2020-01-07

Identification and Characterization of Different Metabolic Subtypes in Cancer

Pervin, Jannat

UNIVERSITY OF CALGARY

Identification and Characterization of Different Metabolic Subtypes in Cancer

by

Jannat Pervin

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN MEDICAL SCIENCE

CALGARY, ALBERTA

JANUARY, 2020

Abstract

Cancer is a leading cause of death worldwide. Genomics based approaches represent a dominant approach in oncological research. However, multiple processes can modify genetic information and impact cancer's phenotype in a non-coding manner such as epigenetic events, transcription of various splice variants, expression of non-coding RNA and miRNA, and post-translational modifications of proteins. Therefore, molecular events that are further downstream of the genome (perhaps reflected by the proteome or the metabolome) may better reflect the tumour phenotype. One feature of cancer is perturbed metabolism. Some of the aberrant metabolic pathways may enhance tumour viability and growth, and these perturbed pathways may be susceptible to pharmacologic inhibition. Thus, our overall goal is to categorize tumours by their metabolic features; to understand the biological implications of these metabolic features, and to identify pathways that could be potentially targeted with drugs. This project involves the development of a workflow to define the metabolic features of a tumour. The workflow will involve the categorization of tumours based on their metabolic features (at the transcriptome level), exploration of associated biological features of each metabolic subtype, and integration of multiple levels of molecular control (including mutation status, copy number variation, methylation, and metabolome). Our work began with breast cancer, which is already well characterized by a large cohort in The Cancer Genome Atlas (TCGA) project. Then we used the same principles to investigate a more complex tumour type, pancreatic cancer, which is characterized by a highly variable degree of stroma infiltration.

Acknowledgements

I would like to thank several people for their contributions to my research and for their continued support over these past few years. First and foremost, I would like to thank Dr. Oliver F. Bathe for being my supervisor. He has been incredibly generous with his time, expertise, and wisdom throughout my research. I am immensely grateful to Dr. Bathe for giving me the opportunity to learn from him and work in his laboratory. I would like to thank my committee members, Dr. Patricia Tang and Dr. Edwin Wang, for their insights and guidance over these past two and a half years. I am sincerely grateful to the past and present members of the Bathe laboratory who have been wonderful laboratory mates and colleagues. I would like to give special thanks to Cynthia stretch, Farshad Farshidfar, Marie Palmnas, Jodi Rattner, Victoria Armstrong, Jennifer Koziak, and Zoe Osborne for the discussions we've had and for their contributions to my research.

Dedication

To my parents Kabir Ahmed and Pervin Akter, my siblings Rowshan, Sharif, Masum, and my nephew Adyan. I could not have done this without you.

Table of Contents

List of Figures and Illustrations

List of Tables

Chapter One: Introduction

1.1 Overview

One hallmark of cancer is the reprogramming of energy metabolism. Cancer cell growth, proliferation, invasion and survival all require access to large amounts of energy (e.g., ATP, GTP), and so these processes are highly dependent on metabolic reprogramming [1-3]. It is possible that inhibition of the metabolic features that characterize cancer cells may also inhibit the manifestation of other hallmarks of cancer.

The perturbed metabolism of cancer cells may also have effects on adjacent non-cancerous cells comprising the tumour stroma. It is well known that tumour cells and stroma cells participate in molecular and functional crosstalk [4-6]. Understanding the functional consequences of this crosstalk on tumour cell metabolism, function and survival may uncover processes that are vulnerable to therapeutic intervention.

In order to target disordered metabolism in cancer, the metabolic features of any tumour must be specifically delineated. Studies on cancer metabolism have identified a number of different metabolic perturbations, including the Warburg effect and other alterations in carbohydrate metabolism, disordered lipid metabolism, altered amino acid and protein metabolism, and disturbances in nucleotide metabolism [7-9]. One consequence of perturbed metabolism is the abnormal accumulation of oncometabolites (e.g., 2-hydroxyglutarate, succinate, fumarate and sarcosine), small molecules that have the potential to initiate or sustain tumour growth and metastasis [10-12]. Importantly, these metabolic features do not appear

uniformly in all tumours. Rather, there is substantial heterogeneity in the metabolic characteristics between tumours [13,14].

Recent efforts by various groups have attempted to devise approaches to subtyping cancers according to their metabolic features [15-18]. This work has been made possible by the large and multidimensional molecular datasets made available by efforts such as The Cancer Genome Atlas and the International Cancer Genomics Consortium. In this review, we discuss potential approaches to identifying clinically meaningful metabolic subtypes of cancers, then identifying features that may be vulnerable to therapeutic attack. We discuss approaches emerging from gene-level data, transcriptional data and metabolomic data. Finally, we provide a step-by-step transcriptome based analytical workflow which we have found provides detailed insight on the metabolic perturbations present in cancer.

1.2 Diverse Metabolic Perturbations in Cancer

Cancer cells are known to have a multitude of metabolic features that distinguish them from normal cells. While it is not the intention of this research to discuss these features in detail, some of the more prominent features are described below.

1.2.1 Warburg Effect

Glucose is the most abundant metabolite in the blood and the primary source of energy in the body [19]. The end product of glucose metabolism is pyruvate, which can be further converted into lactic acid under hypoxic conditions, or to acetyl-CoA in normoxic condition. However, in cancer cells, the rate of glucose uptake increases markedly, and large amounts of lactic acid are produced, even in normoxic conditions. In the 1920s, Otto Heinrich Warburg and

Seigo Minami first observed the metabolic changes in rat liver carcinoma [20]. Afterwards, Warburg and his colleagues tested kidney and heart tissue where they found the same increase in lactate production compared to control tissues (nearly 70-fold higher than healthy liver-tissue). Thus, Warburg proposed that, even in the presence of oxygen, cancer cells utilize glycolysis followed by lactic acid fermentation to produce high rates of energy [21]. This process is known as the "Warburg Effect."

1.2.2 Other Perturbations in Carbohydrate Metabolism

Conversion of glucose is highly linked with other pathways involved in carbohydrate metabolism, as well as to most other metabolic pathways. Besides glycolysis and lactic acid fermentation, the energy derived from glucose involves oxidation to CO_2 and H_2O via the tricarboxylic acid (TCA) cycle [22]. Additionally, glucose conversion to glucose-6-phosphate (G-6-P) can lead to the pentose phosphate pathway (PPP), which runs parallel to glycolysis. The PPP results in the production of NADPH as well as ribose 5-phosphate, a precursor of nucleotide synthesis [23].

Some of the genetic features of cancers are directly responsible for alterations in carbohydrate metabolism. One example is the oncogene KRAS, which enhances glycolysis and also drives the non-oxidative phase of PPP. In contrast, inactivation of KRAS decreases expression of a number of glycolytic enzymes and reduces flux through the non-oxidative phase of PPP [26]. Another example is inactivation of tumour suppressor p53, which enhances glucose uptake in tumour cells, driving increased glycolysis and PPP [27]. The PI3K/Akt/mTOR signaling pathway is frequently disrupted in several different cancer types (e.g. with the oncogene PIKCA and

mutation of tumour suppressor PTEN). The pathway is implicated in the sensitivity of cancer cells to insulin and insulin growth factor 1 (IGF).

The perturbed carbohydrate metabolism in cancer cells has other indirect effects on the cell and the microenvironment. For example, accelerated lactic acid production leads to the production of hydrogen ions which leak into the extracellular compartment, disturbing the pH balance of the microenvironment and disrupting the tumour-stroma interface [28]. This may exacerbate the tumour's capability to invade. Altered carbohydrate metabolism results in perturbations in mitochondrial redox balance [32]. Accelerated glucose metabolism results in increased production of reactive oxygen species (ROS). The overproduction of ROS can subsequently lead to damage to cell structures as well as to cellular DNA, leading to increased autophagy in tumour cells [29-32].

1.2.3 Lipid Metabolism

Altered lipid metabolism has been observed in most cancer types. When lipid metabolism is perturbed, features may include increased uptake of exogenous lipids or de novo lipid synthesis [33]. These perturbations in lipid metabolism are generally considered beneficial to the cancer cell's capability to proliferate, invade, metastasize and survive. Accumulation or increased production of fatty acids is essential in tumour cells for energy production and protein modification [34]. Disrupted fatty acid synthase (FASN) has been associated with more aggressive biological behaviour in various cancers such as breast, colorectal, prostate [35,36]. In a study focusing on colorectal cancer, it was reported that fatty acid synthase is frequently overexpressed in colorectal neoplasms, thus inhibiting the fatty acid synthesis by enzymatic

function with metabolic analogues or decreasing the expression of FASN could be useful strategy to treat colorectal carcinoma [37]. For this study, the expression levels of fatty acid synthase (FASN) was evaluated by immunohistochemistry. Further, fatty acid synthetic activity was quantified using metabolic labelling which showed a significantly increased activity rate (6-16 fold) of FASN in colorectal carcinoma than the serosal fat ($p = 0.01$). Similarly, FASN expression levels were also found to be increased in different colorectal studies [38,39]. Likewise, another study demonstrated increased FASN expression in prostate cancer; where FASN expression was found significantly elevated from low grade to high grade prostatic epithelial neoplasia as well as invasive carcinoma [40]. Most of the critical enzymes involved in fatty acid metabolism are controlled by the sterol regulatory-element binding protein (SREBP) [41,42]. Increased SREBP induces higher FASN expression conferring a more aggressive tumour phenotype [43,44].

ATP-citrate lyase (ACLY) is frequently increased in cancers, including glioblastoma, colorectal cancer, breast cancer, and hepatocellular carcinoma [45]. ACLY catalyzes the conversion of citrate to oxaloacetic acid and acetyl-CoA. The primary function of acetyl-CoA is to deliver the acetyl group to the citric acid cycle; however, it is also pivotal in fatty acid synthesis and lipogenesis (including the synthesis of triglycerides, steroid hormones, cholesterol and bile salts). Acetyl-CoA synthetases catalyze acetyl Co-A conversion from acetate, and therefore acetate is a crucial metabolite for lipid synthesis as well as histone acetylation. Enzymes that are involved in lipogenesis, including ACLY, play a significant role in cancer cell proliferation, development and metastasis. One of the critical regulators of the epithelial-mesenchymal transition process is CTNNB1 (Beta-catenin protein-1). In vitro study confirms that ACLY can promote metastasis by

promoting CTNNB1[46]. Further, the ACLY gene knock-out experiments in different cancer cell lines showed reduced lipid synthesis and poor metastatic ability. A recent study showed that the inhibition of ACLY suppressed tumour growth and apoptosis via increased reactive oxygen species (ROS) production in LNCaP cells [47]. Similarly, knockdown of ACLY expression inhibited tumour growth in a breast cancer study [48]. Furthermore, different studies show that inhibition of ACLY markedly reduces cancer cell proliferation [49,50]. ACLY contributes equally to glucose and lipid metabolism; thus, it is a potential target in cancer treatment.

Cholesterol is essential for various cellular functions and cell growth. Cholesterol can be increased or decreased in different cancers, and the role of cholesterol in tumour biology is not fully understood. In some instances, exogenous cholesterol can enhance tumour progression. For example, in a murine breast cancer model (MMTV-PyMT), a high cholesterol diet increases tumour growth and metastasis [51]. Hypercholesterolemia accelerates the growth of androgen receptor-negative prostate cancer xenografts [52]. Cholesterol has also been shown to activate mTORC1, promoting tumour cell proliferation and metastasis [53]. In hepatocellular carcinoma, mitochondrial cholesterol is elevated due to upregulated steroidogenic acute regulatory (STAR) protein expression, and STAR knockdown results in greater sensitivity to chemotherapeutic agents [54]. In contrast to those observations, in a recent study of pancreatic cancer patients, a cholesterogenic subgroup was identified which had a better survival outcome than other subgroups; tumours with increased glycolytic flux and low levels of cholesterol synthesis had the worst prognosis [55]. Given the lack of consensus in the findings, it is apparent that a better understanding of the impact of cholesterol metabolism on tumour biology is required.

Exosomes are small membrane-bound extracellular vesicles and *have* an intricate structure containing proteins, lipids, and different metabolites [56]. Different studies show that exosomes are secreted from different types of cells, including immune cells, mesenchymal cells, as well as cancer cells. Exosomes derived from cancer cells can alter the tumour microenvironment. *In vitro* experiments show that, in pancreatic and prostate cancer cell lines, the oxygen consumption is highly suppressed by cancer-associated fibroblast secreted exosomes [57]. Lipids, including sterol lipids, sphingolipids, glycerophospholipids, play a regulatory role in exosome formation and release to extracellular microenvironment. Exosome contains a high abundance of ceramide and, depending on the ceramide content, exosomes can modify metabolism in cancer cells. *In vitro* experiments on prostate cancer cell lines demonstrated that exosomes secrete insulin-like growth factors, which activate the PI3/Akt signalling pathway [58]. Interestingly, PI3/Akt activator protein kinase C is associated with ceramide proteins which are transported by the exosomes. Another ceramide derivative is sphingosine-1-phosphate, which is critical to tumour growth as it stimulates angiogenesis [59]. Additionally, a proteomic profiling study of exosomes showed that SW620 (a metastatic human isogenic colorectal cancer cell line) cell-derived exosomes were highly enriched in lipid rafts and associated components [60]. Further, these exosome secreted lipids and proteins that influence different metastatic factors and signalling molecules that are fundamental to tumour progression. Nevertheless, the role of extracellular vesicles in metabolism is complicated and needs to be better understood.

1.2.4. Amino Acid and Protein Metabolism

Altered amino acid metabolism may support the increased energy requirements of cancer cells, and it may also provide a mechanism for cancer cells to escape the immune system [61]. Glutamine is one of the most abundant amino acids in the body, and aside from glucose, cancer cells are mostly dependant on glutamine to produce energy. Several amino acids, including arginine, aspartic acid, alanine, serine, valine, leucine, methionine, contribute in the same manner as glutamine [62].

Cancer cells are characterized by their reprogrammed glutamine metabolism, which may support cancer growth, inhibit apoptosis and induce drug resistance [63]. This may occur by several mechanisms. For instance, the binding of c-Myc to glutamine transporters results in higher glutamate production [64]. Increased glutamine uptake by cancer cells leads to increased glutamate production, which is converted to a-ketoglutarate, and these reactions replenishes the TCA cycle. Consequently, increased glutamine influx accelerates mitochondrial function, and the substrates produced support lipid metabolism. There are additional downstream effects of glutamine dependence. The amidic nitrogen of glutamine can be used in the production of asparagine and other metabolic intermediates such as nucleotides [65]. Moreover, glutamine secreted inorganic ammonia plays a role in inducing autophagy in cancer cells. Glutamate serves as a nitrogen donor which is essential for the synthesis of other non-essential amino acids as well as supports the production of NADPH.

Arginine production is frequently increased in cancer cells [66]. Arginase 1 and 2 convert arginine into ornithine, which is essential for cancer cell proliferation [67]. The role of arginine in the immune system is well established as it affects lymphocyte function; however, it is not well understood whether increased or decreased arginine production is responsible for the alteration in lymphocyte function in cancer. One of the possible reasons for arginine being found altered in many cancer types could be due to the tumour infiltrating macrophages (TIM). TIM contains a higher amount of arginase; thus, it may control the arginine availability in the microenvironment of the tumour. *Ex vivo* experiment shows that circulating arginase-1 significantly increased in glioblastoma patients with T-cell suppression [68]. Another study demonstrates similar outcomes such as, in renal cell carcinoma patients, a high amount of arginase-1 observed, and limiting the T-cell availability [69]. On the contrary, under the arginine depletion, arginase and nitric-oxide synthase get modified so that there is enough accumulation of reactive oxygen species (ROS), which negatively regulates the immune system function [70]. Arginine is a potential therapeutic target based on studies of arginine supplementation and deprivation.

Alanine is an essential source of fuel in cancer cells, and its availability is supported by internal production as well as by external sources. One mechanism by which alanine production is increased in cancer cells is related to the PIK3CA mutation, which is associated with increased expression of alanine aminotransferase (GPT2). GPT2 catalyzes the conversion of glutamate to alanine. One example of an external source of alanine is autophagy-induced production. For example, in PDAC, stroma cells secrete alanine, which is taken up by tumour cells to support their proliferation requirements [71].

Free amino acids released by protein degradation convert into glucose or ketones through the TCA cycle. Decomposed amino acids can also degrade to ammonia and hydrocarbons. High ammonia levels trigger the mTOR complex, thus the activated mTOR complex support tumour cells to proliferate as well as disrupt the urea cycle leading to further metabolic perturbations [72].

1.2.5. Nucleotide Metabolism

Nucleotides, the building blocks of nucleic acids, can be synthesized by salvage synthesis (recycling existing nucleosides) or *de novo* (using amino acids to form the purine and pyrimidine rings). In the oxidative phase of the pentose phosphate pathway (PPP), glucose-6-phosphate (G6P) is oxidized to yield ribulose-5-phosphate; and in the non-oxidative phase, ribulose-5-phosphate is converted to ribose-5-phosphate and xylulose-5-phosphate. Ribose-5-phosphate is necessary for nucleotide synthesis and histidine metabolism [73].

The molecular reprogramming that typifies cancer cells accelerates *de novo* nucleotide synthesis. C-Myc is considered a master regulator of nucleotide synthesis and it is upregulated in more than 50% of cancers [74]. Genes involved in nucleotide metabolism are primarily modulated by Myc [75]. For example, in pancreatic adenocarcinoma, the upregulation of transcriptional activation of ribose-5-phosphate isomerase promotes increased Myc resulting in increased *de novo* synthesis of purines and pyrimidines. Increased mTOR also promotes *de novo* nucleotide biosynthesis [76]. Similarly, highly activated PI3/AKT, ERK/MAPK pathways promote *de novo* pyrimidine biosynthesis, and mutated TP53 increases the gene transcriptions that enhance deoxyribonucleotide triphosphate (dNTP) synthesis [77].

1.2.6. Hypoxia

Hypoxia is a common feature of the tumour microenvironment and can affect the metabolic features of a tumour. A primary regulator of cancer cell proliferation is hypoxia-inducible factor (HIF). HIFs are responsible for forming new blood vessels under hypoxic conditions. HIF has a structure consisting of a helix-loop-helix, which is made of a heterodimeric complex of two subunits - α and β [78]. Subunit α contains three isoforms: HIF-1α, HIF-2α and HIF-3α. HIF-1α degradation is accelerated by oxygen and α-ketoglutarate, which are the intermediates of the tricarboxylic acid cycle (TCA). Several studies showed that different mutations such as p53, RAS, BC12, succinate dehydrogenase, fumarate dehydratase and isocitrate dehydrogenase, cause the activation of hypoxia-inducible factors [79]. The primary functions of HIFs is to regulate transcription of multiple genes that enhance adaptation to hypoxic environments. In cancer, HIF alters various important functions such as glycolysis, nutrient uptake, angiogenesis and apoptosis. For example, HIF-1α induces glycolytic enzymes such as hexokinase I & II, phosphofructokinase-L, and lactate dehydrogenase-A, which further promotes cell survival and migration [80]. HIF also stimulates pyruvate dehydrogenase kinase 1, which inhibits the production of acetyl CoA [81]. Acetyl CoA is essential for fatty acid synthesis. Rabinowitz et al. found an exciting source of acetyl CoA production in a number of cancer cell lines under hypoxic conditions. The major source of acetyl-CoA production is glucose and glutamine; but under hypoxic conditions this contribution was less marked. Mass-spectrometry based ^{13}C-tracer experiments revealed that in those conditions, cancer cells do not depend on amino acids or fatty acids; however, they use acetate to produce Acetyl-CoA (500μM U-^{13}C acetate which is

responsible for a 50%-86% increase in acetyl CoA labelling) [82]. Another study demonstrates that under hypoxic stress, the mitochondrial enzyme SHMT2 is formed. SHMT2 is crucial for sustaining the production of NADPH and redox balance to support tumour cell survival and growth [83]. Cancer cells possess diverse mechanisms to adapt to different microenvironments. With advanced technology such as mass spectrometry, it is now possible to determine how cells use different metabolic pathways to produce energy. For example – a recent study shows that under normal conditions, NADPH is produced mostly by the malic enzyme in mouse adipocytes, whereas hypoxia perturbs the metabolic pathway and switches the primary NADPH source to oxidative phosphorylation (OXPHOS) [83,84]. There are numerous experiments that suggest that HIF-1 plays a vital role in cancer metabolism as a master regulator of various proteins and enzymes associated with energy metabolism.

Pyruvate kinase M2 is an essential metabolite that catalyzes the conversion of phosphoenol-pyruvate to pyruvate by transferring the phosphate group to ADP in glycolysis. After an enormous amount of experiments and studies, it is well established that PKM2 plays a vital role in cancer cell metabolism, transcriptional regulation as well as different extracellular signaling. Expression of PKM2 inhibits the consumption of oxygen in cancer cells by activating the HIF-1 and leading to an excessive amount of lactate production. Thus, PKM2 is crucial for the Warburg effect. Not only that, low expression of PKM2 dimers promotes the accumulation of metabolites involved in glycolysis such as glucose-6-phosphate, 3-phosphoglycerate [85]. Increased glucose-6-phosphate leads to the pentose phosphate pathway, and 3-phosphoglycerate promotes serine synthesis. PKM2 also regulates the activity of different transcriptional factors, including HIF, STAT3, β-

catenin. In cancer cells, PKM2 mediates increased gene expression are required for tumour growth [86,87].

1.2.7. Autophagy

Autophagy refers to a cell's self-digestion process where damaged organelles, proteins, mitochondria, lipids, different cytoplasmic materials are isolated into vesicles for degradation and recycling [88]. The role of autophagy in cancer is complex and not fully understood as it can both promote and suppress tumour growth. The products of autophagy (amino acids, fatty acids, and nucleosides) can be utilized by cancer cells to support their high energy demands. Autophagy can be stimulated upon activation of different signaling pathways, which are activated as a result of energy depletion, stress, hypoxia, and limitation for insulin, growth factors or hormones [89]. As the pancreas is the primary source of digestive enzymes and hormones involved in glucose metabolism, high levels of autophagy are observed in most pancreatic tumours under basal conditions compared to the other epithelial cancers [90]. In such condition, autophagy may contribute to pancreatic tumour progression, as *in vitro* knockout of autophagy-related genes such as ATG5, ATG3 inhibits tumour growth.

Autophagy may also suppress tumor growth in some instances. Microtubule-associated proteins 1A/1B light chain 3B (LC3) is an important regulator of autophagy. Increased levels of LC3 expression are associated with elevated levels of hypoxia markers in tumour cell lines and experiments on primary tumours [91]. Following nutrient deprivation, p53 maintains autophagic homeostasis by regulating the autophagy-related protein LC3, which supports cancer cell survival [92]. Another critical regulator of autophagy is beclin-1 (BECN1). BECN1 is frequently altered in

different cancer types, including gastric and colorectal cancer [93,94]. Studies show that mutation in BECN1 related protein UV-radiation resistance-associated gene (UVRAG) reduces autophagy leading to increased cancer cell proliferation, particularly in colorectal cancer [95]. However, *in vitro*, studies show that knockout of autophagy-related genes significantly inhibits tumour cell growth [96].

1.3 Metabolic Subtypes and Phenotypes

There appears to be considerable diversity in the metabolic features of cancers. For example, the Warburg effect is not universally present in all cancers. It is unknown whether this diversity is clonal in nature, to what degree metabolic diversity exists within a tumor, and whether metabolic features evolve over time. With advanced technology and computational methods, we have begun to appreciate the existence of different metabolic phenotypes. For instance, based on metabolite profiling in cancer cell lines, three distinct metabolic subtypes of pancreatic cancer were described: slow-proliferative, glycolytic and lipogenic [97]. The existence of various metabolic subtypes forms the rationale for applying different metabolic strategies for different tumours.

1.4 Targeting Metabolism

Metabolic targeting for cancer therapy is an emerging field that is currently under investigation to identify the small molecules that may specifically block the fundamental metabolic steps linked to tumour growth. Since, glycolysis plays an essential role in the development of cancer, attenuation or inhibition of glycolysis may be useful for the prevention of abnormal cell proliferation, and invasion, as well as metastasis. Several enzymes are the

primary drivers of glycolysis, such as hexokinase (HK), OFK, and pyruvate kinase. Hexokinase is a potential target for cancer metabolism as several types of cancer exhibit high levels of HK II [98]. For example- in lung, breast or brain cancers, the deletion of HKII is beneficial. Phospho-fructokinase 1 was found to increase in many types of cancer. There are some clinical trials currently underway with small-molecule PFKFB3 inhibitors. Another critical step is when glycolysis derived pyruvate can either be imported into the mitochondria to be oxidized in the TCA cycle or converted to lactate in the cytosol. In mitochondria, pyruvate is converted to acetyl CoA, and the enzyme responsible for regulating this crucial junction in pyruvate metabolism is the pyruvate dehydrogenase complex. Pyruvate dehydrogenase kinase (PDHK) performs inhibitory phosphorylation, which reduces the activity of pyruvate dehydrogenase (PDH), resulting in a decrease in pyruvate flux and an increase in lactate production. The expression of PDHK in various types of cancer makes it a potential therapeutic target [99]. Lactate dehydrogenase complex also plays a crucial role in regulating the fate of pyruvate in cancer. Inhibition of LDHA by small molecule inhibitors or genetic approaches results in slowed cancer cell growth and increased cell death, in hepatocellular carcinoma and breast cancer [100]. Several early-stage clinical trials are underway to assess the efficacy of some LDH inhibitors. The pre-clinical development of inhibitors that have more specificity for LDHA is currently ongoing [101]. Limiting glutamine availability is another attractive strategy. This may be accomplished by inhibition of glutamine transporters. Another approach is to inhibit glutamate dehydrogenase (GDH) to limit glutamine production. Currently, there are no small molecules available to block

the production of GDH, but the initiative to develop some specific GDH inhibitors may allow more effective targeting of glutamine flux into the TCA cycle.

Targeting small metabolites present in the TCA has demonstrated some success. Recently there has been success in pre-clinical and clinical settings with novel compounds that inhibit the gain of function activity of mutant IDH. This mutant isocitrate dehydrogenase (IDH) was shown to dramatically reduce the production of 2-hudroxyglutarate (2HG) and cause cancerous cells to differentiate towards a normal phenotype [102]. There is an early phase trial, ongoing with the small inhibitor of mutant IDH2, AG-221.

Chapter Two: Research Hypothesis and Aims

2.1 Research Hypothesis and Specific Aims

It is apparent that there is considerable metabolic diversity in cancers. If it were possible to delineate the metabolic features of a cancer, then it may also be possible to identify therapeutic vulnerabilities based on these metabolic features. I hypothesize that there are metabolic subtypes of cancer, and the metabolic features of these subtypes influence their clinical and biological behaviour.

Specific aims for this research are –

1. An effective analytical workflow will be developed to identify different metabolic subtypes in cancer based on transcriptome
2. A method will be established to identify different metabolic subtypes in Breast cancer using the bulk tumour
3. A systemic approach will be developed to identify aberrant metabolic pathways within the tumour and stroma compartments focusing on PDAC

To test the hypothesis, I focus on two different types of cancer, including breast cancer (BRCA) and pancreatic ductal adenocarcinoma (PDAC). I present an effective analytical workflow that can be applied to any tumour type to identify metabolic subtypes and their biological features based on the transcriptional features [**Chapter Three**]. I first apply the analytical workflow to breast cancer using principal component analysis to identify metabolic subgroups on the bulk RNA-Seq data [**Chapter Four**]. I then apply the workflow to deconvolved RNA-Seq data from pancreatic cancer patients, which enabled me to determine the relative contributions

of tumour and stroma to phenotype. For this, I used the non-negative matrix factorization (NMF) method to identify metabolic subgroups [**Chapter Five**].

Chapter Three: An Analytical Workflow for The Characterization of Metabolic Subtypes of Cancer

3.1 Introduction

The quintessential metabolic feature of cancer is encapsulated by the Warburg effect, which is the preferential breakdown of glucose using anaerobic pathways, even in normoxic conditions. However, cancers have a wide variety of metabolic features that may correlate with other biological functions. In our efforts to identify metabolic subtypes of cancer, we have devised a bioinformatic analytical workflow that can be used to explore the metabolic features of a particular cancer type. Initially, our approach involves a focused analysis of the transcriptome. Metabolic subtypes of cancer are identified using unsupervised clustering methods. The metabolic subgroups so identified are then used to inform subsequent supervised analyses, including gene set enrichment analysis and pathway analysis (to delineate biological functions), linkage to clinical outcomes, as well as data deconvolution to determine the contribution of various cell types comprising the tumour microenvironment. Other molecular features can be explored based on the metabolic subtypes identified, including mutations, copy number variations, methylation and noncoding RNA. Finally, metabolomic features can be explored based on the metabolic subtypes initially identified. Evaluation of actual metabolites is essential for understanding the net result of the various upstream molecular events. This step-by-step analytical workflow provides detailed insight into the metabolic perturbations. It is possible that identifying the metabolic features that characterize a particular tumour will provide insight into biological pathways that are vulnerable to treatment.

