THE CELL MEMBRANE PROTEOME OF THE SKBR3/HER2+ CELLS AND IMPLICATIONS FOR CANCER TARGETED THERAPIES

Arba Karcini

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Biological Sciences

Iulia M. Lazar (Chair)
Jing Chen
Silke Hauf
Bryan Hsu

May 10, 2023
Blacksburg, Virginia

Keywords: Proteomics, Mass spectrometry, Breast cancer, Signaling pathways, Drug targets, Cancer hallmarks, Phosphoproteomics, Proteogenomics

The cell membrane proteome of the SKBR3/HER2+ cells and implications for cancer targeted therapies

Arba Karcini

ABSTRACT

Breast cancer is the second most common type of cancer among women in the US and the second leading cause of cancer death. HER2+ breast cancers represent ~20% of all cancer types, are highly invasive, and can be treated by using targeted therapies against the HER2 receptor. However, these therapies are challenged by the development of drug resistance, often induced by the presence of mutations in the cell-membrane proteins and receptors and/or by alternative signaling pathways that cross-talk with- or transactivate HER2+ triggered signaling. This study was aimed at investigating the cell membrane proteome of SKBR3 cells, representative of HER2+ breast cancers, and the signaling landscape and cellular responses elicited by the cell membrane receptors when the cells are stimulated with either growth factors or therapeutic drugs. It was hypothesized that the identification of a broad range of cell membrane proteins with roles in cancer progression and signaling crosstalk will lead to a more comprehensive understanding of the biological processes that sustain the proliferation of cancer cells, and will guide the selection of more efficient drug targets. The project was conceptualized in three stages: (1) profiling the cell membrane proteins of SKBR3 cells, (2) determining the functional role of the detected cell membrane proteins in the context of cancer hallmarks and exploring their mutational profile, and (3) analyzing the cellular events that occur in response to treatment with a single therapeutic agent or a combination of drugs. Mass spectrometry technologies were used for performing proteomic and phosphoproteomic profiling of SKBR3 cells, detecting changes in the abundance of the detected proteins, and identifying the presence of mutations in the cell membrane proteins. Orthogonal enrichment methods were developed for profiling the low-abundance cell membrane proteins, for generating a rich landscape of cell membrane receptors with various functional roles and relevance to the cancer hallmarks, and for enabling the detection of potentially new drivers of aberrant proliferation. The analysis of serum-starved, stimulated (with growth factors), or inhibited (with kinase inhibitors) cells revealed alternative protein players and crosstalk activities that determine the fate of cells, and that may fuel the development of resistance to treatment with drugs. The proteome profiles that were generated in

this project expand the opportunities for targeting cancer-relevant processes beyond proliferation, which is commonly attempted, broadening the landscape to also include apoptosis, invasion, and metastasis. Altogether, the findings that emerged from this work will lay the ground for future studies that aim at developing more complex and effective targeted cancer treatment approaches.

The cell membrane proteome of the SKBR3/HER2+ cells and implications for cancer targeted therapies

Arba Karcini

GENERAL AUDIENCE ABSTRACT

Breast cancer is one of the most common cancers among women in the US and the second major contributor to cancer-related deaths. Several therapies that have been developed for the treatment of cancer target the HER2 receptor, which is overexpressed in ~20% of breast cancers and results in a highly invasive cancer phenotype. However, most patients receiving these therapies observe cancer reoccurrence within a year due to the development of resistance to the therapeutic drug. The current challenge stands in identifying novel protein targets, and in developing new therapies that can be used in combination with the existing approaches to eradicate cancer. Research has indicated that proteins located at the cell membrane play crucial roles in cancer progression and invasion due to their involvement in cell response to stimuli and in initiating signaling cascades within the cell. Knowledge about the cell membrane proteins of HER2+ breast cancer cells is limited due to the challenges associated with their isolation. Therefore, this project was aimed at profiling the cell membrane proteins of HER2+ breast cancer cells, and their intra-cellular signaling activity, to provide insights into the behavior of these cells and to support the identification of potentially novel drug targets. The three objectives of the work were to (1) isolate the cell membrane proteins through various approaches using cell culture conditions that would encourage or discourage cancer cell growth, (2) identify the cancer-relevant signaling pathways and processes represented by the detected cell membrane proteins, and (3) investigate the behavior of cancer cells when treated with drugs. To approach these objectives, a powerful analytical technology, called mass spectrometry, was utilized. Mass spectrometry can accurately and simultaneously detect the presence of the proteins in a biological sample. Our study identified cell membrane proteins that are involved in cancer progression through various signaling pathways, and how these proteins interact with each other to drive the behavior of cells. The study also provided insights into how cancer cells respond when they are treated with various drugs, uncovering to the scientific community a variety of proteins with potential therapeutic value. Lastly, this study sheds light on the complex biology of

breast cancer and highlights the importance of continued research to develop more effective treatments.

DEDICATION

Dedicated to my mom for all her sacrifices to provide me the best opportunities in life!

Kushtuar mamit per te gjitha sakrificat e saj per te me siguruar mundesite me te mira ne jete!

ACKNOWLEDGMENTS

I would like to thank Dr. Lazar for her support, advice, and mentoring during this project. I am grateful to her for giving me the opportunity to join her research group! I have grown and learned a lot during my time here, and I am happy to continue a career in research thanks to the inspiration I received during this project! Dr Lazar's wide knowledge of science and engineering has always motivated me to stay a lifelong learner! Thank you!

I would also like to thank my committee members for their time, prompt availability, continued support, and valuable perspectives during this journey. I am grateful for their contribution to my academic advancement!

I am thankful to my family and friends for being by my side and offering much love and strength during this journey!

I want to conclude by expressing deep gratitude to everyone that has contributed to my professional and personal growth!

Table of Contents

Chapter 1. Introduction……1

1.1 Breast cancer……1
1.1.1. Hallmarks of cancer……..1
1.1.2. Molecular subtypes and the associated protein markers……2
1.1.3. HER2+ breast cancer and HER2 protein marker……..4
1.1.4. Mutations……..5
1.1.5. Targeted therapies……6

1.2 Cell membrane proteome……8
1.2.1. Established cell surface receptors……9
1.2.2. Other cell surface proteins……..13

1.3 Signaling pathways……13
1.3.1. MAPK signaling……14
1.3.2. PI3K/Akt signaling……15
1.3.3. Crosstalk and transactivation……..16
1.3.4. Cell cycle in cancer cells……17
1.3.5. Energy production in cancer cells……19
1.3.6. Signal transduction by phosphorylation……..19

1.4 Mass spectrometry: a powerful tool for analyzing the proteome……21
1.4.1. Instrumentation and data acquisition……..21
1.4.2. Chromatography……24
1.4.3. Targeted and untargeted proteomics……25
1.4.4. Proteome Discoverer software……26
1.4.5. Quantification……...……..27

1.5 References……..30

Chapter 2. Purpose of the study……37

2.1 Hypothesis……..37

2.2 Research questions……37

2.3 Objectives……38

2.4 Studies performed towards the objectives……..39

2.5 Significance and impact of the study……40

2.6 References……..41

Chapter 3. Methods……42

3.1 Reagents and materials……42

3.2 Cell culture growing conditions……43

3.3 Cell surface protein labeling and processing……..43

3.4 Cell culture treatments with drugs……..45

3.5 FACS analysis……46

3.6 Cell fractionation: nuclear and cytoplasmic……47

3.7 Phosphopeptide enrichment……47

3.8 Data-dependent acquisition mass spectrometry (DDA-MS) analysis……..48
3.8.1. DDA-MS analysis for cell surface and whole cell extracts……48
3.8.2. DDA-MS analysis for enriched phosphorylated peptides……..49

3.9 Targeted mass spectrometry: Parallel reaction monitoring……49

3.10 Mass spectrometry data processing……..50
3.10.1. Data processing for cell surface, nuclear/cytoplasmic, and phosphorylated proteins…..50
3.10.2. Data processing for mutated proteins……51

3.11 Quantitation and statistical analysis……51
3.11.1. Spectral counting-based quantitation……52
3.11.2. Peak area-based quantitation……52

3.12 In-house databases……..53

3.13 Data analysis, interpretation, and visualization……54

3.14 Validation method: immunofluorescence (IF) microscopy……..54

3.15 Validation method: western blotting……55

3.16 References……58

Chapter 4. Cell membrane proteome……..……..62

4.1 Introduction……63

4.2 Methods……..64

4.3 Results……65
4.3.1. Annotation of the cell surface proteins……65
4.3.2. Efficiency of the cell surface protein enrichment……66
4.3.3. Cell surface protein reproducibility and classification……...68
4.3.4. Biological processes……72
4.3.5. Differential expression……77

4.4 Discussion……..79
4.4.1. Catalytic receptors……..79
4.4.2. G-protein coupled receptors……80
4.4.3. Other cell membrane proteins……..81
4.4.4. Potential drug targets……..83
4.4.5. Differentially expressed proteins……83

4.5 Conclusion……..84

4.6 References……..85

Chapter 5. Cancer hallmarks and mutations……91

5.1 Introduction……91

5.2 Methods……..93

5.3 Results and discussion……..93
5.3.1. Cancer hallmarks……93
5.3.2. Mutated cell surface proteins……..98

5.5 Conclusion……105

5.6 References……106

Chapter 6. Drug treatments……..110

6.1 Introduction……..110

6.2 Methods……112

6.3 Results……...112
6.3.1. Drug treatments and proteome profiling……112
6.3.2. Differentially expressed proteins in the nuclear and cytoplasmic cell fractions……...114
6.3.3. Biological interpretation of the differentially expressed proteins……..116
6.3.4. Phosphorylated peptides enrichment……121
6.3.5. Relevant phosphorylated sites to cellular events……...125

6.4 Discussion……130
6.4.1. Lapatinib/EGF treatment vs EGF stimulation of cells……..131
6.4.2. Lapatinib/ipatasertib/EGF combination treatment vs EGF stimulation of cells……135
6.4.3. Early events represented by the phosphorylated signaling landscape……..137
6.4.4. Addition of ATP to the drug treatments……138
6.4.5. Potential drug targets……139

6.5 Conclusion……140

6.6 References……142

Chapter 7. Conclusion and future work……..153

Appendix A. Validation of some of the detected GPCRs by parallel reaction monitoring (PRM)..……156

Appendix B. Validation of some of the detected up- and down- regulated CS proteins by PRM. ……...……..158

Appendix C. Immunofluorescence labeling of ATP5F1A (elevated in SF) and P2RY2 (elevated in ST) with their respective primary antibodies at 1:100 dilution to assess their intensity profiles in each cell states. ……..……..…160

Appendix D. Zoomed images of surface co-localization of ATP5F1A and UQCRC2 with the cell surface marker E-cadherin. ……..……161

Appendix E. Extended PPI networks of the detected CS proteins with roles in cancer hallmarks. ……..……162

Appendix F. Correlation of the elution times and XCorr scores of peptides from each biological replicate of the cytoplasmic and nuclear fractions of each treatment: replicate 1 (X-axis), replicate 2 (Y-axis), and replicate 3 (color bar). ……..…...163

Appendix G. Volcano plots of nuclear and cytoplasmic fractions of the other treatment comparisons not shown in **Figure 6-3**. ……...163

Appendix H. Differentially expressed proteins validated by western blotting. ……...……..164

Appendix I. Differentially expressed proteins validated by PRM. …….……...……167

Appendix J. Transactivation testing. ……...…..180

Appendix K. Protein List (UniProt description)..……182

List of Figures

Chapter 1

Figure 1-1. The hallmarks of cancer. ……2

Figure 1-2. The molecular subtypes of breast cancer and the expression level of their associated protein markers. ……..4

Figure 1-3. Timeline of preclinical research for the HER2 receptor. ……...……7

Figure 1-4. Drug target kinases. ……...……12

Figure 1-5. MAPK and PI3K pathways upon activation from EGFR. ……...…16

Figure 1-6. Phases of cell cycle and the roles of its components. ……18

Figure 1-7. The core components of a mass spectrometer. ……...…..24

Figure 1-8. Different modes of acquisition for tandem MS and their mode of operation. ……..26

Chapter 4

Figure 4-1. Cell-membrane protein isolation flowchart and database classification. ……..66

Figure 4-2. Cell membrane protein enrichment efficiency. ……67

Figure 4-3. Protein ID Venn and PSM correlation diagrams representing the complementarity and reproducibility of the three labeling methods in detecting cell-membrane proteins matched by at least 2 unique peptides (FDR <3%). ……69

Figure 4-4. Functional categorization of the detected cell-membrane proteins based on GO controlled vocabulary terms (proteins detected by at least 2 unique peptides, FDR <3%). …….70

Figure 4-5. Detectability of cell-membrane receptors and their sequences. ……...72

Figure 4-6. Bar charts of selected functional categories and pathways associated with the detected cell-membrane proteins. ……73

Figure 4-7. Protein–protein interaction networks of selected cell-membrane protein categories. ……74

Figure 4-8. Cancer markers and drug targets detected in the SKBR3 cell-membrane proteome. ……76

Figure 4-9. Proteins with changed abundance in the cell-membrane proteome. ……78

Chapter 5

Figure 5-1. Cancer-supportive processes and drug targeting of 138 cell-membrane hallmark proteins. ……96

Figure 5-2. PPI networks constructed from 138 cell-membrane proteins matched to ten cancer hallmarks. ……..97

Figure 5-3. XCorr vs charge state distribution of mutated peptides detected with at least two tandem mass spectra. ……99

Figure 5-4. Detection of mutated cell surface proteins and peptides. ……..99

Figure 5-5. Relative frequency of the mutated amino acids and categorization by BLOSUM62 matrix. ……101

Figure 5-6. Landscape of 66 detected cell membrane proteins that carry mutations with pathogenic or unknown role categorized by their function. ……...102

Figure 5-7. Biological processes and cancer hallmarks affected by the mutated proteins. …...103

Chapter 6

Figure 6-1. Overview of the drug treatment protocol and proteome profiling results based on spectral counting data. ……113

Figure 6-2. Assessment of data reproducibility. ……...…114

Figure 6-3. Results of the nuclear and cytoplasmic fraction enrichment process and differential expression analysis. ……115

Figure 6-4. Differential expression analysis based on spectral counts and peak areas comparing lapatinib-treated against EGF-treated cells. ……117

Figure 6-5. Differential expression analysis based on spectral counts and peak areas comparing lapatinib/ipatasertib-treated against EGF-treated cells. ……..118

Figure 6-6. Bubble charts representing biological processes descriptive of cancer hallmarks, as represented by the differentially expressed proteins with adjusted p-value ≤0.1. …….120

Figure 6-7. Differentially expressed proteins in the nuclear and cytoplasmic cell fractions aligned with approved (A) and investigational (B) drug targets from the DrugBank database. ……..121

Figure 6-8. Overview of phosphopeptide MS analysis. ……...……..123

Figure 6-9. Detected phosphosites and their corresponding proteins at the 15 min timepoint, categorized by biological processes found in PhosphoSitePlus. ……..…..124

Figure 6-10. Signaling network of phosphorylated proteins created by SIGNOR3.0 with data from Table 6-2. ……...127

Figure 6-11. Detected phosphorylated protein substrates with their specific sites (green) matched to their respective kinases (gray) by SIGNOR3.0. ……128

Figure 6-12. KEA3 analysis of phosphoproteins that displayed a difference of 5 or more in phosphopeptide PSMs between two different drug treatments [15 min exposure to the drug(s)]. ……..129

Figure 6-13. KEA3 analysis of phosphoproteins that displayed a difference of 5 or more in phosphopeptide PSMs between two different drug treatments [30 min exposure to the drug(s)]. ……..130

List of Tables

Chapter 1

Table 1-1. Categories of the well-established cell surface receptors and the emerging cell-surface proteins involved in cell signaling. ……9

Table 1-2. Main pathways induced by EGFR, their protein components, and function. ……14

Table 1-3. Labelling and non-labelling techniques for relative protein quantification: function and advantages or disadvantages. ……27

Chapter 3

Table 3-1. Materials and antibodies used for immunofluorescence imaging to visualize the receptors and the differentially expressed proteins in the cell surface samples. ……..55

Table 3-2. Materials and antibodies used for western blotting to test the cross-talk activities and to validate the differentially expressed proteins in the cytoplasmic/nuclear fractions. ……56

Chapter 4

Table 4-1. Cell-surface protein identification effectiveness with and without enrichment in cell-surface proteins. ……71

Chapter 5

Table 5-1. Cancer hallmarks defined by GO biological processes and examples of cell surface proteins associated with the hallmarks. ……94

Chapter 6

Table 6-1. Tandem MS analysis of phosphoprotein cell fractions and average number of protein IDs calculated from three biological replicates. ……122

Table 6-2. Detected phosphoproteins and specific phosphosites, with their associated function and spectral counts of each phosphosite carrying peptide. ……126

CHAPTER 1. INTRODUCTION

1.1 Breast cancer

Breast cancer is the most common cancer in adolescents and young adults in the USA and the second most common type of cancer worldwide.[1] The risk factors associated with breast cancer include genetic and family history, environmental oestrogens, daily habits such as diet, alcohol consumption, smoking, and other causes related to the female reproductive system such as menarche, childbearing, breastfeeding, or menopause.[1,2] In general, cancer refers to a collection of genetic diseases that can affect any organ in the human body and can manifest through masses known as tumors, which are the result of uncontrolled growth and division of cells. The description of cancer as a genetic disease implies the direct role played by gene alterations in the development and progression of the disease.[3]

1.1.1. Hallmarks of cancer

Striving for a logical understanding of the mechanisms that control the onset of cancer, Hanahan and Weinberg have proposed eight hallmarks that represent distinct biological processes required by cells to successfully lead to cancer. The combination of these hallmarks supports different pathways of cancer development. In **Figure 1-1**, the eight hallmarks are presented as described by Hanahan and Weinberg.[4] Underlying conditions such as genome instability, mutations and inflammation can enhance the hallmark activities and increase the complexity of the disease.[4,5]

While genomic instability is the driving force that leads to genetic changes, **Figure 1-1** depicts the processes that a cell with a mutated genome can undergo. Two of the above hallmarks (deregulated cellular energetics and avoidance of immune destruction) are believed to be emerging hallmarks because they are not fully validated as acquired traits that directly lead to cell survival and proliferation.[4] However, these hallmarks are present and aid in cancer formation and progression. Likewise, there are additional factors that contribute to cancer development that include variations in epigenetic modifications, the presence of senescent cells, and the impact of the human microbiome and of the external microenvironment; all recognized as new dimensions in the cancer landscape in the latest publication by Hanahan.[6]

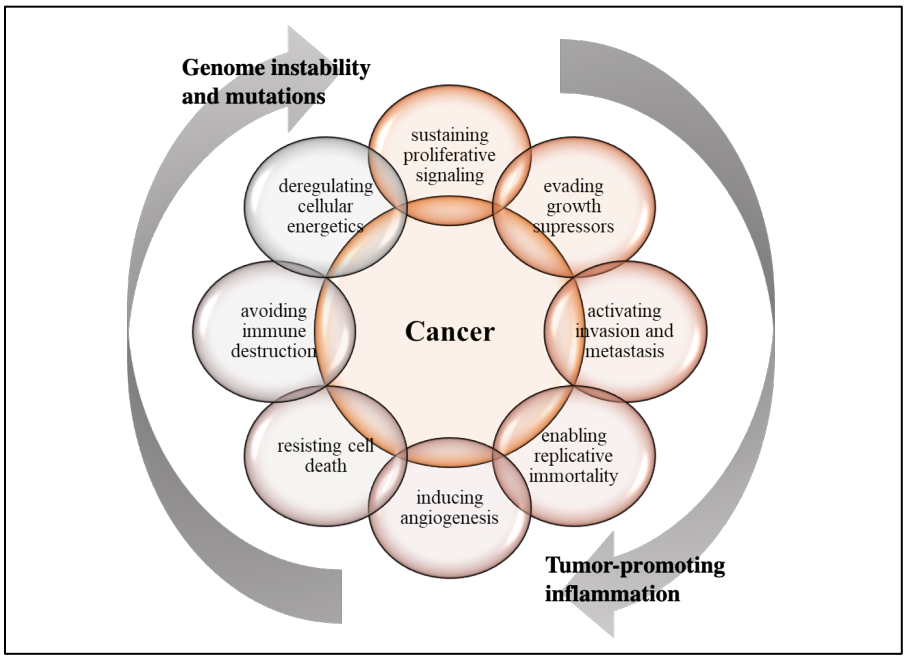


Figure 1-1: The hallmarks of cancer. Each hallmark as described by Hanahan and Weinberg[4] represents a distinct biological trait that a cell should acquire to become tumorigenic. While all these traits are believed to be involved to a certain extent in the development of cancer, a different combination or a different order of events can lead to various underlying mechanisms of cancer progression.

1.1.2. Molecular subtypes and the associated protein markers

It is these characteristics such as growth or gene expression discussed above that have given rise to different cancer type classifications by researchers and clinicians despite the complexity of the disease. These classifications are mostly based on histological or molecular characterizations. Histological assessment refers to categorizing cancer based on structure, growth, and morphology, for example, into the ductal or lobular carcinoma types in breast cancer.[7] Molecular categorization is based on the mRNA expression of genes that present distinct variability between tumors. The most well characterized breast cancer molecular subtypes are classified based on the (over)expression of certain protein cancer markers, among other clinical parameters (**Figure 1-2**): luminal (A and B) cancers that overexpress estrogen receptors (ERs) and to a lesser extent progesterone receptors (PRs), HER2+ cancers that overexpress the human epidermal growth factor receptor 2 (HER2), and basal-like cancers (known as triple negative) that do not overexpress ER, HER2 or progesterone receptors (PRs).[7,8] However, more subtypes and subgroups are also present based on different combinations of these markers and other

proteins because cancer is highly heterogenous, and clinical classifications are always changing and being redefined in the scientific and medical community as described in Weigelt et. al.[7]

Estrogen receptors regulate gene expression, cell growth and differentiation by genomic and non-genomic actions. In the nucleus, via the classical genomic mechanism, ER-α/ER-β act as transcription factors upon binding of- and activation by hormone ligands such as estrogens. Alternatively, the ERs can interact with other transcription factors that bind DNA to affect gene expression. The non-genomic mechanisms manifest themselves by cytoplasmic activities when the ERs localize to the inner side of the plasma membrane. The ERs interact with other signaling proteins in the membrane, such as IGF-I, HER2 or G-protein coupled receptors, to ultimately affect cell proliferation.[9]

The progesterone receptor (PR) is a nuclear ligand-dependent transcription factor. Binding of the progesterone hormone to the inactive PR complex results in a conformation change that enables the PR to dissociate from its multi-protein chaperone complex, homodimerize, and bind to DNA to initiate transcription. In addition to its nuclear activity, the PR was also shown to be implicated in cytoplasmic signaling pathways such as the MAPK.[10]

HER2 is a tyrosine kinase receptor of the ERBB family, as described in the next section. One characteristic of the HER2 receptor is the lack of a specific ligand. HER2 is activated by heterodimerization with one of its other ERBB family members. Overexpression of HER2 increases heterodimer activity and leads to tumor formation due to the stimulation of growth and survival of cells.[11]

Luminal type A and B breast cancers have high levels of estrogen receptor, with type B displaying lower expression levels of progesterone receptor when compared to type A. The HER2 type has high expression levels of HER2 genes and low expression of basal-related genes. On the other hand, basal-like type of cancers express high levels of keratin. They are also known as triple negative type because of their lack or low expression of ER, PR, and HER2, and are associated with BRCA1 gene mutations that lead to the most aggressive cancers.[8] In this project, we will focus on the HER2+ breast cancer type.

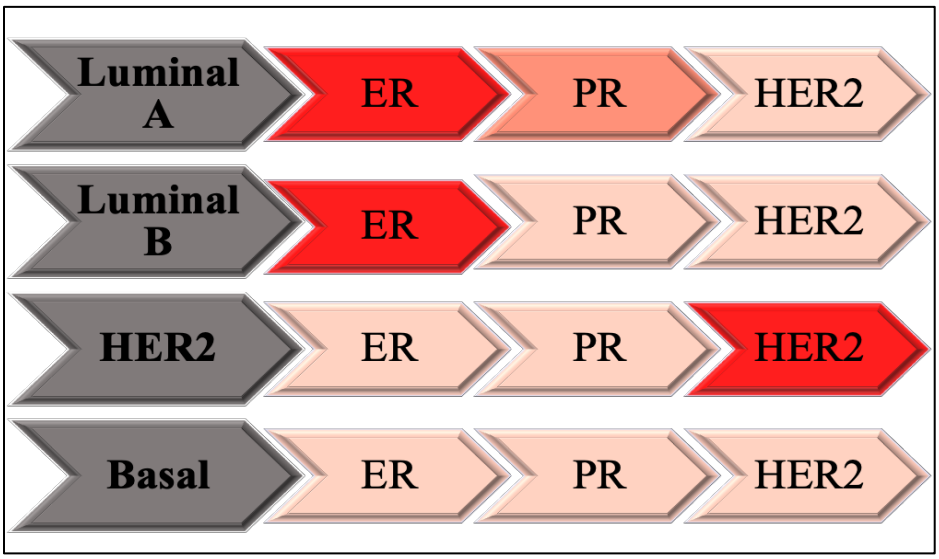


Figure 1-2: The molecular subtypes of breast cancer and the expression level of their associated protein markers. Intensity of the color increases as the expression level of each marker increases.

1.1.3. HER2+ breast cancer and HER2 protein marker

The HER2+ breast cancer type represents up to 20% of breast cancers and around 30% of metastatic breast cancers.[12,13] The HER2 receptor is part of the ERBB family of receptors composed of four members: epidermal growth factor receptor (EGFR)/ERBB1, ERBB2/HER2, ERBB3/HER3 and ERBB4/HER4, which are expressed abundantly in a variety of tissues, including epithelial, and involved in development and most importantly, cancer.[14] As mentioned above, the HER2 receptor is expressed at the cell membrane in an open conformation and no ligand is known to activate it, in contrast to the other family receptors, which upon activation by a variety of ligands (epidermal growth factors/EGF, betacellulin/BTC, herefulins) form homo- and hetero- dimers. Therefore, HER2 is the preferred dimer of the ERBB receptor family as it potentiates and prolongs the downstream signaling pathways initiated by the above ligands, and facilitates the cross-talk and transactivation between the ERBB family receptors through its involvement in signaling with the other ERBB members.[14] These mechanisms explain the oncogenic potential and overexpression of HER2 in cancer, which is estimated to be a 40-100-fold increase at the protein level, with 2 million HER2 receptors expressed on the surface of cancer cells.[15] Due to the strong catalytic signaling it elicits, many downstream pathways are activated (more on the specific pathways in the sections below) with roles in cell proliferation, survival, angiogenesis, and invasion.[12,15,16]

1.1.4. Mutations

In cancer cells, the mutational rate is elevated to confer a selective advantage in the tumor environment, and this happens largely due to the corruption of the genome maintenance and immune systems, loss of telomeric DNA, and heightened sensitivity toward mutagenic factors. Due to these effects and the emerging roles of certain mutations in tumor pathogenesis, Hanahan and Weinberg classified genome instability and mutations as an enabling characteristic of cancer.[4] Since then, the field of cancer genomic studies was dedicated to cataloging these somatic mutations that provide a selective advantage in a cell and lead to a carcinogenic phenotype (driver mutations). The advancement of next-generation sequencing and its capabilities to sequence any random DNA fragment gave rise to the ICGC/TCGA Consortium which aims to systematically categorize somatic mutations in different tumor types.[17] After 10 years of intense collaborations between 20 institutions and more than 11,000 tumors samples from individual patients, more than 2.7 million mutations on 33 different tumor types were catalogued.[18,19] The detected mutations comprise a wide variety of variations not limited to rather small types of mutations (substitution/point mutations, insertion and deletion/indel mutations), but also to the detection of DNA fusions, copy number variations, and other aberrations in DNA sequence, gene expression, epigenetics, protein expression and structures, etc. This wealth of information can be used to inform on the various functional roles and signaling pathways in a cancer cell or to establish mutation signatures.[18] Mutational signatures represent the multitude of mutational processes that lead to or occur during cancer progression and identify and group these processes by their origin.[20] The use of such signatures is hoped to lead to the development of better targeted screening and treatment strategies. The TCGA project and its publicly available datasets have changed the field in major ways by providing a rich, high quality data resource, by boosting the development of computational tools for better mutational analysis, and by expanding the molecular understanding of cancer types, classifications of tumors, and identification of therapeutic targets.[18]

In HER2+ breast cancer cells, tumorigenesis can also be driven by other factors than HER2 overexpression, including mutations in the HER2 gene that account for 3% of cases. In addition, HER2+ breast cancers display co-occurring mutations of HER2 and other ERBB receptor family members, or mutations in other proteins of downstream, compensatory pathways, such as the

case of PI3K mutations found in ~25% of this type of cancer.[21] Another 20% of patients with HER2+ breast cancer, display a truncated active form of HER2 receptors known as p95HER2, which is missing the extracellular domain, but is nevertheless active in transmitting kinase signaling to downstream pathways. Such truncation may be a result of alternative translation start sites or cleaving activity of other cell surface proteins known as matrix metalloproteases.[12] All of these can also cause the development of resistance against targeted therapies,[12,21] which are discussed below.

1.1.5. Targeted therapies

Treatment of breast cancer depends on the type and stage of the disease. For the most advanced stages, treatment includes hormone and targeted therapy, chemotherapy, radiation, surgery, or inclusion in new clinical trials that might be available. A combination of these therapies is often required for achieving promising results.[22] The most sought after and breakthrough therapies are the targeted ones. The HER2 receptor was among the first ones to be targeted due to its overexpression in breast cancer and trastuzumab was the first monoclonal antibody drug to be approved by FDA for HER2+ breast cancers.[12] Trastuzumab binds to the extracellular domain of the HER2 receptor, downregulates its cell surface expression and results in reduced downstream signaling activity. In addition, it facilitates cytotoxicity because of the immune cell recognition of trastuzumab-bound cells.[12,16] Despite the success of trastuzumab treatment alone and in combination with other drugs and chemotherapeutic agents, resistance to such therapy in most patients emerges within a year.[12] At this point, two important tasks emerge for researchers: understanding the resistance mechanisms and developing new targeted therapies. **Figure 1-3** shows a quick summary of the research progression towards the discovery of the HER2 receptor, which led to two Nobel Prize awards **(A)**, and the approval history of targeted therapies, which only boomed after the 2010s **(B)**. It is clear from **Figure 1-3**, how trastuzumab really was a breakthrough therapy used for many years before new advancements in research could lead to alternative options.

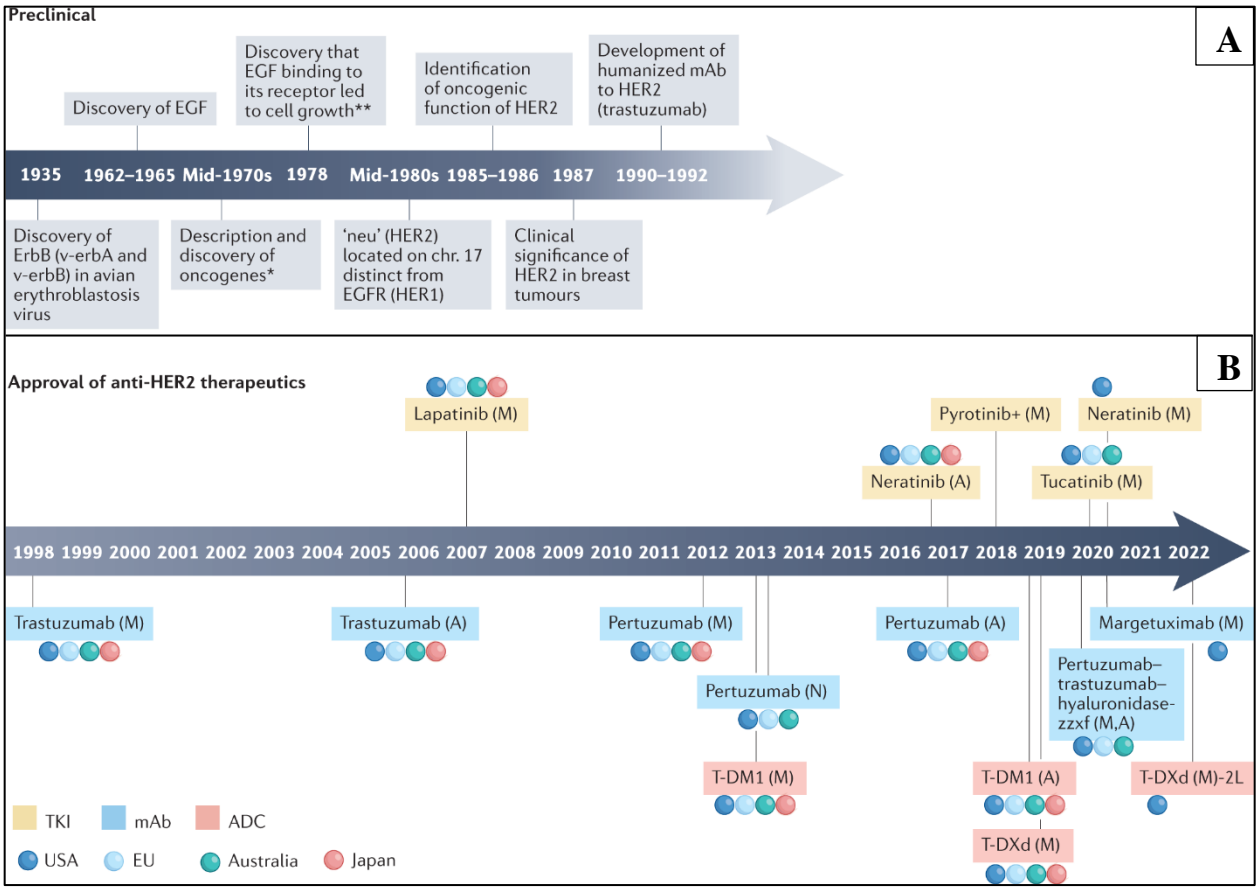


Figure 1-3: Timeline of preclinical research for the HER2 receptor. **(A)** Two Nobel Prize awards: * M. Bishop and H. Varmus (1989), ** S. Cohen and R. Levi-Montalcini (1986) for the biological discoveries. **(B)** Approval history of therapies against HER2 receptor, where the letter in the parenthesis next to the drug name indicates the cancer therapy stage it was approved for: A, adjuvant setting; M, metastatic setting; N, neoadjuvant setting. Each colored circle indicates the country/geo-political area the drug was approved, while the (+) refers to China. The color of the name of the drug indicates the type of the drug: yellow – small molecule tyrosine kinase inhibitor, blue – monoclonal antibody, red – antibody-drug conjugate. Adapted from Swain, et.al. (2022)[21] under the PMC Open Access Subset.

In this study, the HER2 drug used is lapatinib (GW-572016), a small molecule dual inhibitor reversibly binding to the cytoplasmic ATP-binding sites of EGFR and HER2, blocking tyrosine kinase phosphorylation, and inhibiting the downstream signaling pathways, PI3K and MAPK.[23,24] Additionally, lapatinib decreases HER2 ubiquitination leading to inactive HER2 receptor accumulation on the cell surface. In cell culture, lapatinib leads to G1 growth arrest and apoptosis in dosages 1-10 μM from 24-72 h as reported in Tsang, et. al.[24] Combination of lapatinib with other drugs is also common, specifically with trastuzumab, however, the development of resistance is a recurring issue as is with the trastuzumab antibody.[23,24] However,

the mechanisms of resistance in the case of lapatinib differ from those that arise from trastuzumab due to the distinct mode of action of each drug, including cross talk with other receptors such as hormone receptors (estrogen and progesterone), hyperactivation of downstream signaling pathways such as PI3K, and mutations.[24]

In roughly 50% of HER2+ breast cancers the PI3K pathway is altered, and its hyperactivation, as mentioned above, leads to drug resistance against targeted therapies.[25] Therefore, the other drug used in this study is ipatasertib (GDC-0068), a highly selective, ATP-competitive small molecule inhibitor of all 3 AKT isoforms, which share more than 95% sequence similarity in their ATP-binding site. Upon binding, the drug leads to increased phosphorylation of AKT, but decreased phosphorylation and activity of the downstream targets of PI3K pathway, and cell death.[26,27] In cell culture, ipatasertib leads to G1 growth arrest and apoptosis in dosages 1-10 μM from 15-72 h as reported in Lin, et. al.[26] Ipatasertib was used in HER2+ breast cancer therapies in combination with HER2 targeted therapy such as trastuzumab or lapatinib,[25-27] but the combination with the latter one is still in infancy and the efficacy was not yet fully evaluated.

1.2 Cell membrane proteome

Communication between cells or between the inside and outside environment of the cell is essential to a cell's life. Cells do not operate as isolated entities; therefore, they need forms of communication and signal transduction both with other cells/environment and within themselves. Mostly, this communication is realized through receptors, proteins able to receive and recognize signals (mostly from the outside of the cell) and generate a response or change in the activity of the cell. Receptors are diverse and are categorized based on how they receive and transmit signals and the type of change they induce in the cell activity. Receptors also differ based on their cellular location as internal or cell membrane.[28] Here, we focus on the cell membrane proteins, also referenced interchangeably in this work as cell surface (CS) proteins, as they are the main area of interest in this project and possess special interest to drug discovery efforts. What makes cell surface receptors likable drug targets among others is the presence of binding sites and their linkages to disease processes as part of their role in cell signaling, proliferation and differentiation. Even though the pool of potential drug targets is large (estimated between 2,000-3,000), only a few hundreds have been approved so far for effective clinical targeting

applications.[29] The classification of CS receptors is challenging due to their involvement in multiple processes through various mechanisms, but based on literature reviews, a simplified categorization is presented in **Table 1-1** and a snapshot of their activity and role in cancer in the following sections.

Table 1-1: Categories of the well-established cell surface receptors and the emerging cell-surface proteins involved in cell signaling.[28,30-36]

Established Receptors	***Ion channel***	Ligand-gated, voltage-gated
	G-protein coupled	$G\alpha_s$, $G\alpha_i$/$G\alpha_o$, $G\alpha_q$, $G\alpha_{12}$/$G\alpha_{13}$
	Enzyme-linked	Tyrosine kinase, Tyrosine-kinase-associated, Receptor-like tyrosine phosphatases, serine/threonine kinases
Extracellular Matrix (ECM)	***Adhesion***	Integrin, IgSF, Cadherin, Selectin
	Matricellular	TSP1, TSP2, SPARC, Osteopontin, Tenascin-C, Tenascin-X
	MHC	HLA class I, HLA class II
	Scaffold	Adapter, docking, anchoring
	Scavenger	Class A-J
Transporters	***ATP-driven***	P-type ATPases, F-type/V-type ATPases, ATP-binding cassette
	Ion gradient driven	SLC
Secreted	***Cytokines***	Interleukins, interferon, tumor necrosis factor
	Hormones	steroid receptor family

1.2.1. Established cell surface receptors

Based on their mechanism of action, three main categories of established cell surface receptors are known: G-protein-coupled receptors (GPCRs), enzyme-linked receptors, and ion channel

receptors.[28] GPCRs, as indicated by their name, receive and transmit the signal through G-proteins modulation. The GPCRs, also known as 7TM receptors due to their seven transmembrane segments through the cell membrane, receive and transmit the signal through cytoplasmic G-proteins that experience conformational changes to their heterotrimeric structure upon activation. A G-protein contains an alpha, beta and gamma subunit, which are all bound to a GPCR in an inactive state. The alpha subunit is bound to a guanosine diphosphate (GDP), but upon ligand binding, the GDP is replaced by a GTP (guanosine triphosphate) in the alpha subunit of the G-protein and the unit is split in two parts: alpha subunit and beta-gamma subunit. Each of these subunits can interact with other plasma membrane proteins, ion channels, or second messengers to transduce a variety of signals in the cell.[37,38] However, the activation of GPCRs can happen either in the absence or presence of a ligand, or other mechanisms, indicating the functional complexity of the largest family of receptors (~800-1,000 GPCR estimates) and the many signaling pathways it plays a role in. These reasons have made GPCRs one of the most targeted groups by drugs with ~34% of all approved drugs acting on them.[37-39] GPCRs are further classified based on the G-protein that they modulate, or by homology. There are four types of G-protein: GαS, Gαi/Gα0, Gαq, and Gα12/Gα13; each of these subunits bind to other key surface membrane proteins/enzymes such as adenylyl cyclase, phospholipase C, or small monomeric GTPases (a GPCR small membrane protein that can hydrolyze GTP) to regulate different signaling mechanisms in the cell such as calcium levels.[32] By sequence homology, there are five classes of GPCRs: class/family A (rhodopsin family) receptors related to light receptors, class/family B (secretin family) receptors related to glucagon receptors, class/family C (glutamate family) receptors related to neurotransmitters, adhesion family, and frizzled family. Each family can be further divided in a phylogenetic manner.[33,37]

Enzyme linked receptors are receptors that exhibit enzymatic activity upon activation, mostly kinases, 538 of them, whose role is to transfer a phosphate group from ATP to the serine, threonine, or tyrosine residues of a protein in response to a stimulus, therefore initiating a signaling cascade.[40] Phosphorylation by kinases is the most common post-translational event that takes place in a cellular protein affecting all processes of a cell, including growth, differentiation, division, metabolism, immunity, transport, etc. The site of phosphorylation is based on the specificity of kinase and substrate domains, which then determines the specific downstream

responses.[41] All eukaryotic kinase proteins contain an evolutionary conserved ~250 amino acid catalytic domain, which is key to their activity, and other regulatory domains that account for their diversity.[42,43] Based on the categorization proposed by Manning et. al. primarily on sequence homology, domain structure and biological function, the kinome is presented in a phylogenetic tree (**Figure 1-4**) comprised of 8 main groups: CAMK (Ca^{2+}/calmodulin-dependent protein kinase class), AGC (named after the cytoplasmic serine/threonine protein kinase A, G, and C families), CK1 (Casein kinase 1 family), STE (serine/threonine kinases), TK (tyrosine kinases), TKL (tyrosine kinase like, lacking the TK-specific motifs of the TK group), CMGC (named after the initials of some members including cyclin-dependent kinases/CDKs, mitogen-activated protein/MAP kinases, glycogen synthase kinases/GSK, and CDK-like kinases), and other.[44] Kinases can also be classified by domains, where some of the most common domains include Src homology 2 and 3/SH2, SH3 domain binding to phosphotyrosines and proline-rich motifs, and PH domain binding to phospholipids,[43] both related to important signaling events in the cell, such as the MAPK and PI3K pathways discussed in the following sections.

Protein kinases being one of the largest family of proteins, key to many signaling events in the cell, and prone to oncogenic transformation due to mutations and overexpression, such as the case with HER2, are one of the most sought-after targets in cancer therapy, with 68 approved kinase inhibitor drugs in the market.[42] A visual representation of the protein kinase groups and some of the approved drug targets is shown in **Figure 1-4**.

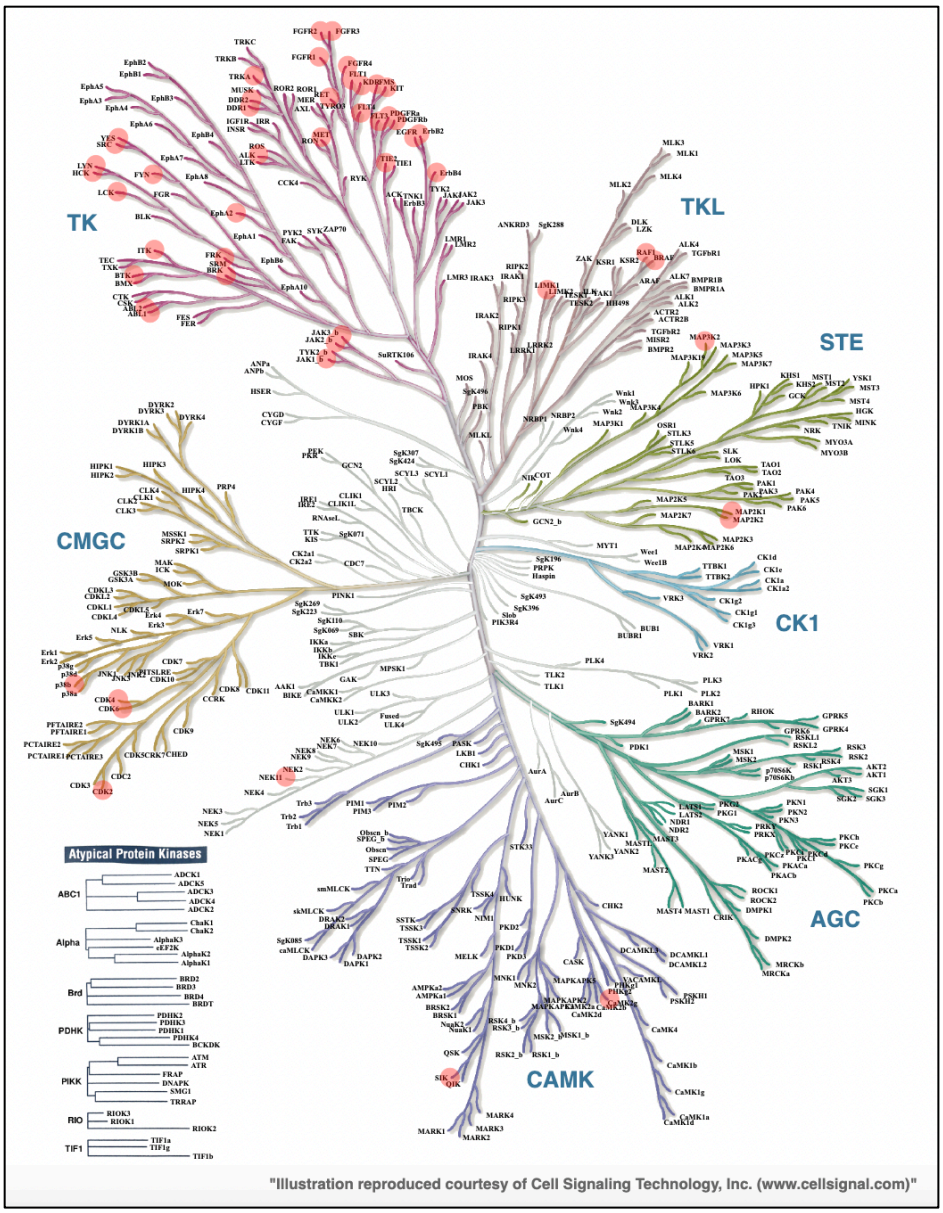


Figure 1-4: Drug target kinases. Protein kinase groups retrieved by KinMap tool developed by Eid, et. al. (2017).[44] Highlighted in red are the drug target kinases, which have an approved drug according to DrugBank.

Ion channels transmit chemical signals through voltage gated and ligand gated mechanisms.[30] Further classifications have arisen based on the type of ligands or localization of these channels, but they are not discussed here since ion channels are not the focus of this project. However, it is noteworthy to mention their involvement in many processes related to cancer, including cell proliferation, migration, and metastasis.[45]

1.2.2. Other cell surface proteins

Other, less studied, non-enzymatic mechanisms involved in signal transmission are supported by adhesion receptors and adapter molecules. These molecules fall under an emerging category of cell surface receptors (**Table 1-1**), and can act as direct receptors or can interact with other receptors by bringing them into proximity and thus initiating the cell signaling pathways through far more complex signaling mechanisms.[46] Below are discussed some of the categories laid out in **Table 1-1** with relevance to cancer formation and progression.

Integrins are cells surface receptors that recognize a variety of cell surface and extracellular matrix (ECM) related ligands, and can bind to more than one ligand. This binding is mediated by alpha and beta chains of heterodimer integrins, but in some cases an activation by outside sources (such as Ca2+ binding) may be required to bind the ligand to integrins. When activated, they engage other scaffolding and cytoplasmic proteins into a vast array of regulation processes involved in cell differentiation, migration and survival. Therefore, integrin alterations and disfunctions are linked with cancer formation and tumor progression.[47]

Adaptors are part of scaffold proteins, and act not only as connectors but are actively involved in signal modulation, transmission, and communication between ligand-protein and protein-protein complexes. The most common adaptor proteins contain the SH2 and SH3 domains mentioned above, linking the receptor tyrosine kinases (RTKs) to downstream targets in signaling pathways crucial to cancer progression such as MAPK.[48] Other cell surface proteins included in **Table 1-1** will be discussed in **Chapter 4**, which presents the results of the cell surface protein enrichment.

1.3 Signaling pathways

Several known cytoplasmic target proteins are linked to the activation of the ERBB family, such as phospholipase C-γ (PLC-γ), Ras, phosphatidylinositol-3 kinase (PI3K), and STAT3. All these effector proteins are part of the traditional downstream pathways that alter gene expression, and their function is briefly described in **Table 1-2**. **Figure 1-5** depicts the signaling mechanisms of the two main pathways discussed here: MAPK and PI3K, which lead to cell cycle progression.

Table 1-2: Main pathways induced by EGFR, their protein components, and function.

Signal receiving cytoplasmic proteins	Downstream signal transducing cytoplasmic proteins	Specific function of the pathway	Ref.
PLC-gamma	Ca^{2+} - CaMK PKC	Modulate cell motility processes by release of intracellular calcium	49, 50
Ras	Raf - MAPK	Cell-cycle progression through G1 phase	49, 51
PI3K	Akt - GSK	Eliminates cell-cycle checkpoints, prevents apoptosis	49, 50
STATs	-	Direct transcriptional activation of genes involved in proliferation, cell-cycle progression	49

1.3.1. MAPK signaling

MAPK signaling is one of the most conserved among eukaryotes, whose multiple pathways play roles in gene expression, cell cycle progression, differentiation, growth and survival, apoptosis, metabolism, and motility. These pathways, which are also known as the conventional MAPK pathways, are distinguished by the sequential activation of three conserved kinases by phosphorylation: MAPKKK → MAPKK → MAPK. MAPK pathways are also comprised of atypical MAPKs that do not abide by the above mentioned three-kinase cascade, but nevertheless utilize phosphorylation of Ser and Thr residues (succeeded by a Pro amino acid) to transmit extracellular signals to downstream substrate and kinase targets following the same specific phosphorylation motif. The broad range of MAPK activity is attributed to several factors including: the variety of substrates such as other protein family kinases (p90 ribosomal S6 kinases/RSKs, mitogen- and stress-activated kinases/MSKs, MAPK-interacting kinases/MNKs, etc.); the specific interaction domains or docking sites with different substrates; and the interaction with scaffolding proteins (**Table 1-1**), which organize the pathway in different structures.[52,53] One of the conventional MAPKs with a central role in cell proliferation are the ERK1/2 MAPKs, which are part of the so-called MAPK-ERK pathway initiated by mitogenic growth factors such as EGF (**Figure 1-5**). In normal cells, continuous activation of this pathway ensures proper cell cycle progression through G1/S, while in cancer cells, this pathway is found to be modified in ~40% of all human cancer types.[52,53] The signaling mechanism is shown in **Figure 1-5** and it involves the activation of cell surface RTKs by auto- or trans- phosphorylation, recruitment of SH2 or PTB domain containing proteins such as GRB2, and G-protein Ras

activation by guanine nucleotide exchange factors (GEF) such as SOS proteins, which mediate the GTP-GDP exchange. Ras is the molecule that interacts directly with the first of the three in the kinase cascade, MAPKKK, which in the ERK pathway is Raf. Both Ras and Raf are targets of extensive mutations in cancer. Next are the MAPKK, MEK1/2 proteins, and the last of the cascade, the MAPK, ERK1/2 proteins. The MAPK-ERK pathway is regulated by a negative feedback loop which makes it challenging to target specific downstream kinase targets such as MEK1/2.[52,53]

1.3.2. PI3K/Akt signaling

The PI3K-AKT signaling pathway is similar to MAPK in the broad range of cellular processes it affects due to its numerous substrates downstream the pathway. This signaling pathway is present in every organ system and cell type in the human genome. In cancer, it plays important roles in cell survival, cell proliferation, metastasis, angiogenesis, and metabolism. Likewise, many of the components of this pathway are mutated or overexpressed in cancer.[54-56] The main protein in this pathway is the serine and threonine kinase Akt or protein kinase B (PKB) acting downstream of the PI3K, which is activated by RTKs, for example EGFR, in response to growth factor stimuli, such as EGF. Activation of Akt requires both phosphorylation of its two conserved residues and translocation to the plasma membrane. In turn, Akt phosphorylates different downstream targets based on a specific phosphorylation motif, where different responses are generated based on the substrate. For example, phosphorylation of the mammalian target of rapamycin (mTOR) substrate activates this protein and its downstream targets, which leads to cell growth and protein synthesis.[54-56] The PI3K-Akt-mTOR axis is also reported to be one of the main supporters of aerobic glycolysis in cancer, discussed below, by providing glucose, lipids, nucleotides, and other nutrition to support the ever-growing needs during cancer progression.[56] On the other hand, phosphorylation of the glycogen synthase kinase 3 (GSK3) and Forkhead Box O (FoxO) transcription factor has an inhibitory role on these substrates, which represses growth and induce cell cycle arrest or apoptosis through downstream effectors, but their inhibition by Akt leads to positive regulation of proliferation and survival. The PI3K-Akt pathway activity is negatively regulated by the tumor suppressor PTEN, which is often downregulated in cancer.[54,55]

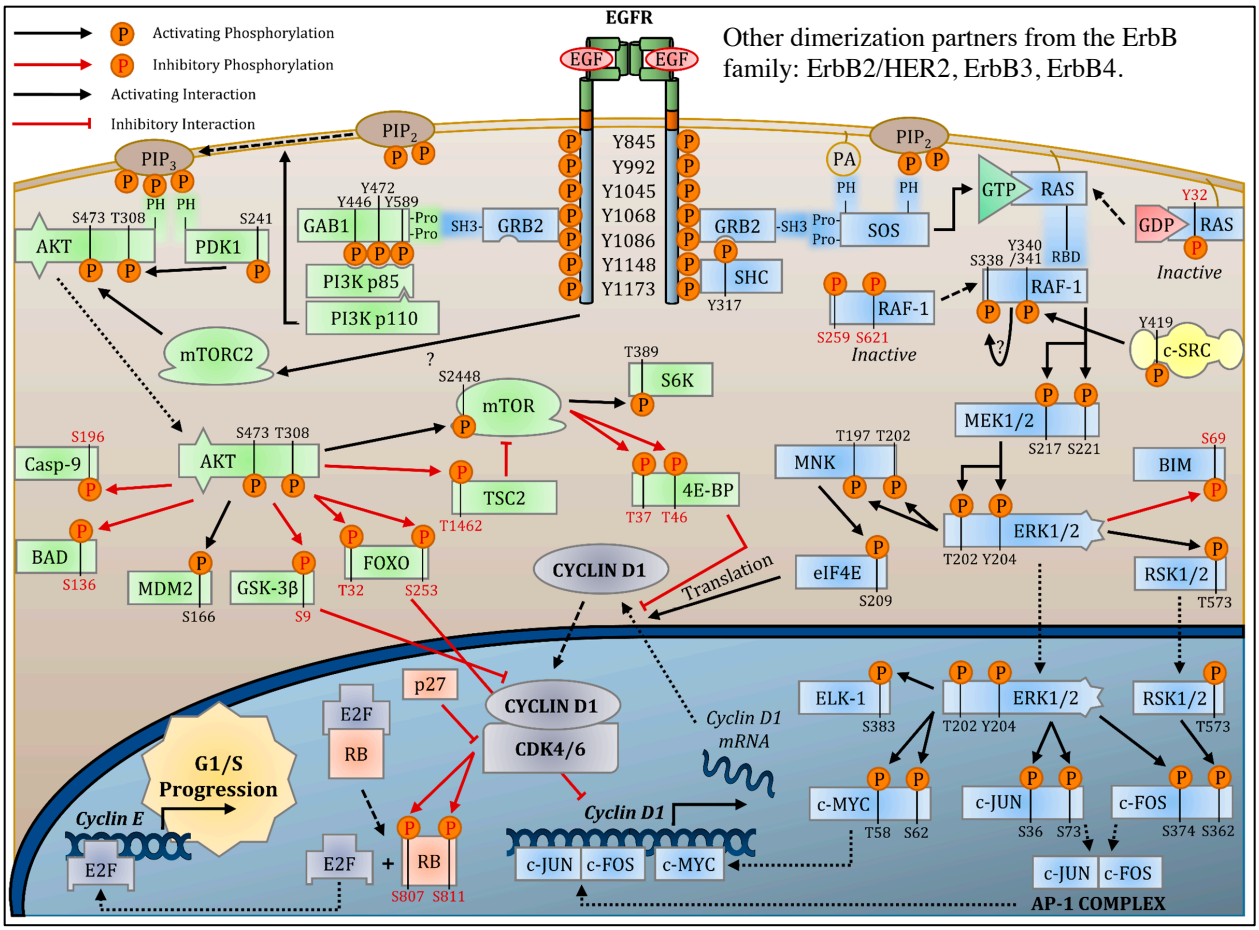


Figure 1-5: MAPK and PI3K pathways upon activation from EGFR. The specific phosphorylation events that take place leading towards G1/S cell cycle progression are shown. Adapted from Wee & Wang (2017)[56] under the CC BY 4.0 license found here: https://creativecommons.org/licenses/by/4.0/legalcode.

1.3.3. Crosstalk and transactivation

Due to the broad implication of both MAPK and Akt pathways in signaling processes, crosstalk and transactivation are common between the components of these pathways, once thought to transduce their signals in a linear fashion. Crosstalk and transactivation, which based on the GO definition represent processes in which one receptor activates another receptor, are also means through which compensatory signaling pathways lead to the development of drug resistance,[12,57] as discussed in **Section 1.1.5**. Some of the mechanisms of crosstalk between the MAPK and Akt pathways include: cross-inhibition, where a component of one pathway has an inhibitory effect on an upstream component of the other pathway and attenuates its signal; cross-activation, where a component of one pathway has an activating effect on an upstream component of the other

pathway and enhances its signal; and convergence, where both pathways act on the same target or group of targets amplifying their response in signaling.[58] Regarding their activation, both pathways can be initiated not only by RTKs, but through other cell membrane receptors, such as GPCRs, either directly or through other proteins acting on the RTKs (transactivation). Many GPCRs are reported to play crucial roles in cancer growth and proliferation, inflammation, and angiogenesis.[59] In addition, the PI3K of the Akt pathway interacts directly with members of the ERBB family (ERBB3 and ERBB4) or indirectly (ERBB1 and ERBB2) through adaptor proteins such as GAB1, or Ras.[56] During signal transduction, ERK can inhibit GAB1, located upstream of the PI3K-Akt pathway and abrogate its signaling as a means of cross-inhibition, or Ras can bind PI3K and activate its signaling as a means of cross-activation mentioned above.[58] One example of pathway convergence is the phosphorylation of the transcription factor FoxO mentioned above by both ERK and AKT in different residues, targeting it for degradation or translocating it to the cytoplasm, respectively, but ultimately inhibiting its function allowing for cell cycle progression.[58] Ultimately, these crosstalk activities provide for a complex landscape in cell signaling, which becomes a challenge for targeted cancer therapies.

1.3.4. Cell cycle in cancer cells

Maintenance of the cell cycle control mechanisms is crucial to a cell's ability to divide correctly, but in cancer cells certain mechanisms are hijacked to allow for continuous cell division. The cell cycle is characterized by two phases: interphase (duplication of genetic material) and M phase (mitosis). During interphase, DNA replication is confined in time in between two gap phases (G1 and G2), which are key decision moments for the cell: G1 for entering the DNA replication phase or exiting and staying in a non-proliferative/quiescence state (G0), and G2 for entering the division or mitosis phase or not. The fine-tuned mechanisms that control this cycle rely mostly on the phosphorylation patterns of cyclin-dependent serine/threonine kinases (CDKs), the expression levels of cyclins, which activate CDKs by binding and translocating them to the nucleus, and the phosphorylation state of the retinoblastoma protein (Rb) as shown in **Figure 1-6**. The activity of CDKs is also controlled by their inhibitors, for example p21, which is transcriptionally activated by the known tumor suppressor p53 and induces cell cycle arrest by binding to cyclins.[60,61] In cancer, the cells continue to progress through the cell cycle without exiting in a quiescence state or being subjected to cell death or apoptosis. This happens due to

the presence of mutations in checkpoint proteins relevant to key phases of the cell cycle that present opportunities for cell cycle exit. There are three main checkpoints that control proper progression through the cell cycle: (1) the G1 checkpoint or restriction checkpoint that blocks entry from G1 into the S-stage of the cell cycle if DNA damage is detected (relevant are the retinoblastoma protein-RB, p107 and p130, that bind the E2F transcription factors preventing them from translocating to the nucleus and initiating the transcription of proteins that enable cell cycle progression, i.e., of checkpoint regulators, cyclins, CDKs, and DNA repair proteins); (2) the G2 checkpoint that surveils the accuracy and completeness of DNA replication and blocks entry into the M-phase if damage is detected (relevant are the ATM/ATR and CHK1/CHK2 proteins that prevent mitotic entry due to DNA damage); and (3) the M or metaphase-to-anaphase transition or spindle checkpoint that checks for correct attachment of sister chromatids to the spindle microtubules (relevant are the spindle assembly checkpoint-SAC proteins, CDC20, and anaphase promoting complex/cyclosome-APC/C proteins).[61] Important throughout the cell cycle is also the master DNA damage repair/tumor suppressor p53 protein and a number of other cell cycle inhibitors such as p21 and p27.

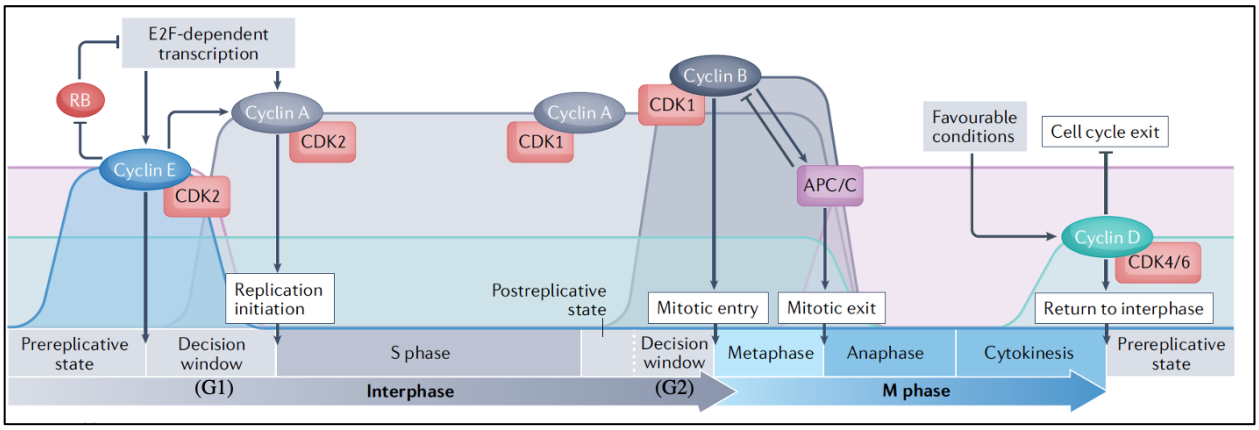


Figure 1-6: Phases of cell cycle and the roles of its components. Upon mitogenic signaling, up-regulated cyclin D binds to CDK4 and CDK6 and initiates the phosphorylation of the RB protein. Hyperphosphorylated RB releases E2F transcription factors enabling them to activate the expression of genes responsible for cell cycle progression, including of cyclin E and A that form a complex with CDK2 and initiate S phase entry. Later in the cell cycle, the cyclin A/B-CDK1 complexes drive cells into mitosis and the APC/C complex is responsible for exiting mitosis and degrading cyclin A and cyclin B, which resets the cell cycle at the end of mitosis. Adapted from Matthews, et. al. (2022)[61] under the Springer Nature license.

1.3.5. Energy production in cancer cells

Adenosine triphosphate (ATP) is the currency of energy in both normal and cancer cells, but the mechanisms of energy production differ. In normal cells, the majority of ATP production relies on (a) anaerobic cytosolic breakdown of glucose into pyruvate via glycolysis (the 1st step of the cellular respiration process), followed by (b) aerobic mitochondrial tricarboxylic acid (TCA) or Krebs cycle and oxidative phosphorylation (OXPHOS) with generation of ATP (the 2nd step of the cellular respiration process). Metabolic rewiring of growing and proliferating cancer cells relies on a much enhanced rate of glycolysis followed by lactic acid production, both occurring in the cytosol under anaerobic conditions, with the result of less ATP production but more biosynthetic precursors production-which are much needed for the fast growth of cancer cells. Confusingly, the process was termed aerobic glycolysis (the Warburg effect) because the path of lactic acid production is favored even when oxygen is present for oxidative phosphorylation. Different types of cancer cells can use, however, both OXPHOS and "aerobic" glycolysis for ATP production. Moreover, ATP can be also produced through glutaminolysis, as part of the TCA cycle fed by glutamine, which similar to glucose is another important fuel to the cancer cells. These processes are known as the Q-effect (Q for glutamine) and the Warburg effect, and research is conducted to identify the therapeutic targets of these rewired mechanisms.[62,63] It has been stated that OXPHOS in cancer cells is deficient due to the numerous mutations and aberrations of both mitochondrial structural and functional related genes, however there is continuous debate in the field about the roles of the OXPHOS components in cancer cells.[63] Cells secrete ATP via different mechanism, and the extracellular ATP (eATP) can act as a signaling molecule with important roles in cancer. eATP can be either degraded by the action of ectonucleotidases or taken by the purinergic receptors on the cell surface such as the P2X ion channels or the P2Y GPCRs. Its activity in cancer is complex because it can promote immune responses in the tumor microenvironment, but can also play a role in metastasis and invasion.[64]

1.3.6. Signal transduction by phosphorylation

As mentioned above, kinases modulate the signaling pathways through phosphorylation, the most common reversible post translational modification (PTM). Phosphorylation affects more than 2/3 of all known proteins encompassing more than 200,000 known phosphorylation sites and an additional ~760,000 predicted sites.[65] Phosphorylation is the transfer by a kinase of the γ-

phosphate group (PO_4) from ATP to the amino acid residues of serine (Ser/S) in 86.4% of the events, threonine (Thr/T) in 11.8% of the events, or tyrosine (Tyr/Y) in 1.8% of the events in eukaryotic cells, even though other unorthodox amino acids such as histidine (His/H) can be phosphorylated as well. Phosphorylation is reversed by the action of phosphatases which remove the phosphate group from the substrate.[65,66] Phosphorylation plays a critical role in all biological functions of a cell, including cell growth, cell cycle, apoptosis, cell motility, adhesion, and signaling. In cancer, phosphorylation sites are reported to carry mutations or have altered expression profiles.[66]

Even though equipped with the knowledge of having detected such a large number of phosphorylation sites, their upstream and downstream mechanisms and functions in signaling remain largely unknown, with more than 95% of phosphorylation sites not being assigned a kinase or a function.[67] This underscores the complex nature of signaling pathways and networks facilitated by phosphorylation events, crosstalk, feedback loops, protein complexes, etc., providing a vast landscape of exploration in both normal and cancer cells. Another challenge in the field remains the pairing between kinases and their downstream targets or substrates. It has been reported that 87% of the known phosphorylated substrates are accounted for only by 20% of the kinases.[66] However, as discussed above in the cases of MAPK and AKT pathways, different kinases have specific substrates, which can range in numbers from several to hundreds. Knowing which substrates and sites to phosphorylate among thousands of potential phosphorylation sites is important for the subsequent signaling function of each kinase. The kinase specificity is determined by the structure of the kinase active site-which categorizes the kinases into Ser/Thr and Tyr kinases; the phosphorylation motifs or the consensus sequence around the phosphorylated residue (p-site); the docking site interactions away from the kinase active- and substrate p- site; and others.[41] The study of phosphorylation events in the signaling networks holds tremendous opportunities for targeted therapies and is most suited to analysis by using mass spectrometry technologies.[66]

1.4 Mass spectrometry: a powerful tool for analyzing the proteome

Proteomics is the field that studies the expression levels of proteins in a system, as a way to characterize it in terms of cellular processes, pathways, and diseases. It is often referred to as the new genomics or functional genomics at the protein level. Proteome refers to all expressed proteins encoded in the genome of an organism. Proteins are always considered the effectors, hence, qualifying and quantifying their levels has the potential to lead to an accurate view of the system. Protein level expression in a cell is not only dependent on the mRNA expression levels, but is also affected by other factors such as translational and post-translational modifications, protein degradation, etc. The impact of such factors cannot be measured by the advanced sequencing technologies of transcriptomics. Moreover, 2D PAGE (two-dimensional polyacrylamide gel electrophoresis)-a technology that is used to separate, identify and quantify proteins, has proven to not be always suitable for the analysis of complex samples with a large range of protein concentrations. Mass spectrometry (MS)-based proteomics, on the other hand, has evolved nowadays into the most powerful tool for comprehensive qualitative and quantitative mapping of the cellular proteome.[68,69]

1.4.1. Instrumentation and data acquisition

Mass spectrometers analyze ions produced from protein/peptide samples, and identify the components based on the mass-to-charge (m/z) ratio of the detected ions. An overview of the components and workflow of a mass spectrometer is given in **Figure 1-7 A**. The instrument contains an ion source responsible for ionizing the neutral molecules through different mechanisms, such as electron ionization (EI) or atmospheric pressure chemical ionization (APCI) for gaseous molecules, matrix assisted laser desorption ionization (MALDI) for samples mixed with a laser excitable matrix, and electrospray ionization (ESI) for samples present in a solution-which is the mode of ionization used in this project, as shown in **Figure 1-7 B**. ESI is used to ionize polar molecules either in a positive or negative ion mode through the application of a high voltage to a liquid analyte solution delivered via a capillary tube to the MS ion source. Based on the polarity of the voltage applied to the capillary, positive or negative ions are formed. A spray of highly charged droplets is produced at the tip of the capillary in a high electric field environment, where under the influence of the electric field, temperature and nebulizing gas, singly or multi-charged ions are produced. The ions enter the mass spectrometer and are guided

by ion optics elements to the analyzer.[70-72] The mass analyzer is the operative center of the MS instrument that enables the separation of ions based on m/z ratios. Its performance is defined by characteristics such as: mass range- the range of m/z values that can be handled by the analyzer; scan rate or scan speed- the rate the analyzer uses to measure the ions over a given mass range in amu/s; mass accuracy- the ability to measure accurately the m/z of ions (i.e., close to the theoretical mass) in ppm units; and resolution- the ability to differentiate between two ions of different m/z, measured as a ratio of mass/Δmass at a specific m/z. Other important parameters referring to the performance of the whole instrument include sensitivity-the ratio of output (ion signal intensity change) over input (sample concentration change), and detection limit-the smallest sample quantity that results in a signal distinguishable from the background noise.[71] A trade-off between these parameters is needed for fast and accurate analysis.

