirror display shows high glucose utilization for the synthesis of protein α-enolase in UM1 cells.

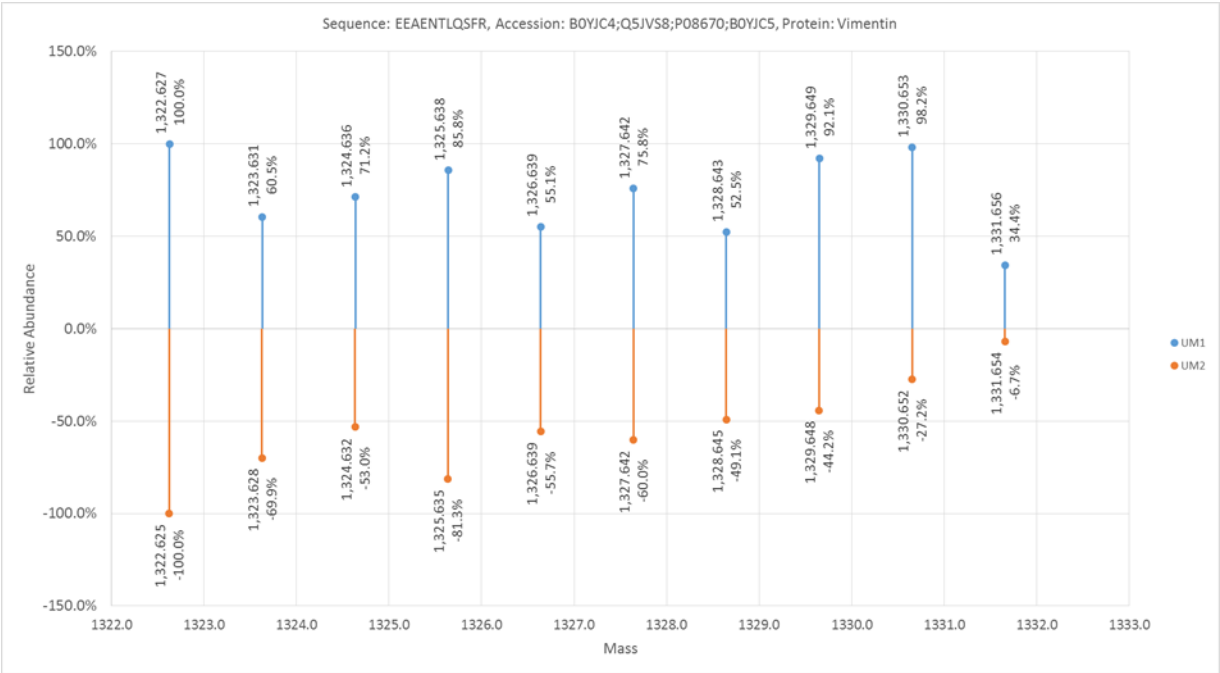


Figure 32: Comparison of the isotopomer distribution in vimentin for UM1 cells (blue) and UM2 cells (orange). The mirror display shows high glucose utilized for the synthesis of protein vimentin in UM1 cells.

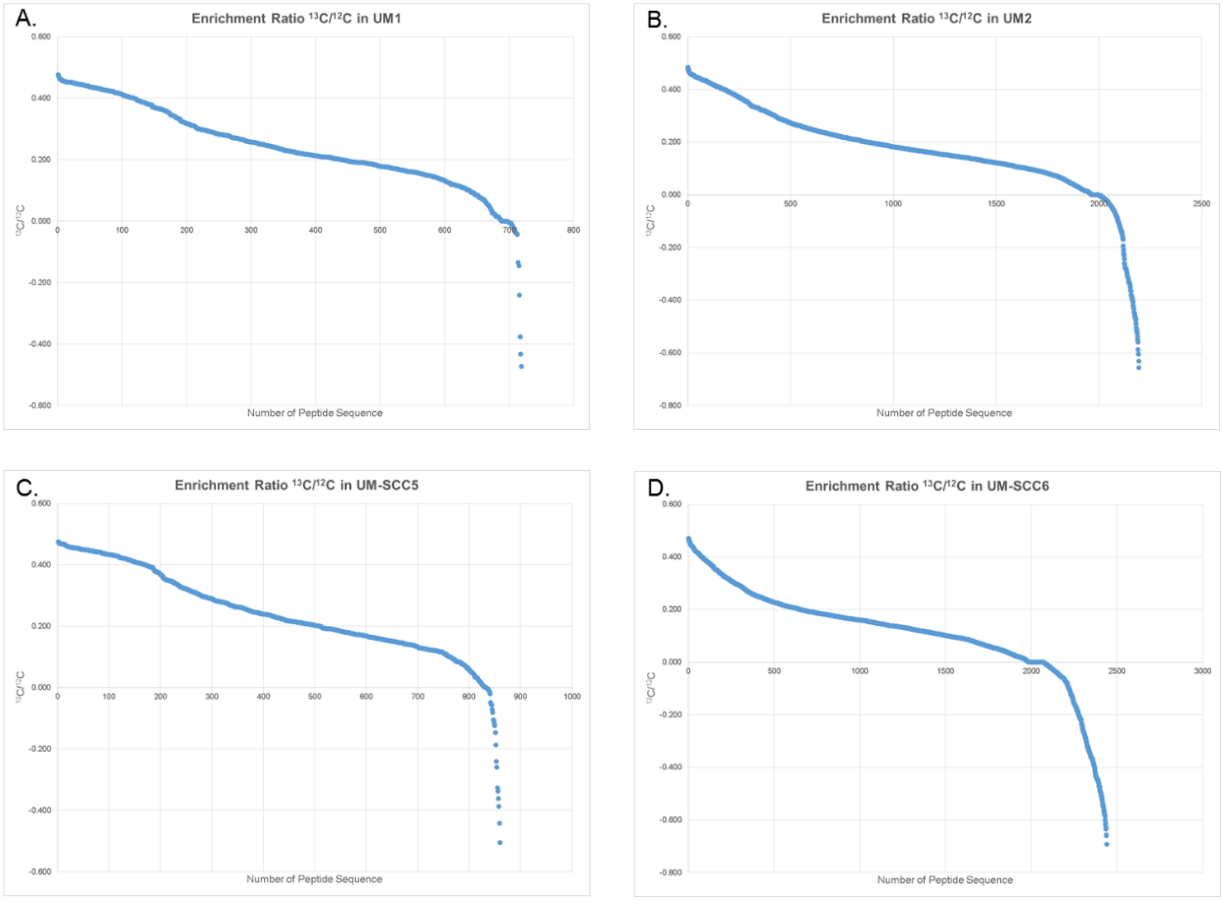
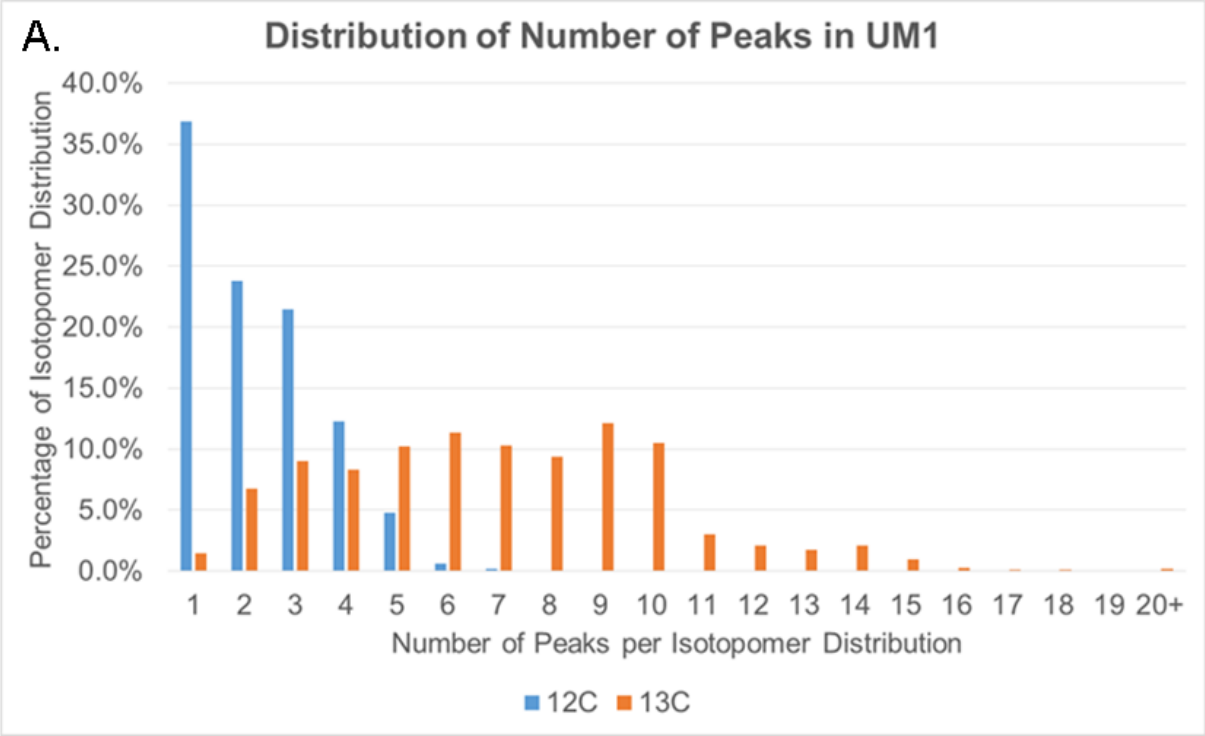


Figure 33: Quantification of enrichment ratio between the cells lines. This diagram represents the number of enrichment ratio of label and unlabeled ($^{13}C/^{12}C$) peptide sequences by percentage of glucose uptake.