3.2 Metabolic Subtypes and Phenotypes

The various metabolic perturbations described above have been variably described in diverse cancer types. However, not all metabolic features are similarly operative in a uniform fashion. Instead, what has become apparent is that there is considerable diversity in the predominant metabolic features of any particular tumour. In some instances, the metabolic features of tumours are linked with alterations in their biology and clinical behaviour. To derive clinical relevance from our knowledge of the range of metabolic derangements that can occur in cancer, it will be imperative to identify metabolic subtypes and to identify mechanisms by which those variants emerge. Metabolic subtypes that confer aggressive biological and clinical behaviour are particularly interesting, since inhibiting metabolic features that correlate with biological aggressivity would most likely have therapeutic benefit. Only when this detailed analysis has been performed will we be able to target the appropriate metabolic derangements in a personalized fashion.

Recently, it has become possible to comprehensively characterize tumours based on their genome, epigenome, transcriptome, proteome, and metabolome. Examples of each have been previously published [103-107]. Each level of biological information contributes to metabolic phenotype. However, the problem is that the phenotype is not a result of linear biological information flow (i.e., from the genome to transcriptome, to proteome, and metabolome). Instead, the metabolic phenotype is a culmination of multiple upstream molecular events that may occur in parallel. In many circumstances, the metabolic phenotype may be the net effect of opposing metabolic features.

The question is, therefore, how best to define metabolic subtypes to ensure that the subtypes accurately reflect the phenotype. Moreover, what approach would provide the greatest insight on pathways that represent vulnerabilities that can be targeted therapeutically? Several approaches have been taken to delineate metabolic subtypes, and these will be described below.

3.2.1 Genome-based Subtyping

One approach to identifying tumour subtypes is to classify based on alterations at the genome level, including mutations and copy number variations. There are several examples of mutations and gene fusions involving genes that have a primary metabolic function, including mutations in genes encoding metabolic enzymes such as SDH and IDH1/2 [108,109]. Mutations in classical oncogenes may also have metabolic effects. Examples include KRAS, BRAF and Myc, which cause secondary alterations in the expression of multiple genes with a direct role in metabolism [110]. There are several problems with a genome-based approach to metabolic subtyping. Most importantly, this approach does not take into account the multiple layers of control that affect gene expression, including epigenetic control mechanisms, miRNA, long noncoding RNA and circular RNA. In addition to alterations at the protein level, genome-based subtyping is unlikely to reflect the metabolic phenotype. Another problem with this approach is that mutations and copy number variations in each gene with a significant metabolic function (such as genes encoding metabolic enzymes) are not that prevalent. Finally, multiple combinations of gene alterations are possible. These latter two problems make it extremely difficult to categorize tumours based on metabolic features. One interesting approach that was recently described was reported by Sinkala et al., who dichotomized tumours based on the

frequency of mutations and copy number variations in metabolic genes [111]. That approach enabled the identification of a subgroup of cancers that were associated with a high rate of mutations in metabolic genes, which was sometimes associated with clinical aggressivity. This approach also suggested that tumours with a high frequency of gene alterations were more sensitive to (which drugs).

3.2.2 Transcriptome- based Subtyping

The transcriptome, like the genome, is subject to modification. Therefore, there can be a discrepancy between mRNA levels and corresponding proteins. However, this discrepancy is not universal (i.e., the correlation between mRNA and protein levels varies). Despite the imperfect correlation of transcription and translation, the pattern of gene expression *as a whole* can paint an accurate picture of the metabolic processes that are perturbed. Several groups have compiled lists of genes that function to modulate metabolism, including genes that have secondary metabolic effects [15,111,112] By interrogating gene sets with specific metabolic functions, it may be possible to subclassify tumours according to their predominant metabolic features. This approach has been used to identify metabolic subtypes of tumours with biologically distinct features [15].

In the context of cancer, glucose metabolism has been most studied, as it has a direct or indirect connection to all other metabolic pathways in the body. Variations in glycolytic subtypes have been described. Follia et al. found four distinct metabolic subtypes of pancreatic ductal adenocarcinoma by focusing their classification on core glycolytic genes (high glycolytic, very high glycolytic, low glycolytic and very low glycolytic) (total patients, n = 275) [113]. They used the

RNA-Seq data from publicly available datasets; they also used metabolomic data of pancreatic cancer cell lines to determine the metabolites that are produced by the identified metabolic subtypes. Further, metabolic proteins were measured in patients before and after chemotherapy using Q-TOF mass spectrometry. The high and very high glycolytic subtypes had a significantly worse prognosis and contained relatively little immune infiltrate. Higher expression levels of glyceraldehyde-3-phosphate dehydrogenase (GAPDH), triosephosphate isomerase 1 (TPI1), Forkhead Box M1 (FOXM1) were observed in the worst subtypes.

Karasinska et al. recently identified four distinct metabolic subtypes of pancreatic ductal adenocarcinoma (quiescent, glycolytic, cholesterogenic and mixed), focusing on glycolytic and cholesterogenic genes [55]. They utilized the genomic, transcriptomic and survival data of 325 resectable and non-resectable PDAC tumours. The glycolytic subtype was associated with the shortest median survival, and the cholesterogenic group had the best overall survival.

Bidkhori et al. identified three distinct metabolic subtypes of hepatocellular carcinoma (HCC) based on the genome-scale metabolic networks [114]. Metabolic subtypes were identified using transcriptomic data, genome-scale metabolic networks and network controllability analysis. Significant differences in overall survival were observed across the identified subtypes. The study highlights some of the critical pathways that were significantly altered in identified subtypes such as kynurenine metabolism, lipid metabolism, WNT/Beta-catenin signalling as well as PI3/AKT/mTOR signalling pathway.

Recently Peng et al. classified almost 10,000 tumors from 33 cancer types based on the pattern of expression of genes involved in metabolism [15]. Metabolic genes were organized into seven major metabolic pathways, and potential master regulators were identified.

3.2.3 Metabolome-based subtyping

The metabolome is the net product of many parallel and competing metabolic pathways. It has been suggested that the metabolome is the closest representation of the metabolic phenotype, as it represents the final manifestation of the many co-mingled transcriptional and translational modifications that characterize a particular tumour. Therefore, it may be argued that achieving a valid metabolic classification of tumours would be better achieved by focussing on metabolites as opposed to transcripts.

One example of using the metabolome to classify tumours was reported by Daemen et al. [97], who identified three distinct subtypes of pancreas cancer based on metabolite profiling (slow-proliferative, glycolytic and lipogenic). These subtypes strongly correlated with previously described pancreatic cancer subtypes (epithelial and mesenchymal variants). From the findings, they provided the essential predictive utility of different metabolic inhibitors that are currently under clinical trials. Similarly, based on metabolite profiling, Haukass et al. identified three distinct metabolic subtypes of breast cancer (Mc1, Mc2 and Mc3) [115]. Metabolic subtypes identified by this approach had distinct upstream molecular features including differences in proteins and genes that are related to the extracellular matrix as well as metabolic pathways.

Recently, Li et al. demonstrated the metabolic diversity of cancer based on metabolite profiling of well characterized cancer cell lines [116]. There were significant links between the

metabolome and genomic alterations such as mutations, copy number variations, and epigenetic features. For instance, one cancer subtype was distinguished by increased kynurenine secretion resulting from the degradation of tryptophan, an essential amino acid. Kynurenine governs inflammation, sustaining immune escape in cancer cells [117].

One challenge with using the metabolome to subtype cancers is related to the limitations of studying the metabolomic variations in tissue. To obtain an accurate description of the intracellular metabolome in cell lines, a quenching step is critical to halt the metabolic activities within a cell without affecting cell membrane integrity [118]. The methods for studying the metabolome in tissue extracts have been well described (119), although the low abundance of metabolites and low sample volumes make the analysis difficult [120,121]. The problem with doing this in tumours is the cellular heterogeneity inherent in whole tumours. That is, a tumour is comprised of cancer cells and multiple stromal elements. Studying whole tumour extracts will not delineate the cellular origin of metabolic processes, which may vary quite widely. To address this, advances in single-cell metabolomics will be required [122], which will also require the development of improved methods of extracting cells from whole tissues in a manner that will not disturb the fundamental metabolic processes. In contrast, methods for single-cell transcriptomic analysis are well established. Moreover, even in whole tissues, the contribution of specific cell compartments can be studied using deconvolution methods [123,124].

There are additional challenges with using metabolomic data to subclassify tumors. First, there is the lack of granularity in metabolomic datasets. No matter what analytical platform is used, it is impossible to fully annotate the metabolome. Even the many features detectable on

mass spectrometry-based platforms are not entirely identifiable as distinct metabolites. With the exception of NMR spectroscopy, metabolomics data are semiquantitative. Often, the differences in abundance of individual metabolites are quite subtle. Because of these technical limitations, we are less enthusiastic about defining metabolic subtypes of tumours using metabolomic data. Rather, we prefer to focus initially on the transcriptome.

3.3 Identification of Subtypes

Unsupervised analysis is the most common initial approach to identifying subtypes based on patterns of molecular features. The subject has been previously been reviewed in detail [125]. In brief, there are multiple clustering methods available. Partitioning methods may be used to divide a dataset into non-overlapping subsets [126]. Hierarchical clustering is one of the partitioning methods that organizes the clusters as a classification tree [127]. Another clustering method, density-based clustering, uses a non-parametric algorithm to group the data points together that are close to eachother and highlights the outliers that lie alone in a low-density area [128]. The choice of clustering method depends to a degree on the type and format of the dataset. Ultimately, whatever method is required, a predictive model will need to be constructed to validate the approach on an external dataset.

One of the most common approaches to identifying clusters involves principal component analysis (PCA)-based hierarchical clustering. PCA reduces the number of variations in a dataset while containing the most information [129]. This approach was taken to classify breast cancer molecular subtypes [130]. Non-negative matrix factorization (NMF), partitioning clustering

method, is another common approach [131]. The NMF method enabled identification of three distinct molecular subtypes of pancreatic cancer (classical, exocrine, quasi-mesenchymal) [106].

3.4 Step-by-Step Analytical Workflow for Metabolic Subtyping

3.4.1 Unsupervised Analysis Focussing on Metabolic Genes

If the goal is to identify metabolic subtypes, then classification should be based on a focussed gene set consisting of genes that have a known function in metabolism. These genes have been annotated in the Reactome and KEGG pathway databases.

The filtration process is an essential step in analyzing the dataset as it allows a more focussed analysis of a smaller, more parsimonious analysis of a smaller gene set that is more likely to contain valuable information. Removing uninformative genes reduces noise. The metabolic gene list can be filtered using various statistical approaches, such as median absolute deviation, which excludes less variable genes from a large dataset.

3.4.2. Define Significant Clusters

To define the significance of clusters, it is important to find out how good the model fits in the given dataset and its predictability. One possible way to determine the best-fit model could be calculating the coefficient determination (R^2). R^2 closer to 0 means that the dependent variable can not be predicted from an independent variable; and R^2 closer to 1 means that the dependant variable can be predicted by the independent variable (perfect fit and reliable model). To test the predictability of identified clusters, the Silhouette algorithm could be used, which validates the consistency within clusters. This method determines how similar an object is to its

cluster comparing other clusters, for example, a higher silhouette score (closer to 1) means that the object is well-matched to its cluster.

Once clusters have been identified, they can be visualized using a heatmap. Two things need to be considered carefully to create heatmap – distance metric selection and linkage method selection [132]. The distance metric is a function that calculates the distance between different groups, basically measuring how close the groups are — examples of distance metric including Pearson, Spearman-rank Correlation, Manhattan, Euclidean.Manhattan Distance: Measuring the distances between two different points as a grid-like path-

$d1 \equiv dSAD:(x,y) \mapsto \|x-y\|1=\sum i=1n|xi-yi|$ [133]

Euclidean Distance: Most common methods out of all, it measures the distances between two data points in a plane-

$$d_2: (x, y) \mapsto \left||x - y|\right|2 = \sqrt{d_{SSD}} = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$ [134]

Linkage clustering methods available include average, complete and single linkage clustering. Average linkage clustering uses the average distance between different clusters. The complete linkage refers to the longest distance between two points in each cluster, while the single linkage clustering refers to the shortest distance [135]. As mentioned earlier, due to variances in the different datasets, it is challenging to say which method is more suitable. However, any of the distance metric and the linkage method could be used to define the significance of the clusters. We tried all of the approaches, and for our particular dataset, Manhattan distance metric and complete linkage clustering turned out to be the best fit.

3.4.3. Determine Clinical Significance

Once the metabolic subtypes have been identified, it is essential to verify clinical relevance. Perhaps the most obvious approach to detect clinically significant biological differences is to determine if the clusters differ in clinical outcomes such as overall survival (OS) and progression-free survival (PFS). Depending on the tumor type, OS and PFS may reflect different outcomes. For example, in tumor types where chemotherapy is administered, PFS may reflect sensitivity or resistance to chemotherapy.

3.4.4. Interrogate Biological Significance

The biological significance of identified subtypes can also be evaluated using supervised analyses, such as gene set enrichment analysis. Gene set enrichment analysis (GSEA) provides positively and negatively core enriched genes involved in specific pathways. Pathways that are highly perturbed in each identified subtype can be further investigated using literature-based informative pathway analysis such as Ingenuity Pathway Analysis (IPA) and KEGG pathways. Additionally, other molecular features such as mutation frequency, copy number variations (CNVs) and methylation patterns, can be assessed for each metabolic subtype.

3.4.5 Validation wth External Datasets

The predictiveness of a model must be tested by validating across a broad set of clinical settings and in different populations. There are many public resources available for validation, including The Cancer Genome Atlas Project (TCGA), International Cancer Genome Consortium (ICGC).

3.4.6 Evaluating the contribution of various cell types

Heterogeneity is one of the most challenging features of a cancer tissue. Bulk tumour RNA-Seq data include RNA from a variety of cell types, including tumour cells and stroma cells. So, the outcome of the analysis may not be accurate as it is a mixed result of cancer and non-cancerous cells. With advanced deconvolution techniques recently developed, it is possible to evaluate the contribution of various cell subsets. DeMixT or Cibersort can perform a three-component deconvolution which provides the expression metrices of tumour, stromal and immune compartments [124,125]. Thus, the identified subtype method can be used on different compartments of a bulk sample, which eventually represents the broad spectrum of different cell functions.

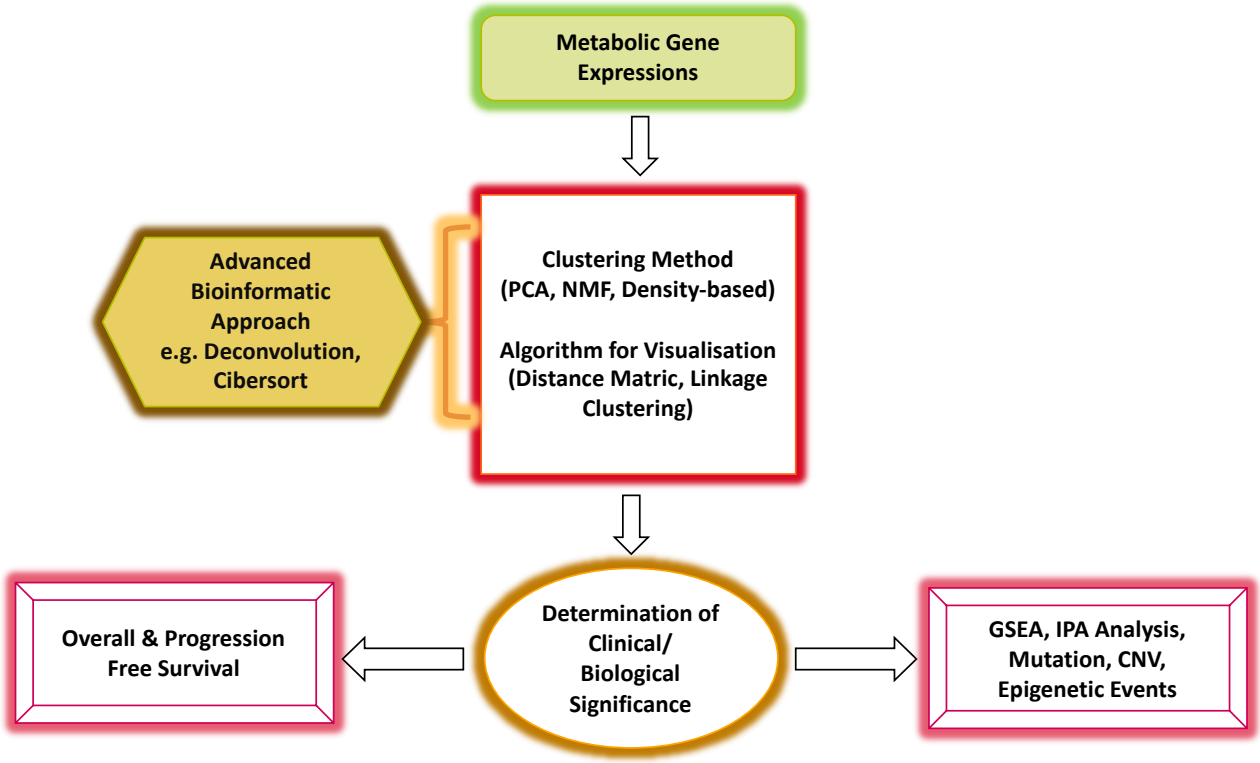


Figure 3.1: Step-by-step analytical workflow based on metabolic gene expressions

3.5 Workflow Overview/ Discussion

Our approach involves a focused analysis of the transcriptome, as methods for whole transcriptome analysis are well established, the data are easily normalized, and the features are sufficiently rich to detect patterns. Our approach involves several steps: unsupervised analysis to identify metabolic subgroups; refinement of the classification by optimizing the number of clusters; verification of clinical and biological relevance; and validation. Validated subgroups can be further characterized by identifying associated mutations, copy number variations, and epigenomic events.

3.6.1 The problem of heterogeneity

One constant feature of cancer is cellular heterogeneity within a given tumour. Heterogeneity could be a result of different genetic, transcriptomic and/or metabolic changes. Metabolic heterogeneity is crucial as it profoundly influences therapeutic susceptibilities. One of the prominent features of cancer cell metabolism is increased glucose consumption and acid production resulting in an acidic microenvironment. Adaption to the changes enables cancer cells to drastically change their behavior, which depends on variations in the tumour cell phenotypes and their non-uniform distribution within the tumour. Previously, *in vitro* experiments on breast cancer cells show that it is possible to revert the cancer cells to healthy cells phenotype only based on the influence of extracellular conditions [136]. The variety of carbon substrates that drive neoplastic cells suggests metabolic heterogeneity, even in tumours that share the same clinical diagnosis. For example, using liquid biopsy-based single-cell metabolic phenotyping on 32 lung adenocarcinoma patients shows diverse heterogeneity of tumour cells that are highly associated with glycolysis and mitochondrial oxidation [137]. There are different subsets of melanoma that found to be more dependent on oxidative phosphorylation than glycolysis [148]. Regulators of the cell cycle promote the metabolism of oxidative phosphorylation, and OXPHOS metabolism supports cancer cells to migrate and invade.

Furthermore, ROS produced by cancer cells has a massive impact on cancer-associated fibroblasts (CAFs) as they provide fuel to tumour cells supporting their growth and survival. Rewiring of energy metabolism system is not only limited to tumour cells but also immune cells, CAFs. Besides glucose, there is increasing evidence that nutrient sources that support malignant

cell functions have a profound effect on the tumour microenvironment. Different cell types within the bulk tumour promote resistance to treatments. Thus, the biology of different stroma cells, as well as the crosstalk between the stroma cells and tumour cells, need to be better understood for the right selection of treatment procedure. Therefore, single-cell analysis could be a potential approach to explore the role of different cells present within the tumour microenvironment.

Chapter Four: Identification of metabolic subtypes in breast cancer (BRCA)

4.1 Introduction

Breast cancer is the second leading cause of cancer-related deaths in women [139]. The mortality rate is worst for locally advanced and metastatic breast cancer. For most of the last decade, hormone receptor status and the expression of surface molecules have dominated treatment options. Previously, in breast cancer, several different subtypes have been proposed based on the clinical, histology and molecular heterogeneity [140-143]. Recent advances have led to the development of therapeutics, including chemotherapy and radiation therapy, that target each of these molecular subtypes. However, not all individuals respond to current treatments and the development of new treatments requires a deeper understanding of breast cancer heterogeneity. In cells, the level of hormones, including estrogen, progesterone and androgen, affect various energy transporters and the expression of multiple metabolic enzymes. The rewiring of energy metabolism is one of the hallmarks of cancer and suggests that there is a connection between disrupted metabolism and drug resistance. This connection means that one possible therapeutic route could be targeting metabolic machinery of different malignant cells. This approach requires a better understanding of the specific metabolic characteristics in any individual.

Over the past two decades, highly dimensional datasets have been generated that have demonstrated the existence of multiple molecular subgroups of breast cancer. For example, Perou et al. identified five molecular subtypes of breast carcinoma based on gene expressions including basal-like, Luminal A, Luminal B, HER2 enriched and normal-like [140]. The identified

molecular subtypes are highly prognostic and represent excellent biological diversity. All these subtypes present different clinical features and therapeutic responses and may utilize metabolic pathways differently to maintain their activity. Thus, some researchers have taken the approach to identify metabolic alterations associated with the molecular subtypes, while others have tried to identify metabolic subtypes based solely on metabolic gene expressions [144-149].

In 2013, Kim et al. explored the role of glutamine related proteins in the molecular subtypes of breast cancer [149]. The study was focused on tissue microarray of 702 breast cancer patients. Using immunohistochemical staining, they found an HER2 enriched molecular subtype that has higher expression of glutamate dehydrogenase (GDH), and amino acid transporter-2 (ASCT) comparing to other molecular subtypes. Overall, the study shows differential expressions of glutamine metabolism-related proteins among the molecular subtypes of breast cancer. Similarly, a mass spectrometry-based study showed alteration in glutamine and beta-alanine metabolism in estrogen receptor-positive (ER+) and negative (ER-) breast cancer [150]. In 2013, Budczies et al. investigated 204 ER+ and 67 ER- breast cancer tissue samples and conducted GC-TOFMS based metabolomics experiment where they found enriched glutamate and reduced glutamine in ER- tumor tissue samples compared to ER+ and healthy breast tissues. This finding suggests that glutaminase inhibitors could be potential targets in ER- breast cancer patients. Furthermore, their analysis highlights a strong correlation between beta-alanine and 4-aminobutyrate aminotransferase (ABAT); low ABAT expression is associated with shortened recurrence-free survival in both ER-positive and negative breast cancer. Another study based on ultra-high-performance liquid chromatography-mass spectrometry (UHPLC-MS) experiment

comprising 267 human breast cancer tissue samples revealed increased palmitate containing phosphatidylcholines in breast tumour samples [151]. Palmitate is a critical metabolite that is produced by fatty acid synthesis and found to be up-regulated early in the tumour progression [152-155]. Similarly, in 2016, Haukass et al. revealed three distinct subtypes (Mc1, Mc2, and Mc3) in breast cancer patients based on metabolic profiling [156]. Identified subtypes show differences in breast cancer-associated proteins and genes that are related to the extracellular matrix as well as metabolic pathways. A combination of the metabolic, transcriptomic and proteomic profiles of breast cancer provide a bigger picture of heterogeneity and reveals metabolic pathways that are susceptible to different metabolically targeted drugs.

Previous studies have described several genes that are expressed differently in breast cancer as well as several transcriptional dysregulations of different metabolic genes [157,158]. Since gene expression has a convincing association between the oncogenic drivers and metabolic phenotypes, it is essential to look into the gene expression profile of different cancer patients to get a better understanding of cancer metabolism. However, distinguishing these significantly expressed genes is challenging. For example, classification of patients based on their significant metabolic gene expressions more precisely is questionable. To address this issue, we determined whether there were systematic differences in expression of genes that were involved in breast cancer metabolism.

4.2 Materials and Methods

4.2.1 Patient samples and gene list

For this analysis, normalized gene expression data (RSEM RNA-Seq) of breast cancer patients (n= 1,192) were downloaded from The Cancer Genome Atlas Project (TCGA). Male patients (n=12) were omitted for the analysis. Overall, 1081 tumour and 111 matched healthy adjacent breast tissue was identified. The metabolic gene list (1,847 genes) was aggregated from the Reactome pathway database. The RNA-Seq dataset was composed of 14,375 genes in total and 1,439 metabolic genes.

4.2.2 RNA-Seq data Normalization

TCGA RNA-Seq data was normalized using winscaling. Winscaling is a normalization method where the data set is auto scaled with a trimmed standard deviation and a trimmed mean. The winscaling method involves identifying the distribution and calculating mean and standard deviation (SD) based on the mid 95 percentile of values.

4.2.3 Metabolic Subgrouping

Principal component analysis based hierarchical clustering was performed on the metabolic genes. Three distinct metabolic subtypes were identified based on the method. Metabolic clusters were defined using the agglomerative clustering algorithm and average linkage used for the hierarchical clustering of the genes. The heatmap was derived using Qlucore software. The linkage criterion determines the distance between sets of samples (variables) as a function of the pairwise distances between variables. If the objects or clusters just merged are indexed by i and j, and if k is any other object or cluster, and if s (i) denotes the

number of elements in cluster i, then the linkage criteria available in Qlucore Omics Explorer for computing the distance d between I + j and k defined as follows.

Average-linkage

$d(i+j,k) = (s(i)*d(i,k) + s(j)*d(j,k)) / (s(i) + s(j))$

4.2.4 Survival Analysis

To determine the clinical relevance of identified subtypes, Kaplan-Meier plots were generated for both overall survival and five-year progression-free survival. PRISM version 8.0.1 was used to perform the survival analysis.

4.2.5 Clinical Data Analysis

All the clinical information was obtained from the TCGA official site. This included age, race, T stage, N stage and M stage. To test if the identified subtypes have any significant association with the following clinical features, individual enrichment tests were performed using IBM SPSS version 24.

4.2.6 Mutation Analysis

Mutation data were collected for all the patients from cBioPortal. For each metabolic subtype, topmost mutated genes were tested using R - Maftools package.

4.2.7 Gene set enrichment analysis and pathway analysis

To investigate the different pathways that were dysregulated in the identified metabolic subtypes, gene set enrichment analysis and literature-based informative pathway analysis was conducted. For the GSEA, seven major metabolic pathways (carbohydrate, TCA cycle, amino acid, nucleotide, energy, vitamin & cofactor, and lipid metabolism) and seven hallmarks of cancer

(angiogenesis, apoptosis, G2M checkpoint, inflammatory response, DNA repair, invasion & metastasis and EMT) were explored. Furthermore, different functions, including cancer hallmark associated functions and canonical pathways, were investigated across the metabolic subtypes. For the canonical pathway analysis, Ingenuity Pathway Analysis (IPA) 2018 was used.

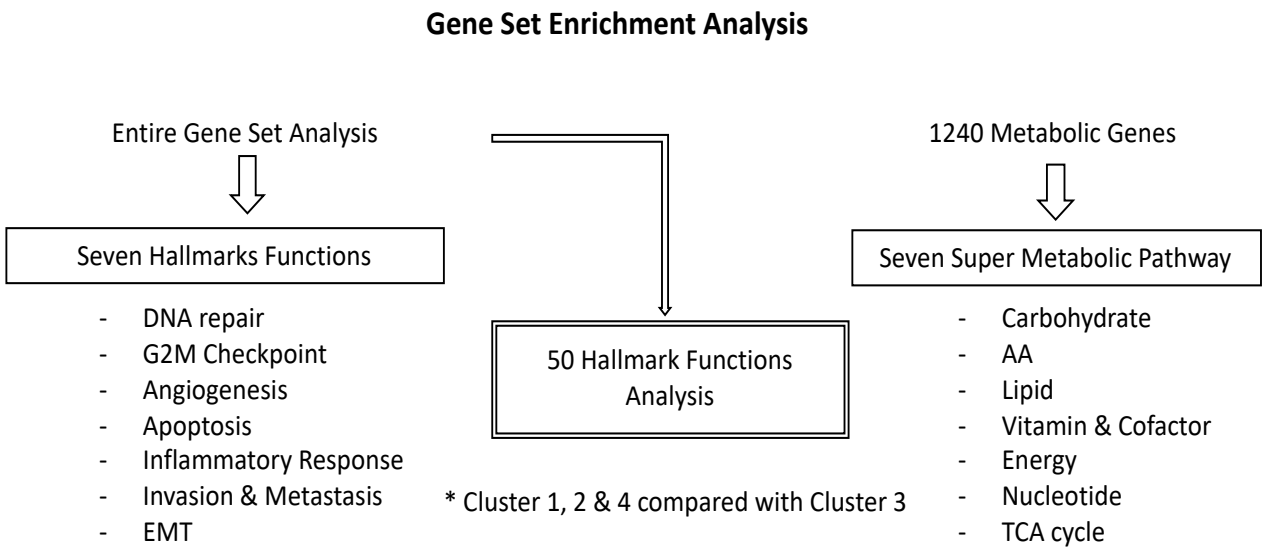


Figure 4.1: Gene set enrichment analysis workflow.