There are 4 basic designs of mass analyzers: quadrupoles, ion traps, time-of-flights, and Fourier transform. The quadrupole is composed of four parallel rods, equally spaced from a central axis, where a beam of ions is fed. The opposing rods form a set of both positively- and negatively-charged elements through the application of direct current (DC) and alternating current (AC), allowing only for a specific range of m/z to pass through the rods, while all the other ions are eliminated either through collisions with the rods or via the vacuum system. Adjusting the DC and AC voltages affects the m/z window that determines which ions will pass to the detector; a narrower m/z window will result in higher resolution.[70,71] The ion trap is similar to a quadrupole but comes in an enclosed shape comprised of two end-cap electrodes that allow for the entrance and exit of ions, and a ring electrode operated by a main radio-frequency (RF) voltage. Supplemental RF voltages are applied to the end caps, and, overall, the combination of DC and RF potentials that are applied to the three electrodes allow for trapping the ions in a 3D quadrupole field. Ions of different m/z enter the 'trap' at the same time, and under the electric field forces, based on their m/z, engage in an oscillating motion of different frequencies. By altering the amplitude of the RF voltage applied to the ring electrode and by applying additional excitation RF voltages to the end-caps, the ions inside the trap can be excited and eliminated from the trap when their frequency is matching the resonant frequency of the end-cap electrodes. As a result, different m/z ions will leave the trap at different times allowing for the generation of a mass spectrum.[70,71] The time-of-flight (TOF) analyzer uses a kinetic principle to measure the

time travelled by ions of different m/z, through a field free region between the ion source and the detector. When stimulated by the same accelerating potential force, ions of lower m/z will reach the detector faster compared to ions of higher m/z. Spatial, temporal, and initial kinetic energy variations result in poor resolution, which can be improved by either introducing a time delay between ion formation and ion extraction from the source or by the use of ion reflectrons.[70,71]
The Fourier transform - ion-cyclotron resonance (FT-ICR) mass spectrometers utilize a cell with a magnetic field to trap ions with the aid of three sets of panels with distinct roles in trapping, exciting and detecting the ions. In addition to the magnetic forces, electric forces are applied to the trapping plates. The ions move in a cyclotron motion (circular orbit) perpendicular to the magnetic field. The frequency of the cyclotron motion is dependent on the mass and the charge of ions. When an RF voltage is applied from the excitation plates, ions with a resonant frequency will be excited into a larger radius, a process that will induce an image current perpendicular to the ion trajectory that will be detected by the receiver/detection plates. FTICR consists in simultaneously exciting all ions by a rapid scan of frequencies and measuring the composite image current on the cell plates. The image current is deconvoluted by a Fourier transform to generate individual frequencies, characteristic to each m/z.[70,71]

The instrument used in this project has an Orbitrap mass analyzer, an ion trap with two outer barrel-shaped electrodes (bended in a trap form) and a central spindle-shaped electrode (**Figure 1-7 C**). The operation of the analyzer relies on the use of electrostatic fields (but no RF or magnetic fields) and properly shaped electrodes that determine the frequencies of ion oscillations (especially of the axial ones along the central electrode) that are used to measure the m/z ratios of ions. Multiple ions that oscillate in the Orbitrap analyzer generate a complex waveform signal that is deconvoluted by using a Fourier transform. The Orbitrap is a high resolving power analyzer and one of the most powerful mass analyzers in the field.[71,73] To achieve better performance, hybrid mass spectrometers combine different types of analyzers to capitalize on their strengths, such as the case of the QExactive instrument used in this project that combines quadrupole with Orbitrap technologies.[71,74]
Lastly, the ions are transferred to a detector that produces a signal proportional to the abundance of ions, which is further analyzed by a computer for displaying the mass spectra and collecting the data (**Figure 1-7 A**).

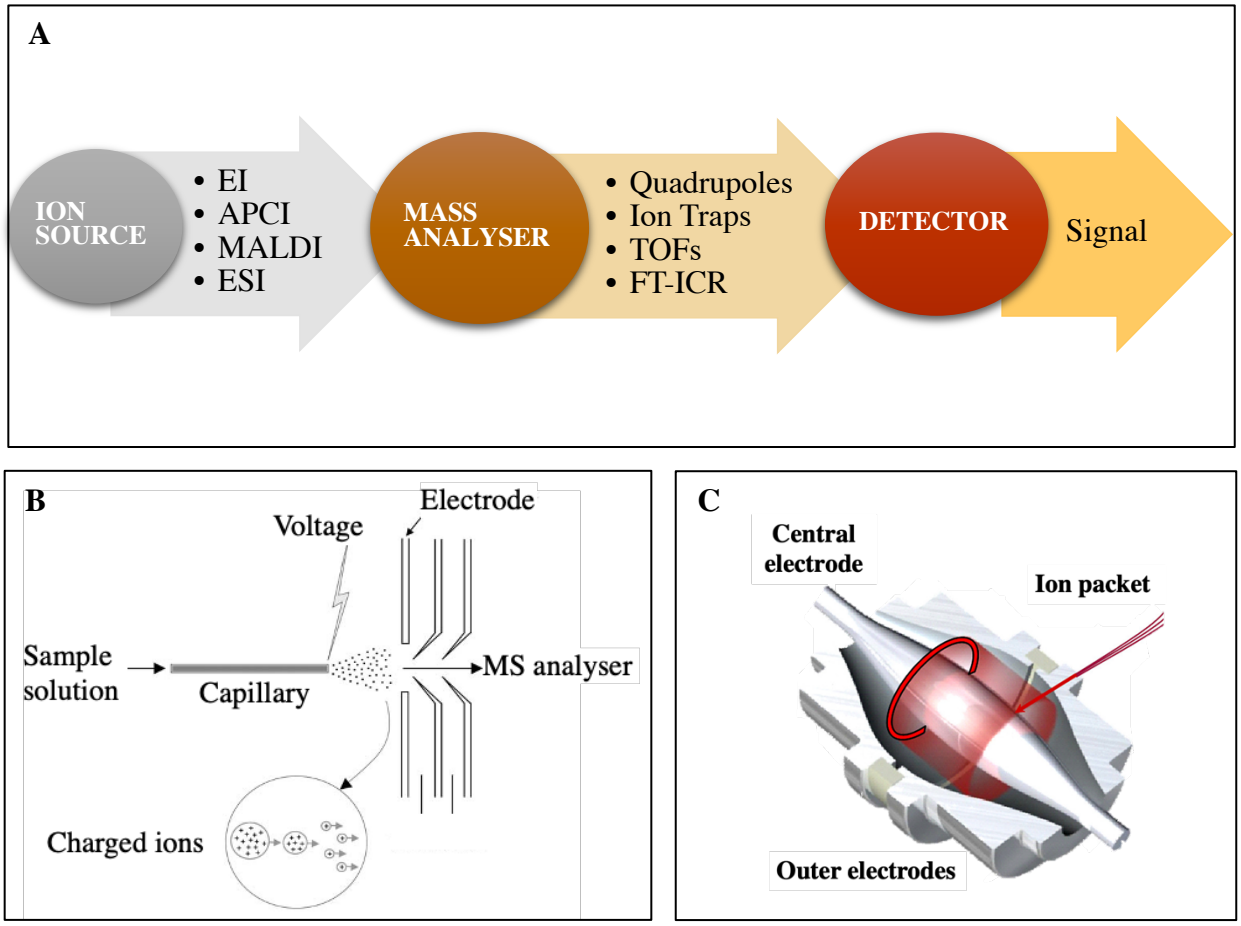


Figure 1-7: The core components of a mass spectrometer. **(A)** ion source-required to convert the protein/peptide sample into gaseous phase ions, mass analyzer-required to separate the ions based on mass-to -charge ratio, and detector-required to produce a signal from the detected ions, **(B)** ESI ion source schematics adapted from Lane (2005)[70] under the Springer Nature license, **(C)** Orbitrap mass analyzer schematic. Adapted with permission from Zubarev & Makarov (2013),[73] copyright 2013 American Chemical Society.

1.4.2. Chromatography

In this project, the peptide mixtures analyzed by the mass spectrometer are prepared by lysing the cells, extracting and denaturing the protein content, digesting the proteins into peptides with trypsin, and passing the peptide mixtures through C18 cartridges to remove the small/polar contaminants, and SCX (strong cation exchange) solid phase extraction cartridges to remove the nonpolar contaminants, before MS analysis. This workflow, even though time consuming, ensures that the samples are of a high purity compared to other methods of sample preparation.[75] Due to the complex nature of the sample, a separation technique such as liquid chromatography

or capillary electrophoresis is used to separate the sample into components prior to MS analysis, either on-line (the sample is eluting from the chromatographic system and is analyzed in real time by the MS) or off-line (the sample is collected after elution and analyzed by the MS at a later time).[71] Chromatographic separation of a sample into its components relies on the use of a mobile and a stationary phase: the mobile phase flows through and interacts with the stationary phase and the sample components that are adsorbed on the stationary phase. A mobile phase gradient generated by the use of two mobile phases (one low and the other high in organic solvent content) enables the distribution and elution of the peptide components from the separation column based on their affinity to the stationary phase: the stronger the interaction with the stationary phase the longer they will be retained in the chromatographic column.[76]

1.4.3. Targeted and untargeted proteomics

The most common way of analyzing a complex peptide mixture is by tandem mass spectrometry, which involves a two-stage mass analysis: first, the detection and isolation of a precursor ion of a specific m/z, and second, fragmentation of the precursor by collision with an inert gas (collision-induced dissociation) to generate charged product ions that will be used for constructing a peptide sequence based on the series of product ions formed. Since the fragmentation occurs along the peptide backbone, the ions series that are produced vary from each other in mass by one amino acid. The fragmentation pattern will depend on the collisional energy, type of collision gas, peptide sequence, charge, and mass. Algorithms will match the generated sequences with the corresponding proteins from the database searches and assign scores to each match.[71,72] Tandem MS or MS/MS analysis can be performed in several acquisition modes (**Figure 1-8**), i.e., in a targeted fashion (multiple reaction monitoring-MRM or parallel reaction monitoring-PRM) when specific precursor ions are selected by the user for data acquisition, or in an un-targeted fashion (data dependent analysis-DDA or data-independent analysis-DIA) when the precursor ions of a certain m/z are selected for fragmentation based on their abundance (such as in DDA) or all the precursor ions within a certain m/z range are selected and fragmented (such as in DIA). The difference between un-targeted and targeted detection lies in the ability of the un-targeted method to theoretically identify all the components of a sample providing breadth of analysis, and of the targeted method to provide higher sensitivity in detecting and quantifying pre-selected precursor ions. An optimal analysis would require a combination of the breadth of

the DDA with the quantitative abilities of the MRM/PRM. The DIA method is also very promising, but the interpretation of the DIA scans is challenging due to the complexity introduced by the fragmentation of all precursor peptides within a given m/z range,[77] as shown in **Figure 1-8**. In this project, the analysis was run in DDA mode, and the validation of the presence of certain proteins was carried out by PRM.

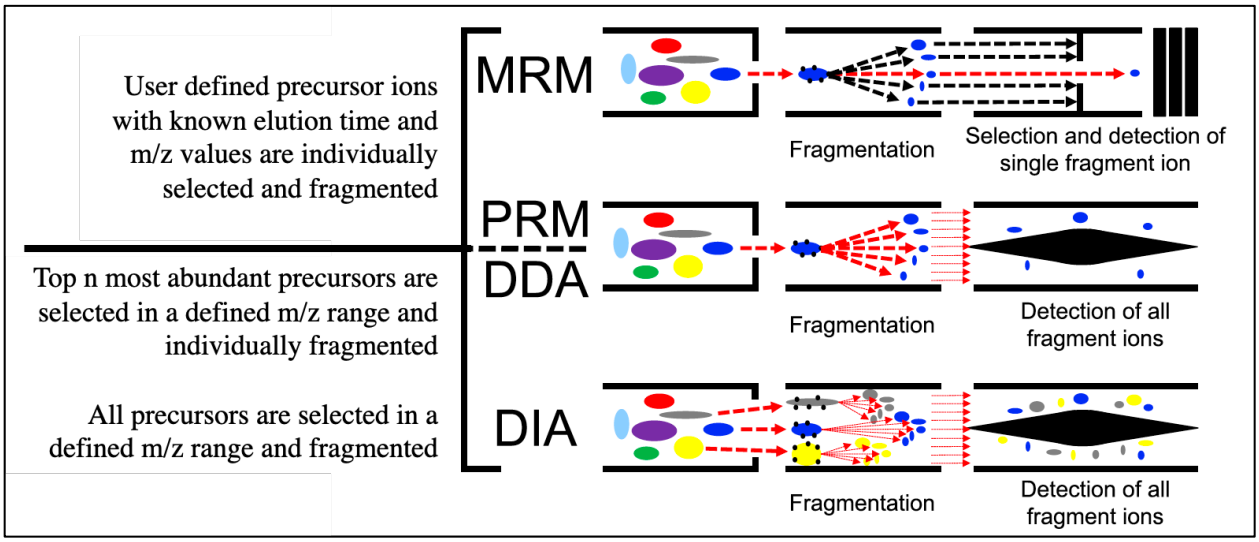


Figure 1-8: Different modes of acquisition for tandem MS and their mode of operation. Adapted from Hu, et. al. (2016)[77] under the CC BY 4.0 license found here: https://creativecommons.org/licenses/by/4.0/legalcode.

1.4.4. Proteome Discoverer software

Peptide sequence identifications and assignments to their parent proteins are performed by algorithms such as SEQUEST, a commercial search engine, that functions as part of of the Proteome Discoverer platform that was used in this project. SEQUEST uses a database to create theoretical tandem mass spectra from precursor ions of a particular m/z, compares the ions detected in the experimental spectra with the theoretical ions, and assigns scores that characterize the quality of the match, among which the cross-correlation score (XCorr) is the most relevant. The highest score indicates the best match and the output is a peptide-spectrum match (PSM). The more matches/PSMs are assigned to a peptide and subsequently to a protein, the more confident the assignment is considered to be. During the database search, several parameters can be defined including the enzyme type that was used for digesting the protein sample, the

expected post-translational modifications, the mass tolerances for the precursor and fragment ions, and the desired false discovery rates (FDRs).[78]

1.4.5. Quantification

Several strategies for quantifying proteins detected by MS have been developed over the years. Relative quantification provides information about the relative change in abundance between samples based on a quantitative ratio. The strategies for relative quantification can be categorized into *in vivo* (metabolically) or *in vitro* (chemically and enzymatically) stable isotope labelling methods, and non-labelling approaches involving intensity-, area- and spectral count measurements. *In vivo* labelling is applicable to cell cultures or whole organisms only, therefore, *in vitro* labelling is better suited for analyzing various sample types including tissues or body fluids. A short description of each strategy is provided in the following table based on the reviews from Lane et. al, Gilany et. al, and Rozanova et. al.[70,72,79]

Table 1-3: Labelling and non-labelling techniques for relative protein quantification: function and advantages or disadvantages.

Quantification Method	Description	Advantages/ Disadvantages
In vivo (metabolic labeling)		
^{15}N	Differentiates the amounts of a protein from two different cell cultures, one grown in ^{15}N and another in ^{14}N containing media.	The entire pool of amino acids is labelled.
SILAC **(Stable Isotope Labelling by Amino Acids in Cell Culture)**	Select essential amino acids are labelled and introduced in the cell culture medium where they are incorporated by the cell into all the proteins. Differentiates based on 'light' or 'heavy' amino acids.	Cell growth, morphology, and differentiation ability or timing are not affected by the labelling. Expensive.
In vitro (chemical labeling)		

ICAT **(Isotope coded affinity tags)**	ICAT reagents are covalently linked to the cysteine residues of proteins. Differentiates based on the mass of affinity tag linker (either heavy or light)	Does not work on proteins with absent cysteine or with modified cysteine residues. Reduced sample complexity.
MCAT **(Mass-Coded Abundance Tagging)**	Uses guanidination of C-terminal lysine residues to transform lysine into a heavier component (homoarginine).	Increases ionization efficiency and does not affect the amino terminus.
iTRAQ (isobaric tags for relative and absolute quantification), TMT (tandem mass tags)	Labels peptides from different samples with an isobaric tag containing an amine-reactive group, a balance/mass normalization group, and a reporter group for a total of same mass. Intensities of the different reporter ions produced in the tandem MS upon fragmentation are proportional to the different original sample abundances.	Enhanced signal intensity and allows for multiplexing (several samples). Expensive and certain software packages are required to interpret the data.
In vitro (enzymatic labeling)		
^{18}O	Each sample is separately digested in $H_2{}^{16}O$ and $H_2{}^{18}O$. Differentiates based on ion intensities of ^{16}O and ^{18}O labelled peptides.	Simple method, but limited use for larger peptides due to isotopic overlap.
Label -free		
Spectral counting	The number of spectra correlates with relative protein abundance between different samples because the more abundant proteins will result in more spectra.	Easy to compute, however normalization and statistical test are still required. Proteins represented by low spectra counts are not reliable.

Area/Intensity based	Signal intensities of ions correspond to their concentrations, therefore peak areas/intensities of either MS or MS/MS (precursor or product ion) are used to relatively quantify between samples	No additional sample preparation steps. Variation between technical replicates (multiple measurement of the same sample) need to be considered. Usually, normalization is required and the use of computational algorithms for statistical analysis.

The availability of a variety of quantification methods allows researchers to choose the most appropriate one based on the sample to be analyzed.

1.5 References

1. Anders, C. K., Johnson, R., Litton, J., Ruddy, K. J., & Bleyer, A. (2017). Breast Cancer Before 40. In A. Bleyer, R. Barr, L. Ries, J. Whelan, A. Ferrari (eds), *Cancer in Adolescents and Young Adults*, 177-202. Cham, Switzerland: Springer. DOI: 10.1007/978-3-319-33679-4_8

2. Key, T. J., Verkasalo, P. K., & Banks, E. (2001). Epidemiology of breast cancer. *The Lancet Oncology*, *2*(3), 133-140. DOI: 10.1016/S1470-2045(00)00254-0

3. National Cancer Institute. (2021). What is Cancer? *Understanding cancer*. Retrieved from https://www.cancer.gov/about-cancer/understanding/what-is-cancer

4. Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of Cancer: The Next Generation. *Cell*, *144*(5), 646-674. DOI: 10.1016/j.cell.2011.02.013

5. Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of Cancer. *Cell*, *100*(1), 57-70. DOI: 10.1016/S0092-8674(00)81683-9

6. Hanahan D. (2022). Hallmarks of cancer: new dimensions. *Cancer discovery, 12*(1), 31-46. DOI: 10.1158/2159-8290.CD-21-1059

7. Weigelt, B., Geyer, F. C., & Reis-Filho, J. S. (2010). Histological types of breast cancer: how special are they?. *Molecular Oncology*, *4*(3), 192-208. DOI: 10.1016/j.molonc.2010.04.004

8. Prat, A., Pineda, E., Adamo, B., Galván, P., Fernández, A., Gaba, L., ... Muñoz, M. (2015). Clinical implications of the intrinsic molecular subtypes of breast cancer. *The Breast*, *24*, S26-S35. DOI: 10.1016/j.breast.2015.07.008

9. Bjornstrom, L., & Sjoberg, M. (2005). Mechanisms of estrogen receptor signaling: convergence of genomic and nongenomic actions on target genes. *Molecular endocrinology*, *19*(4), 833-842. DOI: 10.1210/me.2004-0486

10. Leonhardt, S. A., Boonyaratanakornkit, V., & Edwards, D. P. (2003). Progesterone receptor transcription and non-transcription signaling mechanisms. *Steroids*, *68*(10-13), 761-770. DOI: https://doi.org/10.1016/S0039-128X(03)00129-6

11. Harari, D., & Yarden, Y. (2000). Molecular mechanisms underlying ErbB2/HER2 action in breast cancer. *Oncogene*, *19*(53), 6102. DOI: 10.1038/sj.onc.1203973

12. Nahta, R. (2019). Novel therapies to overcome HER2 therapy resistance in breast cancer. In M.R. Szewczuk, B. Qorri, M. Sambi (eds), *Current Applications for Overcoming Resistance to Targeted Therapies*, 191–221. Cham, Switzerland: Springer. DOI: 10.1007/978-3-030-21477-7_7

13. Bender, L. M., & Nahta, R. (2008). HER2 cross talk and therapeutic resistance in breast cancer. *Frontiers in bioscience, 13*, 3906–3912. DOI: 10.2741/2978

14. Graus-Porta, D., Beerli, R. R., Daly, J. M., & Hynes, N. E. (1997). ErbB-2, the preferred heterodimerization partner of all ErbB receptors, is a mediator of lateral signaling. *The EMBO journal, 16*(7), 1647–1655. DOI: 10.1093/emboj/16.7.1647

15. Gutierrez, C., & Schiff, R. (2011). HER2: biology, detection, and clinical implications. *Archives of pathology & laboratory medicine, 135*(1), 55–62. DOI: 10.5858/2010-0454-RAR.1

16. Iqbal, N., & Iqbal, N. (2014). Human Epidermal Growth Factor Receptor 2 (HER2) in Cancers: Overexpression and Therapeutic Implications. *Molecular biology international,* 852748. DOI: 10.1155/2014/852748

17. The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. (2020). Pan-cancer analysis of whole genomes. *Nature, 578*, 82–93. DOI: 10.1038/s41586-020-1969-6

18. National Cancer Institute. (2019). Outcomes & Impact of The Cancer Genome Atlas. *TCGA*. Retrieved from https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga/history

19. National Cancer Institute. (2022). Data Portal Summary. *GDC Data Portal*. Retrieved from https://portal.gdc.cancer.gov/

20. Alexandrov, L. B., Kim, J., Haradhvala, N. J., Huang, M. N., Tian Ng, A. W., Wu, Y., … PCAWG Consortium (2020). The repertoire of mutational signatures in human cancer. *Nature, 578*(7793), 94–101. DOI: 10.1038/s41586-020-1943-3

21. Swain, S. M., Shastry, M., & Hamilton, E. (2022). Targeting HER2-positive breast cancer: advances and future directions. *Nature reviews drug discovery, 22*(2), 101–126. DOI: 10.1038/s41573-022-00579-0

22. National Cancer Institute. (2022). Breast Cancer Treatment. *Breast Cancer*. Retrieved from https://www.cancer.gov/types/breast/patient/breast-treatment-pdq#link/_515

23. Schlam, I., & Swain, S. M. (2021). HER2-positive breast cancer and tyrosine kinase inhibitors: the time is now. *NPJ breast cancer, 7*(1), 56. DOI: 10.1038/s41523-021-00265-1

24. Tsang, R. Y., Sadeghi, S., & Finn, R. S. (2011). Lapatinib, a dual-targeted small molecule inhibitor of EGFR and HER2, in HER2-amplified breast cancer: from bench to bedside. *Clinical Medicine Insights: Therapeutics*, 3. DOI: 10.4137/CMT.S3783

25. Martorana, F., Motta, G., Pavone, G., Motta, L., Stella, S., Vitale, S. R., … Vigneri, P. (2021). AKT Inhibitors: New Weapons in the Fight Against Breast Cancer? *Frontiers in pharmacology*, *12*, 662232. DOI: 10.3389/fphar.2021.662232

26. Lin, J., Sampath, D., Nannini, M. A., Lee, B. B., Degtyarev, M., Oeh, J., … Lin, K. (2013). Targeting activated Akt with GDC-0068, a novel selective Akt inhibitor that is efficacious in

multiple tumor models. *Clinical cancer research, 19*(7), 1760–1772. DOI: 10.1158/1078-0432.CCR-12-3072

27. Fujimoto, Y., Morita, T. Y., Ohashi, A., Haeno, H., Hakozaki, Y., Fujii, … Mukohara, T. (2020). Combination treatment with a PI3K/Akt/mTOR pathway inhibitor overcomes resistance to anti-HER2 therapy in PIK3CA-mutant HER2-positive breast cancer cells. *Scientific reports, 10*(1), 21762. DOI: 10.1038/s41598-020-78646-y

28. Nature Education. (2014). Cells Receive and Process a Diverse Set of Chemical Signals and Sensory Stimuli. *Cell Biology for Seminars.* Retrieved from https://www.nature.com/scitable/ebooks/cell-biology-for-seminars-14760004/118245184

29. Bull, S. C., & Doig, A. J. (2015). Properties of protein drug target classes. *PloS one, 10*(3), e0117955. DOI: 10.1371/journal.pone.0117955

30. Cooper, G.M. (2000). *The Cell: A Molecular Approach* (2nd ed). Sunderland, MA: Sinauer Associates.

31. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). *Molecular Biology of the Cell* (4th ed). New York, NY: Garland Science.

32. Feher, J. (2012). Cell Signaling. *Quantitative Human Physiology*, 158-170. Boston, MA: Academic Press. DOI: 10.1016/B978-0-12-382163-8.00019-0

33. IUPHAR/BPS Guide to Pharmacology. (2018). *Targets*. Retrieved from http://www.guidetopharmacology.org/targets.jsp

34. Uings, I. J., & Farrow, S. N. (2000). Cell receptors and cell signalling. *Molecular Pathology*, *53*(6), 295-299. DOI: 10.1136/mp.53.6.295

35. Chiodoni, C., Colombo, M. P., & Sangaletti, S. (2010). Matricellular proteins: from homeostasis to inflammation, cancer, and metastasis. *Cancer and Metastasis Reviews*, *29*(2), 295-307. DOI: 10.1007/s10555-010-9221-8

36. Bornstein, P., & Sage, E. H. (2002). Matricellular proteins: extracellular modulators of cell function. *Current opinion in cell biology*, *14*(5), 608-616. DOI: https://doi.org/10.1016/S0955-0674(02)00361-7

37. Gether, U. (2000). Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. *Endocrine Reviews*, *21*(1), 90-113. DOI: 10.1210/edrv.21.1.0390

38. Heldin, C. H., Lu, B., Evans, R., & Gutkind, J. S. (2016). Signals and Receptors. *Cold Spring Harbor perspectives in biology, 8*(4), a005900. DOI: 10.1101/cshperspect.a005900

39. Guo, S., Zhao, T., Yun, Y., & Xie, X. (2022). Recent progress in assays for GPCR drug discovery. *American journal of physiology-cell physiology, 323*(2), C583–C594. DOI: 10.1152/ajpcell.00464.2021

40. Bhullar, K. S., Lagarón, N. O., McGowan, E. M., Parmar, I., Jha, A., Hubbard, B. P., & Rupasinghe, H. P. V. (2018). Kinase-targeted cancer therapies: progress, challenges and future directions. *Molecular cancer, 17*(1), 48. DOI: 10.1186/s12943-018-0804-2

41. Ubersax, J. A., & Ferrell Jr, J. E. (2007). Mechanisms of specificity in protein phosphorylation. *Nature Reviews Molecular Cell Biology*, *8*(7), 530-541. DOI: 10.1038/nrm2203

42. Lee, P. Y., Yeoh, Y., & Low, T. Y. (2022). A recent update on small-molecule kinase inhibitors for targeted cancer therapy and their therapeutic insights from mass spectrometry-based proteomic analysis. *The FEBS journal.* DOI: 10.1111/febs.16442

43. Manning, G., Whyte, D. B., Martinez, R., Hunter, T., & Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science, 298*(5600), 1912–1934. DOI: 10.1126/science.1075762

44. Eid, S., Turk, S., Volkamer, A., Rippmann, F., & Fulle, S. (2017). KinMap: a web-based tool for interactive navigation through human kinome data. *BMC bioinformatics, 18*(1), 16. DOI: 10.1186/s12859-016-1433-7

45. Lang, F., & Stournaras, C. (2014). Ion channels in cancer: future perspectives and clinical potential. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 369*(1638), 20130108. DOI: 10.1098/rstb.2013.0108

46. Buday, L., & Tompa, P. (2010). Functional classification of scaffold proteins and related molecules. *The FEBS journal, 277*(21), 4348–4355. DOI: 10.1111/j.1742-4658.2010.07864.x

47. Hamidi, H., & Ivaska, J. (2018). Every step of the way: integrins in cancer progression and metastasis. *Nature reviews Cancer, 18*(9), 533–548. DOI: 10.1038/s41568-018-0038-z

48. Zeke, A., Lukács, M., Lim, W. A., & Reményi, A. (2009). Scaffolds: interaction platforms for cellular signalling circuits. *Trends in cell biology, 19*(8), 364–374. DOI: 10.1016/j.tcb.2009.05.007

49. Lo, H. W., Hsu, S. C., & Hung, M. C. (2006). EGFR signaling pathway in breast cancers: from traditional signal transduction to direct nuclear translocalization. *Breast Cancer Research and Treatment*, *95*(3), 211-218. DOI: 10.1007/s10549-005-9011-0

50. Price, J. T., Tiganis, T., Agarwal, A., Djakiew, D., & Thompson, E. W. (1999). Epidermal growth factor promotes MDA-MB-231 breast cancer cell migration through a phosphatidylinositol 3′-kinase and phospholipase C-dependent mechanism. *Cancer research, 59*(21), 5475-5478.

51. Downward, J. (2003). Targeting RAS signalling pathways in cancer therapy. *Nature Reviews Cancer, 3*(1), 11-22. DOI: 10.1038/nrc969

52. Cargnello, M., & Roux, P. P. (2011). Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. *Microbiology and molecular biology reviews, 75*(1), 50–83. DOI: 10.1128/MMBR.00031-10

53. Lee, S., Rauch, J., & Kolch, W. (2020). Targeting MAPK Signaling in Cancer: Mechanisms of Drug Resistance and Sensitivity. *International journal of molecular sciences, 21*(3), 1102. DOI: 10.3390/ijms21031102

54. He, Y., Sun, M. M., Zhang, G. G., Yang, J., Chen, K. S., Xu, W. W., & Li, B. (2021). Targeting PI3K/Akt signal transduction for cancer therapy. *Signal transduction and targeted therapy, 6*(1), 425. DOI: 10.1038/s41392-021-00828-5

55. Manning, B. D., & Toker, A. (2017). AKT/PKB Signaling: Navigating the Network. *Cell, 169*(3), 381–405. DOI: 10.1016/j.cell.2017.04.001

56. Wee, P., & Wang, Z. (2017). Epidermal Growth Factor Receptor Cell Proliferation Signaling Pathways. *Cancers, 9*(5), 52. DOI: 10.3390/cancers9050052

57. Binns, D., Dimmer, E., Huntley, R., Barrell, D., O'Donovan, C., & Apweiler, R. (2009). QuickGO: a web-based tool for Gene Ontology searching. *Bioinformatics, 25*(22), 3045–3046. DOI: 10.1093/bioinformatics/btp536

58. Mendoza, M. C., Er, E. E., & Blenis, J. (2011). The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation. *Trends in biochemical sciences, 36*(6), 320–328. DOI: 10.1016/j.tibs.2011.03.006

59. Dorsam, R. T., & Gutkind, J. S. (2007). G-protein-coupled receptors and cancer. *Nature Reviews Cancer, 7*(2), 79–94. DOI: 10.1038/nrc2069

60. Schafer K. A. (1998). The cell cycle: a review. *Veterinary pathology, 35*(6), 461–478. DOI: 10.1177/030098589803500601

61. Matthews, H. K., Bertoli, C., & de Bruin, R. A. M. (2022). Cell cycle control in cancer. Nature reviews. *Molecular cell biology, 23*(1), 74–88. DOI: 10.1038/s41580-021-00404-3

62. Oronsky, B. T., Oronsky, N., Fanger, G. R., Parker, C. W., Caroen, S. Z., Lybeck, M., & Scicinski, J. J. (2014). Follow the ATP: tumor energy production: a perspective. *Anti-cancer agents in medicinal chemistry, 14*(9), 1187–1198. DOI: 10.2174/1871520614666140804224637

63. Seyfried, T. N., Arismendi-Morillo, G., Mukherjee, P., & Chinopoulos, C. (2020). On the Origin of ATP Synthesis in Cancer. *iScience, 23*(11), 101761. DOI: 10.1016/j.isci.2020.101761

64. Kepp, O., Bezu, L., Yamazaki, T., Di Virgilio, F., Smyth, M. J., Kroemer, G., & Galluzzi, L. (2021). ATP and cancer immunosurveillance. *The EMBO journal, 40*(13), e108130. DOI: 10.15252/embj.2021108130

65. Ardito, F., Giuliani, M., Perrone, D., Troiano, G., & Lo Muzio, L. (2017). The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). *International journal of molecular medicine, 40*(2), 271–280. DOI: 10.3892/ijmm.2017.3036

66. Needham, E. J., Parker, B. L., Burykin, T., James, D. E., & Humphrey, S. J. (2019). Illuminating the dark phosphoproteome. *Science signaling, 12*(565), eaau8645. DOI: 10.1126/scisignal.aau8645

67. Johnson, J. L., Yaron, T. M., Huntsman, E. M., Kerelsky, A., Song, J., Regev, A., … Cantley, L. C. (2023). An atlas of substrate specificities for the human serine/threonine kinome. *Nature, 613*(7945), 759–766. DOI: 10.1038/s41586-022-05575-3

68. Cox, J., & Mann, M. (2007). Is proteomics the new genomics? *Cell, 130*(3), 395–398. DOI: 10.1016/j.cell.2007.07.032

69. Chevalier F. (2010). Highlights on the capacities of "Gel-based" proteomics. *Proteome science, 8*, 23. DOI: 10.1186/1477-5956-8-23

70. Lane C. S. (2005). Mass spectrometry-based proteomics in the life sciences. *Cellular and molecular life sciences, 62*(7-8), 848–869. DOI: 10.1007/s00018-005-5006-6

71. Hoffmann, E. D., & Stroobant, V. (2007). *Mass spectrometry: Principles and applications* (3rd ed). Chichester, England: Wiley.

72. Gilany, K., Moens, L., & Dewilde, S. (2010). Mass spectrometry-based proteomics in the life sciences: a review. Journal of Paramedical Sciences, 1 (1), 53-78. DOI: 10.22037/jps.v1i1.1575

73. Zubarev, R. A., & Makarov, A. (2013). Orbitrap mass spectrometry. *Analytical chemistry, 85*(11), 5288–5296. DOI: 10.1021/ac4001223

74. Michalski, A., Damoc, E., Hauschild, J. P., Lange, O., Wieghaus, A., Makarov, A., … Horning, S. (2011). Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Molecular & cellular proteomics, 10*(9), M111.011015. DOI: 10.1074/mcp.M111.011015

75. Raynie, D. E. & Watson, D. W. (2014). Understanding and ImprovingSolid-Phase Extraction. *LC-GC North America, 32*(12), 908-915.

76. Coskun O. (2016). Separation techniques: Chromatography. *Northern clinics of Istanbul, 3*(2), 156–160. DOI: 10.14744/nci.2016.32757

77. Hu, A., Noble, W. S., & Wolf-Yadlin, A. (2016). Technical advances in proteomics: new developments in data-independent acquisition. *F1000Research, 5*, F1000. DOI: 10.12688/f1000research.7042.1

78. Tabb, D. L., Eng, J. K., & Yates, J. R., III. (2001). Protein identification by SEQUEST. In P. James (ed), *Proteome research: mass spectrometry*, 125-142. Berlin, Germany: Springer. DOI: 10.1007/978-3-642-56895-4

79. Rozanova, S., Barkovits, K., Nikolov, M., Schmidt, C., Urlaub, H., & Marcus, K. (2021). Quantitative mass spectrometry-based proteomics: an overview. In K. Marcus, M. Eisenacher, B. Sitek (eds), *Quantitative methods in proteomics* (2nd ed), 85-116. New York, NY: Humana. DOI: 10.1007/978-1-0716-1024-4

CHAPTER 2. PURPOSE OF THE STUDY

This study was focused on the characterization of cell surface proteins in SKBR3/HER2+ cells, the identification of potential sources that enable signaling crosstalk and transactivation, and the assessment of the subsequent effects on signaling pathways in cancer upon targeting selected receptors with therapeutic drugs. Cell surface proteins are the gateway to cell communication, interaction, and response to stimuli;[1] therefore, they represent the most targeted proteins by more than 60% of all cancer drugs.[2] However, cancers have shown to be susceptible to developing drug resistance for a number of reasons including the presence of resistance mutations and pathway compensation through crosstalk or transactivation[3]. The discovery of novel drug targets and the development of combination therapies are essential to addressing the challenges associated with the mechanisms that lead to drug resistance,[3] therefore the study of cell surface proteins and of their interactions and impact on signaling cascades remains crucial in advancing cancer therapies.

2.1 Hypothesis

One of the most reported crosstalk and transactivation mechanisms that leads to cancer proliferation is the transactivation of receptor tyrosine kinases (RTKs), specifically EGFRs, by G-protein coupled receptors (GPCRs). Yet, the mechanism of such signaling processes remains largely unknown due to the large numbers and diverse GPCRs, difficulty in isolating and studying the GPCRs, and the intricate downstream effects that lead to aberrant cancer behavior.[4,5] This research builds on the hypothesis that comprehensive profiling of the cell-membrane proteome will enable the identification of a broad range of cell surface receptors that support potential crosstalk or transactivation activity, including GPCRs, and will lead to a more in-depth understanding of the interactions between these cell-membrane proteins to better inform the choice of therapeutic targets for achieving successful inhibition of cancer cell proliferation.

2.2 Research questions

Based on the current observations and the above hypothesis, several questions are of interest:

1. Which cell surface receptors, including GPCRs, are present on the surface of SKBR3 cells and could be potentially involved in EGFR signaling crosstalk or transactivation?

2. Which other types of cell-membrane proteins, e.g., adhesion proteins, transport proteins, CD antigens, or mutated proteins, could contribute to supporting the uncontrolled proliferation of SKBR3 cells?
3. Which growth factors secreted by cells, or protein ectodomains shed by the action of metalloproteases, could activate the existing cell-surface receptors?
4. How can we assess whether EGFR transactivation will occur or not?
5. How does the treatment with therapeutic drugs affects the above-described signaling processes and the behavior of SKBR3 cells?

2.3 Objectives

The long-term objective of this work is to investigate the mechanisms through which the cell membrane proteome of SKBR3/HER2+ breast cancer cells is implicated in driving the aberrant proliferation of these cells, to explore whether one (or more) GPCRs (or other receptors) can transactivate - or crosstalk with - the proliferative pathways activated by EGFRs, and to assess how the activation of these pathways can be altered by exposing the cells to drug inhibitors. A combination of mass spectrometry and microscopy technologies, as well as molecular biology approaches, will be used to complete the work. The specific objectives include:

1. Generate a reference library of *Homo sapiens* cell surface (CS) proteins and receptors.
2. Develop mass spectrometry and microscopy protocols for detecting and visualizing the presence of cell surface receptors and their mutated counterparts.
3. Perform global proteome analysis of SKBR3 cells, determine which proteins are identifiable in the cell membrane, and select candidate GPCRs that can crosstalk with- or transactivate the EGFR-activated pathways.
4. Induce GPCR activation with selected agonist ligands and assess the effectiveness of this activation on crosstalk/transactivation processes by performing proteomic analysis of various cellular subfractions.
5. Investigate the presence of mutated receptors and proteins in the SKBR3 cell-membrane proteome.

6. Assess the biological relevance of the data and explore the effectiveness of drug cocktails in inhibiting proliferative pathways in SKBR3 cells by targeting protein hubs that lie at the intersection of signaling pathways.

2.4 Studies performed towards the objectives

This study was focused on the SKBR3 cancer cell line, representative of the HER2+ breast cancers. For each of the six objectives above, the following steps were performed:

- **Objective 1:** An in-house CS proteins database was created and maintained regularly based on updates from major databases (ProteinAtlas, UniProt) and new publications in the field. Another database of biological processes representative of cancer hallmarks was created to interpret the role of the cell surface proteins in cancer.
- **Objective 2:** Three different cell surface enrichment methods utilizing mass spectrometry were developed and carried out to identify the surfaceome landscape of SKBR3 cells. Detected proteins and their expression levels were validated by independent methods optimized in the lab such as immunofluorescence labeling and imaging, western blotting, and parallel-reaction monitoring MS. Mutated cell surface proteins were identified based on databases developed previously in our lab (XMan).[6]
- **Objective 3:** Various SKBR3 cellular subfractions (nuclear, cytoplasmic, cell-membrane and secretome) were analyzed to generate a full proteome profile of this cell line. Based on the differentially expressed cell surface receptors under serum-free and serum-treated conditions, a target GPCR, receptor P2Y2, was selected for testing for transactivation. This target was also supported by previous literature reporting transactivation and crosstalk activity in other cancer cells. In addition, drug treatments were designed against protein targets that are part of cell signaling pathways initiated by the EGFR and known to crosstalk during cancer proliferation, such as MAPK and PI3K.
- **Objective 4:** Both ligands for the P2Y2 receptor, ATP and UTP, were tested for their transactivation potential of EGFR and ERBB2 at different timepoints by western blotting. Drugs against the above-mentioned receptors (EGFR, ERBB2, and P2Y2) were tested at different timepoints and concentrations for inhibiting the expression of the receptor and subsequently controlling for and evaluating the transactivation activity. Proteomic analysis of various treatment plans were carried out to assess the effects of different

signaling molecules and drug combinations in several steps: first, EGF stimulation of the canonical EGFR-ERBB2 pathways was inhibited by the dual-inhibitor drug, lapatinib; second, a combination of lapatinib and ipatasertib-a pan-Akt inhibitor was aimed to inhibit the crosstalk between the MAPK and PI3K proliferative pathways; and third, stimulation by ATP-a ligand for GPCRs, in the presence of inhibitory drugs, was aimed at assessing the response of alternative signaling in the cell. The cellular proteomes were analyzed by mass spectrometry, including phosphoprotein-enriched cell fractions that were representative of early signaling events, and nuclear and cytoplasmic fractions that reflected later time-point responses.

- **Objective 5:** The cell surface protein mass spectrometry data were re-analyzed to identify the mutated counterparts of the detected proteins. The mutated proteins were evaluated for their potential role in cancer survival and proliferation.
- **Objective 6:** Qualitative and quantitative proteomic comparisons between different treatment conditions were performed to assess the biological response of SKBR3 cells to the chosen stimulants and drug inhibitors, and the effectiveness of the drug cocktail in inhibiting proliferative signaling.

2.5 Significance and impact of the study

This study generated a comprehensive surfaceome profile of the SKBR3 cell line, representative of aggressive and highly metastatic HER2+ breast cancers that affect ~25% of breast cancer patients.[3] The surfaceome revealed novel potential cancer markers and drug targets, while the drug treatments uncovered alternative metabolic, proliferative, and immunological cancer cell responses that can inform the design of effective combination therapies. The mass spectrometry data contributed with new scientific knowledge that describes the composition of the cell-membrane proteome, the presence of mutations that can potentially induce unrestrained cell proliferation or the development of drug resistance, the presence of phosphorylation sites involved in- and affected by signaling processes, the existence of protein-protein interactions that may affect the fate of the cell, and the response of cancer cells to a previously unused combination of drugs used in breast cancer therapies. This knowledge provides a solid basis for the further development of novel approaches that can combat the progression of breast cancer.

2.6 References

1. Nature Education. (2014). Cells Receive and Process a Diverse Set of Chemical Signals and Sensory Stimuli. *Cell Biology for Seminars.* Retrieved from https://www.nature.com/scitable/ebooks/cell-biology-for-seminars-14760004/118245184

2. Yin, H., & Flynn, A. D. (2016). Drugging Membrane Protein Interactions. *Annual review of biomedical engineering, 18*, 51–76. DOI: 10.1146/annurev-bioeng-092115-025322

3. Nahta, R. (2019). Novel therapies to overcome HER2 therapy resistance in breast cancer. In M.R. Szewczuk, B. Qorri, M. Sambi (eds), *Current Applications for Overcoming Resistance to Targeted Therapies*, 191–221. Cham, Switzerland: Springer. DOI: 10.1007/978-3-030-21477-7_7

4. Wang Z. (2016). Transactivation of Epidermal Growth Factor Receptor by G Protein-Coupled Receptors: Recent Progress, Challenges and Future Research. *International journal of molecular sciences, 17*(1), 95. DOI: 10.3390/ijms17010095

5. Nieto Gutierrez, A., & McDonald, P. H. (2018). GPCRs: Emerging anti-cancer drug targets. *Cellular signalling, 41*, 65–74. DOI: 10.1016/j.cellsig.2017.09.005

6. Yang, X., & Lazar, I. M. (2014). XMAn: a Homo sapiens mutated-peptide database for the MS analysis of cancerous cell states. *Journal of proteome research, 13*(12), 5486–5495. DOI: 10.1021/pr5004467

CHAPTER 3. METHODS

3.1 Reagents and materials

The cell line used in this study, SKBR3, representative of breast cancer cells overexpressing the HER2 receptor, was purchased from ATCC (Manassas, VA). Its supporting reagents: McCoy's 5A (Modified) Medium, trypsin (0.25%)/EDTA (0.53 mM), Dulbecco's Phosphate Buffered Saline (D-PBS), and PenStrep solution were purchased either from ATCC (Manassas, VA) or Gibco (Carlsbad, CA). Other reagents such as Dulbecco's Phosphate Buffered Saline (D-PBS) with calcium, magnesium, fetal bovine serum (FBS), and TrypLE Select Enzyme were from Gibco (Carlsbad, CA). Kit and reagents for labeling the cell surface proteins included: Pierce Cell Surface Biotinylation and Isolation Kit, EasyPep Mini MS Sample Prep Kit, EZ-Link Alkoxyamine-PEG4-Biotin, HALT Protease Inhibitor Cocktail, Pierce Sodium meta-Periodate from Life Technologies (Carlsbad, CA), and aniline from BeanTown Chemical Corporation (Hudson, NH). Reagents used for sample processing such as NaF, Na3VO4, dithiothreitol (DTT), urea, ammonium bicarbonate (NH4HCO3), acetic acid, trifluoroacetic acid (TFA), protease and phosphatase inhibitor cocktails 2 and 3, Triton-X, and the kit for nuclear and cytoplasmic cell fractionation Lytic™ NuCLEAR™ extraction kit were from Sigma (St. Louis, MO). Sequencing grade trypsin and trypsin/LysC were purchased from Promega (Madison, WI). Growth factors and reagents such as Animal-Free Recombinant Human EGF were purchased from PeproTech (Cranbury, NJ), ATP disodium salt from Sigma (St. Louis, MO), and ATP and UTP solution from Life Technologies (Carlsbad, CA). Drugs including Lapatinib ditosylate and Ipatasertib were from Selleck (Houston, TX). For phosphopeptide enrichment, all reagents including Titansphere Phos-TiO/10 μm particles, styrene divinylbenzene (SDB)/graphitic carbon (GC) cartridges, InertSep GC beads, and lactic acid were from GL Sciences (Rolling Hills Estates, CA). Other materials used for sample preparation included Zorbax SB300-C18/5 μm particles, SPEC-PTC18/SPEC-PTSCX cleanup pipette tips, and Bond Elut-C18 cleanup cartridges purchased from Agilent Technologies (Santa Clara, CA). HPLC-grade solvents methanol and acetonitrile were purchased from Fisher Scientific (Fair Lawn, NJ). For FACS, propidium iodide was purchased from Invitrogen (Waltham, MA) and Ribonuclease A (RNase A) was from Sigma (St. Louis, MO). Cell culture materials including Nunc flasks and pipettes

were from Life Technologies (Carlsbad, CA). Water was either produced by a MilliQ Ultrapure water system (Millipore, Bedford, MA) or was distilled in-house from DI water.

3.2 Cell culture growing conditions

SKBR3 cells were cultured in McCoy's 5A medium and FBS (10%) in the presence of Penstrep (0.5%) to prevent bacterial contamination. Cells were grown in Nunc flasks in a water jacketed incubator at 37 °C and supplied with 5% CO2. For each experiment, three biological replicates were harvested from cells grown out of separate cryopreserved vials. SKBR3 cells have undergone cell authentication by short tandem repeats (STR) profiling by ATCC.

3.3 Cell surface protein labeling and processing

To isolate the cell-membrane fraction of SKBR3 cells, a combination of chemical labeling and enzymatic approaches was followed. Based on reported yields and processing times,[1-7] three methods, relying on protein isolation by biotinylation of amino groups and of glycan posttranslational modifications and affinity pulldown, as well as on tryptic shaving of receptors in cell culture, were chosen. Sulfo-NHS-SS-biotin based isolation of proteins enabled the labeling of primary amino groups at the protein N-terminal (α-amino) and Lys (ε-amino) residues, while alkoxamine-PEG4-biotin based isolation, enabled the labeling of carbohydrate moieties that are commonly encountered on the cell-surface proteins. Trypsinization of cells in culture was the least time-consuming method due to minimal processing prior and after sample collection. All reagent solutions were prepared fresh before use, and the reagent and rinse solutions that were used for biotin labeling were cooled to 4 °C before adding to the cells. The first labeling procedure involved the use of EZ-Link Sulfo-NHS-SS-Biotin (0.5 mg/mL) for labeling the protein N-terminal and Lys sidechain amino groups. Cells were rinsed twice with DPBS (+ Ca2+/ Mg2+) and then incubated at 4 °C for 30 min, in the dark, with the biotin reagent. After incubation, the biotin reagent was removed, and each flask was washed twice with 20 mL Tris quenching buffer solution (0.1 M) provided in the kit. The cells were collected by scraping in Tris-buffer (10 mL per flask), and centrifuged for 5 min at 800 x g and 4 °C. The second approach involved the labeling of cell-surface glycoproteins with EZ-Link Alkoxyamine-PEG4-Biotin (0.5 mg/mL) following protocols described by the manufacturer and in previous manuscripts [12], with some modifications. Briefly, the cells were rinsed twice with DPBS (+

Ca2+/ Mg2+) and incubated at 4 °C for 30 min, in the dark, with 20 mL sodium metaperiodate solution (1 mM, pH 6.5) to oxidize the glycan moieties of cell-surface proteins. The cells were rinsed again, twice, with DPBS (+ Ca2+/ Mg2+), and incubated with 12 mL biotin reagent solution in the presence of 10 mM aniline at 4 °C for 30 min, in the dark. After the completion of the labeling reaction, the biotin reagent was removed, and each flask was washed twice with 20 mL DPBS (+ Ca2+/ Mg2+). Cell-surface protein biotinylation of cells was visualized with an inverted Eclipse TE2000-U epi-fluorescence microscope (Nikon, Melville, NY), after staining the cells with Streptavidin Alexa Fluor™ 488 (4 μg/mL) provided by Life Technologies (Carlsbad, CA). The cells were collected by scraping in 10 mL DPBS and centrifuged for 5 min at 800 x g and 4 °C. The labeled cell pellets generated by either procedure were frozen at -80 °C for further processing or subjected to immediate lysis. The third approach consisted of shaving the cell-surface protein ectodomains with TrypLE, a reagent that contains recombinant enzymes for cell dissociation that are free of animal origin trypsin. For this procedure, the SKBR3 cells were washed twice with serum-free medium, and incubated with 10 mL TrypLE solution at 37 °C, with 5% CO2, for 2–4 min. The incubation time was short, to prevent cell detachment. The cell supernatant containing the cell-surface protein ectodomains was then collected, centrifuged for 5 min at 500 x g and 4 °C for the removal of floating cells, and frozen at – 80 °C. The samples generated through the three enrichment procedures will be referred from now on as the amine, glyco, and trypsin samples.

To isolate the cell-surface proteins of the amine-biotinylated samples, the cells were lysed with 500 μL Lysis Buffer (Pierce) supplemented with HALT protease inhibitor cocktail (5 μL), for 30 min, on ice, with intermittent vortexing and sonication. The lysate (~ 500 μL) was collected by centrifugation (15,000 x g, 5 min, 4 °C) and incubated with 250 μL NeutrAvidin beads at room temperature (RT) for 2 h, followed by 4 washes with Wash Buffer (Pierce) and 3 washes with NH_4HCO_3 (100 mM). After each wash, the beads were isolated by centrifugation (1,000 x g, 1 min). Protein recovery from the beads was performed by proteolytic digestion, on the bead, overnight, RT, in 200 μL NH_4HCO_3 (100 mM) supplemented with 25 μL trypsin/Lys C solution (10–12 μg enzyme). After centrifugation (1,000 x g), the beads were further treated with 200 μL DTT (10 mM) for 1 h at RT to recover the di-thiol, covalently bound remaining protein fragments. Both on-bead protein digest and DTT-released fractions were collected and denatured

with urea (8 M) for 1 h at 57 °C (the on-bead digest solution was also added DTT, 5 mM). After dilution with NH_4HCO_3 (100 mM) to reduce the urea concentration to <1 M, the samples were subjected to a second digestion in solution with 10 uL trypsin (~ 5 μg enzyme) for 4 h at 37 °C. After quenching the enzymatic reaction with TFA, the cell-surface peptide extracts were processed for salt and detergent disposal with SPEC-PTC18 and SPEC-SCX cartridges. Isolation of the cell-surface proteins of the biotinylated glyco samples was performed by following a similar procedure to the one that was used for the amine-labeled samples. The cell lysate was incubated with NeutrAvidin beads, the beads were treated with 200 μL DTT (45 mM, 1 h, RT, dark), and after the removal of the DTT solution by centrifugation (1,000 x g, 1 min), on-bead proteolytic digestion was performed overnight, RT, in 200 μL solution of NH_4HCO_3 (100 mM) with 25 μL trypsin/Lys C (10–12 μg enzyme) in the presence of urea (1 M). An additional 4 h digestion at RT was performed by adding to the beads 100 μL NH_4HCO_3 (50 mM) and 10 μL trypsin solution (~ 5 μg enzyme). The collected glycoprotein fraction was then processed with SPEC-PTC18 and SPEC-SCX cartridges. Control samples were prepared from unlabeled cells in the same exact manner. To isolate the cell-surface proteins of the trypsinized samples, the collected solution (~ 10 mL) was digested in a preliminary stage, overnight, with 20 μg trypsin at 37 °C, and concentrated then on a Bond Elut C18 column to remove the large volume of TrypLE solution. The sample was then reconstituted in Tris-buffer (120 μL, 50 mM, pH = 8) and denatured with urea (8 M) and DTT (5 mM) for 1 h at 57 °C. After reducing the urea concentration to <1 M with NH_4HCO_3 (100 mM), the sample was subjected to a second digestion with trypsin (~ 5 μg) for 4 h at 37 °C. After cleanup, all peptide samples were dissolved in 30 μL CH_3CN/ H_2O/TFA (95–98):(2–5):0.01 v/v for LC–MS analysis. Protein concentration measurements for either of these samples, prior to processing and proteolytic digestion, could not be performed due to limited sample availability and low abundance of the cell surface proteins in solution. Three biological replicates were labeled and processed for each cell surface enrichment method.

3.4 Cell culture treatments with drugs

To investigate the role of EGFR/ERBB2 receptors in the cancer signaling pathways and their cross-talk with other receptors, different drug treatments and their combinations were applied to

the cell culture. SKBR3 cells were grown in T175 cm^2 Nunc flaks at conditions described in Section 3.2. Upon reaching ~70–80% confluence, the cells were washed twice with DPBS and incubated in McCoy's 5A for 24 h without any supplements. The next day, the cells were washed once with McCoy's 5A, and either exposed to the drugs for 15 min or just serum-free media (SFM) for the control sample. The cells were exposed to lapatinib (10 μM) alone or to the combination of lapatinib (10 μM) and ipatasertib (1 μM). After the brief incubation with the drug or SFM, the cells were treated as follow:

- **Control**: to SFM was added FBS (10%) + EGF (10 nM)
- **Lapatinib only**: to lapatinib (10 μM) was added FBS (10%) + EGF (10 nM)
- **Drug mix**: to lapatinib/ ipatasertib (10 μM/ 1μM) was added FBS (10%) + EGF (10 nM)
- **Lapatinib only and ATP**: to lapatinib (10 μM) was added FBS (10%) + EGF (10 nM) + ATP (100 μM)
- **Drug mix and ATP**: to lapatinib/ ipatasertib (10 μM/ 1μM) was added FBS (10%) + EGF (10 nM) + ATP (100 μM)

Each of the treatments above were administered to three different T175 cm^2 flasks with cells, which were harvested at different timepoints shown below for various purposes:

- **15 min** and **30 min** by scraping the cells in cold DPBS for investigating the phosphosignaling activity altered by the drugs
- **36 h** by trypsinization for investigating the changes in the proteome after exposure to the drugs

For each treatment (5 in total), 3 independent biological replicates were produced, resulting in a total of 15 samples for each timepoint (15 min, 30 min, 36 h). The 15 min and 30 min samples were subjected to processing protocols for phosphopeptide enrichment as described in **Section 3.7**, and the 36 h samples were subjected to processing protocols for cell fractionation as described in **Section 3.6**.

3.5 FACS analysis

FACS analysis was performed for all samples under the 36 h treatments to observe the cell states across the cell cycle upon different drug regimens. Upon harvesting, the cells were fixed in 70% ethanol and preserved at -20 °C. They were later washed once in DPBS and stained in a freshly

prepared solution of 0.2 mg/mL propidium iodide, 0.2 mg/mL RNAase, and 0.1% Triton X-100 in DPBS. After 30 min of incubation in the dark at room temperature, the samples were analyzed by a FACSCalibur flow cytometer (BD Biosciences, San Jose, CA). All three biological replicates were subject to independent FACS sample preparation and analysis.

3.6 Cell fractionation: nuclear and cytoplasmic

To retrieve cytoplasmic and nuclear cell extracts, the Cell Lytic™ NuCLEAR™ kit was used along with protocols recommended by the manufacturer. The cells were lysed in the presence of a hypotonic buffer supplemented with DTT (1 mM), protease inhibitor cocktail (1% of total lysate volume), and phosphatase inhibitors (Na_3VO_4 and NaF, 1 mM each) on ice. IGEPAL CA-630 solution (0.6% of total lysate volume) was added to aid the lysis process for a better separation between the cytoplasmic and nuclear fraction, without disturbing the nuclear membrane. The cytoplasmic fraction was retrieved after centrifugation at 10,000 x g for 1 min at 4 °C. Then, the nuclear pellet was washed once, and incubated in a high salt extraction buffer for 30 min at 4 °C under constant mixing in a vortex and brief sonication intervals. The nuclear fraction was retrieved after centrifugation at 20,000 x g for 5 min at 4 °C. Protein concentrations in each extract were measured with the Bradford assay (SmartSpec Plus spectrophotometer, Bio-Rad, Hercules, CA) using the Bradford dye reagent and bovine standards (BioRad, Hercules, CA). For MS analysis, the extracts were reduced with urea (8 M)/DTT (5 mM) for 1 h at 56 °C, diluted with NH_4HCO_3 (50 mM), and digested with sequencing grade trypsin at extract: trypsin ratio of 50:1 w/w overnight at 37 °C. The following day, the reaction was quenched with TFA (1% of total volume). Prior to LC/MS analysis, the cell extract peptide mixtures were disposed of salts and detergents with SPEC-PTC18 and SPEC-SCX cartridges. The evaporated samples were reconstituted in CH_3CN/H_2O/TFA solution (98:2:0.01 v/v) and frozen at -80 °C until LC/MS analysis.

3.7 Phosphopeptide enrichment

Cells were lysed by sonication on ice for 5 min in 1 min intervals in a freshly prepared solution of Tris (50 mM) pH=8.2/ NaCl (75 mM)/ urea (8 M) supplemented with DTT (1 mM) and protease and phosphatase inhibitors (1% and 2% of total lysate volume respectively). The protein extract was collected by centrifugation at 10,000 x g for 10 min at 4 °C and its concentration was

measured with Bradford assay. 2 mg protein extract was denatured at 56 °C for 1 h, diluted with NH_4HCO_3 (50 mM), and digested with sequencing grade trypsin at extract: trypsin ratio of 50:1 w/w overnight at 37 °C. The following day, the reaction was quenched with TFA (1% of total volume) and the peptide mixture was cleaned-up with a Bond Elut-C18 cartridge (3 mL) under vacuum. The evaporated peptide mixture was dissolved in CH_3CN:H_2O:TFA (50:50:1%) and processed with inhouse-made spin tips packed with 12 mg TiO_2 beads. Phosphopeptides were adsorbed under acidic conditions (pH <3) in the presence of lactic acid and eluted in basic conditions (pH >10) with NH_4OH solution (5%) and $(CH_2)_4NH$ (5%). The eluted phosphopeptides were cleaned-up with inhouse-made spin tips containing 4 mg of GC and C18 beads respectively. The phosphopeptides were eluted in a strong organic solvent solution CH_3CN/H_2O/TFA (80:20:0.1 v/v) and brought to dryness at room temperature by a vacuum centrifuge, after which they were immediately frozen at -80 °C. The peptides were dissolved in CH_3CN/H_2O/TFA solution (1:99:0.01 v/v) right before LC/MS analysis.

3.8 Data-dependent acquisition mass spectrometry (DDA-MS) analysis

The peptide samples were analyzed with an EASY-nLC 1200 UHPLC system (ThermoFisher Scientific) by using a heated nano-electrospray ionization (ESI) source (2-2.2 kV) and a Q Exactive hybrid quadrupole-Orbitrap mass spectrometer (ThermoFisher Scientific).

3.8.1. Liquid chromatography (LC)-DDA-MS analysis for cell surface and whole cell extracts

An EASY-Spray column ES802A (150 mm long, 75 µm i.d., 3 µm C18/silica particles, ThermoFisher Scientific) for cell surface samples, and an EASY-Spray column ES902 (250 mm long, 75 µm i.d., 2 µm C18/silica particles, ThermoFisher Scientific) for nuclear/cytoplasmic samples was used at flow rates of 250 nL/min. The mobile phases were prepared from H_2O:CH_3CN:TFA, and mixed in ratios of 96:4:0.01 v/v for mobile phase A and 10:90:0.01 v/v for B. For cell surface samples, a separation gradient of 85 m was used with eluent B concentration increasing from 3% to 30% (5-65 min), 45% (65-72 min), 60% (72-73 min), and 90% (73-74 min), where it was kept for 5 min, and then decreased to a final concentration of 3%.[7] For nuclear/cytoplasmic fractionated samples, a separation gradient of 125 min was used with eluent B concentration increasing from 7% to 30% (2-107 min), 45% (107-109 min), 60%

(109-110 min), and 90% (110-111 min), where it was kept for 10 min, and then decreased to a final concentration of 7%.[8] The MS data were acquired over a range of 400-1,600 m/z with resolution set to 70,000, AGC target to 3E6, and maximum IT to 100 ms. Data-dependent MS2 acquisition (dd-MS2) was enabled by using higher-energy collisional dissociation (HCD), isolating the precursor ions with a width of 2.4 m/z, and fragmenting them at 30% normalized collision energy (NCE). dd-MS2 acquisition parameters were set to resolution 17,500, AGC target 1E5 (minimum AGC target 2E3 and intensity threshold 4E4), maximum IT 50 ms, and loop count 20. Charge exclusion was enabled for unassigned and +1 charges, apex trigger was set to 1 to 2 s, dynamic exclusion lasted for 10 s for chromatographic peak widths of 8 s, and features of isotope exclusion and preferred peptide match were turned on. All samples were run three times, which were considered as technical replicates.

3.8.2. DDA-MS analysis for enriched phosphorylated peptides

The phosphopeptide samples were analyzed with an EASY-Spray column ES902 (250 mm long, 75 μm i.d., 2 μm C18/silica particles, ThermoFisher Scientific) at a flow rate of 250 nL/min. Mobile phase A and B were run for 85 min based on the following eluent B concentration: 5% to 30% (5-65 min), 45% (65-72 min), 60% (72-73 min), and 90% (73-74 min), kept for 5 min, and then decreased to a final concentration of 5%. The MS data were acquired over a range of 400-1,600 m/z with resolution set to 70,000, AGC target to 3E6, and maximum IT to 100 ms. For data-dependent MS2 acquisition the precursor ions were isolated with a width of 2.4 m/z, and subjected to fragmentation in a stepped NCE fashion at 20, 25, and 30%, which was shown to increase the number of identifiable phosphopeptides and the confidence of phosphosite localization.[9] dd-MS2 acquisition parameters were set to resolution 17,500, AGC target 1E5 (minimum AGC target 2E3 and intensity threshold 4E4), maximum IT 50 ms, and loop count 20. Charge exclusion was enabled for unassigned and +1 charges, apex trigger was set to 1 to 2 s, dynamic exclusion lasted for 10 s for chromatographic peak widths of 8 s, and features of isotope exclusion and preferred peptide match were turned on.

3.9 Targeted mass spectrometry: Parallel reaction monitoring

Low abundant cell surface proteins such as GPCRs and differentially expressed proteins in the cell surface or drug treated samples were validated by parallel reaction monitoring (PRM). Peptide

selection for PRM analysis was based on a lab-developed framework considering PTMs, XCorr scores, charge states, retention time (RT) relative SDs, and spectra quality. The selected peptides of each sample were searched in a 20 min RT window following the same separation gradient as the original DDA-MS sample analysis. The precursor ions were isolated with a width of 2.0 m/z, and fragmented at 30% normalized collision energy with PRM parameters set as follows: resolution 35,000, AGC target 2E5, and maximum IT 110 ms. The PRM data were processed by the Skyline 20.2 software[10] by using a mass spectral library generated from the DDA-MS searches of the respective samples. For the cell surface samples, the transition settings included ion types: b and y, from ion 2 to last ion with precursor charges of 2 and 3, and ion charges of 1 and 2. For the nuclear/cytoplasmic fractionated samples, the transition settings included ion types: b and y, from ion 1 to last ion with precursor charges of 2, 3 and 4, and ion charges of 1, 2 and 3. The 5 or 10 most intense product ions were picked from the filtered product ions. The presence of a peptide was considered validated when the peptide displayed a minimum of 5 transitions and a dot product (dotp) score >0.9.

3.10 Mass spectrometry data processing

3.10.1. Data processing for cell surface, nuclear/cytoplasmic, and phosphorylated proteins

The MS data were processed by the Proteome Discoverer 2.4/2.5 package (Thermo Fisher Scientific, Waltham, MA) and searched with Sequest HT against a *Homo sapiens* database of 24,433 reviewed, non-redundant UniProt protein sequences (March 2019 download) for the cell surface samples and a *Homo sapiens* database of 20,398 reviewed, non-redundant UniProt protein sequences (August 2022 download) for all the other samples. The processing workflow spectrum filter was set for a peptide precursor mass range of 400-5,000 Da, and the Sequest HT node parameters allowed for the selection of fully tryptic peptides (6-144 aa length) with maximum two missed cleavages, 15 ppm precursor ion tolerance, b/y/a ion fragments with 0.02 Da tolerance, and dynamic modifications on Met (15.995 Da/oxidation) for all samples. For cell surface samples, the following modifications were added: dynamic modification of the N-terminal amino acids (42.011 Da/acetyl), and dynamic modification of the Lys residues and protein N-termini (87.998 Da) for the amine-reactive biotin-labeled samples only. For phosphopeptide enriched samples, dynamic modification of Ser/Thr/Tyr residues (79.9799 Da) was added. Also, for the phosphopeptide enriched samples, the ptm-RS node algorithm was used

for modification site localization allowing only for phosphosites assignments with ≥75% probability. The PSM validator node used target/decoy concatenated databases, with input data of maximum DeltaCn 0.05 and maximum rank 1, and FDR targets of 0.01 (strict) and 0.03 (relaxed). Additional parameters were set in the consensus workflow for both peptide and protein levels. The peptide group modification site probability threshold was set to 75, and the peptide validator node to automatic with peptide level error rate control. The FDRs were set to 0.01 (high) and 0.03 (medium) for PSMs, peptides, and proteins, and the strict parsimony principle was enabled for protein grouping. Lastly, only peptides of at least medium confidence and proteins matched by only rank 1 peptides were retained in the peptide/protein filter node. The peptides were counted only for top scoring proteins.