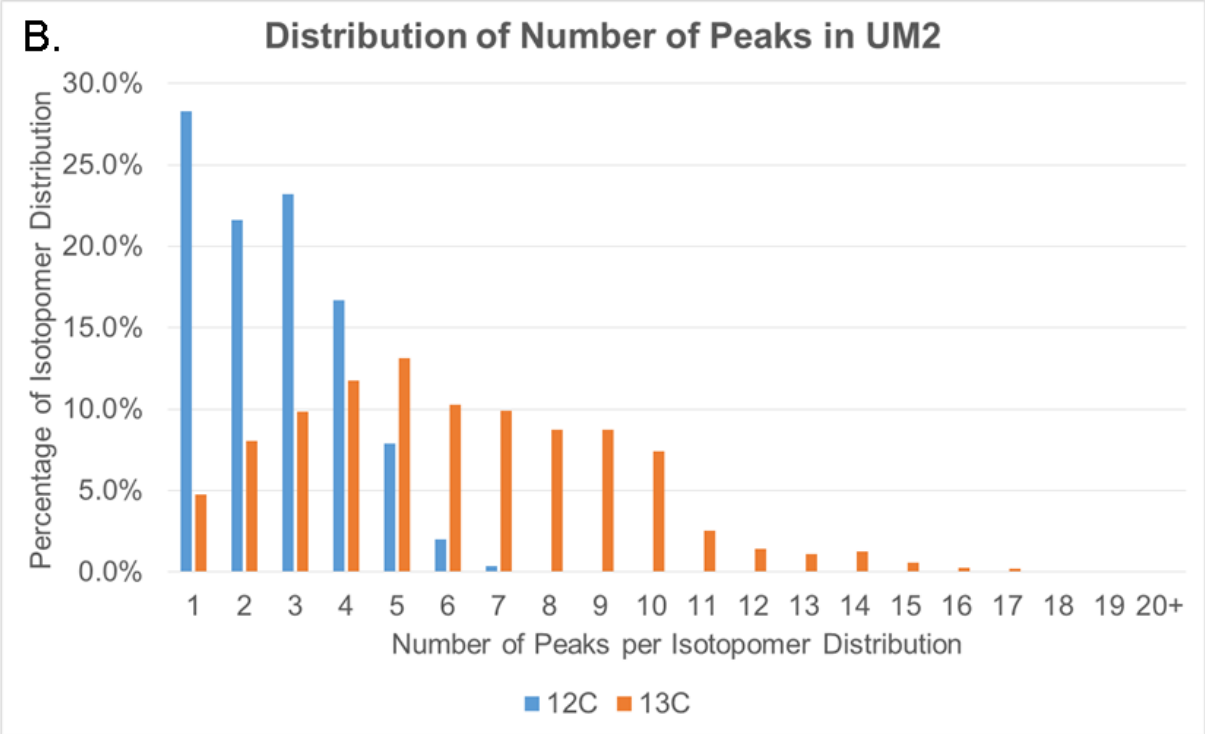
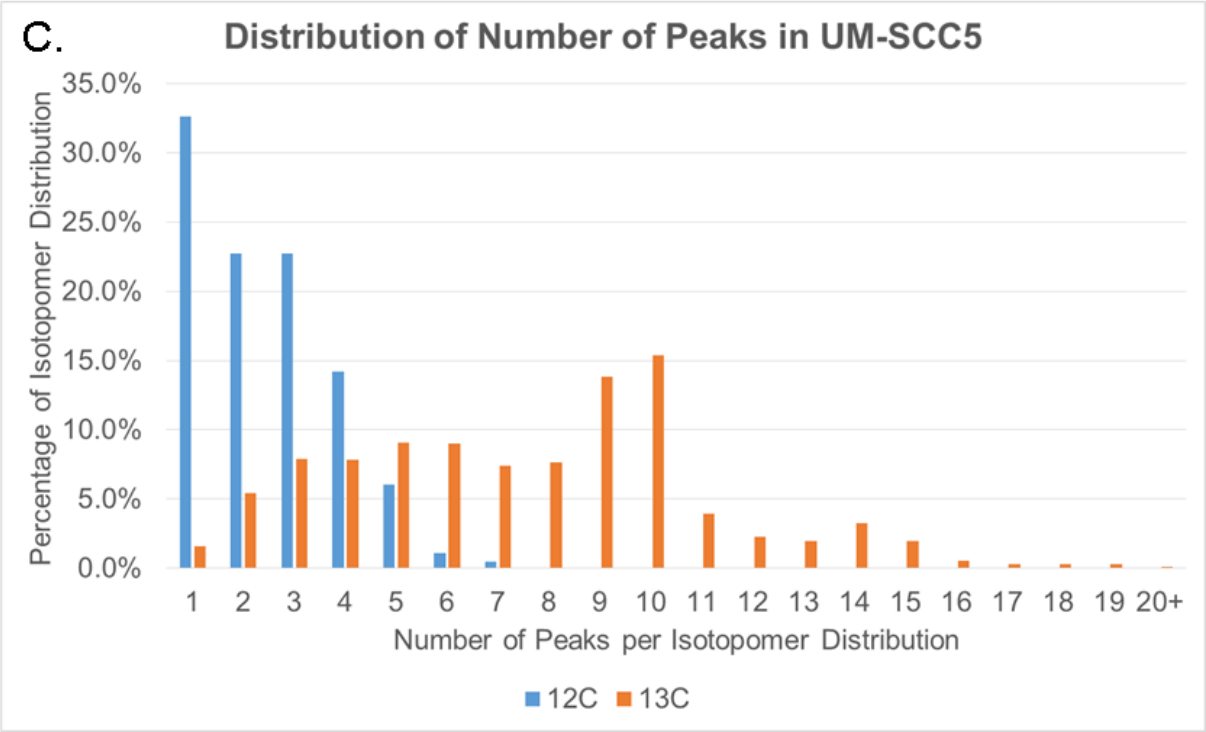


Figure 34 A& B: The histogram shows the distribution of the number of mass peaks for mass isotopomer pattern (12C versus 13C) in UM1 and UM2 cells.



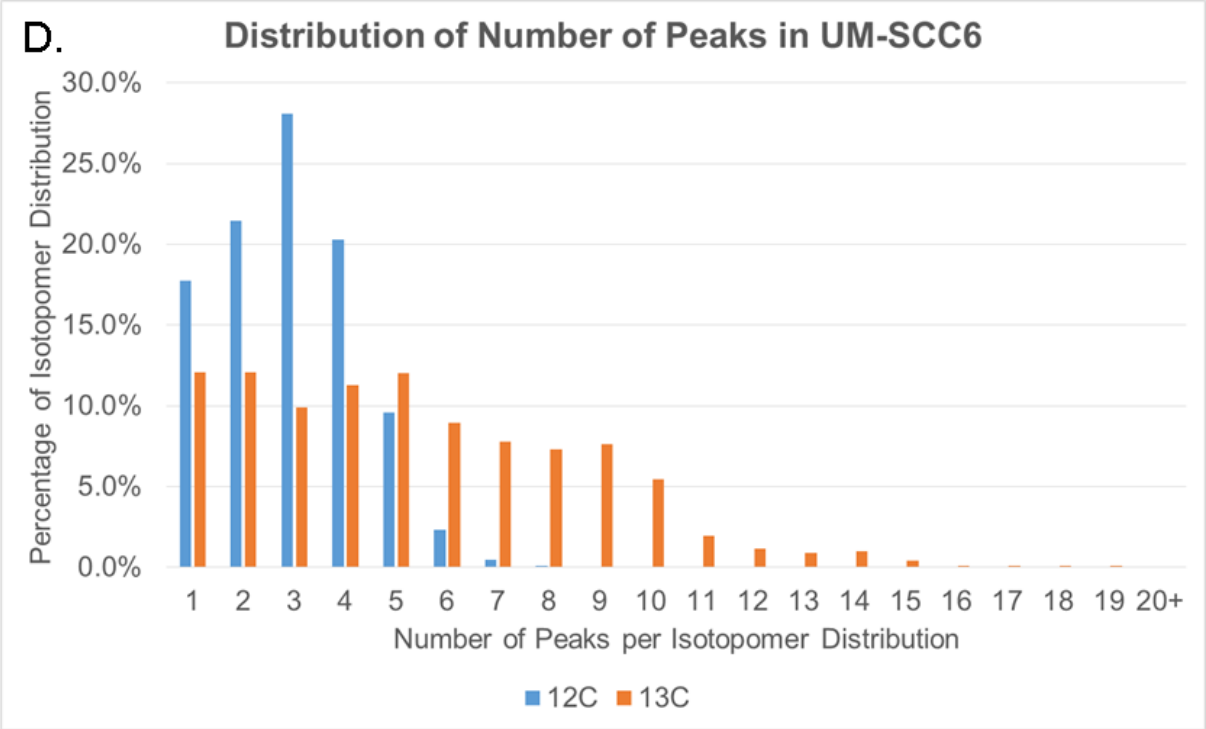


Figure 35 C & D: The histogram shows the distribution of the number of mass peaks for mass isotopomer (^{12}C versus ^{13}C) in UM-SCC5 and UM-SCC6 cells.

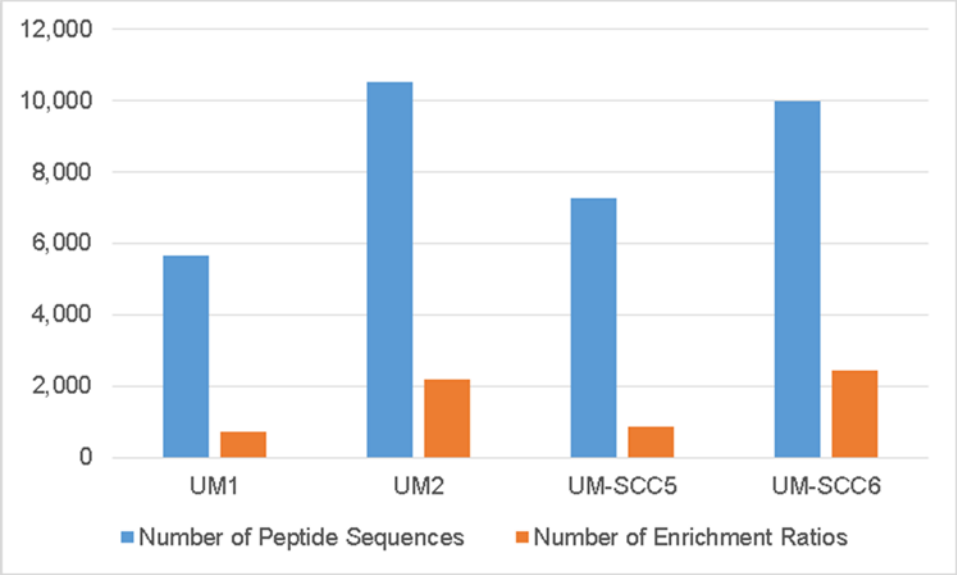


Figure 36: The figure quantifies the total number of peptides vs the total number of peptide sequences with enrichment ratio.