Metabolic genes were used to explore major metabolic pathways and the entire gene set was utilized to investigate seven hallmarks of cancer and 50 hallmark functions. Pathways were considered significantly enriched based on the FDR <0.25 for each metabolic subtype (B1, B2 and B4) compared to the healthy group (B3).

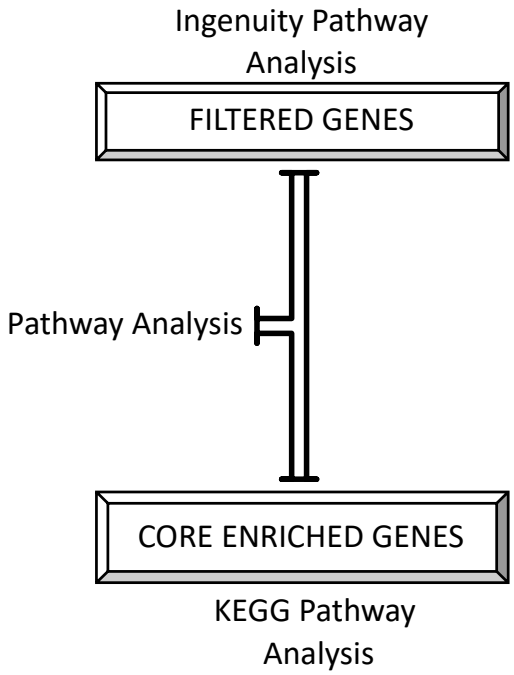


Figure 4.2: Literature-based pathway analysis workflow.

The entire gene set was filtered based on a two-tail t-test and corrected for multiple comparisons (FDR). FDR value and expression fold-changes were used to investigate the functional state of each pathway. Canonical pathways and cellular and molecular functions were explored using Ingenuity Pathway Analysis (IPA). Gene expression fold changes of filtered genes were used for the KEGG pathway analysis (Pathview software). Pathways were considered significantly enriched based on the FDR < 0.25 for each metabolic subtype (B1, B2 and B4) compared to healthy breast tissue group (B3).

4.3 Results

4.3.1 Metabolic gene expression of BRCA patients reveals three distinct subtypes

To identify metabolic subtypes, we utilized the RNA-Seq data from resectable BRCA patients to look at metabolic gene expression. The resectable dataset included 1,081 female patients and 14371 genes. The metabolic genes list was aggregated from different metabolic pathways identified in the Reactome pathway database, including carbohydrate (n=292), lipid (n= 777), amino acid (n= 324), TCA cycle (n= 164), nucleotide (n= 94), energy (n= 88) and vitamin and co-factor (n= 108). We performed an unsupervised analysis to identify different metabolic subgroups of BRCA. Principle component analysis based agglomerative clustering algorithm shows three distinct BRCA metabolic subtypes (B1, B2 and B4); the normal breast tissue group was labelled B3 (n=117). Metabolic gene expressions across the metabolic subtypes are visualized in Figure 4.3.

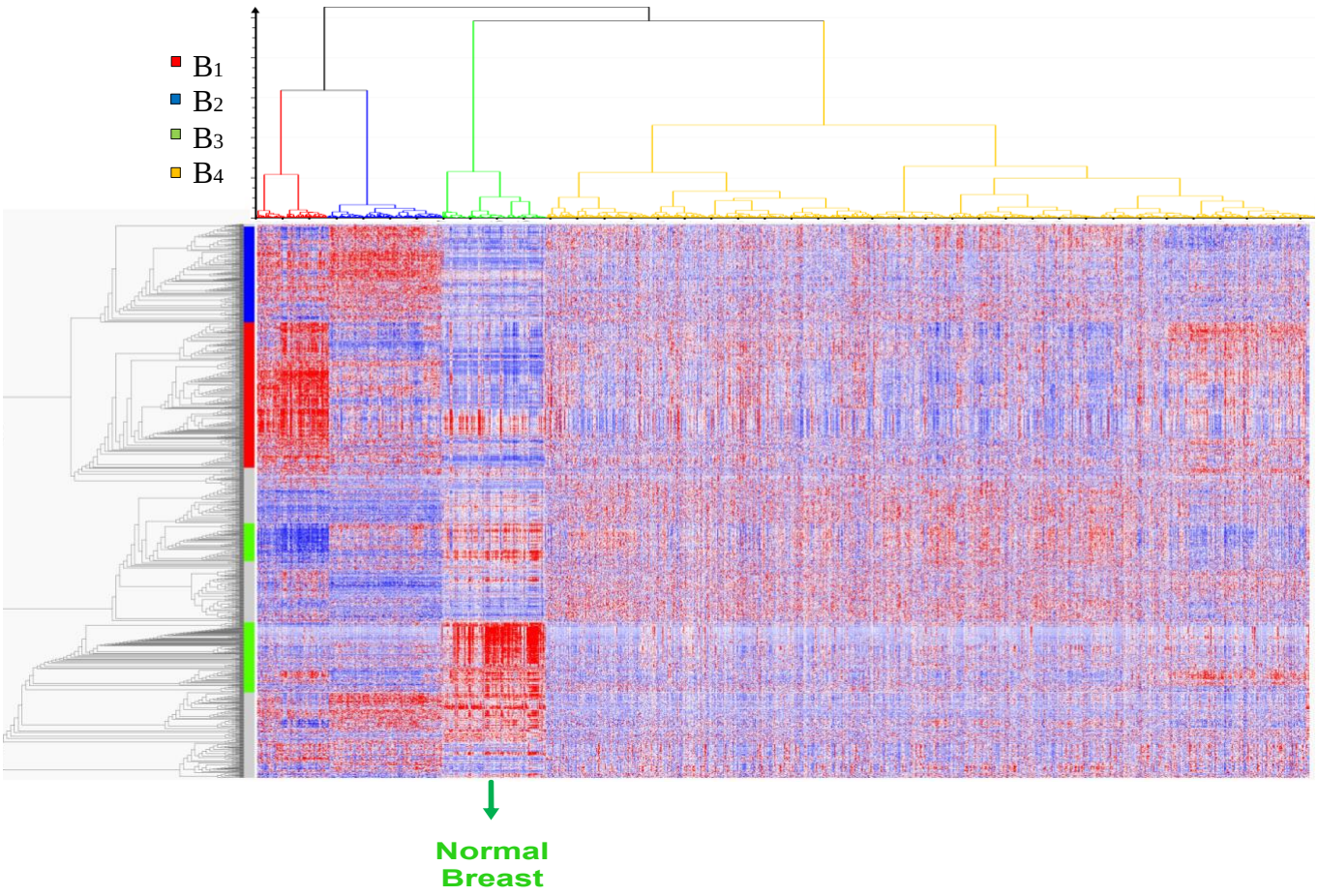


Figure 4.3: Metabolic subtypes of breast cancer.

Heatmap derived using Qlucore software, showing metabolic subtypes identified with principal component analysis (PCA) using Simca software. X-axis represents BRCA patients (n= 1,192) at the top and y-axis represents metabolic genes (n= 1,375). Dendrogram symbolizes four groups; B1 in red, B2 in blue, B3 (normal breast) in green, and B4 in yellow. B1 (n= 81) and B2 (n= 129) mostly comprised basal-like subtypes, and B4 (n= 865) contains higher proportion of luminal A subtypes.

Overall survival and five years progression-free survival of samples belonging to each subtype are not statistically significant (OS; Log-rank (Mantel-cox) test p=0.4206 and PFS; p= 0.2453) [Figure4.4and4.5]

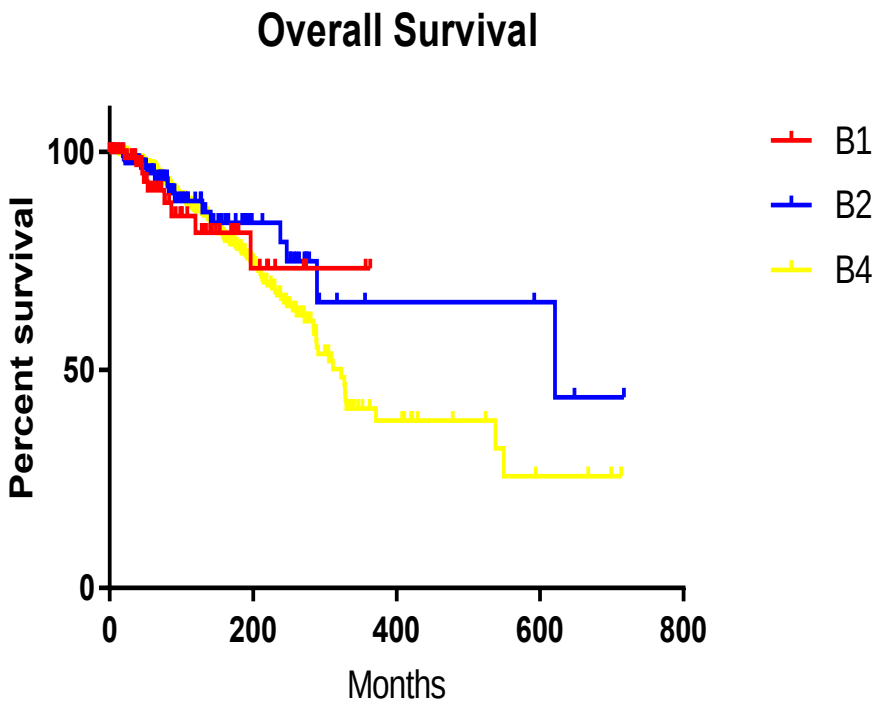


Figure 4.4: Overall survival analysis of breast cancer metabolic subtypes.

Kaplan-Meier plot generated using Prism version 8.0.1. The X-axis represents survival time in months, and the y-axis represents the percentage (%) of patients in each cohort surviving. B1 subtype (n= 81) has the worst prognosis (red curve), B2 subtype (n= 129) has the best prognosis (blue curve), and B4 subtype (n=865) has the second worst subtype (yellow curve).

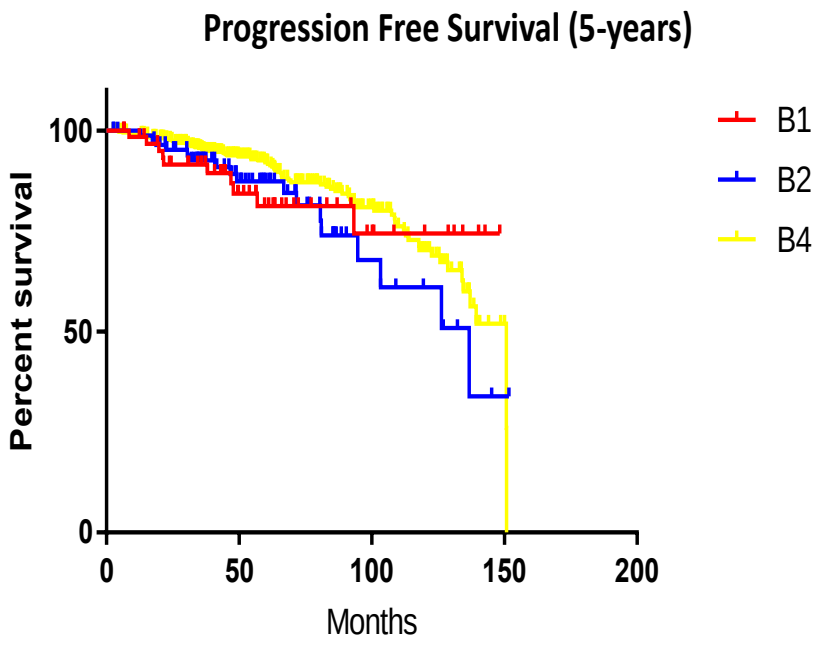


Figure 4.5: Progression-free survival analysis of breast cancer metabolic subtypes.

Kaplan-Meier plot generated using Prism version 8.0.1. The X-axis represents survival in months, and the y-axis represents survival in percentage (%). B1 subtype (n= 81) has the worst prognosis (red curve), B2 subtype (n= 129) has the best prognosis (blue curve), and B4 subtype (n=865) has the second worst subtype (yellow curve).

Identified metabolic subtypes were statistically significant with different clinical features; subtypes vs age (p=.002), subtypes vs race (<.001), subtypes vs T stage (.009), subtypes vs N stage (.007) [Table 3.1]. However, there was no statistical significance observed with the identified subtypes and the metastatic stage (p= .879) [Table 3.1]. Furthermore, our metabolic subtypes were highly associated with the previously identified molecular subtypes of BRCA (PAM50).

Metabolic subtype B1 and B2 show higher relevance with the basal-like subtype where metabolic subtypes B4 shows higher enrichment with the LumA subtypes (Chi-square test, p value < 0.001).

Table 4.1: Clinical features of BRCA metabolic subtypes

Characteristic	B_1	B_2	B_4	P-value*
Mean Age (SD)	58.99± 14.02	55.68± 12.04	59.68± 13.17	**.002**
Race				**>.001**
- White	40(50.0%)	78(64.5%)	619(79.9%)	
- Asian	1 (1.3%)	7 (5.8%)	53 (6.8%)	
- Black	39 (48.8%)	36(29.8%)	103(13.3%)	
T Stage				**.009**
- I	16(19.8%)	24(18.8%)	237(27.6%)	
- II	49(60.5%)	92(71.9%)	477(55.5%)	
- III	14(17.3%)	7(5.5%)	113(13.2%)	
- Iv	2(2.5%)	5(3.9%)	32(3.7%)	
N Stage				**.007**
- 0	44(55.7%)	80(62.0%)	381(45.2%)	
- I	20(25.3%)	35(27.1%)	298(35.3%)	
- II	7(8.9%)	10(7.8%)	101(12.0%)	
- III	8(10.1%)	4(3.1%)	63(7.5%)	
M Stage				**.879**
- 0	50(98.0%)	115(98.3%)	724(97.6%)	
- I	1(2.0%)	2(1.7%)	18(2.4%)	

Oncogenic mutations have been found to be associated with altered metabolic activity in different types of cancer, including breast cancer. However, our analysis shows no significant mutations when comparing breast cancer metabolic subtypes B1, B2 and B4 to normal breast subtype B3 after correcting for multiple comparisons, except subtype B1 (12 events) and B2 (7 events) were different with regards to gene PI3KCA (p=0.0027, p-value adjusted= 0.02) [Figure 4.6 & 4.7].

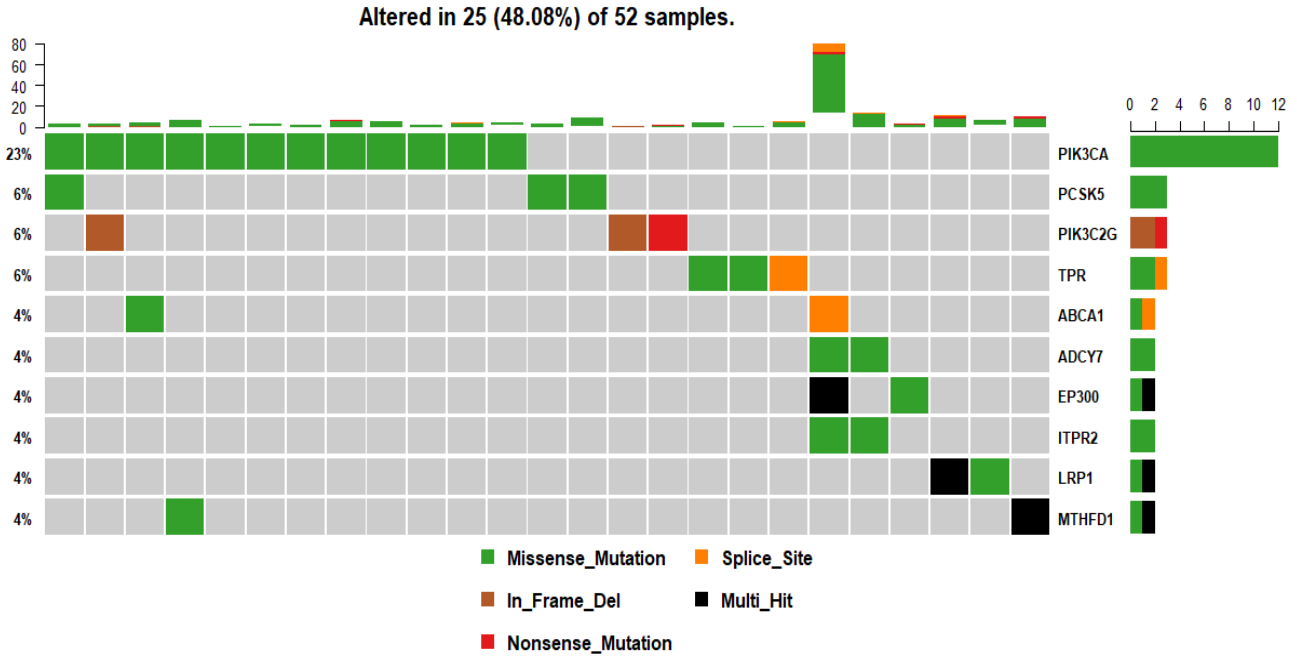


Figure 4.6: Mutational analysis of breast cancer metabolic subtype 1 (B1).

Oncoplot illustrating top ten mutated genes sorted and ordered by decreasing frequency in B1 subtype. Different mutations are represented by different colours in the x-axis including green as missense mutation, orange as splice site, brown as frame deletion, black as multiple hit and red as nonsense mutation. Y-axis represents the percentage of a specific mutation of the selected genes. Analysis and visualisation of the mutations was conducted using the Maftools software. PIK3CA found to be the most frequently mutated gene in this subgroup.

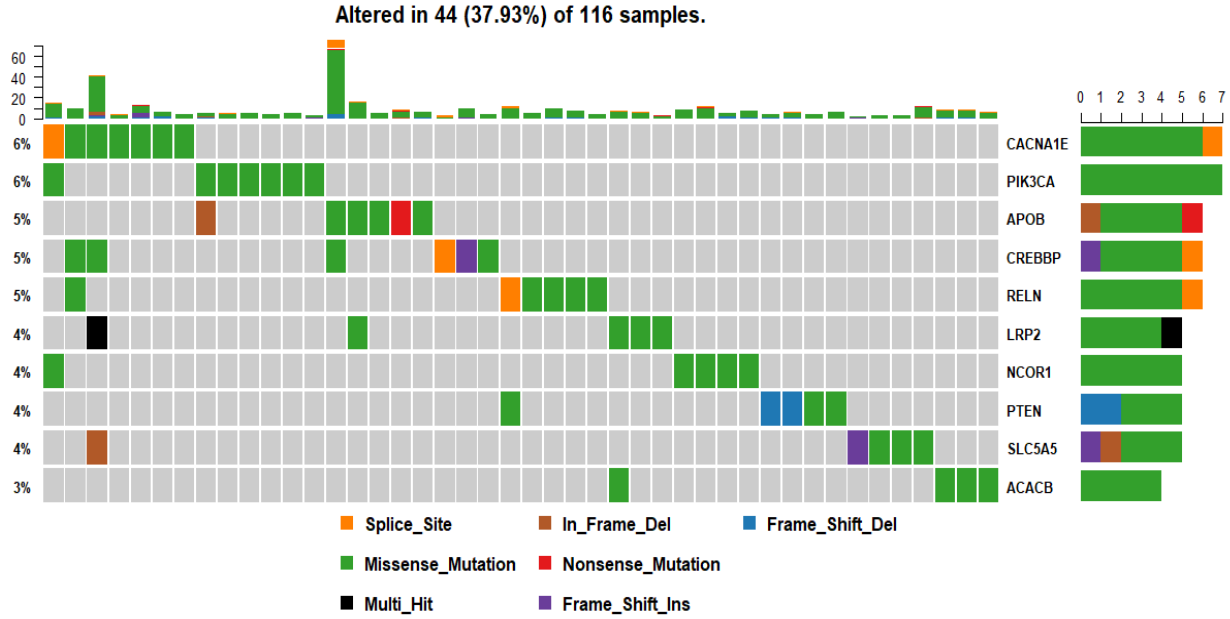


Figure 4.7: Mutational analysis of breast cancer metabolic subtype 2 (B2).

Oncoplot illustrating top ten mutated genes sorted and ordered by decreasing frequency in B2 subtype. Different mutations are represented by different colours in the x-axis including green as missense mutation, orange as splice site, brown as frame deletion, black as multiple hit and red as nonsense mutation. Y-axis represents the percentage of a specific mutation of the selected genes. Analysis and visualisation of the mutations was conducted using the Maftools software. CACNA1E and PIK3CA found to be the most frequently mutated genes in this subgroup.

4.3.2 Characterization of metabolic and biological features of BRCA metabolic subtypes

To understand the complex biological features behind the identified metabolic subtypes of BRCA, we performed a supervised analysis, including gene set enrichment analysis and literature-based informative pathway analysis (IPA). First, we used gene set enrichment analysis (GSEA), which involves testing curated gene sets corresponding to selected biological functions.

We used seven major metabolic pathways including carbohydrate metabolism (n = 286 genes), TCA cycle (n = 148 genes), amino acid metabolism (n = 348 genes), nucleotide metabolism (n = 90 genes), energy metabolism (n = 110 genes), vitamin & cofactor (n = 168 genes) and lipid metabolism (n = 766 genes). Likewise, we performed the hallmark function analysis using seven primary cancer hallmark functions including invasion & metastasis (n = 472 genes), angiogenesis (n = 36 genes), apoptosis (n = 161 genes), G2M checkpoint (n = 200 genes), DNA repair (n = 150 genes), inflammatory responses (n = 200 genes), and epithelial-mesenchymal transition (n = 200 genes). There were only a few overlapping genes between the gene sets which were primarily comprised of unique genes. All the genes represented primary metabolic and hallmark functions. Pathways were considered significantly enriched based on the FDR < 0.25.

GSEA demonstrated that breast cancer subtype B1 is positively enriched in energy metabolism (ES -0.39; FDR 0.154), lipid (ES -0.31; FDR 0.19), amino acid (ES 0.36; FDR 0.144), nucleotide (ES 0.44; FDR 0.097) and TCA cycle (ES 0.52; FDR 0.076). Similarly, GSEA hallmark analysis displays a higher number of genes that were positively enriched in G2M checkpoint (ES 0.5; FDR 0.093) and DNA repair (ES 0.52; FDR 0.076). Metabolic subtype B1 exhibits the most metabolic perturbations compared to other subtypes (B2, and B4). Most of the genes in B2 metabolic subtype were positively enriched in nucleotide (ES 0.44; FDR 0.148), G2M checkpoint (ES 0.75; FDR 0.007), and DNA repair (ES 0.45; FDR 0.062), but negatively enriched in energy metabolism (ES -0.41; FDR 0.155). In the B4 subtype, nucleotide metabolism (ES 0.44; FDR 0.054), G2M checkpoint (ES 0.59; FDR 0.06), and energy metabolism (ES -0.35; FDR 0.251) were found to be highly perturbed. Metabolic and seven hallmark analysis illustrated in figure 4.8.

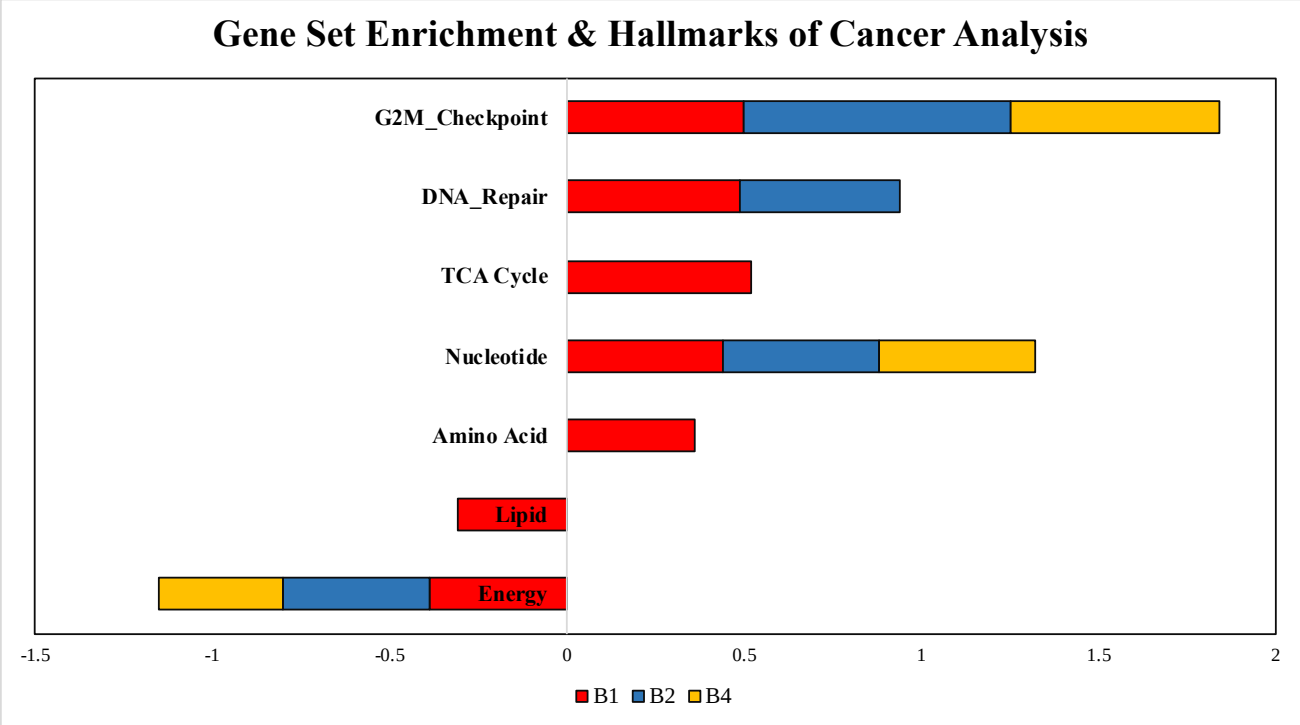


Figure 4.8: Gene set enrichment analysis of identified metabolic subtypes (B1, B2 and B4).

As described in the methods section this analysis was conducted using major metabolic pathways and seven hallmarks of cancer. The X-axis represents enrichment score (ES), and Y-axis embodies major pathways. Red, blue and yellow represent subtype B1, B2 and B4, respectively. Enrichment of nucleotide, energy metabolism and G2M checkpoint observed across all metabolic subtypes.

Further, we identified enriched functions within the identified metabolic subtypes utilizing 50-hallmark functions from the broad institute. 50 hallmark function analysis is a representation of specific biological processes and display of coherent expression. Oxidative phosphorylation (p-value 0.027), DNA repair (p-value 0.010), MYC targets (p-value0.024) and TGF beta signalling (p-value 0.006) found to be significantly enriched in metabolic subtype B1. Likewise, we observed metabolic subtype B2 has enriched functions including E2F targets (p-

value < 0.001), G2M checkpoint (p-value 0.008), mitotic spindle (p-value 0.010) and estrogen response (p-value 0.008). The common function that observed across all identified metabolic subtype is UV response. All of the analyses summarized in figure 4.9. Further, we did a literature-based informative pathway analysis to determine the signalling pathways that may cause the above alterations.