3.10.2. Data processing for mutated proteins

The LC-MS/MS raw files of the cell surface protein samples were re-processed by Sequest HT engine against a Homo sapiens database of 20,433 reviewed, non-redundant UniProt protein sequences (March 2019 download) and an XMan database of 2,539,031 human mutated, non-redundant sequences.[11] The mass range for the parent ion was 400-5,000 Da with 15 ppm mass tolerance, and 0.02 Da fragment mass tolerance. Only top scoring proteins, rank 1 peptides, and equal or higher than 1 XCorr score spectra were accepted. Each peptide was allowed up to two missed tryptic miscleavages and up to four dynamic modifications, which included Met oxidation (+15.995 Da) for all fractions, and modification of Lys (+87.998 Da) for the cell membrane fractions isolated by biotinylation of amine groups. FDR values at peptide and protein level were 0.03 (relaxed) and 0.01 (strict). For each isolation method of the cell membrane fraction, a single result file was generated including all three biological replicates and both cell culture conditions (serum-free and serum-supplemented). For simplicity, they will be referred as glyco, amine, and trypsin. The proteins were assigned as cell membrane proteins based on the cell surface in-house database described in **Section 3.12**.

3.11 Quantitation and statistical analysis

Each of the three biological replicates was analyzed three times by LC-MS/MS and one multiconsensus protein and peptide report was generated from the three technical replicates. The relationship between any two sets of biological replicates was evaluated at the peptide and PSMs

level by the Pearson correlation coefficient "r" between RTs, XCorr scores, and spectral counts. For assessing changes in protein expression, data normalization and quantitation was performed either based on spectral counting or peak areas.

3.11.1. Spectral counting-based quantitation

For cell surface and nuclear/cytoplasmic samples, when quantitation was performed based on the PSM counts, missing PSM values were handled by adding one spectral count to each protein from the dataset. Normalization was performed at the global level by averaging the total spectral counts of the six samples taken into consideration (e.g., three SF and three ST samples for cell surface proteins, or three EGF and three lapatinib samples for nuclear or cytoplasmic proteins) and using the resulting average as a correction factor for adjusting the counts of individual proteins in each sample. For cell surface proteins, a second normalization was performed at the cell-surface protein level by calculating a correction factor based on the spectral counts of only a short list of endogenous cell-surface proteins.[7] Differentially expressed proteins were selected by calculating the Log2 values of the spectral count ratios of the two datasets and using a two-tailed t-test. Proteins matched by either two unique peptides (for cell surface proteins) or three unique peptides (for the nuclear/cytoplasmic proteins of drug-treated cells) with fold change (FC) ≥ 2 and p-value <0.05 were considered for discussion. Benjamini-Hochberg adjusted p-values were calculated for the nuclear and cytoplasmic protein comparisons of drug-treated cells.[12]

3.11.2. Peak area-based quantitation

For the comparisons involving the nuclear/cytoplasmic fractions of drug-treated cells, quantitation was performed based on a Proteome Discoverer 2.5 template workflow of six samples (i.e., three control and three drug-treated samples). Database searching and data processing was performed with a Sequest HT and Percolator-based workflow that uses a semi-supervised learning algorithm and q-value based assessment of statistical significance to differentiate between correct and incorrect PSMs. The FDRs were set at 0.01 (strict) and 0.05 (relaxed) for PSMs. The processing workflow contained the Minora Feature Detector algorithm which detects and matches chromatographic peaks across LC/MS runs and links them to PSMs (PSM confidence was set to High). In the consensus workflow, to account for peptide retention time shifts during multiple sample-runs on the same column, chromatographic alignment was

performed with the Feature Mapper node, enabling a maximum retention time shift of 15 min and maximum mass tolerance of 15 ppm. Protein quantitation was based on summed peptide abundances and pairwise ratio calculations (i.e., drug treatment replicate 1/control replicate 1). Normalization was performed based on total peptide abundance, and modified peptides were excluded from the pairwise ratio-based protein quantifications. Low abundance resampling was chosen as the mode of imputation for missing values. Differentially expressed proteins were selected based on the log2 values of the generated ratios (treatment/control) by applying a t-test. Proteins matched by three unique peptides with FC ≥2 and abundance ratio p-value <0.05 were considered for discussion. Adjusted p-values accounting for multiple testing were calculated based on the Benjamini-Hochberg correction method.

3.12 In-house databases

For analyzing the cell surface samples, an in-house database of cell-membrane proteins was built by extracting relevant entries from the UniProtKB/Swiss-Prot database based on controlled vocabulary terms,[13] from the Human Protein Atlas (HPA) Cellular and Organelle Proteome,[14,15] and from the scientific literature.[5-7,16,17] GeneCards[18] and UniProt[13] were used to assess protein functionality and their classification in four main categories. In addition to the in-house database created for cell surface protein identification, another in-house database was created based on controlled GO annotation vocabulary terms of biological processes representative of the 10 cancer hallmarks.[19] An additional level of data curation was added to the database by including proteins associated with the cancer hallmarks reported in the literature[20-26] and listed in the COSMIC (v94) Cancer Gene Census catalogue (CGC).[27] The detected cell membrane proteins in SKBR3 were then aligned to the above-mentioned curated lists. For the phosphopeptide enriched samples, an in-house database of known phosphosites was assembled from several databases with reported phosphorylated sites in the human proteome such as PhosphoSitePlus,[28] PTMsigDB,[29] and phosphoELM,[30] and utilizing the keyword search in UniProt[13] to extract the reviewed phosphoproteins and their modified residues during the period of September 2021-October 2022. Lastly, another database containing all the detected phosphorylated peptides from all the treatments and replicates was assembled for the in-house analysis and assessment of phosphopeptide enrichment, phosphosite distribution, and specific phosphosite localization probabilities.

3.13 Data analysis, interpretation and visualization

For all the samples, STRING 11.5 was used to build protein–protein interactions (PPI) networks and assess GO enrichment in biological processes,[31] with interaction score confidences set to medium/high and enrichment FDR <0.05. Cytoscape 3.8.2 and 3.9.1 software[32] was utilized to visualize protein networks based on interactomics data exported from STRING. Cancer drug targets were extracted from the DrugBank database in September 2021.[33] RAWGraphs-an open source data visualization framework[34]-was used for building the dendograms, and Protter was used for visualizing the location of a protein relative to the cell-membrane bilayer.[35] Phosphorylated proteins were analyzed by the KEA3 software to assess the top enriched kinases in each sample.[36] Signaling networks by phosphorylation were created with SIGNOR3.0.[37] All other figures were generated either with the Proteome Discoverer software or with Microsoft Excel.

3.14 Validation method: immunofluorescence (IF) microscopy

Materials for IF labeling and imaging, including antibodies, are summarized in **Table 3-1**. The cells were grown in 8-well chamber slides (MatTek, Ashland, MA), rinsed once with DPBS, and fixed with either chilled methanol (-20 °C) for 5 min, or PFA solution (2%) at RT for 15 min. If permeabilized, they were incubated with Triton X-100 solution (0.5% in PBS) at RT for 5 min. The cells were blocked with BSA solution (5% in PBS) at RT for 1 h, followed by three washes with DPBS for 5 min. Incubation with primary antibody was performed overnight at 4 °C, while incubation with secondary antibody was done at RT for 1 h in the dark. Cells were washed after each incubation step with DPBS (5 min, 5 times). Controls were included for each experiment by omitting the primary antibody. At the end, the cells were stained with DAPI solution (1 µg/mL, 5 min, RT, dark) or DAPI mounting media, and left to cure for 24 h at RT in the dark. Images were acquired either with an inverted epi-fluorescence Eclipse TE2000-U microscope (Nikon Instruments Inc, Melville, NY) with a 20X air objective, or by confocal scanning with SoRa mode with a Nikon Eclipse Ti2 with a 40X water objective. The images were processed with the Nikon Denoise.ai and NIS-Elements AR Analysis 5.11.01 software packages.

Table 3-1: Materials and antibodies used for immunofluorescence imaging to visualize the receptors and the differentially expressed proteins in the cell surface samples.

Product	**Supplier**
Bovine Serum Albumin Fraction V	Sigma (St. Louis, MO)
Triton X-100	Sigma (St. Louis, MO)
Paraformaldehyde, reagent grade, crystalline	Sigma (St. Louis, MO)
8-well chamber slide	MatTek (Ashland, MA)
DAPI powder	Cell Signaling Technology (Danvers, MA)
ProLong™ Diamond Antifade Mountant with DAPI solution	Life Technologies (Carlsbad, CA)
Antibodies	**Supplier**
ERBB2 (sc-08) Mouse mAb	Santa Cruz (Dallas, TX)
EGFR (sc-120/sc-101) Mouse mAb	Santa Cruz (Dallas, TX)
mIgG CFL 488 (sc-516176)	Santa Cruz (Dallas, TX)
mIgG CFL 594 (sc-516178)	Santa Cruz (Dallas, TX)
E-cadherin (AF748) Goat pAb	RnD Systems (Minneapolis, MN)
Anti-Goat IgG NL557 (NL001)	RnD Systems (Minneapolis, MN)
P2Y2 (NBP2-94356) Rabbit pAb	Novus Biologicals (Centennial, CO)
ATP5A (NBP2-92928) Rabbit pAb	Novus Biologicals (Centennial, CO)
UQCRC2 (NBP2-97264) Rabbit pAb	Novus Biologicals (Centennial, CO)
Anti-Rabbit IgG (H+L) [DyLight 488] (NBP1-72944)	Novus Biologicals (Centennial, CO)

3.15 Validation method: western blotting

Materials for western blotting, including antibodies, are summarized in **Table 3-2**. The cells were lysed according to their respective treatments (nuclear/cytoplasmic fractions refer to **Section 3.6**, or phosphopeptide enriched fractions refer to **Section 3.7**), or with RIPA buffer supplemented with protease inhibitors (1% of total lysate volume) and phosphatase inhibitors (Na_3VO_4 and NaF, 1 mM each) by sonication on ice. Protein concentrations were measured by the Bradford assay and calculations were performed to load 32 μg protein sample (20 μL) in each lane. The samples were diluted twice in Laemmli sample buffer (2X), RIPA buffer was added to adjust for the total volume, and DTT (100 mM) was used to reduce the disulfide bonds. The samples were boiled for 5 min at 95 °C and immediately loaded in the precast stain-free gel with a protein standard ladder. The SDS-PAGE was run at 200 V for ~ 40-50 min with 1X running buffer prepared from 10X stock Tris (25 mM)/Glycine (192 mM) solution supplemented

with 0.01% SDS (pH ~ 8.3) in a Mini-PROTEAN® electrophoresis cell (Bio-Rad, Hercules, CA). Afterwards, the gel was imaged with a ChemiDoc ™ Imaging System (Bio-Rad, Hercules, CA) to validate the proper separation during SDS-PAGE. Before transfer, all materials including the gel, filter papers, and low fluorescence PVDF membranes were equilibrated for 15 min in 1X transfer buffer prepared from 10X stock Tris (25 mM)/Glycine (192 mM) solution supplemented with 20% (v/v) methanol (pH ~8.3). The transfer stack was assembled based on manufacturer's directions, submerged in ice-cold transfer buffer, and ran at 100 V for 1 h. After transfer, both the gel and the membrane were imaged to verify the protein transfer. The membrane was incubated in blocking solution (3% BSA, 2% NFDM) for 1 h at RT, washed 3 times for 5 min in TBS-Tween-20 (0.1%) wash buffer, and incubated overnight at 4 °C with primary antibody solution (5% BSA). The following day, the membranes were washed 5 times for 5 min in wash buffer and incubated with HRP-linked secondary antibodies for 1 h at RT, followed by the same wash regimen. Membranes were imaged for total protein normalization, then briefly (1-2 min) incubated with ECL substrates for chemiluminescence imaging. Images were acquired on the same BioRad ChemiDoc imaging system using the auto-optimal feature (2x2 bins) or under different manual exposure times ranging from 30 s to 5 min depending on the signal intensity of each individual sample.

Table 3-2: Materials and antibodies used for western blotting to test the cross-talk activities and to validate the differentially expressed proteins in the cytoplasmic/nuclear fractions.

Product	Supplier
Immun-Blot LF PVDF Membrane and Filter papers	BioRad (Hercules, CA)
Mini Trans-Blot Filter paper	BioRad (Hercules, CA)
Clarity Western ECL Substrate	BioRad (Hercules, CA)
2X Laemmli Sample buffer	BioRad (Hercules, CA)
Tween20	BioRad (Hercules, CA)
PrecisionPlus Protein Standards (ladder)	BioRad (Hercules, CA)
Mini-Protean TGX Stain-Free Gels	BioRad (Hercules, CA)
Blotting-Grade blocker (nonfat dry milk - NFDM)	BioRad (Hercules, CA)
Pierce RIPA buffer	Life Technologies (Carlsbad, CA)
Glycine, electrophoresis grade, 99%+	Thermo Fisher Scientific (Haverhill, MA)
Glycine, ≥99% (HPLC)	Sigma (St. Louis, MO)
Bovine Serum Albumin Fraction V	Sigma (St. Louis, MO)
Sodium chloride	Sigma (St. Louis, MO)

Trizma base	Sigma (St. Louis, MO)
Trizma hydrochloride	Sigma (St. Louis, MO)
Sodium dodecyl sulfate (SDS)	Sigma (St. Louis, MO)
Antibodies	**Supplier**
EGFR pY1173 (4407S) Rabbit mAb	Cell Signaling Technology (Danvers, MA)
Total EGFR (11862S) Rabbit mAb	Cell Signaling Technology (Danvers, MA)
EGFR pY1068 (11862S) Rabbit mAb	Cell Signaling Technology (Danvers, MA)
p44/42 MAPK (Erk1/2) pT202/pY204 (9101S) Rabbit pAb	Cell Signaling Technology (Danvers, MA)
Total p44/42 MAPK (Erk1/2) (9102S) Rabbit pAb	Cell Signaling Technology (Danvers, MA)
HER2/ERBB2 pY1248 (2247S) Rabbit pAb	Cell Signaling Technology (Danvers, MA)
Total HER2/ERBB2 (29D8) (2165S) Rabbit mAb	Cell Signaling Technology (Danvers, MA)
CD82 (D7G6H) (12439S) Rabbit mAb	Cell Signaling Technology (Danvers, MA)
B7-H4/VTCN1 (D1M8I) XP® (14572t) Rabbit mAb	Cell Signaling Technology (Danvers, MA)
PDCD4 (D29C6) XP® (9535T) Rabbit mAb	Cell Signaling Technology (Danvers, MA)
Topoisomerase II (D10G9) XP® (12286T) Rabbit mAb	Cell Signaling Technology (Danvers, MA)
Anti-rabbit IgG, HRP-linked Antibody	Cell Signaling Technology (Danvers, MA)

3.16 References

1. Elschenbroich, S., Kim, Y., Medin, J. A., & Kislinger, T. (2010). Isolation of cell surface proteins for mass spectrometry-based proteomics. *Expert review of proteomics, 7*(1), 141–154. DOI: 10.1586/epr.09.

2. Kuhlmann, L., Cummins, E., Samudio, I., & Kislinger, T. (2018). Cell-surface proteomics for the identification of novel therapeutic targets in cancer. *Expert review of proteomics, 15*(3), 259–275. DOI: 10.1080/14789450.2018.1429924

3. Li, Y., Wang, Y., Mao, J., Yao, Y., Wang, K., Qiao, Q., Fang, Z., & Ye, M. (2019). Sensitive profiling of cell surface proteome by using an optimized biotinylation method. *Journal of proteomics, 196*, 33–41. DOI: 10.1016/j.jprot.2019.01.015

4. Wollscheid, B., Bausch-Fluck, D., Henderson, C., O'Brien, R., Bibel, M., Schiess, R., … Watts, J. D. (2009). Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nature biotechnology, 27*(4), 378–386. DOI: 10.1038/nbt.1532

5. Kalxdorf, M., Gade, S., Eberl, H. C., & Bantscheff, M. (2017). Monitoring Cell-surface N-Glycoproteome Dynamics by Quantitative Proteomics Reveals Mechanistic Insights into Macrophage Differentiation. *Molecular & cellular proteomics, 16*(5), 770–785. DOI: 10.1074/mcp.M116.063859

6. Bausch-Fluck, D., Hofmann, A., Bock, T., Frei, A. P., Cerciello, F., Jacobs, A., … Wollscheid, B. (2015). A mass spectrometric-derived cell surface protein atlas. *PloS one, 10*(3), e0121314. DOI: 10.1371/journal.pone.0121314

7. Karcini, A., & Lazar, I. M. (2022). The SKBR3 cell-membrane proteome reveals telltales of aberrant cancer cell proliferation and targets for precision medicine applications. *Scientific reports, 12*(1), 10847. DOI: 10.1038/s41598-022-14418-0

8. Ahuja, S., & Lazar, I. M. (2021). Systems-Level Proteomics Evaluation of Microglia Response to Tumor-Supportive Anti-Inflammatory Cytokines. *Frontiers in immunology, 12*, 646043. DOI: 10.3389/fimmu.2021.646043

9. Diedrich, J. K., Pinto, A. F., & Yates, J. R., 3rd (2013). Energy dependence of HCD on peptide fragmentation: stepped collisional energy finds the sweet spot. Journal of the American Society for Mass Spectrometry, 24(11), 1690–1699. DOI: 10.1007/s13361-013-0709-7

10. MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., … MacCoss, M. J. (2010). Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics, 26*(7), 966–968. DOI: 10.1093/bioinformatics/btq054

11. Flores, M. A., & Lazar, I. M. (2020). XMAn v2-a database of Homo sapiens mutated peptides. *Bioinformatics, 36*(4), 1311–1313. DOI: 10.1093/bioinformatics/btz693

12. Benjamini, Y. and Hochberg, Y. (1995), Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*, 289-300. DOI: 10.1111/j.2517-6161.1995.tb02031.x

13. UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic acids research, 47*(D1), D506–D515. DOI: 10.1093/nar/gky1049

14. Thul, P. J., Åkesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait Blal, … Lundberg, E. (2017). A subcellular map of the human proteome. *Science, 356*(6340), eaal3321. DOI: 10.1126/science.aal3321

15. Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, … Pontén, F. (2015). Proteomics. Tissue-based map of the human proteome. *Science, 347*(6220), 1260419. DOI: 10.1126/science.1260419

16. Bausch-Fluck, D., Goldmann, U., Müller, S., van Oostrum, M., Müller, M., Schubert, O. T., & Wollscheid, B. (2018). The in silico human surfaceome. *Proceedings of the National Academy of Sciences of the United States of America, 115*(46), E10988–E10997. DOI: 10.1073/pnas.1808790115

17. Ramilowski, J. A., Goldberg, T., Harshbarger, J., Kloppmann, E., Lizio, M., Satagopam, V. P., … Forrest, A. R. (2015). A draft network of ligand-receptor-mediated multicellular signalling in human. *Nature communications, 6*, 7866. DOI: 10.1038/ncomms8866

18. Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., … Lancet, D. (2016). The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Current protocols in bioinformatics, 54*, 1.30.1–1.30.33. DOI: 10.1002/cpbi.5

19. Lazar, I. M., Karcini, A., & Haueis, J. R. S. (2022). Mapping the cell-membrane proteome of the SKBR3/HER2+ cell line to the cancer hallmarks. *PloS one, 17*(8), e0272384. DOI: 10.1371/journal.pone.0272384

20. Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of Cancer. Cell, 100(1), 57-70. DOI: 10.1016/S0092-8674(00)81683-9

21. Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of Cancer: The Next Generation. Cell, 144(5), 646-674. DOI: 10.1016/j.cell.2011.02.013

22. Munkley, J., & Elliott, D. J. (2016). Hallmarks of glycosylation in cancer. *Oncotarget, 7*(23), 35478–35489. DOI: 10.18632/oncotarget.8155

23. Muriithi, W., Macharia, L. W., Heming, C. P., Echevarria, J. L., Nyachieo, A., Filho, P. N., & Neto, V. M. (2020). ABC transporters and the hallmarks of cancer: roles in cancer aggressiveness beyond multidrug resistance. *Cancer biology & medicine, 17*(2), 253–269. DOI: 10.20892/j.issn.2095-3941.2019.0284

24. Abcam. (2020). Studying hallmarks of cancer. Retrieved from https://www.abcam.com/cancer/studying-hallmarks-of-cancer

25. Pickup, M. W., Mouw, J. K., & Weaver, V. M. (2014). The extracellular matrix modulates the hallmarks of cancer. EMBO reports, 15(12), 1243–1253. DOI: 10.15252/embr.201439246

26. Gutschner, T., & Diederichs, S. (2012). The hallmarks of cancer: a long non-coding RNA point of view. *RNA biology, 9*(6), 703–719. DOI: 10.4161/rna.20481

27. Sondka, Z., Bamford, S., Cole, C. G., Ward, S. A., Dunham, I., & Forbes, S. A. (2018). The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. *Nature Reviews Cancer, 18*(11), 696–705. DOI: 10.1038/s41568-018-0060-1

28. Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., & Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic acids research, 43*(Database issue), D512–D520. DOI: 10.1093/nar/gku1267
Datasets (non-commercial use) were downloaded from https://www.phosphosite.org/staticDownloads

29. Krug, K., Mertins, P., Zhang, B., Hornbeck, P., Raju, R., Ahmad, R., … Mani, D. R. (2019). A Curated Resource for Phosphosite-specific Signature Analysis. *Molecular & cellular proteomics, 18*(3), 576–593. DOI: 10.1074/mcp.TIR118.000943

30. Diella, F., Gould, C. M., Chica, C., Via, A., & Gibson, T. J. (2008). Phospho.ELM: a database of phosphorylation sites--update 2008. *Nucleic acids research, 36*(Database issue), D240–D244. DOI: 10.1093/nar/gkm772

31. Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., … Mering, C. V. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research, 47*(D1), D607–D613. DOI: 10.1093/nar/gky1131

32. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., … Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research, 13*(11), 2498–2504. DOI: 10.1101/gr.1239303

33. Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., … Wilson, M. (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic acids research, 46*(D1), D1074–D1082. DOI: 10.1093/nar/gkx1037

34. Mauri, M., Elli, T., Caviglia, G., Uboldi, G., & Azzi, M. (2017). RAWGraphs: A visualisation platform to create open outputs. In CHItaly'17, Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter 28:1–28:5. DOI: 10.1145/3125571.3125585

35. Omasits, U., Ahrens, C. H., Müller, S., & Wollscheid, B. (2014). Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics, 30*(6), 884–886. DOI: 10.1093/bioinformatics/btt607

36. Kuleshov, M. V., Xie, Z., London, A. B. K., Yang, J., Evangelista, J. E., Lachmann, A., … Ma'ayan, A. (2021). KEA3: improved kinase enrichment analysis via data integration. *Nucleic acids research, 49*(W1), W304–W316. DOI: 10.1093/nar/gkab359

37. Lo Surdo, P., Iannuccelli, M., Contino, S., Castagnoli, L., Licata, L., Cesareni, G., & Perfetto, L. (2023). SIGNOR 3.0, the SIGnaling network open resource 3.0: 2022 update. *Nucleic acids research, 51*(D1), D631–D637. DOI: 10.1093/nar/gkac883

CHAPTER 4:

The SKBR3 Cell-Membrane Proteome: Role in Aberrant Cancer Cell Proliferation and Resource for Precision Medicine Applications

Arba Karcini and Iulia M. Lazar

BioRxiv, doi.org/10.1101/2021.10.24.465642

Scientific Reports 2022, 12:10847, doi.org/10.1038/s41598-022-14418-0

Department of Biological Sciences, Virginia Tech
1981 Kraft Drive, Blacksburg, VA 24061, USA
*Corresponding author: malazar@vt.edu

Author contributions

Conceptualization, project administration: IML

Experimental design: IML, AK

Experiments and data acquisition: AK

Data analysis: AK, IML

Data visualization: AK

Manuscript writing: IML, AK

CHAPTER 4. CELL MEMBRANE PROTEOME

4.1 Introduction

Breast cancer is a common form of cancer that continues to lead, even in the present day, to a large number of deaths among women worldwide.[1] The different breast cancer subtypes are defined based on the presence of ER, HER2 or PR receptors, whether alone or in combination. HER2+ and triple negative breast cancers have the worst prognosis due to the fact that some HER2+ tumors are either non-responsive or develop resistance to anti-HER2 therapies, while triple negative cancers are non-responsive to hormonal therapies or drugs that target HER2 receptors.[1,2] As a result, focus has been placed on the development of novel therapeutic approaches that rely either on the use of various drug cocktails and treatment regimens that target multiple receptors or compensatory and downstream crosstalk signaling pathways of HER2, or, more recently, on triggering immune system responses that attack the cancer cells.[2]

The heavy interest in the study of cancer cell-membrane receptors has been fueled by their central role in initiating cellular signaling cascades that lead to aberrant cell proliferation, as well as by their potential as cancer markers or drug targets. Cell-membrane receptors include three traditional protein categories, i.e., G-protein-coupled receptors (GPCRs), ion channels, and enzyme-linked receptors, mostly represented by receptor tyrosine kinases (RTKs).[3] GPCRs represent the largest class of receptors,[4] while the enzyme-linked receptors the most studied one,[5] and together they comprise the majority of drug targets. The aberrant activity of these receptors was linked to many diseases including inflammation, metabolic disorders, and cancer.[6] Targeting, for example, HER2 receptors has been at the core of targeting HER2+ tumors. Proteomic analysis of cell-surface (CS) proteins has revealed, however, many important, additional roles for other CS proteins in cancer proliferation.[7] The detection and characterization of these cell-surface targets has been, nevertheless, challenging due to compounding factors such as low abundance, hydrophobicity, presence of post-translational modifications (PTMs), and heterogeneity.[8,9]

Several methods have been developed for the isolation of cell-surface proteins relying mainly on ultra-centrifugation, coating of the plasma membrane with silica-beads, and chemical labeling of N-linked glycosylated proteins or of protein amine, sulfhydryls or aldehyde groups, followed by

affinity pulldown.[8-12] After isolation, the state-of-the art for detecting the enriched CS protein fractions involves mass spectrometry (MS) analysis. The advanced capabilities of the MS technology (i.e., high sensitivity, high mass accuracy and quantification capability) enabled the detection of thousands of proteins per cell line, the compilation of comprehensive cell-surface protein data into interactive databases such as The Cell-Surface Protein Atlas (~1,500 human proteins),[13] and the development of even more comprehensive lists constructed with machine learning based predictor tools (~2,900 human proteins).[14] Altogether, these studies have contributed to the overall knowledge of what has been named the "surfaceome" and its associated signaling networks in humans,[14-15] largely captured in comprehensive public repositories.[16-19]

To capitalize on the wealth of information that can be generated through mass spectrometric analysis, this study was aimed at characterizing the cell-surface proteome of SKBR3/HER2+ breast cancer cells by using orthogonal methods for cell-surface protein enrichment and isolation, categorizing these proteins based on their functional role, identifying key drivers of aberrant proliferation, and exploring the opportunities presented by such cells for the development of effective diagnostic and therapeutic approaches.[20] We also report on the remodeling of the cell-membrane proteome under serum-starved and serum-supplemented conditions, and, lastly, we draw insights into the signaling cascades initiated at the plasma membrane and the potential crosstalk activities that fuel the development of resistance to treatment with therapeutic drugs.

4.2 Methods

HER2+ breast cancer cells, SKBR3, were enriched in the cell surface fraction by three orthogonal methods based on biotinylation of primary amines, biotinylation of oxidized glycans, and shaving of cell receptors in the cell culture as described in detail in **Section 3.3**. The three samples generated by the cell surface enrichment will be referred as amine, glyco, and trypsin in the following sections. These samples were analyzed by mass spectrometry as detailed in **Sections 3.8.1**, and **3.10.1**. Cell surface enrichment was performed in two cell culture conditions: serum-free (SF) and serum-treated (ST) and the identification of proteins with altered abundance levels was based on spectral counts as described in **Section 3.11.1**. The tools used for performing

data analysis and providing the biological interpretation of the cell surface enriched proteins are outlined in **Section 3.13**. An in-house database was created for cell surface identification as detailed in **Section 3.12**. Lastly, PRM and immunofluorescence was used to validate a few of the differentially expressed cell surface proteins following the protocols described in **Sections 3.9** and **3.14**, respectively.

4.3 Results

4.3.1. Annotation of the cell surface proteins

To isolate the cell surface proteins of SKBR3/HER2+ cancer cells, three orthogonal isolation methods were followed as summarized in **Figure 4-1 A,** while to annotate the isolated proteins as cell surface, an in-house database was created comprised of 7,760 theoretical cell surface proteins (**Figure 4-1 B**). The database was an assembly of cell surface proteins found in public databases UniProt (5,440 CS proteins)[16] and ProteinAtlas (2,068 CS proteins),[17,18] and cell surface proteins reported in literature by experimental enrichment methods[12,13] and *in-silico* predictions (4,717 CS proteins).[14,15] Based on the advanced search features of UniProt, cell membrane localization was selected in the Cellular Compartment (CC), Gene Ontology (GO), and Keyword (KW) fields and controlled vocabulary terms were used to construct four major functional categories and their subcategories (**Figure 4-1 B**). IUPHAR was also used to shape the subcategories for the GPCR family.[19] These categories provide a broad frame for more than half of the theoretical surfaceome, while many other cell surface proteins remain without an annotation. It should also be noted that these categories are versatile and multiple proteins can be part of more than one category due to their multiple functional roles in the cell, however this overlap is minimal (10-15%).

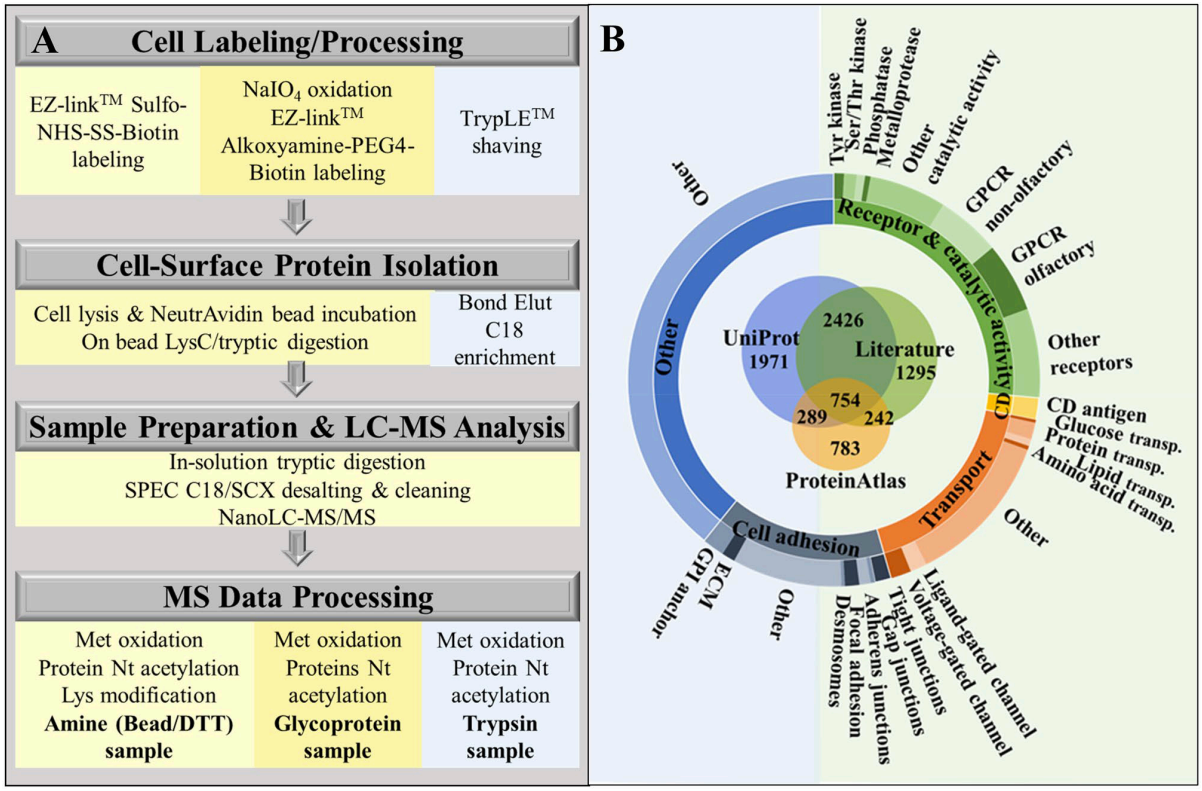


Figure 4-1: Cell-membrane protein isolation flowchart and database classification. **(A)** SKBR3 cell-membrane protein isolation and processing workflow via three distinct methods: biotin labeling of protein primary amine groups, biotin labeling of glycoproteins, and enzymatic shaving. **(B)** In-house built database of 7,760 cell membrane proteins classified based on GO controlled vocabulary terms. Reprinted with permission from Karcini & Lazar (2022).[20]

4.3.2. Efficiency of the cell surface protein enrichment

A separate study was performed to assess the efficiency labeling of the Sulfo-NHS-SS-biotin of lysine (Lys/K) amino acid residue of the primary amines and N-terminal (Nt) in different types of samples including an SKBR3 peptide sample and two isolated BSA protein samples at different sample-to-biotin ratios outlined in **Figure 4-2 A**. Since the Sulfo-NHS-SS-biotin labeling introduces an additional mass of +87.998 Da to each sequence carrying the label, the labeling efficiency was calculated as the percentage of modified peptides out of the total detected peptides. Labeling the SKBR3 peptide sample resulted in 99% of the detected peptides carrying a modification in their Nt and 82% of the detected peptides with a Lys in their sequence carrying a modification in Lys. Meanwhile, labeling the two samples of BSA protein at two different sample-to-biotin ratios (1:560 and 1:56) resulted in a lower labeling efficiency of 29% and 15% for the Nt modification, and 75% and 54% for the Lys modification for each ratio. The lower

efficiency came because of both lowered sample-to-biotin ratio and increased complexity of the labeled sample, from tryptic peptide mixture to protein. In fact, in cell culture the labeling of Lys was only 1.9% (**Figure 4-2 B**) and the labeling of Nt was not detected. The decrease in labeling rate from tryptic peptides to live cell culture is most likely associated with the hindering of the Lys groups in the complex cell surface structure and is also observed by other researchers.[21] Despite the labeling method, all samples carried modifications regarding the oxidation of methionine (Met/M) at a frequency of 4.5% and the acetylation of Nt at a frequency of 0.4% (**Figure 4-2 B**). Another way to assess the labeling efficiency of the biotin would be its imaging in the cells through fluorophore tagged reagents as it was done in the case of alkoxamine-PEG4-biotin (**Figure 4-2 C**), resulting in fluorescent SKBR3 cells throughout the chamber slide.

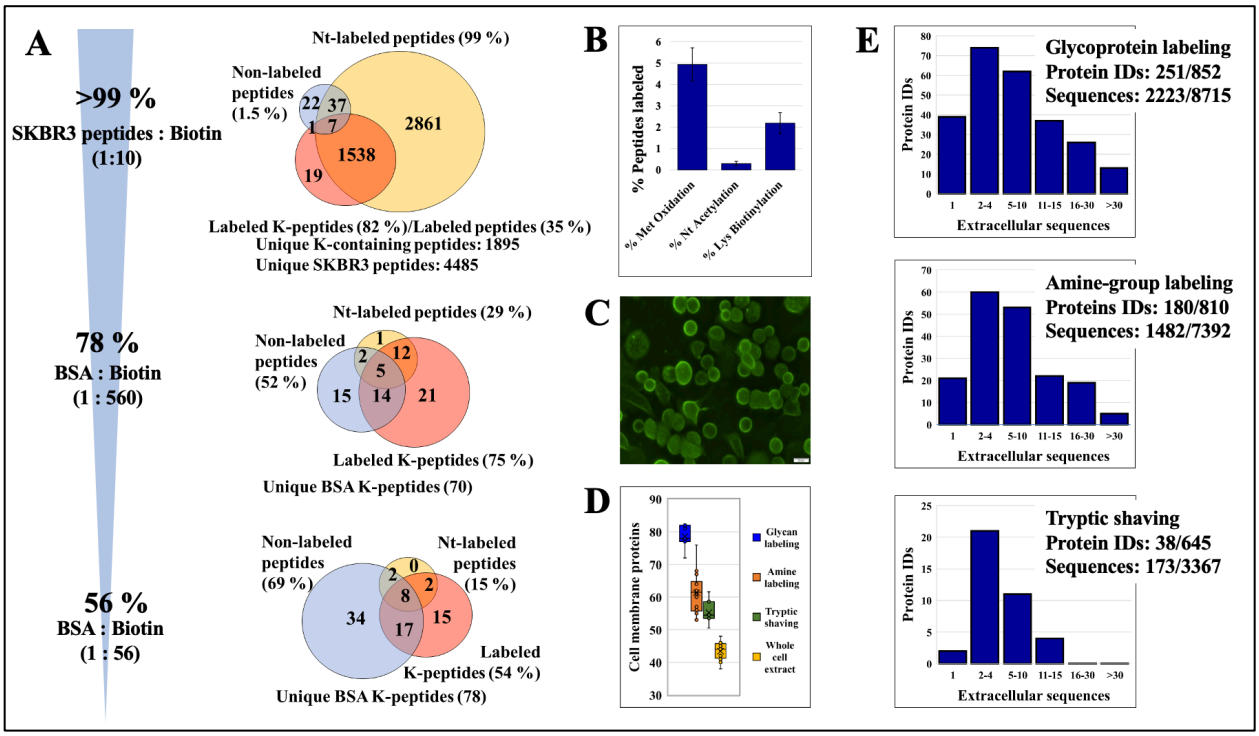


Figure 4-2: Cell membrane protein enrichment efficiency. (**A**) Percent peptides carrying a biotinylation-induced label in an SKBR3 cell extract and in BSA tryptic digests using various peptide/protein:biotin molar ratios (applicable to the amine group labeling method), the Venn Diagrams represent the number of labeled Lys containing peptides (+ 87.998 Da), peptides labeled at the N-terminus (+ 87.998 Da), and the number of non-labeled peptides. (**B**) Percent peptides carrying a PTM: Met oxidation, peptide Nt acetylation, and Lys biotinylation (case of the amine group labeling method), the error bars represent the SD of biological replicates. (**C**) SKBR3 cells labeled by alkoxyamine- PEG4-biotin and conjugated with streptavidin antibody-Alexa Fluor™ 488. (**D**) Cell membrane protein enrichment effectiveness represented by the number of cell membrane proteins in the top 100 most abundant proteins (abundance determined

by the number of matching unique peptides). (**E**) Histograms of protein IDs matched by different numbers of extracellular peptide sequences (with extracellular sequences detected/ total). Extracellular sequence assignments were made based on topological domain information extracted from UniProt. Reprinted with permission from Karcini & Lazar (2022).[20]

Based on the cell surface protein annotation provided by our in-house database (**Figure 4-1 B**), the enrichment efficiency was compared between the three different methods and an unlabeled sample, whole cell extract. This enrichment was calculated as a percentage of the number of the assigned cell surface proteins out of the top 100 proteins based on the peptide counts per protein (**Figure 4-2 D**). It resulted in the glyco method having the most enriched proteins (81%), higher than the amine (64%) and the trypsin (60%) methods, and almost double the unlabeled sample (43%). Another way to assess the detection of cell surface proteins was to determine the extracellular sequences/peptides that were identified by mass spectrometry; however, this analysis is hampered because annotation of extracellular sequences is vastly limited (UniProt contains only 7,546 extracellular sequences corresponding to only 2,923 proteins as of September 2021), and the extracellular peptides might have been missed by the MS analysis. Nevertheless, even by one method of enrichment (glyco), about 1/3 of the known extracellular sequences were identified (**Figure 4-2 E**) confirming again the superior enrichment of this method. Tryptic shaving yielded fewer and shorter extracellular sequences compared to the other two methods. Lastly, a control sample was run in the absence of the biotin reagent to account for the non-specific binding of proteins to the beads, which would result in artefacts, but only 11 proteins were detected with high confidence (FDR <3%, minimum 2 peptides/protein) and assigned as cell surface by the in-house database. These proteins were mostly abundant cytoskeletal or cytosolic proteins with multiple locations in the cell and they provided yet another aspect of the quality of the enrichment experiments.

4.3.3. Cell surface protein reproducibility and classification

Based on the three orthogonal isolation methods used, the annotation of cell surface proteins by our in-house database, and the confidence of the detected peptides and proteins by mass spectrometry with FDR <3%, 2,054 cell surface proteins were detected, of which 1,316 proteins were detected by at least two unique peptides per protein. This subset of 1,316 cell surface proteins detected by all isolation methods in both states: serum-free (SF) and serum-treated (ST)

is used for analysis and biological interpretation due to the confidence and reproducibility of the data shown in Figure **4-3 C, D**.

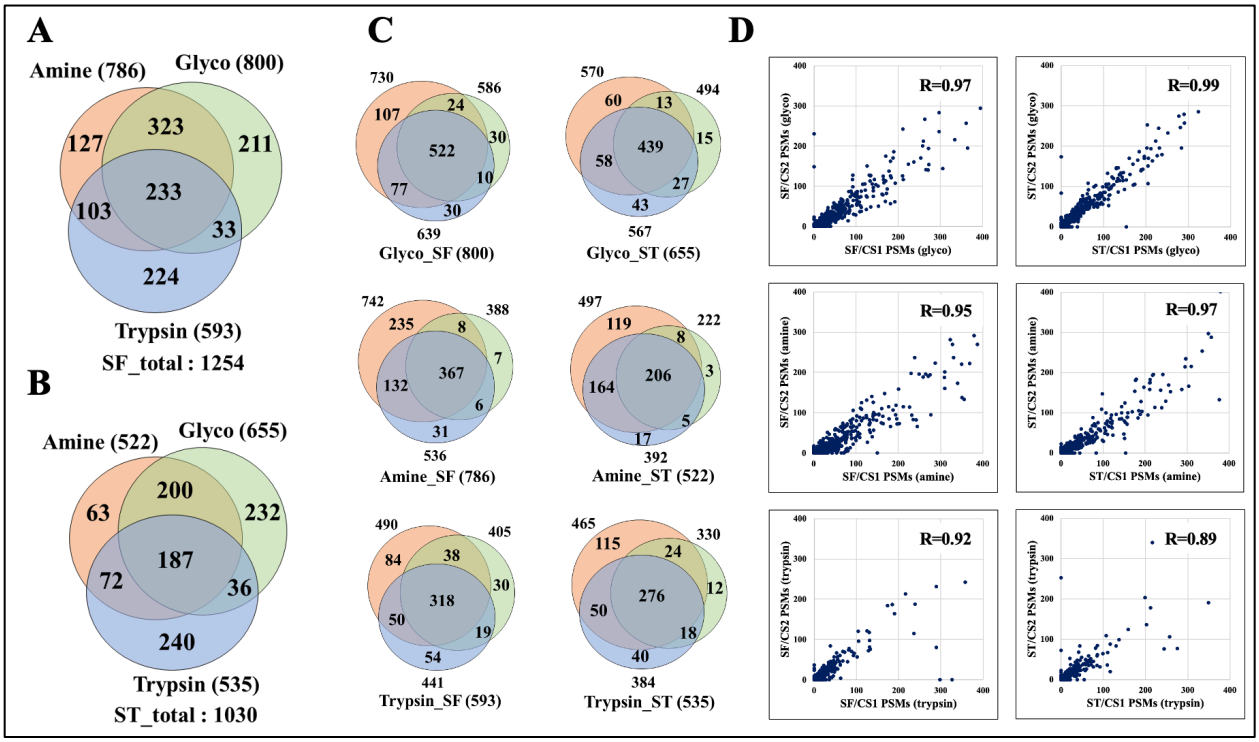


Figure 4-3: Protein ID Venn and PSM correlation diagrams representing the complementarity and reproducibility of the three labeling methods in detecting cell-membrane proteins matched by at least 2 unique peptides (FDR <3%). **(A)** Serum-free cultured cells. **(B)** Serum-treated cells. **(C)** Reproducibility of protein detection between three biological replicates for each labeling method and cell treatment condition (SF and ST). **(D)** PSM correlations between any two biological replicates for each of the three cell-membrane protein enrichment methods; the correlations are shown for the 0–400 PSM range in which most proteins could be found (R = Pearson correlation coefficient). Reprinted with permission from Karcini & Lazar (2022).[20]

The three isolation methods (amine, glyco, trypsin) proved to be complementary to each other in both states (SF and ST) due to the low overlap of the cell surface proteins detected by each of them (**Figure 4-3 A, B**), with each method contributing ~100-200 unique cell surface proteins. Most detected cell surface proteins were isolated by the chemical labeling methods of glyco (800 in SF, 655 in ST) and amine (786 in SF, 522 in ST), followed by trypsin (593 in SF, 535 in ST). The glyco method also displayed the largest reproducibility between the three biological replicates for both SF and ST states as summarized in **Figure 4-3 C**. The reproducibility between any two biological replicates was also shown by the high Pearson correlation coefficient between the spectral counts (≥0.95), with few exceptions (**Figure 4-3 D**). The set of 1,316 cell surface

proteins was aligned with the functional categories and subcategories described above in our in-house database (**Figure 4-1 B**) and 525 proteins, presented in **Figure 4-4 A**, were matched to this classification. Based on the isolation method, different categories of cell surface proteins were detected (**Figure 4-4 B, C, D**), where chemical labeling methods enabled the identification of more receptor kinases and CD antigens, while the trypsin method led to a larger number of detected transporter and adhesion proteins compared to the other categories.

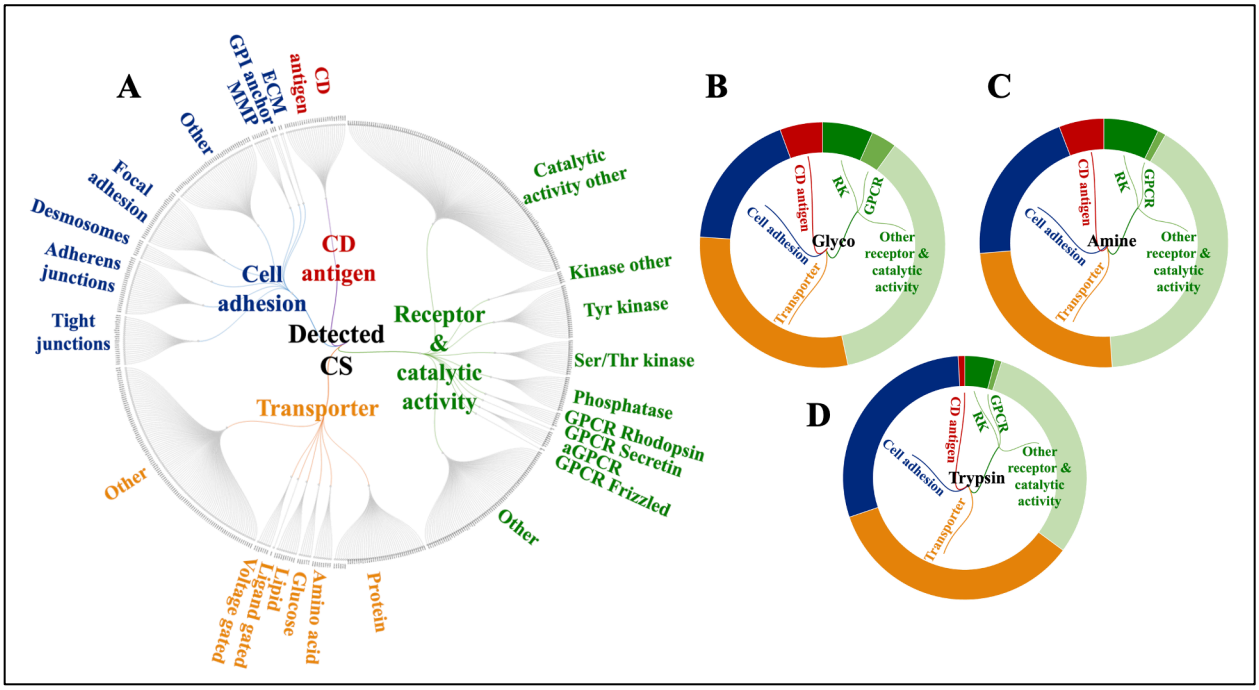


Figure 4-4: Functional categorization of the detected cell-membrane proteins based on GO controlled vocabulary terms (proteins detected by at least 2 unique peptides, FDR <3%). **(A)** Dendrogram of cell-membrane proteins detected by all labeling methods and conditions. **(B)**, **(C)**, and **(D)** Doughnut charts of detected cell-membrane proteins enriched via glycoprotein labeling, amino group labeling, and tryptic shaving methods, respectively. Reprinted with permission from Karcini & Lazar (2022).[20]

To further indicate that certain functional categories described above were indeed represented by a larger number of cell surface proteins as a result of these isolation methods, a comparison is made between the cell surface proteins detected in SF and ST of the whole cell (WC) extracts without any enrichment (three independent biological replicates) and the enriched ones. In addition, cell surface enrichment experiments (three independent biological replicates) using all three isolation methods were conducted in proliferating SKBR3 cells (not exposed to serum starvation) and they resulted in the enrichment of cell surface proteins representative of the same functional categories despite the cell state. These results are shown in **Table 4-1** as percentages

(shown in parenthesis) of the total cell surface proteins detected and annotated by our in-house database because the actual numbers of protein IDs fluctuate among the experimental conditions due to the different number of replicates for each.

Table 4-1: Cell-surface protein identification effectiveness with and without enrichment in cell-surface proteins. The whole protein set (white color columns) or only proteins matched by two unique peptides (2pep) set (gray color column) were considered. Notes: CS-cell surface, WC-whole cell. Bold, red colored values indicate conditions for which a substantial improvement in cell-surface protein enrichment was observed (vs. whole cell, non-enriched samples). Adapted with permission from Karcini & Lazar (2022).[20]

Experimental Conditions	**Not enriched WC (SF + ST)**	**Enriched CS Fraction (SF + ST)**	**Enriched CS Fraction (proliferating)**	**Not enriched WC (SF + ST) 2pep**	**Enriched CS Fraction (SF + ST) 2pep**	**Enriched CS Fraction (proliferating) 2pep**
Matches to the CS proteins (in-house DB)	2821	2054	1359	1108	1316	1175
Receptors	211 (7.5%)	168 (8.2%)	112 (8.2%)	21 (1.9%)	**117 (9%)**	**106 (9%)**
GPCRs	65 (2.3%)	36 (2%)	15 (1.1%)	1 (0.09%)	**15 (1.1%)**	**11 (1%)**
Tyr kinases	30 (1%)	29 (1.4%)	20 (1.5%)	4 (0.4%)	**26 (2%)**	**20 (1.7%)**
Ser/Thr kinases	44 (1.6%)	29 (1.4%)	17 (1.3%)	19 (1.7%)	17 (1.3%)	15 (1.3%)
Transport(ers)	425 (15%)	381 (18.5%)	275 (20%)	163 (14.7%)	279 (21.2%)	247 (21%)
Cell junction/ cell adhesion	421 (15%)	348 (17%)	251 (18%)	158 (14.3%)	255 (19.4%)	231 (19.7%)
GPI anchors	28 (1%)	25 (1.2%)	21 (1.5%)	2 (0.2%)	**17 (1.3%)**	**20 (1.7%)**
CD antigens	84 (3%)	105 (5%)	84 (6%)	13 (1.2%)	**89 (6.8%)**	**80 (6.8%)**

The enriched fraction is composed of more cell surface proteins compared to the non-enriched whole cell extract across the different categories, especially when only proteins with at least two unique peptides are considered indicating an increased confidence in the correct identification of these proteins. In fact, most of the GPCRs, tyrosine kinases, GPI anchors, or CD antigens are only detected when enrichment methods are used, most likely due to their low abundance. The exception are Ser/Thr kinases which are abundantly detected without any enrichment methods. One plausible explanation might be provided by their short extracellular domain in comparison

to their long cytoplasmic region, which is not the case for other types of receptors such as GPCRs or Tyr kinases as illustrated in **Figure 4-5 B** by a representative protein member of each of these classes.

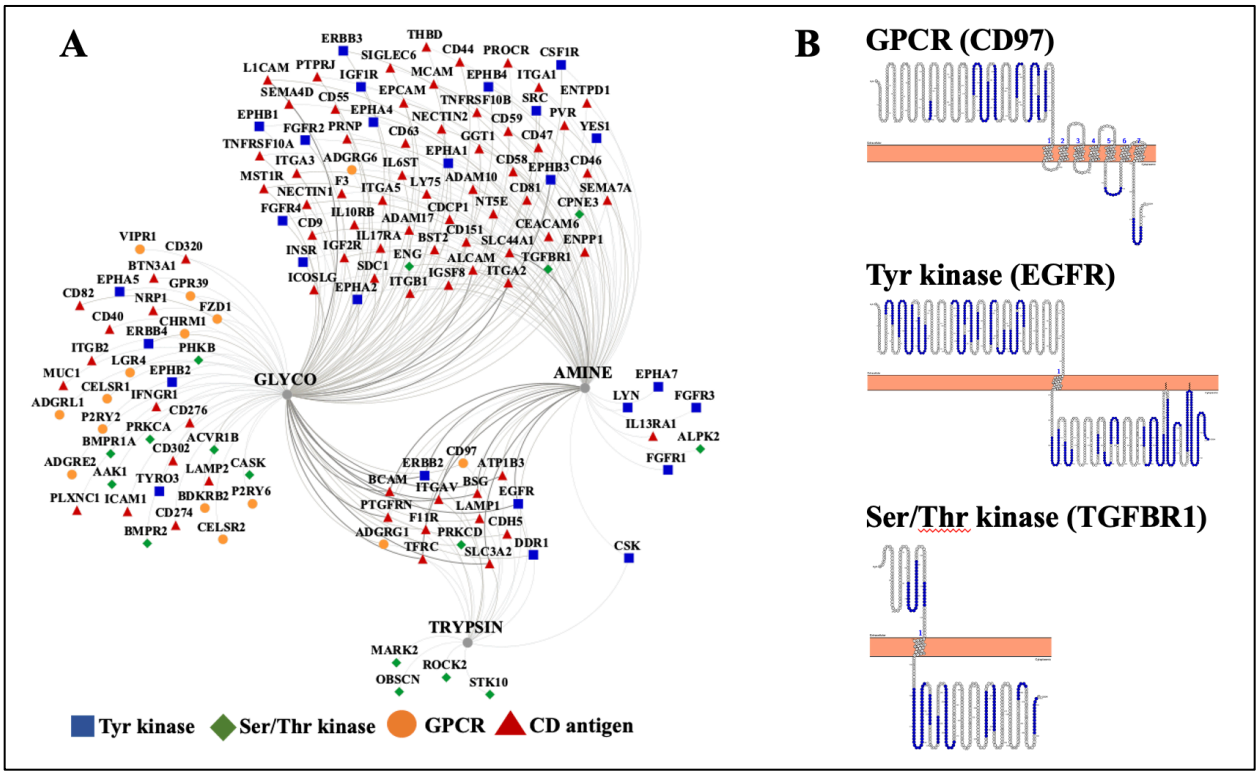


Figure 4-5: Detectability of cell-membrane receptors and their sequences. **(A)** The diagram represents the overlap between the detected proteins produced by the three enrichment methods. Notes: Visualization was performed with Cytoscape, the edge thickness reflects the protein abundance as evidenced by spectral counts. **(B)** Localization of the detected peptide sequences for a representative member of the receptor classes of tyrosine kinases (EGFR), serine/threonine kinases (TGFBR1), and GPCRs (CD97). Note: Visualization was performed with Protter. Adapted with permission from Karcini & Lazar (2022).[20]

In addition, these receptors (Tyr kinases, Ser/Thr kinases, GPCRs, CDs) were mostly detected by glyco method (**Figure 4-5 A**) making it the overall preferred method of enrichment for assessing the cell surface proteins.

4.3.4. Biological processes

The cell surface proteins matched by one of the four main functional categories: receptor and catalytic activity (275), transport molecules (279), cell adhesion and junction molecules (255), and CD antigens (89), were analyzed by STRING to retrieve their enriched biological processes.

These processes are representative of important cell events such as cell communication, growth, differentiation, death, secretion, locomotion, migration, adhesion, and immune response (**Figure 4-6 A**). Some functional categories are more represented in certain processes compared to others such as transport molecules in localization and transport, or CD antigens and RTKs in immune system and cell signaling. In addition, protein-protein interaction (PPI) networks of important categories of detected cell membrane proteins such as receptor kinases (**Figure 4-7 A**), GPCRs (**Figure 4-7 B**), and CD antigens (**Figure 4-7 C**) revealed the dense interplay between the members of these families and the signaling processes they are part of such as MAPK-ERK regulation (EGFR, ERBB2, IGF1R, ephrins), TGFB signaling (TGFBR1, BMPR1A, BMPR2), B-cell and T-cell activation (ADAM17, ITGB1, CD44, CD276), locomotion (ALCAM, EPCAM, CD9, integrins), etc. The detected GPCRs belonged to several different homology classes such as rhodopsin, secretin, and adhesion and their presence was validated by PRM (**Appendix A**) due to their low abundance.

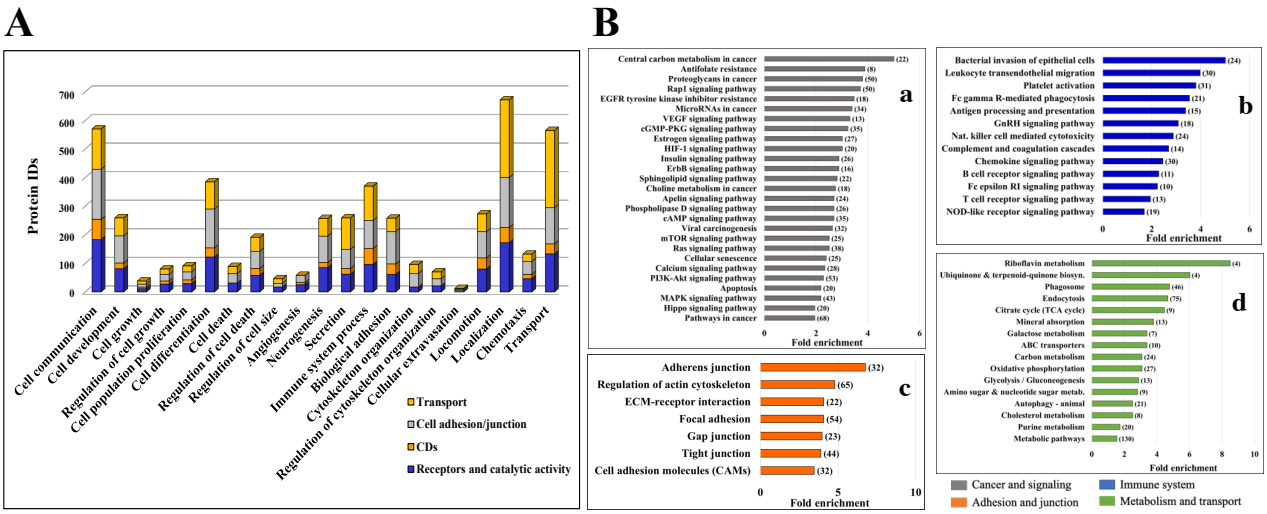


Figure 4-6: Bar charts of selected functional categories and pathways associated with the detected cell-membrane proteins. **(A)** Cancer-relevant enriched biological processes represented by the cell-membrane proteins. **(B)** Enriched KEGG pathways represented by cell-membrane proteins involved in: (Ba) signaling and cancer, (Bb) immune response, (Bc) adhesion/junction, and (Bd) metabolism and transport. Notes: Numbers in parentheses represent the number of proteins matched to each process, the background gene sets were the full set of corresponding proteins in the human proteome. Reprinted with permission from Karcini & Lazar (2022).[20]

To provide a comprehensive overview of the main cell processes carried by the detected cell membrane proteins, all 1316 proteins were subjected to KEGG pathway enrichment analysis, which revealed the most enriched processes related to: cancer and signaling (**Figure 4-6 Ba**)

such as VEGF, estrogen, and ERBB signaling (>3-fold), mTOR, Ras, and PI3K signaling (>2-fold); immune system (**Figure 4-6 Bb**) including platelet activation (~ 4-fold), chemokine, B-cell, and T-cell signaling (>2 -fold); adhesion and junction (**Figure 4-6 Bc**) represented by adherens junction (>6-fold), focal adhesion, gap, and tight junctions (>3-fold); metabolism and transport (**Figure 4-6 Bd**) like endocytosis and TCA cycle (>4-fold), ABC transporters, oxidative phosphorylation, and glycolysis (>3-fold). These processes illustrate the broad and multifunctional role of the proteins found at the cell membrane.

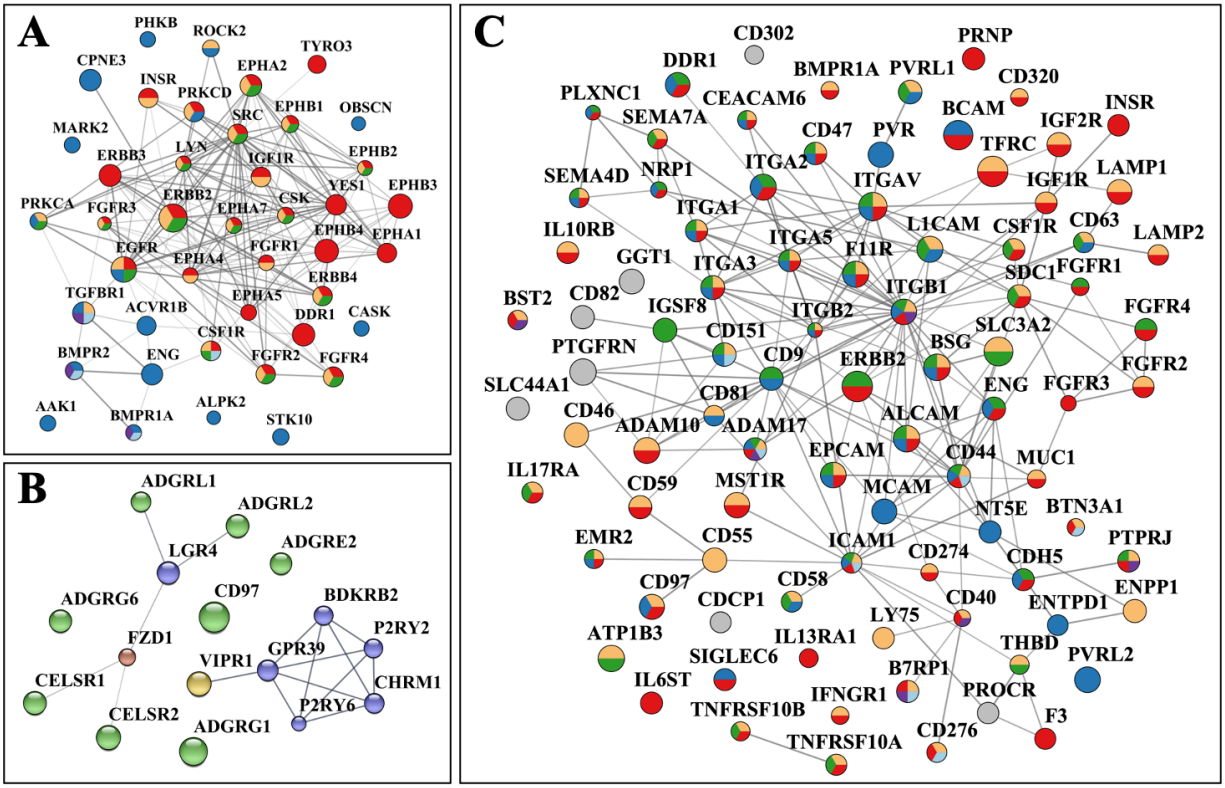


Figure 4-7: Protein–protein interaction networks of selected cell-membrane protein categories. **(A)** Receptor kinases: Red-Tyr kinase receptors, Blue-Ser/Thr kinase receptors, Yellow-MAPK regulation, Green-ERK1/ERK2 regulation, Light blue-Cytokine-cytokine receptor signaling, Magenta-TGFB signaling. **(B)** G-protein coupled receptors: Blue-Class A Rhodopsin, Yellow-Class B1 Secretin, Green-Class B2 Adhesion, Red-Class F Frizzled. **(C)** CD antigens: Yellow-Immune system process, Blue-Biological adhesion, Red-Cell communication, Green-Locomotion, Magenta-B cell activation, Light blue- T-cell activation. Notes: The PPI networks were generated with STRING and visualized with Cytoscape; node size is proportional to the total spectral counts that matched a protein, from <10 (small) to >10,000 (large). Reprinted with permission from Karcini & Lazar (2022).[20]

Several studies have identified certain proteins as suitable tumor markers for diagnosis or therapeutic targeting[22-26] and the detected cell surface proteins proved to be a rich resource of such markers as shown in **Figure 4-8 A and C**, mostly represented by CDs and adhesion proteins such as the tight junctions (TJP1-3). Markers of epithelial, mesenchymal, and stemness, which are indicators of cancer progression and metastatic propensity through the epithelial-mesenchymal transition (EMT) and mesenchymal-epithelial transition (MET) processes,[24] were also detected (**Figure 4-8 B**), with epithelial markers being the dominant ones in the SKBR3 cells surfaceome. The detected cell surface proteins were manually screened against the DrugBank database[27] for matching to the reported targets that have approved drugs against cancer (56 proteins), reported targets against which investigational drugs against cancer are being tested (48 proteins), and reported targets that are part of certain functional categories (transport, adhesion, RTKs, CDs, etc) against which experimental drugs, not necessarily targeting cancer, are being explored (62 proteins). These targets are represented in **Figure 4-8 D** according to their functional groups and level of drug approval.

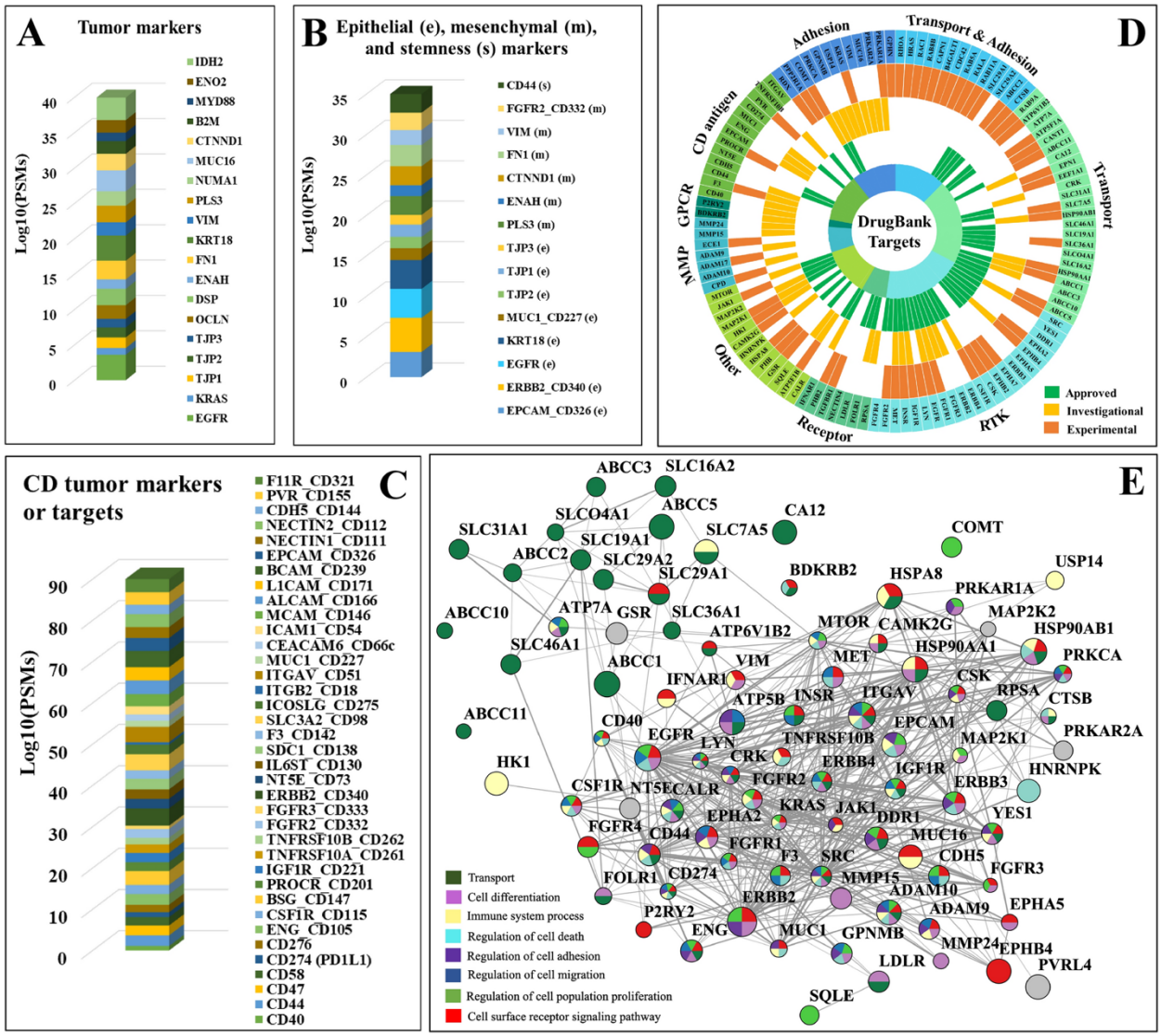


Figure 4-8: Cancer markers and drug targets detected in the SKBR3 cell-membrane proteome. **(A)** Tumor markers. **(B)** Epithelial (e), mesenchymal (m), and stemness (s) cancer markers. **(C)** CD Tumor markers and drug targets. **(D)** Sunburst chart representing receptors by functional categories identified in DrugBank as potential cancer therapeutic drug targets based on approved (green), investigational (yellow), and experimental (orange) status levels. **(E)** PPI network of the approved and investigational cancer drug targets. Note: Node size is proportional to the log10(SC), and edge thickness reflects the STRING interaction score. Reprinted with permission from Karcini & Lazar (2022).[20]

Most of the cell surface proteins being targeted by drugs are part of the kinase receptors and transport categories, while a vast opportunity for exploration can be offered by CD antigens, GPCRs, MMPs, and adhesion cell surface proteins. The dense PPI network between targets with approved drugs and those with investigational drugs shown in **Figure 4-8 E** provides opportunities for combined therapy exploration based on their interaction partners and the

processes they affect. Most of the solute carrier proteins (SLCs) and the ATP-binding cassette transporters (ABCs) with roles in transport form a separate cluster which interacts with members of the bigger cluster (receptor kinases, CD antigens) involved in regulation of cell processes such as: proliferation, adhesion, migration, death, and immune system.