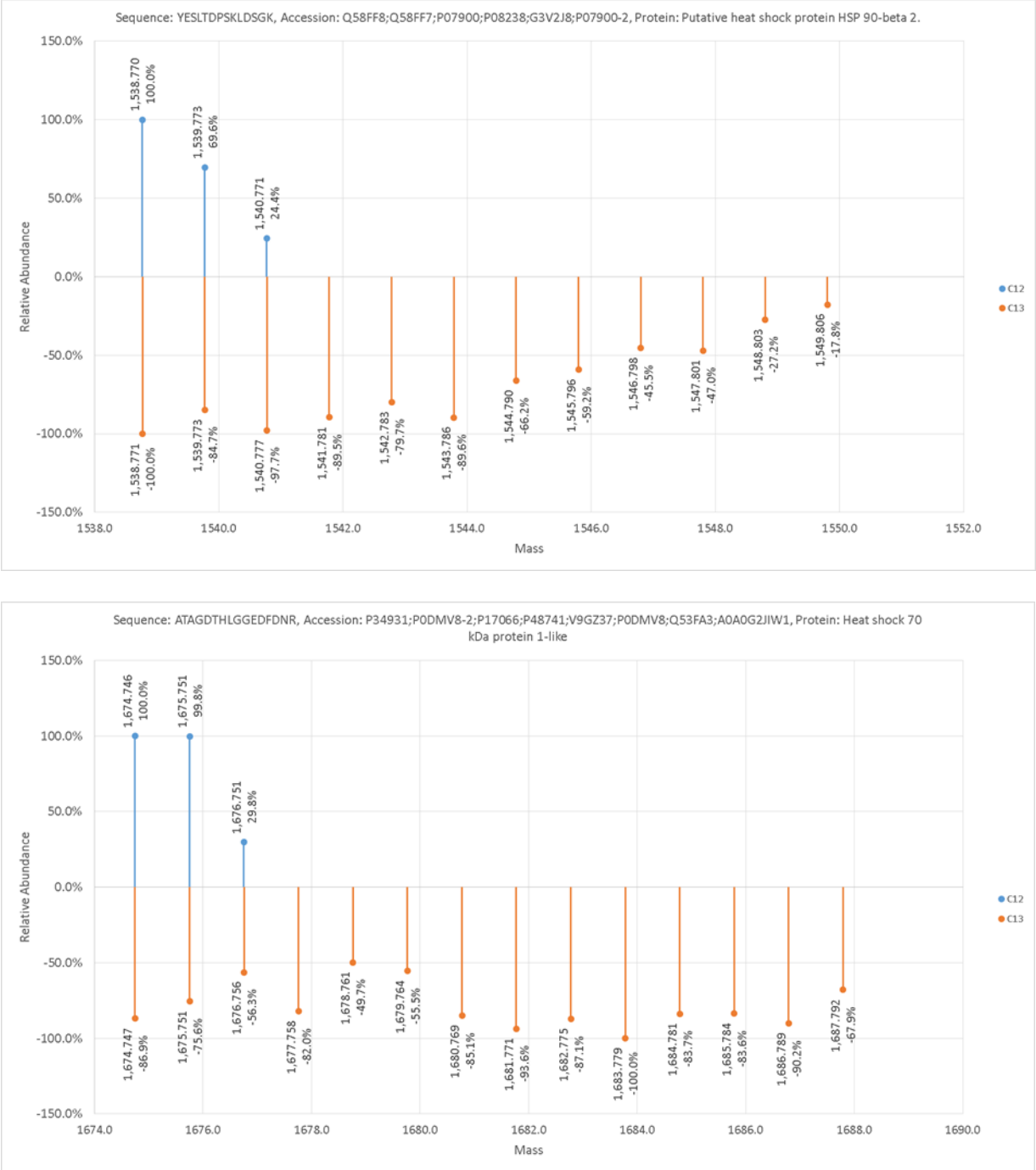


Figure 37: Comparison of ^{12}C unlabeled (blue) and ^{13}C labeled (orange) peptide isotopomer distribution for HSP90 (**A**) and HSP70 (**B**) in UM1 cells.

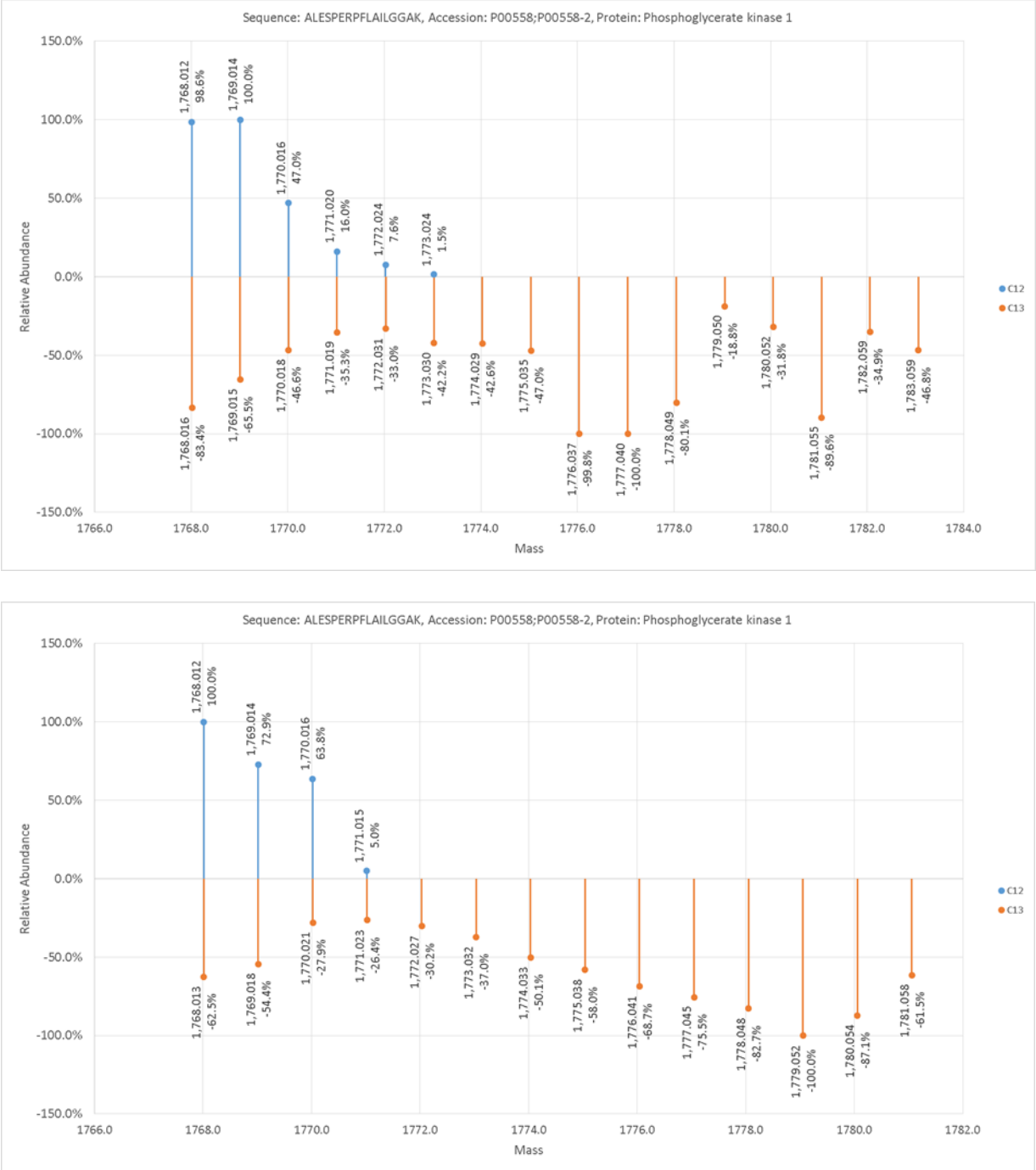


Figure 38: Comparison of ^{12}C unlabeled (blue) and ^{13}C labeled (orange) peptide isotopomer distribution for PGK-1 in UM2 (**C**) and UM-SCC5 (**D**) cells.

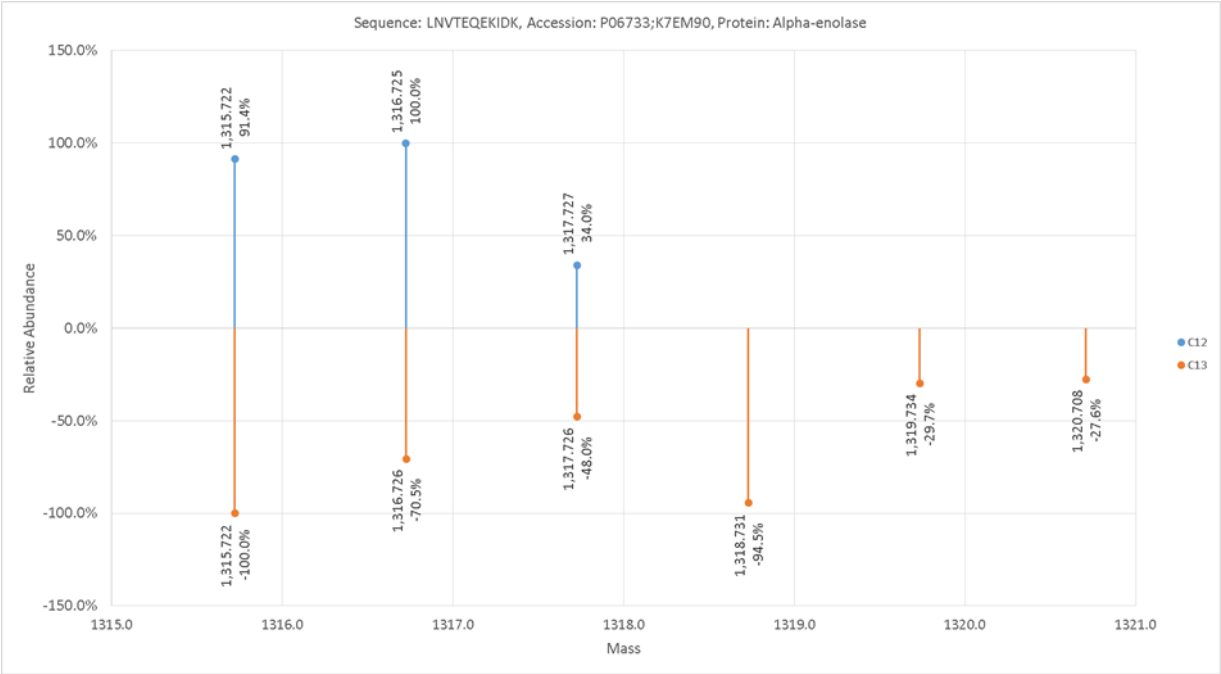


Figure 39: Comparison of ^{12}C unlabeled (blue) and ^{13}C labeled (orange) peptide isotopomer distribution of α- enolase in UM2 cells.

Table 7: ^{13}C labeled proteins in UM1/UM2 cells. Note: Abundance sum is the sum of relative abundance in an isotopomer distribution. Abundance sum difference is the difference between the abundance sums of two cell lines.