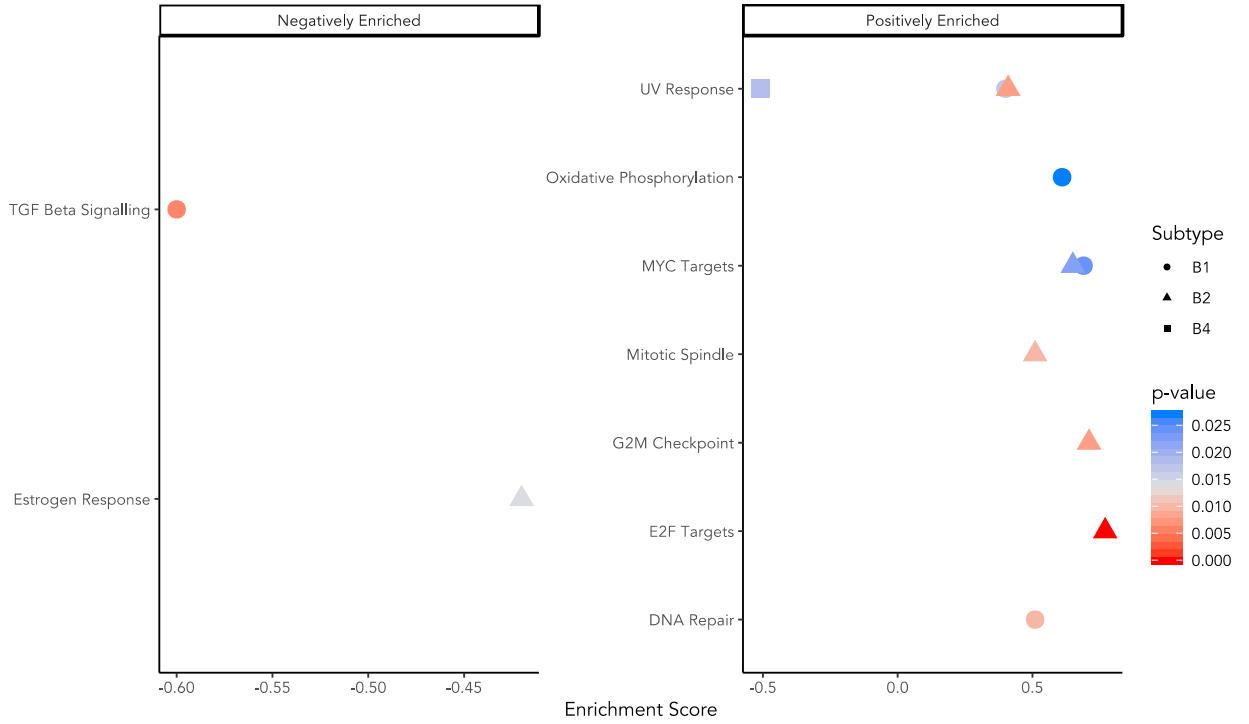


Figure 4.9: 50 hallmark functions analysis of BRCA subtypes (n = 1,192 samples and n = 14,735 genes)

The x-axis represents enrichment score, and the y-axis signifies the altered functions. The legend's different colour code shows the significance level of enrichment. Each metabolic subtype was compared with the healthy breast tissue group. Most of the functions found altered in metabolic subtypes B1 (circle shape) and B2 (triangle shape). Alterations in UV response is the shared feature found across all metabolic subtypes.

Further, we investigated perturbed canonical pathways in each subtype that may provide a better explanation of different metabolic and cellular perturbations. In B1 subtypes, most of the genes are poorly expressed, resulting in the down-regulation of the PTEN signalling (B-H p-value = 0.039) pathway. Further analysis reveals highly activated protein synthesis (B-H p-value = 2.39E-06) and RNA post-transcriptional modification (B-H p-value = 5.05E-06) [Figure 4.10].

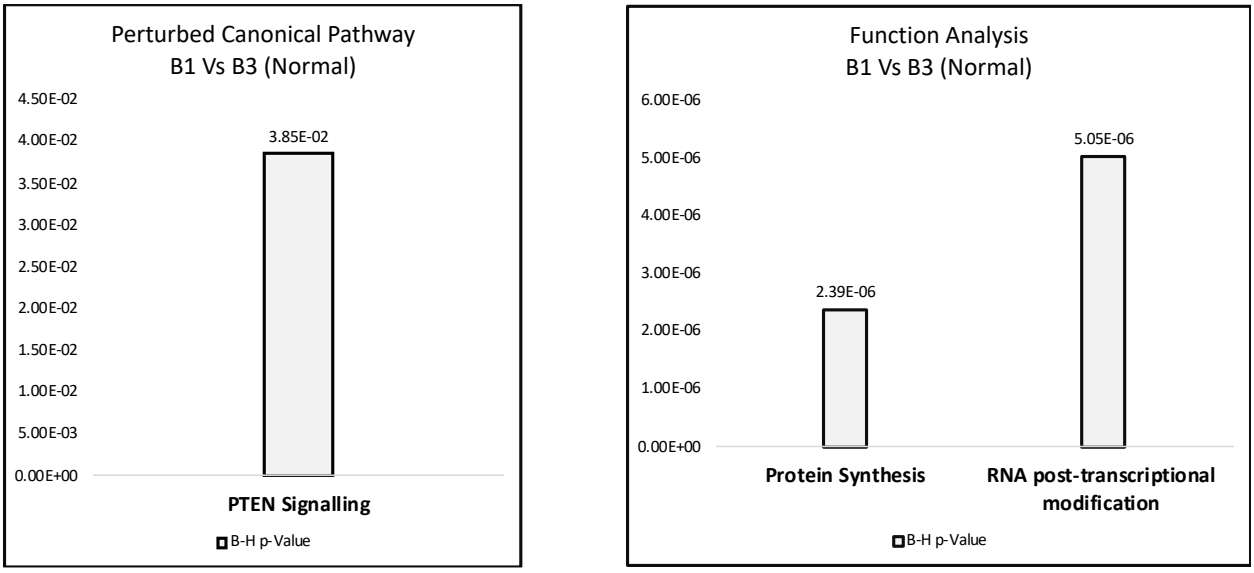


Figure 4.10: Perturbed canonical pathways and functions in B1.

The analysis was performed using filtered gene set and compared B1 subtype (n = 81 samples) to the healthy breast tissue (B3) (n = 117 samples). The X-axis in the bar chart denotes pathways and functions, and Y-axis is the p-value. Benjamini-Hochberg test (B-H) was performed to correct the multiple comparisons. This analysis shows that substantial downregulation of PTEN signalling is highly associated with the disrupted protein synthesis and RNA post-transcriptional modification in the B1 breast cancer metabolic subtype.

In the B2 subtype, highly activated mitotic roles of polo-like kinase signalling pathway (B-H p-value 0.01) were found. Moreover, altered polo-like kinase signalling was found to affect some major functions within the subtype. For example, cell cycle (B-H p-value 1.85E-06) and DNA replication (B-H p-value 6.83E-06) processes were identified to be significantly increased in B2 along with decreased cell death (B-H p-value 2.42E-07) [Figure 4.11].

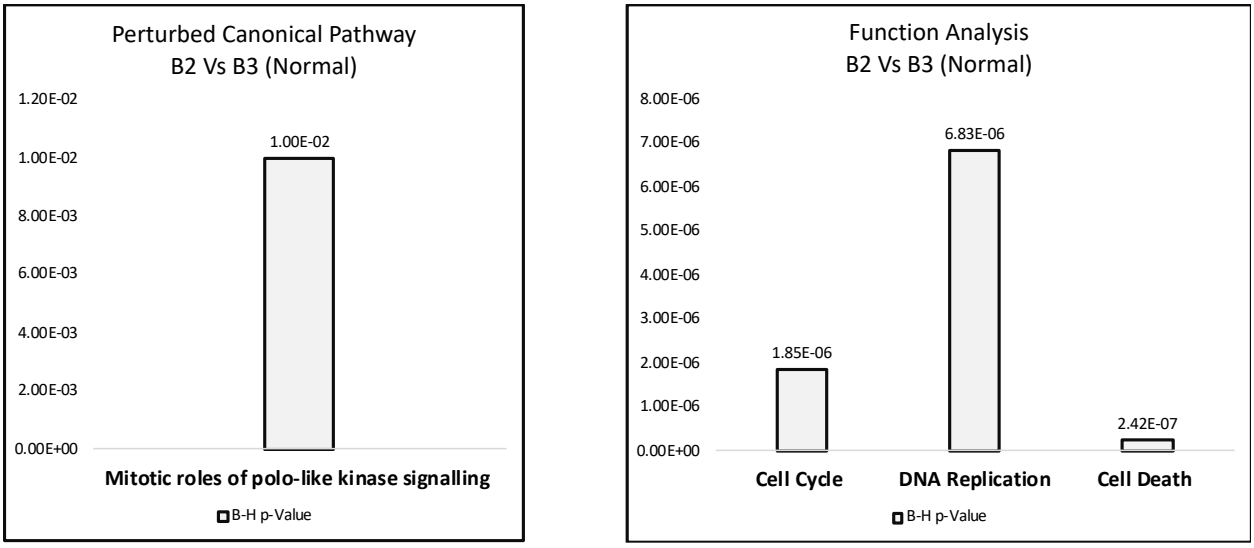


Figure 4.11: Perturbed canonical pathways and functions in B2.

The analysis was performed using a filtered gene set and compared B2 subtype (n = 129 samples) to the healthy breast tissue (B3) (n = 117 samples). The X-axis in the bar chart denotes pathways and functions, and Y-axis is the p-value. Benjamini-Hochberg test (B-H) was performed to correct the multiple comparisons. This analysis shows that significantly altered polo-like kinase signalling affects the cell cycle, DNA replication and cell survival in B2 breast cancer metabolic subtype.

Finally, most of the canonical pathways and primary functions were found to be significantly downregulated in the B4 subtype compared to the normal breast. Perturbed

canonical pathways included Integrin signalling (B-H p-value 0.01), P-21 activated kinases (PAK) signalling pathway (B-H p-value 0.02), and Fibroblast growth factor signalling (FGF) signalling pathway (B-H p-value 0.05). All the down-regulated signalling was associated with decreased cell to cell interaction (B-H p-value 2.97E-05), cellular growth and proliferation (B-H p-value 2.26E-06) [Figure 4.12].

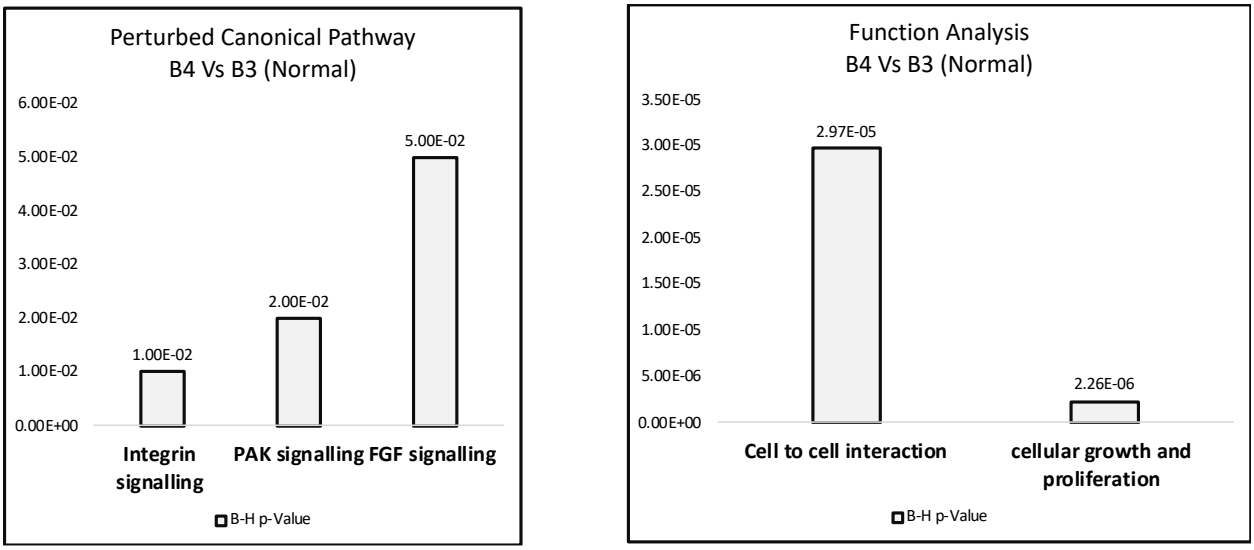


Figure 4.12: Perturbed canonical pathways and functions in B4.

The analysis was performed using filtered gene set and compared subtype B4 (n = 865 samples) to the healthy breast tissue (B3) (n = 117 samples). The X-axis in the bar chart denotes pathways and functions, and Y-axis is the p-value. Benjamini-Hochberg test (B-H) was performed to correct the multiple comparisons. This analysis shows that there is increased integrin signalling, PAK signalling, FGF signalling highly affects cell to cell interaction and cellular growth and proliferation in the B4 breast cancer metabolic subtype.

4.4 Discussion

One of the critical hallmarks of cancer is the reprogramming of the energy system. A well-established cancer cell metabolic abnormality is the Warburg effect which is the elevation of glycolysis in the presence of oxygen [24]. Over the years, scientists have discovered that besides the Warburg effect, other crucial metabolic alterations help tumour cells to proliferate and survive. However, metabolism is a complex process that is not fully characterized. It is important for the field of cancer research that the metabolic alterations of cancer cells are better understood and defined. Based on our TCGA breast cancer patient cohort's transcriptomic data profiling, we found that differential metabolic gene expression predominantly reflects crucial metabolic activities within the cancer cells and biological functions within the patients.

One of the main aims of this study was to identify meaningful metabolic subtypes. Thus, we focused on metabolic gene expressions to classify breast cancer patients and explored different biochemical and biological functions associated with these identified subtypes, combining gene set enrichment analysis and informative pathway analysis. One of the advantages of this systemic approach is that it gives a better understanding of the functional state of different activities (both metabolic and biological functions), showing whether they are up-regulated, down-regulated, and/or affected/neutral. Throughout our approach, we found that metabolic gene expression subtypes correlated with each other. For instance, nucleotide and energy metabolic perturbations were observed more frequently in all the identified subtypes

compared to the normal breast tissue group. Focusing on identified subtypes, we also showed that it is possible to investigate different oncogenic mutations and copy number variations that might be associated with the alterations affecting metabolism.

In this study, we demonstrated potential clinical features that significantly correspond with identified BRCA subtypes. The identified metabolic subtypes highly correspond with the age, race, T stage and N stage of the tumour. Furthermore, we observed that identified metabolic subtypes show distinct patterns. For example, increased nucleotide metabolism, particularly purine metabolism, was found to be highly associated with the bad prognosis and decreased pyrimidine and carbohydrate metabolism was found to be mostly associated with better prognosis. Likewise, we demonstrated that the underlying biological consequences are associated with the altered metabolic activity across the metabolic subtypes. Our result suggests that both positively and negatively enriched nucleotide metabolism highly affects the G2/M checkpoint, thus affecting the cell cycle. This indicates that metabolic status has the potential to help inform treatment selection.

There are well-established studies that show dysregulated metabolic functions helps cancer cell to proliferate abnormally and survive, which suggest that inhibiting the altered metabolic functions may pose an advantage in preventing abnormal cancer cell proliferation. Recently, researchers have made efforts to understand cancer cell metabolism and its effectiveness in terms of therapy. While there has been limited success looking at this aspect up until this point, our systemic based approach presents a promising standard therapeutic strategy. For instance, in bad prognosis subtypes, up-regulated functions may indicate susceptibility to

therapy that targets their regulatory factors. Identifying the key regulatory factors in specific subtypes may reverse the functional state of a particular activity, changing up-regulated subtypes to down-regulated subtypes. The impact that this systemic approach could have on treatment decisions could provide a potential survival benefit and improvement in health condition to patients

It is clear that the discovery of metabolic subtypes in breast cancer could have major implications in cancer research. Even though there are no measurable changes in survival between metabolic subtypes, the existence of metabolic subtypes may provide new therapeutic targets. That is, it may be possible to affect cell viability by targeting the specific metabolic features of cancer. The existence of these metabolic subtypes and their relationship to clinical outcomes will have to be validated in other large datasets.

Chapter Five: Different Metabolic Subtypes of Pancreatic Ductal Adenocarcinoma

5.1 Introduction

Pancreatic ductal adenocarcinoma (PDAC) is highly lethal and it is the fourth leading cause of cancer death [159]. The five-year overall survival rate is less than 7% [160,161]. Only about 20% of patients are eligible for surgery [162]. Even after resection, the disease recurs in up to 80% [163,164]. The aggressive tumour biology as well as the involvement of non-cancerous cells in PDAC make it challenging to treat, and it is resistant to all currently available therapies such as chemotherapy and radiotherapy [165-167]. There is a need for novel therapeutic targets in PDAC. Furthermore, the biology of the tumour and stromal cells needs to be better understood to design more effective subtype-specific treatment.

Changes in oncogenes and tumour suppressors highly affect gene expression levels, which subsequently cause metabolic reprogramming in cells [168-172]. Various groups have made efforts to devise approaches to subtyping cancers according to their metabolic features [8,173-175]. Large and multidimensional molecular datasets like those developed by The Cancer Genome Atlas and the International Cancer Genomics Compendium have made this work possible. These studies focused on metabolic subtyping and show a broad transcriptional disruption of metabolic genes. However, the main challenge is how to categorize cancer patients more effectively into different subtypes based on the metabolic gene expressions as well as their utilization in clinical practice. To address those challenges, we developed a systematic

transcriptome-based bioinformatics approach to identify and characterize different metabolic subtypes in PDAC patients.

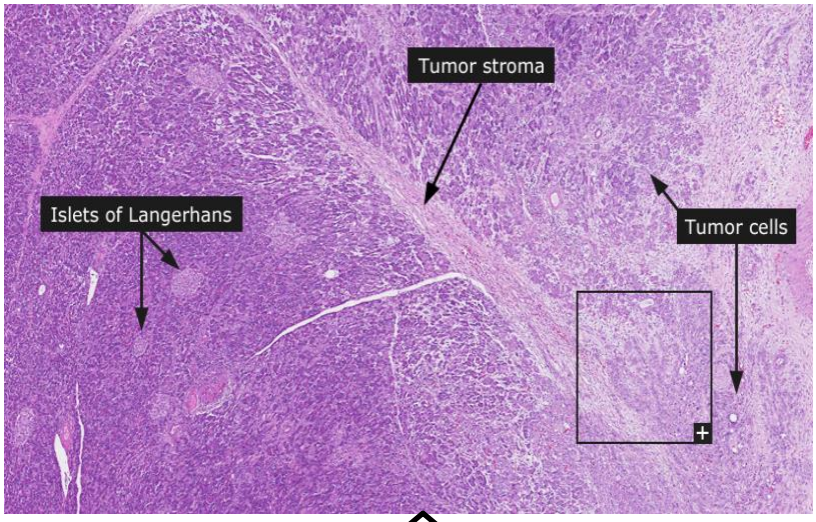
5.2 Methods

5.2.1 Overview

We downloaded normalized RSEM RNA-Seq patient (n=137) data from the TCGA portal. The dataset contained 13,218 genes, including 1240 metabolic genes. Metabolic genes identification, dataset normalization, was performed in the same way as described in chapter four.

5.2.2. Deconvolution of the original PDAC RNA-Seq dataset

Dr. Wenyi Wang's team at the MD Anderson Cancer Centre performed the deconvolution of the dataset. They developed a two-component deconvolution method termed as DeMixT. This algorithm separates the tumour cells and stroma cells from bulk tumour tissue. DeMixT reconstitutes the expression profile for all the components for each gene and each sample as well as estimates all distribution parameters and cellular properties.



5.2.3 Metabolic Subgrouping

We applied the non-negative matrix factorization (NMF) method to our metabolic gene set to identify the metabolic subtypes. NMF method factorizes a data matrix (V) into two matrices (W and H), and no negative value exists in all three matrices .

$V \approx W \times H$

Based on this approach, a four-cluster model was identified as the best fit for the given dataset. We then performed the hierarchical analysis using the Manhattan distance metric and complete-linkage clustering to visualize the metabolic subtypes. The complete linkage covers the longest distance between two points in each cluster and Manhattan distance measures distances between two different points as a grid-like path as follows –

$d1 \equiv dSAD:(x,y) \mapsto \|x-y\|1 = \sum i=1n |xi-yi|$

5.2.4 Survival Analysis

For the overall survival analysis, we generated a Kaplan-Meier plot using PRISM version 8.0.1. For all the downstream analyses, each metabolic subtype was compared to the best prognosis group (M3).

5.2.5 Clinical Data Analysis

All the additional clinical data including age, race, pathologic tumour stage, histological grade was downloaded from the cBioportal. We performed descriptive statistics comparing identified metabolic subtypes and age and used a Pearson chi-square test to determine the significance between metabolic subtypes and other clinical characteristics. All of the analyses were performed using IBM SPSS version 24.

5.2.6 Mutation and Copy Number Variations Analysis

All mutational and CNV data of the TCGA patients were derived from the cBioPortal. We performed Kruskal Wallis Test to determine the significance of different mutations and copy number alterations in identified metabolic subtypes.

5.2.7 Gene Set Enrichment Analysis and Ingenuity pathway analysis

Gene set enrichment analysis, and pathway analysis was performed the same as described in chapter four.

5.3 Results

5.3.1 Metabolic gene expressions identify four distinct subtypes of PDAC

We have established four distinct metabolic subtypes (M1, M2, M3 and M4) of PDAC based on patterns of 1,240 metabolic gene expression using the same method described in chapter two to identify genes. The study was conducted using 137 PDAC TCGA patients. We used a non-negative matrix factorization (NMF) and hierarchical clustering (HCL) method to identify the distinct subtypes. M3 subtype was = the largest group (n = 43, 0.31%), followed by the M2 subtype (n = 42, 0.30%), M1 Subtype (n = 29, 0.21%) and M4 subtype (n = 23, 0.16%) [Figure 5.1].

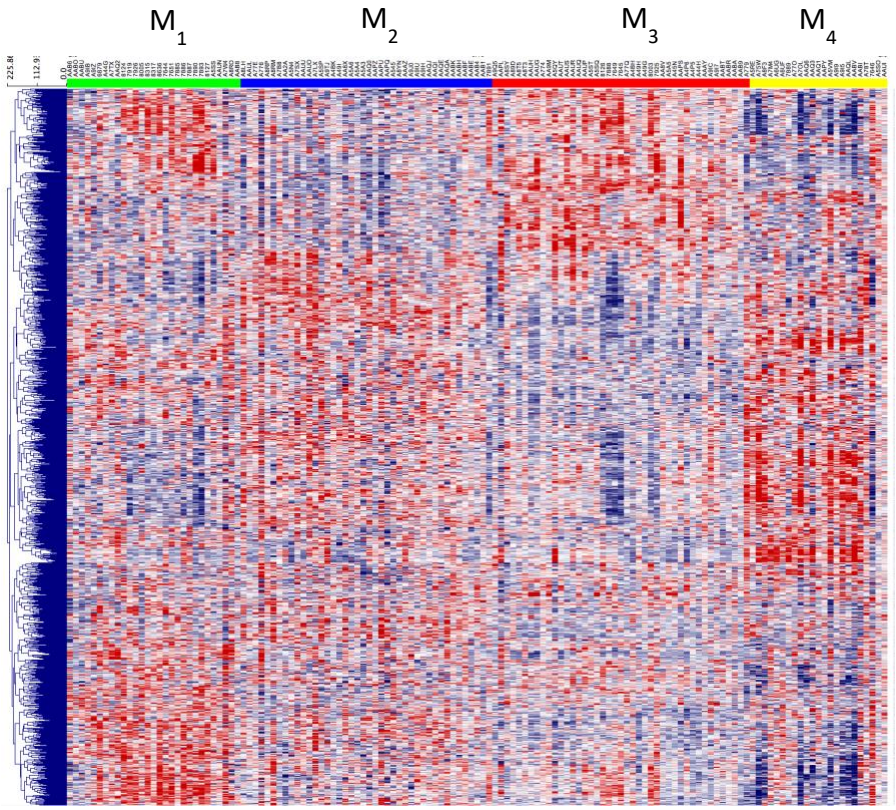


Figure 5.1: Metabolic subtypes of pancreatic ductal adenocarcinoma (PDAC).

Four distinct metabolic subtypes identified based on the non-negative matrix factorization (NMF) method. The heatmap created using Manhattan distance and complete linkage clustering algorithm. Colour bars at the top (x-axis) represent the groups of patients: M1 is red, M2 is blue, M3 is red, and M4 is yellow. Y-axis symbolises metabolic genes (n= 1240).

The patient outcomes for each metabolic subtype have significant clinical relevance (p-value < 0.001) [Figure 5.2]. The Kaplan-Meier plot shows M1 is the worst prognosis group with a median survival of 10.7 months, followed by the intermediate group, M4 with an average survival

of 15.7 months. Subtype M3 and M2 are the best prognosis groups, with an average survival of 24.6 and 23.4 months, respectively.

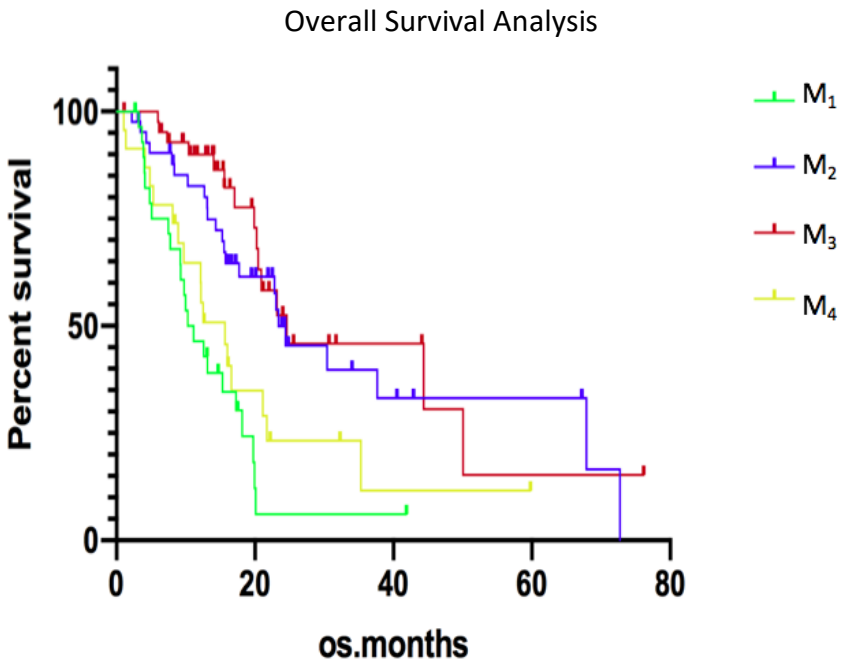


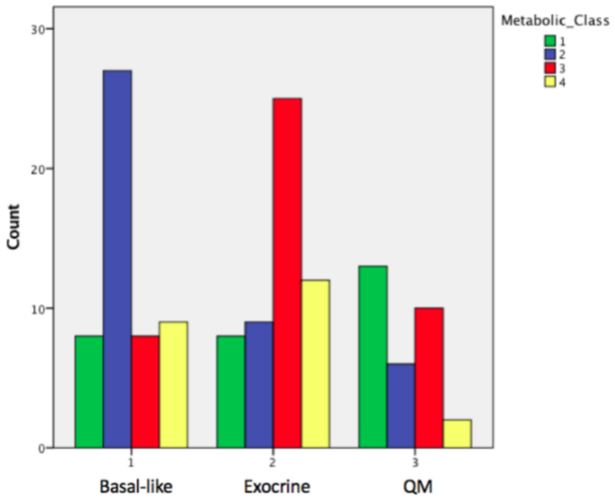
Figure 5.2: Overall survival analysis of identified metabolic subtypes (n = 137 patients).

Kaplan-Meier plot generated for the overall survival analysis. The X-axis symbolizes the overall survival of patients in months, and the y-axis indicates survival in percentage (%). The different colours indicate different subtypes: M1 is green, M2 is blue, M3 is red and M4 is yellow. Overall survival is statistically significant across identified metabolic subtypes (p-value < 0.001). Based on the analysis, M1 is the worst, followed by M4, and M3 is the best, followed by M2.