4.3.5. Differential expression

Since two different treatments, serum-free (SF) and serum-treated (ST), were used to isolate the cell surface proteins of SKBR3 cell line, label-free quantitation analysis between treatments was performed based on protein abundances as determined by the spectral counts. This analysis was performed for the set of 852 cell surface proteins detected by the glyco method since it proved to be the most effective enrichment method and yield the highest number of detected cell surface proteins (**Figures 4-2 D, 4-3, 4-5 A**). The protein abundances compared were normalized once globally to account for spectral count variability between samples due to processing, and a second time against a specific detected list of known cell surface proteins to account for the presence of proteins from cell compartments other than the cell surface. The correction factors for the first normalization were calculated based on the spectral counts of all proteins being compared as explained in **Section 3.11.1**, while the correction factors for the second normalization were calculated based on the spectral counts of the selected cell surface list. This normalization list contained 10 cell surface proteins that were detected in every biological replicate of every enrichment method, associated only with the cell surface fraction, not with other cell compartments, and had a transmembrane domain (SLC2A1, SLC3A2, SLC16A3, CDH5, PCDH1, DSG2, F11R, ITGB5, GNAS, SUSD2). This analysis resulted in 49 upregulated and 42 downregulated proteins in the SF treatment by ≥2-fold change with a p-value ≥0.05. **Figure 4-9** summarizes the findings and shows a summary of the enriched biological processes .

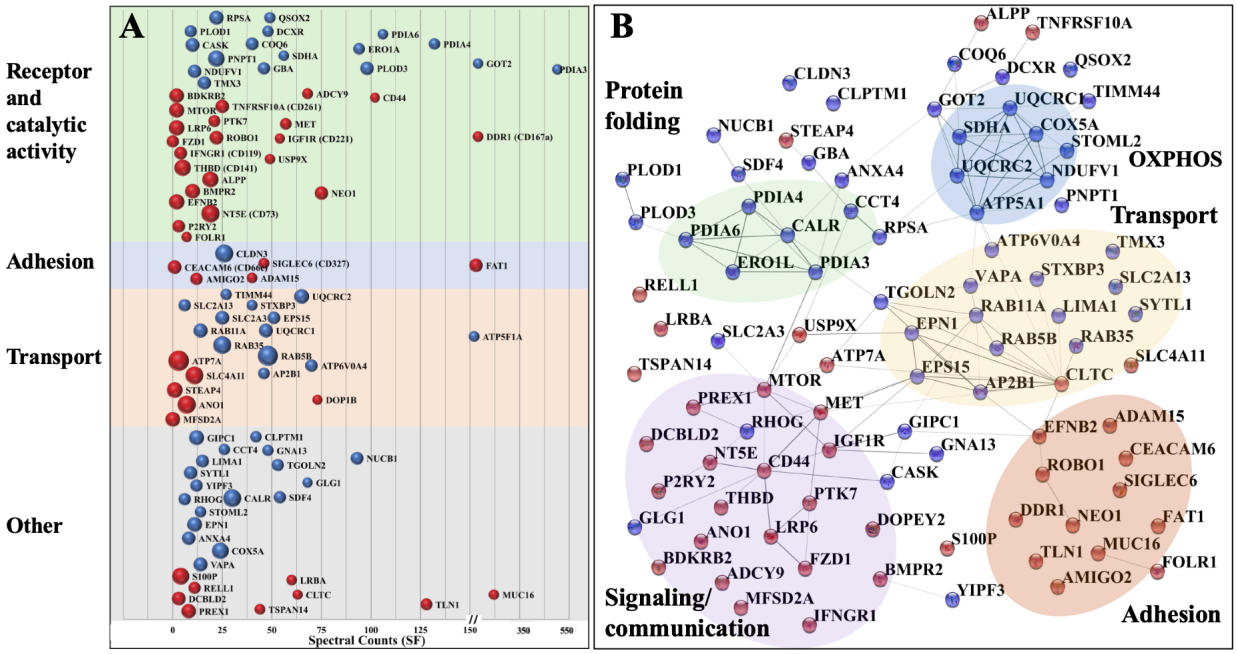


Figure 4-9: Proteins with changed abundance in the cell-membrane proteome. **(A)** Proteins with increased (red) and decreased (blue) abundance in ST vs. SF cells, categorized by function (y axis) and spectral counts in SF cells (x-axis). The sphere size is proportional to the log2(FC) in protein spectral counts. **(B)** Cytoscape visualization of the STRING PPI network created with the proteins that displayed a change in abundance (same color scheme as in A). Reprinted with permission from Karcini & Lazar (2022).[20]

The detected cell surface proteins with altered expression were part of all the major functional groups such as transport, adhesion, and receptors and catalytic activity, which had the highest number of differentially expressed proteins. Transport proteins displayed the largest fold changes for either serum-free or serum-treated upregulated proteins (**Figure 4-9 A**). The upregulated serum-free proteins were part of several enriched biological processes including: intracellular transport, endocytosis, or plasma membrane to endosome transport mediated by Rab GTPase proteins (RAB35, RAB11A, RAB5B); protein folding and processing in ER mediated by protein disulfide isomerases (PDIA3, PDIA4, PDIA6) and calreticulin (CALR); secretion regulated by different categories of receptors with catalytic activity such as CASK, or transport proteins such as SLC2A3; and oxidative phosphorylation (OXPHOS) facilitated by mitochondrial complex proteins (NDUFV1, SDHA, UQCRC1/2, COX5A, ATP5A1) detected at the cell membrane. As these mitochondrial proteins were detected at the cell membrane, validation by other methods such as PRM and immunofluorescence imaging of their cell surface presence and upregulation is presented in **Appendices B, C, D**. The upregulated serum-treated proteins were part of enriched

processes including cell adhesion by their respective adhesion proteins like CEACAM6, SIGLEC6, ADAM15, and signaling and communication carried by receptors with catalytic activity such as MET, MTOR, IGF1R, CD44. These processes show the different responses initiated by the cell surface proteins in the presence or absence of the serum. A rich environment with different growth factors found in the serum promote signal transduction and growth, while a lack of these conditions induces stress responses that alter the metabolic processes of the cell.

4.4 Discussion

The combination of three cell surface protein enrichment methods and the creation of an in-house database for assigning cell surface proteins (**Figure 4-1**) resulted in the identification of 1,317 cell surface proteins (<3% FDR with at least 2 unique peptides) in both serum-free and serum-treated cells. Assessments of reproducibility and enrichment efficiency (**Figures 4-2, 4-3**) showed that the glyco method performed the best and resulted in the highest number of receptor kinases, GPCRs, and CD antigens (**Figure 4-5 A**). The proteins identified in these groups played important roles in MAPK-ERK proliferative pathways, TGFB signaling, immune system, adhesion, and locomotion (**Figure 4-7**).

4.4.1. Catalytic receptors

Receptor kinases due to their role in signal transmission, cell proliferation, differentiation, and development, and their reported overexpression and aberrant activity in many diseases including inflammation, metabolic disorders, and cancer are widespread drug targets as shown by our results (**Figure 4-8 D, E**) and reported in the literature.[6] Among the detected catalytic receptors, the most relevant to cancer growth and proliferation are the ERBBs, fibroblast growth factor receptors (FGFRs), and ephrin receptors (**Figure 4-7 A**) part of the receptor tyrosine kinases (RTKs). Ephrin receptors are the largest family of RTKs and play multi-functional roles in cancer.[28] In particular, ephrin and FGF receptors are part of interaction networks of adhesion/motility and angiogenesis hallmarks since they are involved in vascular and epithelial development and remodeling.[29,30] In addition, the involvement of these catalytic receptors in immune response and inflammation (**Figure 4-6 Bb**) presents novel therapeutic opportunities such as the advancement of chimeric antigen receptor- T cell therapy (CAR-T) in solid tumors. In the last years, different research groups have taken interest in cell surface proteins such as

chemokines and their ligands for their ability to recruit immune cells and advance such therapy,[31] and the detected ephrin receptors, both EPH-A and EPH-B family, play an important role in chemokine signaling and T cell trafficking.[32]

Despite the RTKs, another major group of catalytic receptors are the serine/threonine kinase receptors such as the detected bone morphogenetic receptors (BMPRs), activin receptors (ACVRs), and transforming growth factor-beta receptor type 1 (TGFRB1), all involved in cytokine signaling and directly interacting with each other (**Figure 4-7 A**). These multifunctional cytokine receptors, which respond to cytokines such as TGFB and BMP, are reported for aberrant expression and dual roles in tumorigenesis.[19,33] The detected BMPR2, BMPR1A, ACVR1, ACVR1B, and TGFBR1 proteins affect growth and apoptosis, and depending on the cell line and context can have an inhibitory or stimulatory effect on tumor progression, but the specific factors driving these dual responses remain unknown.[33-35] In addition, these receptors, like the RTKs, have expansive roles in signaling and cancer by interacting with and affecting other detected cell surface proteins such as CD44, upregulated by BMPR2 to advance metastasis in breast cancer,[35,36] and neuropilin-1 (NRP1) to support angiogenesis.[37] The presence of these bone morphogenetic receptors in breast cancer represents a novel therapeutic opportunity due to the metastatic propensity of breast cancer in the bone tissue.[33] In addition, the presence of mutations or changes in expression of these serine/threonine kinases (TGFRBs, BMPRs, ACVRs) can be an indicator of their activity in cancer and leveraged therapeutically.[33,34] Therefore, this study only opens the door for further exploration into this family of cell surface receptors, which act not only through the canonical SMAD pathway to regulate transcription,[33] but also play a role in proliferation and apoptosis through other pathways, such as MAPK or Hippo, as indicated by **Figure 4-6 Ba**.

4.4.2. G-protein coupled receptors

One of the most interesting findings of this study was the identification of the cell surface G-protein coupled receptors across their homology classes, especially the adhesion class (aGPCR), detected mainly by the glyco method probably due to the known glycosylation rate of aGPCRs at a specific region of their structure.[38] Adhesion GPCRs, such as the detected ADGRG1 (GPR56) or ADGRE5 (CD97), have recently gained attention due to their newly discovered roles in

cancer progression and metastasis and their reported increased expression in various cancers.[39] Despite being one of the largest families of cell surface and druggable receptors,[40] few of them are being investigated as drug targets against cancer (**Figure 4-8 D**) probably due to the high sequence homology between many GPCRs at the binding site of ligands and a lack of understanding of the mechanism of action of many aGPCRs.[38,41] In addition, they entail a complex signaling landscape in part due to their numerous downstream targets and in part due to their crosstalk and transactivation activity of catalytic receptors (including both RTKs and serine/threonine receptors) whose specific players, receptors and ligands, and mechanisms remain unexplored.[40,42,43] The detected GPCRs play important roles in various cancer types and depending on the context their overexpression is reported to promote cell proliferation and invasion, or suppress growth and inflammation due to their involvement in critical pathways such as Wnt signaling (FZD1, LGR4),[44,45] purinergic signaling (P2RY2, P2RY6),[46,47] and adhesion (CD97, ADGRG1.[48,49] Interestingly, the last two proteins have the highest spectral counts among the detected GPCRs and are reported as the highest expressed aGPCRs among the cancer types that they have been studied (melanoma, gastric, esophageal, and thyroid).[49] However, further studies are needed in other cancer types to establish their impact in tumor progression. The detection of these GPCRs (CD97 and ADGRG1) by our study in a HER2+ breast cancer cell line contributes to their recognition as potential targets for cancer therapy.

4.4.3. Other cell membrane proteins

The data analysis of the detected cell surface proteins highlighted the role of the non-catalytic receptors such as (CD antigens, adhesion, transporters) in processes and pathways initiated by catalytic receptors such as MAPK, PI3K-Akt, mTOR signaling (**Figure 4-6**). These receptors do not strictly adhere to a certain functional group and there are many overlaps between them, such as integrins or proteins with adhesion properties also being CD antigens (ITGAV/CD51, ITGB2/CD18, EPCAM/CD326, etc.). The immune checkpoint markers CD274 (programmed cell death 1 ligand/ PDL1) and CD276, and their interactors in the PPI networks (**Figure 4-7 C**) may suggest possible mechanisms of immune response evasion. Both markers are part of the same family modulating T-cell response and their overexpression in many cancer types has driven multiple investigations into therapeutical interventions including the combinational of antibodies targeting CD274 and CD276.[50] In addition to T-cell immune response, natural killer

(NK) mediated cytotoxicity is another mechanisms of tumor immunosurveillance, which can be modulated by another detected CD antigen, poliovirus receptor (PVR/ CD155).[51] CD antigens were not only involved in immune signaling (**Figure 4-7 C**), but also in epithelial-mesenchymal transition (EMT) with both epithelial (CD326/EPCAM, CD227/MUC1) and stemness (CD44) markers being present in the SKBR3 cells, an indication of the heterogeneity of these cancer cells. The proteins that shape the EMT transition, and subsequently the tumor microenvironment, such as MUC1 and CD44, are also involved in immune escape and can be potential targets for immunotherapy.[52] Despite many of the detected CDs being tumor markers (**Figure 4-8 C**), only few of them are being investigated as drug targets (**Figure 4-8 D**); thus, our detected CD antigens expand the current landscape of potential therapeutic agents. CD antigens are also integral to proliferative signaling (**Figure 4-6**) such as the detected immunoglobulin superfamily of cell adhesion molecules (CAM) - ALCAM/ CD166, L1CAM/ CD171, and BCAM/ CD239 - modulating the signaling of integrins and catalytic receptors such as EGFR.[53]

Integrins play complex roles in regulating cell growth, survival, and migration by engaging other scaffolding and catalytic receptors, such as ADAM and ephrin receptors (**Figure 4-6, 4-7 C**). ADAM10 is part of the top100 most abundant cell surface proteins detected and is often considered a novel player in the breast cancer landscape due to the shedding of other surface proteins such as CD44, cadherins, and other cell adhesion molecules, leading to cancer progression.[54] Our data also indicates its involvement in immune responses and adhesion processes (**Figure 4-7 C**) and supports its further investigation as a drug target (**Figure 4-8 D**). Transporters are mainly involved in metabolic and glucose-related processes including gluconeogenesis, glycolysis, and carbohydrate mechanisms, represented by the detected solute carrier family proteins (SLCs) such as SLC2A3, SLC23A2, SLC39A14, SLC2A1, SLC7A5 (**Figure 4-6**). SLCs are reported to be involved in metabolic rewiring, metastasis, and T-cell differentiation and activation.[55,56] SLCs are found to be differentially expressed in many breast cancer types and serve as carries for important substrates such as hormones (oestrogen) and pharmacological agents, becoming an important factor in designing cancer therapies[56]. Other classes of detected cell surface receptors with a transport activity include ion channels, such as ORAI1, ANO1, and PANX1, which are tumor markers involved in evasion and cancer

metastasis,[57] and ATP-binding cassette (ABC) transporters such as ABCC1/2/3/5/10 associated with multidrug-resistance due to their overexpression leading to efflux of the drugs.[58]

4.4.4. Potential drug targets

The detected cell surface proteins were explored for their potential as drug targets (**Figure 4-7 D**) and a group of key catalytic receptors, adhesion proteins, CD antigens, and ABC transporters, such as EGFR, ERBB, EPCAM, and ABCC1/2/3/5/10, have already approved drugs against them in various cancers. The complex protein signatures and network architectures presented here (**Figure 4-7**) allow for broad targeting of interconnected RTKs or CDs and can uncover weaknesses that may trigger failures within established or extended networks. Targeting orthogonal pathways by using combinations of signal transduction and angiogenesis inhibitors, regulators of apoptosis, or inhibitors of drug efflux proteins can selectively attack or eliminate cancer cells and combat the strategies used by cancer cells to resist drugs.

4.4.5. Differentially expressed proteins

The most interesting behavior between the serum-free and serum-treated cells was the presence and upregulation of the mitochondrial respiratory proteins part of the OXPHOS system (NDUFV1, SDHA, UQCRC1/2, COX5A, ATP5A1) in the surface of the serum-free cells which can be a result of protein transport and re-localization to the membrane or elevated abundances of the proteins. OXPHOS upregulation was only observed in the cell surface enriched fractions, but not in the whole cell extracts comparing serum-free and serum-treated conditions, indicating a two-fold purpose of such behavior. OXPHOS proteins even though previously reported in the cell surface,[59-61] remain largely unknown for their role in the membrane with some studies suggesting eATP generation.[60] In fact, eATP is overexpressed in various cancers triggering different responses, including tumor progression and invasion, cell death, or immune responses,[62,63] and the cell-surface ATP synthase components (F_1F_0 ATP) have been linked to aggressive, late-stage cancers[64] prompting for development of targeted therapies. OXPHOS and aerobic glycolysis are responsible for ATP production in normal and cancer cells respectively, therefore, an overexpression of the OXPHOS proteins might suggest a heavier reliance of cancer cells on this process for energy production suggesting possible mechanisms of cancer survival under nutrient-deprived conditions, also reported by few other studies.[65,66] OXPHOS inhibitors

have also been suggested as a targeting method for high OXPHOS cancers.[60,67] In the presence of nutrients, processes that support cancer cell proliferation, migration, and propensity for tissue invasion and metastasis are facilitated by proteins involved in cell growth, survival, and cell differentiation including PI3K, mTOR, IGF1R, MET, and CD44 receptors (**Figure 4-6**). These multifunctional proteins are also found to drive resistance to targeted therapies such as EGFR tyrosine kinase inhibitors (TKIs),[68,69] further contributing to cancer recurrence and posing future challenges for developing more effective combinational therapies.

4.5 Conclusion

The cell surface proteome of SKBR3 cells revealed a diverse range of cell-surface receptors, immune response factors, adhesion molecules, and transporter proteins that support interactions between cancer cells and the tumor environment. Analysis of protein-protein interaction networks demonstrated that these cell-membrane proteins work together to nourish abnormal cell growth, invasion, and metastasis by using signaling mechanisms represented not only by the well-studied catalytic receptors, but also through other functional groups such as cell adhesion molecules and transporters. Metalloproteinases, nectins, ephrins, and bone morphogenetic proteins were among the significant protein groups with newly discovered functions in cancer development. Proteins that support or regulate angiogenesis, cell-cell or cell-extracellular matrix interactions, and epithelial-mesenchymal transition processes assisted in cell migration, invasion, metastasis, and immune responses. Changes in the abundance of specific cell-surface proteins due to serum removal shed new light into the existing studies on how cancer cells manipulate metabolic mechanisms to continue their growth and proliferation. Drug targeting propensity was revealed for a broad range of receptors beyond the traditional targeting of multi-drug resistance proteins such as ABC transporters or catalytic receptors such as EGFR family. The availability of a vast array of multifunctional cell-membrane proteins offered promising prospects for developing more effective combination therapies that target cell growth, autocrine/survival, apoptotic, angiogenic, and cell migratory pathways. Overall, this study highlighted the untapped opportunities for cancer diagnosis, prognosis, and evaluating the likelihood of recurrence after therapy.

4.6 References

1. Watkins E. J. (2019). Overview of breast cancer. *JAAPA : official journal of the American Academy of Physician Assistants, 32*(10), 13–17. DOI: 10.1097/01.JAA.0000580524.95733.3d

2. Nahta, R. (2019). Novel therapies to overcome HER2 therapy resistance in breast cancer. In M.R. Szewczuk, B. Qorri, M. Sambi (eds), Current Applications for Overcoming Resistance to Targeted Therapies, 191–221. Cham, Switzerland: Springer. DOI: 10.1007/978-3-030-21477-7_7

3. Cooper, G.M. (2000). The Cell: A Molecular Approach (2nd ed). Sunderland, MA: Sinauer Associates.

4. Insel, P. A., Sriram, K., Wiley, S. Z., Wilderman, A., Katakia, T., McCann, T., … Murray, F. (2018). GPCRomics: GPCR Expression in Cancer Cells and Tumors Identifies New, Potential Biomarkers and Therapeutic Targets. *Frontiers in pharmacology, 9*, 431. DOI: 10.3389/fphar.2018.00431

5. Lemmon, M. A., & Schlessinger, J. (2010). Cell signaling by receptor tyrosine kinases. *Cell, 141*(7), 1117–1134. DOI: 10.1016/j.cell.2010.06.011

6. Nieto Gutierrez, A., & McDonald, P. H. (2018). GPCRs: Emerging anti-cancer drug targets. Cellular signalling, 41, 65–74. DOI: 10.1016/j.cellsig.2017.09.005

7. Ziegler, Y. S., Moresco, J. J., Tu, P. G., Yates, J. R., III, & Nardulli, A. M. (2014). Plasma membrane proteomics of human breast cancer cell lines identifies potential targets for breast cancer diagnosis and treatment. *PloS one, 9*(7), e102341. DOI: 10.1371/journal.pone.0102341

8. Elschenbroich, S., Kim, Y., Medin, J. A., & Kislinger, T. (2010). Isolation of cell surface proteins for mass spectrometry-based proteomics. Expert review of proteomics, 7(1), 141–154. DOI: 10.1586/epr.09.

9. Kuhlmann, L., Cummins, E., Samudio, I., & Kislinger, T. (2018). Cell-surface proteomics for the identification of novel therapeutic targets in cancer. Expert review of proteomics, 15(3), 259–275. DOI: 10.1080/14789450.2018.1429924

10. Li, Y., Wang, Y., Mao, J., Yao, Y., Wang, K., Qiao, Q., Fang, Z., & Ye, M. (2019). Sensitive profiling of cell surface proteome by using an optimized biotinylation method. *Journal of proteomics, 196*, 33–41. DOI: 10.1016/j.jprot.2019.01.015

11. Wollscheid, B., Bausch-Fluck, D., Henderson, C., O'Brien, R., Bibel, M., Schiess, R., … Watts, J. D. (2009). Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nature biotechnology, 27*(4), 378–386. DOI: 10.1038/nbt.1532

12. Kalxdorf, M., Gade, S., Eberl, H. C., & Bantscheff, M. (2017). Monitoring Cell-surface N-Glycoproteome Dynamics by Quantitative Proteomics Reveals Mechanistic Insights into

Macrophage Differentiation. *Molecular & cellular proteomics, 16*(5), 770–785. DOI: 10.1074/mcp.M116.063859

13. Bausch-Fluck, D., Hofmann, A., Bock, T., Frei, A. P., Cerciello, F., Jacobs, A., … Wollscheid, B. (2015). A mass spectrometric-derived cell surface protein atlas. *PloS one, 10*(3), e0121314. DOI: 10.1371/journal.pone.0121314

14. Bausch-Fluck, D., Goldmann, U., Müller, S., van Oostrum, M., Müller, M., Schubert, O. T., & Wollscheid, B. (2018). The in silico human surfaceome. Proceedings of the National Academy of Sciences of the United States of America, 115(46), E10988–E10997. DOI: 10.1073/pnas.1808790115

15. Ramilowski, J. A., Goldberg, T., Harshbarger, J., Kloppmann, E., Lizio, M., Satagopam, V. P., … Forrest, A. R. (2015). A draft network of ligand-receptor-mediated multicellular signalling in human. *Nature communications, 6*, 7866. DOI: 10.1038/ncomms8866

16. UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic acids research, 47*(D1), D506–D515. DOI: 10.1093/nar/gky1049

17. Thul, P. J., Åkesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait Blal, … Lundberg, E. (2017). A subcellular map of the human proteome. *Science, 356*(6340), eaal3321. DOI: 10.1126/science.aal3321

18. Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, … Pontén, F. (2015). Proteomics. Tissue-based map of the human proteome. *Science, 347*(6220), 1260419. DOI: 10.1126/science.1260419

19. Alexander, S. P. H., Christopoulos, A., Davenport, A. P., Kelly, E., Mathie, A., Peters, J. A., … CGTP Collaborators (2019). THE CONCISE GUIDE TO PHARMACOLOGY 2019/20: G protein-coupled receptors. *British journal of pharmacology, 176*(Suppl 1), S21–S141. DOI: 10.1111/bph.14748

20. Karcini, A., & Lazar, I. M. (2022). The SKBR3 cell-membrane proteome reveals telltales of aberrant cancer cell proliferation and targets for precision medicine applications. *Scientific reports, 12*(1), 10847. DOI: 10.1038/s41598-022-14418-0

21. Hörmann, K., Stukalov, A., Müller, A. C., Heinz, L. X., Superti-Furga, G., Colinge, J., & Bennett, K. L. (2016). A Surface Biotinylation Strategy for Reproducible Plasma Membrane Protein Purification and Tracking of Genetic and Drug-Induced Alterations. *Journal of proteome research, 15*(2), 647–658. DOI: 10.1021/acs.jproteome.5b01066

22. Barriere, G., Fici, P., Gallerani, G., Fabbri, F., Zoli, W., & Rigaud, M. (2014). Circulating tumor cells and epithelial, mesenchymal and stemness markers: characterization of cell subpopulations. *Annals of translational medicine, 2*(11), 109. DOI: 10.3978/j.issn.2305-5839.2014.10.04

23. Krawczyk, N., Meier-Stiegen, F., Banys, M., Neubauer, H., Ruckhaeberle, E., & Fehm, T. (2014). Expression of stem cell and epithelial-mesenchymal transition markers in circulating tumor cells of breast cancer patients. *BioMed research international*, 415721. DOI: 10.1155/2014/415721

24. Ribatti, D., Tamma, R., & Annese, T. (2020). Epithelial-Mesenchymal Transition in Cancer: A Historical Overview. *Translational oncology, 13*(6), 100773. DOI: 10.1016/j.tranon.2020.100773

25. Liao, T. T., & Yang, M. H. (2020). Hybrid Epithelial/Mesenchymal State in Cancer Metastasis: Clinical Significance and Regulatory Mechanisms. *Cells, 9*(3), 623. DOI: 10.3390/cells9030623

26. National Cancer Institute. (2021). Tumor Markers in Common Use. Retrieved from https:// www. cancer. gov/ about- cancer/ diagn osis- stagi ng/ diagn osis/ tumor- marke rs- list/

27. Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., … Wilson, M. (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic acids research, 46*(D1), D1074–D1082. DOI: 10.1093/nar/gkx1037

28. Mosch, B., Reissenweber, B., Neuber, C., & Pietzsch, J. (2010). Eph receptors and ephrin ligands: important players in angiogenesis and tumor angiogenesis. *Journal of oncology*, 135285. DOI: 10.1155/2010/135285

29. Salvucci, O., & Tosato, G. (2012). Essential roles of EphB receptors and EphrinB ligands in endothelial cell function and angiogenesis. *Advances in cancer research, 114*, 21–57. DOI: 10.1016/B978-0-12-386503-8.00002-8

30. Latko, M., Czyrek, A., Porębska, N., Kucińska, M., Otlewski, J., Zakrzewska, M., & Opaliński, Ł. (2019). Cross-Talk between Fibroblast Growth Factor Receptors and Other Cell Surface Proteins. *Cells, 8*(5), 455. DOI: 10.3390/cells8050455

31. Umut, Ö., Gottschlich, A., Endres, S., & Kobold, S. (2021). CAR T cell therapy in solid tumors: a short review. *Memo, 14*(2), 143–149. DOI: 10.1007/s12254-021-00703-7

32. Darling, T. K., & Lamb, T. J. (2019). Emerging Roles for Eph Receptors and Ephrin Ligands in Immunity. *Frontiers in immunology, 10*, 1473. DOI: 10.3389/fimmu.2019.01473

33. Bach, D. H., Park, H. J., & Lee, S. K. (2017). The Dual Role of Bone Morphogenetic Proteins in Cancer. *Molecular therapy oncolytics, 8*, 1–13. DOI: 10.1016/j.omto.2017.10.002

34. Qiu, W., Kuo, C. Y., Tian, Y., & Su, G. H. (2021). Dual Roles of the Activin Signaling Pathway in Pancreatic Cancer. *Biomedicines, 9*(7), 821. DOI: 10.3390/biomedicines9070821

35. Ehata, S., & Miyazono, K. (2022). Bone Morphogenetic Protein Signaling in Cancer; Some Topics in the Recent 10 Years. *Frontiers in cell and developmental biology, 10*, 883523. DOI: 10.3389/fcell.2022.883523

36. Huang, P., Chen, A., He, W., Li, Z., Zhang, G., Liu, Z., … Wang, J. (2017). BMP-2 induces EMT and breast cancer stemness through Rb and CD44. *Cell death discovery, 3*, 17039. DOI: 10.1038/cddiscovery.2017.39

37. Niland, S., & Eble, J. A. (2019). Neuropilins in the Context of Tumor Vasculature. *International journal of molecular sciences, 20*(3), 639. DOI: 10.3390/ijms20030639

38. Bassilana, F., Nash, M., & Ludwig, M. G. (2019). Adhesion G protein-coupled receptors: opportunities for drug discovery. *Nature Reviews Drug Discovery, 18*(11), 869–884. DOI: 10.1038/s41573-019-0039-y

39. Gad, A. A., & Balenga, N. (2020). The Emerging Role of Adhesion GPCRs in Cancer. *ACS pharmacology & translational science, 3*(1), 29–42. DOI: 10.1021/acsptsci.9b00093

40. Guo, S., Zhao, T., Yun, Y., & Xie, X. (2022). Recent progress in assays for GPCR drug discovery. *American journal of physiology - cell physiology, 323*(2), C583–C594. DOI: 10.1152/ajpcell.00464.2021

41. Schalop, L. & Allen, J. (2017) GPCRs, Desirable Therapeutic Targets in Oncology. *Drug Discovery and Development*. Retrieved from https://www. drugdiscoverytrends.com/gpcrs-desirable-therapeutic-targets-in-oncology/

42. Cattaneo, F., Guerra, G., Parisi, M., De Marinis, M., Tafuri, D., Cinelli, M., & Ammendola, R. (2014). Cell-surface receptors transactivation mediated by g protein-coupled receptors. *International journal of molecular sciences, 15*(11), 19700–19728. DOI: 10.3390/ijms151119700

43. Heldin, C. H., Lu, B., Evans, R., & Gutkind, J. S. (2016). Signals and Receptors. Cold Spring Harbor perspectives in biology, 8(4), a005900. DOI: 10.1101/cshperspect.a005900

44. Zeng, C. M., Chen, Z., & Fu, L. (2018). Frizzled Receptors as Potential Therapeutic Targets in Human Cancers. *International journal of molecular sciences, 19*(5), 1543. DOI: 10.3390/ijms19051543

45. Yue, Z., Yuan, Z., Zeng, L., Wang, Y., Lai, L., Li, J., … Liu, M. (2018). LGR4 modulates breast cancer initiation, metastasis, and cancer stem cells. *FASEB journal, 32*(5), 2422–2437. DOI: 10.1096/fj.201700897R

46. Campos-Contreras, A. D. R., Díaz-Muñoz, M., & Vázquez-Cuevas, F. G. (2020). Purinergic Signaling in the Hallmarks of Cancer. *Cells, 9*(7), 1612. DOI: 10.3390/cells9071612

47. Woods, L. T., Forti, K. M., Shanbhag, V. C., Camden, J. M., & Weisman, G. A. (2021). P2Y receptors for extracellular nucleotides: Contributions to cancer progression and therapeutic implications. *Biochemical pharmacology, 187*, 114406. DOI: 10.1016/j.bcp.2021.114406

48. Aust, G., Zheng, L., & Quaas, M. (2022). To Detach, Migrate, Adhere, and Metastasize: CD97/ADGRE5 in Cancer. *Cells, 11*(9), 1538. DOI: 10.3390/cells11091538

49. Wu, V., Yeerna, H., Nohata, N., Chiou, J., Harismendy, O., Raimondi, F., … Gutkind, J. S. (2019). Illuminating the Onco-GPCRome: Novel G protein-coupled receptor-driven oncocrine networks and targets for cancer immunotherapy. *The Journal of biological chemistry, 294*(29), 11062–11086. DOI: 10.1074/jbc.REV119.005601

50. Liu, S., Liang, J., Liu, Z., Zhang, C., Wang, Y., Watson, A. H., … Wang, X. (2021). The Role of CD276 in Cancers. *Frontiers in oncology, 11*, 654684. DOI: 10.3389/fonc.2021.654684

51. Lupo, K. B., & Matosevic, S. (2020). CD155 immunoregulation as a target for natural killer cell immunotherapy in glioblastoma. *Journal of hematology & oncology, 13*(1), 76. DOI: 10.1186/s13045-020-00913-2

52. Terry, S., Savagner, P., Ortiz-Cuaran, S., Mahjoubi, L., Saintigny, P., Thiery, J. P., & Chouaib, S. (2017). New insights into the role of EMT in tumor immune escape. *Molecular oncology, 11*(7), 824–846. DOI: 10.1002/1878-0261.12093

53. Friedl, P., & Alexander, S. (2011). Cancer invasion and the microenvironment: plasticity and reciprocity. *Cell, 147*(5), 992–1009. DOI: 10.1016/j.cell.2011.11.016

54. Mullooly, M., McGowan, P. M., Kennedy, S. A., Madden, S. F., Crown, J., O' Donovan, N., & Duffy, M. J. (2015). ADAM10: a new player in breast cancer progression? *British journal of cancer, 113*(6), 945–951. DOI: 10.1038/bjc.2015.288

55. Sutherland, R., Meeson, A., & Lowes, S. (2020). Solute transporters and malignancy: establishing the role of uptake transporters in breast cancer and breast cancer metastasis. *Cancer metastasis reviews, 39*(3), 919–932. DOI: 10.1007/s10555-020-09879-6

56. Ren, W., Liu, G., Yin, J., Tan, B., Wu, G., Bazer, F. W., … Yin, Y. (2017). Amino-acid transporters in T-cell activation and differentiation. *Cell death & disease, 8*(3), e2655. DOI: 10.1038/cddis.2016.222

57. Klumpp, L., Sezgin, E. C., Eckert, F., & Huber, S. M. (2016). Ion Channels in Brain Metastasis. *International journal of molecular sciences, 17*(9), 1513. DOI: 10.3390/ijms17091513

58. Xiao, H., Zheng, Y., Ma, L., Tian, L., & Sun, Q. (2021). Clinically-Relevant ABC Transporter for Anti-Cancer Drug Resistance. *Frontiers in pharmacology, 12*, 648407. DOI: 10.3389/fphar.2021.648407

59. Moser, T. L., Kenan, D. J., Ashley, T. A., Roy, J. A., Goodman, M. D., Misra, U. K., … Pizzo, S. V. (2001). Endothelial cell surface F1-F0 ATP synthase is active in ATP synthesis and is inhibited by angiostatin. *Proceedings of the National Academy of Sciences of the United States of America, 98*(12), 6656–6661. DOI: 10.1073/pnas.131067798

60. Kim, B. W., Lee, C. S., Yi, J. S., Lee, J. H., Lee, J. W., Choo, H. J., … Ko, Y. G. (2010). Lipid raft proteome reveals that oxidative phosphorylation system is associated with the plasma membrane. *Expert review of proteomics, 7*(6), 849–866. DOI: 10.1586/epr.10.87

61. Chang, H. Y., Huang, T. C., Chen, N. N., Huang, H. C., & Juan, H. F. (2014). Combination therapy targeting ectopic ATP synthase and 26S proteasome induces ER stress in breast cancer cells. *Cell death & disease, 5*(11), e1540. DOI: 10.1038/cddis.2014.504

62. Vultaggio-Poma, V., Sarti, A. C., & Di Virgilio, F. (2020). Extracellular ATP: A Feasible Target for Cancer Therapy. *Cells, 9*(11), 2496. DOI: 10.3390/cells9112496

63. Kepp, O., Bezu, L., Yamazaki, T., Di Virgilio, F., Smyth, M. J., Kroemer, G., & Galluzzi, L. (2021). ATP and cancer immunosurveillance. *The EMBO journal, 40*(13), e108130. DOI: 10.15252/embj.2021108130

64. Speransky, S., Serafini, P., Caroli, J., Bicciato, S., Lippman, M. E., & Bishopric, N. H. (2019). A novel RNA aptamer identifies plasma membrane ATP synthase beta subunit as an early marker and therapeutic target in aggressive cancer. *Breast cancer research and treatment, 176*(2), 271–289. DOI: 10.1007/s10549-019-05174-3

65. Antico Arciuch, V. G., Elguero, M. E., Poderoso, J. J., & Carreras, M. C. (2012). Mitochondrial regulation of cell cycle and proliferation. *Antioxidants & redox signaling, 16*(10), 1150–1180. DOI: 10.1089/ars.2011.4085

66. Liu, Z., Sun, Y., Tan, S., Liu, L., Hu, S., Huo, H., Li, M., Cui, Q., & Yu, M. (2016). Nutrient deprivation-related OXPHOS/glycolysis interconversion via HIF-1α/C-MYC pathway in U251 cells. *Tumour biology, 37*(5), 6661–6671. DOI: 10.1007/s13277-015-4479-7

67. Ashton, T. M., McKenna, W. G., Kunz-Schughart, L. A., & Higgins, G. S. (2018). Oxidative Phosphorylation as an Emerging Target in Cancer Therapy. *Clinical cancer research, 24*(11), 2482–2490. DOI: 10.1158/1078-0432.CCR-17-3070

68. Huang, L., & Fu, L. (2015). Mechanisms of resistance to EGFR tyrosine kinase inhibitors. *Acta pharmaceutica Sinica. B, 5*(5), 390–401. DOI 10.1016/j.apsb.2015.07.001

69. Xu, H., Niu, M., Yuan, X., Wu, K., & Liu, A. (2020). CD44 as a tumor biomarker and therapeutic target. *Experimental hematology & oncology, 9*(1), 36. DOI: 10.1186/s40164-020-00192-0

CHAPTER 5. CANCER HALLMARKS AND MUTATIONS

5.1 Introduction

The cancer hallmarks were first described by Hanahan and Weinberg,[1,2] and evolved to include ten major biological capabilities, i.e., sustained proliferative signaling, insensitivity to anti-growth signals, evasion of apoptosis, limitless replicative potential, sustained angiogenesis, tissue invasion and metastasis, reprogramming of energy metabolism, evasion of immune destruction, genome instability, and inflammation. The tumor cell microenvironment is a constantly changing landscape, hosting various growth factors and proteins with critical roles in inter-cellular communication, immune recognition, cell-adhesion, and motility.[3] The aberrant or mutated expression of cell-membrane receptors that interact with this environment is often the driver of uncontrolled cell proliferation. In cancer cells, the mutational rate is elevated, conferring a selective advantage to the tumor cells. This occurs largely due to the corruption of the genome maintenance and immune systems, loss of telomeric DNA, and heightened sensitivity toward mutagenic factors. Mutations can be either heritable (germline), or random (somatic), some of them with driver characteristics that can lead to a carcinogenic phenotype.[2]

Currently, the ICGC/TCGA Consortium reports more than 2.7 million mutations identified in 33 different tumor types.[4,5] The detected mutations are not limited to small-scale substitutions/point mutations or insertions/deletions, but also include larger-scale DNA fusions, copy number variations, and other aberrations in DNA sequence, gene expression, epigenetics, and protein expression or structure.[4] A number of databases exist that catalogue and assign molecular function and clinical relevance to these mutations. Some of the largest curated databases with relevance to cancer and disease include the Human Gene Mutation database (HGMD)[6] with more than 350,000 entries focused on disease-related germline mutations, the HuVarBase (774,000 entries) and MutHTP (200,000 membrane protein entries) databases from the Gromiha lab that integrate mutation information from publicly available data at gene and protein level into an easy-to-use searchable web database,[8] and the Catalogue of Somatic Mutations in Cancer (COSMIC) which is the world's largest catalogue on somatic mutations with more than 23 million genetic variations from targeted gene and genome-wide screens, including the TCGA findings.[9]

So far, most of the publications related to mutations describe data generated by sequencing experiments, while thousands of publicly available proteomics datasets remain unexplored. Scientists are yet unable to mine the full depth of mutational information hidden in the proteome. The links between certain point mutations and protein function remain largely unexplored due to the complexity of protein roles, presence of post-translational modifications (PTMs), and the randomness of mutations. Therefore, the attempts to examine the proteome-level mutational landscapes have lagged behind, and the knowledge about the mutational status of cell membrane proteins-which are at the forefront of signaling events in the cell and one of the main drug targets, is largely missing. Nevertheless, the field of proteogenomics is looking at prosperous findings as technological advancements in computation and modeling, and freely available datasets of mass spectrometry studies on protein expression and phosphorylation activity, are ever expanding. Zhang and Krug have demonstrated that technologies such as reverse-phase protein arrays (RPPA) and mass spectrometry (MS) can generate high-throughput protein data which can offer invaluable insights into tumor proliferation, cell cycle, signaling, and prospective personalized treatment.[10,11] The identified mutations can be assessed for their impact on protein function either experimentally or by using prediction servers such as the Mutation Assessor.[12] Our lab has built a freely available database (XMAn) that can be used by any search engine that can process raw mass spectrometry data to identify mutated peptides.[13] Our database and other similar ones are used in massive computing endeavors to map and profile thousands of PTMs and mutations,[14] highlighting the value of proteogenomic analysis enabled by the availability of mass spectrometry tools.

Having profiled the SKBR3/HER2+ surfaceome, in this work we focused on correlating the cell surface proteins with the cancer hallmarks, identifying their point mutations, and evaluating these mutations in the context of relevant biological processes such that these proteins can be further explored for their suitability for therapeutic targeting. To our knowledge, this is the first work to provide mass spectrometry evidence for mutations found in the cell membrane proteins of SKBR3 cells.

5.2 Methods

HER2+ breast cancer cells, SKBR3, were enriched in cell surface proteins by using three orthogonal methods based on the biotinylation of primary amines, biotinylation of oxidized glycans, and shaving of cell receptors in cell culture, as described in detail in **Section 3.3**. The three samples generated by the cell surface protein enrichment methods will be referred in the following sections as the amine, glyco, and trypsin samples. These cell-membrane protein samples were analyzed by mass spectrometry, as detailed in **Section 3.8.1**. The tools used for performing data analysis and providing the biological interpretation of the cell surface enriched proteins are outlined in **Section 3.13**. An in-house database was created to represent the biological processes of the ten cancer hallmarks (**Table 5-1**).[15] The creation of the databases is detailed in **Section 3.12**. The detected cell membrane proteins in SKBR3 were then aligned to the above-mentioned curated lists of cancer hallmarks and the numbers are reported in **Table 5-1**. The mutated peptides and proteins were identified by using the XMAn database when performing the database searches of raw files, as detailed in **Section 3.10.2**. The mutated proteins were assigned to the cell membrane based on the cell surface in-house database described in **Section 3.12**.

5.3 Results and Discussion

5.3.1. Cancer hallmarks

To assess the involvement of cell surface proteins in cancer promoting processes, an in-house database of cancer hallmarks was built based on GO annotations and aligned with the detected cell surface proteins, as described in **Table 5-1**. An excess of 1,000 cell surface proteins were found to be implicated in hallmark processes. The hallmarks matched by the most proteins included proliferative signaling, invasion and metastasis, immune evasion, and insensitivity to anti-growth signals, with more than 600 cell surface protein matches each. The least represented hallmarks were angiogenesis and limitless replicative potential, with less than 100 proteins. Several research groups[1,2,16-20] had undertaken efforts to categorize certain proteins under the hallmarks of cancer, and most of their findings were in-line with our in-house database assignments. The same is shown for the cancer hallmark proteins from the COSMIC (v94) Cancer Gene Census catalogue (CGC),[21] that are accompanied by their role in each hallmark as promoting, suppressing, or both (**Table 5-1**). The data highlights certain cell surface protein

families such as ERBB, fibroblast growth factor receptor (FGFR), MET, Ras, MTOR, and NOTCH1/2, which are part of multiple cancer hallmarks. A limited group of 138 cell surface proteins that could be mapped to the in-house built hallmark database and at least one additional source such as CGC or literature[1,2,16-21] was subjected to STRING analysis for identifying the biological pathways and processes represented by these proteins (**Figure 5-1 A**).

Table 5-1: Cancer hallmarks defined by GO biological processes and examples of cell surface proteins associated with the hallmarks. Note: Proteins highlighted in bold/italic are part of other hallmark categories, according to GO annotations, than reported in the literature. Reproduced with permission from Lazar, et. al.[15]

Cancer hallmarks[1,2]	**Biological processes associated with the cancer hallmarks based on UniProt/GO annotations**	**Protein IDs -Total/ genome -Cell-membrane -Detected**	**Hallmark proteins associated with the cell membrane [1,2,16-20]**	**Cell membrane proteins present in the CGC with cancer promoting (↑), suppressing (↓), or dual role (↑↓)[21]**
Sustained proliferative signaling	• Cell communication (& regulation) • Signaling (& regulation) • Signaling receptor activity (& regulation)	6,871 4,058 847	RRAS, NRAS, HRAS, KRAS, EGFR, ERBB2, FGFR1/2, MAP2K1/2, MTOR, MET, IGF1R, PIK3CA, PDGFRA, GRB2, ***ABCC1***	NRAS ↑, HRAS ↑, RAC1 ↑, CALR ↑, ACVR1 ↑, DNM2 ↑, EPS15 ↑, ERBB3 ↑, ERBB4 ↑, ERBB2 ↑, EZR ↑, FGFR3 ↑, FGFR1 ↑, NDRG1 ↑, NOTCH1 ↑, NOTCH2 ↑, EGFR ↑, IL6ST ↑, KRAS ↑, PRKAR1A ↑, GNA11 ↑, GNAQ ↑, GNAS ↑, JAK1 ↑, MTOR ↑, MET ↑, FGFR2 ↑, FGFR4 ↑, MYD88 ↑, CTNNB1 ↑, PLCG1 ↑, PIK3CA ↑, PDGFRA ↑, RET ↑
Tissue invasion and metastasis	• Cell adhesion (& regulation) • Cell motility (& regulation) • Actin cytoskeleton organization (& regulation) • ECM organization (& regulation) • Secretion • EMT (& regulation)	4,650 2,646 742	CD44, EPCAM, POSTN, TNC, LGALS1, ABCC1, PIK3CA, ***ABCB1***	RHOA ↑↓, NRAS ↑, NRAS ↑, RAC1 ↑, CALR ↑, ACVR1 ↑, DDX3X ↑↓, ATP1A1 ↑, BMPR1A ↑, DNM2 ↑, ERBB3 ↑, ERBB2 ↑, EZR ↑, FAT1 ↑↓, FGFR1 ↑, FHIT ↓, NDRG1 ↓, NOTCH2 ↑, EGFR ↑, KRAS ↑, PRKAR1A ↓, MTOR ↑, MET ↑, FGFR4 ↑, MYD88 ↑, APC ↑, CTNNB1 ↑, BCORL1 ↑, PLCG1 ↑, PIK3CA ↑, PDGFRA ↑, RET ↑
Evasion of immune destruction	• Immune system process • Innate and adaptive immune response (& regulation) • Antigen processing and presentation (& regulation) • Autophagy (& regulation) • Cellular senescence • Cellular response to hypoxia (& regulation) • Chemotaxis (& regulation) • Chemokine mediated signaling pathway (& regulation) • Chemokine production (& regulation) • (Myeloid/B-cell/T-cell activation and regulation are included)	4,055 2,272 631	CD274, ***SIGLEC6***	NRAS ↓, RAC1 ↑, B2M ↓, EGFR ↑, JAK1 ↓, MET ↑, CTNNB1 ↑, RET ↑

Insensitivity to anti-growth signals (evading growth suppressors)	• Cell cycle, cell division, cell growth (& regulation) • Cell population proliferation (& regulation) • Protein catabolic process (& regulation) • Protein folding (& regulation) • Protein targeting (& regulation) • Signaling pathways & their regulation (PI3K, TOR, Wnt, signal transduction by p53 class mediator)	4,925 1,951 625	APC	RHOA ↑, ATP2B3 ↑, DDX3X ↑, ERBB4 ↑, FAT1 ↑, NDRG1 ↑, NOTCH2 ↑, GNAS ↑, APC ↑, CTNNB1 ↑, PDGFRA ↑
Genome instability	• Cellular response to DNA damage stimulus (& regulation) • Disease variant	4,030 1,659 513	COL7A1, FEN1, MSH6	USP8 ↓, RHOA ↓, NRAS ↑, RAC1 ↓, CLTC ↓, FHIT ↓, LMNA ↓, ERCC4 ↑↓, MET ↑, ATRX ↓, APC ↓, CTNNB1 ↓
Evasion of apoptosis (resisting cell death)	• Cell death (& regulation) • Cell aging (& regulation) • Autophagy (& regulation) • (Includes apoptosis, senescence, necrosis)	2,602 1,139 374	ITGB1, ITGB4, ***ABCB1***	RHOA ↑↓, NRAS ↑, NRAS ↑, RAC1 ↑, CALR ↑, ACVR1 ↑, DDX3X ↓, ATP1A1 ↑, ERBB3 ↑, ERBB4 ↑↓, ERBB2 ↑, EZR ↑, FAT1 ↑, FGFR3 ↑, NOTCH1 ↑, NOTCH2 ↑, EGFR ↑, KRAS ↑↓, PRKAR1A ↑↓, JAK1 ↓, MTOR ↑, MET ↑↓, FGFR2 ↑, MYD88 ↑, APC ↓, CTNNB1 ↑, PLCG1 ↑, PIK3CA ↑, PDGFRA ↓, RET ↑↓
Inflammation	• Inflammatory response (& regulation) • NFKB signaling (& regulation) • I-kappaB kinase/NF-kappaB signaling (& regulation) • Receptor signaling pathw. via STAT (& regulation) • MAPK cascade (& regulation) • TNF mediated signaling pathway (& regulation) • Macrophage activation (& regulation) • Cellular response to cytokine stimulus (& regulation)	2,391 1,323 345	COL1A1, ABCC1/2/3, CFTR, ITGB4, ABCC6, ***DDR1***	RHOA ↓, NRAS ↑, IL6ST ↑, KRAS ↑, MYD88 ↑, RET ↑
Deregulating cellular energetics (reprogramming of energy metabolism)	• Gluconeogenesis (& regulation) • Glycolytic process (& regulation) • Canonical glycolysis • Pyruvate oxidation • TCA cycle • Oxidative phosphorylation • Electron transport chain • ATP metabolic process (& regulation) Carbohydrate metabolic process (& regulation) • Lipid metabolic process (& regulation) • One carbon metabolic process • Choline metabolic process • Cellular response to hypoxia (& regulation) • Peroxisome proliferator activated receptor signaling pathway (& regulation) • Energy homeostasis	2,300 880 292	ATP1B1, GAPDH, IDH2, PFKM, ATP6V1B1, SLC2A1, ***VDAC1***	USP8 ↑, DDX3X ↑, ERBB2 ↑, NOTCH1 ↑, EGFR ↑, PICALM ↑, PAFAH1B2 ↑, KRAS ↑, SDHA ↑, MTOR ↑, CTNNB1 ↑
Sustained angiogenesis	• Angiogenesis (& regulation) • VEGF signaling (& regulation)	504 294 98	HSPG2, THBS1, FLT1, ***NRAS, HRAS, KRAS***,	NRAS ↑, RAC1 ↑, CALR ↑, DNM2 ↑, NOTCH1 ↑, EGFR ↑, KRAS ↑, SDHA ↓, MTOR ↑, MET ↑,

			AGRN*, *SDC1/4*, *LGALS1*, *ABCB1	CTNNB1 ↑, PLCG1 ↑, PIK3CA ↑, PDGFRA ↑↓
Limitless replicative potential	• DNA replication (& regulation) • Chromosome organization (& regulation) • Signal transduction by p53 class mediator (& regulation) • Telomere maintenance (& regulation)	1,497 251 87	***RAP1A*, *ITGB1*, *TPP1***	NRAS ↑, FGFR1 ↑, NDRG1 ↓, NOTCH1 ↑, KRAS ↑, CTNNB1 ↑

Different sources of analysis such as Wiki, Reactome, and KEGG found these cell surface hallmark proteins to represent processes related to disease, proliferation, and cancer-supportive signaling pathways such as Ras-Raf-MAPK, PI3K-Akt-mTOR, as shown in **Figure 5-1 A**.

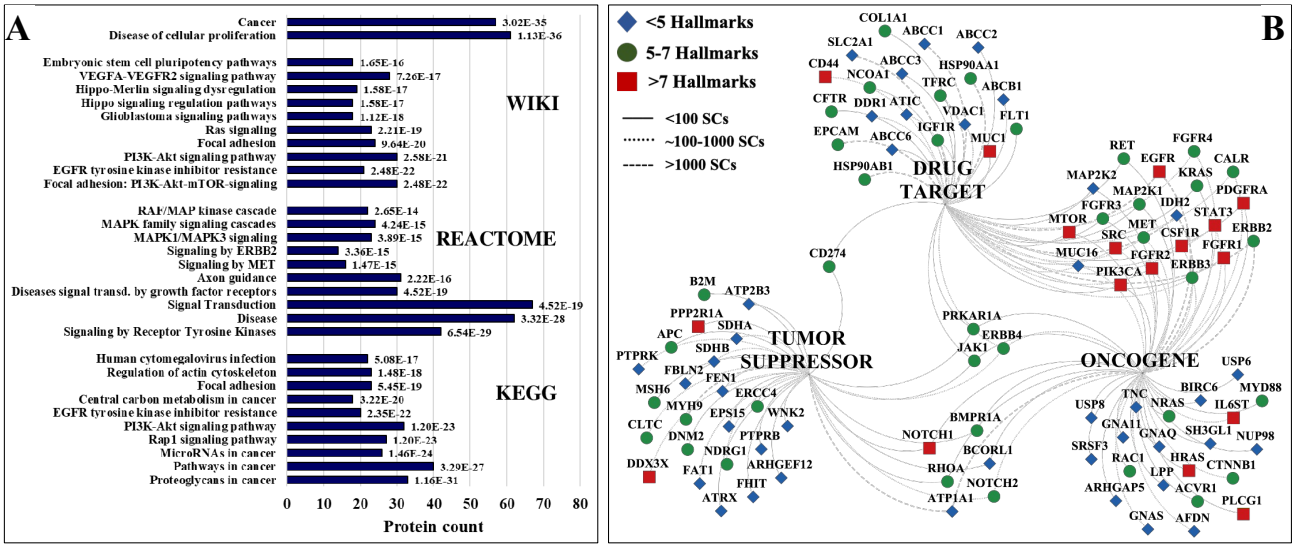


Figure 5-1: Cancer-supportive processes and drug targeting of 138 cell-membrane hallmark proteins. **(A)** Bar chart of top biological processes and pathways represented by the 138 subset of hallmark proteins. Note: Bar chart labels indicate FDRs. **(B)** Overlapping proteins between SKBR3 cell-membrane oncogenes, tumor suppressors, and drug targets. Top/left legend indicates the symbols of proteins matched to a particular number of hallmarks. Line pattern indicates the peptide spectral counts (SCs) matched to each identified protein. Reproduced with permission from Lazar, et. al.[15]

The subset of 138 hallmark proteins comprised 42 oncogenes, 24 tumor suppressors, and 9 proteins with dual roles as oncogenes and tumor suppressors as categorized by the CGC.[21] Many of the cell-membrane oncogenes and tumor suppressors represent cancer drug targets from the DrugBank database[22] either with approved or investigational status, especially the proteins with roles in multiple hallmarks such as EGFR, mTOR, SRC, PIK3CA, CD44 (**Figure 5-1 B**). To broaden the field of hallmark proteins that can be targeted with drugs, PPI networks were constructed with STRING tools for each hallmark from either the 138 subset (**Figure 5-2**) or

from the whole set of cell surface proteins (>1,000) matched to the cancer hallmark database (**Appendix E**).

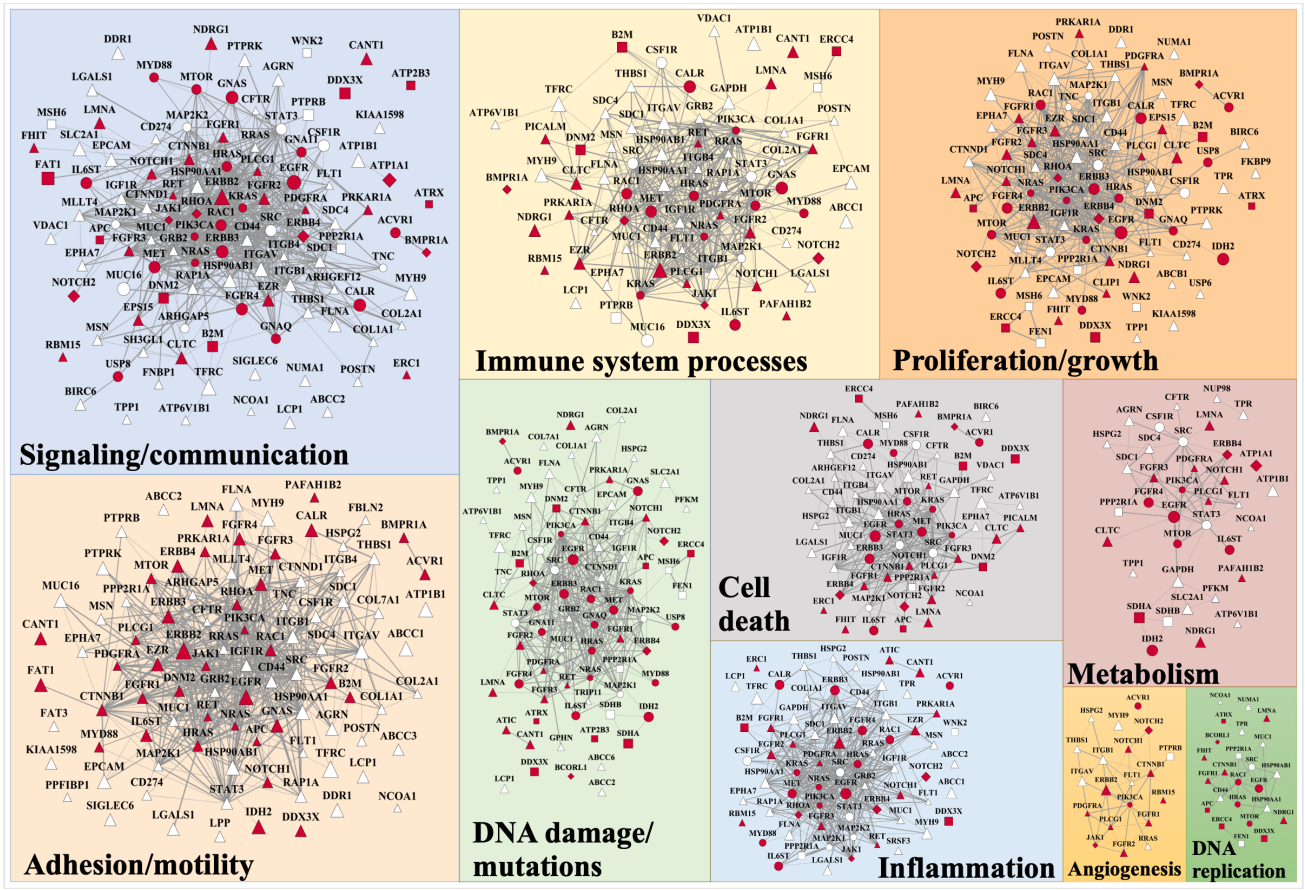


Figure 5-2: PPI networks constructed from 138 cell-membrane proteins matched to ten cancer hallmarks. CGC proteins are represented as oncogenes (◯), suppressors (□), or oncogenes and suppressors (◊). Red icons indicate whether the protein was annotated as a cancer hallmark protein by CGC, other hallmark proteins (Δ). Node size is proportional to the protein abundance based on PSMs and edge thickness reflects the STRING interaction score (≥0.7). Reproduced with permission from Lazar, et. al.[15]

The detected cell surface proteins that were supportive of the various cancer hallmarks were representative of all major classes of cell-membrane proteins, i.e., receptors/catalyst, CDs, adhesion/junction proteins, and transporters.[15] Their presence was actually observed in every hallmark category. Some of the highest abundance proteins had multiple functional roles, were involved in five or more hallmarks (ERBB family, EPCAM/CD326, TFRC/CD71, ATP1A1, ITGAV, ITGB1/CD29, CD44, IGF1R, MET), and were already used as drug targets (**Figure 5-1 B, 5-2**). The hallmarks of cell communication/signaling, adhesion/motility, immune response, and cell cycle/growth encompassed the most complex PPIs, indicating the close and complex

involvement of cell surface proteins in these hallmarks. These PPI networks revealed a broad landscape of interacting proteins that can be used to identify the components of protein complexes, novel regulatory protein functions, as well as novel drug targeting candidates. For example, the RTK receptors interact with each-other (CSF1R, IGF1R, FGFR1-4, ERBB1-4) and with a variety of other cell surface receptors such as CD antigens (CD44), integrins (ITGAV, ITGB1), and the highly conserved non-catalytic NOTCH receptors (NOTCH 1/2), as shown in **Figure 5-2**. An interesting feature of the detected RTKs is their involvement in chemotaxis, a crucial event in the tumor microenvironment that supports metastasis,[23] and immune evasion. Yet, therapies against invasion and dissemination are challenging because chemotaxis occurs during the early stages of cancer metastasis and might be unfeasible to target.[23] The tumor microenvironment and the extracellular matrix (ECM) are also remodeled by the adhesion proteins, such as integrins and CD44, further contributing to invasion and metastasis and becoming potential targets for therapy.[24,25] NOTCH receptors can be an attractive drug target due to implications in proliferation, adhesion, inflammation, and angiogenesis, but their dual role as tumor suppressors and oncogenes (**Figure 5-1 B**) challenges the conventional ways of drug discovery and requires additional studies regarding their involvement in specific pathways in various cancers.[26] The detection of a broad range of proteins interacting with the RTKs, and of major players involved in immune invasion and chemotaxis, will inform the drug discovery efforts that target specifically tumor dissemination, in addition to growth and proliferation targeted strategies, providing yet unexplored opportunities for early detection of cancers.

5.3.2. Mutated cell surface proteins

The detection of mutated peptide sequences by mass spectrometry was enabled by the prior development of a mutated peptide database (XMAn) in our laboratory.[27,28] High-confidence selection of cell surface mutated peptides was performed by using a set of data filtering criteria that included the XCorr vs. charge state parameter with thresholds set to XCorr=2 for 2+ , XCorr=3 for 3+, and XCorr=4 for 4+ charged peptides, and limiting the ion isolation interference to less than 10.1% (a more stringent value than the suggested cutoff of 30%).[29] **Figure 5-3** shows the distribution of the XCorr scores for each charge state as described above.

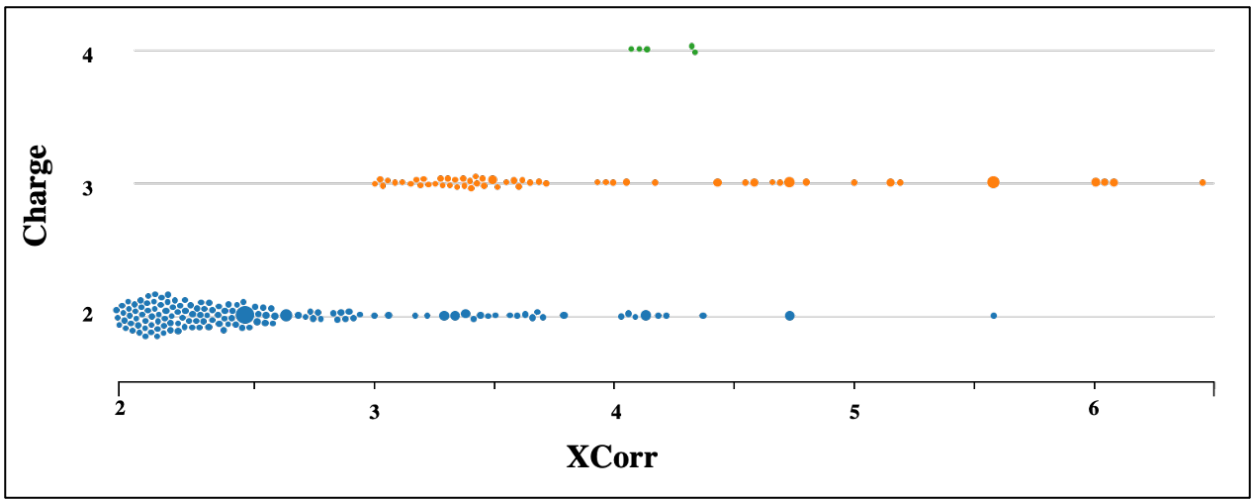
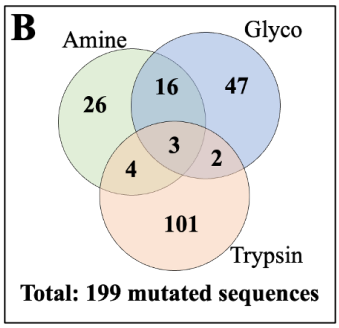


Figure 5-3. XCorr vs charge state distribution of mutated peptides detected with at least two tandem mass spectra. The nodes size is proportional to the abundance of the peptides as represented by the combined spectral counts from each enrichment method.

Mass spectrometry is unable to differentiate between certain substitutions that have very similar masses, such as I↔L (I=L=113.16). These mutations were removed from data interpretation. Based on these criteria, the number of mutated proteins and peptides that resulted from each cell surface protein enrichment method is provided in **Figure 5-4**. The three enrichment methods proved to be complementary.

A	Amine	Glyco	Trypsin	Combined
Mutated proteins (all)	43	55	102	169
Mutated sequences (all)	49	68	110	199
Mutated sequences (pathogenic)	22	39	65	115
Mutated sequences (neutral)	19	24	39	68
Mutated sequences (unknown)	8	5	6	16



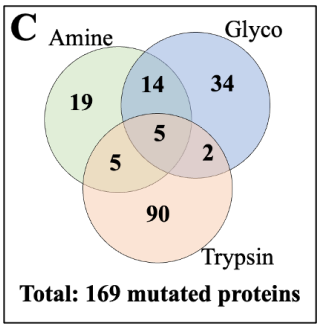


Figure 5-4: Detection of mutated cell surface proteins and peptides. **(A)** Number of mutated proteins and peptides. Overlap of mutated peptides **(B)** and mutated proteins **(C)** between the three cell surface protein enrichment methods.

The COSMIC data that were used to build the XMAn mutated peptide database defined each mutation as having a pathogenic, neutral or unknown function. Our analysis was focused on non-neutral mutations detected by at least two tandem mass spectra across the three enrichment

methods, which resulted in a total of 73 missense and nonsense mutations representing 66 proteins.

Figure 5-5 A provides an overview of the 73 mutations with pathogenic or unknown function normalized to the known frequency of each original amino acid.[30] Among the most frequently mutated amino acids in the SKBR3 cell surfaceome were methionine (Met/M), valine (Val/V), isoleucine (Ile/I), glutamine (Gln/Q), and asparagine (Asn/N). Studies on transmembrane protein mutations have found that the most frequent substitutions related to disease are from non-polar residues to non-polar and charged amino acids.[31] The cell membrane comprises mostly non-polar amino acids (M, V, I) that maintain transmembrane hydrophobicity, therefore, mutations resulting in amino acids with charged side chains, such as the ones detected (e.g., M→R, N→D, Q→K), may affect protein folding due to the polarity they introduce in the transmembrane region of the protein.[31]

A further categorization of the 73 missense mutations based on their BLOSUM62 matrix scores is provided in **Figure 5-5 B**. BLOSUM (Blocks Substitution Matrix) is one of several scoring systems that have been developed to evaluate the likelihood of amino acid substitutions in aligned protein sequences. BLOSUM62 is a commonly used matrix built based on the frequency of substitutions in conserved regions of protein families with less than 62% similarity in local sequence alignments. The scores range from -4 to +11, with a score of zero indicating an equal frequency of observed and expected substitutions, negative scores lower frequencies, and positive scores higher frequencies[32]. The lower the score, the less likely is for the substitution to be preserved in conserved protein regions, reflecting thus significant changes in the amino acid physicochemical properties with higher functional impact. Studies have suggested that cancer driver mutations will have lower BLOSUM scores[33]. The map of mutations from **Figure 5-5 B**, reflects functional impact (color) and occurrence (size). As found in our previous work for mutations that have been identified in whole protein cell extracts,[34] the BLOSUM62 scores calculated for the detected mutations in the cell membrane proteins indicated that most substitutions cluster around zero and one, but a group of negative scores (-1, -2, and -3) reflected non-conservative mutations that could have profound effects on protein structure and function (e.g., I→T, M→R, M→T, R→S, S→C, S→N, L↔R, S→L, I→N, L↔P).