Sequence	Accessions	ProteinDescriptions	# Peaks UM1	# Peaks UM2	# Peaks Difference	Abundance Sum UM1	Abundance Sum UM2	Abundance Sum Difference	Chi-square normalized
ADTLTDEINFLR	P02538;P48668;P04259	Keratin, type II cytoskeletal	10	9	-1	7.486	2.715	4.772	0.872
aGTVVLDDVELR	P25205	DNA replication licensing factor MCM3	9	2	-7	6.932	1.674	5.257	0.701
AQYEEIAQR	P35908;Q01546;P19013;P02538;P48668;P04259	Keratin, type II cytoskeletal 2 epidermal |Keratin, type II cytoskeletal 2 oral	10	9	-1	6.235	2.047	4.188	0.700
VTLELGGK	O94788;O94788-3;O94788-2;P47895;P30837;P05091;P05091-2;H0YMG7;H0Y2X5;P00352;O94788-4	Retinal dehydrogenase 2|Aldehyde dehydrogenase family 1 member A3	3	4	1	1.651	1.663	-0.012	0.683
NKLEGLEDALQK	P02538;P48668;P04259	Keratin, type II cytoskeletal 6A	14	10	-4	9.951	3.384	6.566	0.677
TIVEEVQDGK	Q04695;F5GWP8	Keratin, type I cytoskeletal 17	9	1	-8	6.796	1.000	5.796	0.671
TAAENEFVTLK	P12035;P02538;P48668;P04259	Keratin, type II cytoskeletal 3	10	5	-5	7.843	2.303	5.540	0.662
ISLGLPVGAVINcADNTGAK	B9ZVP7;C9JD32;J3KT29;P62829	60S ribosomal protein L23	12	1	-11	8.441	1.000	7.441	0.661
AQFAQPEILIGTIPGAGGTQR	P30084	Enoyl-CoA hydratase, mitochondrial	12	3	-9	9.254	2.102	7.152	0.650
ASLENSLEETK	P02533;P08779;K7ENW6	Keratin, type I cytoskeletal 14	10	5	-5	7.324	2.036	5.288	0.649
SYELPDGQVITIGNER	Q562R1;Q9BYX7;Q6S8J3;A5A3E0;P62736;P60709;P68032;P63261;P63267;P68133;P63267-2	Beta-actin-like protein 2	10	3	-7	8.140	1.935	6.205	0.639
FAMEPEEFDSDTLR	P13667	Protein disulfide-isomerase A4	8	1	-7	5.649	1.000	4.649	0.628
VQEQVHTLLSQDQAQAAR	Q07065	Cytoskeleton-associated protein 4	14	2	-12	8.631	1.956	6.675	0.627
LQSSNIFTVAK	Q10567;Q10567-4;Q10567-3;Q10567-2	AP-1 complex subunit beta-1	10	2	-8	7.348	1.937	5.411	0.613
NLANTVTEEILEK	O60506-2;O60506-5;O60506;O60506-3;O60506-4;B7Z645	Isoform 2 of Heterogeneous nuclear ribonucleoprotein Q	11	1	-10	7.475	1.000	6.475	0.607
NADPILISLK	Q9ULV4;Q9ULV4-2;Q9ULV4-3;B4E3S0;H0YHL7	Coronin-1C	9	2	-7	6.799	1.862	4.937	0.604
LLEGEDAHLSSSQFSSGSQSSR	P02533	Keratin, type I cytoskeletal 14	15	4	-11	9.270	2.593	6.677	0.598
AVTEQGAELSNEER	P27348;E9PG15	14-3-3 protein theta	10	2	-8	7.064	1.435	5.629	0.597
AVYcIIGR	Q8N1F7;H3BPA9;H3BVG0;Q8N1F7-2	Nuclear pore complex protein Nup93	10	1	-9	6.971	1.000	5.971	0.597
DAEEWFFTK	P02533	Keratin, type I cytoskeletal 14	10	10	0	7.226	3.711	3.514	0.587
ETADTDTADQVMASFK	H0YJW3;P12814;P12814-2;P12814-3;P12814-4;H9KV75	Alpha-actinin-1	10	1	-9	6.843	1.000	5.843	0.584
LVILANNcPALR	E5RJH3;P62888;E5RI99	60S ribosomal protein	9	5	-4	7.591	2.420	5.172	0.584
DAGVIAGLNVLR	P34931;P0DMV8-2;P0DMV8;Q53FA3;A0A0G2JIW1	Heat shock 70 kDa protein 1-like	10	2	-8	6.257	1.439	4.819	0.576
SRAEAESWYQTK	P02538;P48668;O95678;P04259	Keratin, type II cytoskeletal 6A	14	7	-7	8.749	2.926	5.823	0.567
LLAPEAGSHPQQTQK	Q8NFQ8	Torsin-1A-interacting protein 2	2	12	10	1.365	7.725	-6.359	0.567
VLPPPAGYVPIR	O75533	Splicing factor 3B subunit 1	8	1	-7	5.344	1.000	4.344	0.565
AASADSTTEGTPADGFTVLSTK	O75367;O75367-2;O75367-3	Core histone macro-H2A.1	3	11	8	2.033	8.513	-6.480	0.560
aYAYLFK	E9PKL7;P61019;P61019-2	Ras-related protein Rab-2A	1	9	8	1.000	5.299	-4.299	0.554
DGTAGIPNLQLYDVK	Q9BY44;Q9BY44-3;C9IZE1;F8WAE5;H7C5Q3	Eukaryotic translation initiation factor 2A	8	3	-5	6.072	2.803	3.269	0.545
GQTLTLTQQQR	P98194;P98194-7;P98194-2;P98194-5;B4E2Q0;H0Y8X9;H0Y9S7;H0Y9V7;P98194-3;P98194-4;P98194-6;P98194-9;P98194-8	Calcium-transporting ATPase type 2C member 1	2	9	7	1.803	6.776	-4.973	0.544
AILVDLEPGTMDSVR	G3V2N6;G3V2R8;G3V3R4;G3V5W4;Q5JP53;A0A0B4J269;Q13885;Q9BVA1;Q13509;P07437	isoform CRA_d|Tubulin beta	12	2	-10	8.042	1.473	6.569	0.540
EGMNIVEAMER	P62937;P62937-2	Peptidyl-prolyl cis-trans isomerase A	10	2	-8	6.807	1.778	5.029	0.539

Table 8: ^{13}C labeled proteins in UM-SCC5/UM-SCC6 cells.

Sequence	Accessions	Protein Description	# Peaks UM-SCC5	# Peaks UM-SCC6	# Peaks Difference	Abundance Sum UM-SCC5	Abundance Sum UM-SCC6	Abundance Sum Difference	Chi-squared Normalized
LLSESAQPLKK	Q9H3N1	Thioredoxin-related transmembrane protein 1	2	1	-1	1.246	1.000	0.246	0.957
AQYEEIAQR	P35908;Q01546;P19013;P02538;P48668;P04259	Keratin, type II cytoskeletal 2 epidermal	10	9	-1	6.502	2.027	4.474	0.815
TAAENEFVTLKK	P12035;P02538;P48668;P04259	Keratin, type II cytoskeletal 3	14	13	-1	9.702	3.820	5.881	0.737
EQFLDGDGWTSR	K7EJB9;P27797	Calreticulin	10	2	-8	8.424	1.401	7.023	0.696
DLKPENLLLASK	Q13554;Q13557;Q13555;E7EQE4;E9PBE8;H7BXS4;Q9UQM7;Q13557;;Q13555;Q5SWX3;Q8WU40;Q13554	Calcium/calmodulin-dependent protein kinase type II subunit beta	2	1	-1	1.473	1.000	0.473	0.689
aTNIEQIFR	P18583;J3QSZ5	Protein SON	9	1	-8	5.643	1.000	4.643	0.673
FGDPVVQSDMK	P0DMV8-2;P0DMV8;A0A0G2JIW1	Isoform 2 of Heat shock 70 kDa protein 1A	9	6	-3	7.051	2.040	5.011	0.672
ASLENSLEETK	P02533;P08779;K7ENW6	Keratin, type I cytoskeletal 14	10	7	-3	7.406	2.033	5.373	0.670
GVPAGNSDTEGGQPGR	Q9UKV3;Q9UKV3-5;E7EQT4;G3V3B0;G3V3T3;Q9UKV3-2;Q9UKV3-3;S4R3H4	Apoptotic chromatin condensation inducer in the nucleus	8	1	-7	6.146	1.000	5.146	0.669
SINPDEAVAYGAAVQAAILSGDK	P11142-2;E9PKE3;E9PNE6;P11142;A8K7Q2	Isoform 2 of Heat shock cognate 71 kDa protein	2	3	1	1.822	1.717	0.106	0.665
FFDEESYSLLR	A6NLM8;P51571	Translocon-associated protein subunit delta	10	2	-8	7.611	1.226	6.386	0.662
NVPYGNIQSR	Q10471	Polypeptide N-acetylgalactosaminyltransferase 2	8	1	-7	6.284	1.000	5.284	0.661
APFDLFENR	P07900;P07900-2	Heat shock protein HSP 90-alpha	9	2	-7	7.199	1.267	5.932	0.654
TTAENEFVMLK	P13647;F8W0C6;H0YIN9	Keratin, type II cytoskeletal 5	10	2	-8	7.869	1.328	6.541	0.653
TVLcGTcGQPADK	P02545-4;P02545-3;P02545-5;P02545-6;P02545;A0A0C4DGC5	Isoform 4 of Prelamin-A/C	10	2	-8	6.623	1.893	4.730	0.647
ELSDFISYLQR	P30101	Protein disulfide-isomerase A3	9	2	-7	6.641	1.621	5.020	0.637
IETNENNLESAK	Q07065	Cytoskeleton-associated protein 4	10	1	-9	7.324	1.000	6.324	0.634
DAEEWFFTK	P02533	Keratin, type I cytoskeletal 14	10	10	0	7.500	3.300	4.200	0.626
NPDTQWITKPVHK	E7ENU7;P61313;E7EQV9	Ribosomal protein L15	8	1	-7	5.954	1.000	4.954	0.619
EQAEAEVASLNR	D6R904;Q5HYB6;P06753;P06753;A0A087WWU8	Tropomyosin alpha-3 chain	3	9	6	2.291	7.088	-4.797	0.617
NMQDLVEDFK	P13647;F8W0C6	Keratin, type II cytoskeletal 5	10	2	-8	7.545	1.333	6.212	0.610
YEELQVTAGR	P02538;P48668;O95678	Keratin, type II cytoskeletal 6A	10	2	-8	7.008	1.291	5.718	0.605
NLSSTTDDEAPR	O00515;E9PDI4;A0A0J9YW63;A0A0J9YVY2;A0A0J9YYF6	Ladinin-1	10	1	-9	6.818	1.000	5.818	0.605
GSTDNLMDDIER	P50990-2;P50990-3;P50990	Isoform 2 of T-complex protein 1 subunit theta	9	1	-8	6.378	1.000	5.378	0.604

Table 9: The table shows different peptide sequence of HSP 90 protein in the four cancer cell lines.

Cell Line	Sequence	Accessions	# Peaks C12	# Peaks C13	# Peaks Diff	Abundance Sum C12	Abundance Sum C13	Abundance Sum Difference	Chi-squared	Chi-square normalized	Enrichment Ratio
UM1	YESLTDPSKLDSGK	Q58FF8;Q58FF7;P07900;P08238;G3V2J8	3	12	9	1.940	8.042	6.102	6.127	0.511	0.267
UM2	YESLTDPSKLDSGK	Q58FF8;Q58FF7;P07900;P08238;G3V2J8		6			3.129				0.000
UM-SCC5	not found										0.000
UM-SCC6	YESLTDPSKLDSGK	Q58FF8;Q58FF7;P07900;P08238;G3V2J8	1			1.000					0.000
Sequence 2											0.000
UM1	no found										0.000
UM2	NPDDITQEEYGEFYK	P08238	2	9	7	1.566	6.159	4.593	4.734	0.526	0.434
UM-SCC5	NPDDITQEEYGEFYK	P08238		10			7.823				0.000
UM-SCC6	NPDDITQEEYGEFYK	P08238	5			2.773					0.000
Sequence 3											0.000
UM1	APFDLFENR	P07900		10			5.489				0.000
UM2	APFDLFENR	P07900		1			1.000				0.000
UM-SCC5	APFDLFENR	P07900	1	9	8	1.000	7.199	6.199	6.201	0.689	0.000
UM-SCC6	APFDLFENR	P07900	3	2	-1	2.363	1.267	-1.096	1.159	0.386	-0.387
Sequence 4											0.000
UM1	DQVANSAFVER	P07900		8			5.868				0.000
UM2	DQVANSAFVER	P07900	3			1.719					0.000
UM-SCC5	DQVANSAFVER	P07900		11			6.720				0.000
UM-SCC6	DQVANSAFVER	P07900	4	2	-2	2.072	1.147	-0.925	1.135	0.284	-0.434

References

Abbassi-Ghadi, N., S. Kumar, J. Huang, R. Goldin, Z. Takats and G. B. Hanna (2013). "Metabolomic profiling of oesophago-gastric cancer: A systematic review." European Journal of Cancer 49(17): 3625-3637.

Ahsan, A., S. G. Ramanand, C. Whitehead, S. M. Hiniker, A. Rehemtulla, W. B. Pratt, S. Jolly, C. Gouveia, K. Truong, C. Van Waes, D. Ray, T. S. Lawrence and M. K. Nyati (2012). "Wild-type EGFR Is Stabilized by Direct Interaction with HSP90 in Cancer Cells and Tumors." Neoplasia (New York, N.Y.) 14(8): 670-677.