5.3.2 Association of metabolic subtypes with known PDAC molecular and immune subtypes

In 2011, Collisson et al. identified three molecular subtypes in PDAC based on 62 gene signatures (Basal-like, exocrine and quasi-mesenchymal). In 2015, Moffit et al. identified two tumour subtypes basal, classical and two stroma subtypes as normal and activated. In both studies, the authors found the basal-like subtype is associated with poor survival outcomes. To explore whether the identified metabolic expression patterns across the identified subtypes could underline the differences between previously identified molecular subtypes, we determined the previously identified PDAC subtypes for each sample and investigated their degree of overlap with the metabolic phenotypes. Statistical analysis shows significant correlation between identified metabolic subtypes and Collisson subtypes (p-value < 0.001); however, we found no significant correlation found the Moffit subtypes (p-value = 0.158). From the enrichment analysis, we found that the basal-like subtype highly corresponds with the M2 subtype (64.3%), exocrine subtype with M3 (58.1%), and quasi-mesenchymal subtype with the worst metabolic subtype, M1 (44.8%) [Figure 5.3].

Metabolic Subtypes Vs. Collisson Subtypes



Metabolic Subtypes Vs. Moffit Subtypes

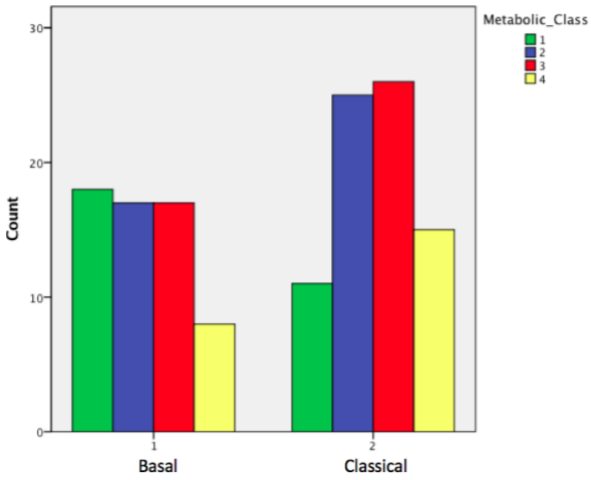


Figure 5.3: Enrichment analysis between metabolic subtypes and previously identified molecular subtypes.

Based on the chi-square test, identified metabolic subtypes highly correspond to the Collisson subtypes: basal-like, exocrine and quasi-mesenchyme (p-value < 0.001). No statistical significance was observed with Moffit subtypes: basal-like and classical (p-value 0.158). The different colours indicate our identified subtypes: M1 is green, M2 is blue, M3 is red and M4 is yellow.

A recent paper published in Immunity on almost 10,000 cases of 33 cancer types described six patterns of immune response (C1, C2, C3, C4, C5 and C6) [108].C1 had higher expression of angiogenic genes with increased cell proliferation, C2 had the highest M1/M2 macrophage polarization and the most T cell receptor diversity, C3 had elevated T helper 17 and 1 genes and low to moderate cell proliferation, C4 showed Th1 suppressed and high M2 response,

C5 had the lowest lymphocyte and highest macrophage responses, and finally, C6 showed the highest TGF-β signatures. To investigate whether any significance exists between the identified metabolic subtypes and immune subtypes, we performed an enrichment test. The analysis shows only four immune subtypes (C1, C2, C3 and C6) out of six correspond highly with the identified metabolic subtypes. We found that the worst subtype M1 has a high correlation with C1 (65.5%), while the best subtype M3 corresponds with the C3 (46.5%). Interestingly, the immune subtype C3 was the only immune group that had low to moderate tumour cell proliferation [Figure 5.4]. Furthermore, the crosstabulation test supports the classification of M1 as the worse prognosis group, as it is linked with the higher cell proliferative group (C1).

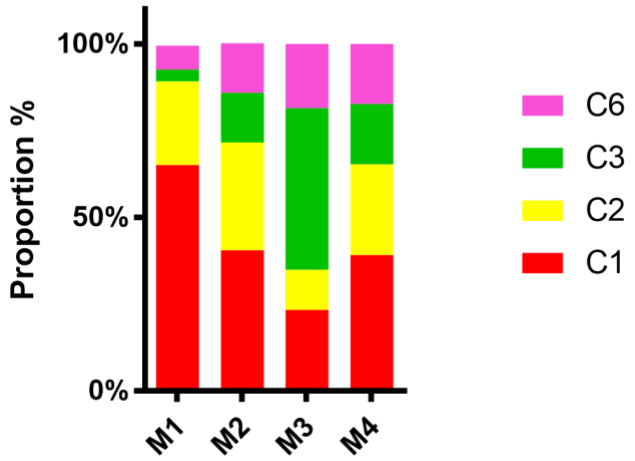


Figure 5.4: Enrichment test between metabolic subtypes and immune subtypes.

Metabolic subtypes (M1, M2, M3 and M4) are labelled on the x-axis and the proportion of the immune subtypes is labelled on the y-axis as a percentage (%). This enrichment test demonstrates that the metabolic subtypes are highly correlated with the immune subtypes (Pearson chi-square < 0.001). It shows that the worst metabolic subtype (M1) highly corresponds with immune subtype

C1 (worst immune subtype) and the best prognosis group (M3) corresponds with immune subtype C3 (good immune subtype).

5.3.3 Deconvolved tumour metabolic expressions reveals diverse functional alterations

To determine the biochemical and biological relevance of metabolic subtypes, we inspected various pathways by gene set enrichment analysis (GSEA) based on the RNA expression (p-value < 0.05 and FDR < 0.25). The analysis included seven major metabolic pathways (carbohydrate, TCA cycle, lipid, amino acid, nucleotide, energy, vitamin & cofactor) and seven functional cancer hallmarks (apoptosis, angiogenesis, epithelial-mesenchymal transition, G2M checkpoint, inflammatory response, DNA repair, invasion & metastasis). From the analysis, we found that each of the metabolic subtypes has an extensive pathway-level functional effect. The GSEA analysis reveals the worst metabolic subtype has perturbed carbohydrate metabolism (p-value 0.012; FDR 0.068), which results alterations in G2M checkpoint (p-value 0.024; FDR 0.03) and invasion & metastasis (p-value 0.024; FDR 0.157) [Figure 5.5]. Further pathway analysis confirmed that the worst M1 subtype is highly dependent on glycolysis to produce the energy supporting cell proliferation, showing elevated expression of glyceraldehyde-3-phosphate dehydrogenase (GAPDH), fructose-bisphosphate A (ALDOA), enolase 1(ENO1), phosphoglycerate kinase 1 (PGK1). From the 50 hallmarks analysis, we found that most of the critical functions are positively enriched, including glycolysis, TGF beta signalling, NOTCH signalling, MTORC1 signalling etc. across all identified metabolic subtypes. The worst metabolic subtype M1 exhibits the most perturbed functions [Figure 5.6].

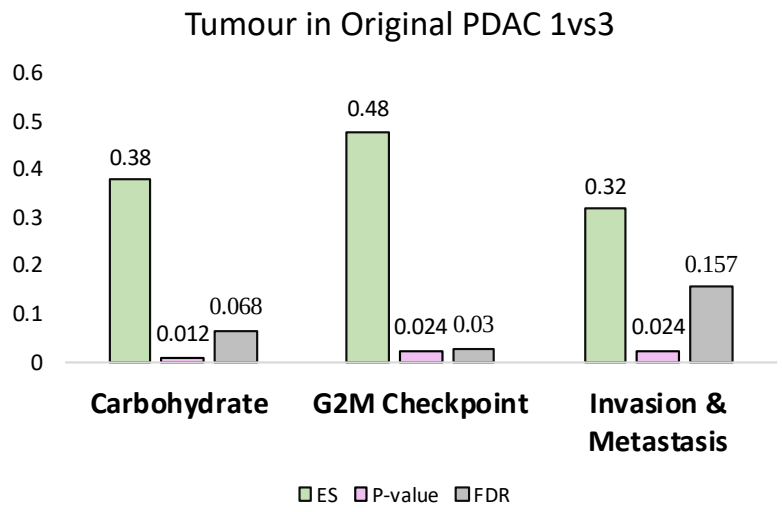


Figure 5.5: Gene set enrichment analysis of the deconvolved tumour M1.

The green bar represents enrichment score (ES), the purple colour indicates p-value, and grey colour signifies FDR value. GSEA shows that a positively enriched carbohydrate metabolic pathway is significantly associated with a positively enriched G2M checkpoint and Invasion & metastasis.

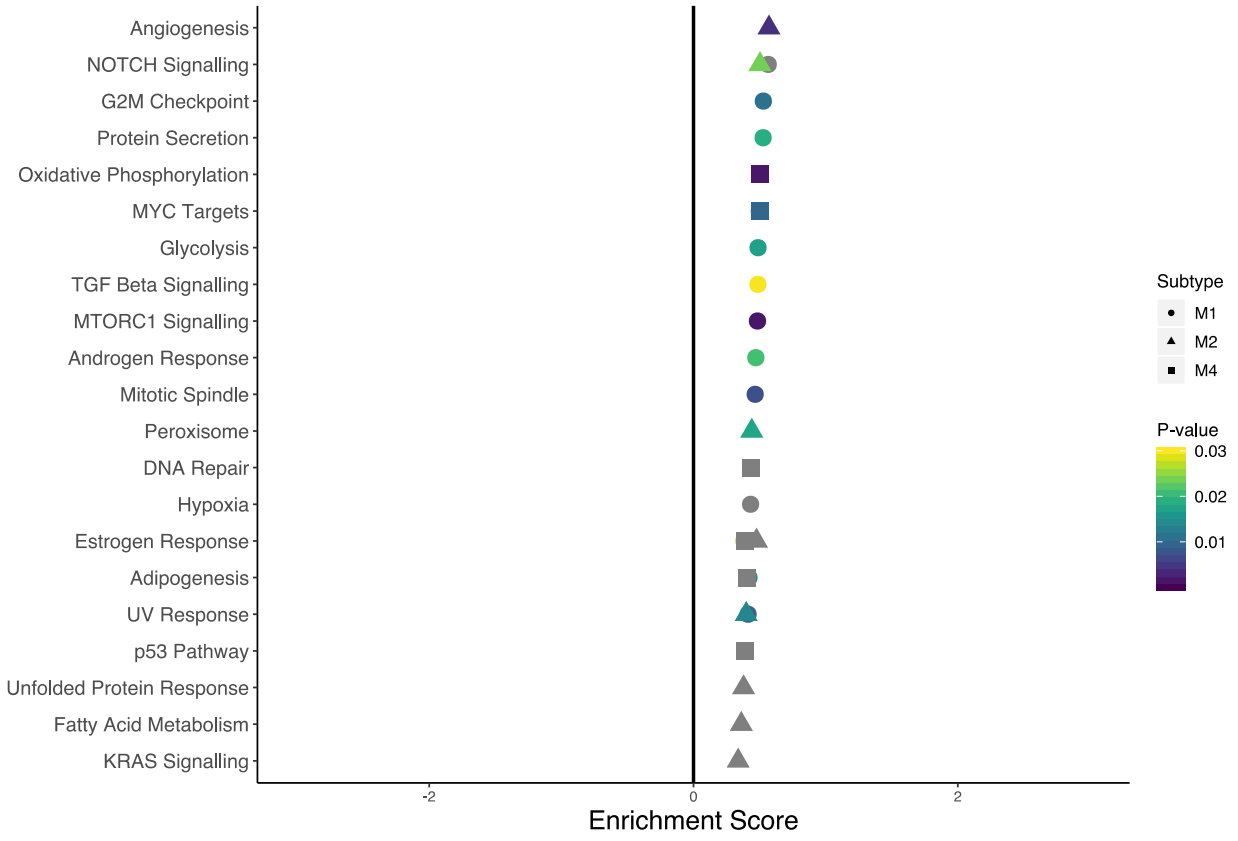


Figure 5.6: 50 hallmark functions analysis of PDAC deconvolved tumour.

The x-axis represents enrichment score, and the y-axis signifies the altered functions. Different colour code shows the significance level of enrichment. All the functions found positively enriched in the tumour compartment.

Most of the primary functions including cell survival (p-value 2.62E-07), post translational modification (p-value 2.36E-06) , protein synthesis (p-value 5.24E-05), DNA repair (p-value 7.61E-05) and cell cycle (p-value 6.58E-03) were found to be significantly up-regulated in M1. Additionally, most critical pathways, such as insulin-like growth factors receptor signalling (p-value 8.29E-07), interleukin-2 signalling (p-value 8.01E-06), and nuclear factor kappa B signalling (p-value 2.45E-04), were observed to be highly increased in M1 subtype, while p53 signalling (p-value 9.99E-06), and PTEN signalling (p-value 1.3E-03) were decreased [Figure 5.7].

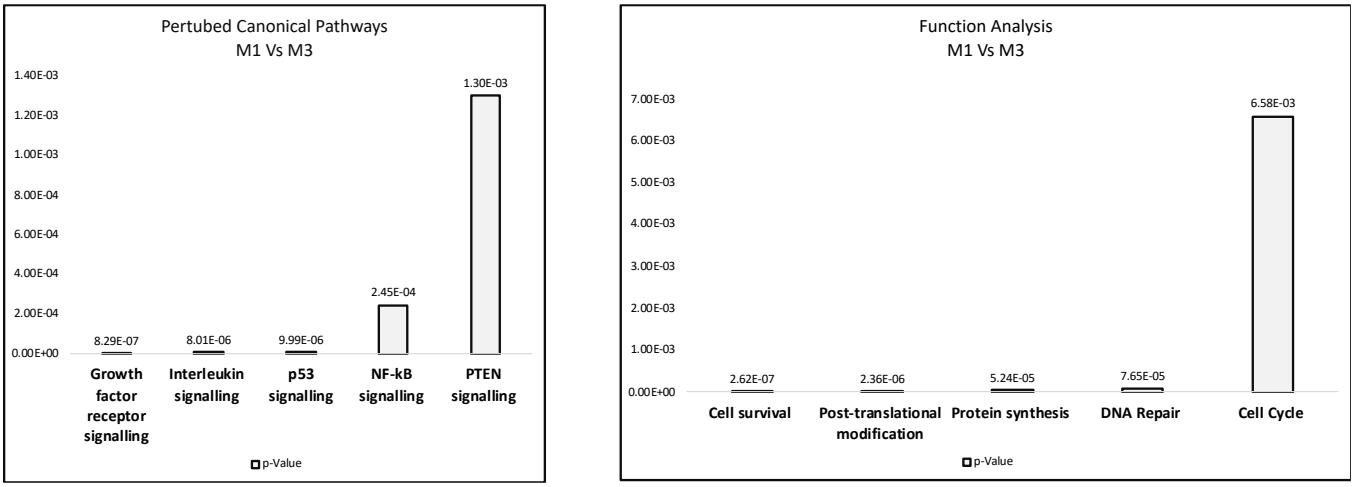


Figure 5.7: Altered canonical pathways and functional analysis of deconvolved tumour M1.

The x-axis in the bar chart denotes pathways and functions, and the y-axis represents p-value. We performed the analysis by comparing each subtype with the best prognosis group (M3). Most of the critical signalling pathways in the worst subtype were highly perturbed, resulting in alterations in major cellular and molecular functions including cell survival, protein synthesis and DNA repair.

In contrast, the intermediate group (M4) appeared to be dependent on amino acid metabolism for energy production. GSEA shows altered amino acid (p-value 0.002; FDR 0.001), and TCA cycle (p-value 0.00; FDR 0.008) effecting angiogenesis (p-value 0.004; FDR 0.031), and DNA repair process (p-value 0.025; FDR 0.067) [Figure 5.8]. Further analysis revealed up-regulation in glycine and serine metabolism, and notably increased 3-phosphate hydroxy pyruvate to serine conversion and sarcosine to glycine and glycine to 2-amino acetate conversion. Impaired mTOR complex and elevated LC3, ATG3, ATG2, ATG5, VPS34 indicates upregulated autophagy in M4.

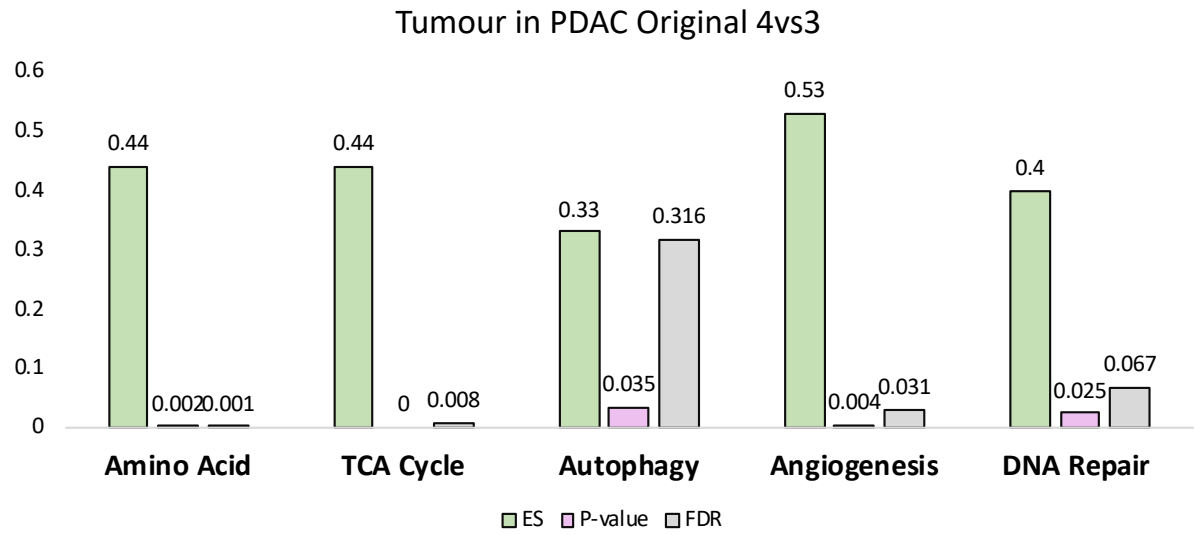
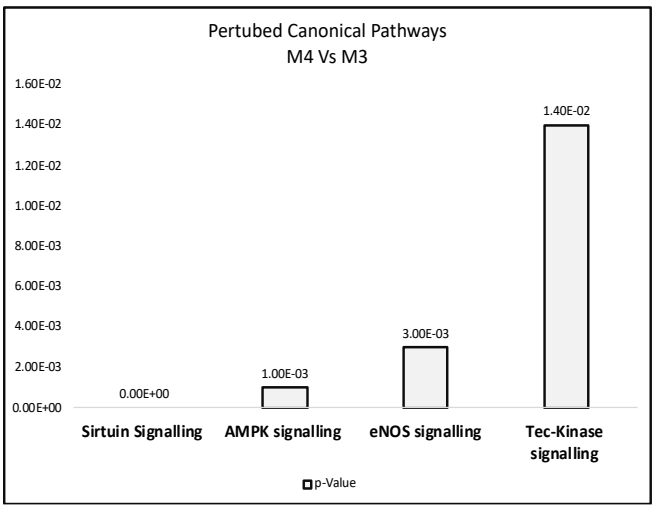


Figure 5.8: Gene set enrichment analysis of deconvolved tumour M4.

The green bar represents enrichment score (ES), purple indicates p-value and grey signifies FDR value. Enriched pathways include amino acid, TCA cycle, autophagy, angiogenesis and DNA repair.

We identified post-translational modification (p-value < 0.001), protein synthesis (p-value 0.001) and cell cycle (p-value 0.019) functions as significantly increased in subtype M4. Altered pathways in this subtype include Sirtuin signalling (p-value < 0.001), AMP-activated protein kinase signalling (p-value 0.001), endothelial nitric oxide synthase (p-value 0.003), and Tec kinase signalling (p-value 0.014) [Figure 5.9].



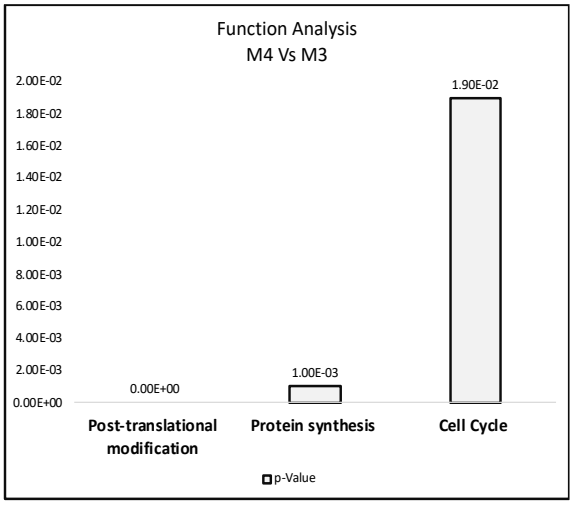


Figure 5.9: Altered canonical pathways and functional analysis of deconvolved tumour M4.

The X-axis in the bar chart denotes pathways and functions, and Y-axis represents p-value. We performed the analysis by comparing each subtype with the best prognosis group (M3). Most of the critical signalling pathways were highly perturbed in the worst subtype, resulting in alterations in major cellular and molecular functions including cell cycle and protein synthesis.

Gene set enrichment analysis revealed that the good prognosis group, M2, utilizes lipid and nucleotide metabolism to yield energy [Figure 5.10]. No primary functions were altered except decreased protein synthesis (p-value 0.001) in M2, which could be due to the decreased

PI3/Akt signalling (p-value 0.007) and increased insulin receptor (p-value < 0.001) and ERK/MAPK (p-value < 0.001) signalling pathways [Figure 5.11]. Overall, the results show how these subtypes are metabolically diverse and have a strong correlation with vital hallmarks of cancer, as well as hallmark related functions.

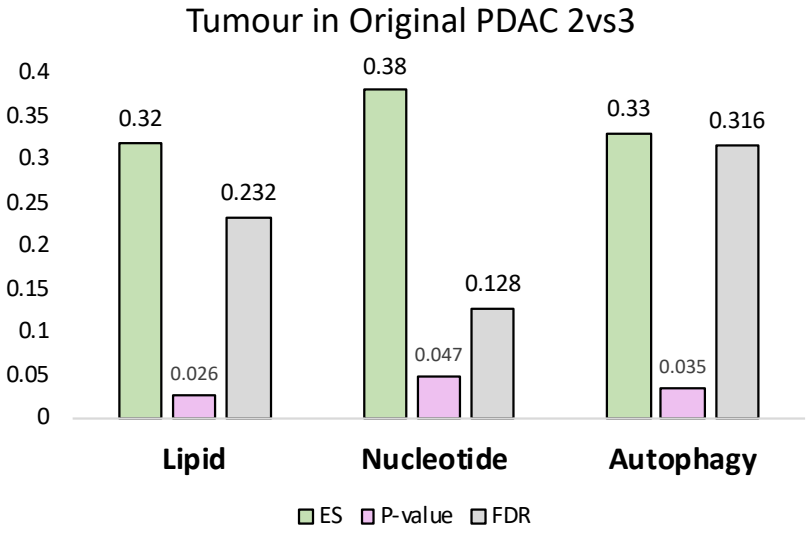
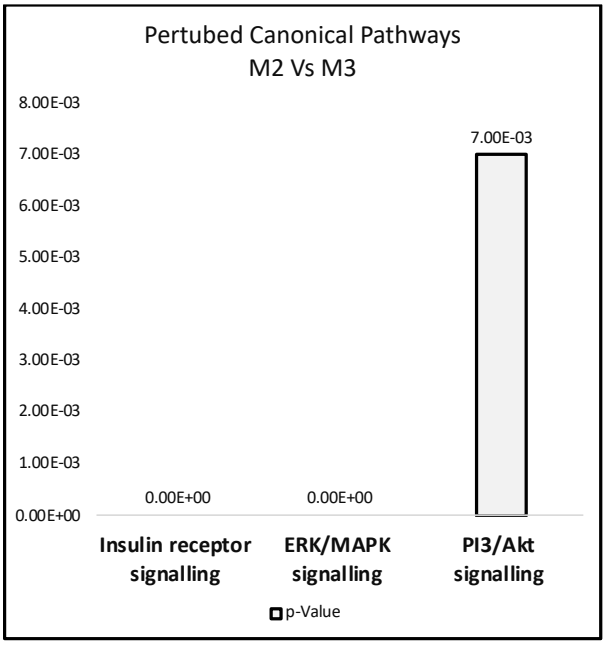


Figure 5.10: Gene set enrichment analysis of deconvolved tumour M2.

The green bar represents enrichment score (ES); purple indicates p-value and grey signifies FDR value. Lipid and Nucleotide metabolic pathways were significantly enriched in M2. No hallmark function was enriched.



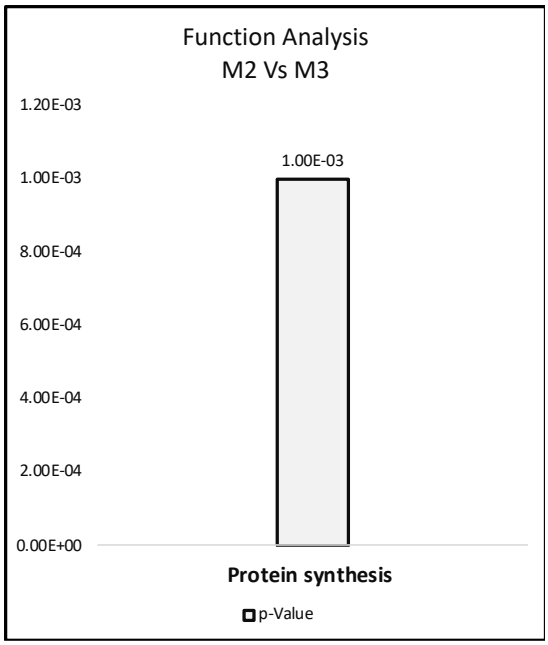


Figure 5.11: Altered canonical pathways and functional analysis of deconvolved tumour M2.

The X-axis in the bar chart denotes pathways and functions, and the Y-axis signifies p-value. We performed the analysis by comparing each subtype with the best prognosis group (M3). Protein synthesis was affected in this subgroup.

In the worst subtypes (M1 & M4), most of the critical signalling pathways were highly perturbed, resulting in alterations in major cellular and molecular functions.

5.3.4 Association of identified metabolic subtypes with tumour genomic profile

One of the potent drivers of tumour initiation is mutant KRAS oncogene. In PDAC, KRAS mutation is frequently observed. Likewise, mutations in different tumour suppressor genes, including Tp53, SMAD4, CDKN2A, are normal in PDAC. Evidence suggests that all these mutations cause aggressive PDAC tumour growth. To investigate different oncogenic events, we determined the mutation and copy number alteration frequency across the identified metabolic

subtypes. From the mutation analysis, we found significant enrichment of KRAS mutation (p-value 4.62E-08, q-value 4.509E-04) among the metabolic subtypes [Figure 5.12]; however, the copy number alteration frequency was not significant [Figure 5.13]. Our finding, significant KRAS mutation in the worst metabolic subtype, is compatible with different studies that show KRAS drive tumour metabolism in the direction of glycolysis in PDAC. Furthermore, this has been supported by our pathway analysis, where we found that the M1 subtype has increased glycolysis. Thus, PDAC tumours with KRAS mutation are dependent on glucose utilization and vulnerable to the glycolytic inhibition.

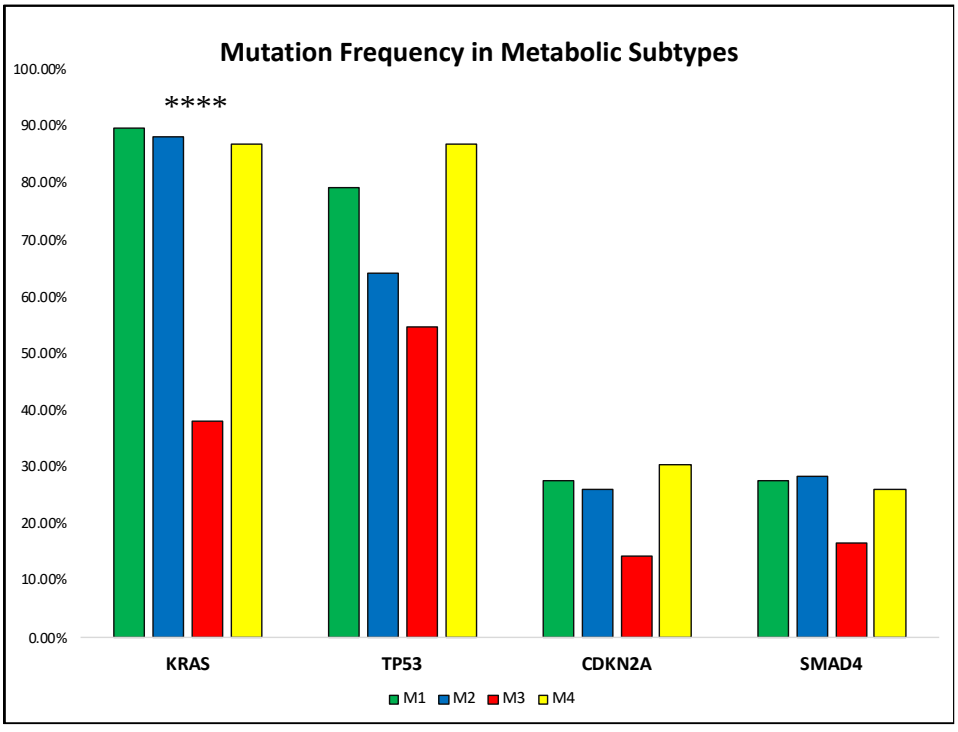


Figure 5.12: Mutational analysis of identified PDAC metabolic subtypes.