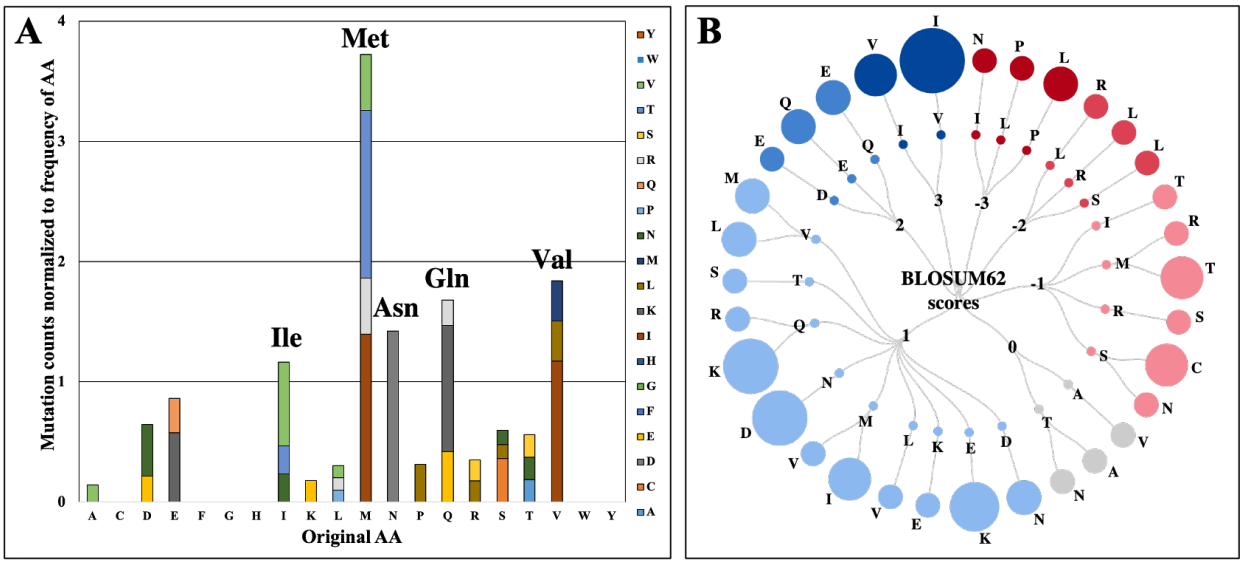


Figure 5-5: Relative frequency of the mutated amino acids and categorization by BLOSUM62 matrix. **(A)** Each stacked column indicates the mutated residues and their counts from the original amino acid. The counts are normalized to the frequency of the original amino acid in the human proteome. **(B)** Mutations with pathogenic and unknown roles categorized by their BLOSUM62 score. The node size is proportional to the number of peptides carrying each mutation, the color is representative of the scores in the BLOSUM62 matrix (0-grey, <0-red, >0-blue).

The landscape of all 66 proteins affected by the 73 mutations is displayed in **Figure 5-6**. Most of the mutated sequences carried missense mutations that are pathogenic, with only a few carrying nonsense mutations (*) or having unknown function (white highlights).

Altogether, many of the proteins carrying the pathogenic mutations are known for their supporting role in cancer development, the detected mutations having possible implications in disease progression and development of therapeutic resistance. These proteins included receptor tyrosine or serine/threonine kinases (ERBB2, ERBB4, DAPK1), CD antigen (BCAM), adhesion proteins (FN1, CXADR, ANXA5), and transport proteins (SLC3A2, ATP8A1, ATP1A3, TRPV5).

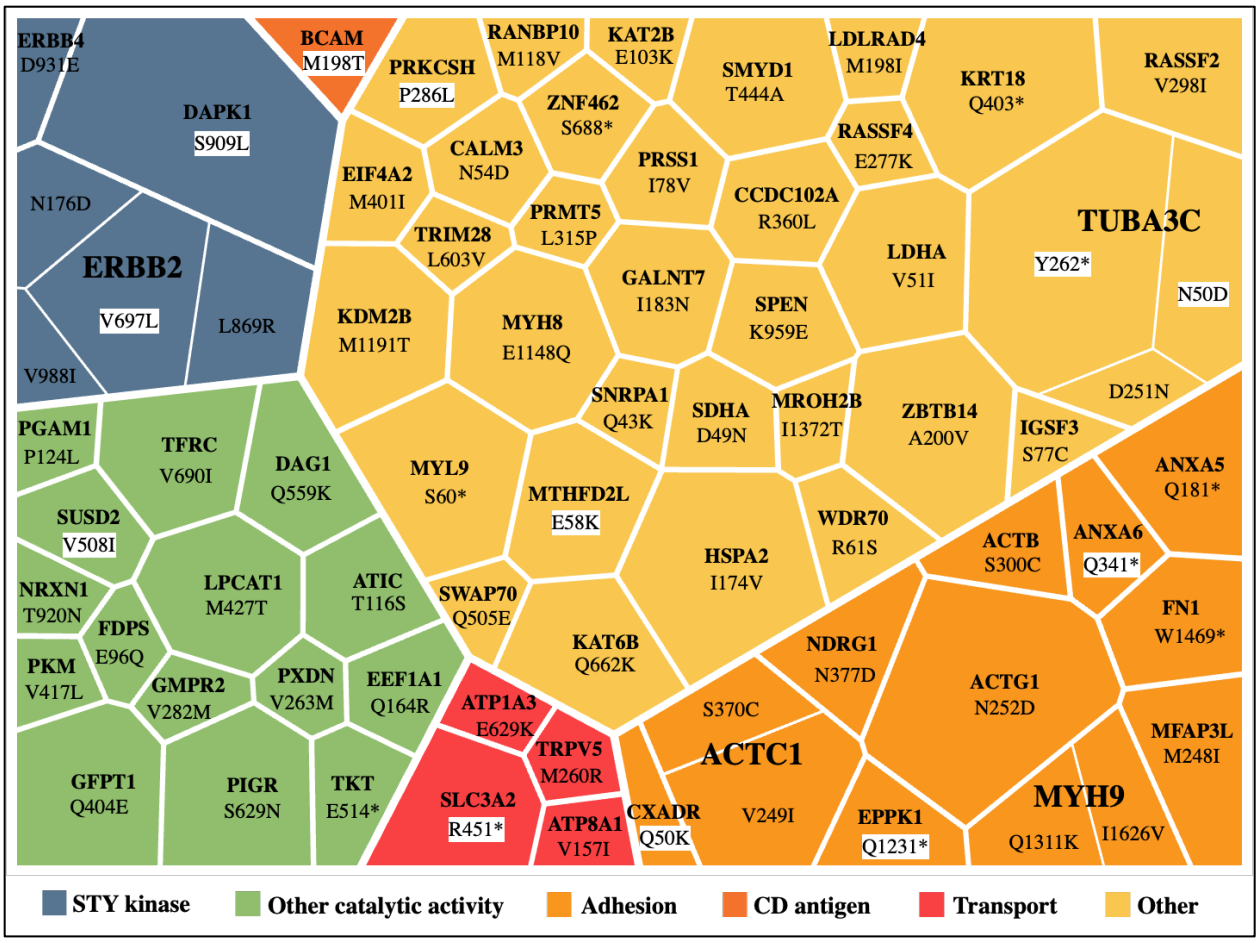


Figure 5-6: Landscape of 66 detected cell membrane proteins that carry mutations with pathogenic or unknown role categorized by their function. The size is proportional to the protein abundance, as reflected by the total spectral counts from all enrichment methods. Mutations highlighted in white indicate mutations with unknown function based on the COSMIC classification.

The PPI network of the 66 mutated proteins uncovered their biological roles in proliferation, cell activation, adhesion, cell-cell junction, carbon metabolism, and also in cancer-supportive pathways such as MAPK (ERBB2, ERBB4, FN1) and mutated Ras (CALM3, FN1) signaling, as well as drug resistance (SLC3A2) (**Figure 5-7 A**).

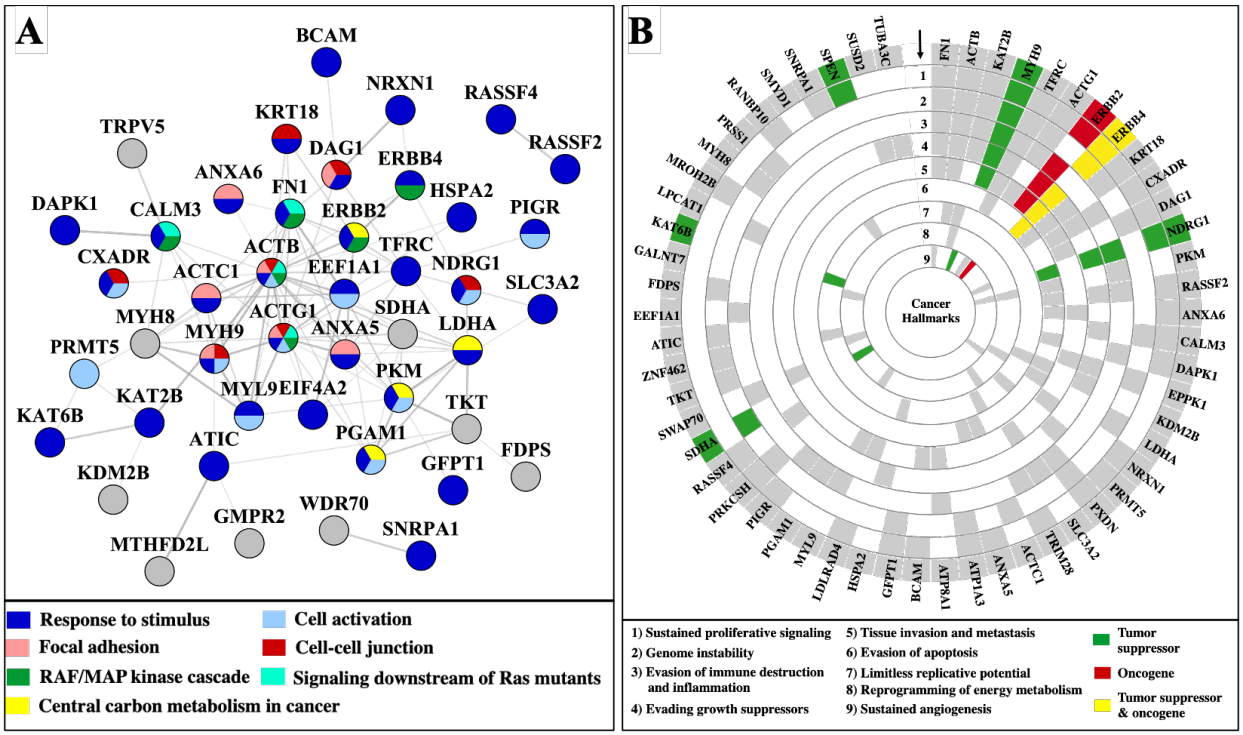


Figure 5-7: Biological processes and cancer hallmarks affected by the mutated proteins. **(A)** PPI network of 66 proteins carrying mutations with pathogenic or unknown role, representing cancer-relevant biological processes and signaling pathways (color-coded). **(B)** Circular chart representing mutated cell-membrane proteins supportive of the cancer hallmarks, with highlighted oncogenes (red), tumor suppressors (green), and dual oncogene/tumor suppressor role (yellow). Note: the immune evasion and inflammation hallmarks are shown together due to the protein overlaps.

Of the 66 proteins, 57 were found to be part of at least one cancer hallmark, almost half were part of three or more hallmarks, and several ones were multifunctional proteins supportive of five or more hallmarks, including ERBB2, ERBB4, FN1, TRFC/CD71 (**Table 5-1** and **Figure 5-7 B**). The number of mutated tumor suppressors (NDRG1, MYH9, SPEN, SDHA, SUSD2) was larger than the number of mutated oncogenes or of proteins with dual oncogene/tumor suppressor roles (ERBB2, ERBB4), inviting further research into the mechanisms through which these proteins exert their role in cancer development or progression. Investigating the mutated surfaceome presents a two-fold research interest, first, due to the importance of cell surface proteins as drug targets and gateway to signaling processes in the cell, and second, due to the impact of mutations on altering the function of proteins and subsequently the downstream signaling pathways.

The protein with the most mutated sites was the receptor kinase ERBB2 (N176D, V697L, L869R, V988I). Other kinases such as EGFR and ERBB4 contained several potential Leu/Ile substitutions. Due to the identical mass of Leu and Ile, this substitution cannot be determined by mass spectrometry, however, if present, can have important implications. Since both Leu (L) and Ile (I) are hydrophobic amino acids, this common substitution is not known to alter protein function but rather alter binding and substrate recognition.[35] Potential presence in receptor kinases such as EGFR/ERBB, which are known to act upon ligand binding and phosphorylate various substrates, can be an indicator of altered binding specificity towards the various ligands and substrates that they modulate, but further research is required to validate the presence of this mutation and elucidate its potential role in cancer. On the other hand, the substitutions of Leu to either Arg (R) or Pro (P), such as in ERBB2 L869R and PRMT5 L315P, are among the ones with the lowest BLOSUM62 scores (**Figure 5-5 B**), leading to potentially significant changes in protein function.

In HER2+ breast cancers, the HER2 receptor amplification or overexpression is a prognostic marker, but mutations in the HER2 protein may have additional role in activating signaling pathways and can be therefore exploited for therapeutic purposes.[36] Among these, the L896R substitution is an activating mutation with roles in ERBB2-ERBB3 dimerization, PI3K-Akt signaling, and cell growth.[37] Due to oncogenic significance, such mutations represent an untapped potential compared to the current targeted therapeutic approaches against HER2 amplification or overexpression that use monoclonal antibodies and small molecule inhibitors. An in-silico study has recently reported the L315P substitution in PRMT5 as a key driver mutation for the PRMT5 enzymatic activity.[38] The protein arginine methyltransferase 5 (PRMT5) has multiple functions in DNA repair, cell cycle regulation, and signal transduction, and is a sought-after drug target with inhibitors currently being in clinical trials[38,39]. As PRMT5 is already an important drug target, further experimental studies are required to validate the mutation impact of L315P in cancer progression or drug resistance. BLOSUM62 scores are one way to assess changes introduced by point mutations, yet they are not informative enough in regard to the potential impact of mutations on the functional role of proteins because no single approach is sufficient for capturing the total activity of a protein[35]. Here, we presented a landscape of SKBR3 mutated cell surface proteins, filtered through multiple parameters (XCorr,

isolation interference, pathogenicity, multiple spectral counts), that are associated with cancer-driving biological processes and hallmarks, with the hope of guiding future targeted therapies against oncogenic point mutations.

5.4 Conclusion

This study presents the results of a proteogenomic approach to detect protein-level mutations in the SKBR3 cell-membrane proteome. The detected proteins were supportive of cancer hallmark processes, being implicated in both the development and metastatic capabilities of these cancer cells. The list of mutated peptides represents highly interconnected, multifunctional proteins with roles in many cancer hallmarks that are current or promising drug targets, and whose mutations can further inform the developmental efforts of potential inhibitors. With the advance of high-throughput MS instrumentation and various PTM and mutation databases, the ability to generate mutated cell-membrane profiles will change the face of targeted therapeutics, highlighting the need for experimental studies that can validate the role of detected mutations. High confidence detection of cell surface proteins and their mutations will facilitate the selection of patient-tailored drug targets for achieving superior therapeutic outcomes in the various stages of cancer development.

5.5 References

1. Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of Cancer: The Next Generation. *Cell, 144*(5), 646-674. DOI: 10.1016/j.cell.2011.02.013

2. Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of Cancer: The Next Generation. *Cell, 144*(5), 646-674. DOI: 10.1016/j.cell.2011.02.013

3. McAllister, S. S., & Weinberg, R. A. (2014). The tumour-induced systemic environment as a critical regulator of cancer progression and metastasis. *Nature cell biology, 16*(8), 717–727. DOI: 10.1038/ncb3015

4. National Cancer Institute. (2019). Outcomes & Impact of The Cancer Genome Atlas. *TCGA*. Retrieved from https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga/history

5. National Cancer Institute. (2022). Data Portal Summary. *GDC Data Portal*. Retrieved from https://portal.gdc.cancer.gov/

6. Stenson, P. D., Mort, M., Ball, E. V., Chapman, M., Evans, K., Azevedo, L., … Cooper, D. N. (2020). The Human Gene Mutation Database (HGMD®): optimizing its use in a clinical diagnostic or research setting. *Human genetics, 139*(10), 1197–1207. DOI: 10.1007/s00439-020-02199-3

7. Kulandaisamy, A., Binny Priya, S., Sakthivel, R., Tarnovskaya, S., Bizin, I., Hönigschmid, P., … Gromiha, M. M. (2018). MutHTP: mutations in human transmembrane proteins. *Bioinformatics, 34*(13), 2325–2326. DOI: 10.1093/bioinformatics/bty054

8. Popov, P., Bizin, I., Gromiha, M., A, K., & Frishman, D. (2019). Prediction of disease-associated mutations in the transmembrane regions of proteins with known 3D structure. *PloS one, 14*(7), e0219452. DOI: 10.1371/journal.pone.0219452

9. Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., … Forbes, S. A. (2019). COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic acids research, 47*(D1), D941–D947. DOI: 10.1093/nar/gky1015

10. Zhang, Y., Kwok-Shing Ng, P., Kucherlapati, M., Chen, F., Liu, Y., Tsang, Y. H., … Creighton, C. J. (2017). A Pan-Cancer Proteogenomic Atlas of PI3K/AKT/mTOR Pathway Alterations. *Cancer cell, 31*(6), 820–832.e3. DOI: 10.1016/j.ccell.2017.04.013

11. Krug, K., Jaehnig, E. J., Satpathy, S., Blumenberg, L., Karpova, A., Anurag, M., … Clinical Proteomic Tumor Analysis Consortium (2020). Proteogenomic Landscape of Breast Cancer Tumorigenesis and Targeted Therapy. *Cell, 183*(5), 1436–1456.e31. DOI: 10.1016/j.cell.2020.10.036

12. Reva, B., Antipin, Y., & Sander, C. (2011). Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic acids research, 39*(17), e118. DOI: 10.1093/nar/gkr407

13. Flores, M. A., & Lazar, I. M. (2020). XMAn v2-a database of Homo sapiens mutated peptides. *Bioinformatics, 36*(4), 1311–1313. DOI: 10.1093/bioinformatics/btz693

14. Prakash, A., Taylor, L., Varkey, M., Hoxie, N., Mohammed, Y., Goo, Y. A., … Orsburn, B. C. (2021). Reinspection of a Clinical Proteomics Tumor Analysis Consortium (CPTAC) Dataset with Cloud Computing Reveals Abundant Post-Translational Modifications and Protein Sequence Variants. *Cancers, 13*(20), 5034. DOI: 10.3390/cancers13205034

15. Lazar, I. M., Karcini, A., & Haueis, J. R. S. (2022). Mapping the cell-membrane proteome of the SKBR3/HER2+ cell line to the cancer hallmarks. *PloS one, 17*(8), e0272384. DOI: 10.1371/journal.pone.0272384

16. Munkley, J., & Elliott, D. J. (2016). Hallmarks of glycosylation in cancer. *Oncotarget, 7*(23), 35478–35489. DOI: 10.18632/oncotarget.8155

17. Muriithi, W., Macharia, L. W., Heming, C. P., Echevarria, J. L., Nyachieo, A., Filho, P. N., & Neto, V. M. (2020). ABC transporters and the hallmarks of cancer: roles in cancer aggressiveness beyond multidrug resistance. *Cancer biology & medicine, 17*(2), 253–269. DOI: 10.20892/j.issn.2095-3941.2019.0284

18. Abcam. (2020). Studying hallmarks of cancer. Retrieved from https://www.abcam.com/cancer/studying-hallmarks-of-cancer

19. Pickup, M. W., Mouw, J. K., & Weaver, V. M. (2014). The extracellular matrix modulates the hallmarks of cancer. *EMBO reports, 15*(12), 1243–1253. DOI: 10.15252/embr.201439246

20. Gutschner, T., & Diederichs, S. (2012). The hallmarks of cancer: a long non-coding RNA point of view. *RNA biology, 9*(6), 703–719. DOI: 10.4161/rna.20481

21. Sondka, Z., Bamford, S., Cole, C. G., Ward, S. A., Dunham, I., & Forbes, S. A. (2018). The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. *Nature Reviews Cancer, 18*(11), 696–705. DOI: 10.1038/s41568-018-0060-1

22. Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., … Wilson, M. (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic acids research, 46*(D1), D1074–D1082. DOI: 10.1093/nar/gkx1037

23. Roussos, E. T., Condeelis, J. S., & Patsialou, A. (2011). Chemotaxis in cancer. *Nature Reviews Cancer, 11*(8), 573–587. DOI: 10.1038/nrc3078

24. Hamidi, H., & Ivaska, J. (2018). Every step of the way: integrins in cancer progression and metastasis. *Nature Reviews Cancer, 18*(9), 533–548. DOI: 10.1038/s41568-018-0038-z

25. Senbanjo, L. T., & Chellaiah, M. A. (2017). CD44: A Multifunctional Cell Surface Adhesion Receptor Is a Regulator of Progression and Metastasis of Cancer Cells. *Frontiers in cell and developmental biology, 5*, 18. DOI: 10.3389/fcell.2017.00018

26. Majumder, S., Crabtree, J. S., Golde, T. E., Minter, L. M., Osborne, B. A., & Miele, L. (2021). Targeting Notch in oncology: the path forward. *Nature Reviews Drug discovery, 20*(2), 125–144. DOI: 10.1038/s41573-020-00091-3

27. Yang, X., & Lazar, I. M. (2014). XMAn: a Homo sapiens mutated-peptide database for the MS analysis of cancerous cell states. *Journal of proteome research, 13*(12), 5486–5495. DOI: 10.1021/pr5004467

28. Flores, M. A., & Lazar, I. M. (2020). XMAn v2-a database of Homo sapiens mutated peptides. *Bioinformatics, 36*(4), 1311–1313. DOI: 10.1093/bioinformatics/btz693

29. Sandberg, A., Branca, R. M., Lehtiö, J., & Forshed, J. (2014). Quantitative accuracy in mass spectrometry based proteomics of complex samples: the impact of labeling and precursor interference. *Journal of proteomics, 96*, 133–144. DOI: 10.1016/j.jprot.2013.10.035

30. UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic acids research, 47*(D1), D506–D515. DOI: 10.1093/nar/gky1049. Retrieved from ….

31. Molnár, J., Szakács, G., & Tusnády, G. E. (2016). Characterization of Disease-Associated Mutations in Human Transmembrane Proteins. *PloS one, 11*(3), e0151760. DOI: 10.1371/journal.pone.0151760

32. Henikoff, S., & Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences of the United States of America, 89*(22), 10915–10919. DOI: 10.1073/pnas.89.22.10915

33. Espinosa, O., Mitsopoulos, K., Hakas, J., Pearl, F., & Zvelebil, M. (2014). Deriving a mutation index of carcinogenicity using protein structure and protein interfaces. *PloS one, 9*(1), e84598. DOI: 10.1371/journal.pone.0084598

34. Lazar, I. M., Karcini, A., Ahuja, S., & Estrada-Palma, C. (2019). Proteogenomic Analysis of Protein Sequence Alterations in Breast Cancer Cells. *Scientific reports, 9*(1), 10381. DOI: 10.1038/s41598-019-46897-z

35. Betts, M.J., & Russell, R.B. (2003). Amino-Acid Properties and Consequences of Substitutions. In M. R. Barnes, I. C. Gray (eds), *Bioinformatics for Geneticists*, 289-316. John Wiley & Sons. DOI: 10.1002/0470867302.ch14

36. Connell, C. M., & Doherty, G. J. (2017). Activating HER2 mutations as emerging targets in multiple solid cancers. *ESMO open, 2*(5), e000279. DOI: 10.1136/esmoopen-2017-000279

37. Joshi, S. K., Keck, J. M., Eide, C. A., Bottomly, D., Traer, E., Tyner, J. W., … Druker, B. J. (2020). ERBB2/HER2 mutations are transforming and therapeutically targetable in leukemia. *Leukemia, 34*(10), 2798–2804. DOI: 10.1038/s41375-020-0844-7

38. Rasheed, S., Bouley, R. A., Yoder, R. J., & Petreaca, R. C. (2023). Protein Arginine Methyltransferase 5 (PRMT5) Mutations in Cancer Cells. *International journal of molecular sciences, 24*(7), 6042. DOI: 10.3390/ijms24076042

39. Stopa, N., Krebs, J. E., & Shechter, D. (2015). The PRMT5 arginine methyltransferase: many roles in development, cancer and beyond. *Cellular and molecular life sciences, 72*(11), 2041–2059. DOI: 10.1007/s00018-015-1847-9

CHAPTER 6. DRUG TREATMENTS

6.1 Introduction

Breast cancer is a heterogeneous disease, and the HER2 positive breast cancer subtype is characterized by the overexpression of the HER2 receptor in ~20% of breast cancers.[1] The HER2 receptor was among the first ones to be targeted due to its overexpression in breast cancer, and trastuzumab was the first monoclonal antibody drug to be approved by FDA for HER2+ breast cancer. Despite the success of trastuzumab treatment alone or in combination with other drugs and chemotherapeutic agents, resistance to such therapy emerges within a year in most patients.[2] Many other therapeutic agents have emerged since the discovery of trastuzumab in 1998.[3] For example, lapatinib and ipatasertib are two drugs that have been developed to target the HER2 and pan-AKT receptors, respectively, and their downstream MAPK and PI3K-Akt signaling pathways.[4,5] Lapatinib is a dual kinase inhibitor that targets both the HER2 and EGFR receptors and acts as an ATP-binding competitive small molecule in the intracellular catalytic region,[4] while ipatasertib acts as ATP competitive small molecule inhibitor against the AKT kinase.[6] Each of these drugs has shown promise in treating HER2+ breast cancer alone or in combination with other therapeutic agents.[7,8] In around 50% of HER2+ breast cancers, the PI3K-Akt pathway is altered, and its hyperactivation leads to the development of drug resistance.[7] Therefore, drugs that inhibit the kinases in this pathway like ipatasertib are used in HER2+ breast cancer treatments in combination with HER2 targeted therapy such as trastuzumab and lapatinib,[5,7,8] but the combination with the latter one has not been widely researched, especially through mass spectrometry technologies. Combining lapatinib and ipatasertib has been shown to be beneficial in treating HER2+/PI3KCA-mutant breast cancer cells as it enhances the anti-proliferative effects of HER2 targeted therapies through acting on multiple signaling pathways.[8]

The MAPK and PI3K-Akt pathways extensively cross-talk with each other through various mechanisms including cross-activation, cross-inhibition, and convergence,[9] indicating the complexity of signaling networks that act within a cell. This presents challenges, but also untapped opportunities for the investigation and advancement of targeted therapies. In addition, small signaling molecules that can activate other cell surface receptors, mainly GPCRs, can transactivate the MAPK and PI3K-Akt pathways, supporting and further driving cancer cell

proliferation and dissemination.[10,11] Therefore, understanding the crosstalk between the MAPK and PI3K-Akt pathways is crucial for the development of effective cancer therapies that bypass the causes that lead to the development of drug resistance. Ultimately, investigating the role of growth factors and the inhibition of their respective receptors by drugs that act on the MAPK and PI3K-Akt pathways will lead to an improved understanding of cancer progression and survival. This study was aimed at providing new insights into the combinatorial effects of two small molecule inhibitors, lapatinib and ipatasertib, in HER2+ breast cancer cells, by using mass spectrometry technologies. ATP signaling was investigated as a potential crosstalk mechanism to attenuate the effects of kinase inhibitors through the GPCR purinergic receptors, as a result of previous findings that reported the involvement of ATP in the transactivation of RTKs.[12,13] Our previous work on profiling the SKBR3 cell membrane proteome has also pointed to the potential role of ATP in such receptor transactivation processes.[14] However, ATP signaling is complex and triggers conflicting response in the cancer cells, by either promoting or inhibiting tumor progression, depending on the context.[15,16] Nevertheless, the therapeutical implications of ATP and purinergic signaling remain an area of extensive research.[17] Mass spectrometry was utilized for its specificity, sensitivity, quantitative ability, and comprehensive profiling capability of both the proteome and phosphoproteome.[18-20] The response of SKBR3/HER2+ cells, nuclear and cytoplasmic cell fractions, to short (15-30 min) and prolonged (36 h) inhibition with drugs was explored, by performing phosphoproteomic (for 15-30 min treatments) and proteomic (for 36 h treatments) profiling of cells. Mass spectrometric analysis of posttranslational modifications (PTM) has greatly advanced in recent years, providing additional data on the function and role of proteins in a cell.[21] PhosphoSitePlus, the largest repository of phosphorylated sites detected by mass spectrometry, hosts more than 200,000 reported PTM sites with only a fraction having assigned biological function,[22] creating thus a challenge in connecting explicit site modifications with specific cellular events. In this study, the phosphorylation profiles that were altered within the first 15-30 min to drug exposure were mapped to the pathways that affected the long-term response of cells after 36 h of exposure to the drugs, providing insights into the broad landscape of proteins and biological processes that can be explored for the development of future therapeutics and co-targeting strategies.

6.2 Methods

The SKBR3 cells were treated with drugs as described in **Section 3.4** for 15 min, 30 min, and 36 h, to investigate the cellular responses and signaling cascades of early events and changes in protein expression after prolonged exposure. FACS analysis was performed as described in **Section 3.5**. Cells collected after 15 min and 30 min were enriched in phosphorylated peptides based on the procedures outlined in **Section 3.7**, and analyzed by mass spectrometry as described in **Section 3.8.2**. Cells collected after 36 h were separated in nuclear and cytoplasmic fractions as detailed in **Section 3.6** and analyzed by mass spectrometry (**Section 3.8.1**). Data analysis was performed by using in-house built databases and bioinformatics tools outlined in **Sections 3.12** and **3.13**. Differentially expressed proteins were quantified by label free methods based on spectral counting (**Section 3.11.1**) and peak areas (**3.11.2**). Proteins matched by three unique peptides with fold change (FC) ≥2 in spectral counts or peak areas and p-value <0.05, from either quantification method, were considered for discussion. Some of the differentially expressed proteins were validated by parallel reaction monitoring (PRM) as described in **Section 3.9** or by western blot as described in **Section 3.15**.

6.3 Results

6.3.1. Drug treatments and proteome profiling

The SKBR3 drug treatment strategy consisted of a multi-step process (**Figure 6-1 A**): (a) EGF stimulation of cells to assess canonical growth; (b) lapatinib treatment alone; (c) lapatinib and experimental drug ipatasertib treatment; (d) ATP addition to either lapatinib alone or lapatinib/ipatasertib combined to inspect for alternative altered signaling pathways. FBS was added to all cell cultures. The drug treatments were performed for 36 h to allow for the observation of clear changes in the proteome profiles of drug-treated cells vs EGF control. After 36 h, changes in cell morphology and cell adhesion were observable. The proteome profiles of the respective cytoplasmic and nuclear cell fractions yielded ~3,800-4,600 proteins per fraction, treatment, and replicate (**Figure 6-1 Ba**). More than half of the detected proteins were identified by at least 2 peptides (**Figure 6-1 Bb**)-group that was further considered for analysis. The protein identification reproducibility between the three biological replicates was high (>75%) (**Figure 6-1 Bc**). As expected, and as shown by FACS, the presence of drugs led the cells towards an increasingly pronounced G1 phase after 36 h treatment (**Figure 6-2 A**) when compared to EGF

treatment without any drug. Contrary to expectations, the addition of the ATP signaling molecule could not drive the drug-treated cells towards a proliferative behavior. Therefore, major changes in cell response could be inferred from the comparisons of lapatinib and lapatinib/ipatasertib treatments to EGF control.

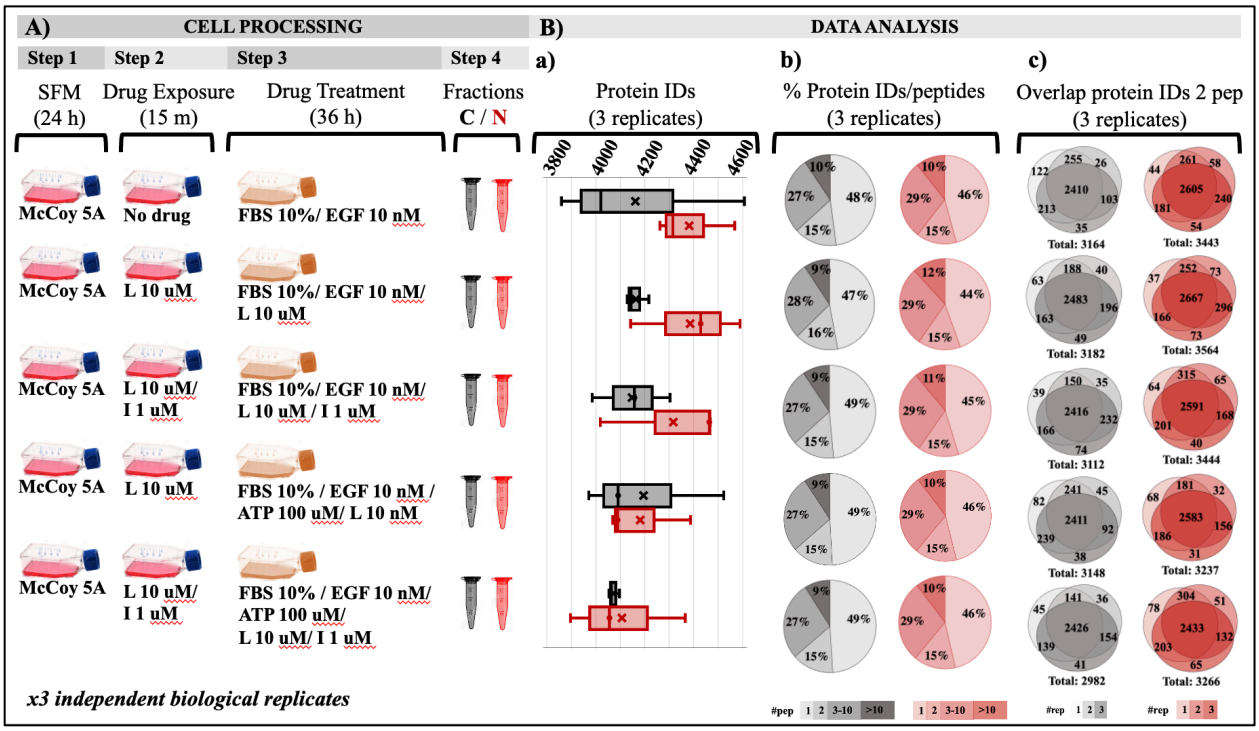


Figure 6-1: Overview of the drug treatment protocol and proteome profiling results based on spectral counting data. **(A)** Drug treatment layout and steps; **(Ba)** Range of detected proteins in the three replicates of each cell fraction; (**Bb**) Pie charts representing the distribution of detected proteins based on the number of unique peptides; (**Bc**) Protein identification reproducibility between three biological replicates (only proteins detected by at least 2 unique peptides were considered). Note: L- lapatinib, I- ipatasertib, C- cytoplasmic fraction (grey), N- nuclear fraction (red).

In addition, the peptide and PSM detection reproducibility was assessed for each treatment based on retention times, XCorr scores, and spectral counts. **Figures 6-2 B** and **6-2 C** show one treatment (EGF) for both cytoplasmic and nuclear fractions as an example of the peptides that co-elute at the same time or have the same XCorr scores across the three different biological replicates represented by X-, Y- axis and color as the third dimension. PSMs reproducibility for any two biological replicates of each treatment and fraction (**Figure 6-2 D, E**) has a correlation score of 0.95 or higher. The relations for the other treatments and replicates can be found in **Appendix F**.

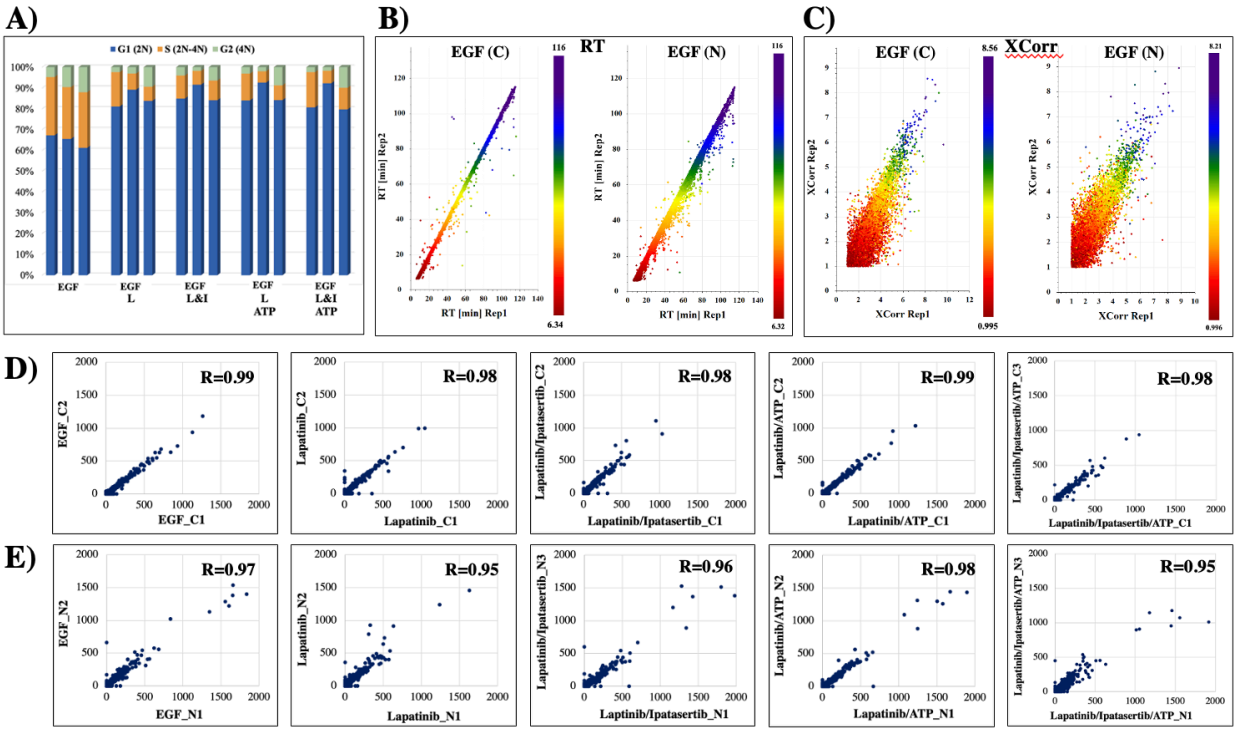


Figure 6-2: Assessment of data reproducibility. **(A)** FACS distribution of cells after 36 h drug treatment in the G1, S, and G2 phases of the cell cycle. **(B)** Correlation of the elution times of peptides from each biological replicate of the cytoplasmic and nuclear fractions of the EGF treatment: replicate 1 (X-axis), replicate 2 (Y-axis), and replicate 3 (color bar). **(C)** Correlation of the XCorr scores of peptides from each biological replicate of the cytoplasmic and nuclear fractions of the EGF treatment: replicate 1 (X-axis), replicate 2 (Y-axis), and replicate 3 (color bar). **(D)** Correlation of the spectral counts for any two biological replicates represented by X-axis (replicate 1) and Y-axis (replicate 2 or 3) for the cytoplasmic fractions of all the treatments. **(E)** Correlation of the spectral counts for any two biological replicates represented by X-axis (replicate 1) and Y-axis (replicate 2 or 3) for the nuclear fractions of all the treatments. Note: L- lapatinib, I- ipatasertib, C- cytoplasmic fraction, N- nuclear fraction.

6.3.2. Differentially expressed proteins in the nuclear and cytoplasmic cell fractions

The nuclear and cytoplasmic fractionation of cells collected from each treatment resulted in the detection of two large protein groups, where each fraction was enriched by ~70% or more in their top100 peptide number-based selected proteins related to their specific location (nuclear or cytoplasmic), as determined by the CC location assignment in UniProt (**Figure 6-3 A**). To identify the group of proteins that were differentially expressed under each treatment and fraction, label-free quantitative analysis based on spectral counts and peak areas was performed for the treatment vs. control groups. Five comparisons were performed (**Figure 6-3 B**). Spectral count and peak area comparisons provided complementary results, and both were used in the

analysis to increase the overall detection effectiveness of differentially expressed proteins, and to also provide confidence in reporting differentially expressed proteins that are not discriminated by statistical computations. The complementarity of these two methods has been documented in literature, where the main advantage of the spectral counting method is the sampling of a larger range of protein abundances, while that of the peak area measurement method is related to greater accuracy in measuring the protein abundance ratios with overlapping peptide ions.[23]

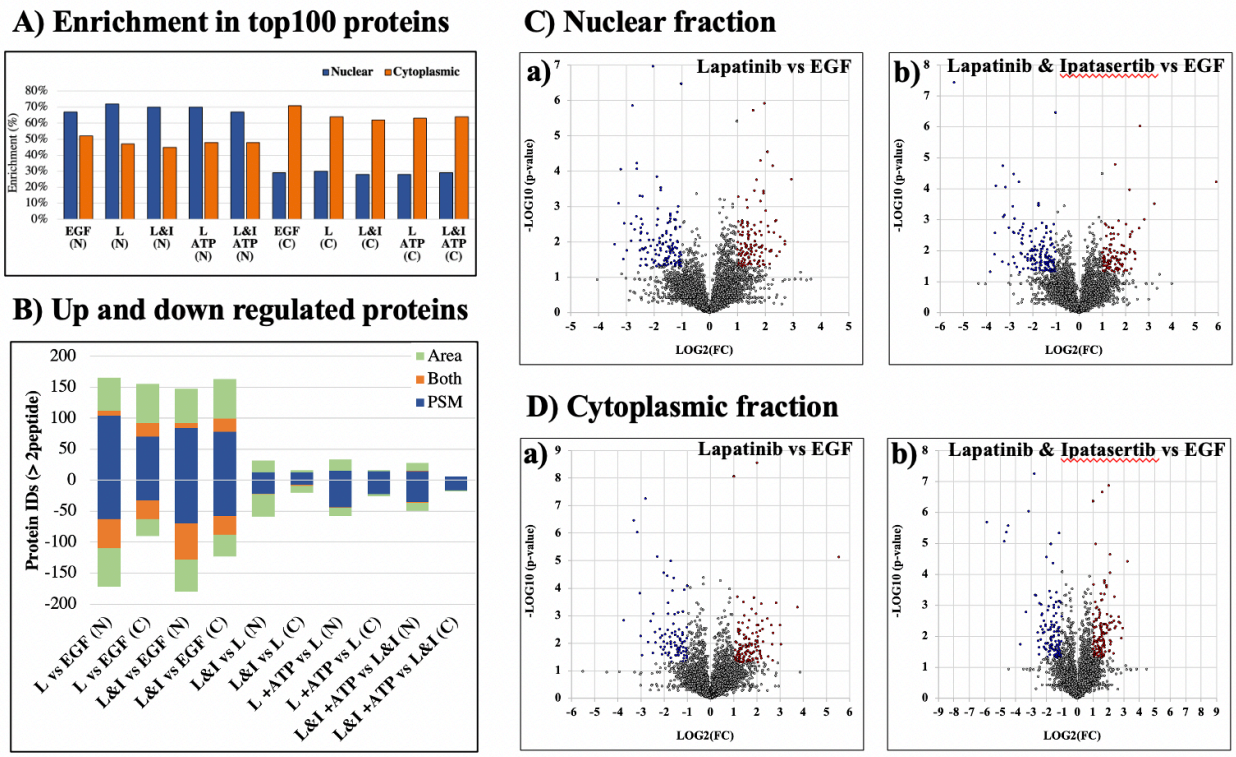


Figure 6-3: Results of the nuclear and cytoplasmic fraction enrichment process and differential expression analysis. **(A)** Nuclear and cytoplasmic protein enrichment among the top 100 most abundant proteins-as represented by total peptide matches per protein, calculated based on the CC location provided by UniProt. **(B)** Counts of differentially expressed proteins by each label-free quantification method for all comparisons and fractions. Up-regulated proteins are represented by positive values, while the down-regulated proteins are represented by negative values. **(C)** and **(D)** Volcano plots of nuclear and cytoplasmic fractions, respectively, of Lapatinib (Ca, Da) and Lapatinib&Ipatasertib (Cb, Db) treatments compared to EGF treatment control. Differentially expressed proteins are indicated in red (up-regulated) and blue (down-regulated), and display ≥2-fold change abundance ratio with p-value ≤0.05. Note: L- Lapatinib, I- Ipatasertib, C- cytoplasmic fractions, N- nuclear fractions.

To take advantage of the sensitivity of spectral counting and the accuracy of area measurement methods, the differentially expressed proteins detected by at least 3 unique peptides by each

method were combined, and constituted the pool of proteins further presented in this study (**Figure 6-3 B**). Due to the higher sampling capacity of spectral counting (**Figure 6-3 B**), this method was utilized to construct the volcano plots, providing a better visual representation of the up-and down-regulated proteins in each fraction and treatment comparison (**Figure 6-3 C** and **Appendix G**).

6.3.3. Biological interpretation of the differentially expressed proteins

The lapatinib vs. EGF and lapatinib/ipatasertib vs. EGF comparisons yielded the largest number of proteins that displayed a change in expression level (**Figure 6-3**). These proteins were representative of enriched biological processes that are outlined in **Figures 6-4** and **6-5**, as captured by STRING and Gene Ontology (GO) analysis. The other comparisons that included the use of ATP in the treatment cocktail, or that represented the lapatinib/ipatasertib vs lapatinib treatment, resulted in a lower count of differentially expressed proteins, of which, however, several had important roles in signaling and proliferation-to be discussed in the next sections.

Treatment by drugs, either alone or in combination, resulted mostly in the downregulation of cell cycle-related processes including G1/S and G2/M transition, chromosome segregation, and spindle assembly, as largely represented by the nuclear proteins. These processes along with other regulatory events of cell cycle progression were also evident in the cytoplasmic fractions, as mainly represented by proteins with dual localization such as cyclin-dependent kinases (CDKs) or adapter proteins with broad implications in signaling such as 14-3-3 protein sigma. Other down-regulated processes by the drug treatments related to protein folding and adhesion, as mainly documented by cytoplasmic proteins, and translation, metabolism, and signaling events in the nuclear and cytoplasmic fraction, as documented by kinases, adapters, and regulatory proteins. Since the nuclear fractions displayed a greater variety of up-and down-regulated processes, their protein-protein interaction (PPI) networks are shown to inform about the association of the nuclear proteins with the respective cellular processes affected by their change in abundance.

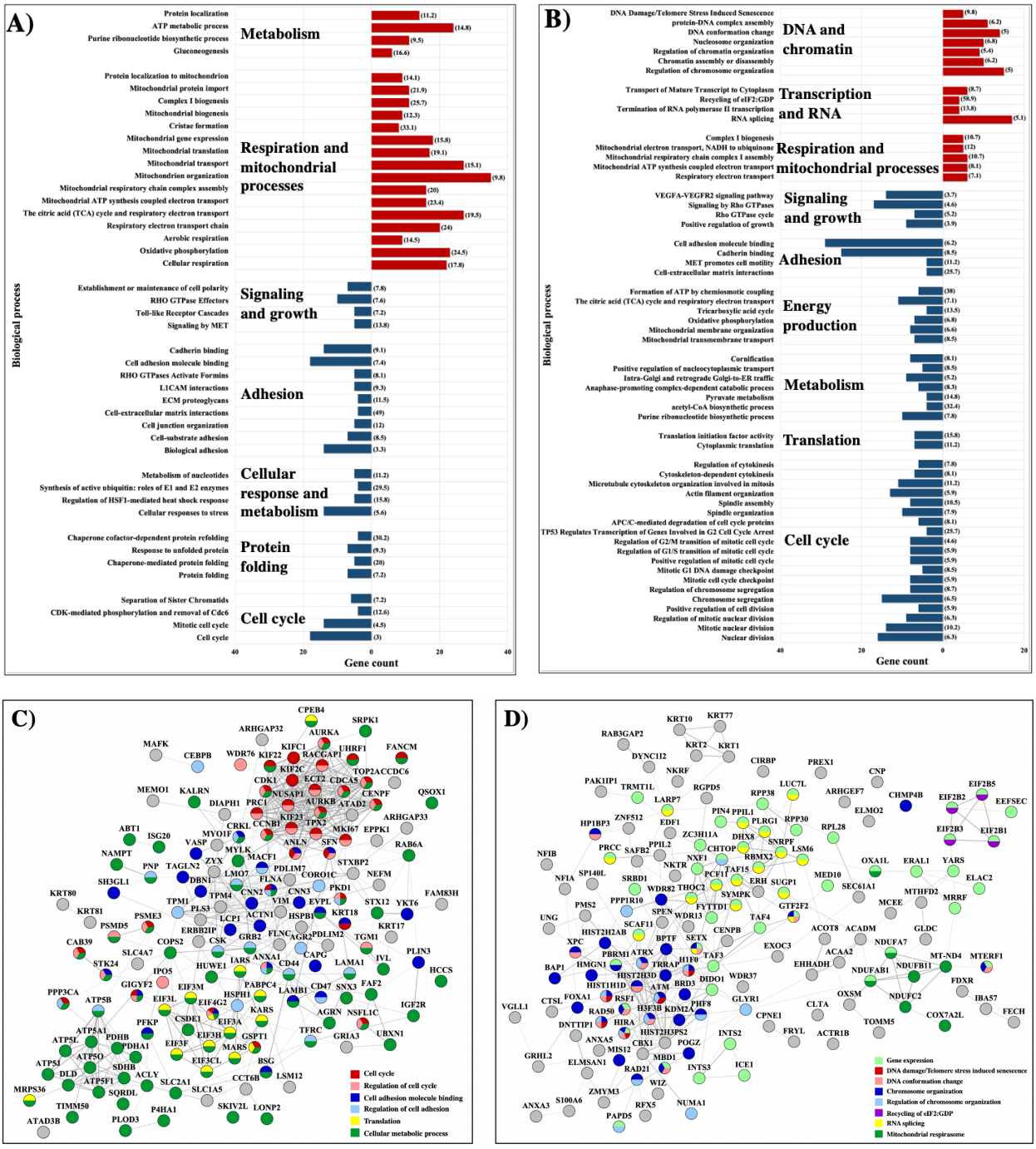


Figure 6-4: Differential expression analysis based on spectral counts and peak areas comparing lapatinib-treated against EGF-treated cells. **(A)** Up- (red) and down- (blue) regulated processes retrieved by STRING/GO for the cytoplasmic fraction. **(B)** Up- (red) and down- (blue) regulated processes retrieved by STRING/GO for the nuclear fraction. **(C)** STRING PPI network of down-regulated nuclear proteins visualized in Cytoscape 3.9.0. **(D)** STRING PPI network of up-regulated nuclear proteins visualized in Cytoscape 3.9.0. Notes: The selected STRING processes have an FDR ≤0.05, at least 3-fold enrichment, and are represented by at least 4 genes/category. The numbers in parenthesis indicates fold-enrichment.

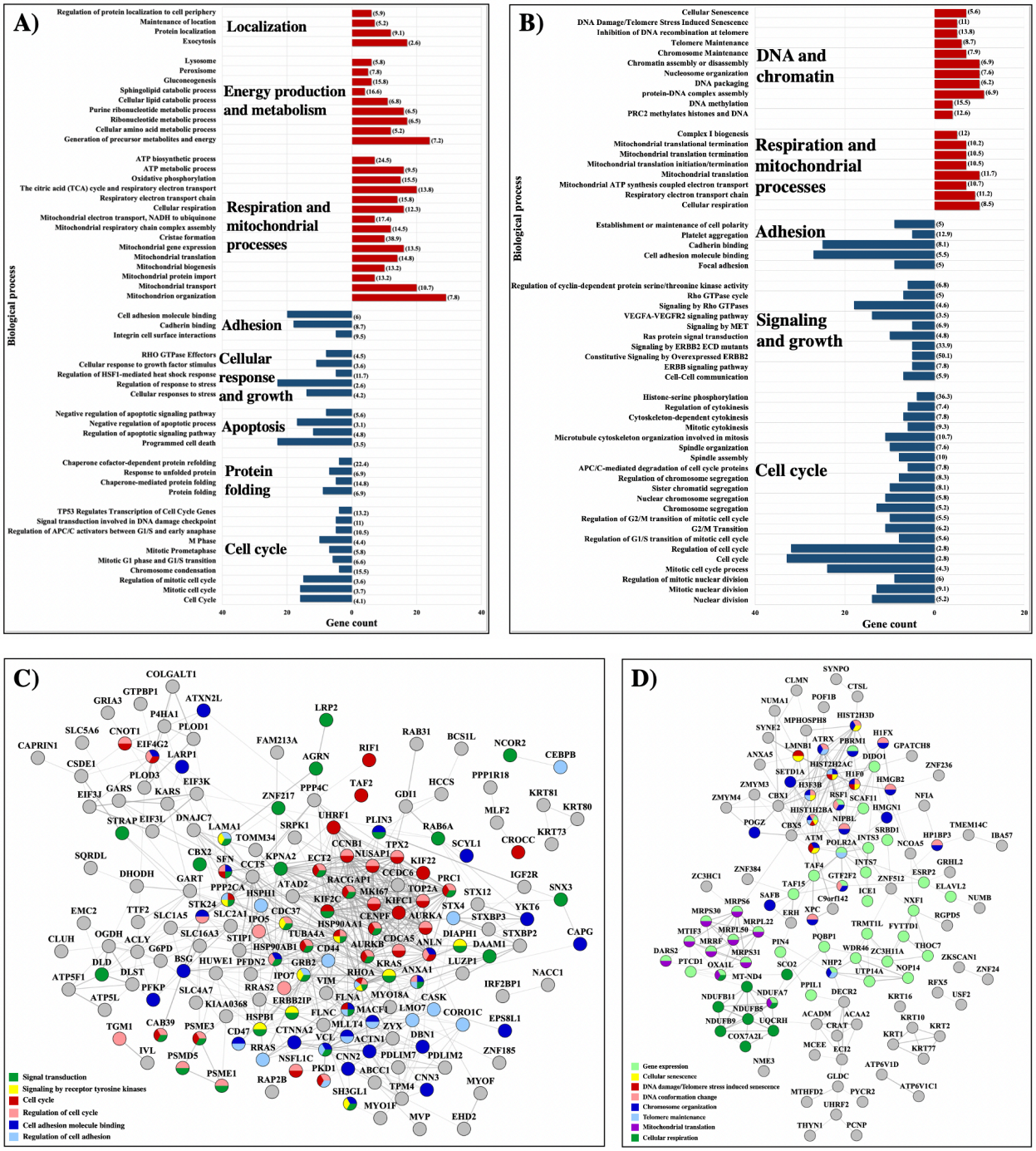


Figure 6-5: Differential expression analysis based on spectral counts and peak areas comparing lapatinib/ipatasertib-treated against EGF-treated cells. **(A)** Up- (red) and down- (blue) regulated processes retrieved by STRING/GO for the cytoplasmic fraction. **(B)** Up- (red) and down- (blue) regulated processes retrieved from STRING/GP for the nuclear fraction. **(C)** STRING PPI network of down-regulated nuclear proteins visualized in Cytoscape 3.9.0. **(D)** STRING PPI network of up-regulated nuclear proteins visualized in Cytoscape 3.9.0. Notes: The selected STRING processes have an FDR ≤0.05, at least 3-fold enrichment, and are represented by at least 4 genes/category. The number in parenthesis indicates fold-enrichment.

In addition to inhibiting major events related to cell cycle progression and signal transduction, the treatment with drugs also resulted in the up-regulation of certain processes that included chromatin organization and DNA conformation, telomere maintenance, DNA damage, and cellular senescence in the nuclear fractions, and broader cellular events related to metabolism, energy production, respiration, and mitochondrial processes in both nuclear and cytoplasmic fractions (**Figure 6-4, 6-5**). When assessing the added effect of ipatasertib on the affected biological processes, certain activities were found to be more enriched than in the lapatinib treatment alone, including negative regulation of apoptosis and ERBB signaling in the down-regulated processes, and protein localization, exocytosis, and cellular senescence in the up-regulated processes (**Figure 6-5 A, B**).

Some of the processes described above are representative of several cancer hallmarks, and the in-house built hallmark database was used to pinpoint the differentially expressed proteins from these categories, identified by either the spectral counting or peak area measurement methods with adjusted p-value ≤0.1, further to be discussed in **Section 6.4** (**Figure 6-6**). Differences in fold-change abundances between the lapatinib and lapatinib/ipatasertib cell treatments are also inferable from this figure.

To confirm the change in abundance between treatments, certain proteins representative of the above processes were validated by independent methods such as parallel-reaction monitoring (PRM) or western blotting (**Appendices H, I**).

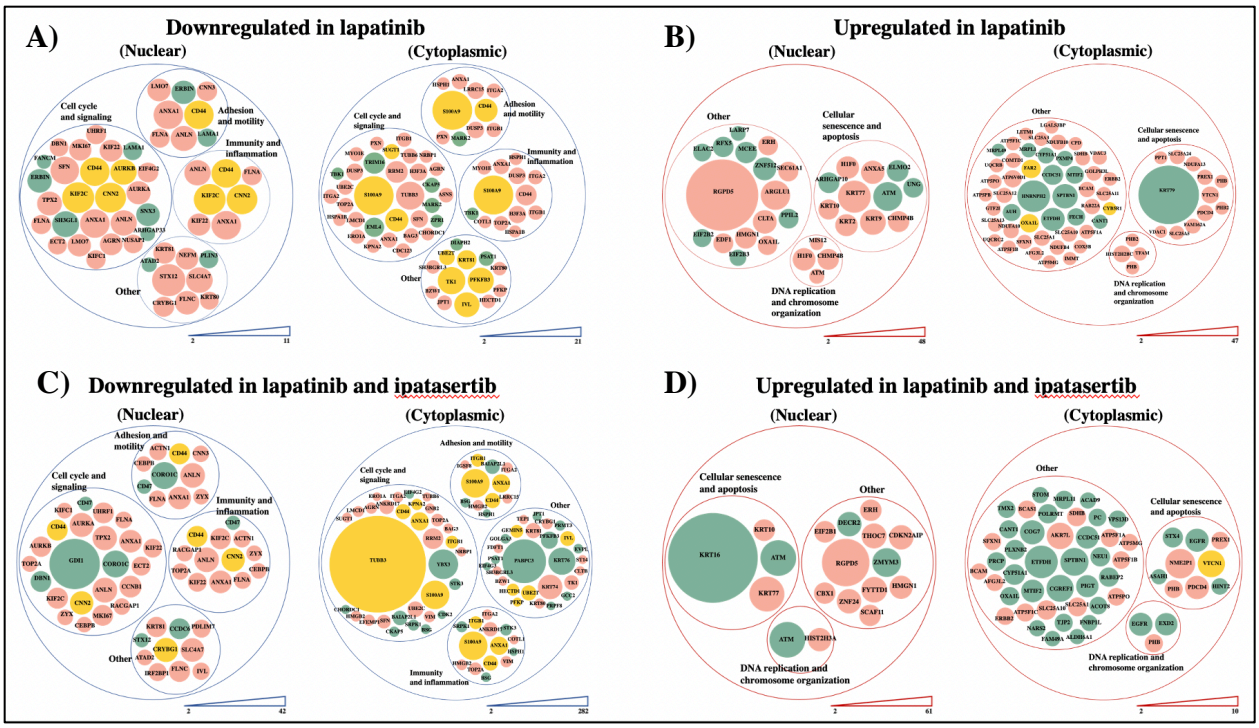


Figure 6-6: Bubble charts representing biological processes descriptive of cancer hallmarks, as represented by the differentially expressed proteins with adjusted p-value ≤0.1. The detected proteins were found to be downregulated **(A/C)**, or upregulated **(B/D)** in the drug vs EGF control comparisons. Notes: The color represents the quantitation method by which a protein was found to be differentially expressed: area-based (light pink), PSM-based (green), both (yellow). Each protein is represented by a node, with the size being indicative of the fold-change in the expression of that protein. If the protein was detected by both quantitation methods, the highest fold-change is represented. The range of protein fold-changes for each comparison is shown by the bar at the bottom of each diagram.

To further assess the importance of the up-and down-regulated proteins that emerged from the drug treatments described above, the differentially expressed proteins were screened for their cancer drug target potential, based on information extracted from the DrugBank database of approved and investigational targets (**Figure 6-7**).

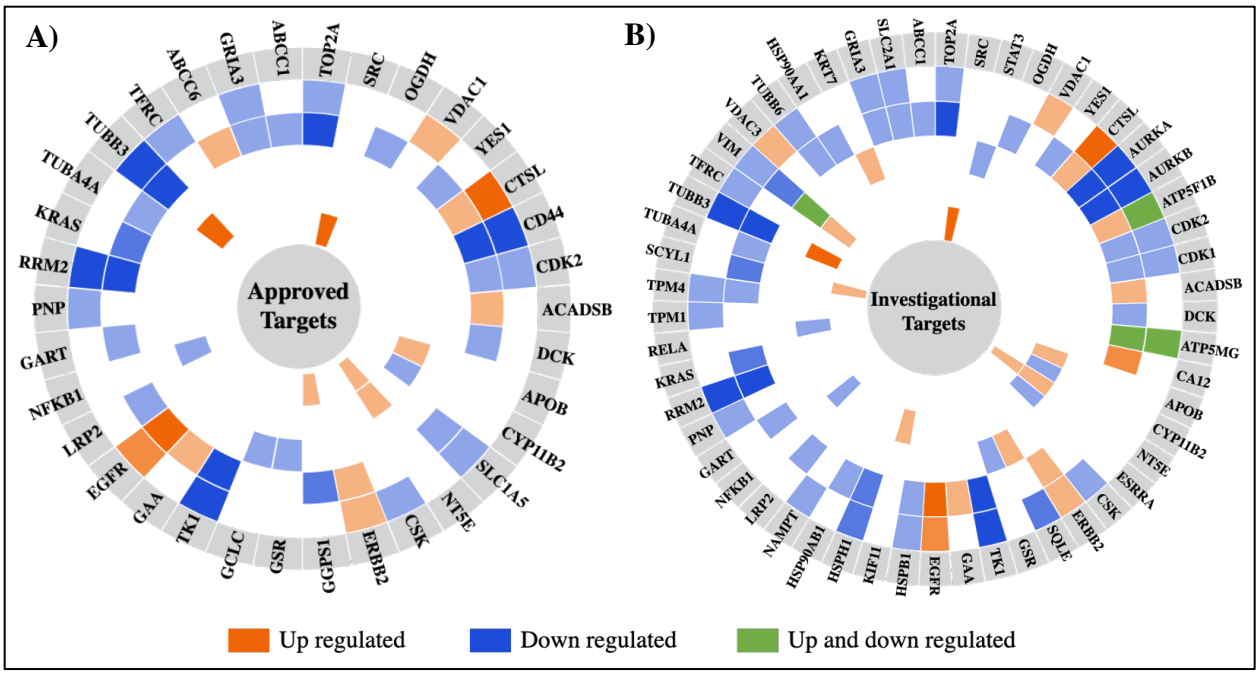


Figure 6-7: Differentially expressed proteins in the nuclear and cytoplasmic cell fractions aligned with approved **(A)** and investigational **(B)** drug targets from the DrugBank database. From the outer to the inner circle the comparisons are: lapatinib vs EGF, lapatinib/ipatasertib vs EGF, lapatinib/ipatasertib vs lapatinib, lapatinib/ATP vs lapatinib, lapatinib/ipatasertib/ATP vs lapatinib/ipatasertib. The different colors represent proteins that are up-regulated (orange) or down-regulated (blue) in either the nuclear or cytoplasmic cell fractions, or up- in one fraction and down-regulated in the other (green). The color intensity represents the fold change ranging from 2-3 (lightest), >3-4 (darker), and >4 (darkest). The green colored differentially expressed proteins are excluded due to the different fold changes in the nuclear and cytoplasmic fractions.

Our analysis revealed that the differentially expressed proteins represent not only currently approved drug targets but mainly investigational targets, still undergoing research about the potential impact on cancer development upon treatment with a drug. The investigational targets included cell cycle-relevant proteins (AURKA, AURKB, CDK1, CDK2, TOP2A) – emerging from the drug to EGF treatment comparisons, transcription factors such as STAT3 – from combination to single drug treatment comparisons, or the transcriptional regulator heterodimeric RELA-NFKB1 complex - from the added effect of ATP to the drug treatments.

6.3.4. Phosphorylated peptides enrichment

The MS2 analysis of the phosphopeptide-enriched cell extracts yielded an additional subset of proteins (**Table 6-1**) representative of early stages of cellular signaling processes (15 min- 30 min drug treatments). This protein subset was evaluated based on the detected phosphorylated

peptides and amino acid residues. The enrichment performed at the peptide level resulted in 80% or greater percentage of detected phosphopeptides (**Figure 6-8 A**), the majority of which were phosphorylated only on one single residue, mostly Serine (S) (**Figure 6-8 B, C**). Tyrosine (Y) modifications were the most difficult to detect, as evidenced by less than 1% detection among the phosphosites, result that otherwise is consistent with the reported literature. This is likely due to the extremely short half-lives of these modified residues (few seconds), yet the phosphorylation of these sites is essential to signal transduction and oncogenic activity.[24]

Table 6-1: Tandem MS analysis of phosphoprotein cell fractions and average number of protein IDs calculated from three biological replicates.

	EGF		**Lapatinb**		**Lapatinib & Ipatasertib**		**Lapatinib ATP**		**Lapatinib & Ipatasertib ATP**	
Timepoints	*15 min*	*30 min*	*15 min*	*30 min*	*15 min*	*30 min*	*15 min*	*30 min*	*15 min*	*30 min*
# Proteins	2603	2755	2534	2630	2610	2536	2598	2588	2617	2463
# Phosphorylated proteins	2017	2153	2021	2053	2056	2033	2074	2048	2064	1954
# Phophorylated proteins (2 peptide)	998	1075	979	1000	1009	992	1008	995	1008	951

Specific site reproducibility among the replicates of each drug treatment was greater than 65% for sites that were detected by at least two spectra per replicate (**Figure 6-8 D**). An in-house phosphosite database was built based on several data repositories of detected phosphorylated sites, as shown in **Figure 6-8 E**. The database comprised 247,967 modified S, T, and Y sites, but only 238,179 phosphorylated sites that corresponded to reviewed *Homo sapiens* proteins were considered (isoforms and unreviewed entries being excluded). This comprehensive database provided the necessary framework for analyzing the experimentally detected phosphosites from all drug treatments and replicates, resulting in the identification of 9718 unique phosphosites that could be matched to the entries in the database.

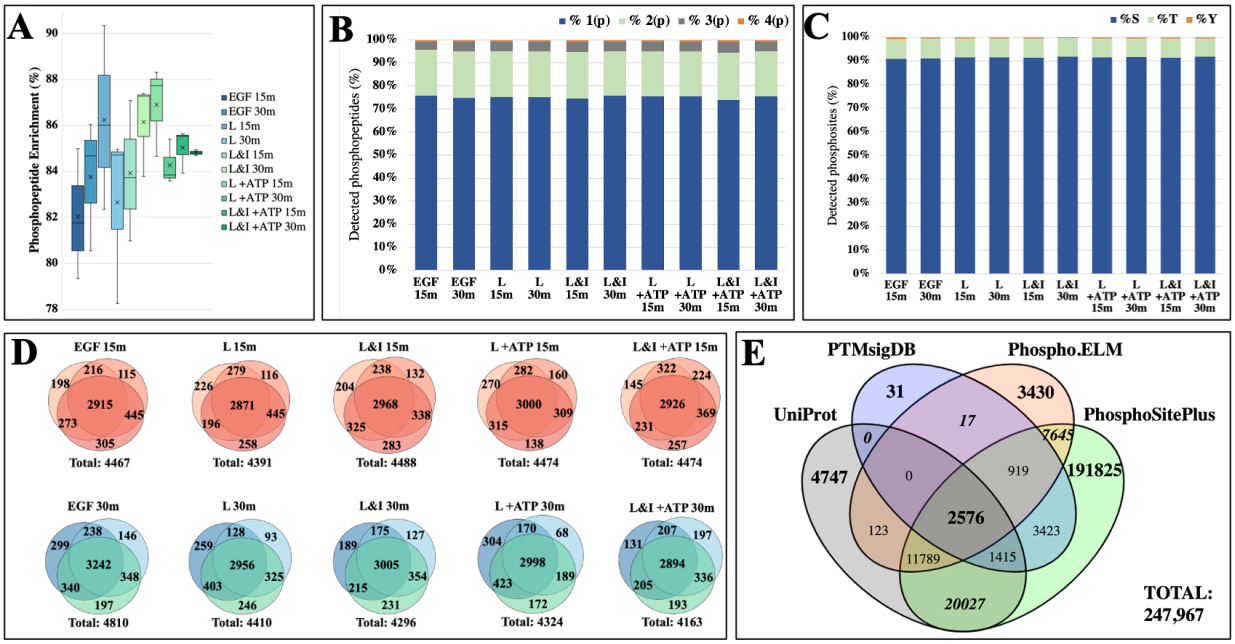


Figure 6-8: Overview of phosphopeptide MS analysis. **(A)** Detected phosphopeptides shown as a percentage of the total number of detected peptides for all treatments and timepoints. **(B)** Detected peptides with one, two, three, and four phosphorylation site modification(s), shown as a percentage out of the total number of detected phosphopeptides for all treatments and timepoints. **(C)** Detected phosphorylation sites on serine (S), threonine (T), and tyrosine (Y), shown as a percentage out of the total number of detected phosphosites. **(D)** Reproducibility of detecting each phosphosite with at least two spectra (PSMs) in each biological replicate, for each drug treatment and timepoint. **(E)** Venn diagram representing the contribution of the public databases of experimentally detected phosphosites to building the in-house phosphosite database.