Alazawi, W., L. Morris, M. Stanley, D. Garrod and N. Coleman (2003). "Altered expression of desmosomal

components in high-grade squamous intraepithelial lesions of the cervix." Virchows Archiv **443**(1): 51-56.

Allen, J., H. M. Davey, D. Broadhurst, J. K. Heald, J. J. Rowland, S. G. Oliver and D. B. Kell (2003). "High-throughput classification of yeast mutants for functional genomics using metabolic footprinting." Nature biotechnology **21**(6): 692-696.

Arellano-Garcia, M. E., R. Li, X. Liu, Y. Xie, X. Yan, J. A. Loo and S. Hu (2010). "Identification of Tetranectin as a Potential Biomarker for Metastatic Oral Cancer." International Journal of Molecular Sciences **11**(9): 3106-3121.

Beliakoff, J. and L. Whitesell (2004). "Hsp90: an emerging target for breast cancer therapy." Anti-Cancer Drugs **15**(7): 651-662.

Benjamin, D. I., B. F. Cravatt and D. K. Nomura (2012). "Global profiling strategies for mapping Dysregulated Metabolic Pathways in Cancer." Cell metabolism **16**(5): 565-577.

Bennett, D. (2005). "Growing pains for metabolomics." The Scientist **19**(8): 25-28.

Berens, E. B., G. M. Sharif, M. O. Schmidt, G. Yan, C. W. Shuptrine, L. M. Weiner, E. Glasgow, A. T. Riegel and A. Wellstein (2016). "Keratin-associated protein 5-5 controls cytoskeletal

function and cancer cell vascular invasion." Oncogene.

Bonner, J. A., N. J. Maihle, B. R. Folven, T. J. H. Christianson and K. spain (1994). "The interaction of epidermal growth factor and radiation in human head and neck squamous cell carcinoma cell lines with vastly different radiosensitivities." International Journal of Radiation Oncology*Biology*Physics **29**(2): 243-247.

Boren, J., M. Cascante, S. Marin, B. Comín-Anduix, J. J. Centelles, S. Lim, S. Bassilian, S. Ahmed, W.-N. P. Lee and L. G. Boros (2001). "Gleevec (STI571) influences metabolic enzyme activities and glucose carbon flow towards nucleic acid and fatty acid synthesis in myeloid tumor cells." Journal of Biological Chemistry.

Boren, J., W.-N. P. Lee, S. Bassilian, J. J. Centelles, S. Lim, S. Ahmed, L. G. Boros and M. Cascante (2003). "The Stable Isotope-based Dynamic Metabolic Profile of Butyrate-induced HT29 Cell Differentiation." Journal of Biological Chemistry **278**(31): 28395-28402.

Boring, C. C., T. S. Squires, T. Tong and S. Montgomery (1994). "Cancer statistics, 1994." CA: A Cancer Journal for Clinicians **44**(1): 7-26.

Boros, L. G., M. Cascante and W.-N. Paul Lee (2002). "Metabolic profiling of cell growth and death in cancer: applications in drug discovery." Drug Discovery Today **7**(6): 364-372.

Brenner, J. C., M. P. Graham, B. Kumar, L. M. Saunders, R. Kupfer, R. H. Lyons, C. R. Bradford and T. E. Carey (2010). "Genotyping of 73 UM-SCC head and neck squamous cell carcinoma cell lines." Head & neck **32**(4): 417-426.

Brown, S. D., D. J. Rhodes and B. J. Pritchard (2007). "A validated SPME-GC–MS method for simultaneous quantification of club drugs in human urine." Forensic Science International **171**(2): 142-150.

Brunelli, C., C. Bicchi, A. Di Stilo, A. Salomone and M. Vincenti (2006). "High-speed gas chromatography in doping control: Fast-GC and fast-GC/MS determination of β-adrenoceptor ligands and diuretics." Journal of Separation Science **29**(18): 2765-2771.

Byers, H. R. and J. Bhawan (1998). "PATHOLOGIC PARAMETERS IN THE DIAGNOSIS AND

PROGNOSIS OF PRIMARY CUTANEOUS MELANOMA." Hematology/Oncology Clinics of North America **12**(4): 717-735.

Cairns, R. A., I. S. Harris and T. W. Mak (2011). "Regulation of cancer cell metabolism." Nat Rev Cancer **11**(2): 85-95.

Calderwood, S. K. (2010). "Heat shock proteins in breast cancer progression–A suitable case for treatment?" International Journal of Hyperthermia **26**(7): 681-685.

Caraglia, M., A. Abbruzzese, A. Leardi, S. Pepe, A. Budillon, G. Baldassare, C. Selleri, S. D. Lorenzo, A. Fabbrocini, G. Giuberti, G. Vitale, G. Lupoli, A. R. Bianco and P. Tagliaferri (1999). "Interferon-alpha induces apoptosis in human KB cells through a stress-dependent mitogen activated protein kinase pathway that is antagonized by epidermal growth factor." Cell Death Differ **6**(8): 773-780.

Cargile, B. J., J. L. Bundy, A. M. Grunden and J. L. Stephenson (2004). "Synthesis/Degradation Ratio Mass Spectrometry for Measuring Relative Dynamic Protein Turnover." Analytical Chemistry **76**(1): 86-97.

Cascante, M., V. Selivanov and A. Ramos-Montoya (2012). Application of tracer-based metabolomics and flux analysis in targeted cancer drug design. The Handbook of Metabolomics. W.-M. T. Fan, N. A. Lane and M. R. Higashi. Totowa, NJ, Humana Press**:** 299-320.

Chardy, P., M. Glemarec and A. Laurec (1976). "Application of inertia methods to benthic marine ecology: Practical implications of the basic options." Estuarine and Coastal Marine Science **4**(2): 179-205.

Chokkathukalam, A., D.-H. Kim, M. P. Barrett, R. Breitling and D. J. Creek (2014). "Stable isotope-labeling studies in metabolomics: new insights into structure and dynamics of metabolic networks." Bioanalysis **6**(4): 511-524.

Choo, A. Y., S. G. Kim, M. G. V. Heiden, S. J. Mahoney, H. Vu, S.-O. Yoon, L. C. Cantley and J. Blenis (2010). "Glucose addiction of TSC-null cells is caused by failed mTORC1-dependent balancing of metabolic demand with supply." Molecular cell **38**(4): 487-499.

Chung, C. H., J. S. Parker, G. Karaca, J. Wu, W. K. Funkhouser, D. Moore, D. Butterfoss, D. Xiang, A. Zanation, X. Yin, W. W. Shockley, M. C. Weissler, L. G. Dressler, C. G. Shores, W. G. Yarbrough and C. M. Perou (2004). "Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression." Cancer cell **5**(5): 489-500.

Ciocca, D. R., D. J. Adams, D. P. Edwards, R. J. Bjercke and W. L. McGuire (1983). "Distribution of an estrogen-induced protein with a molecular weight of 24,000 in normal and malignant human tissues and cells." Cancer Research **43**(3): 1204.

Cohen-Kerem, R., M. A. Rahat, W. Madah, E. Greenberg, E. Sabo and I. Elmalah (2004). "Cytokeratin-17 as a Potential Marker for Squamous Cell Carcinoma of the Larynx." Annals of Otology, Rhinology & Laryngology **113**(10): 821-827.

Cottrell, J. S. and U. London (1999). "Probability-based protein identification by searching sequence databases using mass spectrometry data." electrophoresis **20**(18): 3551-3567.

Crowe, D. L., G. E. Milo and C. E. Shuler (1999). "Keratin 19 Downregulation by Oral Squamous Cell Carcinoma Lines Increases Invasive Potential." Journal of Dental Research **78**(6): 1256-1263.

Cui, T., Y. Chen, L. Yang, T. Knösel, K. Zöller, O. Huber and I. Petersen (2011). "DSC3 expression is regulated by p53, and methylation of DSC3 DNA is a prognostic marker in human colorectal cancer." British journal of cancer **104**(6): 1013-1019.

De Paepe, B. (2012). "Mitochondrial Markers for Cancer: Relevance to Diagnosis, Therapy, and Prognosis and General Understanding of Malignant Disease Mechanisms." ISRN Pathology **2012**: 15.

DeBerardinis, R. J., N. Sayed, D. Ditsworth and C. B. Thompson (2008). "Brick by brick: metabolism and tumor cell growth." Current Opinion in Genetics & Development **18**(1): 54-61.

Deer, E. L., J. Gonzalez-Hernandez, J. D. Coursen, J. E. Shea, J. Ngatia, C. L. Scaife, M. A. Firpo and S. J. Mulvihill (2010). "Phenotype and Genotype of Pancreatic Cancer Cell Lines." Pancreas **39**(4): 425-435.

Dyce, O. H., A. F. Ziober, R. S. Weber, K. Miyazaki, S. S. Khariwala, M. Feldman and B. L. Ziober (2002). "Integrins in Head and Neck Squamous Cell Carcinoma Invasion." The Laryngoscope **112**(11): 2025-2032.

Elias, J. E., W. Haas, B. K. Faherty and S. P. Gygi (2005). "Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations." Nat Meth **2**(9): 667-675.

Fan, T. W. M., P. K. Lorkiewicz, K. Sellers, H. N. B. Moseley, R. M. Higashi and A. N. Lane (2012). "Stable isotope-resolved metabolomics and applications for drug development." Pharmacology & Therapeutics **133**(3): 366-391.

Faubert, B., G. Boily, S. Izreig, T. Griss, B. Samborska, Z. Dong, F. Dupuy, C. Chambers, Benjamin J. Fuerth, B. Viollet, Orval A. Mamer, D. Avizonis, Ralph J. DeBerardinis, Peter M. Siegel and Russell G. Jones (2013). "AMPK is a negative regulator of the Warburg effect and suppresses tumor growth in vivo." Cell Metabolism **17**(1): 113-124.

Fell, D. A. (2005). "Enzymes, metabolites and fluxes." Journal of Experimental Botany **56**(410): 267-272.

Feng, Z. and A. J. Levine (2010). "The regulation of energy metabolism and the IGF-1/mTOR pathways by the p53 protein." Trends in Cell Biology **20**(7): 427-434.

Ferreira, L. M. R. (2010). "Cancer metabolism: The Warburg effect today." Experimental and Molecular Pathology **89**(3): 372-380.

Fiehn, O. (2002). "Metabolomics–the link between genotypes and phenotypes." Plant molecular biology **48**(1-2): 155-171.

Friedman, J. A., S. C. Wise, M. Hu, C. Gouveia, R. V. Broek, C. Freudlsperger, V. R. Kannabiran, P. Arun, J. B. Mitchell, Z. Chen and C. Van Waes (2013). "HSP90 Inhibitor SNX5422/ 2112 Targets the Dysregulated Signal and Transcription Factor Network and Malignant Phenotype of Head and Neck Squamous Cell Carcinoma." Translational Oncology **6**(4): IN1-IN5.

Gatenby, R. A. and R. J. Gillies (2004). "Why do cancers have high aerobic glycolysis?" Nat Rev Cancer **4**(11): 891-899.