The bar chart represents the most common mutant genes in PDAC. Colour bars at the x-axis represent the groups of patients: M1 is red, M2 is blue, M3 is red, and M4 is yellow. Y-axis

represents percentage of mutations in each metabolic group. Kruskal-Wallis test shows that mutant KRAS is significantly enriched across identified metabolic subtypes of PDAC.

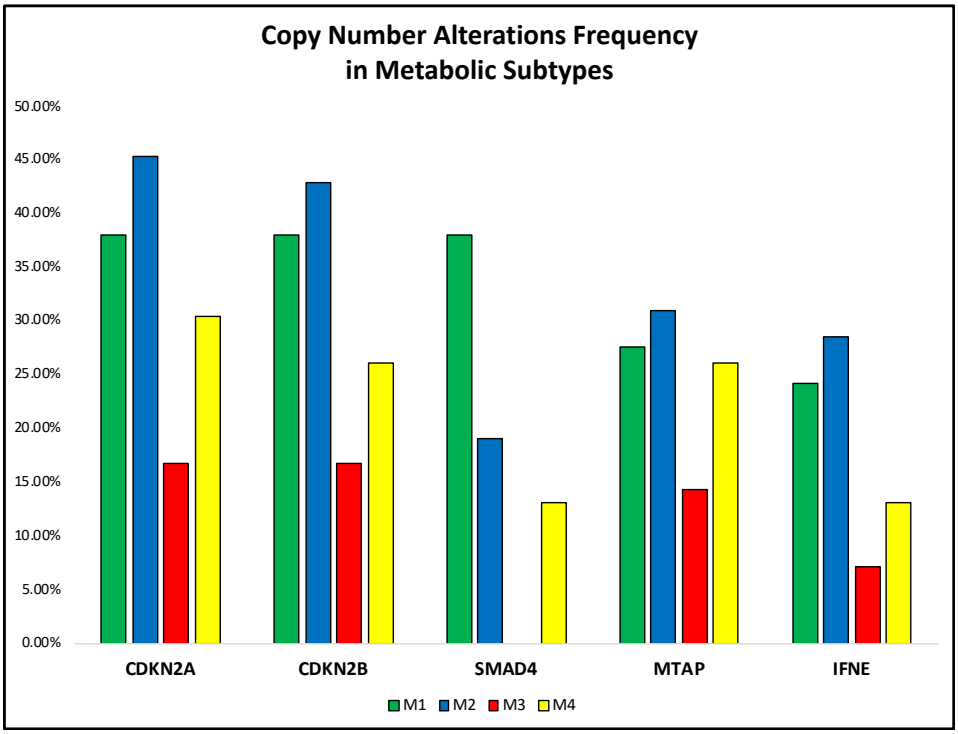


Figure 5.13: Copy number alteration analysis of identified PDAC metabolic subtypes.

The bar chart represents the most common CNV frequency in PDAC. Colour bars at the x-axis represent the groups of patients: M1 is red, M2 is blue, M3 is red, and M4 is yellow. Y-axis represents percentage of alterations in each metabolic group. No significant alteration observed across identified metabolic subtypes.

5.3.5 Metabolic expression subtypes are informative to understand the role of stroma cells in PDAC

Dense stroma is one of the prominent features of PDAC. The stroma is composed of different extracellular matrix proteins, immune cells, stellate cells, endothelial cells, fibroblasts, but the biology of stroma cells is not clearly understood. To address this, we investigated the role of stroma in the identified subtypes in a similar way as we assessed the tumour compartment. From the analysis, no significant metabolic activity and major hallmarks of cancer were altered. The exception was an increase in fatty acid metabolism (p-value 0.006; FDR 0.134), particularly beta-oxidation process, in the worst subtype (M1). However, we looked at the 50 hallmark functions and found positively enriched functions, including G2M checkpoint (p-value < 0.001), E2F signalling (p-value 0.006), MTORC1 signalling (p-value 0.006) and Glycolysis (p-value 0.045) in the worst metabolic subtype, M1. In the intermediate group M4, we found similar altered functions. However, in the good prognosis group M2 we observed enriched G2M checkpoint (p-

value 0.008), myogenesis (p-value 0.016) and JAK/STAT signalling (p-value 0.033) [Figure 5.14].

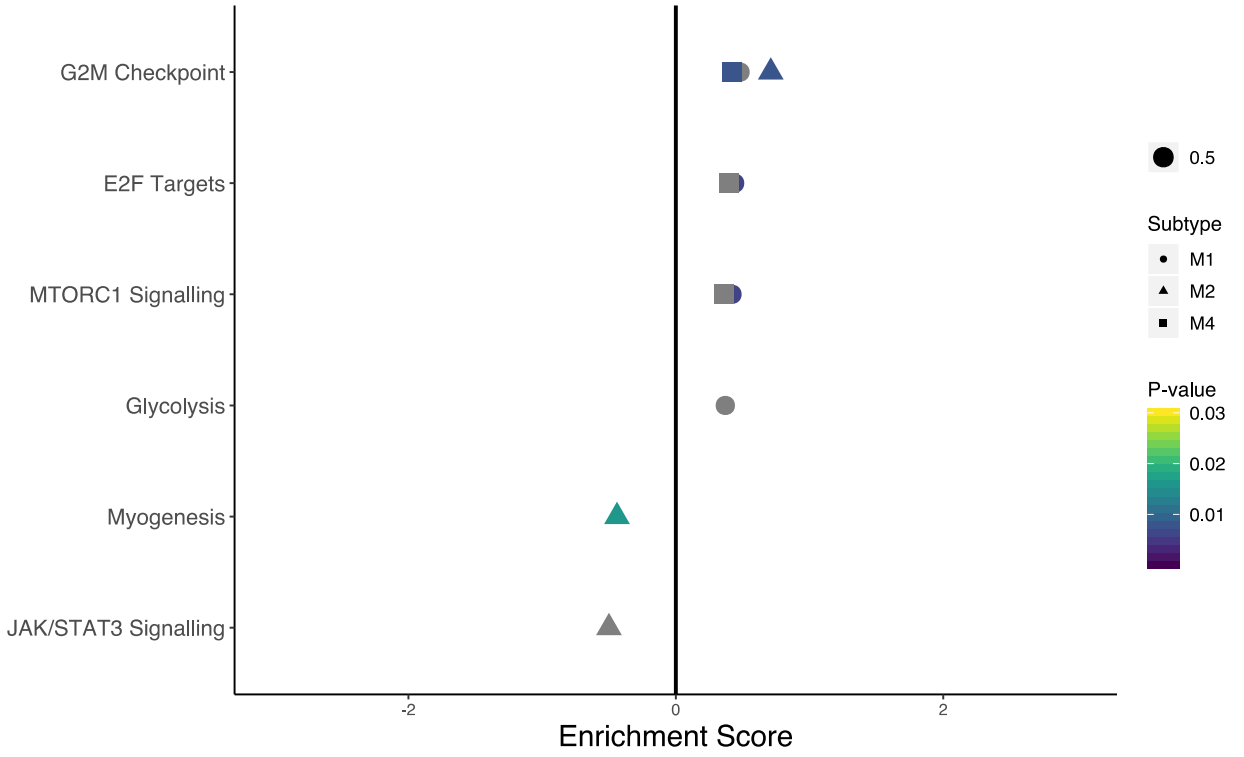


Figure 5.14: 50 hallmark functions analysis of PDAC deconvolved stroma.

The x-axis represents enrichment score, and the y-axis signifies the altered functions. Different colour code shows the significance level of enrichment. Metabolic subtypes M1 and M4 shared similar alterations (positively enriched) except glycolysis, while metabolic subtype M2 has negatively enriched JAK/STAT signalling and myogenesis.

Furthermore, we found different signalling pathways disrupted in the stroma compartment of all metabolic subtypes. In M1, there was elevated JAK/Stat (p-value 0.013), telomerase (p-value 0.01), integrin (p-value 0.04) and micropinocytosis (p-value 0.05) signalling, but decreased p53 (p-value 0.03) and death receptor (p-value 0.05) signalling [Figure 5.15]. Genes

including receptor tyrosine kinases (RTK), heat shock protein 70 & 72 (HSP70 & 72), transforming growth factor-beta receptor (TGFβR) were found to highly expressed in the stroma compartment of the worst subtype. From previous studies, it is well established that integrins are crucial for the remodelling of the extracellular matrix. Similarly, JAK/STAT signalling pathways found to play an important role in the functional regulation of pancreatic stellate cells (PSC). Our study demonstrates similar signalling alterations in the stroma compartment. Our results indicate that in subtype M1, stroma cells behave more aggressively compared to other subtypes, which could contribute to its classification as the worst prognostic group.

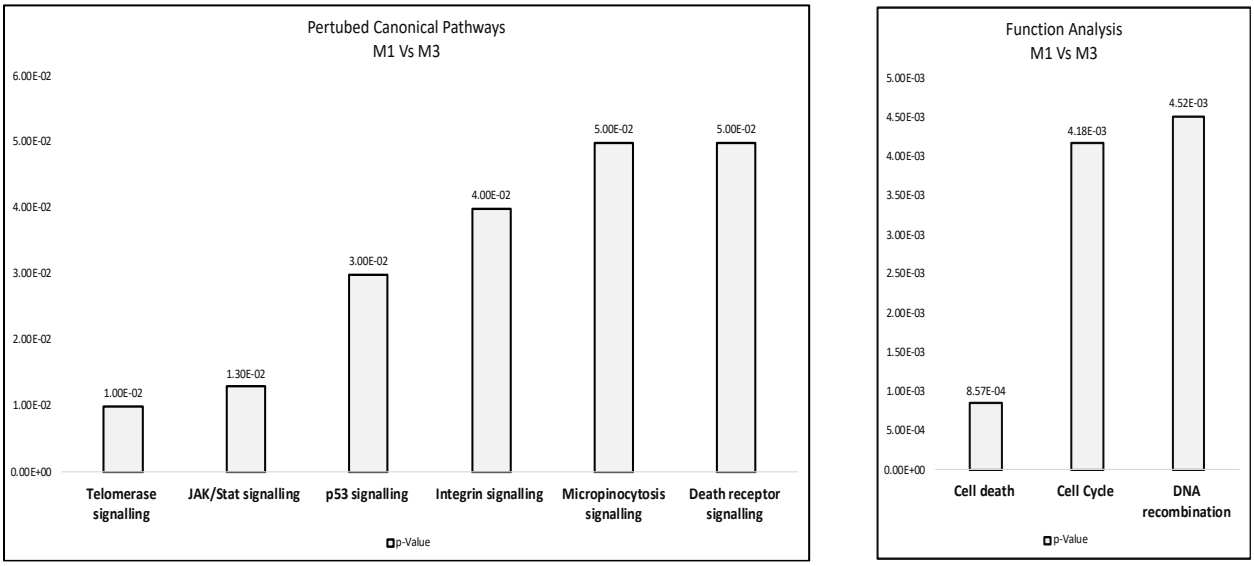
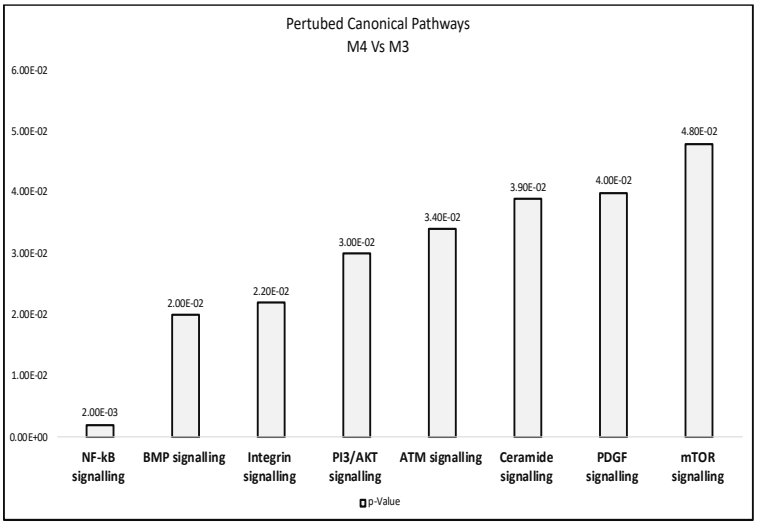


Figure 5.15: Pathway analysis of deconvolved stroma M1.

Bar chart representing most perturbed canonical pathways and associated altered function in M1 PDAC metabolic subtype compared to the best prognosis group (M3). The X-axis in the bar chart denotes pathways and functions, and the Y-axis represents p-value. Identified

dysregulated canonical pathways in stroma cells result in significant disruption in the cell cycle process, DNA repair process and cellular survival in the worst metabolic subtype.

In intermediate group (M4), most of the signalling pathways were significantly down-regulated: NF-kB (p-value 0.002), BMP (p-value 0.02), Integrin (p-value 0.022), PI3/AKT (p-value 0.03), ATM (P-value 0.034), PDGF (p-value 0.04), ceramide (p-value 0.039), and mTOR (p-value 0.048) signalling. All these perturbations lead to abnormal cell cycle progression (p-value 0.001), DNA repair process (p-value < 0.001) and apoptosis (p-value 0.01) [Figure 5.16]. The alteration of the functions is well explained from the gene expression levels, such as increased TSC complex subunit 2 (TSC2), splicing factor 3B (SF3B), Aquarius intron-binding spliceosomal factor (AQR), chemokine ligand 27 (CCL27), and decreased G-protein coupled receptor (GPCR), serine/threonine kinase 11 (LKB1), TSC complex subunit 1 (TSC1), bone morphogenetic protein 6 &15 (BMP 6 & 15), interferon-alpha and beta receptor subunit 2 (IFNAR2). Interestingly, subtype M1 and M4 dysregulate pathways with similar functions, like cell progression, DNA repair process, and cell death, but exhibit opposite signalling alterations with few exceptions. All the canonical pathways found in subtype M1 are significantly up-regulated except two down-regulated pathways, but in subtype M4, all the pathways found to be down-regulated. Thus, our result demonstrates how stroma cells perturb major functions within cells by exploiting different signalling pathways resulting in a more aggressive disease phenotype.



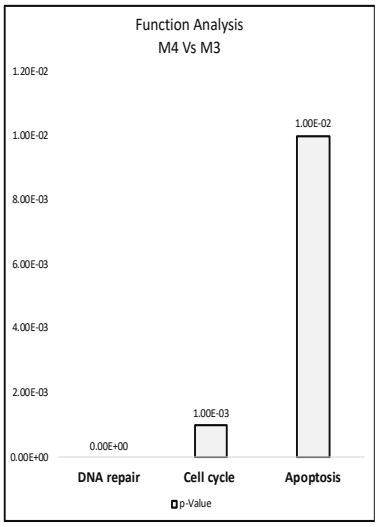


Figure 5.16: Pathway analysis of deconvolved stroma M4.

Bar chart representing most perturbed canonical pathways and associated altered function in intermediate group (M4) PDAC metabolic subtype compared to the best prognosis group (M3). The X-axis in the bar chart denotes pathways and functions, and the Y-axis indicates p-value. Major Identified dysregulated canonical pathways in stroma cells result in significant disruption in the cell cycle process, DNA repair process and cellular survival in the intermediate metabolic subtype.

Lastly, in the good prognosis group, there were no significant changes observed in the major canonical pathways. However, the RNA post-transcriptional modification (p-value < 0.001) was affected. The source of these changes in the RNA modification remains unclear in our analysis.

However, stroma compartment of each PDAC metabolic subtype reveals diverse biological alterations. Our results show that stroma cells in PDAC possess a unique mechanism of affecting most of the significant signalling cascades.

5.4 Discussion

We developed a systematic approach to identify metabolic subtypes in PDAC that will help us to determine associated aberrant biochemical and biological functions. Potentially these subtypes could instruct on novel therapeutic targets. Our approach has unveiled diverse metabolic activity in different cell compartments (tumour and stroma). The main objective of this study was the identification and characterization of meaningful metabolic subtypes in PDAC. One of the main advantages of this study is the classification of metabolic subtypes that allows investigating functions in different compartments of a bulk tumour as well as their functional state, for example, upregulation or downregulation of a specific pathway.

The identified metabolic subtypes in this study have potential clinical implications. From our analysis, we found that increased carbohydrate metabolism and amino acid metabolism are associated with the worst prognosis, where perturbations in lipid metabolism were associated with the opposite outcome. We found that the worst subtype relies on glycolysis in tumour cells and fatty acid beta-oxidation in stroma cells to produce energy. Immune subtype correlation analysis indicates the worst metabolic subtype has an increased tumour cell proliferation rate (due to elevated angiogenic gene expressions). The normal functions of immune cells is highly impeded within the worst subtype, which is further supported by the pathway analysis. Interestingly, the intermediate group shows a different metabolic dependency of tumour cells

involving high levels of amino acid secretion in association with autophagy, as well as increased TCA cycle. Conversely, tumour cells in the good prognosis group exhibit dependency on lipid metabolism. Most of the previous studies (e.g. Follia et al., Karasinska et al.,) done in PDAC were focused on a specific alteration such as carbohydrate or lipid metabolism. This study provides a more comprehensive picture.

PDAC is a stroma rich cancer, and one of the aims of this study was to explore the functions of stroma cells in the identified metabolic subtypes. The metabolic activity of stroma cells remains unclear; however, we did identify perturbed signalling pathways in the stroma in association with the metabolic subtypes. These pathways affect cellular proliferation, differentiation and survival. For example, in vivo study shows that STAT signalling significantly upregulated in stroma cells which eventually help stroma cells to change their shape/structure and form tumour [176]. Likewise, p53 and PTEN signalling are crucial to controlling the abnormal cell proliferation. Our study displayed increased JAK/STAT and decreased p53 and PTEN signalling which is a clear indication that stroma cells contribute to the progression of PDAC. For instance, in the worst metabolic subtype (M1), stroma cells manage to inhibit tumour suppressing signalling pathways, which indicates that stroma cells are contributing to abnormal cell proliferation, thus affecting the cell cycle process. Furthermore, in our study we also found higher expression levels of genes involved in integrin and transforming growth factor beta signalling resulting in remodelling of the extracellular matrix. One of the interesting observations of the stroma cells is their distinct activity across the metabolic subtypes. In one subtype, they altered the signalling pathways by both activating and inhibiting them, while in the other subtype, they

nearly shut down the signalling pathways by inhibiting them from maintaining their functions. Our results suggest that the stroma cells maintain a distinctive mechanism of altering signalling pathways in PDAC.

Overall, our study shows that PDAC is metabolically diverse. Each metabolic subtype is dependent on different metabolic and signalling pathways to maintain their activities. Focusing on the clinical outcome, it is clear that some of downstream effects of the pathways are less harmful compared to others. Thus, identification of the metabolic subtypes provides a better insight into different biological pathways that are vulnerable to treatment.

Chapter Six: Concluding Remarks

In patients with cancer, chemotherapy is an essential mainstay of treatment. However, chemotherapy is toxic, and it is never known whether chemotherapy will be effective on any individual. Since only a proportion of individuals respond to chemotherapy, that means that a proportion of patients must endure the toxicities of chemotherapy without seeing any benefits. It is for this reason that substantial effort is being made to individualize treatments for cancer.

In recent years, it has become possible to perform whole-genome sequencing of a tumour, making it possible to identify all of the mutations contained within a tumour. Many groups have utilized this information to predict the types of drugs that might be effective in an individual. Indeed, some notable successes observed using this approach. However, this approach is not helpful for all individuals.

Given the link between metabolism and clinical progression, we hypothesize that metabolic features dictate the clinical and biological behaviour in different types of cancer, particularly in pancreatic ductal adenocarcinoma and breast cancer. We aimed to develop a systemic approach using the transcriptome to better understand the specific metabolic characteristics of individuals with these two cancer types. Thus, we developed a practical systemic approach, which is transcriptional based analytical workflow (**Chapter three**); then, we applied that approach to the two different types of cancer to test the effectiveness of it. First, we applied the workflow using a PCA based clustering method on a bulk tumour focusing breast cancer on identifying meaningful metabolic subtypes (**Chapter four**); then, we used the workflow using a different clustering method, NMF, for the identification of metabolic subtypes in PDAC

(**Chapter five**). We also explored the metabolic and biological functions of both tumour and stroma cells in identified metabolic subtypes (**compartmental analysis; Chapter five**). Our result suggests that the NMF clustering method is more powerful and useful in identifying metabolic subgroups compared to PCA. The proposed step-by-step analytical workflow provides detailed insight into the metabolic perturbations. Through this approach, we were able to determine the diverse signalling alterations that occur within tumour and stroma cells and how they affect the critical cellular and molecular functions. Not only that, but based on this approach, it is possible to determine critical mutations, copy number alterations, and different epigenetic events. Identifying the metabolic features that characterize a particular tumour provides insight into biological pathways that are vulnerable to treatment. This novel approach to enhance clinical decisions using molecular information may alter the way we individualize cancer care in the future, minimizing drug exposure to patients who will not benefit and enhancing the likelihood that they will benefit from any particular drug.

There are several interesting future directions to follow up on this research:

1. Validation - use a different, larger cohort of patients.

2. Single-cell analysis - to determine whether the identified metabolic subtypes appear uniformly throughout the tumour or there is a strong presence of heterogeneity.

3. Metabolic profiling – identify candidate proteins that may be targets of systemic therapy using liquid-chromatography-mass spectrometry (LC-MS) in conjunction with the transcriptomic information.

References

1. Dang, C. (2012). Links between metabolism and cancer. Genes & Development, 26(9), pp.877-890.
2. Cantor, J. and Sabatini, D. (2012). Cancer Cell Metabolism: One Hallmark, Many Faces. Cancer Discovery, 2(10), pp.881-898.
3. Kato, Y., Maeda, T., Suzuki, A. and Baba, Y. (2018). Cancer metabolism: New insights into classic characteristics. Japanese Dental Science Review, 54(1), pp.8-21.
4. Neumann, C., von Hörschelmann, E., Reutzel-Selke, A., Seidel, E., Sauer, I., Pratschke, J., Bahra, M. and Schmuck, R. (2018). Tumour–stromal cross-talk modulating the therapeutic response in pancreatic cancer. Hepatobiliary & Pancreatic Diseases International, 17(5), pp.461-472.
5. Criscitiello, C., Esposito, A. and Curigliano, G. (2014). Tumour–stroma crosstalk. Current Opinion in Oncology, 26(6), pp.551-555.
6. Bremnes, R., Dønnem, T., Al-Saad, S., Al-Shibli, K., Andersen, S., Sirera, R., Camps, C., Marinez, I. and Busund, L. (2011). The Role of Tumour Stroma in Cancer Progression and Prognosis: Emphasis on Carcinoma-Associated Fibroblasts and Non-small Cell Lung Cancer. Journal of Thoracic Oncology, 6(1), pp.209-217.
7. Morandi, A., Giannoni, E. and Chiarugi, P. (2016). Nutrient Exploitation within the Tumour–Stroma Metabolic Crosstalk. Trends in Cancer, 2(12), pp.736-746.

8. Mira, A., Morello, V., Céspedes, M., Perera, T., Comoglio, P., Mangues, R. and Michieli, P. (2017). Stroma-derived HGF drives metabolic adaptation of colorectal cancer to angiogenesis inhibitors. Oncotarget, 8(24).
9. Lacaria, M., Gu, W. and Lupski, J. (2013). A functional role for structural variation in metabolism. Adipocyte, 2(1), pp.55-57.
10. Wishart, D. (2015). Is Cancer a Genetic Disease or a Metabolic Disease?. EBioMedicine, 2(6), pp.478-479.
11. Fan, J., Teng, X., Liu, L., Mattaini, K., Looper, R., Vander Heiden, M. and Rabinowitz, J. (2014). Human Phosphoglycerate Dehydrogenase Produces the Oncometabolite d-2-Hydroxyglutarate. ACS Chemical Biology, 10(2), pp.510-516.
12. Zhou, Z., Ibekwe, E. and Chornenkyy, Y. (2018). Metabolic Alterations in Cancer Cells and the Emerging Role of Oncometabolites as Drivers of Neoplastic Change. Antioxidants, 7(1), p.16.
13. Sharick, J., Jeffery, J., Karim, M., Walsh, C., Esbona, K., Cook, R. and Skala, M. (2019). Cellular Metabolic Heterogeneity In Vivo Is Recapitulated in Tumour Organoids. Neoplasia, 21(6), pp.615-626.
14. Gentric, G., Mieulet, V. and Mechta-Grigoriou, F. (2017). Heterogeneity in Cancer Metabolism: New Concepts in an Old Field. Antioxidants & Redox Signaling, 26(9), pp.462-485.

15. Peng, X., Chen, Z., Farshidfar, F., Xu, X., Lorenzi, P., Wang, Y., Cheng, F., Tan, L., Mojumdar, K., Du, D., Ge, Z., Li, J., Thomas, G., Birsoy, K., Liu, L., Zhang, H., Zhao, Z., Marchand, C., Weinstein, J., Bathe, O., Liang, H. (2018). Molecular Characterization and Clinical Relevance of Metabolic Expression Subtypes in Human Cancers. Cell Reports, 23(1), pp.255-269.e4.
16. Rosario, S., Long, M., Affronti, H., Rowsam, A., Eng, K. and Smiraglia, D. (2018). Pan-cancer analysis of transcriptional metabolic dysregulation using The Cancer Genome Atlas. Nature Communications, 9(1).
17. Xiao, Z., Dai, Z. and Locasale, J. (2019). Metabolic landscape of the tumour microenvironment at single cell resolution. Nature Communications, 10(1).
18. Haukaas, T., Euceda, L., Giskeødegård, G., Lamichhane, S., Krohn, M., Jernström, S., Aure, M., Lingjærde, O., Schlichting, E., Garred, Ø., Due, E., Mills, G., Sahlberg, K., Børresen-Dale, A. and Bathen, T. (2016). Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes. Cancer & Metabolism, 4(1).
19. Gillies, R., Robey, I. and Gatenby, R. (2008). Causes and Consequences of Increased Glucose Metabolism of Cancers. Journal of Nuclear Medicine, 49(Suppl_2), pp.24S-42S.
20. Warburg, O. (1956). On the Origin of Cancer Cells. Science, 123(3191), pp.309-314.
21. Burns, J. and Manda, G. (2017). Metabolic Pathways of the Warburg Effect in Health and Disease: Perspectives of Choice, Chain or Chance. International Journal of Molecular Sciences, 18(12), p.2755.

22. Fadaka, A., Ajiboye, B., Ojo, O., Adewale, O., Olayide, I. and Emuowhochere, R. (2017). Biology of glucose metabolization in cancer cells. Journal of Oncological Sciences, 3(2), pp.45-51.

23. Patra, K. and Hay, N. (2014). The pentose phosphate pathway and cancer. Trends in Biochemical Sciences, 39(8), pp.347-354.

24. Singh, D., Arora, R., Kaur, P., Singh, B., Mannan, R. and Arora, S. (2017). Overexpression of hypoxia-inducible factor and metabolic pathways: possible targets of cancer. Cell & Bioscience, 7(1).

25. Robey, I., Lien, A., Welsh, S., Baggett, B. and Gillies, R. (2005). Hypoxia-Inducible Factor-1α and the Glycolytic Phenotype in Tumours. Neoplasia, 7(4), pp.324-330.

26. Ying, H., Kimmelman, A., Lyssiotis, C., Hua, S., Chu, G., Fletcher-Sananikone, E., Locasale, J., Son, J., Zhang, H., Coloff, J., Yan, H., Wang, W., Chen, S., Viale, A., Zheng, H., Paik, J., Lim, C., Guimaraes, A., Martin, E., Chang, J., Hezel, A., Perry, S., Hu, J., Gan, B., Xiao, Y., Asara, J., Weissleder, R., Wang, Y., Chin, L., Cantley, L. and DePinho, R. (2012). Oncogenic Kras Maintains Pancreatic Tumours through Regulation of Anabolic Glucose Metabolism. Cell, 149(3), pp.656-670.