Each treatment and replicate yielded an average of ~4,200 detected phosphosites that were present, and an additional ~500 phosphosites that were not present in the in-house built database, providing a potential pool of novel modified sites with relevance to specific signaling processes. A careful investigation of each detected phosphosite is required before claiming these modified residues as being novel phosphorylated sites. The challenge is two-fold because first, these sites are detected by one experiment and one method only, i.e., high-throughput MS (yet such data constitute the basis of more than 50% of the phosphosites found in major repositories such as PhosphoSitePlus),[22] and, second, their detection does not necessarily correlate to a cellular function unless further studies are conducted. In fact, only a fraction of all known phosphosites is paired with a kinase and has an assigned function.[20] This was also evident from the PhosphoSitePlus database, where only 8,812 out of >200,000 reported sites were associated with a biological function or process.[22] The combined treatments and replicates of this study resulted

in 744 unique sites that could be matched to a molecular function or process based on information provided by PhosphoSitePlus. Some of these biological processes are represented in **Figure 6-9** together with the detected sites. Some entries are shown repeatedly due to the same sites being part of multiple processes. For example, T185 phosphorylation of ERK2 is related to induced carcinogenesis, induced cell motility, induced cell growth, and induced and inhibited (both) transcription.

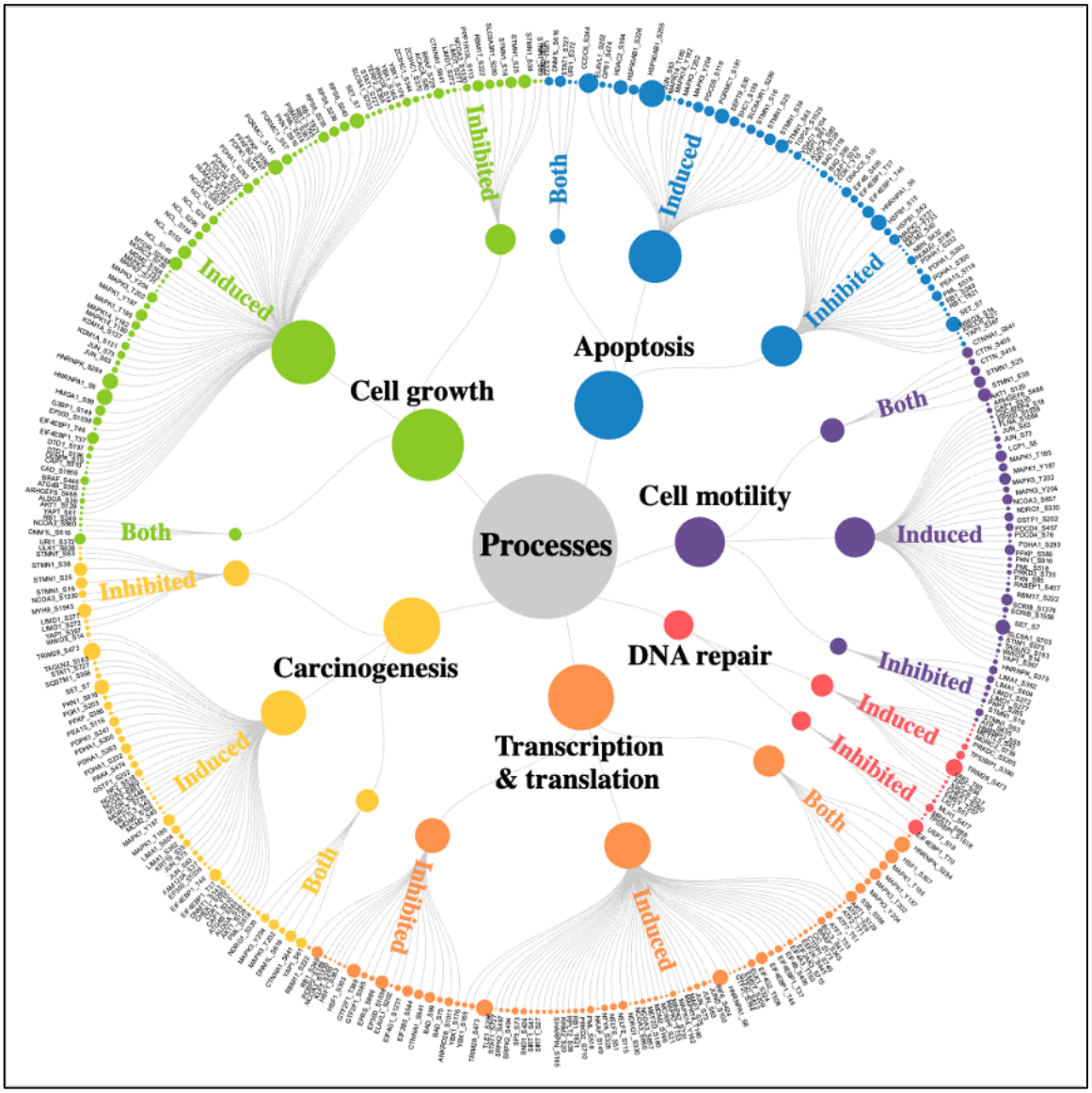


Figure 6-9: Detected phosphosites by all treatments and their corresponding proteins at the 15 min timepoint, categorized by biological processes found in PhosphoSitePlus. There were no major differences in the number of detected phosphosties between the 15 min and 30 min timepoints.

The functional roles of specific phosphorylation sites mostly depend on context, i.e., the presence of other interacting proteins, other phosphosites, or signaling events. **Figure 6-9** represents ~348 phosphosite entries (not unique) detected during the 15 min treatments. Few of these sites were specific to one drug treatment only, limiting thus the information that could be extracted from the existing differences because abundance levels at specific timepoints could not be inferred accurately from PSM measurements alone. Nevertheless, the detection of these sites enabled the identification of major signaling pathways affected by the drug treatments.

6.3.5. Relevant phosphorylated sites to cellular events

Among the detected phosphopeptides, the phosphorylation sites of highest interest were those known to affect the proliferation pathways of MAPK and PI3K. **Table 6-2** shows the detected phosphosites with relevant roles in signaling, as described by previous studies, and their respective spectral counts corresponding to the performed treatments and time points. Due to the interconnectivity of these proteins and known kinase-substrate relationships, the detection of these phosphosites can inform us about the activated pathways and signaling dynamics in the cell.

Table 6-2: Detected phosphoproteins and specific phosphosites, with their associated function and spectral counts of each phosphosite carrying peptide.

Gene	Site	Function	EGF 15m	EGF 30m	L 15m	L 30m	L&I 15m	L&I 30m	L ATP 15m	L ATP 30m	L&I ATP 15m	L&I ATP 30m	Ref
EGFR	Y1110	initiates signal propagation	6										25
EGFR	Y1172	initiates signal propagation	4	3									26
EGFR	Y1197	initiates signal propagation	8	9									25
ERBB2	Y1248	initiates signal propagation		3									27
PLCG1	Y1253	drives carcinogenesis, prognostic factor	6	3									26,28
GAB1	Y627	bi-phosphorylated site required for ERK2 (MAPK1) activation	9	9					1				29
GAB1	Y659	bi-phosphorylated site required for ERK2 (MAPK1) activation	8	5			3						29
MAPK1	T185	key signaling molecule with hundreds of downstream substrates	43	44	27	26	30	26	27	25	31	26	26
MAPK1	Y187	key signaling molecule with hundreds of downstream substrates	71	74	51	48	56	47	56	48	55	52	26
MAPK3	T202	key signaling molecule with hundreds of downstream substrates	27	28	12	12	13	10	19	9	14	9	26
MAPK3	Y204	key signaling molecule with hundreds of downstream substrates	57	59	38	33	38	31	47	32	44	34	26
RPS6KA1	S363	phosphorylated by ERK1/2 to promote cell proliferation and growth	2	1		1	1		2		4		26
RPS6KA1	S380	phosphorylated by ERK1/2 to promote cell proliferation and growth	32	23	10	10	11	7	13	9	15	8	26
RPS6KB2	S417	first step for its activation, involved in transcription regulation	2	3					2		1		30
RPS6KB2	S423	first step for its activation, involved in transcription regulation	2	3					2		1		30
RPS6KB2	S416	first step for its activation, involved in transcription regulation	2	3					2		1		30
MTOR	S2448	mTOR activation, increased cyclin D1 activity leading to G1/S progression	5	1	5		2	1	1	1	3	3	26
TSC2	S939	tumor suppressor, deactivated by phosphorylation from AKT		4									26
RAF1	S259	inhibitory site, keeps RAF-1 in an inactive form	1		8	6	4	7	6	8	4	2	26
RAF1	S621	inhibitory site, keeps RAF-1 in an inactive form	21	13	15	19	21	21	22	19	18	21	26
BCL2L11	S77	phosphorylated by ERK1/2 to reduce its pro-apoptotic activity	12	8	2	1	6		1		3		31
BAD	S75	anti-apoptotic, but versatile role depending on kinases and pathways	33	34	30	27	30	29	34	32	35	31	32
BAD	S99	anti-apoptotic, but versatile role depending on kinases and pathways	20	23	16	16	16	8	13	9	13	12	32
BAD	S118	anti-apoptotic, but versatile role depending on kinases and pathways	17	22	18	17	22	23	29	24	27	20	32
EIF4EBP1	S65	phosphorylated by multiple kinases, stimulates protein synthesis and growth		2					1		1		33
EIF4EBP1	T37	phosphorylated by multiple kinases, stimulates protein synthesis and growth	9	12	12		30	4	7		16	10	33
EIF4EBP1	T46	phosphorylated by multiple kinases, stimulates protein synthesis and growth	58	15	5	40	115	59	98	39	25	32	33
EIF4EBP1	T70	phosphorylated by multiple kinases, stimulates protein synthesis and growth	8	2	7	6	2	5	7	9	3	2	33
MDM2	S166	phosphorylated by AKT, downregulates p53	7	4	7	1	8	2	9	4	9	3	26
FOS	S374	member of the AP-1 complex transcription factor driving G1/S progression		9									26
FOSL2	S320	member of the AP-1 complex transcription factor driving G1/S progression	4		2		1						25
JUN	S63	member of the AP-1 complex transcription factor driving G1/S progression	16	25	14	16	17	13	17	21	16	17	26
JUN	S73	member of the AP-1 complex transcription factor driving G1/S progression	12	18	12	12	11	13	12	16	14	14	26
JUND	S259	member of the AP-1 complex transcription factor driving G1/S progression	28	27	22	18	21	15	27	19	30	12	25
ATF2	S112	member of the AP-1 complex transcription factor driving G1/S progression	8	5	7	2	9	2	4	6	4		25
ATF2	T69	member of the AP-1 complex transcription factor driving G1/S progression	8	10	7	4	7	1	6	8	5	4	25
ATF2	T71	member of the AP-1 complex transcription factor driving G1/S progression	8	10	7	4	7	2	6	8	6	4	25
MYCN	S62	induces cyclin D1 transcription leading to G1/S progression		5			2			1		1	26
MYCN	T58	induces cyclin D1 transcription leading to G1/S progression		4			2						26
RB1	S807	phosphorylation leads to G1/S transition	18	12	21	12	21	18	13	3	14	3	26
RB1	S811	phosphorylation leads to G1/S transition	14	8	14	9	18	11	12		12	1	26
STAT1	S727	oncogenic signaling	9	15	6	10	7	2	9	12	9	8	27
STAT5B	Y699	oncogenic signaling	5	1									27

Figure 6-10 represents a SIGNOR3.0-generated signaling network created from the detected phosphoproteins carrying functional phosphosites with either stimulatory or inhibitory role. For a limited but relevant subset of proteins, the network highlights the signaling relationship between kinases and target substrates, the targets that can be phosphorylated by several kinases, and the dual effect of many kinases (such as the inhibitory or stimulatory effect of MAPK1 and MAPK3). These relationships become only more complex with the addition of phosphatases that turn off phosphorylation processes.

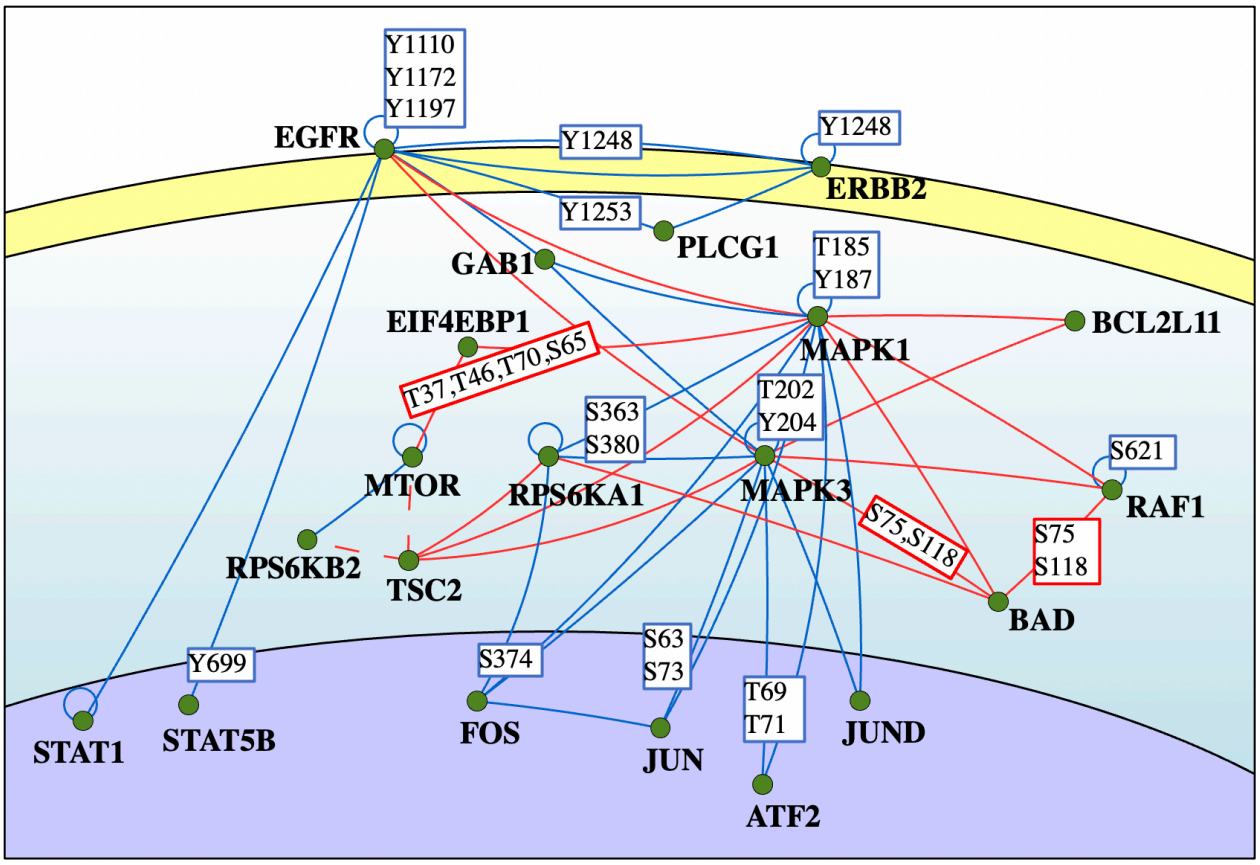


Figure 6-10: Signaling network of phosphorylated proteins created by SIGNOR3.0 with data from Table 6-2. Solid and dashed lines represent directly or indirectly connected proteins, respectively, with either stimulatory (blue) or inhibitory (red) activity.

Out of the 9,718 detected unique phosphosites from all treatments and time points, 6,661 were detected at least by two PSMs, of which 882 were assigned matching kinases, with an associated confidence score, according to the SIGNOR3.0 database. Kinase-substrate matches with high-confidence scores (>0.7) were inputted in Cytoscape to generate a PPI network (**Figure 6-11**). This network conveys a broad signaling landscape, beyond what is known and characterized in the literature (**Figure 6-10**). Despite the lack of knowledge about the function of specific phosphosites, their detection is crucial to motivate further research about their signaling role in cellular processes.

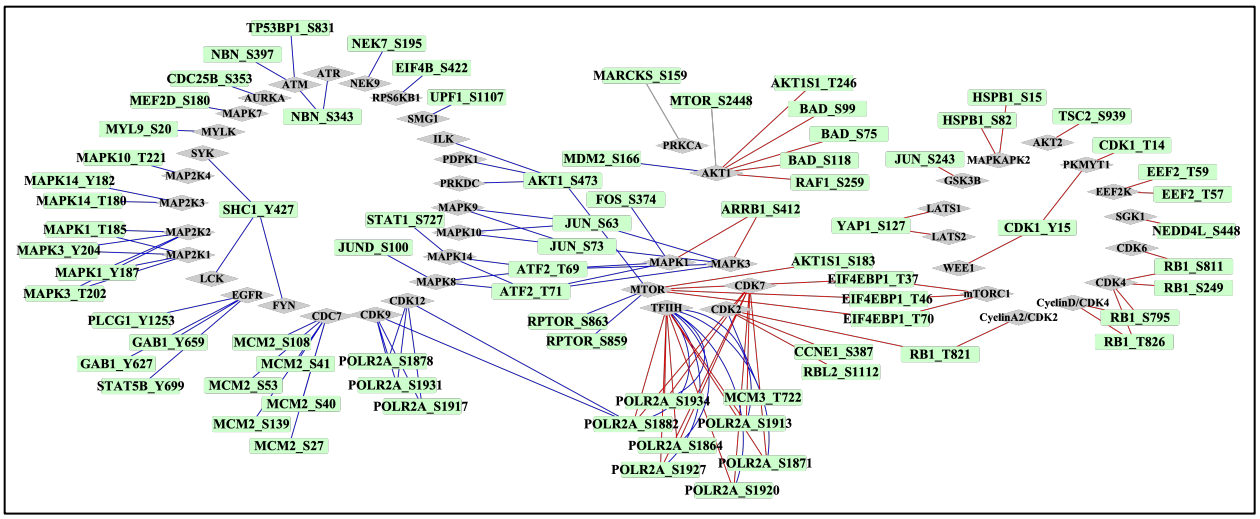


Figure 6-11: Detected phosphorylated protein substrates with their specific sites (green) matched to their respective kinases (gray) by SIGNOR3.0. The inter-connecting lines represent stimulatory/activation (blue) or inhibitory (red) signaling.

Biological interpretation of the phosphoproteins. To provide a better understanding of the biological processes that could have been affected by each drug treatment, kinase enrichment analysis with KEA3 tools was performed for the phosphoproteins that displayed a difference of 5 or more in phosphopeptide PSMs between two different drug treatments. The treatments that were compared included: lapatinib vs EGF, lapatinib/ipatasertib vs EGF, lapatinib/ATP vs lapatinib, and lapatinib/ipatasertib&ATP vs lapatinib/ipatasertib, as indicated in **Figures 6-12** and **6-13**. Although most of the enriched kinases that emerged from this analysis were the same for all treatments, the substrates that were phosphophorylated by these kinases were different. As a result, the enriched biological processes that were retrieved from the PPI networks constructed from the substrates and their respective kinases showed that different signaling pathways were affected by the drug treatments (**Figures 6-12, 6-13**). The 15 min signaling events represented a more diverse landscape of the cellular dynamics when compared to the 30 min signaling events that were mostly characterized by mRNA processing biological categories and pathways. As a result, the 15 min results will be discussed in more detail in the following sections.

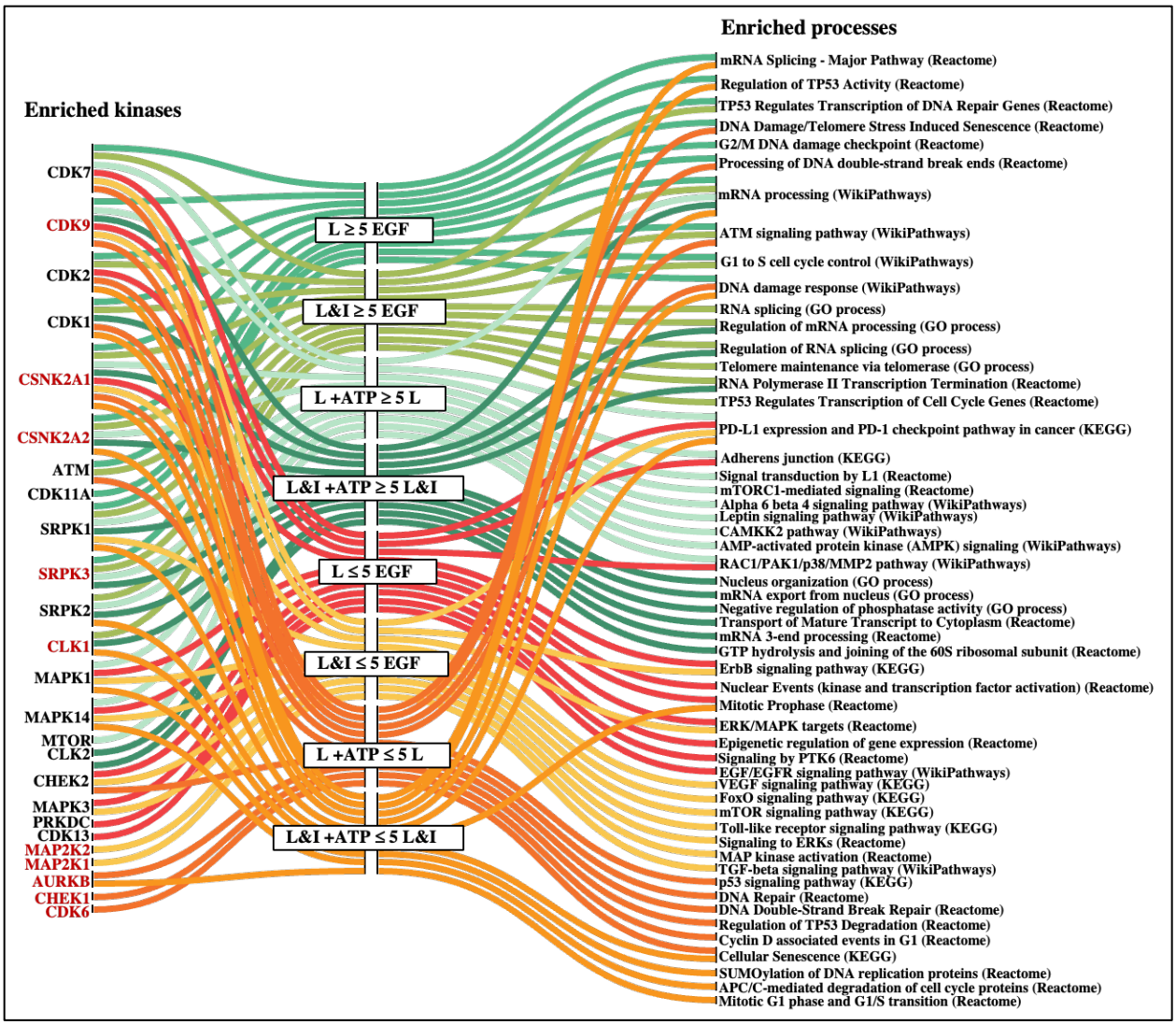


Figure 6-12: KEA3 analysis of phosphoproteins that displayed a difference of 5 or more in phosphopeptide PSMs between two different drug treatments [15 min exposure to the drug(s)]. The detected phosphorylated proteins were found to be enriched in the kinases shown on the left (black font). Kinases that were not detected in the dataset, but suggested by KEA3 as kinases of the detected phosphorylayted proteins, are shown in red font. The selected GO/KEGG/REACTOME/WIKI processes that were differentially affected in each comparison were at least 10-fold enriched, with an FDR ≤0.05, and were represented by at least 4 genes/category.

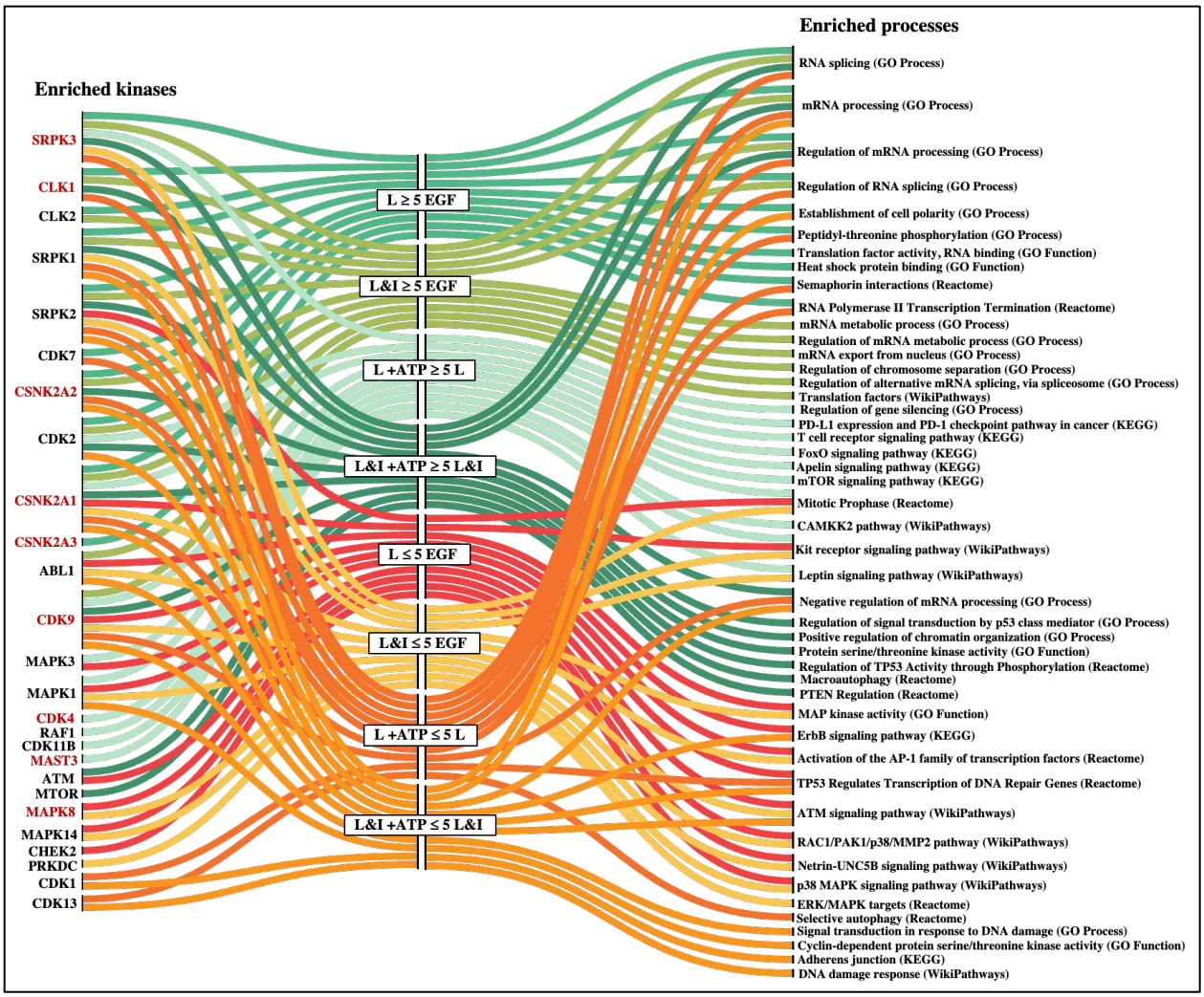


Figure 6-13: KEA3 analysis of phosphoproteins that displayed a difference of 5 or more in phosphopeptide PSMs between two different drug treatments [30 min exposure to the drug(s)]. The detected phosphorylated proteins were found to be enriched in the kinases shown on the left (black font). Kinases that were not detected in the dataset, but suggested by KEA3 as kinases of the detected phosphorylayted proteins, are shown in red font. The selected GO/KEGG/REACTOME/WIKI processes that were differentially affected in each comparison were at least 10-fold enriched, with an FDR ≤0.05, and were represented by at least 4 genes/category.

6.4 Discussion

Proteomics analysis of early and late cellular events affected by the treatments of SKBR3/HER2+ breast cancer cells with ERBB2/EGFR receptor kinase inhibitor lapatinib and pan-AKT inhibitor ipatasertib revealed many biological processes (**Figures 6-4, 6-5, 6-12, 6-13**) relevant to cancer progression and hallmarks, such as growth, apoptosis, immune response, and metabolism. Lapatinib treatment led to the downregulation of cell cycle, proliferative signaling,

and growth, while the addition of ipatasertib contributed to the downregulation of AKT downstream targets and pathways, such as mTOR and FoxO, and negative regulation of apoptotic processes. The addition of the ATP signaling molecule to the drug-treated cells did not result in the detection of entire up/down-regulated biological processes or pathways, but revealed instead a few potential drug targets of relevance to the development of more effective cancer therapies (**Figure 6-7**). A combination of label-free quantitation methods (spectral count- and peak area-based) relying on three criteria for classifying a protein as differentially expressed, i.e., abundance change greater than 2-fold, p-value ≤0.05-0.1, and requirement of at least 3 unique detected peptides per protein, was used for assessing the changes that were induced by the treatment of cells with drugs. Several of these proteins were validated by independent methods, such as PRM and western blotting (**Appendices H, I**).

6.4.1. Lapatinib/EGF treatment vs EGF stimulation of cells

Down regulated proteins

Cell cycle and signaling. In the nucleus, lapatinib treatment led to the downregulation of the cell proliferation and cancer prognosis maker, Ki67, and of cell cycle related processes, such as G1/S and G2/M transition, mitotic nuclear division, chromosome segregation, spindle assembly and organization, and cytokinesis, as represented by a highly interconnected network of nuclear proteins (AURKA, AURKB, CDK1, CCNB1, PKD1, TOP2A, SFN, TPX2, KIF2C, KIFC1, KIF22, KIF23, NUSAP1, PRC1, ECT2, RACGAP1, FANCM, ANLN, UHRF1) (**Figure 6-4**). Proteins that displayed a high fold change in abundance included AURKA, AURKB, and NUSAP1. These proteins are overexpressed in multiple cancers and are linked to poor patient prognosis, tumor growth, invasion, and metastasis through signaling pathways such as MAPK, PI3K/AKT, and AMPK.[34,35]

Cytoplasmic proteins with regulatory roles in cell cycle progression were also downregulated by lapatinib, including the proliferation marker TK1, PFKFB3 (regulator of CDK1 detected in the nuclear fraction), SFN (dually located), ZPR1, UBE2C, microtubule-associated protein EML4, microtubule affinity regulating kinase MARK2, and SUGT1, required for G1/S, G2/M transition. An interesting finding was the downregulation of involucrin (IVL), an early differentiation marker of keratinocytes regulated by EGFR, which in contrast to our study was reported to be

upregulated by lapatinib.[36] Downregulation by lapatinib of the proteins with high fold-change like TK1 and PFKFB3 agreed with previous studies.[37,38]

Immune response and inflammation. The proteins down-regulated by lapatinib with the highest fold-change abundance were S100A9 in the cytoplasm and KIF2C in the nuclear fraction, both involved in cell cycle progression and inflammatory and immune response processes. Even though these processes were not enriched as a whole in the STRING analysis (**Figure 6-5**), several additional proteins with roles in immunity and inflammation were detected (S100A9, TBK1, ANXA1, COTL1, DUSP3, KIF2C), indicating an effect of this drug on the immune response of the SKBR3 cells. Overexpression of S100A9 has been implicated in the development and progression of many cancer types, including breast cancer, and correlated with high expression of Ki67 and HER2.[39] The involvement of S100A9 in multiple cellular processes and signaling cascades like invasion (through H-Ras pathway), apoptosis (through p53 dependent or independent pathways), or inflammation (through RAGE pathway) has made S100A9 a protein of interest to many studies, which have led to conflicting reports about its role in tumorigenesis.[39,40] MAPK and JAK1/2 inhibitors, and STAT3 silencing have led to S100A9 expression inhibition,[39,41] but to our knowledge no direct link was yet established between lapatinib and downregulation of S100A9 expression.

KIF2C is a multifunctional protein with roles in cell cycle, chromosome segregation, and genomic stability, and its overexpression in many cancers has led to proliferation, invasion, and metastasis via PI3K/AKT, mTORC1, and MAPK/ERK signaling.[42-44] In accordance with these findings, it was reported that KIF2C suppression inhibits mitosis.[42] Yet, an important and less studied aspect is the involvement of KIF2C in immune response, where the role of KIF2C in immune cell infiltration leading to tumor inhibition is either positively or negatively associated, depending on the cancer type.[43,44] Liu et al. reported that in breast cancer, increased expression of KIF2C associated with higher immune cell infiltration.[44] As the treatment of cells with lapatinib downregulated KIF2C expression, with likely cell cycle implications, it would be of interest to investigate the behavior of infiltrating immune cells under lapatinib treatment and consider the effects when combined with immunotherapy. Lastly, both S100A9 and KIF2C, as being an

integral part of the tumor microenvironment, have been reported as suitable markers for evaluating immunotherapy response.[44-46]

Adhesion and signaling. Lapatinib treatment led to many downregulated proteins representative of biological adhesion processes in both the nuclear and cytoplasmic fractions, such as CNN2, CNN3, SH3GL1, CD44, PLIN3, LAMA1, FLNA, LMO7, DBN1, ANXA1 in the nuclear fractions, and CD44, ITGB1, ITGA2, TUBB3, TUBB6, PXN, ANXA1, DIAPH2, MARK2, LRRC15, BAG3, BZW1 in the cytoplasm. Affected biological processes included cadherin binding, cell adhesion molecule binding, cell-ECM interactions, cell junction organization, and L1CAM interaction processes (**Figure 6-4**). Cell-adhesion molecules were described for their shuttling between transmembrane regions or cytosol to the nucleus,[47] therefore the localization of some of the above-mentioned proteins could vary between different cell compartments. The proteins with the highest fold-change such as CD44, CNN2, and SH3GL1 display multiple functional roles in the cell (**Figure 6-6**). CD44 is a multifunctional protein that promotes tumor progression, invasion, metastasis, and epithelial-mesenchymal transition[48] through various mechanisms including nuclear translocation.[49] Furthermore, CD44 has been demonstrated to play a role in the development of drug resistance in breast cancer cells treated with trastuzumab, lapatinib, and tamoxifen,[50-52] therefore targeting CD44 in conjunction with other potential drug targets remains of high interest. The homologous genes CNN2 and CNN3 affect cell proliferation, adhesion, migration, and metastasis.[53] They are overexpressed in many cancer types, but reports about their function in the cell are confounding depending on the cancer type,[54-59] although most of the above studies state that CNN2/3 overexpression promotes tumorigenesis and invasion probably through MAPK/ERK signaling.[57] In our studies, these proteins were downregulated by the lapatinib treatment, calling attention to the potential role of this drug in mediating adhesion processes. SH3GL1 (SH3 Domain Containing GRB2 Like 1, Endophilin A2) protein levels were linked to tumor progression, metastasis, and worse patient outcomes in many cancer types.[60,61] Endophilin silencing in SKBR3 cells led to impaired HER2 receptor internalization and reduced downstream signaling in cells treated with Trastuzumab, and reduced cytotoxicity in HER2+ cells treated with TDM-1.[61] To our knowledge, the downregulation of these proteins by lapatinib in SKBR3 cells is firstly reported here, providing

novel opportunities for future research into the mechanisms of action of this drug and its effects on biological processes such as adhesion, signaling, and migration.

Up regulated proteins

Chromosome organization, DNA repair, and apoptosis. Lapatinib treatment also led to the upregulation of several nuclear proteins involved in DNA conformation change, DNA damage, chromosome organization, and regulation of chromosome organization like ATM, MIS12, H1F0, CHMP4B, ATRX, XPC, NUMA1, RSF1, FOXA1, and TRRAP, all shown to interact with each other (**Figure 6-5**). Among these, the proteins that displayed greater abundance fold-changes were the ATM kinase and the transcription factor ZNF512. The latter one is not well described in the literature, but based on the Kaplan-Meier analysis it is a prognostic marker in breast cancer, whose overexpression leads to favorable patient survival outcome.[62] Recently, it was identified as a binding partner of the histone deacetylase (HDAC) inhibitor panobinostat,[63] which acts on cell-cycle progression and apoptosis.[64] To our knowledge, this is the first report for the overexpression of this gene in HER2+ breast cancer cells treated with lapatinib. Its further investigation can help elucidate the mechanisms of lapatinib action in the cell and inform about possible off-target effects. Kinase ataxia-telangiectasia mutated (ATM) protein is an integral part of the DNA damage repair (DDR) response and regulates growth inhibition in concert with another upregulated protein, NUMA1, whose overexpression was associated with increased survival and metastatic-free survival in patient cohorts.[65-67] Here, we also report the upregulation of chromatin remodeler ATRX, a controversial protein regarding its role in telomere lengthening and tumor progression as being either positively or negatively correlated,[68-69] which invites further research into its role in cancer.

In the cytoplasm, HNRNPH2-an RNA binding protein, had one of the highest fold-changes in abundance. Even though a nuclear protein, it shuttles between the nucleus and cytoplasm to control pre-mRNA splicing,[70] which suggests an influence of lapatinib on RNA splicing in agreement with previous findings.[71] However, in a pan-cancer analysis, HNRNPH2 was found to be involved in a wider array of cellular functions including G2/M checkpoint, JAK/STAT3 signaling, and DNA repair[72] raising further interest in this protein and the need for future studies.

Unknown function. The most abundant up-regulated nuclear protein (47-fold enrichment) was the RANBP2-like and GRIP domain-containing protein 5 (RGPD5), a GTP-binding protein of the Ras family, that has various cellular roles through its association with the nuclear membrane and the many protein interactions that it forms, but with yet unknown function in cancer. It has been reported that this protein was found to be mutated, regulated by miRNA, and methylated in breast cancer,[73-75] but conclusive data on its role in tumorigenesis are lacking. Its upregulation by lapatinib will spur further interest in investigating its potential role in signaling and cancer. Overall, lapatinib has proven to be an efficient drug leading to cell cycle arrest and apoptosis and shown to enhance the effects of other combinational therapies.[76] This work corroborated previous findings, led to the identification of novel proteins affected by the exposure of cells to lapatinib, and provided a more comprehensive view of the biological landscape that is altered by the treatment of cells with this drug.

6.4.2. Lapatinib/ipatasertib/EGF combination treatment vs EGF stimulation of cells

Statistical comparisons of spectral counts and areas between the lapatinib/ipatasertib and lapatinib-alone treatments resulted in the detection of only a few differentially expressed proteins. The addition of ipatasertib to lapatinib downregulated the nuclear abundance of several transcription factors including ZBTB21, TAF2/5, GTF3C2, MAFK, and STAT3, involved in DNA-binding, initiation of RNA polymerase II/III-dependent transcription, and transcription regulation.[77] Decreased abundance of STAT3 is of high interest because STAT3 and PI3K/AKT signaling are two distinct pathways, which lately have been implicated in crosstalk.[78] The changes in the STAT3 status by inhibitors of PI3K/AKT, either measured by phosphorylation activity or by total protein abundance, are confounding, as revealed by studies that indicate both downregulation or upregulation.[78,79] Nevertheless, the impact of adding ipatasertib to the drug treatment of cells was more evident from comparing directly the lapatinib/ipatasertib/EGF treatment to EGF stimulation alone (**Figures 6-4, 6-5, 6-6**). The addition of ipatasertib resulted in several enriched biological processes such as negative regulation of apoptosis (downregulated), localization, and exocytosis (upregulated), as shown in **Figure 6-5**. Furthermore, the differentially expressed proteins by the drug combination treatment displayed a higher-fold change difference than that induced by lapatinib alone (except for the up-regulated cytoplasmic proteins) (**Figure 6-6**).

Down regulated proteins

The downregulated proteins by the combination of drug treatments were in accordance, and part of the same processes, with what was found to be downregulated by lapatinib. However, the cell cycle and signaling processes displayed a high fold-change of two specific proteins, GDI1 and TUBB3. GDI1 is a prognostic marker in breast cancer whose overexpression leads to unfavorable results in patient survival.[80] As a GDP dissociation inhibitor, GDI1 modulates Rho GTPases activation and localization, however, the studies regarding its expression levels in different cancer types are confounding.[81-83] Its overexpression in colorectal and hepatocellular cancer cells was found to lead to telomere dysfunction, and cell proliferation and migration, respectively, through the PI3K/Akt pathway.[83,84] To our knowledge, the ipatasertib effect in GDI1 expression has not been reported before. TUBB3, a microtubule protein with important roles in chromosome segregation during mitosis and meiosis, has been found overexpressed in many cancer types, including breast, in which its overexpression was linked to decreased sensitivity to taxane drugs, advanced tumor stage, and HER2 amplification in patient cohorts.[85] However, in response to PI3K/Akt pathway inhibitors, TUBB3 expression was reported to be either downregulated or upregulated in various cancer types.[86-89] There was no report on the expression of TUBB3 upon ipatasertib treatment. Another downregulated protein by the use of the drug combination, with a high fold-change, was PABPC3, an RNA-binding protein with role in mRNA translation initiation. This protein was also found to be highly mutated in breast cancer,[90] and associated with tumor progression in colorectal cancer.[91] Due to the existence of only few studies characterizing its role in cancer, it would be of interest to further investigate the expression of PABPC3 in the presence of kinase inhibitors and its relation to the PI3K/Akt signaling pathway. For instance, another family member of the conserved gene family PABPC, PABPC1, was found to contribute to trastuzumab resistance in SKBR3 cells.[92]

Up regulated proteins

Some proteins involved in DNA damage and apoptosis were found to be more abundantly expressed in the lapatinib/ipatasertib treatment based on their fold-change, such as ATM and PDCD4, which also showed the same trend in the PRM experiments. PDCD4 is a protein translation inhibitor and tumor suppressor frequently downregulated in many cancer types,[93]

including breast cancer, where HER2 activation downregulates PDCD4 through the MAPK and PI3K/AKT pathways.[94] Overexpression of PDCD4 was also reported to promote apoptosis in other breast cancer cell lines in the presence of the HER2 inhibitor trastuzumab.[95] The mechanisms of PDCD4 inhibition of proliferation remain under research, but one possible mode of action is the binding of PDCD4 with the translation initiation factor eIF4A, and blocking of its interaction with EIF4G which is required for cap-dependent protein translation.[94,96] Interestingly, EIF4G2 and EIF4G3 were found to be downregulated by the drug treatments in the cytoplasm. EIF4G2 was also found to be downregulated in the nuclear fraction, most likely due to translocation as reported before.[97]

6.4.3. Early events represented by the phosphorylated signaling landscape

The biological processes that were affected by either the lapatinib treatment alone or by the combination with ipatasertib are consistent with some of the processes that emerged from the differential expression analysis of the cell lysates, indicating the same trends in cell behavior, spanning from early events in minutes to prolonged effects in days. The phosphoproteins that carried more phosphosites in the lapatinib or lapatinib&ipatasertib treatments represented cellular processes related to DNA damage and repair, G1/S cell cycle control and TP53 regulation, ATM signaling, and RNA splicing and processing, all indicative of cell cycle arrest and apoptosis. On the contrary, when compared to EGF stimulation, the phosphoproteins that carried less phosphosites in the lapatinib or lapatinib/ipatasertib treatments represented processes that reflected PD-L1 expression, transcription factor activation, adherens junction, and signaling pathways such as ERBB, EGF, ERK, MAPK, VEGF, FoxO, mTOR, TGF-beta signaling. The last four pathways were found to be enriched only when the drug combination treatment was used and compared to EGF stimulation, providing evidence for the potential added benefit of ipatasertib. In fact, since ipatasertib is an AKT inhibitor, it is expected to affect AKT-related pathways and downstream targets such as FoxO and mTOR[27]. These results also indicated that the ipatasertib effects may be better captured by the analysis of cell phosphorylation profiles rather than differential expression of total proteins, as was the case with another Akt inhibitor, MK2206.[98]

6.4.4. Addition of ATP to the drug treatments

ATP addition to the drug treatments did not lead to enriched biological processes, however, changes in the expression of a few proteins were observable, mainly the upregulation of GDI1 and of several genes that were found to be overexpressed in-, or associated with cancer progression and poor patient prognosis, such as the cancer stem cell marker ALDH1A3, nucleoporin POM121C, and proto-oncogene Src.[99-102] This study could be indicative of the ATP role as a signaling molecule that promotes tumorigenesis, but the cellular mechanisms remain unknown and subject to future studies.

The proteins that displayed changes in the number or abundance of phosphosites were involved in cellular senescence, DNA damage and repair, p53 signaling and TP53 regulation, and mitotic prophase and G1/S transition, suggesting combined effects from the drugs in reducing cell cycle activities and most likely from the ATP in reducing apoptosis. It was interesting to note that ATP appeared to have a more pronounced effect in the presence of only lapatinib than in the presence of both lapatinib and ipatasertib drugs. The enhanced effect of ATP when added to the lapatinib treatment was evident from the increased phosphorylation of proteins involved mTORC1, CAMKK2, and AMPK signaling. Detected phosphoproteins such as ULK1, TSC2, and RPTOR interconnect these pathways affecting cell growth, autophagy, cellular energy, and mitochondrial homeostasis.[103] In the lapatinib/ipatasertib drug treatment, the addition of ATP resulted in an increase of phosphorylation of proteins that were implicated in processes such as mRNA processing, RNA splicing and their regulation, similar to the processes represented by the phosphoproteins involved in the lapatinib/ipatasertib treatment when compared to EGF. **Figure 6-12** illustrates the intersections of these processes.

Lastly, western blot experiments could not confirm whether ATP can transactivate the ERBB2 receptor through the action of the P2RY2 receptor or other cell membrane receptors. The schematic of the proposed transactivation mechanism and some of the representative experimental results are shown in **Appendix J**. In brief, the transactivation experiment assessed whether ATP, a ligand of the P2RY2 receptor, could induce ERBB2-initiated downstream signaling through the MAPK/ERK pathway, as previously shown for EGFR in other cancer cells.[12,13] The ERBB2/Tyr 1248 is the most well-studied phosphorylation site, which activates

downstream signaling.[104] Lapatinib was used to block the direct, endogenous activation of the ERBB2 receptor and the phosphorylation of this site. The phosphorylation of the site would have indicated ERBB2 activation by the addition of ATP to the culture medium and would have confirmed the transactivation process. Similarly, the phosphorylation of the well-known activator Thr 202/Tyr 204 sites of the ERK1 and ERK2 receptors[105] would have indicated either cross talk (no ERBB2 phosphorylation) or transactivation (ERBB2 phosphorylation). The detection of these phosphorylation sites in our experiments occurred, however, in a sporadic and non-reproducible manner, indicating the need for further studies for elucidating the molecular mechanisms of the ERBB2 transactivation process.

6.4.5. Potential drug targets

An important objective of this study was to reveal differentially expressed proteins that could guide the choice of potential therapeutical treatments against breast cancer. Many of the detected proteins (34 proteins) represent known approved drug targets in the DrugBank database, and yet many more (53 proteins) are investigational targets (**Figure 6-7**). We report here that the downregulated proteins in either the nuclear or cytoplasmic fractions (ABCC1, SLC2A1, GRIA3, TUBB3/6, VIM, TFRC, SCYL1, TPM1/4, RRM2, PNP, GART, LRP2, NAMPT, HSPH1, HSPB1, TK1, CSK, DCK, OGDH), or the upregulated ones (VDAC1/3, CTSL, ATP51B, ACADSB, GAA, CA12), constitute an emerging pool of potentially novel targets of interest to be further studied. Other differentially expressed proteins and putative drug targets, as also suggested by a few other studies, included SRPK1-a protein whose inhibition would simultaneously affect diverse and distinct processes (e.g., migration, metastasis, apoptosis, sensitivity to chemotherapy in breast cancer),[106] and the multifunctional TBK1 protein (involved in processes such as proliferation, autophagy, innate immunity, and mitochondrial metabolism)-whose targeting would suppress proliferation and activate T-cell mediated immune responses.[107] Additional potential targets included the upregulated SORL1-due to its role in promoting resistance to HER-2 targeted therapies in breast cancer,[108] and the tumor suppressor PTPRD-a tyrosine phosphatase implicated in cell growth inhibition.[109,110]

Recently, a particular focus was placed on the development of cancer immunotherapy approaches, and some of the differentially expressed proteins included previously suggested

targets such as the above mentioned TBK1-with a suggested role in immune evasion, the downregulated AGRN protein-due to its promotion of cell growth and invasion through the PI3K/AKT pathway, and immune cell infiltration,[111] and the upregulated VTCN1-due to its role in T-cell responses.[112]

Last, another emerging group of targets against cancer included the mitochondrial proteins involved in apoptosis. Mitochondrial ribosomal proteins (MRPs) have important roles in mitochondrial translation and protein synthesis necessary for oxidative phosphorylation, cellular respiration, and energy production. Their abnormal expression has been reported is several cancers, including breast cancer, and was linked to apoptosis and cell death initiation, tumorigenesis, and other processes that derive from the complexity and variety of the differentially expressed MRPs. Taken altogether, mitochondrial dysfunction and apoptosis constitute hallmarks of cancer that are of importance to targeted therapies.[113-115] Several MRPs that were upregulated upon drug treatment are known to promote apoptosis and could provide for potential therapeutic targets. These included the death-associated protein 3 (DAP3/MRPS29)-an Akt substrate that binds to apoptotic factors and TRAIL receptors,[113] MRPS30/mL65-which induces apoptosis independent of the death receptor-induced extrinsic pathway,[114] and MRPL11-which indirectly activates p53 and leads to cell cycle arrest.[116] The vast landscape of the differentially expressed MRPs invites further investigation of their intertwined functions in metabolism, energy, and cancer progression.

6.5 Conclusion

The detection of a conglomerate of up- and down-regulated proteins in both the nuclear and the cytoplasmic fractions unveiled a complex picture of the effects of the drugs on the cellular proteome. Lapatinib treatment had a significant impact on the cell cycle and cell proliferation pathways in SKBR3 cells. Treatment with lapatinib resulted in the downregulation of several nuclear and cytoplasmic proteins involved in the regulation of cell cycle progression, including Ki67, CDK1, CCNB1, TOP2A, and KIF2C, among others. These proteins play critical roles in mitotic nuclear division, spindle assembly and organization, cytokinesis, and DNA replication. The downregulation of these proteins was consistent with cell cycle arrest and apoptosis, which was further supported by the upregulation of proteins involved in processes related to DNA

damage and repair, and chromosome organization, such as ATM, PDCD4, and ZNF512. Adding ipatasertib resulted in several enriched biological processes, including negative regulation of apoptosis and localization, and exocytosis-as represented by changes in the expression of nuclear and cytoplasmic proteins, and of signaling pathways FoxO, mTOR, and TGFB-as represented by changes in the phosphorylation profile of proteins. The addition of ATP to the drug treatments was indicative of ATP's potential role as a signaling molecule promoting tumorigenesis through the upregulation of some proteins involved in cell cycle and signaling, which in turn were found to be downregulated by the drugs. Lastly, the qualitative assessment of the phosphoproteomics data provided phosphosites with relevance to signal transduction and cancer progression. Further quantitative assessment would be needed to elucidated their specific role in cancer cell proliferation. The therapeutical targeting potential was highlighted by differentially expressed proteins representing a vast pool of investigational drug targets, potential immunotherapeutic targets (TBK1, AGRN, VTCN1), and apoptosis-inducing MRPs (DAP3, MRPS30, MRPL11). Overall, these findings provide valuable insights into the molecular effects underlying the therapeutic effects of lapatinib and ipatasertib in breast cancer treatment and expand the opportunities for novel and combinatorial targeted therapies.

6.6 References

1. Martínez-Sáez, O., & Prat, A. (2021). Current and Future Management of HER2-Positive Metastatic Breast Cancer. *JCO oncology practice, 17*(10), 594–604. DOI: 10.1200/OP.21.00172

2. Nahta, R. (2019). Novel therapies to overcome HER2 therapy resistance in breast cancer. In M.R. Szewczuk, B. Qorri, M. Sambi (eds), *Current Applications for Overcoming Resistance to Targeted Therapies*, 191–221. Cham, Switzerland: Springer. DOI: 10.1007/978-3-030-21477-7_7

3. Swain, S. M., Shastry, M., & Hamilton, E. (2022). Targeting HER2-positive breast cancer: advances and future directions. *Nature reviews drug discovery, 22*(2), 101–126. DOI: 10.1038/s41573-022-00579-0

4. Tsang, R. Y., Sadeghi, S., & Finn, R. S. (2011). Lapatinib, a dual-targeted small molecule inhibitor of EGFR and HER2, in HER2-amplified breast cancer: from bench to bedside. *Clinical Medicine Insights: Therapeutics*, 3. DOI: 10.4137/CMT.S3783

5. Lin, J., Sampath, D., Nannini, M. A., Lee, B. B., Degtyarev, M., Oeh, J., … Lin, K. (2013). Targeting activated Akt with GDC-0068, a novel selective Akt inhibitor that is efficacious in multiple tumor models. *Clinical cancer research, 19*(7), 1760–1772. DOI: 10.1158/1078-0432.CCR-12-3072

6. Manning, B. D., & Toker, A. (2017). AKT/PKB Signaling: Navigating the Network. *Cell, 169*(3), 381–405. DOI: 10.1016/j.cell.2017.04.001

7. Martorana, F., Motta, G., Pavone, G., Motta, L., Stella, S., Vitale, S. R., … Vigneri, P. (2021). AKT Inhibitors: New Weapons in the Fight Against Breast Cancer? *Frontiers in pharmacology, 12*, 662232. DOI: 10.3389/fphar.2021.662232

8. Fujimoto, Y., Morita, T. Y., Ohashi, A., Haeno, H., Hakozaki, Y., Fujii, … Mukohara, T. (2020). Combination treatment with a PI3K/Akt/mTOR pathway inhibitor overcomes resistance to anti-HER2 therapy in PIK3CA-mutant HER2-positive breast cancer cells. *Scientific reports, 10*(1), 21762. DOI: 10.1038/s41598-020-78646-y

9. Mendoza, M. C., Er, E. E., & Blenis, J. (2011). The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation. Trends in biochemical sciences, 36(6), 320–328. DOI: 10.1016/j.tibs.2011.03.006

10. New, D. C., & Wong, Y. H. (2007). Molecular mechanisms mediating the G protein-coupled receptor regulation of cell cycle progression. *Journal of molecular signaling, 2*, 2. DOI: 10.1186/1750-2187-2-2

11. Kilpatrick, L. E., & Hill, S. J. (2021). Transactivation of G protein-coupled receptors (GPCRs) and receptor tyrosine kinases (RTKs): Recent insights using luminescence and fluorescence technologies. *Current opinion in endocrine and metabolic research, 16*, 102–112. DOI: 10.1016/j.coemr.2020.10.003

12. de Araújo, J. B., Kerkhoff, V. V., de Oliveira Maciel, S. F. V., & de Resende E Silva, D. T. (2021). Targeting the purinergic pathway in breast cancer and its therapeutic applications. *Purinergic signalling, 17*(2), 179–200. DOI: 10.1007/s11302-020-09760-9

13. Woods, L. T., Jasmer, K. J., Muñoz Forti, K., Shanbhag, V. C., Camden, J. M., Erb, L., … Weisman, G. A. (2020). P2Y2 receptors mediate nucleotide-induced EGFR phosphorylation and stimulate proliferation and tumorigenesis of head and neck squamous cell carcinoma cell lines. *Oral oncology, 109*, 104808. DOI: 10.1016/j.oraloncology.2020.104808

14. Karcini, A., & Lazar, I. M. (2022). The SKBR3 cell-membrane proteome reveals telltales of aberrant cancer cell proliferation and targets for precision medicine applications. *Scientific reports, 12*(1), 10847. DOI: 10.1038/s41598-022-14418-0

15. Woods, L. T., Forti, K. M., Shanbhag, V. C., Camden, J. M., & Weisman, G. A. (2021). P2Y receptors for extracellular nucleotides: Contributions to cancer progression and therapeutic implications. *Biochemical pharmacology, 187*, 114406. DOI: 10.1016/j.bcp.2021.114406

16. Vultaggio-Poma, V., Sarti, A. C., & Di Virgilio, F. (2020). Extracellular ATP: A Feasible Target for Cancer Therapy. *Cells, 9*(11), 2496. DOI: 10.3390/cells9112496

17. Kepp, O., Bezu, L., Yamazaki, T., Di Virgilio, F., Smyth, M. J., Kroemer, G., & Galluzzi, L. (2021). ATP and cancer immunosurveillance. *The EMBO journal, 40*(13), e108130. DOI: 10.15252/embj.2021108130

18. Cox, J., & Mann, M. (2007). Is proteomics the new genomics? *Cell, 130*(3), 395–398. DOI: 10.1016/j.cell.2007.07.032

19. Chevalier F. (2010). Highlights on the capacities of "Gel-based" proteomics. *Proteome science, 8*, 23. DOI: 10.1186/1477-5956-8-23

20. Needham, E. J., Parker, B. L., Burykin, T., James, D. E., & Humphrey, S. J. (2019). Illuminating the dark phosphoproteome. *Science signaling, 12*(565), eaau8645. DOI: 10.1126/scisignal.aau8645

21. Olsen, J. V., & Mann, M. (2013). Status of large-scale analysis of post-translational modifications by mass spectrometry. *Molecular & cellular proteomics, 12*(12), 3444–3452. DOI: 10.1074/mcp.O113.034181

22. Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., & Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic acids research, 43*(Database issue), D512–D520. DOI: 10.1093/nar/gku1267

23. Old, W. M., Meyer-Arendt, K., Aveline-Wolf, L., Pierce, K. G., Mendoza, A., Sevinsky, J. R., … Ahn, N. G. (2005). Comparison of label-free methods for quantifying human proteins by

shotgun proteomics. Molecular & cellular proteomics, 4(10), 1487–1502. DOI: 10.1074/mcp.M500084-MCP200

24. Hunter T. (2014). The genesis of tyrosine phosphorylation. *Cold Spring Harbor perspectives in biology, 6*(5), a020644. DOI: 10.1101/cshperspect.a020644

25. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., & Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell, 127*(3), 635–648. DOI: 10.1016/j.cell.2006.09.026

26. Wee, P., & Wang, Z. (2017). Epidermal Growth Factor Receptor Cell Proliferation Signaling Pathways. *Cancers, 9*(5), 52. DOI: 10.3390/cancers9050052

27. Shi, Z., Wulfkuhle, J., Nowicka, M., Gallagher, R. I., Saura, C., Nuciforo, P. G., … Isakoff, S. J. (2022). Functional Mapping of AKT Signaling and Biomarkers of Response from the FAIRLANE Trial of Neoadjuvant Ipatasertib plus Paclitaxel for Triple-Negative Breast Cancer. *Clinical cancer research, 28*(5), 993–1003. DOI: 10.1158/1078-0432.CCR-21-2498

28. Lattanzio, R., Iezzi, M., Sala, G., Tinari, N., Falasca, M., Alberti, S., … Piantelli, M. (2019). PLC-gamma-1 phosphorylation status is prognostic of metastatic risk in patients with early-stage Luminal-A and -B breast cancer subtypes. *BMC cancer, 19*(1), 747. DOI: 10.1186/s12885-019-5949-x

29. Cunnick, J. M., Mei, L., Doupnik, C. A., & Wu, J. (2001). Phosphotyrosines 627 and 659 of Gab1 constitute a bisphosphoryl tyrosine-based activation motif (BTAM) conferring binding and activation of SHP2. *The Journal of biological chemistry, 276*(26), 24380–24387. DOI: 10.1074/jbc.M010275200

30. Pardo, O. E., & Seckl, M. J. (2013). S6K2: The Neglected S6 Kinase Family Member. *Frontiers in oncology, 3*, 191. DOI: 10.3389/fonc.2013.00191

31. Shao, Y., & Aplin, A. E. (2012). ERK2 phosphorylation of serine 77 regulates Bmf pro-apoptotic activity. *Cell death & disease, 3*(1), e253. DOI: 10.1038/cddis.2011.137

32. Bui, N. L., Pandey, V., Zhu, T., Ma, L., Basappa, & Lobie, P. E. (2018). Bad phosphorylation as a target of inhibition in oncology. *Cancer letters, 415*, 177–186. DOI: 10.1016/j.canlet.2017.11.017

33. Qin, X., Jiang, B., & Zhang, Y. (2016). 4E-BP1, a multifactor regulated multifunctional protein. *Cell cycle, 15*(6), 781–786. DOI: 10.1080/15384101.2016.1151581

34. Borah, N. A., & Reddy, M. M. (2021). Aurora Kinase B Inhibition: A Potential Therapeutic Strategy for Cancer. *Molecules, 26*(7), 1981. DOI: 10.3390/molecules26071981

35. Qiu, J., Xu, L., Zeng, X., Wu, Z., Wang, Y., Wang, Y., … Du, Z. (2021). NUSAP1 promotes the metastasis of breast cancer cells via the AMPK/PPARγ signaling pathway. *Annals of translational medicine, 9*(22), 1689. DOI: 10.21037/atm-21-5517

36. Joly-Tonetti, N., Ondet, T., Monshouwer, M., & Stamatas, G. N. (2021). EGFR inhibitors switch keratinocytes from a proliferative to a differentiative phenotype affecting epidermal development and barrier function. *BMC cancer, 21*(1), 5. DOI: 10.1186/s12885-020-07685-5

37. Kim, H. P., Yoon, Y. K., Kim, J. W., Han, S. W., Hur, H. S., Park, J., … Kim, T. Y. (2009). Lapatinib, a dual EGFR and HER2 tyrosine kinase inhibitor, downregulates thymidylate synthase by inhibiting the nuclear translocation of EGFR and HER2. *PloS one, 4*(6), e5933. DOI: 10.1371/journal.pone.0005933

38. O'Neal, J., Clem, A., Reynolds, L., Dougherty, S., Imbert-Fernandez, Y., Telang, S., … Clem, B. F. (2016). Inhibition of 6-phosphofructo-2-kinase (PFKFB3) suppresses glucose metabolism and the growth of HER2+ breast cancer. *Breast cancer research and treatment, 160*(1), 29–40. DOI: 10.1007/s10549-016-3968-8

39. Markowitz, J., & Carson, W. E., 3rd (2013). Review of S100A9 biology and its role in cancer. *Biochimica et biophysica acta, 1835*(1), 100–109. DOI: 10.1016/j.bbcan.2012.10.003

40. Srikrishna G. (2012). S100A8 and S100A9: new insights into their roles in malignancy. *Journal of innate immunity, 4*(1), 31–40. DOI: 10.1159/000330095

41. Rodriguez-Barrueco, R., Yu, J., Saucedo-Cuevas, L. P., Olivan, M., Llobet-Navas, D., Putcha, P., … Silva, J. M. (2015). Inhibition of the autocrine IL-6-JAK2-STAT3-calprotectin axis as targeted therapy for HR-/HER2+ breast cancers. *Genes & development, 29*(15), 1631–1648. DOI: 10.1101/gad.262642.115

42. Mo, S., Fang, D., Zhao, S., Thai Hoa, P. T., Zhou, C., Liang, T., … Han, C. (2022). Down regulated oncogene KIF2C inhibits growth, invasion, and metastasis of hepatocellular carcinoma through the Ras/MAPK signaling pathway and epithelial-to-mesenchymal transition. *Annals of translational medicine, 10*(3), 151. DOI: 10.21037/atm-21-6240

43. Huang, X., Zhao, F., Wu, Q., Wang, Z., Ren, H., Zhang, Q., … Xu, J. (2023). KIF2C Facilitates Tumor Growth and Metastasis in Pancreatic Ductal Adenocarcinoma. *Cancers, 15*(5), 1502. DOI: 10.3390/cancers15051502

44. Liu, S., Ye, Z., Xue, V. W., Sun, Q., Li, H., & Lu, D. (2023). KIF2C is a prognostic biomarker associated with immune cell infiltration in breast cancer. *BMC cancer, 23*(1), 307. DOI: 10.1186/s12885-023-10788-4

45. Wagner, N. B., Weide, B., Gries, M., Reith, M., Tarnanidis, K., Schuermans, V., … Gebhardt, C. (2019). Tumor microenvironment-derived S100A8/A9 is a novel prognostic biomarker for advanced melanoma patients and during immunotherapy with anti-PD-1 antibodies. *Journal for immunotherapy of cancer, 7*(1), 343. DOI: 10.1186/s40425-019-0828-1

46. Helfen, A., Schnepel, A., Rieß, J., Stölting, M., Gerwing, M., Masthoff, M., … Eisenblätter, M. (2021). S100A9-Imaging Enables Estimation of Early Therapy-Mediated Changes in the Inflammatory Tumor Microenvironment. *Biomedicines, 9*(1), 29. DOI: 10.3390/biomedicines9010029

47. Zheng, H. C., & Jiang, H. M. (2022). Shuttling of cellular proteins between the plasma membrane and nucleus (Review). *Molecular medicine reports, 25*(1), 14. DOI: 10.3892/mmr.2021.12530

48. Xu, H., Niu, M., Yuan, X., Wu, K., & Liu, A. (2020). CD44 as a tumor biomarker and therapeutic target. *Experimental hematology & oncology, 9*(1), 36. DOI: 10.1186/s40164-020-00192-0

49. Senbanjo, L. T., & Chellaiah, M. A. (2017). CD44: A Multifunctional Cell Surface Adhesion Receptor Is a Regulator of Progression and Metastasis of Cancer Cells. *Frontiers in cell and developmental biology, 5*, 18. DOI: 10.3389/fcell.2017.00018

50. Boulbes, D. R., Chauhan, G. B., Jin, Q., Bartholomeusz, C., & Esteva, F. J. (2015). CD44 expression contributes to trastuzumab resistance in HER2-positive breast cancer cells. *Breast cancer research and treatment, 151*(3), 501–513. DOI: 10.1007/s10549-015-3414-3

51. Lesniak, D., Sabri, S., Xu, Y., Graham, K., Bhatnagar, P., Suresh, M., & Abdulkarim, B. (2013). Spontaneous epithelial-mesenchymal transition and resistance to HER-2-targeted therapies in HER-2-positive luminal breast cancer. *PloS one, 8*(8), e71987. DOI: 10.1371/journal.pone.0071987

52. Hiscox, S., Baruha, B., Smith, C., Bellerby, R., Goddard, L., Jordan, N., … Gee, J. (2012). Overexpression of CD44 accompanies acquired tamoxifen resistance in MCF7 cells and augments their sensitivity to the stromal factors, heregulin and hyaluronan. *BMC cancer, 12*, 458. DOI: 10.1186/1471-2407-12-458

53. Liu, R., & Jin, J. P. (2016). Calponin isoforms CNN1, CNN2 and CNN3: Regulators for actin cytoskeleton functions in smooth muscle and non-muscle cells. *Gene, 585*(1), 143–153. DOI: 10.1016/j.gene.2016.02.040

54. Nair, V. A., Al-Khayyal, N. A., Sivaperumal, S., & Abdel-Rahman, W. M. (2019). Calponin 3 promotes invasion and drug resistance of colon cancer cells. *World journal of gastrointestinal oncology, 11*(11), 971–982. DOI: 10.4251/wjgo.v11.i11.971

55. Hu, J., Xie, W., Shang, L., Yang, X., Li, Q., Xu, M., … Wu, Y. (2017). Knockdown of calponin 2 suppressed cell growth in gastric cancer cells. *Tumour biology, 39*(7). DOI: 10.1177/1010428317706455

56. Dai, F., Luo, F., Zhou, R., Zhou, Q., Xu, J., Zhang, Z., … Song, L. (2020). Calponin 3 is associated with poor prognosis and regulates proliferation and metastasis in osteosarcoma. *Aging, 12*(14), 14037–14049. DOI: 10.18632/aging.103224

57. Kang, X., Wang, F., Lan, X., Li, X., Zheng, S., Lv, … Zhou, S. (2018). Lentivirus-mediated shRNA Targeting CNN2 Inhibits Hepatocarcinoma in Vitro and in Vivo. *International journal of medical sciences, 15*(1), 69–76. DOI: 10.7150/ijms.21113

58. Zheng, L., Zhao, P., Liu, K., & Kong, D. L. (2020). Ectopic expression of CNN2 of colon cancer promotes cell migration. *Translational cancer research, 9*(2), 1063–1069. DOI: 10.21037/tcr.2019.12.61

59. Yang, C., Zhu, S., Feng, W., & Chen, X. (2021). Calponin 3 suppresses proliferation, migration and invasion of non-small cell lung cancer cells. *Oncology letters, 22*(2), 634. DOI: 10.3892/ol.2021.12895

60. Baldassarre, T., Watt, K., Truesdell, P., Meens, J., Schneider, M. M., Sengupta, S. K., & Craig, A. W. (2015). Endophilin A2 Promotes TNBC Cell Invasion and Tumor Metastasis. *Molecular cancer research, 13*(6), 1044–1055. DOI: 10.1158/1541-7786.MCR-14-0573

61. Baldassarre, T., Truesdell, P., & Craig, A. W. (2017). Endophilin A2 promotes HER2 internalization and sensitivity to trastuzumab-based therapy in HER2-positive breast cancers. *Breast cancer research, 19*(1), 110. DOI: 10.1186/s13058-017-0900-z

62. Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., … Ponten, F. (2017). A pathology atlas of the human cancer transcriptome. *Science, 357*(6352), eaan2507. DOI: 10.1126/science.aan2507. Gene pathology retrieved from: https://www.proteinatlas.org/ENSG00000243943-ZNF512/pathology

63. Perrin, J., Werner, T., Kurzawa, N., Rutkowska, A., Childs, D. D., Kalxdorf, M., … Bergamini, G. (2020). Identifying drug targets in tissues and whole blood with thermal-shift profiling. *Nature biotechnology, 38*(3), 303–308. DOI: 10.1038/s41587-019-0388-4

64. Moore D. (2016). Panobinostat (Farydak): A Novel Option for the Treatment of Relapsed Or Relapsed and Refractory Multiple Myeloma. *P & T : a peer-reviewed journal for formulary management, 41*(5), 296–300.