Gill, S. S., D. G. T. Thomas, N. V. Bruggen, D. G. Gadian, C. J. Peden, J. D. Bell, I. J. Cox, D. K. Menon, R. A. Iles, D. J. Bryant and G. A. Coutts (1990). "Proton MR Spectroscopy of Intracranial Tumours: In Vivo and In Vitro Studies." Journal of Computer Assisted Tomography **14**(4).

Gottlieb, E. and I. P. M. Tomlinson (2005). "Mitochondrial tumour suppressors: a genetic and biochemical update." Nat Rev Cancer **5**(11): 857-866.

Gowda, G. A. N. and D. Djukovic (2014). Overview of Mass Spectrometry-Based Metabolomics: Opportunities and Challenges. Mass Spectrometry in Metabolomics: Methods and Protocols. D. Raftery. New York, NY, Springer New York**:** 3-12.

Griffin, J. L. (2003). "Metabonomics: NMR spectroscopy and pattern recognition analysis of body fluids and tissues for characterisation of xenobiotic toxicity and disease diagnosis." Current opinion in chemical biology **7**(5): 648-654.

Griffin, J. L. and R. A. Kauppinen (2007). "A metabolomics perspective of human brain tumours." FEBS Journal **274**(5): 1132-1139.

Griffin, J. L. and J. P. Shockcor (2004). "Metabolic profiles of cancer cells." Nat Rev Cancer **4**(7): 551-561.

Gross, S., R. A. Cairns, M. D. Minden, E. M. Driggers, M. A. Bittinger, H. G. Jang, M. Sasaki, S. Jin, D. P. Schenkein, S. M. Su, L. Dang, V. R. Fantin and T. W. Mak (2010). "Cancer-associated metabolite 2-hydroxyglutarate accumulates in acute myelogenous leukemia with isocitrate dehydrogenase 1 and 2 mutations." The Journal of Experimental Medicine **207**(2): 339-344.

Hamidov, Z., A. Altendorf-Hofmann, Y. Chen, U. Settmacher, I. Petersen and T. Knösel (2011). "Reduced expression of desmocollin 2 is an independent prognostic biomarker for shorter patients survival in pancreatic ductal adenocarcinoma." Journal of Clinical Pathology **64**(11): 990-994.

Hay, N. (2016). "Reprogramming glucose metabolism in cancer: can it be exploited for cancer therapy?" Nature Reviews Cancer.

Hellerstein, M. K. and R. A. Neese (1999). "Mass isotopomer distribution analysis at eight years:

theoretical, analytic, and experimental considerations." American Journal of Physiology - Endocrinology And Metabolism **276**(6): E1146.

Hicks, M. J. and C. M. Flaitz (2000). "Oral mucosal melanoma: epidemiology and pathobiology." Oral Oncology **36**(2): 152-169.

Hori, D., Y. Hasegawa, M. Kimura, Y. Yang, I. C. Verma and S. Yamaguchi (2005). "Clinical onset and prognosis of Asian children with organic acidemias, as detected by analysis of urinary organic acids using GC/MS, instead of mass screening." Brain and Development **27**(1): 39-45.

Hu, S., J. Wang, E. H. Ji, T. Christison, L. Lopez and Y. Huang (2015). "Targeted Metabolomic Analysis of Head and Neck Cancer Cells Using High Performance Ion Chromatography Coupled with a Q Exactive HF Mass Spectrometer." Analytical Chemistry **87**(12): 6371-6379.

Humphries, J. D., A. Byron and M. J. Humphries (2006). "Integrin ligands at a glance." Journal of Cell Science **119**(19): 3901.

Jellum, E., I. Bjørnson, R. Nesbakken, E. Johansson and S. Wold (1981). "Classification of human cancer cells by means of capillary gas chromatography and pattern recognition analysis." Journal of Chromatography A **217**: 231-237.

Jemal, A., F. Bray, M. M. Center, J. Ferlay, E. Ward and D. Forman (2011). "Global cancer statistics." CA: A Cancer Journal for Clinicians **61**(2): 69-90.

Jordan, K. W., J. Nordenstam, G. Y. Lauwers, D. A. Rothenberger, K. Alavi, M. Garwood and L. L. Cheng (2009). "Metabolomic Characterization of Human Rectal Adenocarcinoma With Intact Tissue Magnetic Resonance Spectroscopy." Diseases of the colon and rectum **52**(3): 520-525.

Jung, Y., Y. Shiozawa, J. Wang, J. Wang, Z. Wang, E. A. Pedersen, C. H. Lee, C. L. Hall, P. J. Hogg, P. H. Krebsbach, E. T. Keller and R. S. Taichman (2009). "Expression of PGK1 by prostate cancer cells induces bone formation." Molecular Cancer Research **7**(10): 1595-1604.

Kamangar, F., G. M. Dores and W. F. Anderson (2006). "Patterns of Cancer Incidence, Mortality, and Prevalence Across Five Continents: Defining Priorities to Reduce Cancer Disparities in Different Geographic Regions of the World." Journal of Clinical Oncology **24**(14): 2137-2150.

Kanojia, D., S. S. Sawant, A. M. Borges, A. D. Ingle and M. M. Vaidya (2012). "Alterations in keratins and associated proteins during 4- Nitroquinoline-1-oxide induced rat oral carcinogenesis." Journal of Carcinogenesis **11**: 14.

Karantza, V. (2011). "Keratins in health and cancer: more than mere epithelial cell markers." Oncogene **30**(2): 127-138.

Katz, J. and N. Grunnet (1979). Estimation of metabolic pathways in steady state in vitro: rates of tricarboxylic acid and pentose cycles, Elsevier (North-Holland).

Katz, J., W. Lee, P. Wals and E. Bergner (1989). "Studies of glycogen synthesis and the Krebs cycle by mass isotopomer analysis with [U-13C] glucose in rats." Journal of Biological Chemistry **264**(22): 12994-13004.

Katz, J. and W. N. Lee (1991). "Application of mass isotopomer analysis for determination of pathways of glycogen synthesis." American Journal of Physiology - Endocrinology and Metabolism **261**(3): E332-E336.

Katz, J., P. Wals and W. Lee (1993). "Isotopomer studies of gluconeogenesis and the Krebs cycle with 13C-labeled lactate." Journal of Biological Chemistry **268**(34): 25509-25521.

Katz, J., P. A. Wals and W. N. Lee (1991). "Determination of pathways of glycogen synthesis and the dilution of the three-carbon pool with [U-13C]glucose." Proceedings of the National Academy of Sciences of the United States of America **88**(6): 2103-2107.

Keller, A., A. I. Nesvizhskii, E. Kolker and R. Aebersold (2002). "Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search." Anal Chem **74**(20): 5383-5392.

Kim, I.-Y., S.-H. Suh, I.-K. Lee and R. R. Wolfe (2016). "Applications of stable, nonradioactive isotope tracers in in vivo human metabolic research." Experimental & Molecular Medicine **48**(1): e203.

Kim, J.-w. and C. V. Dang (2006). "Cancer's molecular sweet tooth and the Warburg effect." Cancer Research **66**(18): 8927-8930.

Kitamura, R., T. Toyoshima, H. Tanaka, S. Kawano, T. Kiyosue, R. Matsubara, Y. Goto, M. Hirano, K. Oobu and S. Nakamura (2012). "Association of cytokeratin 17 expression with differentiation in oral squamous cell carcinoma." Journal of Cancer Research and Clinical Oncology **138**(8): 1299-1310.

Koenig, T., B. H. Menze, M. Kirchner, F. Monigatti, K. C. Parker, T. Patterson, J. J. Steen, F. A. Hamprecht and H. Steen (2008). "Robust Prediction of the MASCOT Score for an Improved Quality Assessment in Mass Spectrometric Proteomics." Journal of Proteome Research **7**(9): 3708-3717.

Kowalczyk, A. P. and K. J. Green (2013). "Structure, Function and Regulation of Desmosomes." Progress in molecular biology and translational science **116**: 95-118.

Kruger, N. J. and A. von Schaewen (2003). "The oxidative pentose phosphate pathway: structure and organisation." Current Opinion in Plant Biology **6**(3): 236-246.

Lazaris, A. C., E. B. Chatzigianni, D. Panoussopoulos, G. N. Tzimas, P. S. Davaris and B. C. Golematis (1997). "Proliferating cell nuclear antigen and heat shock protein 70 immuno-localization in invasive ductal breast cancer not otherwise specified." Breast Cancer Research and Treatment **43**(1): 43-51.

Lazaris, A. C., G. E. Theodoropoulos, P. S. Davaris, D. Panoussopoulos, L. Nakopoulou, C. Kittas and B. C. Golematis (1995). "Heat shock protein 70 and HLA-DR molecules tissue expression: Prognostic implications in colorectal cancer." Diseases of the Colon & Rectum **38**(7): 739-745.

Lee, W.-N. P., L. G. Boros and V.-L. W. Go (2012). Metabolic Pathways as Targets for Drug Screening, INTECH Open Access Publisher.

Lee, W., L. Boros, J. Puigjaner, S. Bassilian, S. Lim and M. Cascante (1998). "Investigation of the pentose cycle using [1, 2-13C2]-glucose and mass isotopomer analysis: estimation of transketolase and transaldolase activities." Am. J. Physiol **274**: E843-E851.

Lee, W. N. P. (2006). "Characterizing phenotype with tracer based metabolomics." Metabolomics **2**: 31-39.

Lee, W. N. P., P. Guo, S. Lim, S. Bassilian, S. T. Lee, J. Boren, M. Cascante, V. L. W. Go and L. G. Boros (2004). "Metabolic sensitivity of pancreatic tumour cell apoptosis to glycogen phosphorylase inhibitor treatment." Br J Cancer **91**(12): 2094-2100.

Lee, W. N. P., P. N. Wahjudi, J. Xu and V. L. Go (2010). "Tracer-based Metabolomics: Concepts and Practices." Clinical biochemistry **43**(16-17): 1269-1277.

Levine, A. J. and A. M. Puzio-Kuter (2010). "The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes." Science **330**(6009): 1340-1344.

Li, M., H. Wang, F. Li, W. E. Fisher, C. Chen and Q. Yao (2005). "Effect of cyclophilin A on gene expression in human pancreatic cancer cells." The American Journal of Surgery **190**(5): 739-745.

Li, M., Q. Zhai, U. Bharadwaj, H. Wang, F. Li, W. E. Fisher, C. Chen and Q. Yao (2006). "Cyclophilin A is overexpressed in human pancreatic cancer cells and stimulates cell proliferation through CD147." Cancer **106**(10): 2284-2294.

Lin, S. M., L. Zhu, A. Q. Winter, M. Sasinowski and W. A. Kibbe (2005). "What is mzXML good for?" Expert Review of Proteomics **2**(6): 839-845.

Lippman , S. M. and W. K. Hong (2001). "Molecular Markers of the Risk of Oral Cancer." New England Journal of Medicine **344**(17): 1323-1326.

Liu, L.-K., X.-Y. Jiang, X.-X. Zhou, D.-M. Wang, X.-L. Song and H.-B. Jiang (2009). "Upregulation of vimentin and aberrant expression of E-cadherin/[beta]-catenin complex in oral squamous cell carcinomas: correlation with the clinicopathological features and patient outcome." Mod Pathol **23**(2): 213-224.