27. Patra, K. and Hay, N. (2014). The pentose phosphate pathway and cancer. Trends in Biochemical Sciences, 39(8), pp.347-354.

28. Andersen, A., Moreira, J. and Pedersen, S. (2014). Interactions of ion transporters and channels with cancer cell metabolism and the tumour microenvironment. Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1638), p.20130098.

29. Kongara, S. and Karantza, V. (2012). The interplay between autophagy and ROS in tumourigenesis. Frontiers in Oncology, 2.

30. Li, Z., Yang, Y., Ming, M. and Liu, B. (2011). Mitochondrial ROS generation for regulation of autophagic pathways in cancer. Biochemical and Biophysical Research Communications, 414(1), pp.5-8.

31. Liu, E. and Ryan, K. (2012). Autophagy and cancer - issues we need to digest. Journal of Cell Science, 125(10), pp.2349-2358.

32. White, E. (2015). The role for autophagy in cancer. Journal of Clinical Investigation, 125(1), pp.42-46.

33. Zhang, F. (2012). Dysregulated lipid metabolism in cancer. World Journal of Biological Chemistry, 3(8), p.167.

34. Currie, E., Schulze, A., Zechner, R., Walther, T. and Farese, R. (2013). Cellular Fatty Acid Metabolism and Cancer. Cell Metabolism, 18(2), pp.153-161.

35. Hilvo, M. and Orešič, M. (2012). Regulation of lipid metabolism in breast cancer provides diagnostic and therapeutic opportunities. Clinical Lipidology, 7(2), pp.177-188.

36. Zaytseva, Y., Rychahou, P., Le, A., Scott, T., Flight, R., Kim, J., Harris, J., Liu, J., Wang, C., Morris, A., Sivakumaran, T., Fan, T., Moseley, H., Gao, T., Lee, E., Weiss, H., Heuer, T., Kemble, G. and Evers, M. (2018). Preclinical evaluation of novel fatty acid synthase inhibitors in primary colorectal cancer cells and a patient-derived xenograft model of colorectal cancer. Oncotarget, 9(37).

37. Rashid,, A., Pizer, E., Moga, M., Milgraum, L., Zahurak, M., Pasternack, G., Kuhajda, F. and Hamilton, S. (1997). Elevated Expression of Fatty Acid Synthase and Fatty Acid Synthetic Activity in Colorectal Neoplasia. American Journal of Pathology, 150(1).

38. Zaytseva, Y., Harris, J., Mitov, M., Kim, J., Butterfield, D., Lee, E., Weiss, H., Gao, T. and Evers, B. (2015). Increased expression of fatty acid synthase provides a survival advantage to colorectal cancer cells via upregulation of cellular respiration. Oncotarget, 6(22).

39. Lu, T., Sun, L., Wang, Z., Zhang, Y., He, Z. and Xu, C. (2019). Fatty acid synthase enhances colorectal cancer cell proliferation and metastasis via regulating AMPK/mTOR pathway . OncoTargets and Therapy, Volume 12, pp.3339-3347.

40. Swinnen, J., Roskams, T., Joniau, S., Van Poppel, H., Oyen, R., Baert, L., Heyns, W. and Verhoeven, G. (2002). Overexpression of fatty acid synthase is an early and common event in the development of prostate cancer. International Journal of Cancer, 98(1), pp.19-22.

41. Joseph, S., Laffitte, B., Patel, P., Watson, M., Matsukuma, K., Walczak, R., Collins, J., Osborne, T. and Tontonoz, P. (2002). Direct and Indirect Mechanisms for Regulation of Fatty Acid Synthase Gene Expression by Liver X Receptors. Journal of Biological Chemistry, 277(13), pp.11019-11025.

42. Yang, Y., Morin, P., Han, W., Chen, T., Bornman, D., Gabrielson, E. and Pizer, E. (2003). Regulation of fatty acid synthase expression in breast cancer by sterol regulatory element binding protein-1c. Experimental Cell Research, 282(2), pp.132-137.

43. Wu, S. and Näär, A. (2019). SREBP1-dependent de novo fatty acid synthesis gene expression is elevated in malignant melanoma and represents a cellular survival trait. Scientific Reports, 9(1).

44. Lee, M., Moon, J., Park, S., Koh, Y., Ahn, Y. and Kim, K. (2008). KLF5 enhances SREBP-1 action in androgen-dependent induction of fatty acid synthase in prostate cancer cells. Biochemical Journal, 417(1), pp.313-322.

45. Zaidi, N., Swinnen, J. and Smans, K. (2012). ATP-Citrate Lyase: A Key Player in Cancer Metabolism. Cancer Research, 72(15), pp.3709-3714.

46. Wen, J., Min, X., Shen, M., Hua, Q., Han, Y., Zhao, L., Liu, L., Huang, G., Liu, J. and Zhao, X. (2019). ACLY facilitates colon cancer cell metastasis by CTNNB1. Journal of Experimental & Clinical Cancer Research, 38(1).

47. Migita, T., Okabe, S., Ikeda, K., Igarashi, S., Sugawara, S., Tomida, A., Taguchi, R., Soga, T. and Seimiya, H. (2013). Inhibition of ATP Citrate Lyase Induces an Anticancer Effect via Reactive Oxygen Species. The American Journal of Pathology, 182(5), pp.1800-1810.

48. Wang, D., Yin, L., Wei, J., Yang, Z. and Jiang, G. (2017). ATP citrate lyase is increased in human breast cancer, depletion of which promotes apoptosis. Tumour Biology, 39(4), p.101042831769833.

49. Khwairakpam, A., Shyamananda, M., Sailo, B., Rathnakaram, S., Padmavathi, G., Kotoky, J. and Kunnumakkara, A. (2015). ATP Citrate Lyase (ACLY): A Promising Target for Cancer Prevention and Treatment. Current Drug Targets, 16(2), pp.156-163.

50. Granchi, C. (2018). ATP citrate lyase (ACLY) inhibitors: An anti-cancer strategy at the crossroads of glucose and lipid metabolism. European Journal of Medicinal Chemistry, 157, pp.1276-1291.

51. Llaverias, G., Danilo, C., Mercier, I., Daumer, K., Capozza, F., Williams, T., Sotgia, F., Lisanti, M. and Frank, P. (2011). Role of Cholesterol in the Development and Progression of Breast Cancer. The American Journal of Pathology, 178(1), pp.402-412.

52. 2. Moon, H., Ruelcke, J., Choi, E., Sharpe, L., Nassar, Z., Bielefeldt-Ohmann, H., Parat, M., Shah, A., Francois, M., Inder, K., Brown, A., Russell, P., Parton, R. and Hill, M. (2015). Diet-induced hypercholesterolemia promotes androgen-independent prostate cancer metastasis via IQGAP1 and caveolin-1. Oncotarget, 6(10).

53. 3. Castellano, B., Thelen, A., Moldavski, O., Feltes, M., van der Welle, R., Mydock-McGrane, L., Jiang, X., van Eijkeren, R., Davis, O., Louie, S., Perera, R., Covey, D., Nomura, D., Ory, D. and Zoncu, R. (2017). Lysosomal cholesterol activates mTORC1 via an SLC38A9–Niemann-Pick C1 signaling complex. Science, 355(6331), pp.1306-1311.

54. 4. Montero, J., Morales, A., Llacuna, L., Lluis, J., Terrones, O., Basanez, G., Antonsson, B., Prieto, J., Garcia-Ruiz, C., Colell, A. and Fernandez-Checa, J. (2008). Mitochondrial Cholesterol Contributes to Chemotherapy Resistance in Hepatocellular Carcinoma. Cancer Research, 68(13), pp.5246-5256.

55. Karasinska, J., Topham, J., Kalloger, S., Jang, G., Denroche, R., Culibrk, L., Williamson, L., Wong, H., Lee, M., O'Kane, G., Moore, R., Mungall, A., Moore, M., Warren, C., Metcalfe, A., Notta, F., Knox, J., Gallinger, S., Laskin, J., Marra, M., Jones, S., Renouf, D. and

Schaeffer, D. (2019). Metabolic gene expression heterogeneity and survival in PDAC. *Oncotarget*.

56. Skotland, T., Sandvig, K. and Llorente, A. (2017). Lipids in exosomes: Current knowledge and the way forward. Progress in Lipid Research, 66, pp.30-41.
57. Elsherbini, A. and Bieberich, E. (2018). Ceramide and Exosomes: A Novel Target in Cancer Biology and Therapy. Advances in Cancer Research, pp.121-154.
58. Zhao, H., Yang, L., Baddour, J., Achreja, A., Bernard, V., Moss, T., Marini, J., Tudawe, T., Seviour, E., San Lucas, F., Alvarez, H., Gupta, S., Maiti, S., Cooper, L., Peehl, D., Ram, P., Maitra, A. and Nagrath, D. (2016). Tumour microenvironment derived exosomes pleiotropically modulate cancer cell metabolism. eLife, 5.
59. Ji, H., Greening, D., Barnes, T., Lim, J., Tauro, B., Rai, A., Xu, R., Adda, C., Mathivanan, S., Zhao, W., Xue, Y., Xu, T., Zhu, H. and Simpson, R. (2013). Proteome profiling of exosomes derived from human primary and metastatic colorectal cancer cells reveal differential expression of key metastatic factors and signal transduction components. PROTEOMICS, 13(10-11), pp.1672-1686.
60. Brzozowski, J., Jankowski, H., Bond, D., McCague, S., Munro, B., Predebon, M., Scarlett, C., Skelding, K. and Weidenhofer, J. (2018). Lipidomic profiling of extracellular vesicles derived from prostate and prostate cancer cell lines. Lipids in Health and Disease, 17(1).
61. Timosenko, E., Hadjinicolaou, A. and Cerundolo, V. (2017). Modulation of cancer-specific immune responses by amino acid degrading enzymes. Immunotherapy, 9(1), pp.83-97.

62. Vučetić, M., Cormerais, Y., Parks, S. and Pouysségur, J. (2017). The Central Role of Amino Acids in Cancer Redox Homeostasis: Vulnerability Points of the Cancer Redox Code. Frontiers in Oncology, 7.

63. Zhang, J., Pavlova, N. and Thompson, C. (2017). Cancer cell metabolism: the essential role of the nonessential amino acid, glutamine. The EMBO Journal, 36(10), pp.1302-1315.

64. Bott, A., Peng, I., Fan, Y., Faubert, B., Zhao, L., Li, J., Neidler, S., Sun, Y., Jaber, N., Krokowski, D., Lu, W., Pan, J., Powers, S., Rabinowitz, J., Hatzoglou, M., Murphy, D., Jones, R., Wu, S., Girnun, G. and Zong, W. (2015). Oncogenic Myc Induces Expression of Glutamine Synthetase through Promoter Demethylation. Cell Metabolism, 22(6), pp.1068-1077.

65. Zhu, Y., Li, T., Ramos da Silva, S., Lee, J., Lu, C., Eoh, H., Jung, J. and Gao, S. (2017). A Critical Role of Glutamine and Asparagine γ-Nitrogen in Nucleotide Biosynthesis in Cancer Cells Hijacked by an Oncogenic Virus. mBio, 8(4).

66. Albaugh, V., Pinzon-Guzman, C. and Barbul, A. (2016). Arginine-Dual roles as an onconutrient and immunonutrient. Journal of Surgical Oncology, 115(3), pp.273-280.

67. Patil, M., Bhaumik, J., Babykutty, S., Banerjee, U. and Fukumura, D. (2016). Arginine dependence of tumour cells: targeting a chink in cancer's armor. Oncogene, 35(38), pp.4957-4972.

68. Sippel, T., White, J., Nag, K., Tsvankin, V., Klaassen, M., Kleinschmidt-DeMasters, B. and Waziri, A. (2011). Neutrophil Degranulation and Immunosuppression in Patients with

GBM: Restoration of Cellular Immune Function by Targeting Arginase I. Clinical Cancer Research, 17(22), pp.6992-7002.

69. Rodriguez, P., Ernstoff, M., Hernandez, C., Atkins, M., Zabaleta, J., Sierra, R. and Ochoa, A. (2009). Arginase I–Producing Myeloid-Derived Suppressor Cells in Renal Cell Carcinoma Are a Subpopulation of Activated Granulocytes. Cancer Research, 69(4), pp.1553-1560.
70. Aulak, K., Miyagi, M., Yan, L., West, K., Massillon, D., Crabb, J. and Stuehr, D. (2001). Proteomic method identifies proteins nitrated in vivo during inflammatory challenge. Proceedings of the National Academy of Sciences, 98(21), pp.12056-12061.
71. Sousa, C., Biancur, D., Wang, X., Halbrook, C., Sherman, M., Zhang, L., Kremer, D., Hwang, R., Witkiewicz, A., Ying, H., Asara, J., Evans, R., Cantley, L., Lyssiotis, C. and Kimmelman, A. (2016). Pancreatic stellate cells support tumour metabolism through autophagic alanine secretion. Nature, 536(7617), pp.479-483.
72. Merhi, A., Delrée, P. and Marini, A. (2017). The metabolic waste ammonium regulates mTORC2 and mTORC1 signaling. Scientific Reports, 7(1).
73. Stincone, A., Prigione, A., Cramer, T., Wamelink, M., Campbell, K., Cheung, E., Olin-Sandoval, V., Grüning, N., Krüger, A., Tauqeer Alam, M., Keller, M., Breitenbach, M., Brindle, K., Rabinowitz, J. and Ralser, M. (2014). The return of metabolism: biochemistry and physiology of the pentose phosphate pathway. Biological Reviews, 90(3), pp.927-963.

74. Miller, D., Thomas, S., Islam, A., Muench, D. and Sedoris, K. (2012). c-Myc and Cancer Metabolism. Clinical Cancer Research, 18(20), pp.5546-5553.

75. Liu, Y., Li, F., Handler, J., Huang, C., Xiang, Y., Neretti, N., Sedivy, J., Zeller, K. and Dang, C. (2008). Global Regulation of Nucleotide Biosynthetic Genes by c-Myc. PLoS ONE, 3(7), p.e2722.

76. Saxton, R. and Sabatini, D. (2017). mTOR Signaling in Growth, Metabolism, and Disease. Cell, 168(6), pp.960-976.

77. Villa, E., Ali, E., Sahu, U. and Ben-Sahra, I. (2019). Cancer Cells Tune the Signaling Pathways to Empower de Novo Synthesis of Nucleotides. Cancers, 11(5), p.688.

78. Wu, D., Potluri, N., Lu, J., Kim, Y. and Rastinejad, F. (2015). Structural integration in hypoxia-inducible factors. Nature, 524(7565), pp.303-308

79. Maxwell, P. (2001). Activation of the HIF pathway in cancer. Current Opinion in Genetics & Development, 11(3), pp.293-299.

80. Chen, J. and Russo, J. (2012). Dysregulation of glucose transport, glycolysis, TCA cycle and glutaminolysis by oncogenes and tumour suppressors in cancer cells. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, 1826(2), pp.370-384.

81. Courtnay, R., Ngo, D., Malik, N., Ververis, K., Tortorella, S. and Karagiannis, T. (2015). Cancer metabolism and the Warburg effect: the role of HIF-1 and PI3K. Molecular Biology Reports, 42(4), pp.841-851.

82. Kamphorst, J., Chung, M., Fan, J. and Rabinowitz, J. (2014). Quantitative analysis of acetyl-CoA production in hypoxic cancer cells reveals substantial contribution from acetate. Cancer & Metabolism, 2(1).

83. Ye, J., Fan, J., Venneti, S., Wan, Y., Pawel, B., Zhang, J., Finley, L., Lu, C., Lindsten, T., Cross, J., Qing, G., Liu, Z., Simon, M., Rabinowitz, J. and Thompson, C. (2014). Serine Catabolism Regulates Mitochondrial Redox Control during Hypoxia. Cancer Discovery, 4(12), pp.1406-1417.

84. Liu, L., Shah, S., Fan, J., Park, J., Wellen, K. and Rabinowitz, J. (2016). Malic enzyme tracers reveal hypoxia-induced switch in adipocyte NADPH pathway usage. Nature Chemical Biology, 12(5), pp.345-352.

85. Luo, W. and Semenza, G. (2012). Emerging roles of PKM2 in cell metabolism and cancer progression. Trends in Endocrinology & Metabolism, 23(11), pp.560-566.

86. Hsu, M. and Hung, W. (2018). Pyruvate kinase M2 fuels multiple aspects of cancer cells: from cellular metabolism, transcriptional regulation to extracellular signaling. Molecular Cancer, 17(1).

87. Luo, W., Hu, H., Chang, R., Zhong, J., Knabel, M., O'Meally, R., Cole, R., Pandey, A. and Semenza, G. (2011). Pyruvate Kinase M2 Is a PHD3-Stimulated Coactivator for Hypoxia-Inducible Factor 1. Cell, 145(5), pp.732-744.

88. Nazio, F., Bordi, M., Cianfanelli, V., Locatelli, F. and Cecconi, F. (2019). Autophagy and cancer stem cells: molecular mechanisms and therapeutic applications. Cell Death & Differentiation, 26(4), pp.690-702.

89. Tang, D., Kang, R., Zeh, H. and Lotze, M. (2011). High-Mobility Group Box 1, Oxidative Stress, and Disease. Antioxidants & Redox Signaling, 14(7), pp.1315-1335.

90. Yang, S. and Kimmelman, A. (2011). A critical role for autophagy in pancreatic cancer. Autophagy, 7(8), pp.912-913.

91. Perera, R., Stoykova, S., Nicolay, B., Ross, K., Fitamant, J., Boukhali, M., Lengrand, J., Deshpande, V., Selig, M., Ferrone, C., Settleman, J., Stephanopoulos, G., Dyson, N., Zoncu, R., Ramaswamy, S., Haas, W. and Bardeesy, N. (2015). Transcriptional control of autophagy–lysosome function drives pancreatic cancer metabolism. Nature, 524(7565), pp.361-365.

92. Scherz-Shouval, R., Weidberg, H., Gonen, C., Wilder, S., Elazar, Z. and Oren, M. (2010). p53-dependent regulation of autophagy protein LC3 supports cancer cell survival under prolonged starvation. Proceedings of the National Academy of Sciences, 107(43), pp.18511-18516.

93. Coppola, D., Khalil, F., Eschrich, S., Boulware, D., Yeatman, T. and Wang, H. (2008). Down-regulation of Bax-interacting factor-1 in colorectal adenocarcinoma. Cancer, 113(10), pp.2665-2670.

94. Woo Lee, J., Goo Jeong, E., Hwa Soung, Y., Woo Nam, S., Young Lee, J., Jin Yoo, N. and Hyung Lee, S. (2006). Decreased expression of tumour suppressor Bax-interacting factor-1 (Bif-1), a Bax activator, in gastric carcinomas. Pathology, 38(4), pp.312-315.

95. He, S., Zhao, Z., Yang, Y., O'Connell, D., Zhang, X., Oh, S., Ma, B., Lee, J., Zhang, T., Varghese, B., Yip, J., Dolatshahi Pirooz, S., Li, M., Zhang, Y., Li, G., Ellen Martin, S.,

Machida, K. and Liang, C. (2015). Truncating mutation in the autophagy gene UVRAG confers oncogenic properties and chemosensitivity in colorectal cancers. Nature Communications, 6(1).

96. Yang, A., Rajeshkumar, N., Wang, X., Yabuuchi, S., Alexander, B., Chu, G., Von Hoff, D., Maitra, A. and Kimmelman, A. (2014). Autophagy Is Critical for Pancreatic Tumour Growth and Progression in Tumours with p53 Alterations. Cancer DISCOVERY, 4(8), pp.905-913.
97. Daemen, A., Peterson, D., Sahu, N., McCord, R., Du, X., Liu, B., Kowanetz, K., Hong, R., Moffat, J., Gao, M., Boudreau, A., Mroue, R., Corson, L., O'Brien, T., Qing, J., Sampath, D., Merchant, M., Yauch, R., Manning, G., Settleman, J., Hatzivassiliou, G. and Evangelista, M. (2015). Metabolite profiling stratifies pancreatic ductal adenocarcinomas into subtypes with distinct sensitivities to metabolic inhibitors. Proceedings of the National Academy of Sciences, 112(32), pp.E4410-E4417.
98. Keijer, J. and van Dartel, D. (2014). Reprogrammed Metabolism of Cancer Cells as a Potential Therapeutic Target. *Current Pharmaceutical Design*, 20(15), pp.2580-2594.
99. Dang, C., Hamaker, M., Sun, P., Le, A. and Gao, P. (2011). Therapeutic targeting of cancer cell metabolism. *Journal of Molecular Medicine*, 89(3), pp.205-212.
100. Ward, P., Patel, J., Wise, D., Abdel-Wahab, O., Bennett, B., Coller, H., Cross, J., Fantin, V., Hedvat, C., Perl, A., Rabinowitz, J., Carroll, M., Su, S., Sharp, K., Levine, R. and Thompson, C. (2010). The Common Feature of Leukemia-Associated IDH1 and IDH2

Mutations Is a Neomorphic Enzyme Activity Converting α-Ketoglutarate to 2-Hydroxyglutarate. *Cancer Cell*, 17(3), pp.225-234.

101. Miao, P., Sheng, S., Sun, X., Liu, J. and Huang, G. (2013). Lactate dehydrogenase a in cancer: A promising target for diagnosis and therapy. *IUBMB Life*, 65(11), pp.904-910.

102. Golub, D., Iyengar, N., Dogra, S., Wong, T., Bready, D., Tang, K., Modrek, A. and Placantonakis, D. (2019). Mutant Isocitrate Dehydrogenase Inhibitors as Targeted Cancer Therapeutics. *Frontiers in Oncology*, 9.

103. Comprehensive molecular portraits of human breast tumours. (2012). Nature, 490(7418), pp.61-70.

104. Kuijjer, M., Paulson, J., Salzman, P., Ding, W. and Quackenbush, J. (2018). Cancer subtype identification using somatic mutation data. British Journal of Cancer, 118(11), pp.1492-1501.

105. Figueroa, M., Lugthart, S., Li, Y., Erpelinck-Verschueren, C., Deng, X., Christos, P., Schifano, E., Booth, J., van Putten, W., Skrabanek, L., Campagne, F., Mazumdar, M., Greally, J., Valk, P., Löwenberg, B., Delwel, R. and Melnick, A. (2010). DNA Methylation Signatures Identify Biologically Distinct Subtypes in Acute Myeloid Leukemia. Cancer Cell, 17(1), pp.13-27.

106. Collisson, E., Sadanandam, A., Olson, P., Gibb, W., Truitt, M., Gu, S., Cooc, J., Weinkle, J., Kim, G., Jakkula, L., Feiler, H., Ko, A., Olshen, A., Danenberg, K., Tempero, M., Spellman, P., Hanahan, D. and Gray, J. (2011). Subtypes of pancreatic ductal

adenocarcinoma and their differing responses to therapy. Nature Medicine, 17(4), pp.500-503.

107. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, Ziv E, Culhane AC, Paull EO, Sivakumar IKA, Gentles AJ, Malhotra R, Farshidfar F, Colaprico A, Parker JS, Mose LE, Vo NS, Liu J, Liu Y, Rader J, Dhankani V, Reynolds SM, Bowlby R, Califano A, Cherniack AD, Anastassiou D, Bedognetti D, Mokrab Y, Newman AM, Rao A, Chen K, Krasnitz A, Hu H, Malta TM, Noushmehr H, Pedamallu CS, Bullman S, Ojesina AI, Lamb A, Zhou W, Shen H, Choueiri TK, Weinstein JN, Guinney J, Saltz J, Holt RA, Rabkin CS; Cancer Genome Atlas Research Network, Lazar AJ, Serody JS, Demicco EG, Disis ML, Vincent BG, Shmulevich I. (2018). The Immune Landscape of Cancer. Immunity, 48(4), pp.812-830.e14.

108. Parker, S. and Metallo, C. (2015). Metabolic consequences of oncogenic IDH mutations. Pharmacology & Therapeutics, 152, pp.54-62.

109. Lussey-Lepoutre, C., Hollinshead, K., Ludwig, C., Menara, M., Morin, A., Castro-Vega, L., Parker, S., Janin, M., Martinelli, C., Ottolenghi, C., Metallo, C., Gimenez-Roqueplo, A., Favier, J. and Tennant, D. (2015). Loss of succinate dehydrogenase activity results in dependency on pyruvate carboxylation for cellular anabolism. Nature Communications, 6(1).

110. Oermann, E., Wu, J., Guan, K. and Xiong, Y. (2012). Alterations of metabolic genes and metabolites in cancer. Seminars in Cell & Developmental Biology, 23(4), pp.370-380.

111. Sinkala, M., Mulder, N. and Patrick Martin, D. (2019). Metabolic gene alterations impact the clinical aggressiveness and drug responses of 32 human cancers. Communications Biology, 2(1).

112. van der Knaap, J. and Verrijzer, C. (2016). Undercover: gene control by metabolites and metabolic enzymes. Genes & Development, 30(21), pp.2345-2369.

113. Follia, L., Ferrero, G., Mandili, G., Beccuti, M., Giordano, D., Spadi, R., Satolli, M., Evangelista, A., Katayama, H., Hong, W., Momin, A., Capello, M., Hanash, S., Novelli, F. and Cordero, F. (2019). Integrative Analysis of Novel Metabolic Subtypes in Pancreatic Cancer Fosters New Prognostic Biomarkers. Frontiers in Oncology, 9.

114. Bidkhori, G., Benfeitas, R., Klevstig, M., Zhang, C., Nielsen, J., Uhlen, M., Boren, J. and Mardinoglu, A. (2018). Metabolic network-based stratification of hepatocellular carcinoma reveals three distinct tumour subtypes. Proceedings of the National Academy of Sciences, 115(50), pp.E11874-E11883.

115. Haukaas, T., Euceda, L., Giskeødegård, G., Lamichhane, S., Krohn, M., Jernström, S., Aure, M., Lingjærde, O., Schlichting, E., Garred, Ø., Due, E., Mills, G., Sahlberg, K., Børresen-Dale, A. and Bathen, T. (2016). Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes. Cancer & Metabolism, 4(1).

116. Li, H., Ning, S., Ghandi, M., Kryukov, G., Gopal, S., Deik, A., Souza, A., Pierce, K., Keskula, P., Hernandez, D., Ann, J., Shkoza, D., Apfel, V., Zou, Y., Vazquez, F., Barretina, J., Pagliarini, R., Galli, G., Root, D., Hahn, W., Tsherniak, A., Giannakis, M., Schreiber, S.,

Clish, C., Garraway, L. and Sellers, W. (2019). The landscape of cancer cell line metabolism. Nature Medicine, 25(5), pp.850-860.

117. Triplett, T., Garrison, K., Marshall, N., Donkor, M., Blazeck, J., Lamb, C., Qerqez, A., Dekker, J., Tanno, Y., Lu, W., Karamitros, C., Ford, K., Tan, B., Zhang, X., McGovern, K., Coma, S., Kumada, Y., Yamany, M., Sentandreu, E., Fromm, G., Tiziani, S., Schreiber, T., Manfredi, M., Ehrlich, L., Stone, E. and Georgiou, G. (2018). Reversal of indoleamine 2,3-dioxygenase–mediated cancer immune suppression by systemic kynurenine depletion with a therapeutic enzyme. Nature Biotechnology, 36(8), pp.758-764.

118. Lu, W., Su, X., Klein, M., Lewis, I., Fiehn, O. and Rabinowitz, J. (2017). Metabolite Measurement: Pitfalls to Avoid and Practices to Follow. Annual Review of Biochemistry, 86(1), pp.277-304.

119. Zukunft, S., Prehn, C., Röhring, C., Möller, G., Hrabě de Angelis, M., Adamski, J. and Tokarz, J. (2017). High-throughput extraction and quantification method for targeted metabolomics in murine tissues. Metabolomics, 14(1).

120. Duncan, K., Fyrestam, J. and Lanekoff, I. (2019). Advances in mass spectrometry based single-cell metabolomics. The Analyst, 144(3), pp.782-793.

121. Elsevier.com. (2019). Single-Cell Omics - 1st Edition. [online] Available at: https://www.elsevier.com/books/single-cell-omics/barh/978-0-12-817532-3 [Accessed 4 Aug. 2019].

122. Duncan, K., Fyrestam, J. and Lanekoff, I. (2019). Advances in mass spectrometry based single-cell metabolomics. The Analyst, 144(3), pp.782-793.