65. Salvador Moreno, N., Liu, J., Haas, K. M., Parker, L. L., Chakraborty, C., Kron, S. J., … Vidi, P. A. (2019). The nuclear structural protein NuMA is a negative regulator of 53BP1 in DNA double-strand break repair. *Nucleic acids research, 47*(6), 2703–2715. DOI: 10.1093/nar/gkz138

66. Palazzo, L., Della Monica, R., Visconti, R., Costanzo, V., & Grieco, D. (2014). ATM controls proper mitotic spindle structure. *Cell cycle, 13*(7), 1091–1100. DOI: 10.4161/cc.27945

67. Stucci, L. S., Internò, V., Tucci, M., Perrone, M., Mannavola, F., Palmirotta, R., & Porta, C. (2021). The ATM Gene in Breast Cancer: Its Relevance in Clinical Practice. *Genes, 12*(5), 727. DOI: 10.3390/genes12050727

68. Hussien, M. T., Shaban, S., Temerik, D. F., Helal, S. R., Mosad, E., Elgammal, S., … Ibrahim, A. (2020). Impact of DAXX and ATRX expression on telomere length and prognosis of breast cancer patients. *Journal of the Egyptian National Cancer Institute, 32*(1), 34. DOI: 10.1186/s43046-020-00045-1

69. VandenBussche, C. J., Allison, D. B., Graham, M. K., Charu, V., Lennon, A. M., Wolfgang, C. L., … Heaphy, C. M. (2017). Alternative lengthening of telomeres and ATRX/DAXX loss can be reliably detected in FNAs of pancreatic neuroendocrine tumors. *Cancer cytopathology, 125*(7), 544–551. DOI: 10.1002/cncy.21857

70. Bain, J. M., Cho, M. T., Telegrafi, A., Wilson, A., Brooks, S., Botti, C., … Chung, W. K. (2016). Variants in HNRNPH2 on the X Chromosome Are Associated with a Neurodevelopmental Disorder in Females. *American journal of human genetics, 99*(3), 728–734. DOI: 10.1016/j.ajhg.2016.06.028

71. Imami, K., Sugiyama, N., Imamura, H., Wakabayashi, M., Tomita, M., Taniguchi, M., … Ishihama, Y. (2012). Temporal profiling of lapatinib-suppressed phosphorylation signals in EGFR/HER2 pathways. *Molecular & cellular proteomics, 11*(12), 1741–1757. DOI: 10.1074/mcp.M112.019919

72. Li, H., Liu, J., Shen, S., Dai, D., Cheng, S., Dong, X., … Guo, X. (2020). Pan-cancer analysis of alternative splicing regulator heterogeneous nuclear ribonucleoproteins (hnRNPs) family and their prognostic potential. *Journal of cellular and molecular medicine, 24*(19), 11111–11119. DOI: 10.1111/jcmm.15558

73. Tosun Yildirim, H., Aktas, S., Diniz, G., Aktas, T. C., Baran, B., Bayrak, S., … Olgun, N. (2019). Scanning all chromosomal abnormalities with microarray-based comparative genomic hybridization in differential diagnosis of pediatric cancers. *International journal of clinical and experimental pathology, 12*(8), 3140–3148.

74. Zhang, J., Fan, J., Zhou, C., & Qi, Y. (2017). miR-363-5p as potential prognostic marker for hepatocellular carcinoma indicated by weighted co-expression network analysis of miRNAs and mRNA. *BMC gastroenterology, 17*(1), 81. DOI: 10.1186/s12876-017-0637-2

75. Dedeurwaerder, S., Desmedt, C., Calonne, E., Singhal, S. K., Haibe-Kains, B., Defrance, M., … Fuks, F. (2011). DNA methylation profiling reveals a predominant immune component in breast cancers. *EMBO molecular medicine, 3*(12), 726–741. DOI: 10.1002/emmm.201100801

76. Bilancia, D., Rosati, G., Dinota, A., Germano, D., Romano, R., & Manzione, L. (2007). Lapatinib in breast cancer. *Annals of oncology, 18*, Suppl 6, vi26–vi30. DOI: 10.1093/annonc/mdm220

77. Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., … Lancet, D. (2016). The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Current protocols in bioinformatics, 54*, 1.30.1–1.30.33. DOI: 10.1002/cpbi.5

78. Vogt, P. K., & Hart, J. R. (2011). PI3K and STAT3: a new alliance. *Cancer discovery, 1*(6), 481–486. DOI: 10.1158/2159-8290.CD-11-0218

79. Bian, C., Liu, Z., Li, D., & Zhen, L. (2018). PI3K/AKT inhibition induces compensatory activation of the MET/STAT3 pathway in non-small cell lung cancer. *Oncology letters, 15*(6), 9655–9662. DOI: 10.3892/ol.2018.8587

80. Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., … Ponten, F. (2017). A pathology atlas of the human cancer transcriptome. *Science, 357*(6352), eaan2507. DOI: 10.1126/science.aan2507. Gene pathology retrieved from: https://www.proteinatlas.org/ENSG00000203879-GDI1/pathology

81. Haga, R. B., & Ridley, A. J. (2016). Rho GTPases: Regulation and roles in cancer cell biology. *Small GTPases, 7*(4), 207–221. DOI: 10.1080/21541248.2016.1232583

82. Porter, A. P., Papaioannou, A., & Malliri, A. (2016). Deregulation of Rho GTPases in cancer. *Small GTPases, 7*(3), 123–138. DOI: 10.1080/21541248.2016.1173767

83. Wang, H., Wang, B., Liao, Q., An, H., Li, W., Jin, X., … Zhao, L. (2014). Overexpression of RhoGDI, a novel predictor of distant metastasis, promotes cell proliferation and migration in hepatocellular carcinoma. *FEBS letters, 588*(3), 503–508. DOI: 10.1016/j.febslet.2013.12.016

84. Huang, D., Lu, W., Zou, S., Wang, H., Jiang, Y., Zhang, X., … Fang, L. (2017). Rho GDP-dissociation inhibitor α is a potential prognostic biomarker and controls telomere regulation in colorectal cancer. *Cancer science, 108*(7), 1293–1302. DOI: 10.1111/cas.13259

85. Lebok, P., Öztürk, M., Heilenkötter, U., Jaenicke, F., Müller, V., Paluchowski, P., … Quaas, A. (2016). High levels of class III β-tubulin expression are associated with aggressive tumor features in breast cancer. *Oncology letters, 11*(3), 1987–1994. DOI: 10.3892/ol.2016.4206

86. Levallet, G., Bergot, E., Antoine, M., Creveuil, C., Santos, A. O., Beau-Faller, M., … Intergroupe Francophone de Cancérologie Thoracique (IFCT) (2012). High TUBB3 expression, an independent prognostic marker in patients with early non-small cell lung cancer treated by preoperative chemotherapy, is regulated by K-Ras signaling pathway. *Molecular cancer therapeutics, 11*(5), 1203–1213. DOI: 10.1158/1535-7163.MCT-11-0899

87. McCarroll, J. A., Gan, P. P., Erlich, R. B., Liu, M., Dwarte, T., Sagnella, S. S., … Kavallaris, M. (2015). TUBB3/βIII-tubulin acts through the PTEN/AKT signaling axis to promote tumorigenesis and anoikis resistance in non-small cell lung cancer. *Cancer research, 75*(2), 415–425. DOI: 10.1158/0008-5472.CAN-14-2740

88. Sekino, Y., Han, X., Kawaguchi, T., Babasaki, T., Goto, K., Inoue, S., … Matsubara, A. (2019). TUBB3 Reverses Resistance to Docetaxel and Cabazitaxel in Prostate Cancer. *International journal of molecular sciences, 20*(16), 3936. DOI: 10.3390/ijms20163936

89. Martinez, E., Vazquez, N., Lopez, A., Fanniel, V., Sanchez, L., Marks, R., … Keniry, M. (2020). The PI3K pathway impacts stem gene expression in a set of glioblastoma cell lines. *Journal of cancer research and clinical oncology, 146*(3), 593–604. DOI: 10.1007/s00432-020-03133-w

90. Midha, M. K., Huang, Y. F., Yang, H. H., Fan, T. C., Chang, N. C., Chen, T. H., … Chen, C. J. (2020). Comprehensive Cohort Analysis of Mutational Spectrum in Early Onset Breast Cancer Patients. *Cancers, 12*(8), 2089. DOI: 10.3390/cancers12082089

91. Wang, Y., Chen, Y., Xiao, S., & Fu, K. (2020). Integrated Analysis of the Functions and Prognostic Values of RNA-Binding Proteins in Colorectal Cancer. *Frontiers in cell and developmental biology, 8*, 595605. DOI: 10.3389/fcell.2020.595605

92. Dong, H., Wang, W., Mo, S., Liu, Q., Chen, X., Chen, R., … Hu, J. (2018). Long non-coding RNA SNHG14 induces trastuzumab resistance of breast cancer via regulating PABPC1 expression through H3K27 acetylation. *Journal of cellular and molecular medicine, 22*(10), 4935–4947. DOI: 10.1111/jcmm.13758

93. Wang, Q., & Yang, H. S. (2018). The role of Pdcd4 in tumour suppression and protein translation. *Biology of the cell*. DOI: 10.1111/boc.201800014

94. Cai, Q., Yang, H. S., Li, Y. C., & Zhu, J. (2022). Dissecting the Roles of PDCD4 in Breast Cancer. *Frontiers in oncology, 12*, 855807. DOI: 10.3389/fonc.2022.855807

95. Afonja, O., Juste, D., Das, S., Matsuhashi, S., & Samuels, H. H. (2004). Induction of PDCD4 tumor suppressor gene expression by RAR agonists, antiestrogen and HER-2/neu antagonist in breast cancer cells. Evidence for a role in apoptosis. *Oncogene, 23*(49), 8135–8145. DOI: 10.1038/sj.onc.1207983

96. Matsuhashi, S., Manirujjaman, M., Hamajima, H., & Ozaki, I. (2019). Control Mechanisms of the Tumor Suppressor PDCD4: Expression and Functions. *International journal of molecular sciences, 20*(9), 2304. DOI: 10.3390/ijms20092304

97. Liu, Y., Cui, J., Hoffman, A. R., & Hu, J. F. (2023). Eukaryotic translation initiation factor eIF4G2 opens novel paths for protein synthesis in development, apoptosis and cell differentiation. *Cell proliferation, 56*(3), e13367. DOI: 10.1111/cpr.13367

98. Wolf, D. M., Yau, C., Wulfkuhle, J., Brown-Swigart, L., Gallagher, R. I., Magbanua, M. J. M., … van 't Veer, L. (2020). Mechanism of action biomarkers predicting response to AKT inhibition in the I-SPY 2 breast cancer trial. *NPJ breast cancer, 6*, 48. DOI: 10.1038/s41523-020-00189-2

99. McLean, M. E., MacLean, M. R., Cahill, H. F., Arun, R. P., Walker, O. L., Wasson, M. D., … Marcato, P. (2023). The Expanding Role of Cancer Stem Cell Marker ALDH1A3 in Cancer and Beyond. *Cancers, 15*(2), 492. DOI: 10.3390/cancers15020492

100. Wang, T., Sun, H., Bao, Y., En, R., Tian, Y., Zhao, W., & Jia, L. (2020). POM121 overexpression is related to a poor prognosis in colorectal cancer. *Expert review of molecular diagnostics, 20*(3), 345–353. DOI: 10.1080/14737159.2020.1707670

101. Ma, H., Li, L., Jia, L., Gong, A., Wang, A., Zhang, L., … Tang, G. (2019). POM121 is identified as a novel prognostic marker of oral squamous cell carcinoma. *Journal of Cancer, 10*(19), 4473–4480. DOI: 10.7150/jca.33368

102. Luo, J., Zou, H., Guo, Y., Tong, T., Ye, L., Zhu, C., … Li, P. (2022). SRC kinase-mediated signaling pathways and targeted therapies in breast cancer. *Breast cancer research, 24*(1), 99. DOI: 10.1186/s13058-022-01596-y

103. Mihaylova, M. M., & Shaw, R. J. (2011). The AMPK signalling pathway coordinates cell growth, autophagy and metabolism. *Nature cell biology, 13*(9), 1016–1023. DOI: 10.1038/ncb2329

104. Hayashi, N., Iwamoto, T., Gonzalez-Angulo, A. M., Ferrer-Lozano, J., Lluch, A., Niikura, N., … Ueno, N. T. (2011). Prognostic impact of phosphorylated HER-2 in HER-2+ primary breast cancer. *The oncologist, 16*(7), 956–965. DOI: 10.1634/theoncologist.2010-0409

105. Roskoski R., Jr (2012). ERK1/2 MAP kinases: structure, function, and regulation. *Pharmacological research, 66*(2), 105–143. DOI: 10.1016/j.phrs.2012.04.005

106. Nikas, I. P., Themistocleous, S. C., Paschou, S. A., Tsamis, K. I., & Ryu, H. S. (2019). Serine-Arginine Protein Kinase 1 (SRPK1) as a Prognostic Factor and Potential Therapeutic Target in Cancer: Current Evidence and Future Perspectives. *Cells, 9*(1), 19. DOI: 10.3390/cells9010019

107. Runde, A. P., Mack, R., S J, P. B., & Zhang, J. (2022). The role of TBK1 in cancer pathogenesis and anticancer immunity. *Journal of experimental & clinical cancer research, 41*(1), 135. DOI: 10.1186/s13046-022-02352-y

108. Al-Akhrass, H., Pietilä, M., Lilja, J., Vesilahti, E. M., Anttila, J. M., Haikala, H. M., … Ivaska, J. (2022). Sortilin-related receptor is a druggable therapeutic target in breast cancer. *Molecular oncology, 16*(1), 116–129. DOI: 10.1002/1878-0261.13106

109. Veeriah, S., Brennan, C., Meng, S., Singh, B., Fagin, J. A., Solit, D. B., … Chan, T. A. (2009). The tyrosine phosphatase PTPRD is a tumor suppressor that is frequently inactivated and mutated in glioblastoma and other human cancers. *Proceedings of the National Academy of Sciences of the United States of America, 106*(23), 9435–9440. DOI: 10.1073/pnas.0900571106

110. Kim, M., Morales, L. D., Jang, I. S., Cho, Y. Y., & Kim, D. J. (2018). Protein Tyrosine Phosphatases as Potential Regulators of STAT3 Signaling. *International journal of molecular sciences, 19*(9), 2708. DOI: 10.3390/ijms19092708

111. Han, L., Shi, H., Ma, S., Luo, Y., Sun, W., Li, S., … Gong, Y. (2022). Agrin Promotes Non-Small Cell Lung Cancer Progression and Stimulates Regulatory T Cells via Increasing IL-6 Secretion Through PI3K/AKT Pathway. *Frontiers in oncology, 11*, 804418. DOI: 10.3389/fonc.2021.804418

112. Podojil, J. R., & Miller, S. D. (2017). Potential targeting of B7-H4 for the treatment of cancer. *Immunological reviews, 276*(1), 40–51. DOI: 10.1111/imr.12530

113. Huang, G., Li, H., & Zhang, H. (2020). Abnormal Expression of Mitochondrial Ribosomal Proteins and Their Encoding Genes with Cell Apoptosis and Diseases. *International journal of molecular sciences, 21*(22), 8879. DOI: 10.3390/ijms21228879

114. Kim, H. J., Maiti, P., & Barrientos, A. (2017). Mitochondrial ribosomes in cancer. *Seminars in cancer biology, 47*, 67–81. DOI: 10.1016/j.semcancer.2017.04.004

115. Pecoraro, A., Pagano, M., Russo, G., & Russo, A. (2021). Ribosome Biogenesis and Cancer: Overview on Ribosomal Proteins. *International journal of molecular sciences, 22*(11), 5496. DOI: 10.3390/ijms22115496

116. Zhang, Y., Wolf, G. W., Bhat, K., Jin, A., Allio, T., Burkhart, W. A., & Xiong, Y. (2003). Ribosomal protein L11 negatively regulates oncoprotein MDM2 and mediates a p53-dependent ribosomal-stress checkpoint pathway. *Molecular and cellular biology, 23*(23), 8902–8912. DOI: 10.1128/MCB.23.23.8902-8912.2003

CHAPTER 7. CONCLUSION AND FUTURE WORK

This project was aimed at profiling the cell membrane proteome of SKBR3/HER2+ breast cancer cells and exploring the interactions between the cell membrane proteins and the signaling pathways that are triggered or altered by the treatment of cells with growth factors and/or inhibitory drugs targeted against cell surface receptors. The project was completed in three stages with focus on: (1) isolating and profiling by mass spectrometry the cell surface (CS) proteins of SKBR3 cells cultured in the presence or absence of serum, to advance the understanding of the complex and dynamic surfaceome; (2) determining the functional role of the detected cell surface proteins in the context of cancer hallmarks, and exploring their mutational profile, to expand the drug targeting opportunities in HER2+ cells; and (3) analyzing the cellular events that occur in response to treatment with a single agent or a drug combination, to provide insights into the combined mechanism of action of two therapeutic drugs that target receptors with critical role in supporting cancer cell proliferation. This investigation was carried out by mass spectrometry approaches leveraging the fields of proteomics, phosphoproteomics, and proteogenomics.

Due to low abundance, hydrophobic properties, and the presence of post-translational modifications (PTMs), cell surface proteins are challenging to analyze. Therefore, multiple enrichment methods, including amine and glycan biotin labeling, and enzymatic shaving, were successfully employed to recover the CS proteins from complex cellular extracts, enhance the ability to detect these proteins, and expand the pool of known CS proteins in SKBR3 cells. The identification of a diverse range of cell-surface proteins such as kinase receptors, CD antigens, adhesion molecules, and transporters demonstrated the critical role that the cell membrane plays in promoting and sustaining the growth and survival of cancer cells. The analysis of protein-protein interaction networks further revealed the complexity of signaling mechanisms that sustain abnormal cell growth, invasion, metastasis, and immune escape, initiated not only from the well-established receptors and CD antigens, but also from the diverse protein families of metalloproteinases, nectins, ephrins, and bone morphogenetic receptors. The changes in the abundance of specific cell-surface proteins due to serum removal highlighted potential mechanisms that alter cancer cell metabolism and energy rewiring for continued proliferation. Overall, the analysis of the SKBR3 surfaceome underscored the complex landscape and

multifunctional role of cell-membrane proteins, revealing novel prospects for developing therapeutic approaches that target cell growth, survival, apoptotic, angiogenic, and migration pathways.

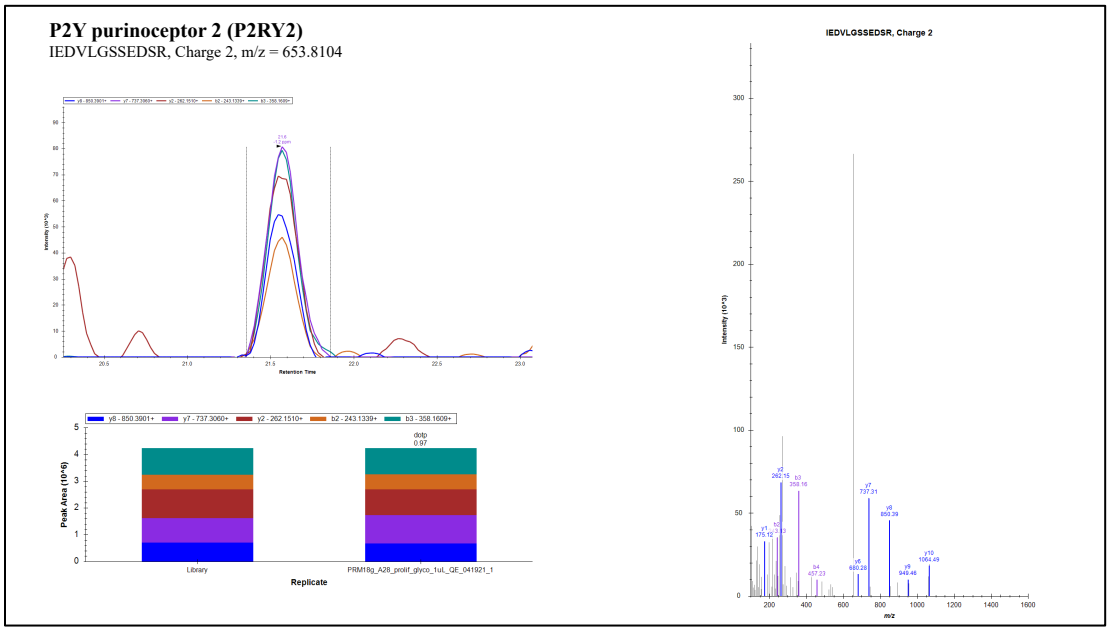
Mapping the detected cell surface proteins to the cancer hallmarks uncovered highly interconnected PPI networks between the representative proteins of each hallmark, indicating the complex scenery of critical players that promote and sustain the development and metastatic propensity of these cancer cells. The identification of numerous oncogenes and tumor suppressors, and of many of their protein interactors, exposed new opportunities for devising more effective combination therapies that target the sustainability of various cancer hallmarks. The proteogenomic approach used in this study identified pathogenic mutations in cell surface proteins across all functional categories (catalytic receptors, CDs, adhesion, and transport molecules), and provided insights into the biological pathways and hallmark processes that could be affected by the mutated proteins. The detection of mutated sites in the cell surface protein drug targets (for example, of ERBB2) provided a better understanding of the resistance mechanisms that evolve in response to current targeted therapeutic approaches, aiding the development of future cancer treatment strategies that can act more effectively against targetable mutations. Altogether, the availability of tumor cell-membrane protein profiles with their corresponding mutations provides a significant stepping stone for advancing the field of targeted therapeutics, but further research is needed in apprehending and experimentally validating the role of specific mutations in oncogenesis and drug resistance.

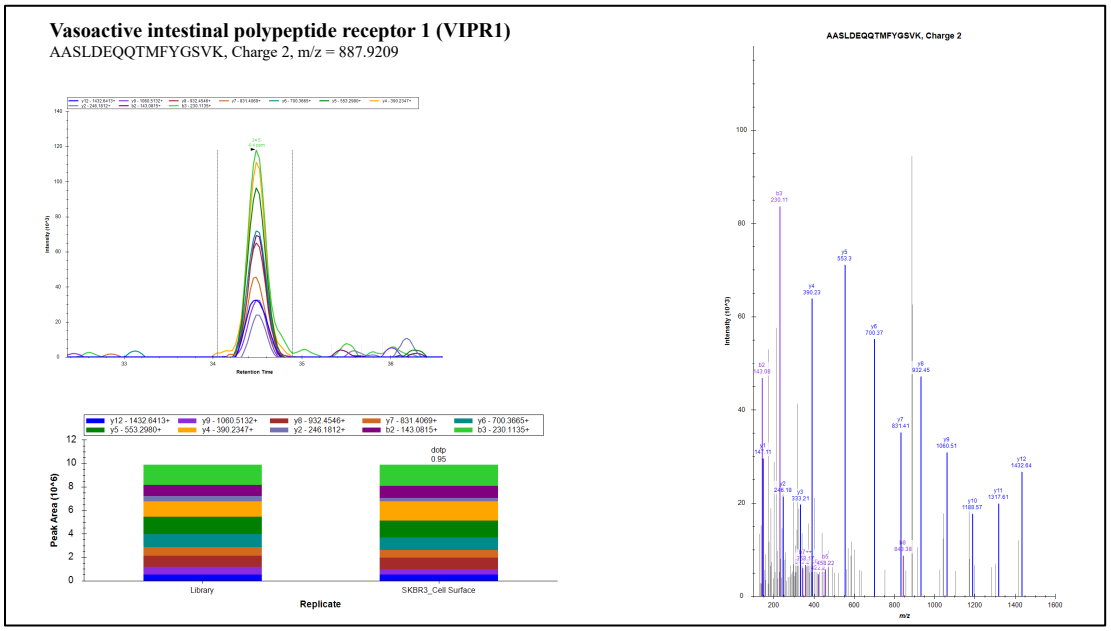
Last, this study revealed a number of cellular processes that reflect the therapeutic effects of small molecule inhibitors, i.e., of lapatinib targeting EGFR and HER2 cell surface receptors, and of ipatasertib targeting the Akt receptors. Lapatinib treatment had a significant impact on the cell cycle and cell proliferation pathways in SKBR3 cells, resulting in the inhibition of cell cycle progression and the upregulation of many proteins involved in cell cycle arrest and DNA repair. Combination with ipatasertib triggered the upregulation of additional biological processes that included, among others, regulation of apoptosis, localization, exocytosis, and signaling pathways such as FoxO, mTOR, and TGF-beta. The qualitative assessment of the phosphoproteome profiles associated with these drug treatments enabled the identification of phosphosites of

relevance to specific signal transduction networks. The differentially expressed proteins represented a vast pool of investigational drug targets, potential immunotherapeutic targets, and apoptosis-inducing MRPs, expanding further the opportunities for novel and combinatorial targeted therapies. Future work will focus on quantifying the changes in the phosphorylation of specific sites and exploring the therapeutic potential of lapatinib and ipatasertib in combination with other drugs to improve, overall, breast cancer treatment outcomes.

This study generated extensive proteome profiles of the various cellular fractions and states of the SKBR3/HER2+ breast cancer cell line, including cell surface, nuclear, cytoplasmic, phosphorylated, and mutated profiles, under different treatment regimens with growth factors and therapeutic drugs. The analysis explored the landscape of early signaling events and late cellular changes in protein expression, as well as the crosstalk between the associated signaling pathways, and mapped the detected proteins to relevant cancer hallmarks and representative processes. Many new potential drug targets for combination therapeutics were identified, with mechanisms of action that affect signal transduction, immune response, and metabolism in SKBR3 cancer cells. This study has enabled the generation of high-quality mass spectrometry proteomics data that will be available to the scientific community to help pave the way for future investigations into specific cell-membrane protein drug targets, for identifying mutations that lead to drug resistance, for elucidating the involvement of these proteins in the cancer hallmarks, and for advancing personalized therapeutic approaches.

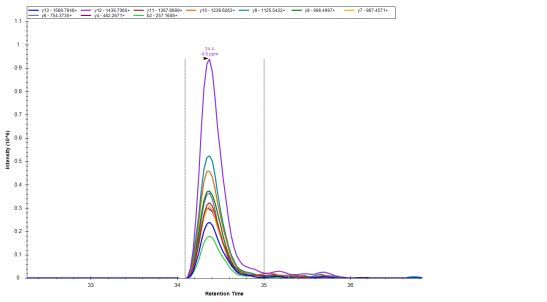
Appendix A. Validation of some of the detected GPCRs by parallel reaction monitoring (PRM). A representative protein of each homology class (**Figure 4-7 B**) is shown. The full dataset of validated proteins can be found in Karcini & Lazar, 2022. Reprinted with permission.[20]

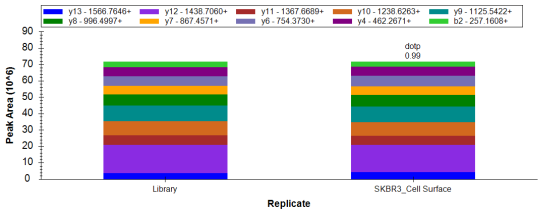


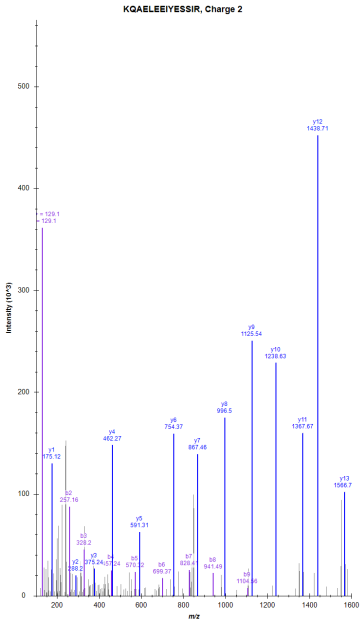


CD97 antigen (CD97)

KQAELEEIYESSIR, Charge 2, m/z = 847.9334

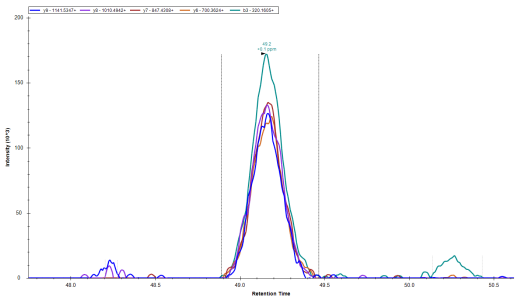


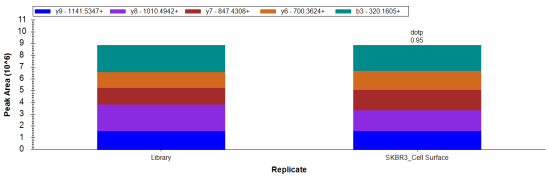


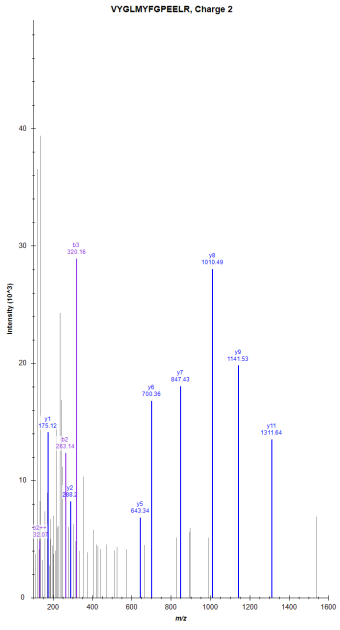


Frizzled-1 (FZD1)

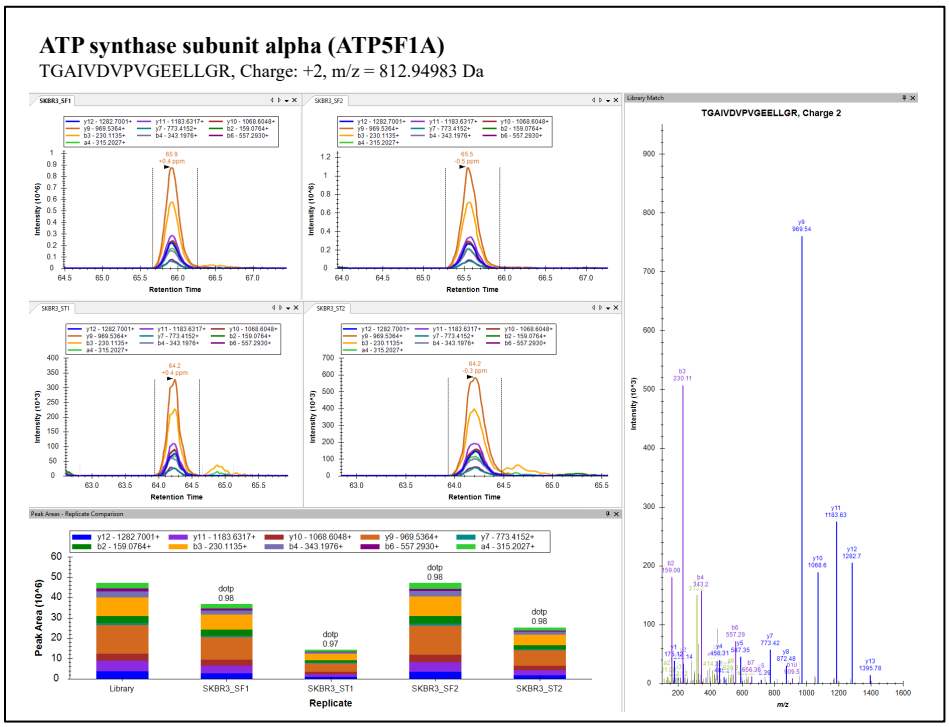
VYGLMYFGPEELR, Charge 2, m/z = 787.3892

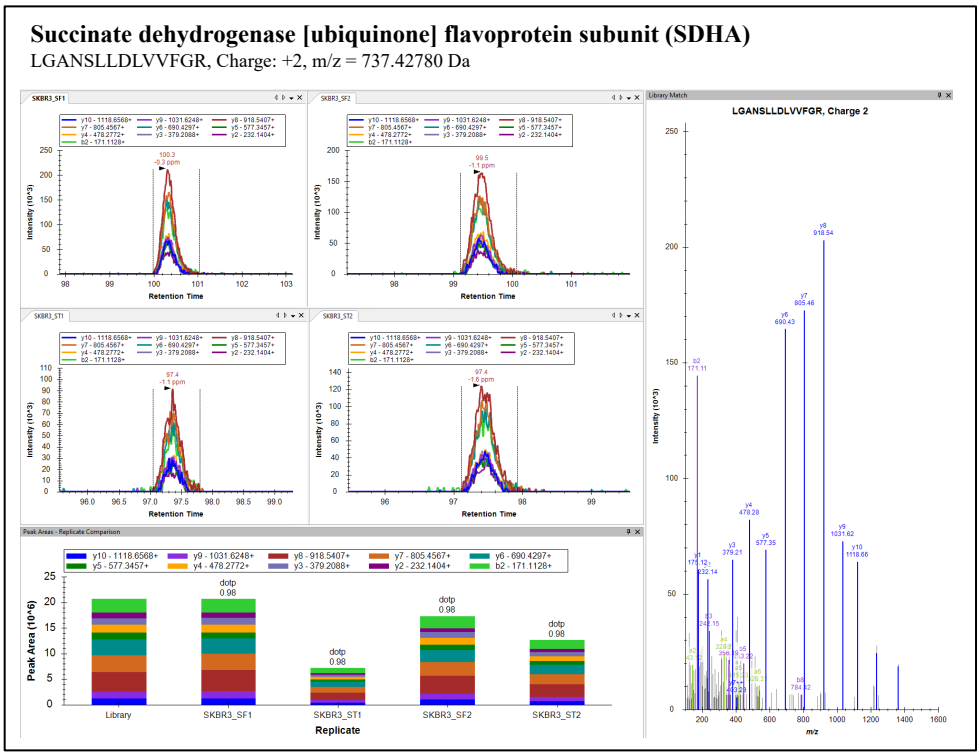






Appendix B. Validation of some of the detected up- and down- regulated CS proteins by PRM. Two representative proteins are shown for each state: elevated expression in serum-free/ SF (ATP5F1A, SDHA) and elevated expression in serum-treated/ ST (P2RY2, CD44). The full dataset of validated proteins can be found in Karcini & Lazar, 2022. Reprinted with permission.[20]



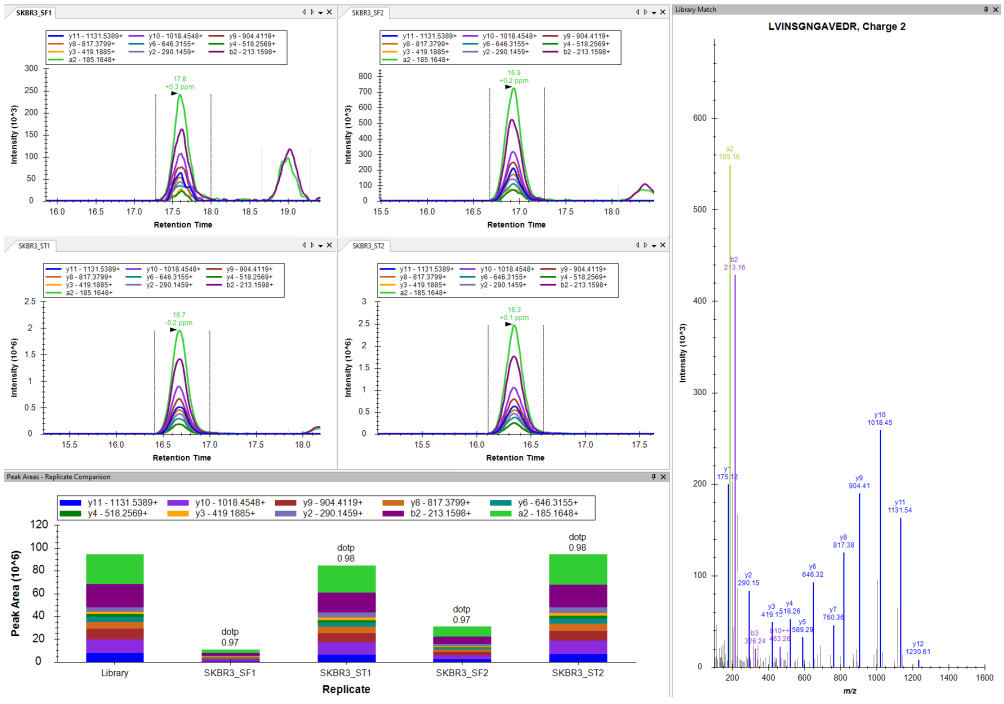


P2Y purinoceptor 2 (P2RY2)

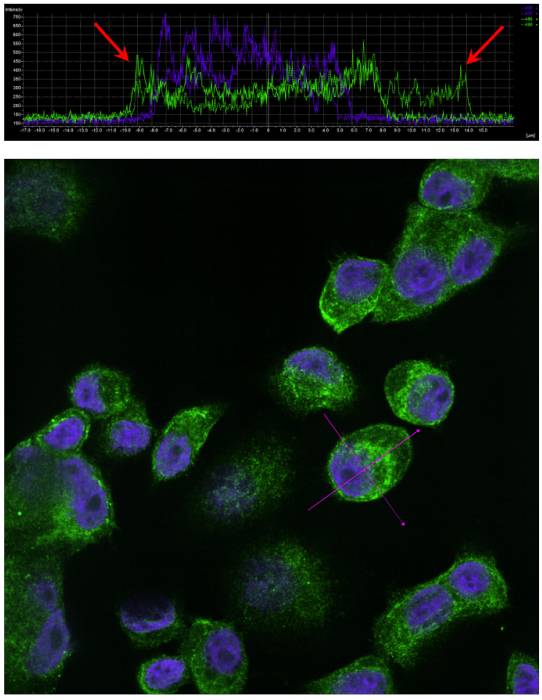
TLYYSFR, Charge: +2, m/z = 475.24411 Da

CD44 antigen (CD44)

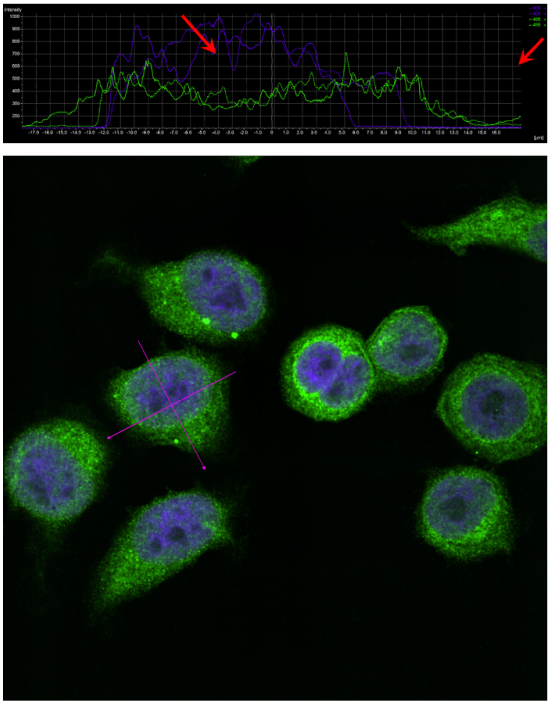
LVINSGNGAVEDR, Charge: +2, m/z = 672.34949 Da



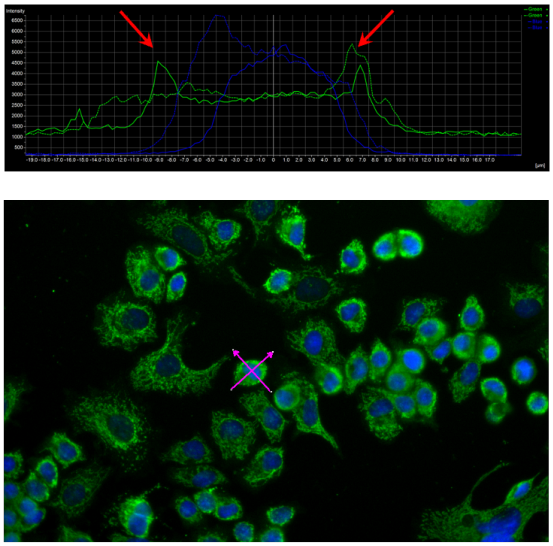
Appendix C. Immunofluorescence labeling of ATP5F1A (elevated in SF) and P2RY2 (elevated in ST) with their respective primary antibodies at 1:100 dilution to assess their intensity profiles in each cell states. Images were acquired either with a 40X water objective in confocal scanning with SoRa mode, exposed at 80 ms DAPI (blue), and 200 ms FITC 488 (green) for ATP5F1A in a Nikon Eclipse Ti2; or with a 20X air objective and exposed at: 10 ms DAPI (blue), 600 ms FITC 488 (green) for P2RY2 in a Nikon Eclipse TE2000-U. Images were processed with NIS-Elements AR Analysis software 5.11.01. Reprinted with permission.[20]



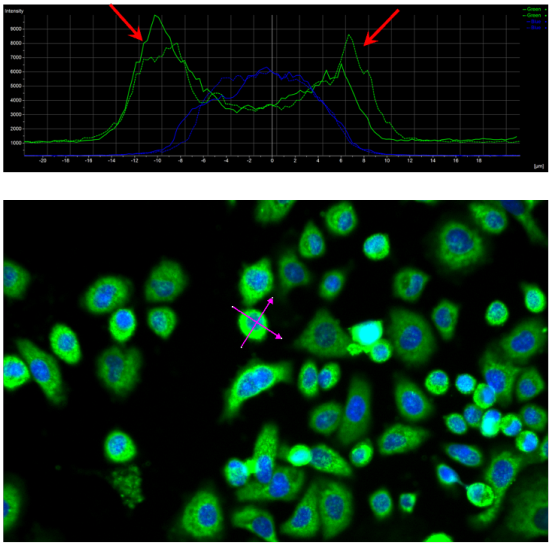
ATP5F1A_SKBR3_48h arrest



ATP5F1A_SKBR3_48h arrest/24h release



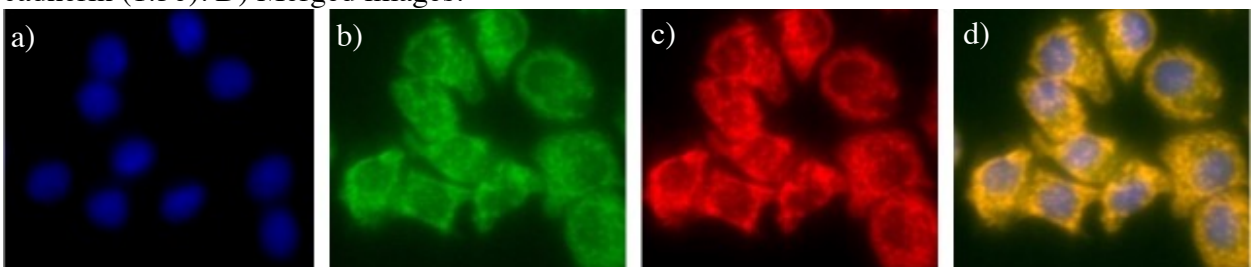
P2RY2_SKBR3_48h arrest



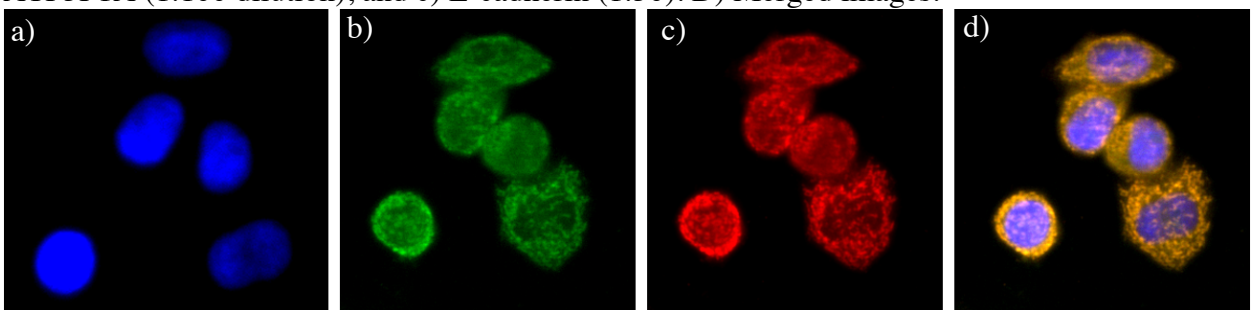
P2RY2_SKBR3_48h arrest/24h release

Appendix D. Zoomed images of surface co-localization of ATP5F1A and UQCRC2 with the cell surface marker E-cadherin. The immunofluorescence labeling was performed as described in **Section 3.14**. The SKBR3 cells were fixed by methanol and non-permeabilized. The primary antibody (Ab) are listed below and the dilutions are provided in parenthesis. The images were acquired with a 20X air objective and exposed at: 20 ms DAPI (blue), 500 ms FITC 488 (green), 800 ms TRITC (red).

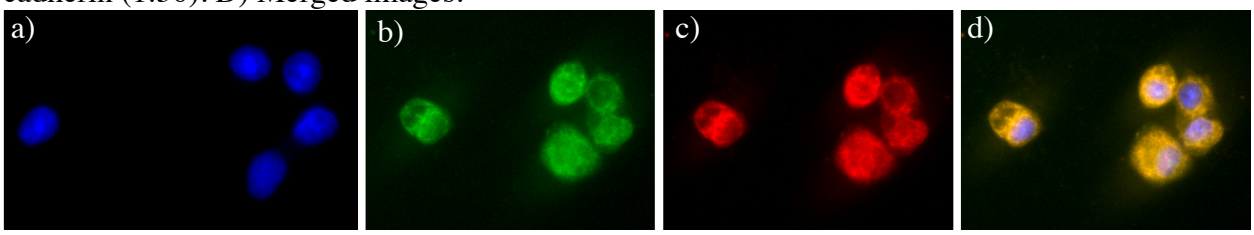
SKBR3 cells conditions: 48 h serum-free, stained with a) DAPI, b) ATP5F1A (1:100), and c) E-cadherin (1:50). D) Merged images.



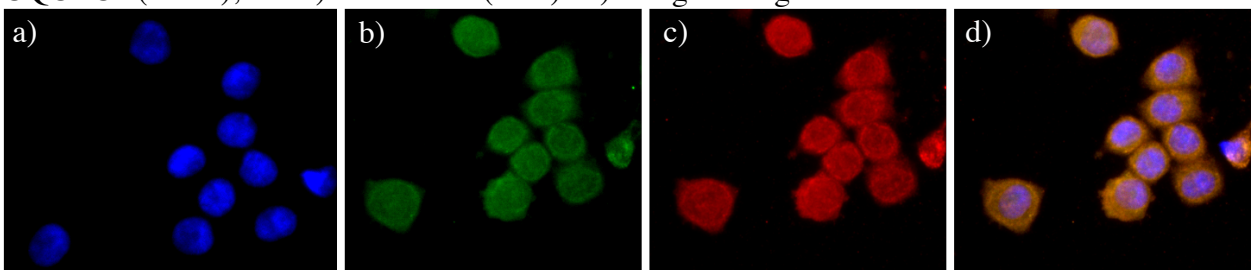
SKBR3 cells conditions: 48 h serum-free/24 h serum-treated, stained with a) DAPI, b) ATP5F1A (1:100 dilution), and c) E-cadherin (1:50). D) Merged images.



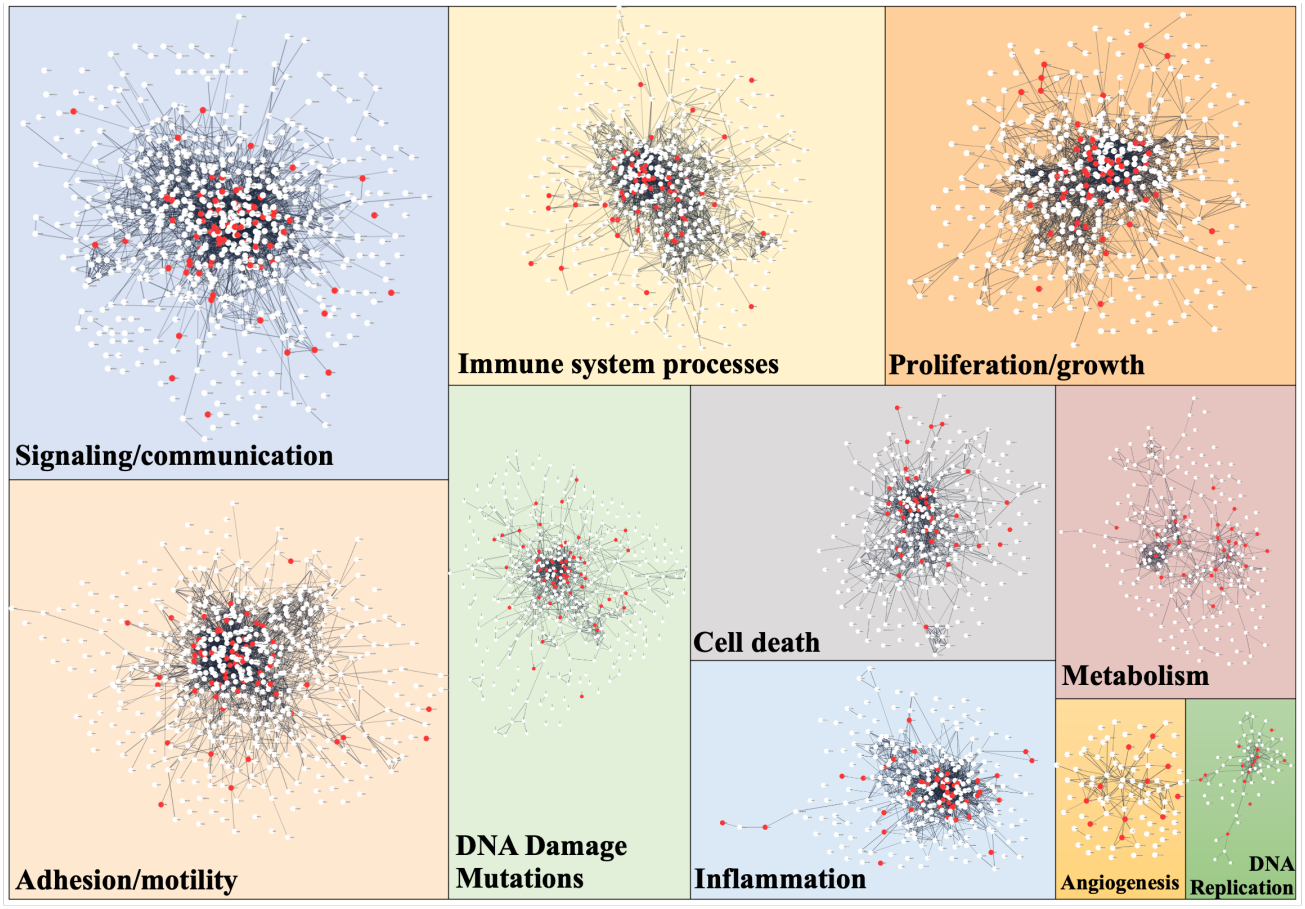
SKBR3 cells conditions: 48 h serum-free, stained with a) DAPI, b) UQCRC2 (1:100), and c) E-cadherin (1:50). D) Merged images.



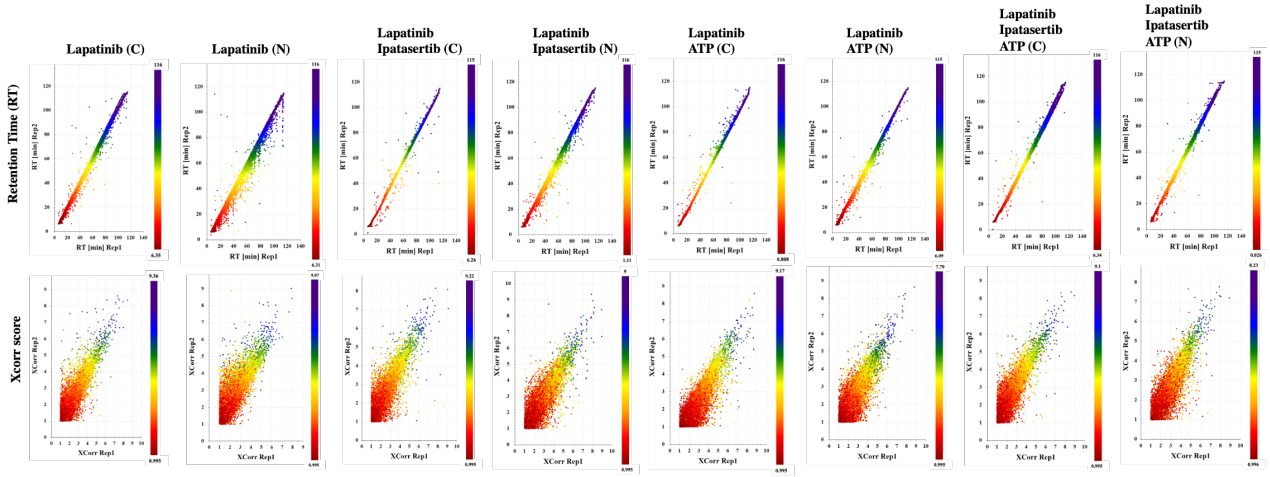
SKBR3 cells conditions: 48 h serum-free/24 h serum-treated, stained with a) DAPI, b) UQCRC2 (1:100), and c) E-cadherin (1:50). D) Merged images.



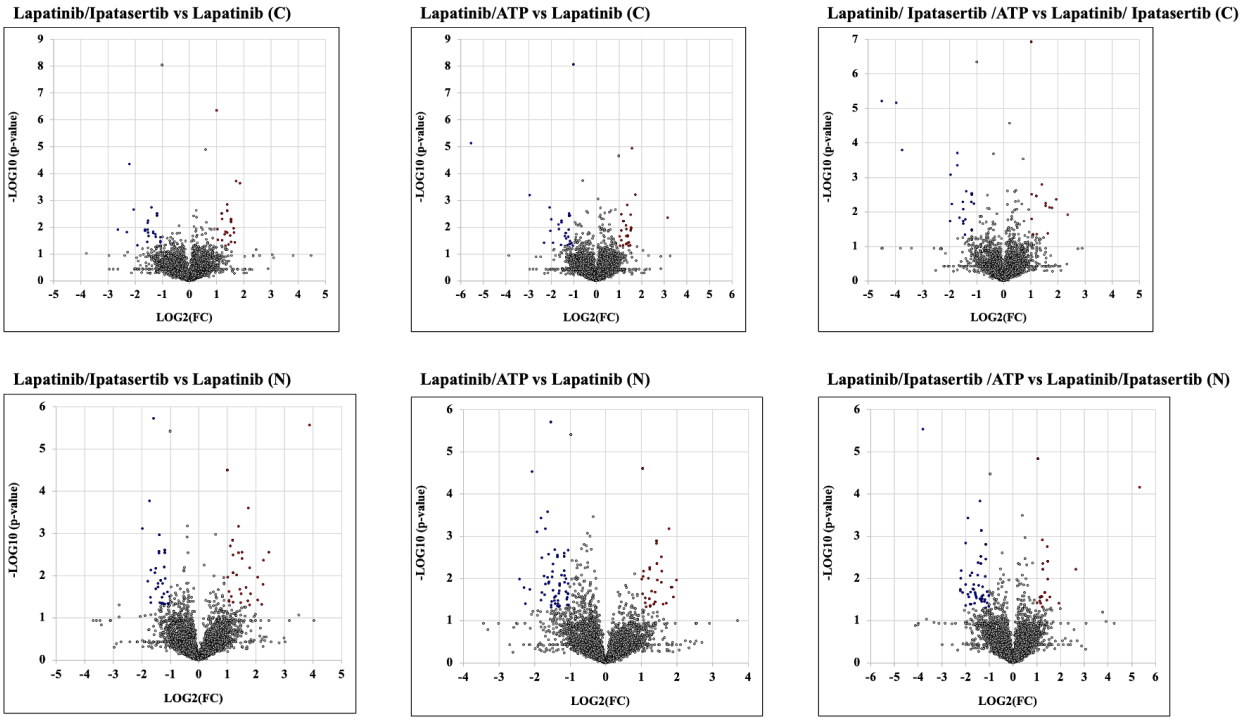
Appendix E. Extended PPI networks of the detected CS proteins with roles in cancer hallmarks. Red icons indicate whether the protein was annotated as a cancer hallmark protein by CGC. These complex networks provide a vast array of potential proteins that can be therapeutically targeted. Reprinted with permission.[15]



Appendix F. Correlation of the elution times and XCorr scores of peptides from each biological replicate of the cytoplasmic and nuclear fractions of each treatment: replicate 1 (X-axis), replicate 2 (Y-axis), and replicate 3 (color bar).



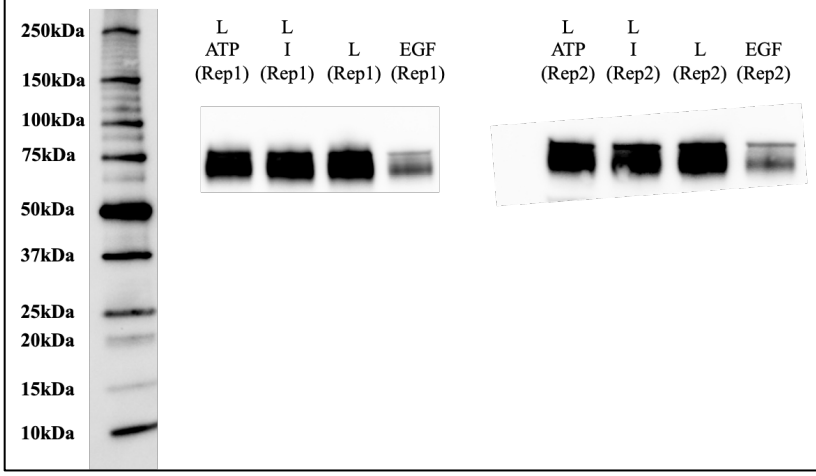
Appendix G. Volcano plots of nuclear and cytoplasmic fractions of the other treatment comparisons not shown in **Figure 6-3**. Differentially expressed proteins are indicated in red (up-regulated) and blue (down-regulated), and display greater than or equal to 2-fold change abundance ratio with p-value less than or equal to 0.05.



Appendix H. Differentially expressed proteins validated by western blotting. The experimental procedure was described in **Section 3.15**.

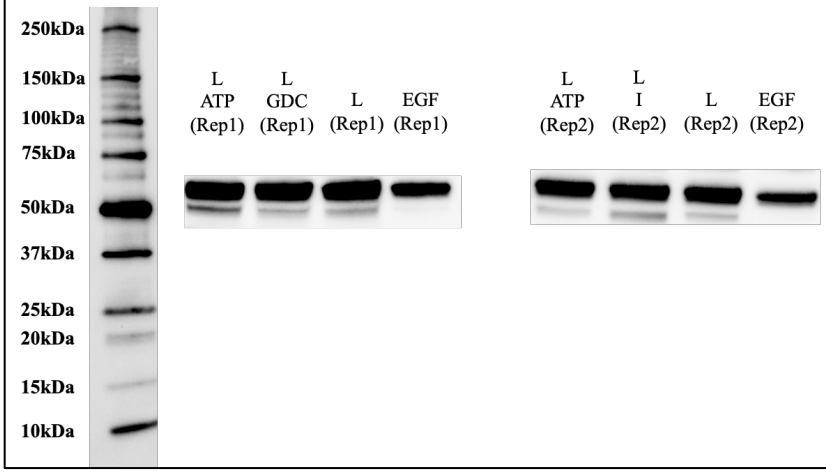
Protein: VTCN1/ B7-H4/ Q7Z7D3, MW = 75 kDa, upregulated in L vs EGF and L/I vs EGF (cytoplasmic fraction)
Conditions:
- Primary Ab –VTCN1 (CST-14572) 1:1000
- Secondary Ab – Anti Rabbit HRP conjugate (CST- 7074S) 1:2000

Note: L-lapatinib, I-ipatasertib



Protein: PDCD4/ Q53EL6, MW = 60 kDa, upregulated in L vs EGF and L/I vs EGF (cytoplasmic fraction)
Conditions:
- Primary Ab –PDCD4 (CST-9535) 1:1000
- Secondary Ab – Anti Rabbit HRP conjugate (CST- 7074S) 1:2000

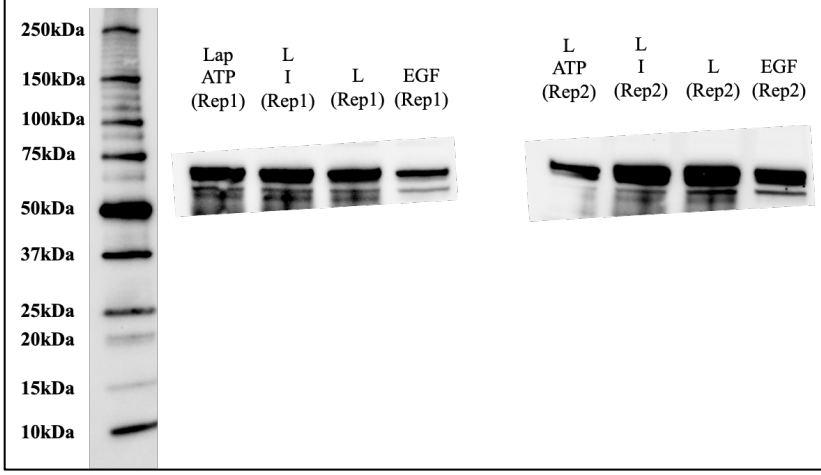
Note: L-lapatinib, I-ipatasertib



Protein: CD82/ P27701, MW = 30-90 kDa, upregulated in L vs EGF and L/I vs EGF (cytoplasmic fraction)
Conditions:

- Primary Ab –CD82 (CST-12439) 1:1000
- Secondary Ab – Anti Rabbit HRP conjugate (CST- 7074S) 1:2000

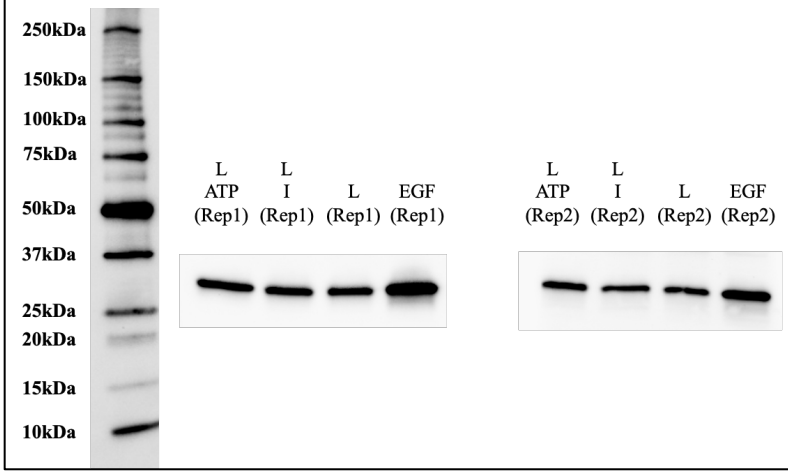
Note: L-lapatinib, I-ipatasertib



Protein: 14-3-3 sigma/ P31947, MW = 28 kDa, downregulated in L vs EGF and L/I vs EGF (cytoplasmic fraction)
Conditions:

- Primary Ab –14-3-3 sigma (RD-AF4424) 1:200
- Secondary Ab – Anti Goat HRP conjugate (RD- HAF017) 1:1000

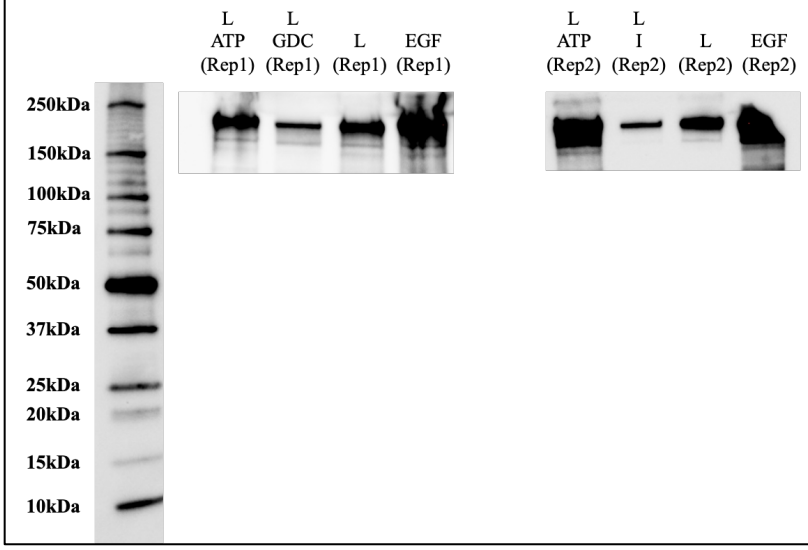
Note: L-lapatinib, I-ipatasertib



Protein: TOP2A/ P11388, MW = 190 kDa, downregulated in L vs EGF and L/I vs EGF (nuclear fraction)
Conditions:

- Primary Ab –TOP2A (CST-12286) 1:1000
- Secondary Ab – Anti Rabbit HRP conjugate (CST- 7074S) 1:2000

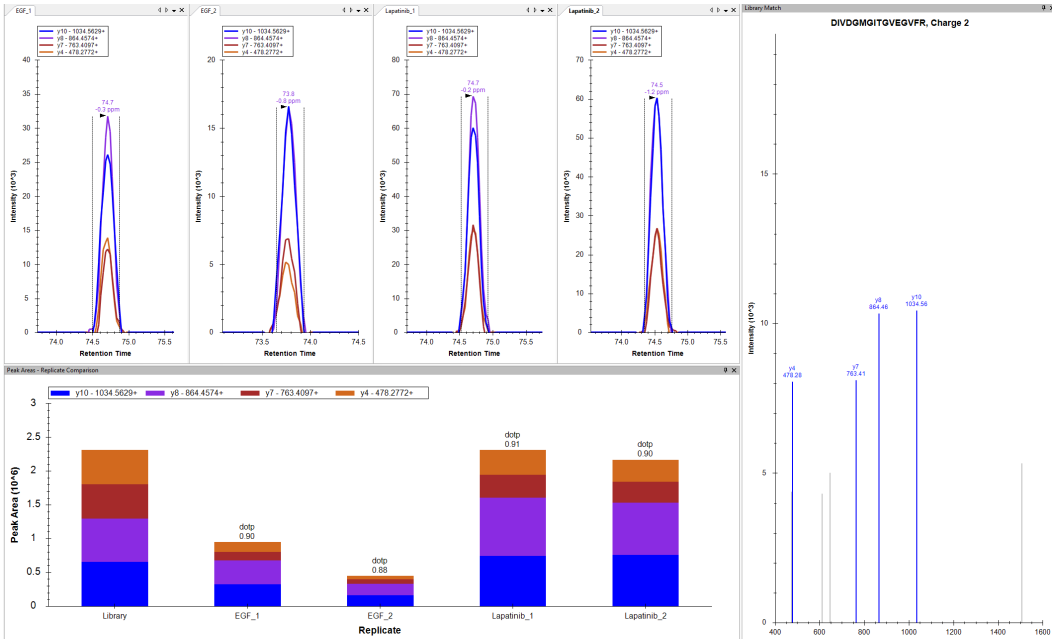
Note: L-lapatinib, I-ipatasertib



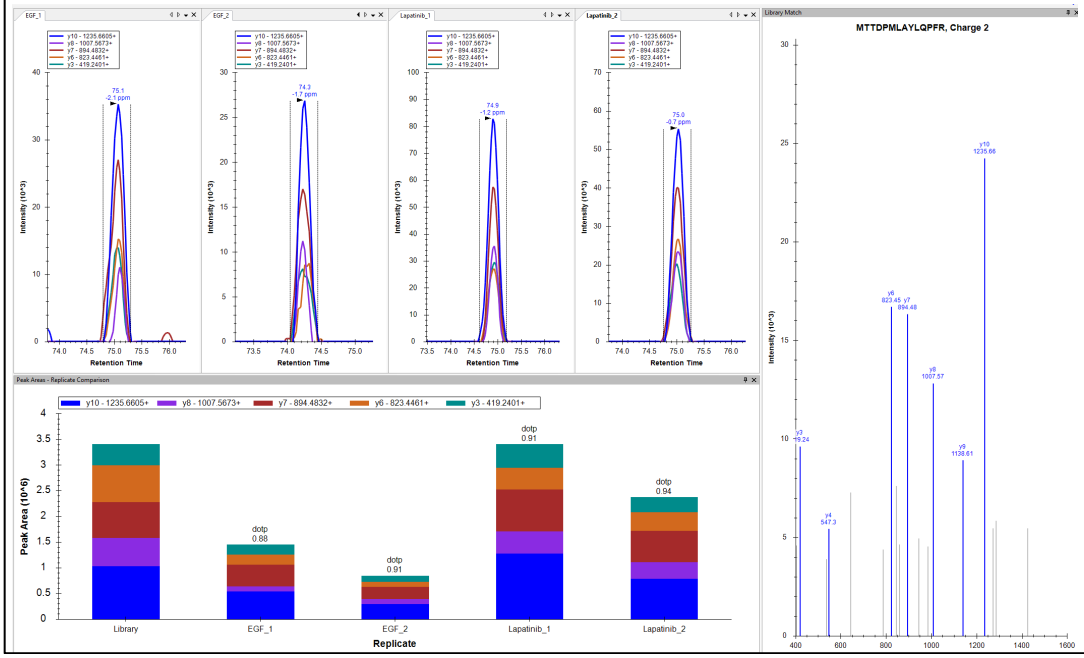
Appendix I. Differentially expressed proteins validated by PRM. The experimental procedure was described in **Section 3.9**.