Liu, Y.-Q., Z.-G. Huang, G.-N. Li, J.-L. Du, Y.-P. Ou, X.-N. Zhang, T.-T. Chen and Q.-L. Liang (2015). "Effects of α-enolase (ENO1) over-expression on malignant biological behaviors of AGS cells." International Journal of Clinical and Experimental Medicine **8**(1): 231-239.

Lu, W., J. Gao, J. Yang, Y. Cao, L. Jiang, M. Li, Y. Zhang, J. Zhou and Y. Liu (2015). "Down-regulated phosphoglycerate kinase 1 expression is associated with poor prognosis in patients

with gallbladder cancer." Medicine **94**(49): e2244.

Luo, A., J. Kong, G. Hu, C.-C. Liew, M. Xiong, X. Wang, J. Ji, T. Wang, H. Zhi, M. Wu and Z. Liu (2003). "Discovery of Ca2+-relevant and differentiation-associated genes downregulated in esophageal squamous cell carcinoma using cDNA microarray." Oncogene **23**(6): 1291-1299.

Mamas, M., W. B. Dunn, L. Neyses and R. Goodacre (2011). "The role of metabolites and metabolomics in clinically applicable biomarkers of disease." Archives of Toxicology **85**(1): 5-17.

Marin-Valencia, I., C. Yang, T. Mashimo, S. Cho, H. Baek, X.-L. Yang, K. N. Rajagopalan, M. Maddie, V. Vemireddy, Z. Zhao, L. Cai, L. Good, B. P. Tu, K. J. Hatanpaa, B. E. Mickey, J. M. Matés, J. M. Pascual, E. A. Maher, C. R. Malloy, R. J. DeBerardinis and R. M. Bachoo (2012). "Analysis of tumor metabolism reveals mitochondrial glucose oxidation in genetically diverse, human glioblastomas in the mouse brain in vivo." Cell Metabolism **15**(6): 827-837.

Martin, F. P. J., S. Collino and S. Rezzi (2011). "1H NMR-based metabonomic applications to decipher gut microbial metabolic influence on mammalian health." Magnetic Resonance in Chemistry **49**(S1): S47-S54.

Meyer, M. R., F. T. Peters and H. H. Maurer (2010). "Automated mass spectral deconvolution and identification system for GC-MS screening for drugs, poisons, and metabolites in urine." Clinical Chemistry **56**(4): 575-584.

Misiorek, J. O., I. A. K. Lähdeniemi, J. H. Nyström, V. M. Paramonov, J. A. Gullmets, H. Saarento, A. Rivero-Müller, T. Husøy, P. Taimen and D. M. Toivola (2016). "Keratin 8-deletion induced colitis predisposes to murine colorectal cancer enforced by the inflammasome and IL-22 pathway." Carcinogenesis.

Misuno, K., X. Liu, S. Feng and S. Hu (2013). "Quantitative proteomic analysis of sphere-forming stem-like oral cancer cells." Stem Cell Research & Therapy **4**(6): 156.

Moll, R., M. Divo and L. Langbein (2008). "The human keratins: biology and pathology." Histochemistry and Cell Biology **129**(6): 705-733.

Muñoz-Pinedo, C., N. El Mjiyad and J. E. Ricci (2012). "Cancer metabolism: current perspectives and future directions." Cell Death & Disease **3**(1): e248.

Nakayama, S., A. Sasaki, H. Mese, R. E. Alcalde and T. Matsumura (1998). "Establishment of high and low metastasis cell Lines derived from a human tongue squamous cell carcinoma." Invasion and Metastasis **18**(5-6): 219-228.

Nesvizhskii, A. I., A. Keller, E. Kolker and R. Aebersold (2003). "A statistical model for identifying proteins by tandem mass spectrometry." Anal Chem **75**(17): 4646-4658.

Nicholson, J. K., J. C. Lindon and E. Holmes (1999). "'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data." Xenobiotica **29**(11): 1181-1189.

O'Connell, T. M. (2012). "Recent advances in metabolomics in oncology." Bioanalysis **4**(4): 431-451.

Oliver, S. G., M. K. Winson, D. B. Kell and F. Baganz (1998). "Systematic functional analysis of the yeast genome." Trends in biotechnology **16**(9): 373-378.

Orth, J. D., I. Thiele and B. O. Palsson (2010). "What is flux balance analysis?" Nat Biotech **28**(3): 245-248.

Oshiro, M. M., C. J. Kim, R. J. Wozniak, D. J. Junk, J. L. Muñoz-Rodríguez, J. A. Burr, M. Fitzgerald, S. C. Pawar, A. E. Cress and F. E. Domann (2005). "Epigenetic silencing of DSC3 is a common event in human breast cancer." Breast Cancer Research **7**(5): 1.

Ouyang, H., L.-j. Mou, C. Luk, N. Liu, J. Karaskova, J. Squire and M.-S. Tsao (2000). "Immortal Human Pancreatic Duct Epithelial Cell Lines with Near Normal Genotype and Phenotype." The American Journal of Pathology **157**(5): 1623-1631.

Papageorgopoulos, C., K. Caldwell, H. Schweingrubber, R. A. Neese, C. H. L. Shackleton and M. Hellerstein (2002). "Measuring synthesis rates of muscle creatine kinase and myosin with stable isotopes and mass spectrometry." Analytical Biochemistry **309**(1): 1-10.

Papageorgopoulos, C., K. Caldwell, C. Shackleton, H. Schweingrubber and M. K. Hellerstein

(1999). "Measuring Protein Synthesis by Mass Isotopomer Distribution Analysis (MIDA)." Analytical Biochemistry **267**(1): 1-16.

Papagerakis, S., A.-H. Shabana, B. H. Pollock, P. Papagerakis, J. Depondt and A. Berdal (2009). "Altered desmoplakin expression at transcriptional and protein levels provides prognostic information in human oropharyngeal cancer." Human pathology **40**(9): 1320-1329.

Pappin, D. J. C., P. Hojrup and A. J. Bleasby (1993). "Rapid identification of proteins by peptide-mass fingerprinting." Current Biology **3**(6): 327-332.

Park, B. J., S. I. Chiosea and J. R. Grandis (2010). "Molecular changes in the multistage pathogenesis of head and neck cancer." Cancer biomarkers : section A of Disease markers **9**(1-6): 325-339.

Patti, G. J., R. Tautenhahn, B. R. Fonslow, Y. Cho, A. Deutschbauer, A. Arkin, T. Northen and G. Siuzdak (2011). "Meta-analysis of global metabolomics and proteomics data to link alterations with phenotype." Spectroscopy **26**(3).

Patti, G. J., R. Tautenhahn, D. Rinehart, K. Cho, L. Shriver, M. Manchester, I. Nikolskiy, C. H. Johnson, N. G. Mahieu and G. Siuzdak (2013). "A View from Above: Cloud Plots to Visualize Global Metabolomic Data." Analytical chemistry **85**(2): 798-804.

Paul, T. M. L. and M. M. G. Margaret (2016). "Cyclophilin function in cancer; lessons from virus replication." Current Molecular Pharmacology **9**(2): 148-164.

Pedrioli, P. G. A., J. K. Eng, R. Hubley, M. Vogelzang, E. W. Deutsch, B. Raught, B. Pratt, E. Nilsson, R. H. Angeletti, R. Apweiler, K. Cheung, C. E. Costello, H. Hermjakob, S. Huang, R. K. Julian Jr, E. Kapp, M. E. McComb, S. G. Oliver, G. Omenn, N. W. Paton, R. Simpson, R. Smith, C. F. Taylor, W. Zhu and R. Aebersold (2004). "A common open representation of mass spectrometry data and its application to proteomics research." Nat Biotech **22**(11): 1459-1466.

Priante, A. V. M., A. L. Carvalho and L. P. Kowalski (2010). "Segundo tumor primário em pacientes com câncer de vias aerodigestivas superiores." Brazilian Journal of Otorhinolaryngology **76**: 251-256.

Prince, M. E., R. Sivanandan, A. Kaczorowski, G. T. Wolf, M. J. Kaplan, P. Dalerba, I. L. Weissman, M. F. Clarke and L. E. Ailles (2007). "Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma." Proceedings of the National Academy of Sciences **104**(3): 973-978.

Raamsdonk, L. M., B. Teusink, D. Broadhurst, N. Zhang, A. Hayes, M. C. Walsh, J. A. Berden, K. M. Brindle, D. B. Kell and J. J. Rowland (2001). "A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations." Nature biotechnology **19**(1): 45-50.

Rajagopalan, K. N. and R. J. DeBerardinis (2011). "Role of Glutamine in Cancer – Therapeutic and Imaging Implications." Journal of Nuclear Medicine **52**(7): 1005-1008.

Rasola, A., L. Neckers and D. Picard (2014). "Mitochondrial oxidative phosphorylation TRAP(1)ped in tumor cells." Trends in Cell Biology **24**(8): 455-463.

Reaves, M. L. and J. D. Rabinowitz (2011). "Metabolomics in systems microbiology." Current Opinion in Biotechnology **22**(1): 17-25.

Rhee, E. P. and R. E. Gerszten (2012). "Metabolomics and Cardiovascular Biomarker Discovery." Clinical Chemistry **58**(1): 139-147.

Robertson, D. G., P. B. Watkins and M. D. Reily (2011). "Metabolomics in Toxicology: Preclinical and Clinical Applications." Toxicological Sciences **120**(suppl 1): S146-S170.

Rodríguez-Enríquez, S., P. A. Vital-González, F. L. Flores-Rodríguez, A. Marín-Hernández, L. Ruiz-Azuara and R. Moreno-Sánchez (2006). "Control of cellular proliferation by modulation of oxidative phosphorylation in human and rodent fast-growing tumor cells." Toxicology and Applied Pharmacology **215**(2): 208-217.

Safadi, R. A., D. H. Bader, N. I. Abdullah and M. A. Sughayer (2016). "Immunohistochemical expression of keratins 6, 7, 8, 14, 16, 18, 19, and MNF-116 pancytokeratin in primary and metastatic melanoma of the head and neck." Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology **121**(5): 510-519.

Sandulache, V. C., T. J. Ow, C. R. Pickering, M. J. Frederick, G. Zhou, I. Fokt, M. Davis-Malesevich, W. Priebe and J. N. Myers (2011). "Glucose, not glutamine, is the dominant energy source required for proliferation and survival of head and neck squamous carcinoma cells." Cancer **117**(13): 2926-2938.

Santarosa, M., D. Favaro, M. Quaia and E. Galligioni (2005). "Expression of heat shock protein 72 in renal cell carcinoma: Possible role and prognostic implications in cancer patients." European Journal of Cancer **33**(6): 873-877.

Savorani, F., M. A. Rasmussen, M. S. Mikkelsen and S. B. Engelsen (2013). "A primer to nutritional metabolomics by NMR spectroscopy and chemometrics." Food Research International **54**(1): 1131-1145.

Scalbert, A., L. Brennan, O. Fiehn, T. Hankemeier, B. S. Kristal, B. van Ommen, E. Pujos-Guillot, E. Verheij, D. Wishart and S. Wopereis (2009). "Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research." Metabolomics **5**(4): 435-458.

Shah, S. H., W. E. Kraus and C. B. Newgard (2012). "Metabolomic Profiling for the Identification of Novel Biomarkers and Mechanisms Related to Common Cardiovascular Diseases." Form and Function **126**(9): 1110-1120.

Siegel, R. L., K. D. Miller and A. Jemal (2017). "Cancer statistics, 2017." CA: A Cancer Journal for Clinicians **67**(1): 7-30.