123. Wang, Z., Cao, S., Morris, J., Ahn, J., Liu, R., Tyekucheva, S., Gao, F., Li, B., Lu, W., Tang, X., Wistuba, I., Bowden, M., Mucci, L., Loda, M., Parmigiani, G., Holmes, C. and Wang, W. (2018). Transcriptome Deconvolution of Heterogeneous Tumour Samples with Immune Infiltration. iScience, 9, pp.451-460.

124. Chen, B., Khodadoust, M., Liu, C., Newman, A. and Alizadeh, A. (2018). Profiling Tumour Infiltrating Immune Cells with CIBERSORT. Methods in Molecular Biology, pp.243-259.

125. Le Van, T., van Leeuwen, M., Carolina Fierro, A., De Maeyer, D., Van den Eynden, J., Verbeke, L., De Raedt, L., Marchal, K. and Nijssen, S. (2016). Simultaneous discovery of cancer subtypes and subtype features by molecular data integration. Bioinformatics, 32(17), pp.i445-i454.

126. Gandhi, G. and Srivastava, R. (2014). Review Paper: A Comparative Study on Partitioning Techniques of Clustering Algorithms. International Journal of Computer Applications, 87(9), pp.10-13.

127. Hartigan, J. (2001). Statistical Clustering. International Encyclopedia of the Social & Behavioral Sciences, pp.15014-15019.

128. Risch, T., Canli, T., Khokhar, A., Yang, J., Munagala, K., Silberstein, A., Chrysanthis, P., Pitoura, E., Ganti, V., Deshpande, A., Do, H., Caverlee, J., Li, N., Gruenwald, L., Popa, L., Kansal, A., Zhao, F., Poulovassilis, A., Herzog, M., Song, I., Han, J., Abadi, D., Hansen, D., O'Keefe, C., Johnson, T., Gupta, A., Batini, C., Sattler, K., Scannapieco, M., Domingo-

Ferrer, J., Zhang, R., Kemme, B., Zhang, Q., Lin, X., Bouganim, L., Gibbons, P., Golab, L., Ahmad, Y., Çetintemel, U., Gupta, A., Prabhakar, S., Cheng, R., Hinterberger, H., Song, I., Golfarelli, M., Eder, J., Wiggisser, K., Vassiliadis, P., Blanco, C., Fernández-Medina, E., Trujillo, J., Piattini, M., Murphy, S., Song, I., Chen, E., Li, C., Li, X., Mattoso, M., Gyssens, M., Mylopoulos, J., Madden, S., Goda, K., Helland, P., Amza, C., Bertossi, L., Hainaut, J., Henrard, J., Englebert, V., Roland, D., Hick, J., Ferrari, E., Panda, B., Gehrke, J., Riedewald, M., Berndtsson, M., Mellin, J., Chaudhuri, S., Narasayya, V., Weikum, G., Bruno, N., Chaudhuri, S., Weikum, G., Chaudhuri, S., Weikum, G., Gupta, A., Karvounarakis, G., Gehrke, J., Gehrke, J., Gokhale, A., Gokhale, A., Cohen, E., Lin, T., Dobra, A., Dobra, A., Roscoe, T., Lin, T., Goda, K., Kaushik, R., Chang, K., Moro, M., Tsotras, V., Keim, D., Bak, P., Schäfer, M., Ester, M., Borgida, A., Batini, C., Maurino, A., Mitra, P., Batini, C., Sirangelo, C., Moore, R., Janée, G., Floriani, L., Magillo, P., Srinivasan, V., Yang, S., Fox, E., Sion, R., Carminati, B., Pedersen, T., Shen, H., Ding, C., Clementini, E., Goda, K., Blackwell, A., Costabile, M., Wada, K., Domingo-Ferrer, J., Järvelin, K., Kekäläinen, J., Dustdar, S., Platzer, C., Krämer, B., Garofalakis, M., Ahn, G., Boncz, P., Goda, K., Sheelagh, C., Risch, T., Weske, M., Garofalakis, M., Tan, K., Tan, K., Elnikety, S., Tok, W., Galuba, W., Girdzijauskas, S., Sattler, K., Bressan, S., Sattler, K., Tan, K., Kalnis, P., Tok, W., Amati, G., Munson, E., Zhao, Y., Karypis, G., Tompa, F., Plachouras, V., He, B., Plachouras, V., V. Munson, E., Caverlee, J., Mitra, P., Laarsgard, M., Cook, D. and Matera, M. (2009). Density-based Clustering. Encyclopedia of Database Systems, pp.795-799.

129. de Souto, M., Costa, I., de Araujo, D., Ludermir, T. and Schliep, A. (2008). Clustering cancer gene expression data: a comparative study. BMC Bioinformatics, 9(1).

130. Raj-Kumar, P., Liu, J., Hooke, J., Kovatich, A., Kvecher, L., Shriver, C. and Hu, H. (2019). PCA-PAM50 improves consistency between breast cancer intrinsic and clinical subtyping reclassifying a subset of luminal A tumours as luminal B. Scientific Reports, 9(1).

131. Zeng, Z., Vo, A., Mao, C., Clare, S., Khan, S. and Luo, Y. (2019). Cancer classification and pathway discovery using non-negative matrix factorization. Journal of Biomedical Informatics, 96, p.103247.

132. Ncss-wpengine.netdna-ssl.com. (2019). [online] Available at: https://ncss-wpengine.netdna-ssl.com/wp content/themes/ncss/pdf/Procedures/NCSS/Clustered_Heat_Maps-Double_Dendrograms.pdf [Accessed 14 Aug. 2019].

133. Datanovia. (2019). Clustering Distance Measures - Datanovia. [online] Available at: https://www.datanovia.com/en/lessons/clustering-distance-measures/ [Accessed 14 Aug. 2019].

134. Jaskowiak, P., Campello, R. and Costa, I. (2014). On the selection of appropriate distances for gene expression data clustering. BMC Bioinformatics, 15(S2).

135. Muthukalathi, S., Ramanujam, R. and Thalamuthu, A. (2014). Consensus Clustering for Microarray Gene Expression Data. Bonfring International Journal of Data Mining, 4(4), pp.26-33.

136. Howlett, A., Petersen, O., Bissell, M. and Steeg, P. (1994). A Novel Function for the nm23-H1 Gene: Overexpression in Human Breast Carcinoma Cells Leads to the Formation of Basement Membrane and Growth Arrest. JNCI Journal of the National Cancer Institute, 86(24), pp.1838-1844.

137. Li, Z., Wang, Z., Tang, Y., Lu, X., Chen, J., Dong, Y., Wu, B., Wang, C., Yang, L., Guo, Z., Xue, M., Lu, S., Wei, W. and Shi, Q. (2019). Liquid biopsy-based single-cell metabolic phenotyping of lung cancer patients for informative diagnostics. Nature Communications, 10(1).

138. Vazquez, F., Lim, J., Chim, H., Bhalla, K., Girnun, G., Pierce, K., Clish, C., Granter, S., Widlund, H., Spiegelman, B. and Puigserver, P. (2013). PGC1α Expression Defines a Subset of Human Melanoma Tumours with Increased Mitochondrial Capacity and Resistance to Oxidative Stress. Cancer Cell, 23(3), pp.287-301.

139. Breast Cancer Research Foundation. (2019). *Breast Cancer Statistics And Resources*. [online] Available at: https://www.bcrf.org/breast-cancer-statistics-and-resources [Accessed 9 Nov. 2019].

140. Perou, C., Sørlie, T., Eisen, M., van de Rijn, M., Jeffrey, S., Rees, C., Pollack, J., Ross, D., Johnsen, H., Akslen, L., Fluge, Ø., Pergamenschikov, A., Williams, C., Zhu, S., Lønning, P., Børresen-Dale, A., Brown, P. and Botstein, D. (2000). Molecular portraits of human breast tumours. Nature, 406(6797), pp.747-752.

141. Milioli, H., Tishchenko, I., Riveros, C., Berretta, R. and Moscato, P. (2017). Basal-like breast cancer: molecular profiles, clinical features and survival outcomes. *BMC Medical Genomics*, 10(1).

142. Netanely, D., Avraham, A., Ben-Baruch, A., Evron, E. and Shamir, R. (2016). Expression and methylation patterns partition luminal-A breast tumours into distinct prognostic subgroups. *Breast Cancer Research*, 18(1).

143. Sønderstrup, I., Jensen, M., Ejlertsen, B., Eriksen, J., Gerdes, A., Kruse, T., Larsen, M., Thomassen, M. and Lænkholm, A. (2019). Subtypes in BRCA-mutated breast cancer. *Human Pathology*, 84, pp.192-201.

144. Walsh, A., Cook, R., Manning, H., Hicks, D., Lafontant, A., Arteaga, C. and Skala, M. (2013). Optical Metabolic Imaging Identifies Glycolytic Levels, Subtypes, and Early-Treatment Response in Breast Cancer. *Cancer Research*, 73(20), pp.6164-6174.

145. Dong, T., Kang, X., Liu, Z., Zhao, S., Ma, W., Xuan, Q., Liu, H., Wang, Z. and Zhang, Q. (2015). Altered glycometabolism affects both clinical features and prognosis of triple-negative and neoadjuvant chemotherapy-treated breast cancer. *Tumour Biology*, 37(6), pp.8159-8168.

146. More, T., RoyChoudhury, S., Christie, J., Taunk, K., Mane, A., Santra, M., Chaudhury, K. and Rapole, S. (2017). Metabolomic alterations in invasive ductal carcinoma of breast: A comprehensive metabolomic study using tissue and serum samples. *Oncotarget*, 9(2).

147. Ruckhäberle, E., Rody, A., Engels, K., Gaetje, R., von Minckwitz, G., Schiffmann, S., Grösch, S., Geisslinger, G., Holtrich, U., Karn, T. and Kaufmann, M. (2007). Microarray analysis of altered sphingolipid metabolism reveals prognostic significance of sphingosine kinase 1 in breast cancer. *Breast Cancer Research and Treatment*, 112(1), pp.41-52.

148. Choi, J., Kim, D., Jung, W. and Koo, J. (2013). Metabolic interaction between cancer cells and stromal cells according to breast cancer molecular subtype. *Breast Cancer Research*, 15(5).

149. Kim, S., Kim, D., Jung, W. and Koo, J. (2013). Expression of glutamine metabolism-related proteins according to molecular subtype of breast cancer. *Endocrine-Related Cancer*, 20(3), pp.339-348.

150. Budczies, J., Brockmöller, S., Müller, B., Barupal, D., Richter-Ehrenstein, C., Kleine-Tebbe, A., Griffin, J., Orešič, M., Dietel, M., Denkert, C. and Fiehn, O. (2013). Comparative metabolomics of estrogen receptor positive and estrogen receptor negative breast cancer: alterations in glutamine and beta-alanine metabolism. *Journal of Proteomics*, 94, pp.279-288.

151. Hilvo, M., Denkert, C., Lehtinen, L., Muller, B., Brockmoller, S., Seppanen-Laakso, T., Budczies, J., Bucher, E., Yetukuri, L., Castillo, S., Berg, E., Nygren, H., Sysi-Aho, M., Griffin, J., Fiehn, O., Loibl, S., Richter-Ehrenstein, C., Radke, C., Hyotylainen, T., Kallioniemi, O., Iljin, K. and Oresic, M. (2011). Novel Theranostic Opportunities Offered

by Characterization of Altered Membrane Lipid Metabolism in Breast Cancer Progression. *Cancer Research*, 71(9), pp.3236-3245.

152. Louie, S., Roberts, L., Mulvihill, M., Luo, K. and Nomura, D. (2013). Cancer cells incorporate and remodel exogenous palmitate into structural and oncogenic signaling lipids. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1831(10), pp.1566-1572.

153. Kim, S., Lee, Y. and Koo, J. (2015). Differential Expression of Lipid Metabolism-Related Proteins in Different Breast Cancer Subtypes. *PLOS ONE*, 10(3), p.e0119473.

154. Cappelletti, V., Iorio, E., Miodini, P., Silvestri, M., Dugo, M. and Daidone, M. (2017). Metabolic Footprints and Molecular Subtypes in Breast Cancer. *Disease Markers*, 2017, pp.1-19.

155. Hilvo, M., Denkert, C., Lehtinen, L., Muller, B., Brockmoller, S., Seppanen-Laakso, T., Budczies, J., Bucher, E., Yetukuri, L., Castillo, S., Berg, E., Nygren, H., Sysi-Aho, M., Griffin, J., Fiehn, O., Loibl, S., Richter-Ehrenstein, C., Radke, C., Hyotylainen, T., Kallioniemi, O., Iljin, K. and Oresic, M. (2011). Novel Theranostic Opportunities Offered by Characterization of Altered Membrane Lipid Metabolism in Breast Cancer Progression. *Cancer Research*, 71(9), pp.3236-3245.

156. Haukaas, T., Euceda, L., Giskeødegård, G., Lamichhane, S., Krohn, M., Jernström, S., Aure, M., Lingjærde, O., Schlichting, E., Garred, Ø., Due, E., Mills, G., Sahlberg, K., Børresen-Dale, A. and Bathen, T. (2016). Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes. *Cancer & Metabolism*, 4(1).

157. Vander Heiden, M. & DeBerardinis, R. Understanding the Intersections between Metabolism and Cancer Biology. Cell 168, 657-669 (2017).

158. Rhodes, D. et al. Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. Proceedings of the National Academy of Sciences 101, 9309-9314 (2004).

159. Cancer.Net. (2019). Pancreatic Cancer - Statistics. [online] Available at: https://www.cancer.net/cancer-types/pancreatic-cancer/statistics [Accessed 24 Oct. 2019].

160. Hirshberg Foundation for Pancreatic Cancer Research. (2019). Prognosis - Hirshberg Foundation for Pancreatic Cancer Research. [online] Available at: http://pancreatic.org/pancreatic-cancer/about-the-pancreas/prognosis/ [Accessed 24 Oct. 2019].

161. Rawla, P., Sunkara, T. and Gaduputi, V. (2019). Epidemiology of Pancreatic Cancer: Global Trends, Etiology and Risk Factors. World Journal of Oncology, 10(1), pp.10-27.

162. Kaur, S., Baine, M., Jain, M., Sasson, A. and Batra, S. (2012). Early diagnosis of pancreatic cancer: challenges and new developments. Biomarkers in Medicine, 6(5), pp.597-612.

163. Torre, M., Nigri, G., Conte, A., Mazzuca, F., Tierno, S., Salaj, A., Marchetti, P., Ziparo, V. and Ramacciato, G. (2014). Is a Preoperative Assessment of the Early Recurrence of Pancreatic Cancer Possible after Complete Surgical Resection?. Gut and Liver, 8(1), pp.102-108.

164. Kimura, K., Amano, R., Nakata, B., Yamazoe, S., Hirata, K., Murata, A., Miura, K., Nishio, K., Hirakawa, T., Ohira, M. and Hirakawa, K. (2014). Clinical and pathological features of five-year survivors after pancreatectomy for pancreatic adenocarcinoma. World Journal of Surgical Oncology, 12(1), p.360.

165. Garrido-Laguna, I., Uson, M., Rajeshkumar, N., Tan, A., de Oliveira, E., Karikari, C., Villaroel, M., Salomon, A., Taylor, G., Sharma, R., Hruban, R., Maitra, A., Laheru, D., Rubio-Viqueira, B., Jimeno, A. and Hidalgo, M. (2011). Tumour Engraftment in Nude Mice and Enrichment in Stroma- Related Gene Pathways Predict Poor Survival and Resistance to Gemcitabine in Patients with Pancreatic Cancer. Clinical Cancer Research, 17(17), pp.5793-5800.

166. Weizman, N., Krelin, Y., Shabtay-Orbach, A., Amit, M., Binenbaum, Y., Wong, R. and Gil, Z. (2013). Macrophages mediate gemcitabine resistance of pancreatic adenocarcinoma by upregulating cytidine deaminase. Oncogene, 33(29), pp.3812-3819.

167. Zheng, X., Carstens, J., Kim, J., Scheible, M., Kaye, J., Sugimoto, H., Wu, C., LeBleu, V. and Kalluri, R. (2015). Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. Nature, 527(7579), pp.525-530.

168. Liu, W., Le, A., Hancock, C., Lane, A., Dang, C., Fan, T. and Phang, J. (2012). Reprogramming of proline and glutamine metabolism contributes to the proliferative and metabolic responses regulated by oncogenic transcription factor c-MYC. Proceedings of the National Academy of Sciences, 109(23), pp.8983-8988.

169. Kerr, E., Gaude, E., Turrell, F., Frezza, C. and Martins, C. (2016). Mutant Kras copy number defines metabolic reprogramming and therapeutic susceptibilities. Nature, 531(7592), pp.110-113.

170. Min, H. and Lee, H. (2018). Oncogene-Driven Metabolic Alterations in Cancer. Biomolecules & Therapeutics, 26(1), pp.45-56.

171. Jia, S., Liu, Z., Zhang, S., Liu, P., Zhang, L., Lee, S., Zhang, J., Signoretti, S., Loda, M., Roberts, T. and Zhao, J. (2016). Erratum: Corrigendum: Essential roles of PI(3)K–p110β in cell growth, metabolism and tumourigenesis. Nature, 533(7602), pp.278-278.

172. Elstrom, R., Bauer, D., Buzzai, M., Karnauskas, R., Harris, M., Plas, D., Zhuang, H., Cinalli, R., Alavi, A., Rudin, C. and Thompson, C. (2004). Akt Stimulates Aerobic Glycolysis in Cancer Cells. Cancer Research, 64(11), pp.3892-3899.

173. Rosario, S., Long, M., Affronti, H., Rowsam, A., Eng, K. and Smiraglia, D. (2018). Pan-cancer analysis of transcriptional metabolic dysregulation using The Cancer Genome Atlas. Nature Communications, 9(1).

174. Xiao, Z., Dai, Z. and Locasale, J. (2019). Metabolic landscape of the tumour microenvironment at single cell resolution. Nature Communications, 10(1).

175. Haukaas, T., Euceda, L., Giskeødegård, G., Lamichhane, S., Krohn, M., Jernström, S., Aure, M., Lingjærde, O., Schlichting, E., Garred, Ø., Due, E., Mills, G., Sahlberg, K., Børresen-Dale, A. and Bathen, T. (2016). Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes. Cancer & Metabolism, 4(1).

176. Wörmann, S., Song, L., Ai, J., Diakopoulos, K., Kurkowski, M., Görgülü, K., Ruess, D., Campbell, A., Doglioni, C., Jodrell, D., Neesse, A., Demir, I., Karpathaki, A., Barenboim, M., Hagemann, T., Rose-John, S., Sansom, O., Schmid, R., Protti, M., Lesina, M. and Algül, H. (2016). Loss of P53 Function Activates JAK2–STAT3 Signaling to Promote Pancreatic Tumour Growth, Stroma Modification, and Gemcitabine Resistance in Mice and Is Associated With Patient Survival. Gastroenterology, 151(1), pp.180-193.e12.

Appendix

Introduction

Bulk tumour contains many different malignant cells; however, it also consists of non-malignant cells known as stroma cells. A vast proportion of stroma contains fibroblasts, immune cells, endothelial cells, stellate cells, and cancer-associated fibroblasts (CAFs). One of the prominent features of pancreatic cancer is dense stroma. Previously, different studies show that cancer-associated fibroblasts help tumour cells to grow and spread, resulting in the aggressive disease phenotype. However, the role of stroma cells in pancreatic cancer is complicated and needs to be better understood. Earlier, we demonstrated different biological features of stroma cells in our identified metabolic subtypes of PDAC (**Chapter Five**). However, we found that identified metabolic subtypes are primarily dominated by the tumour cell metabolism and there was no stroma metabolic activity observed except fatty acid beta oxidation process. To investigate the stroma metabolism, we executed a supervised analysis known as orthogonal partial least square discriminant analysis (OPLS-DA) to identify particular stroma subtypes that correspond with the original metabolic subtypes. Additionally, to identify some potential therapeutic targets in identified metabolic subtypes, we used pancreatic cancer cell line data. We took the same supervised approach to categorize the cell lines into identified metabolic subtypes.

Methods

i. Data Normalisation

Data normalization performed in the same manner described in chapter four.

ii. Prediction Model

We performed OPLS-DA using SIMCA software to categorize deconvolved stroma into identified original metabolic subtypes. OPLSA-DA is a powerful statistical tool – it's a prediction and regression method.

iii. CCLE Samples and Dataset

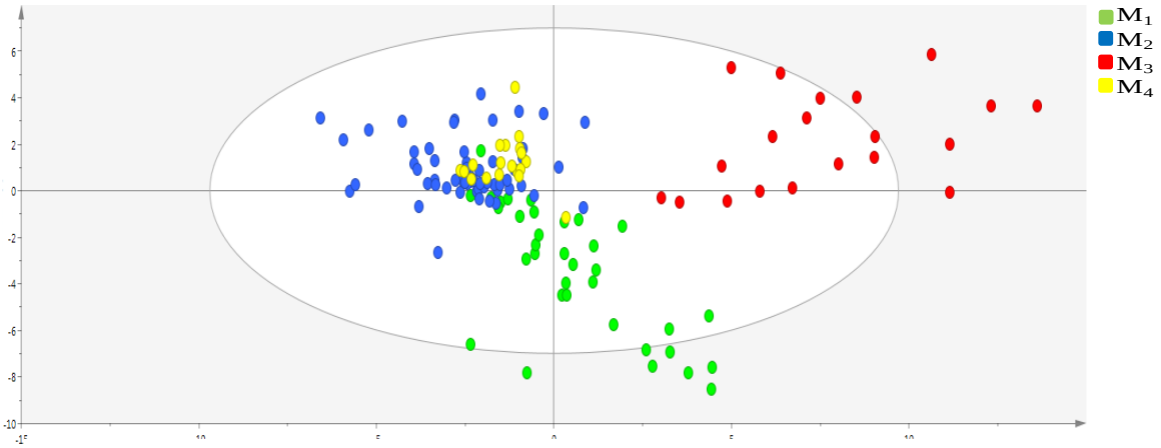
All the cell line RNA-Seq data was derived from the Cancer Cell Line Encyclopaedia (CCLE). A total of 41 pancreatic cell lines were used for the analysis.

iv. Batch Correction

The batch correction was performed on the original PDAC RNA-seq data and CCLE cell line data. Combat tools in gene patterns used for the batch correction.

Results

We performed a supervised analysis to identify stroma subtypes that relate to the original metabolic subtypes. Only metabolic genes were used for this analysis. We found two distinct stroma metabolic subtypes that highly correlate with the original subtypes. The model is statistically significant (p-value 1.11208e-35) and the given dataset fits the model well, based on R^2Y value (0.606). We identified one distinct stroma metabolic subtype that significantly correlates with the worst original metabolic subtype (M1) and another stroma metabolic subtype that correlates with the best prognostic metabolic subtype (M3) [Figure 7.1].



$R^2Y = 0.606$, $Q^2 = 0.454$, p-value = 1.11208e-35

Figure 7.1: OPLS-DA score plot of different stroma metabolic subtypes.

Further, we looked at the functional features of stroma subtypes in same way as described in **chapter four and five**. Stroma subtypes that correspond with the worst original metabolic subtype show increased lipid (p-value 5.03E-20), carbohydrate (p-value 5.03E-20) and nucleotide metabolism (7.04E-05) [Figure 6.2] compared to the best prognostic predicted stroma

subtype. Amino acid metabolism was significantly affected (p-value 5.52E-16) in this subtype; however, we were not able to determine its functional state [Figure 7.2]. Our result demonstrates that even though the stroma compartment does not exhibit any metabolic activity within the identified metabolic subtypes, it is possible that some independent stroma subtypes might exist that utilise metabolic pathways in a different manner.

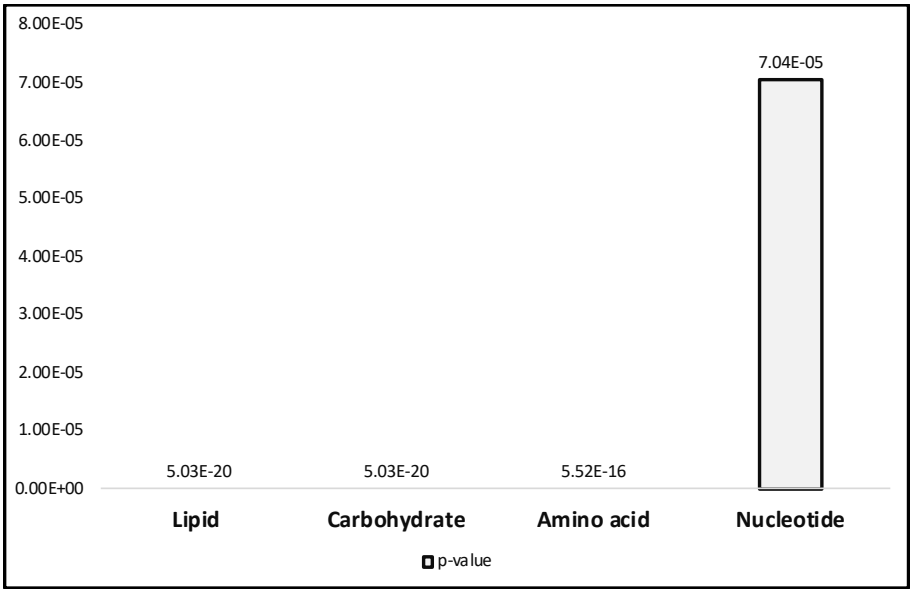


Figure 7.2: Altered metabolic pathways in stroma subtype 1.

Further, we tried to categorize the cell lines into identified metabolic subtypes. Both datasets were combined for the batch correction. After the batch correction, the dataset was normalized using the winscaling method. To categorize the cell lines, we performed the same supervised analysis. Interestingly, all cell lines clustered as one group in the middle of the OPLS-DA score plot [Figure 7.3]. Though the score plot was significant based on the p-value 0, the R^2Y (0.388) and Q^2 (0.306) values were relatively weak, indicating that the identified model may not be the best model for the given dataset.

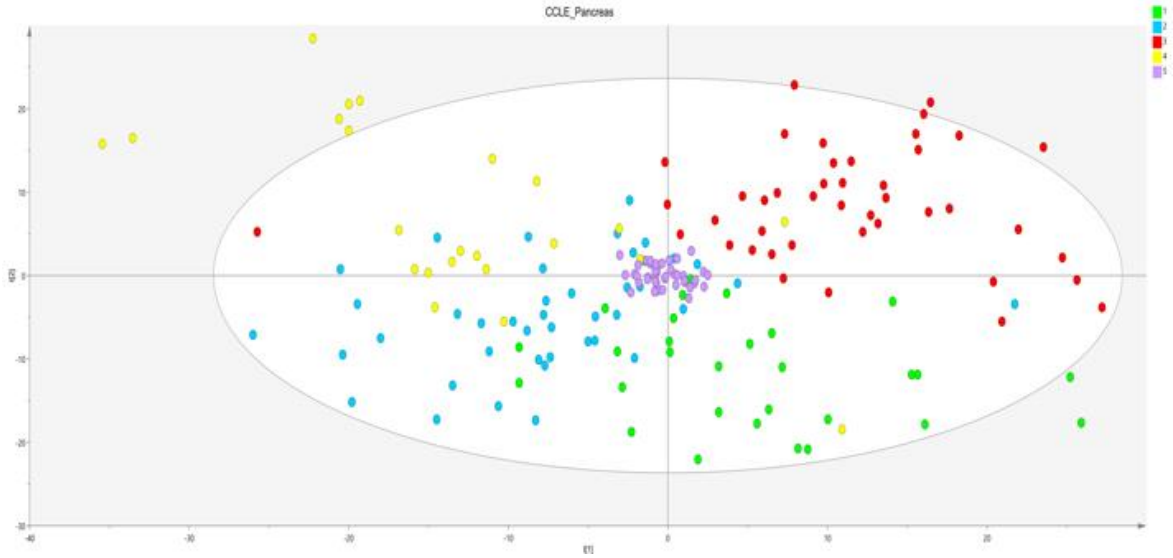


Figure 7.3: OPLS-DA score of CCLE pancreatic cell lines.

Green represents the worst prognostic group (M1), blue is the good prognostic group (M2), red is the best prognostic group (M3), yellow is the second worst prognostic group (M4), and violet represents 41 pancreatic cell lines.

Based on our CCLE analysis, the tumour cell lines cannot be categorized into our identified metabolic subtypes. We were not able to definitively determine why, but one possible reason could be the lack of stromal influence in the cell lines. To test the hypothesis, different types of analysis, like single-cell analysis, need to be done.