Stewart, B. and C. P. Wild (2016). "World cancer report 2014." World.

Sun, S., X. Liang, X. Zhang, T. Liu, Q. Shi, Y. Song, Y. Jiang, H. Wu, Y. Jiang, X. Lu and D. Pang (2015). "Phosphoglycerate kinase-1 is a predictor of poor survival and a novel prognostic biomarker of chemoresistance to paclitaxel treatment in breast cancer." Br J Cancer **112**(8): 1332-1339.

Takahashi, H., N. Shikata, H. Senzaki, M. Shintaku and A. Tsubura (1995). "Immunohistochemical staining patterns of keratins in normal oesophageal epithelium and

carcinoma of the oesophagus." Histopathology **26**(1): 45-50.

Tavassol, F., O. F. Starke, H. Kokemüller, G. Wegener, C. C. M. Müller-Tavassol, N.-C. Gellrich and A. Eckardt (2011). "Prognostic significance of heat shock protein 70 (HSP70) in patients with oral cancer." Head & Neck Oncology **3**: 10-10.

Theodoridis, G. A., H. G. Gika, E. J. Want and I. D. Wilson (2012). "Liquid chromatography–mass spectrometry based global metabolite profiling: A review." Analytica Chimica Acta **711**: 7-16.

Trock, B. J. (2011). "Application of metabolomics to prostate cancer." Urologic Oncology: Seminars and Original Investigations **29**(5): 572-581.

Usenius, J.-P. R., R. A. Kauppinen, P. A. Vainio, J. A. Hernesniemi, M. P. Vapalahti, L. A. Paljärvi and S. Soimakallio (1994). "Quantitative Metabolite Patterns of Human Brain Tumors: Detection by 1H NMR Spectroscopy In Vivo and In Vitro." Journal of Computer Assisted Tomography **18**(5): 705-713.

Vaitheesvaran, B., J. Xu, J. Yee, Q.-Y. Lu, V. Go, G. Xiao and W.-N. Lee (2014). "The Warburg effect: a balance of flux analysis." Metabolomics: 1-10.

Vaupel, P. (2004). "The Role of Hypoxia-Induced Factors in Tumor Progression." The Oncologist **9**(suppl 5): 10-17.

Vogt, J. A., C. Hunzinger, K. Schroer, K. Hölzer, A. Bauer, A. Schrattenholz, M. A. Cahill, S. Schillo, G. Schwall, W. Stegmann and G. Albuszies (2005). "Determination of fractional synthesis rates of mouse hepatic proteins via metabolic 13C-Labeling, MALDI-TOF MS and analysis of relative isotopologue abundances using average masses." Analytical Chemistry **77**(7): 2034-2042.

Vogt, J. A., K. Schroer, K. Hölzer, C. Hunzinger, M. Klemm, K. Biefang-Arndt, S. Schillo, M. A. Cahill, A. Schrattenholz, H. Matthies and W. Stegmann (2003). "Protein abundance quantification in embryonic stem cells using incomplete metabolic labelling with 15N amino acids, matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry, and analysis of relative isotopologue abundances of peptides." Rapid Communications in Mass Spectrometry

17(12): 1273-1282.

Wang, J., T. T. Christison, K. Misuno, L. Lopez, A. F. Huhmer, Y. Huang and S. Hu (2014). "Metabolomic Profiling of Anionic Metabolites in Head and Neck Cancer Cells by Capillary Ion Chromatography with Orbitrap Mass Spectrometry." Analytical Chemistry **86**(10): 5116-5124.

Wang, J., X. Zhang, D. Ma, W.-N. P. Lee, J. Xiao, Y. Zhao, V. L. Go, Q. Wang, Y. Yen, R. Recker and G. G. Xiao (2013). "Inhibition of transketolase by oxythiamine altered dynamics of protein signals in pancreatic cancer cells." Experimental Hematology & Oncology **2**: 18-18.

Wang, L., J. Chen, L. Chen, P. Deng, Q. bu, P. Xiang, M. Li, W. Lu, Y. Xu, H. Lin, T. Wu, H. Wang, J. Hu, X. Shao, X. Cen and Y.-L. Zhao (2013). "1H-NMR based metabonomic profiling of human esophageal cancer tissue." Molecular Cancer **12**(1): 25.

Warburg, O. (1956). "On the Origin of Cancer Cells." Science **123**(3191): 309-314.

Wei, J., G. Xie, Z. Zhou, P. Shi, Y. Qiu, X. Zheng, T. Chen, M. Su, A. Zhao and W. Jia (2011). "Salivary metabolite signatures of oral cancer and leukoplakia." International Journal of Cancer **129**(9): 2207-2217.

Weiss, R. H. and K. Kim (2012). "Metabolomics in the study of kidney diseases." Nat Rev Nephrol **8**(1): 22-33.

Wetzels, R. H. W., H. E. Schaafsma, I. M. Leigh, E. B. Lane, S. M. Troyanovsky, S. S. C. Wagenaar, G. P. Vooijs and F. C. S. Ramaekers (1992). "Laminin and type VII collagen distribution in different types of human lung carcinoma: correlation with expression of keratins 14, 16, 17 and 18." Histopathology **20**(4): 295-303.

Wicha, M. S., S. Liu and G. Dontu (2006). "Cancer Stem Cells: An Old Idea—A Paradigm Shift." Cancer Research **66**(4): 1883-1890.

Williams, M. D., R. Reeves, L. S. Resar and H. H. Hill (2013). "Metabolomics of colorectal cancer: past and current analytical platforms." Analytical and Bioanalytical Chemistry **405**(15): 5013-5030.

Wishart, D., R. Mandal, A. Stanislaus and M. Ramirez-Gaona (2016). "Cancer metabolomics and

the human metabolome database." Metabolites **6**(1): 10.

Wishart, D. S., T. Jewison, A. C. Guo, M. Wilson, C. Knox, Y. Liu, Y. Djoumbou, R. Mandal, F. Aziat, E. Dong, S. Bouatra, I. Sinelnikov, D. Arndt, J. Xia, P. Liu, F. Yallou, T. Bjorndahl, R. Perez-Pineiro, R. Eisner, F. Allen, V. Neveu, R. Greiner and A. Scalbert (2013). "HMDB 3.0—The Human Metabolome Database in 2013." Nucleic Acids Research **41**(Database issue): D801-D807.

Wu, H., R. Xue, C. Lu, C. Deng, T. Liu, H. Zeng, Q. Wang and X. Shen (2009). "Metabolomic study for diagnostic model of oesophageal cancer using gas chromatography/mass spectrometry." Journal of Chromatography B **877**(27): 3111-3117.

Wysocki, V. H., K. A. Resing, Q. Zhang and G. Cheng (2005). "Mass spectrometry of peptides and proteins." Methods **35**(3): 211-222.

Xiao, G. G., M. Garg, S. Lim, D. Wong, V. L. Go and W.-N. P. Lee (2008). "Determination of protein synthesis in vivo using labeling from deuterated water and analysis of MALDI-TOF spectrum." Journal of Applied Physiology **104**(3): 828-836.

Xu, J., Y. Chen, R. Zhang, Y. Song, J. Cao, N. Bi, J. Wang, J. He, J. Bai, L. Dong, L. Wang, Q. Zhan and Z. Abliz (2013). "Global and Targeted Metabolomics of Esophageal Squamous Cell Carcinoma Discovers Potential Diagnostic and Therapeutic Biomarkers." Molecular & Cellular Proteomics **12**(5): 1306-1318.

Yan, S.-K., B.-J. Wei, Z.-Y. Lin, Y. Yang, Z.-T. Zhou and W.-D. Zhang (2008). "A metabonomic approach to the diagnosis of oral squamous cell carcinoma, oral lichen planus and oral leukoplakia." Oral Oncology **44**(5): 477-483.

Yang, C., J. Sudderth, T. Dang, R. G. Bachoo, J. G. McDonald and R. J. DeBerardinis (2009). "Glioblastoma cells require glutamate dehydrogenase to survive impairments of glucose metabolism or Akt signaling." Cancer research **69**(20): 7986-7993.

Yang, L., Y. Chen, T. Cui, T. Knösel, Q. Zhang, K. F. Albring, O. Huber and I. Petersen (2012). "Desmoplakin acts as a tumor suppressor by inhibition of the Wnt/β-catenin signaling pathway

in human lung cancer." Carcinogenesis.

Yang, Y., L. Wang, S. Wang, S. Liang, A. Chen, H. Tang, L. Chen and F. Deng (2013). "Study of metabonomic profiles of human esophageal carcinoma by use of high-resolution magic-angle spinning 1H NMR spectroscopy and multivariate data analysis." Analytical and Bioanalytical Chemistry **405**(10): 3381-3389.

Yonezawa, K., S. Nishiumi, J. Kitamoto-Matsuda, T. Fujita, K. Morimoto, D. Yamashita, M. Saito, N. Otsuki, Y. Irino, M. Shinohara, M. Yoshida and K. Nibu (2013). "Serum and tissue metabolomics of head and neck cancer." Cancer Genomics Proteomics **10**(5): 233-238.

Yoshizawa, K., S. Nozaki, A. Okamune, H. Kitahara, T. Ohara, K. Kato, S. Kawashiri and E. Yamamoto (2009). "Loss of maspin is a negative prognostic factor for invasion and metastasis in oral squamous cell carcinoma." Journal of Oral Pathology & Medicine **38**(6): 535-539.

Yuneva, M. O., T. W. M. Fan, T. D. Allen, R. M. Higashi, D. V. Ferraris, T. Tsukamoto, J. M. Matés, F. J. Alonso, C. Wang, Y. Seo, X. Chen and J. M. Bishop (2012). "The Metabolic Profile of Tumors Depends on both the Responsible Genetic Lesion and Tissue Type." Cell Metabolism **15**(2): 157-170.

Zamboni, N., S.-M. Fendt, M. Ruhl and U. Sauer (2009). "13C-based metabolic flux analysis." Nat. Protocols **4**(6): 878-892.

Zhang, J., L. Liu, S. Wei, G. A. Nagana Gowda, Z. Hammoud, K. A. Kesler and D. Raftery (2011). "Metabolomics study of esophageal adenocarcinoma." The Journal of Thoracic and Cardiovascular Surgery **141**(2): 469-475.e464.

Zhao, Y., E. B. Butler and M. Tan (2013). "Targeting cellular metabolism to improve cancer therapeutics." Cell Death & Disease **4**(3): e532.

Zhao, Y., W.-N. P. Lee, S. Lim, V. L. Go, J. Xiao, R. Cao, H. Zhang, R. Recker and G. G. Xiao (2009). "Quantitative Proteomics: Measuring Protein Synthesis Using (15)N Amino Acids Labeling in Pancreas Cancer Cells." Analytical chemistry **81**(2): 764-771.

Zheng, J. (2012). "Energy metabolism of cancer: Glycolysis versus oxidative phosphorylation

(Review)." Oncology letters **4**(6): 1151-1157.

Zieker, D., S. Bühler, Z. Üstündag, I. Königsrainer, S. Manncke, K. Bajaeifer, J. Vollmer, F. Fend, H. Northoff, A. Königsrainer and J. Glatzle (2013). "Induction of tumor stem cell differentiation—novel strategy to overcome therapy resistance in gastric cancer." Langenbeck's Archives of Surgery **398**(4): 603-608.