Upregulated in Lapatinib

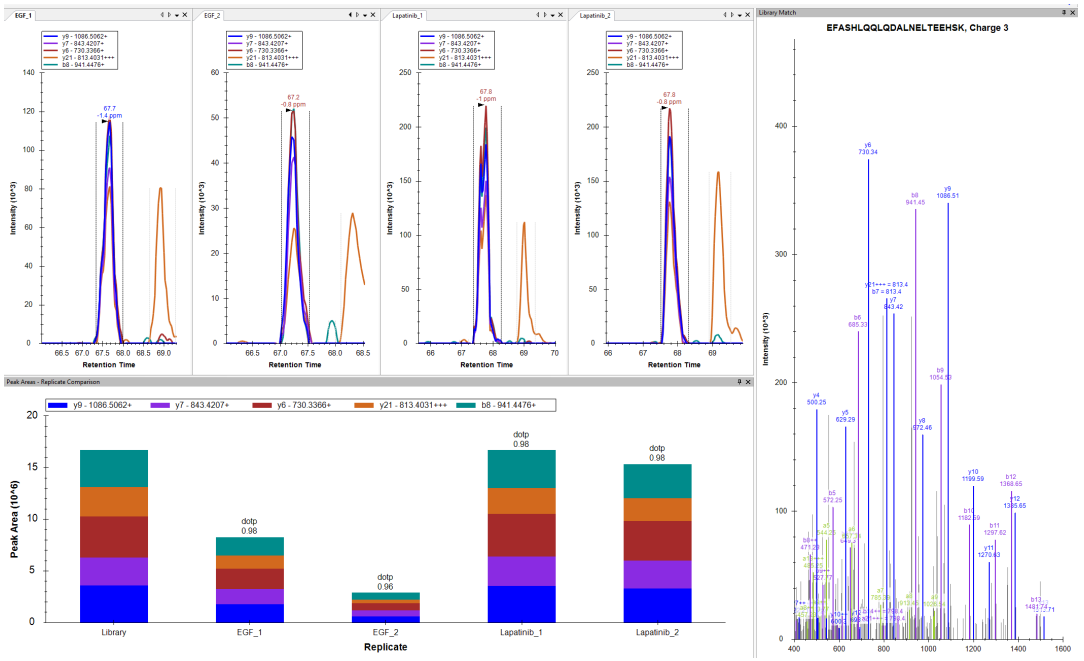
Serine-protein kinase ATM (ATM)
DIVDGMGITGVEGVFR, Charge +2, m/z = 832.9193 Da



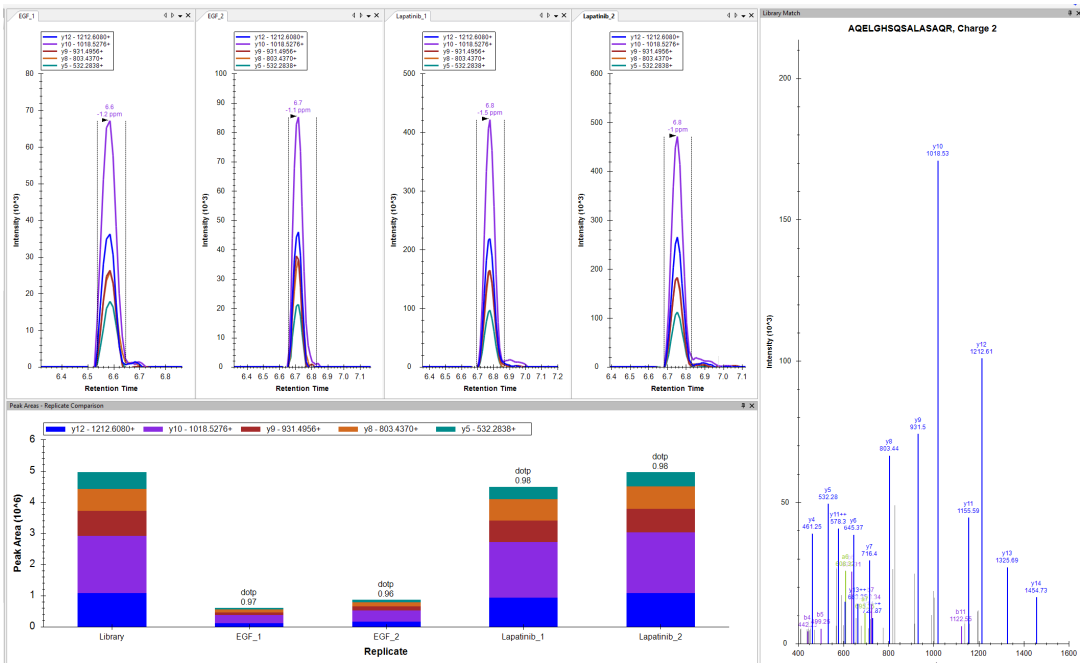
Serine-protein kinase ATM (ATM)
MTTDPMLAYLQPFR, Charge +2, m/z = 842.4153 Da



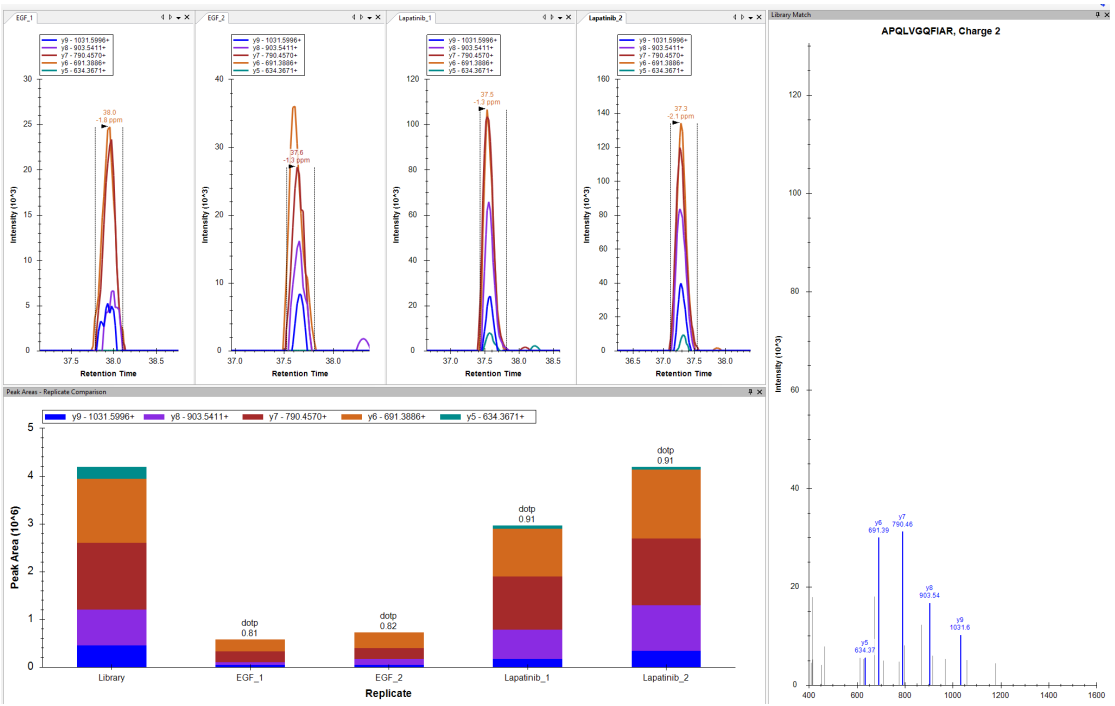
Nuclear mitotic apparatus protein 1 (NUMA1)
EFASHLQQLQDALNELTEEHSK, Charge +3, m/z = 856.4173 Da



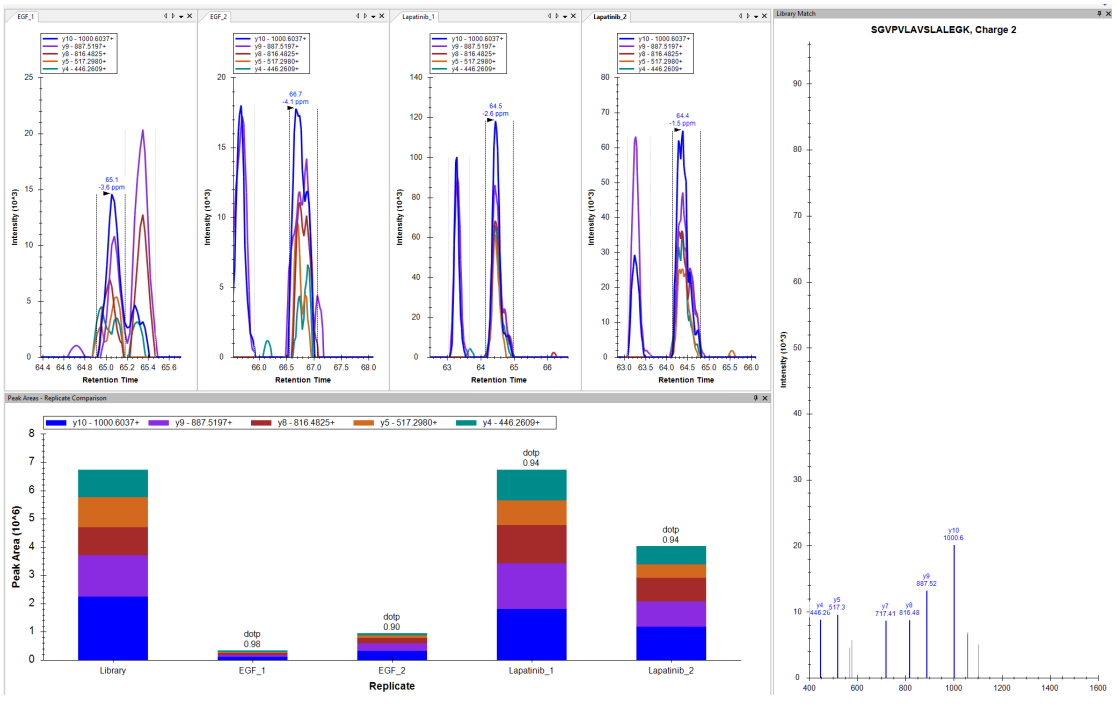
Nuclear mitotic apparatus protein 1 (NUMA1)
AQELGHSQSALASAQR, Charge +2, m/z = 827.4188 Da



Programmed cell death protein 4 (PDCD4)
APQLVGQFIAR, Charge +2, m/z = 600.3484 Da



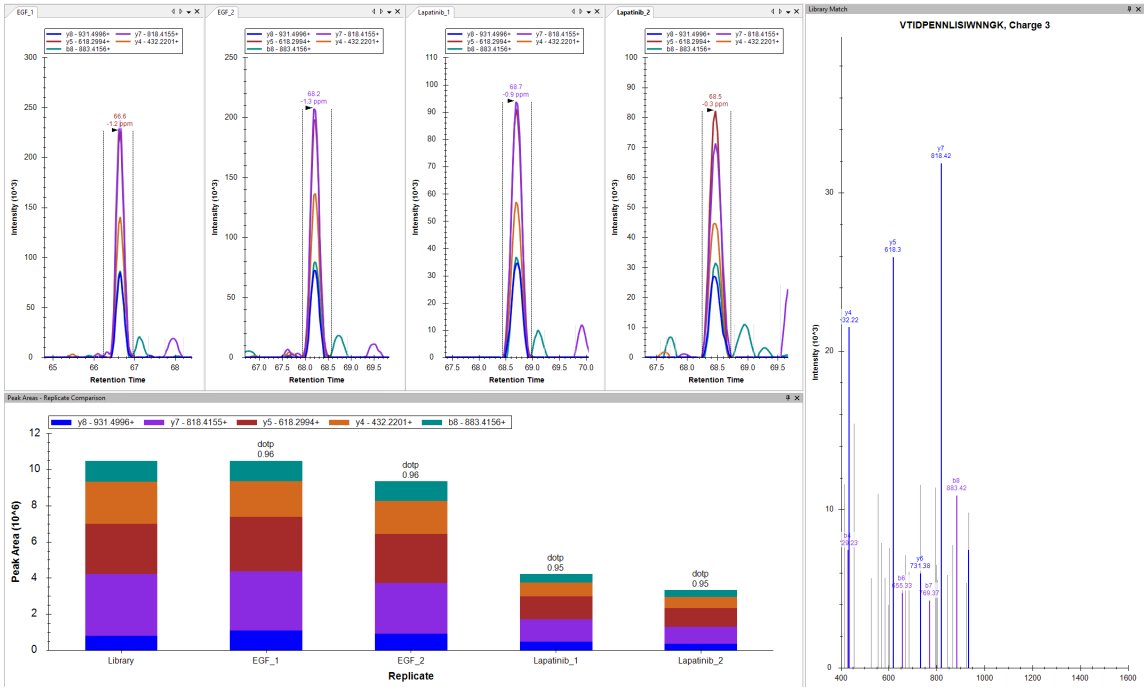
Programmed cell death protein 4 (PDCD4)
SGVPVLAVSLALEGK, Charge +2, m/z = 720.4270 Da



Downregulated in Lapatinib

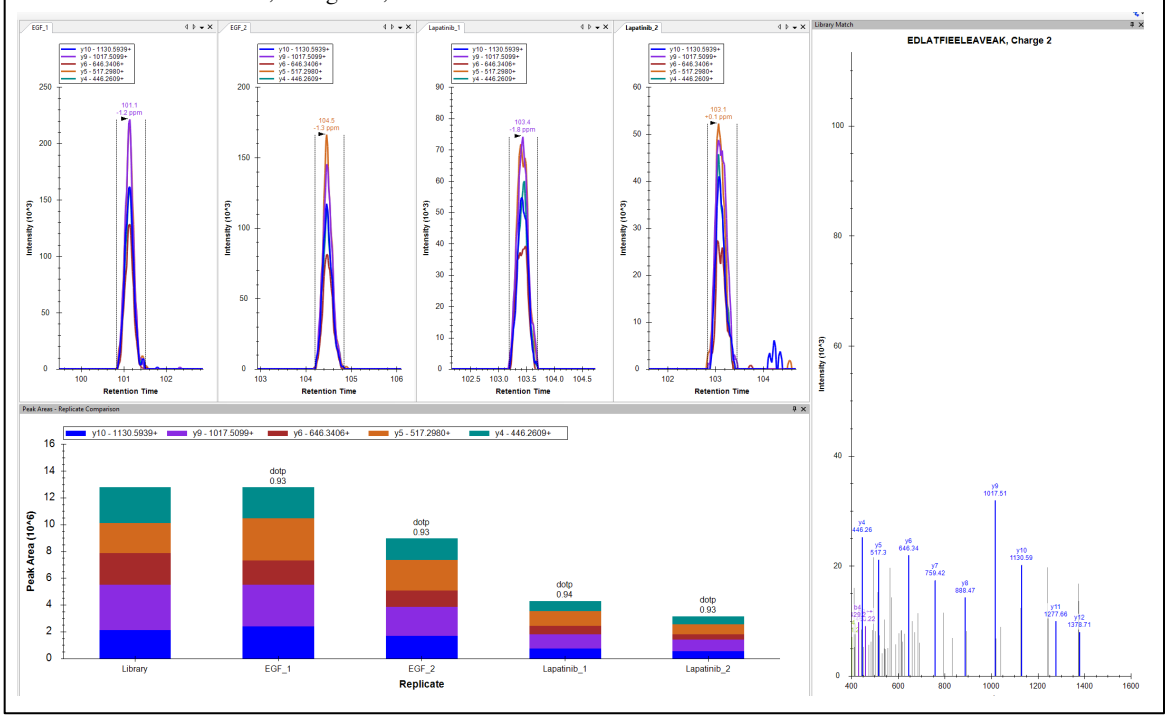
DNA topoisomerase 2-alpha (TOP2A)

VTIDPENNLISIWNNGK, Charge +3, m/z = 643.0022 Da

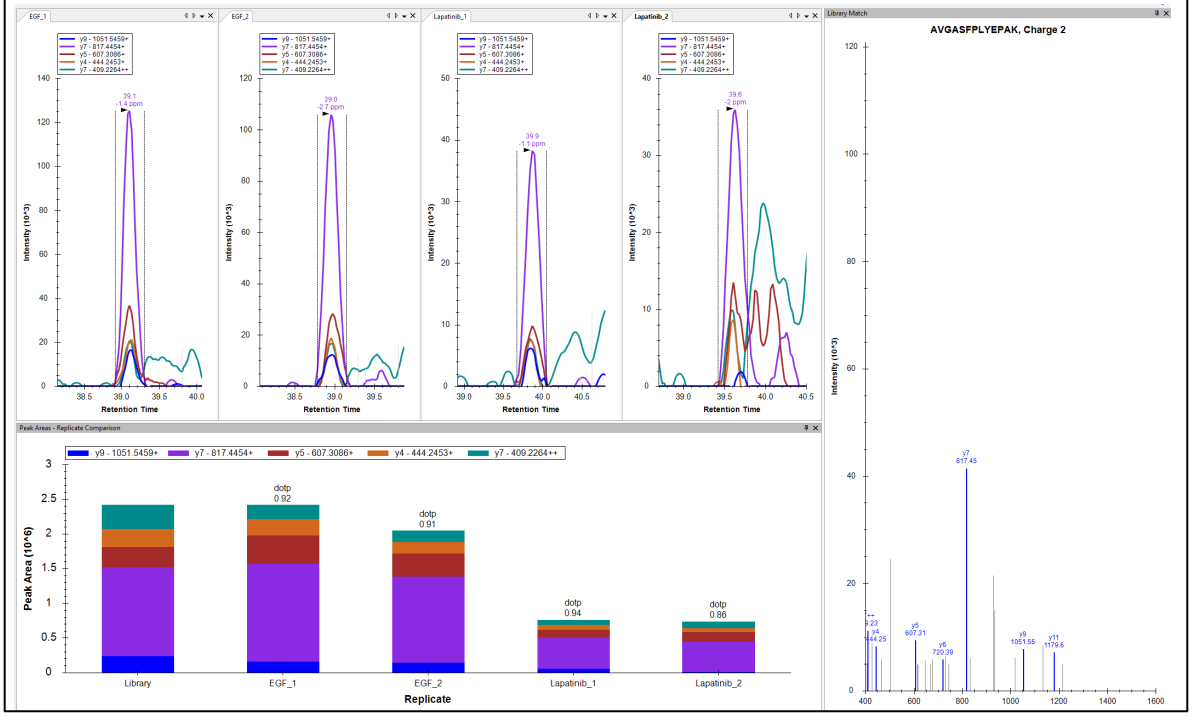


DNA topoisomerase 2-alpha (TOP2A)

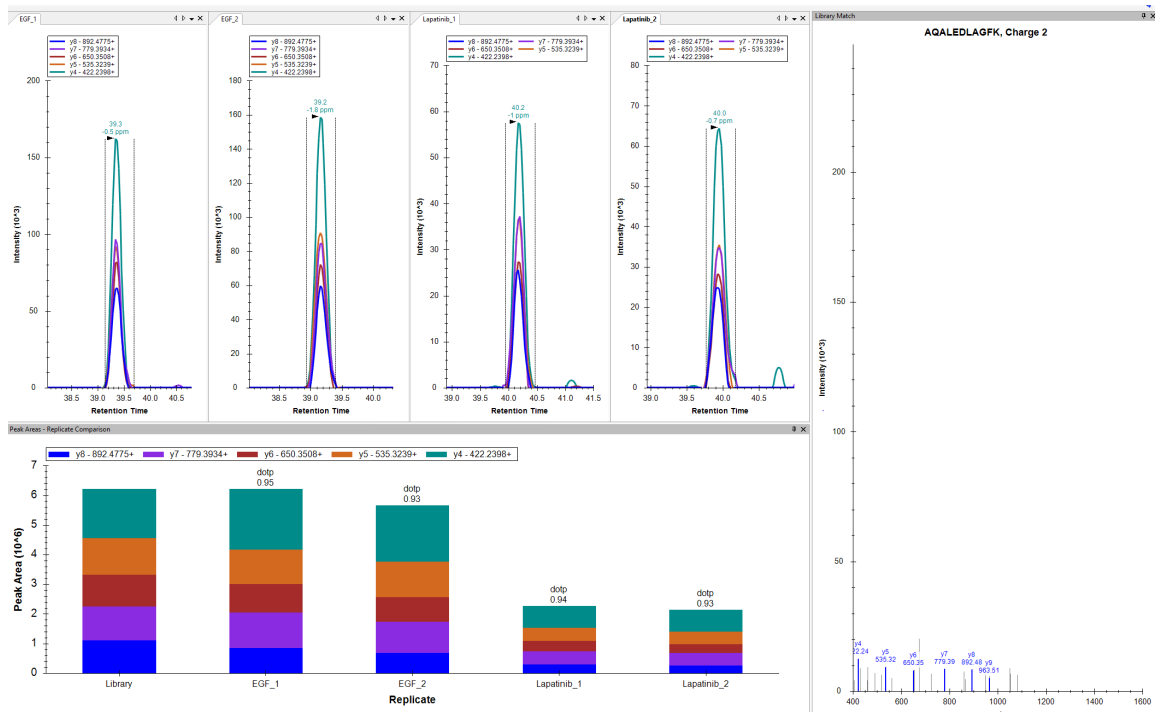
EDLATFIEELEAVEAK, Charge +2, m/z = 903.9540 Da



Proliferation marker protein Ki-67 (MKI67)
AVGASFPLYEPAK, Charge +2, m/z = 675.3586 Da

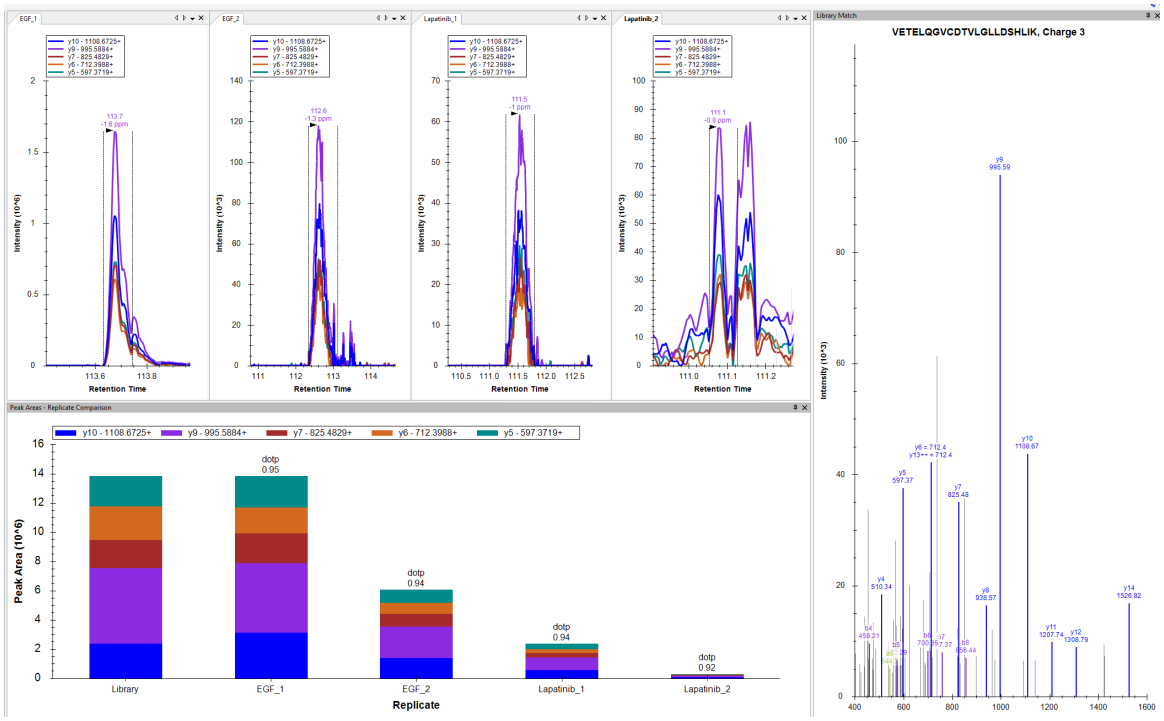


Proliferation marker protein Ki-67 (MKI67)
AQALEDLAGFK, Charge +2, m/z = 581.8088 Da



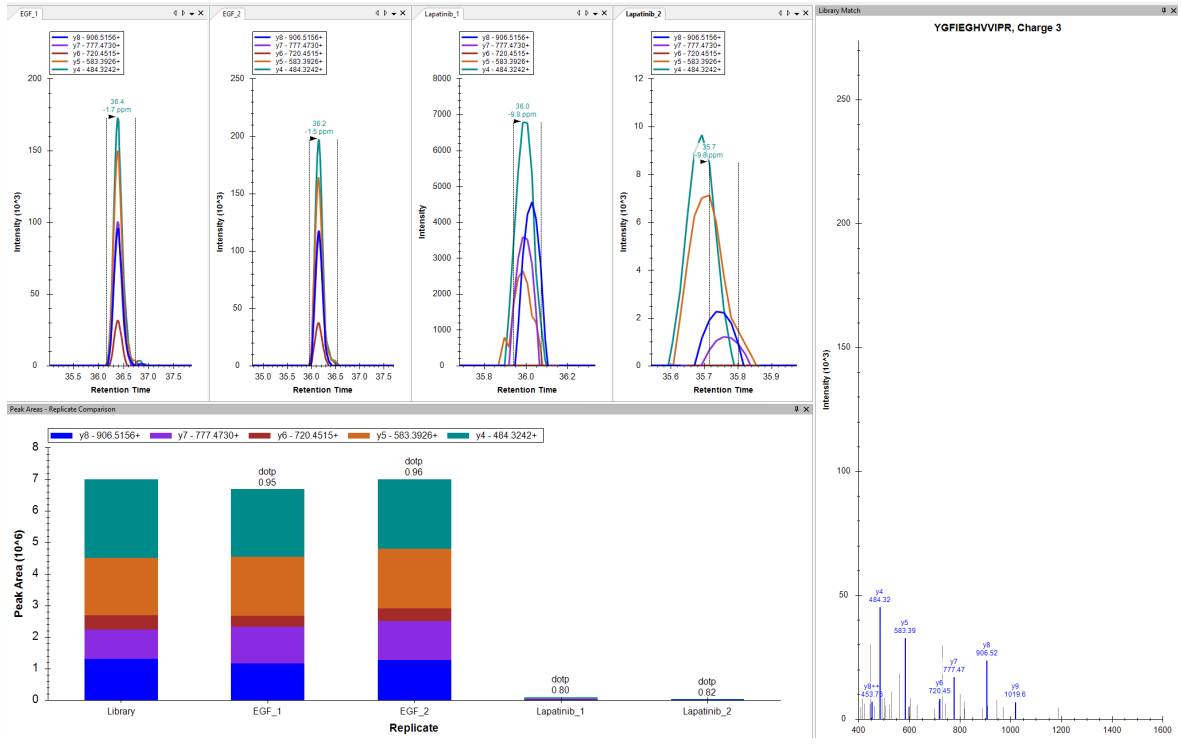
14-3-3 protein sigma (SFN)

VETELQGVCDTVLGLLDSHLIK, Charge +3, m/z = 794.7577 Da



CD44 antigen (CD44)

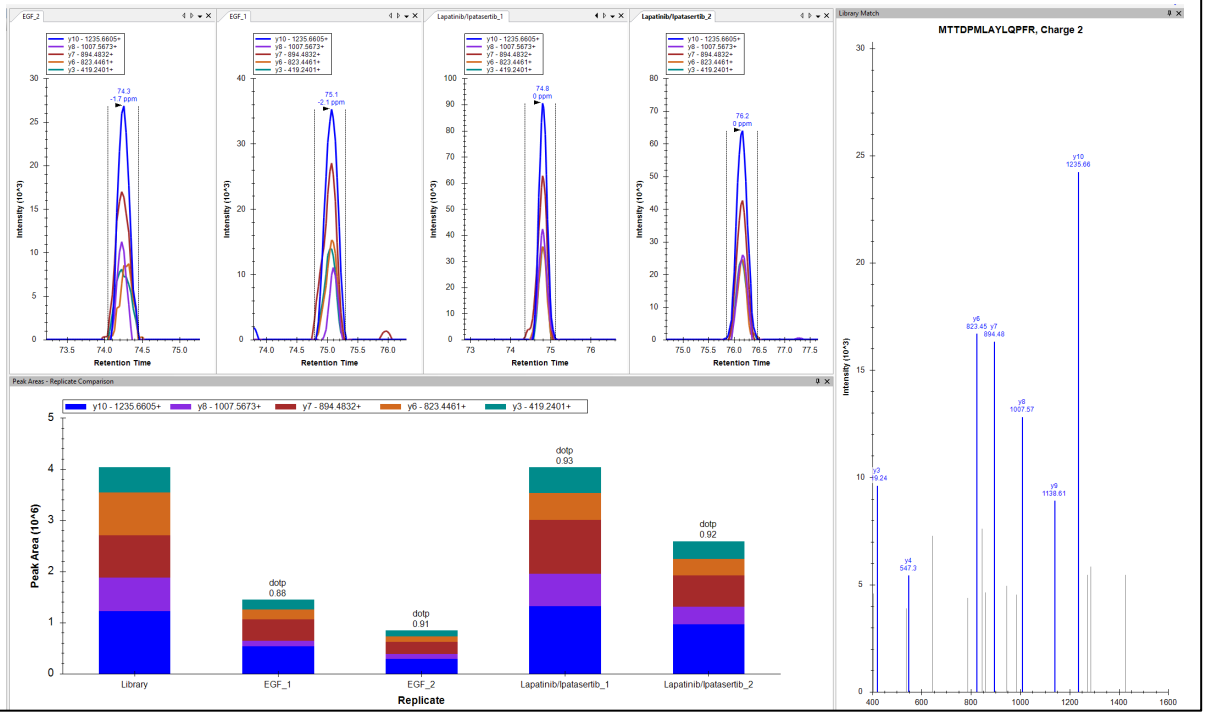
YGFIEGHVVIPR, Charge +3, m/z = 462.9225 Da



Upregulated in Lapatinib/Ipatasertib

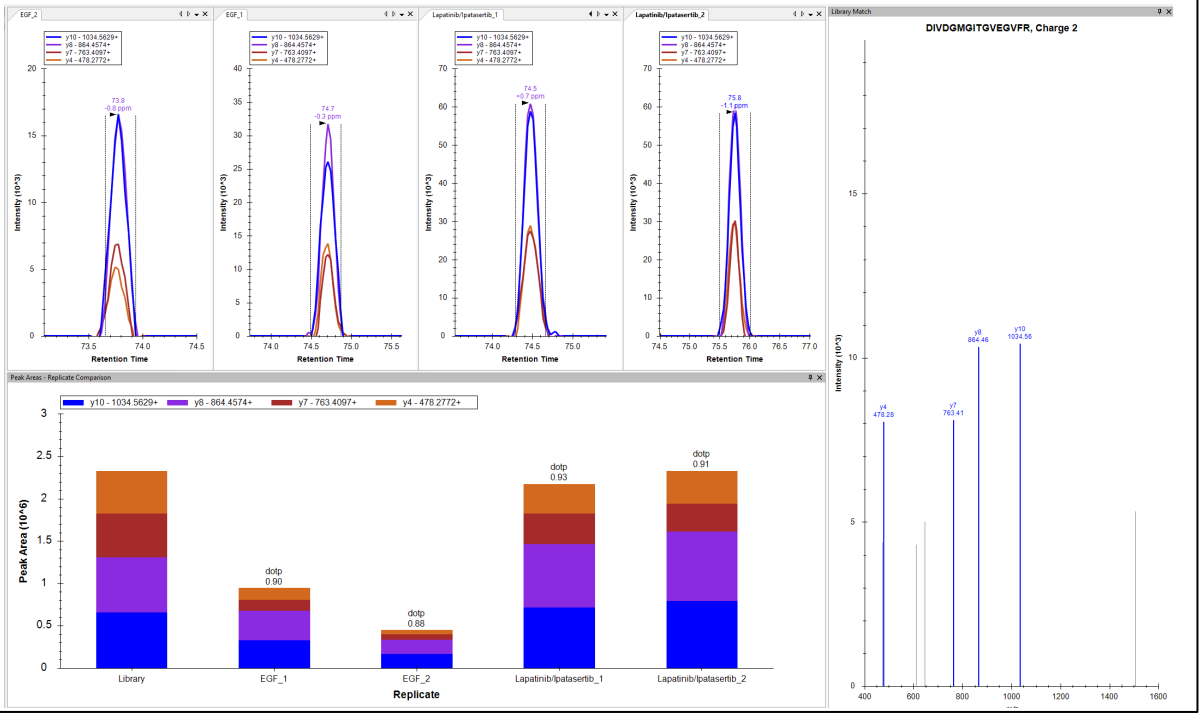
Serine-protein kinase ATM (ATM)

MTTDPMLAYLQPFR, Charge +2, m/z = 842.4153 Da

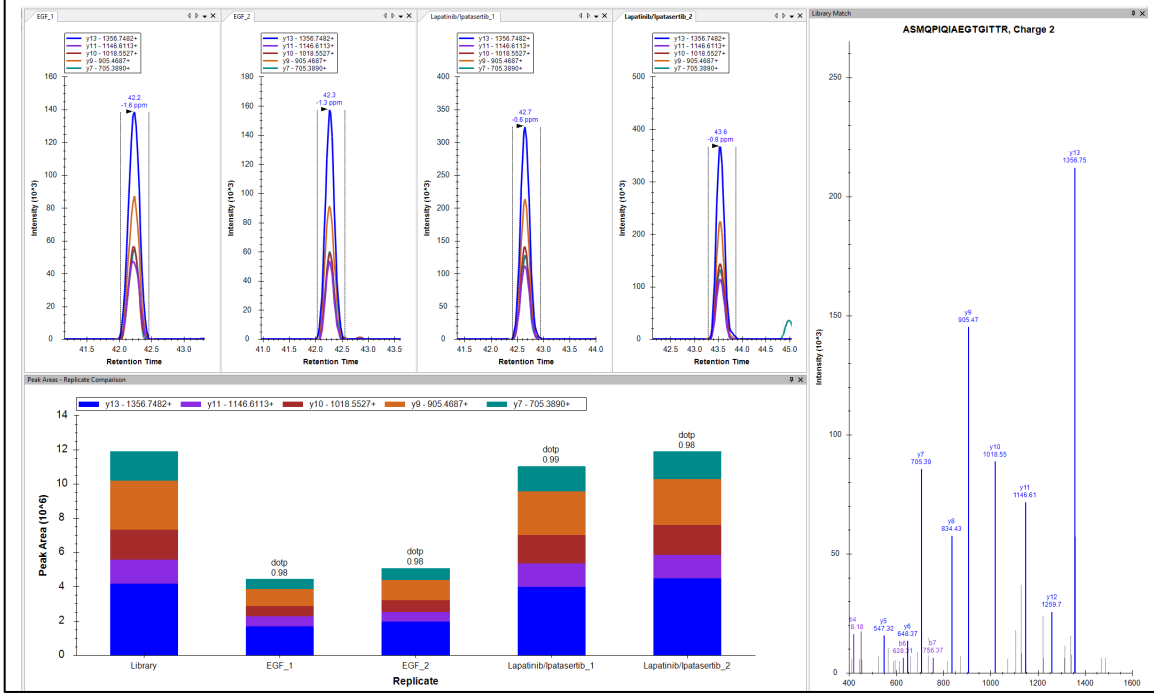


Serine-protein kinase ATM (ATM)

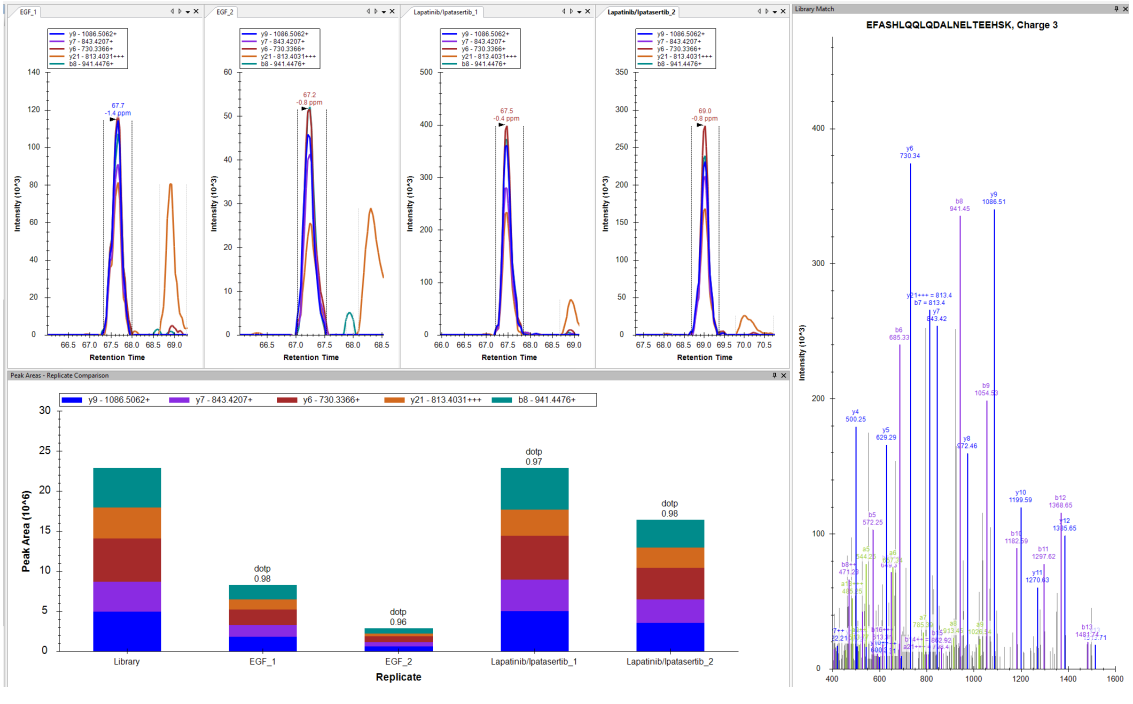
DIVDGMGITGVEGVFR, Charge +2, m/z = 832.9193 Da



Nuclear mitotic apparatus protein 1 (NUMA1)
ASMQPIQIAEGTGITTR, Charge +2, m/z = 887.4618 Da

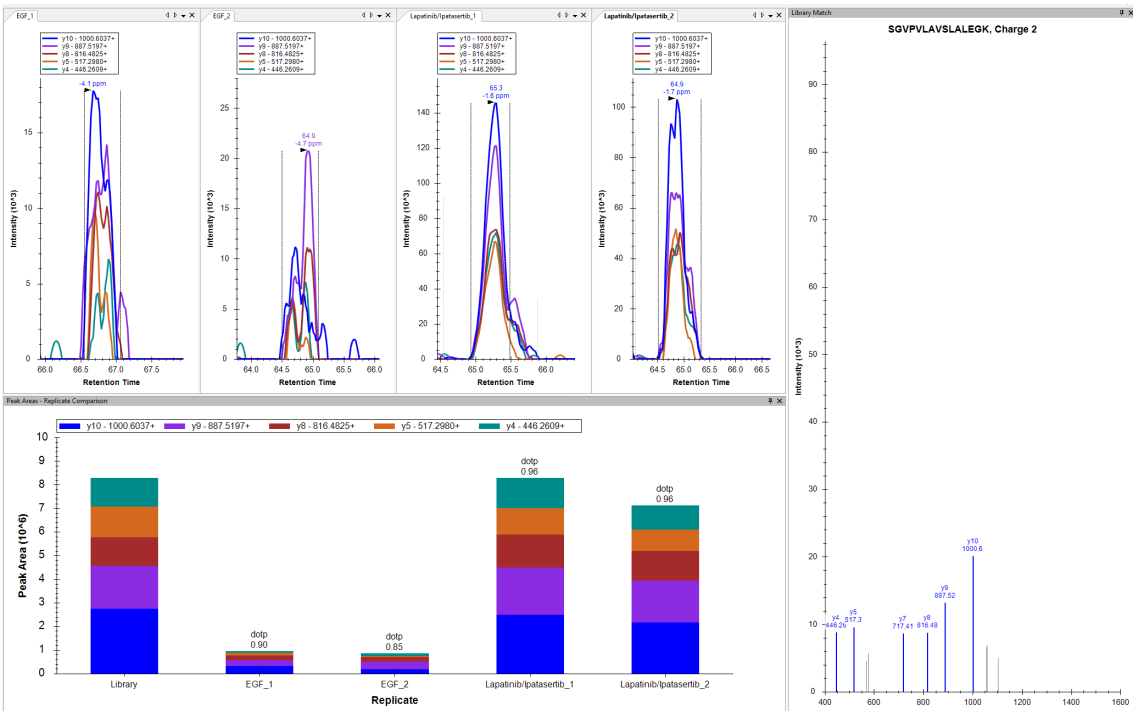


Nuclear mitotic apparatus protein 1 (NUMA1)
EFASHLQQLQDALNELTEEHSK, Charge +3, m/z = 856.4173 Da



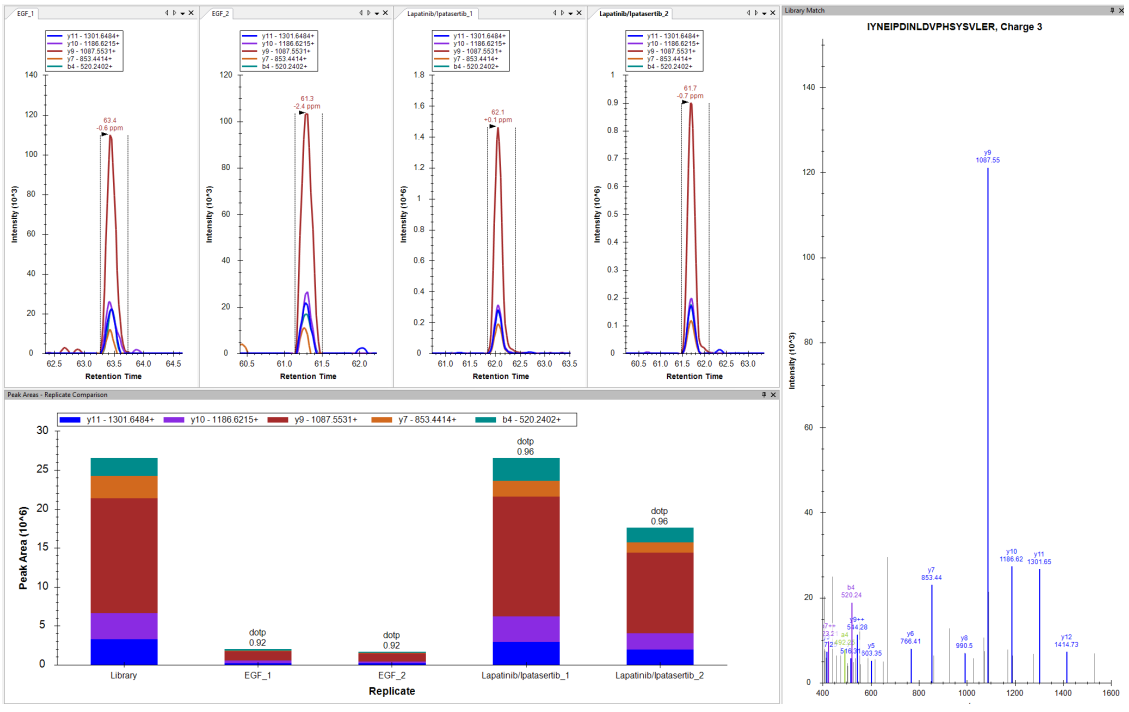
Programmed cell death protein 4 (PDCD4)

SGVPVLAVSLALEGK, Charge +2, m/z = 720.4270 Da



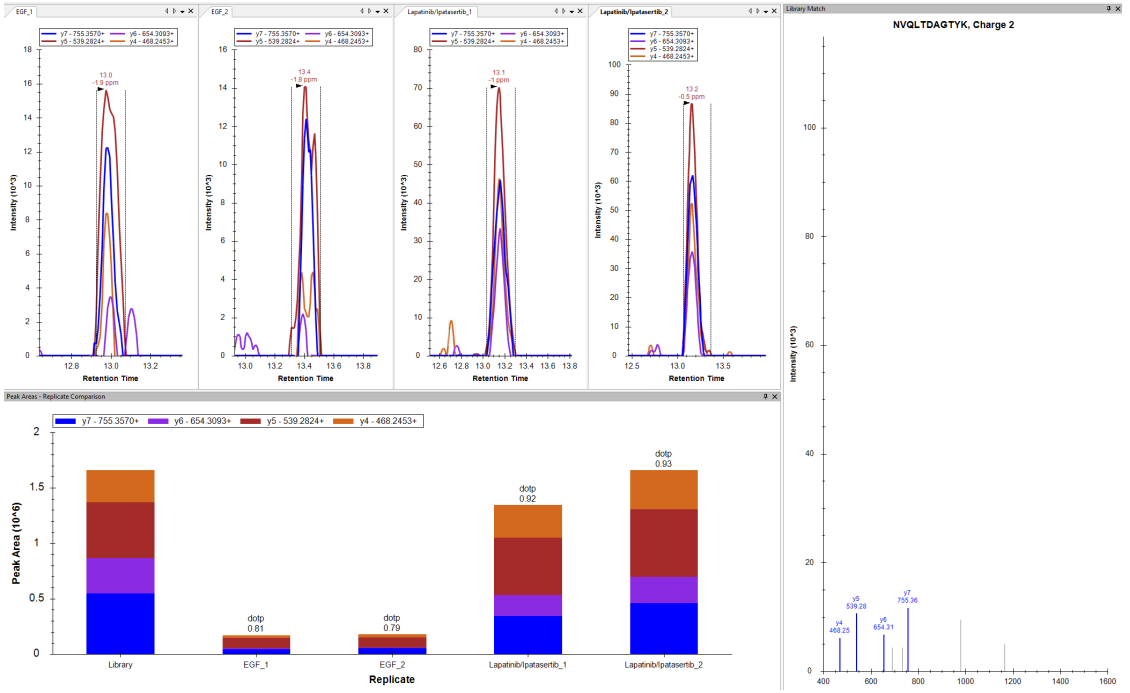
Programmed cell death protein 4 (PDCD4)

IYNEIPDINLDVPHSYSVLER, Charge +3, m/z = 829.4236 Da



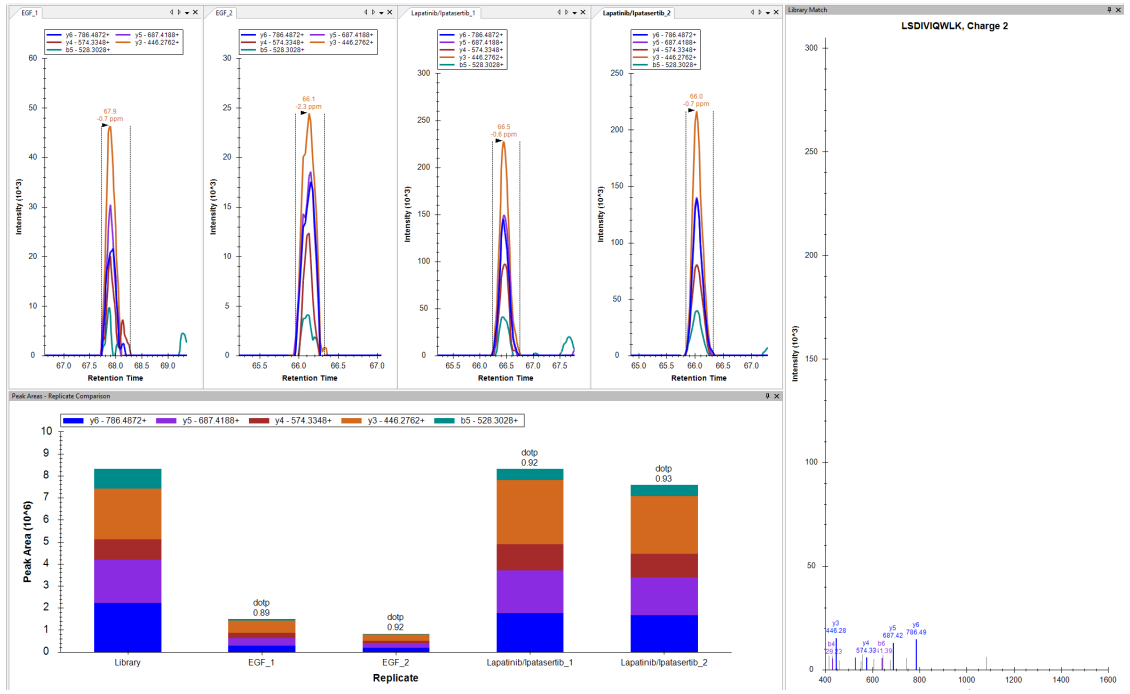
V-set domain-containing T-cell activation inhibitor 1 (VTCN1)

NVQLTDAGTYK, Charge +2, m/z = 605.3091 Da



V-set domain-containing T-cell activation inhibitor 1 (VTCN1)

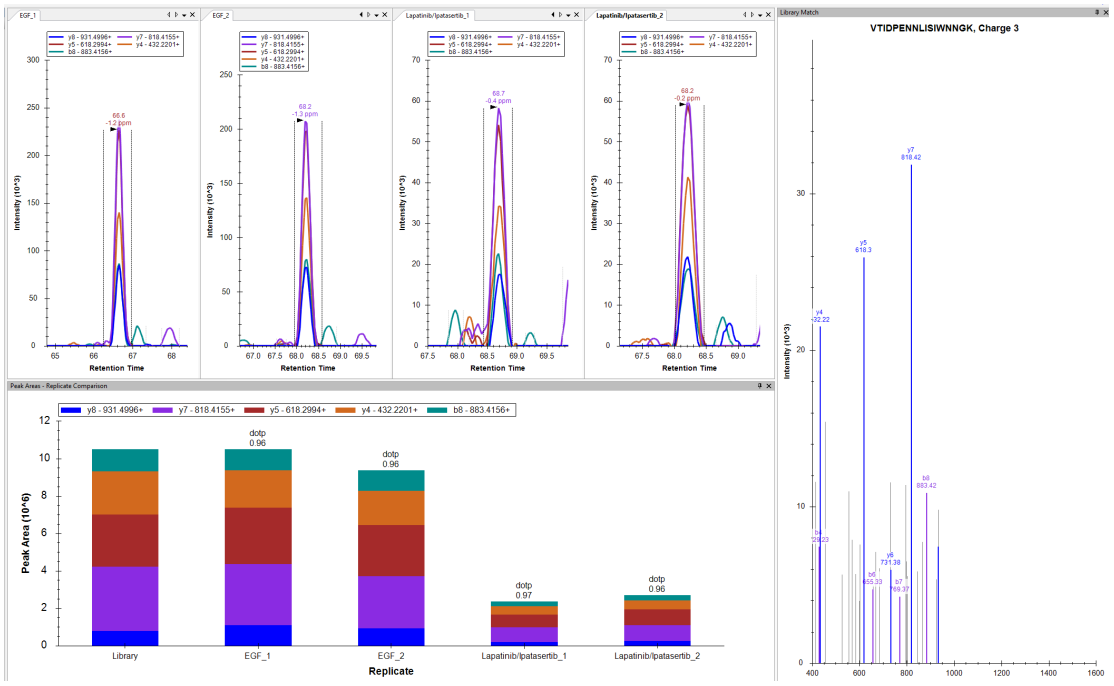
LSDIVIQWLK, Charge +2, m/z = 607.8608 Da



Downregulated in Lapatinib/Ipatasertib

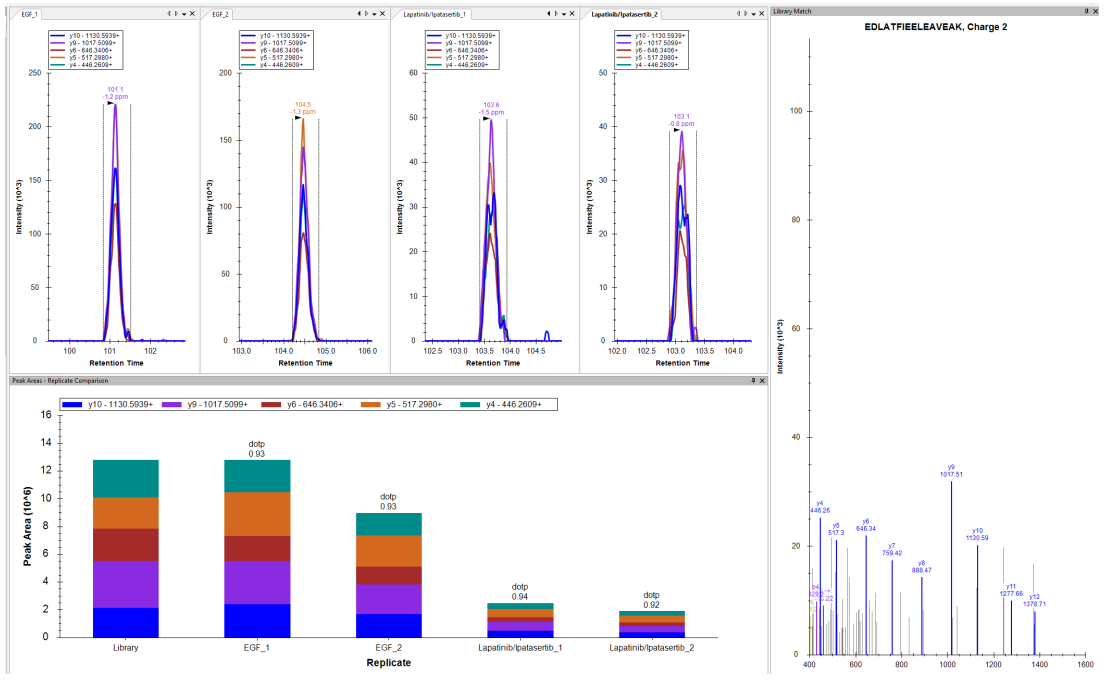
DNA topoisomerase 2-alpha (TOP2A)

VTIDPENNLISIWNNGK, Charge +3, m/z = 643.0022 Da



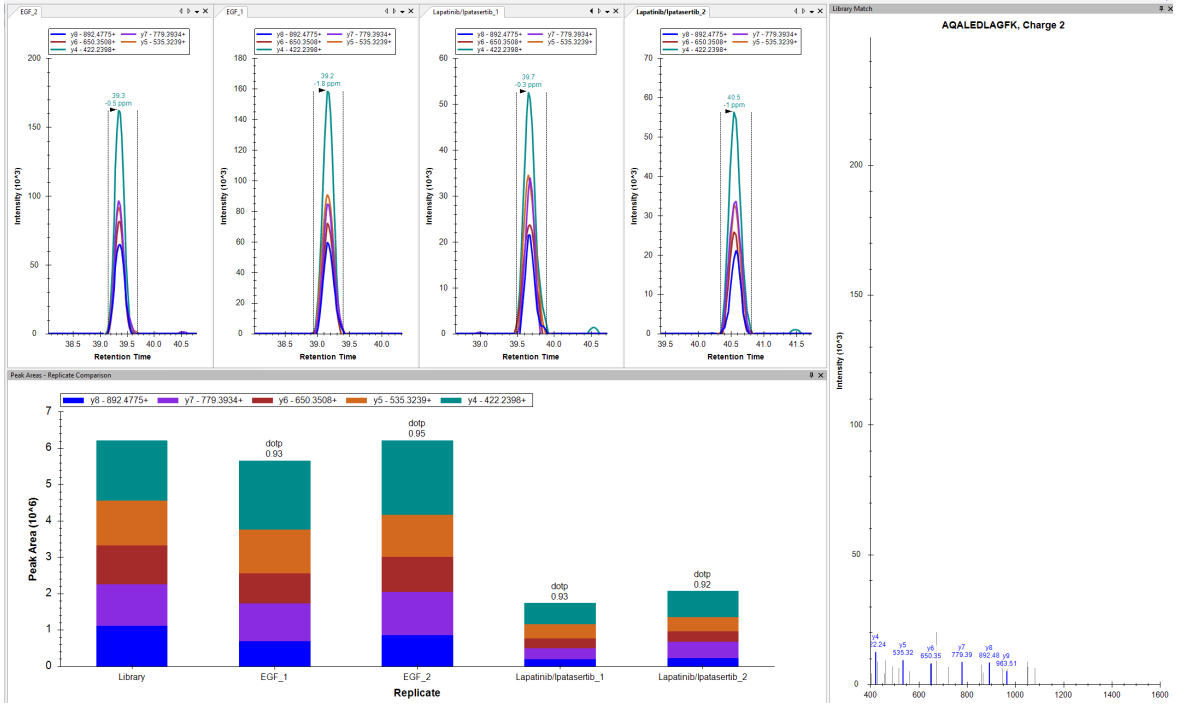
DNA topoisomerase 2-alpha (TOP2A)

EDLATFIEELEAVEAK, Charge +2, m/z = 903.9540 Da



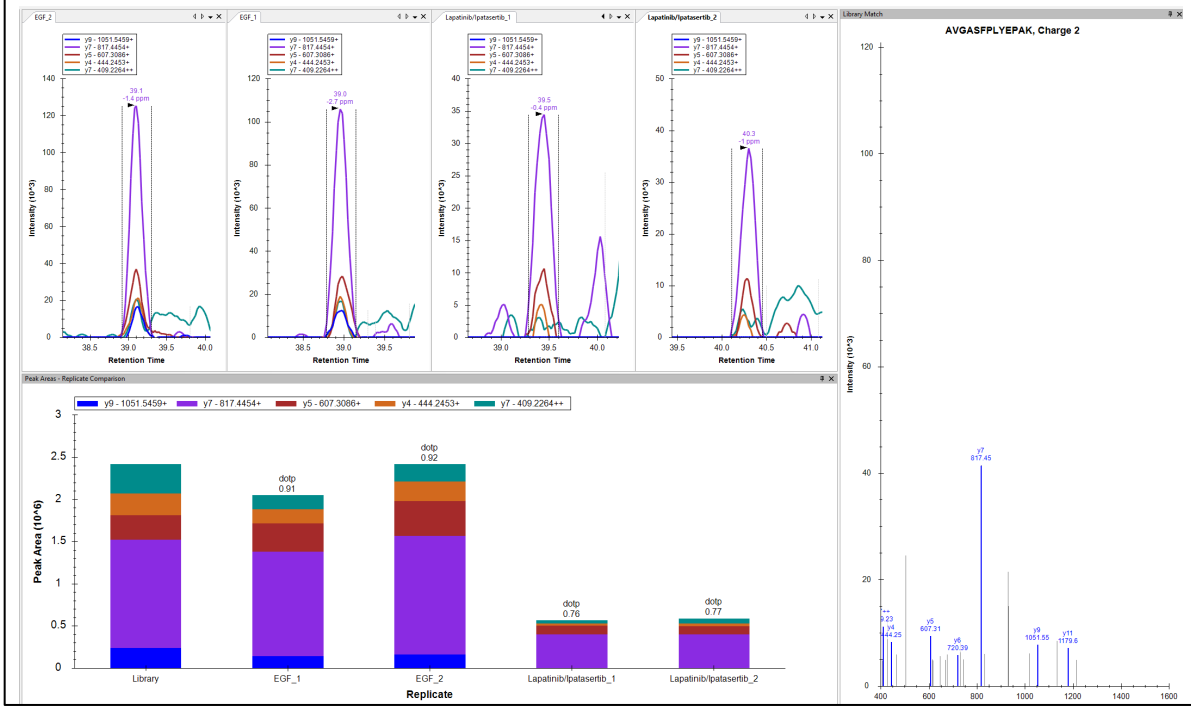
Proliferation marker protein Ki-67 (MKI67)

AQALEDLAGFK, Charge +2, m/z = 581.8088 Da



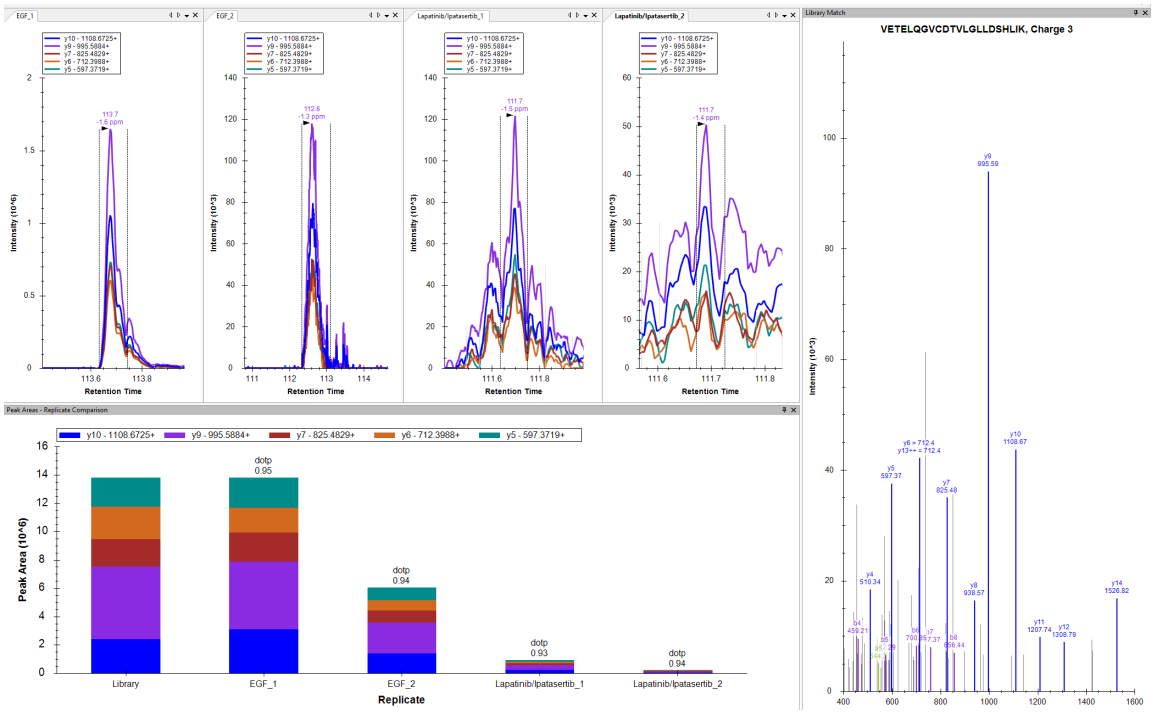
Proliferation marker protein Ki-67 (MKI67)

AVGASFPLYEPAK, Charge +2, m/z = 675.3586 Da



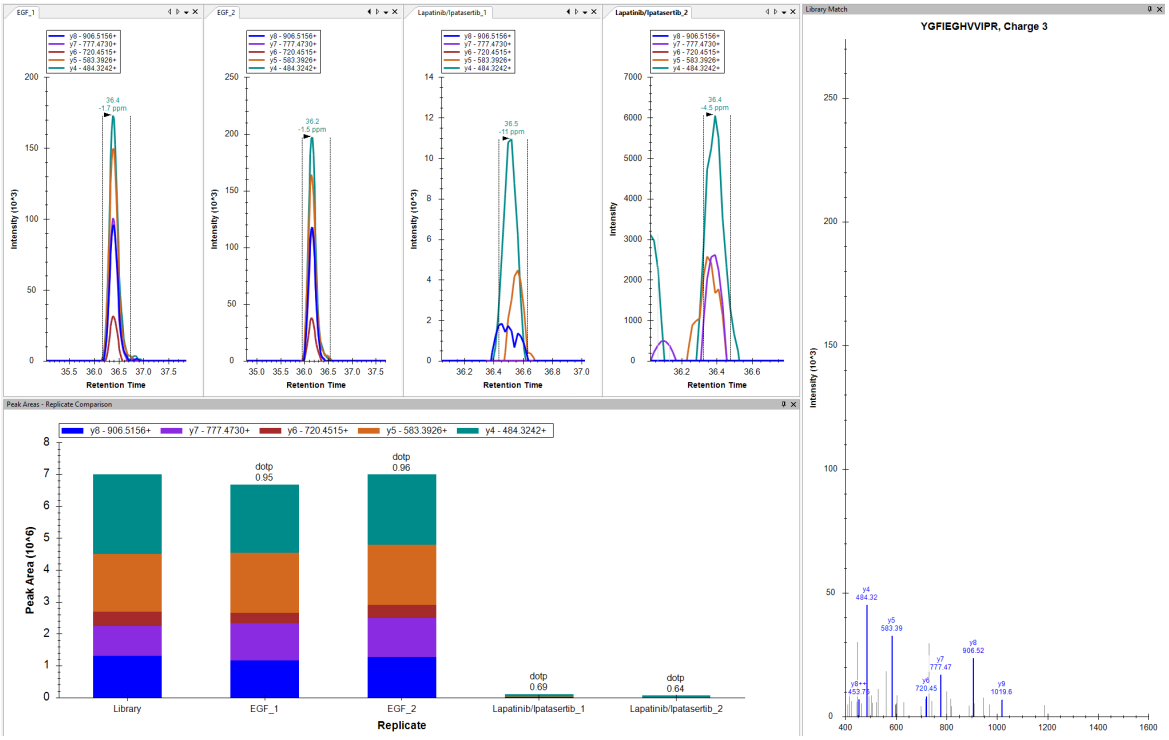
14-3-3 protein sigma (SFN)

VETELQGVCDTVLGLLDSHLIK, Charge +3, m/z = 794.7577 Da

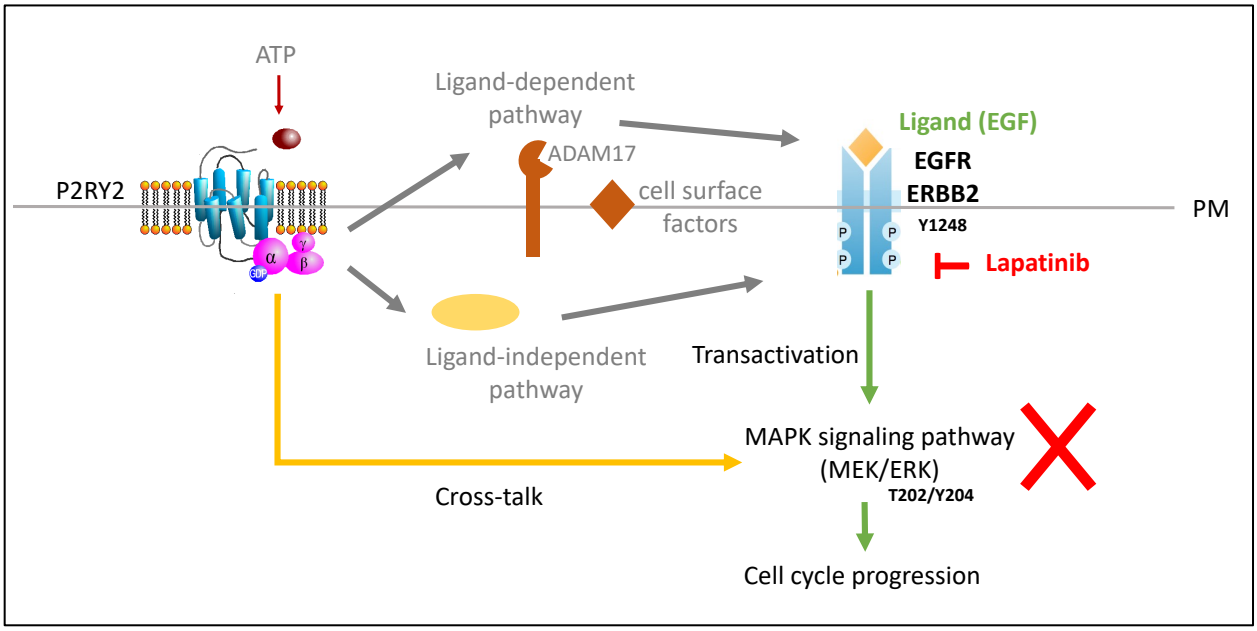


CD44 antigen (CD44)

YGFIEGHVVIPR, Charge +3, m/z = 462.9225 Da



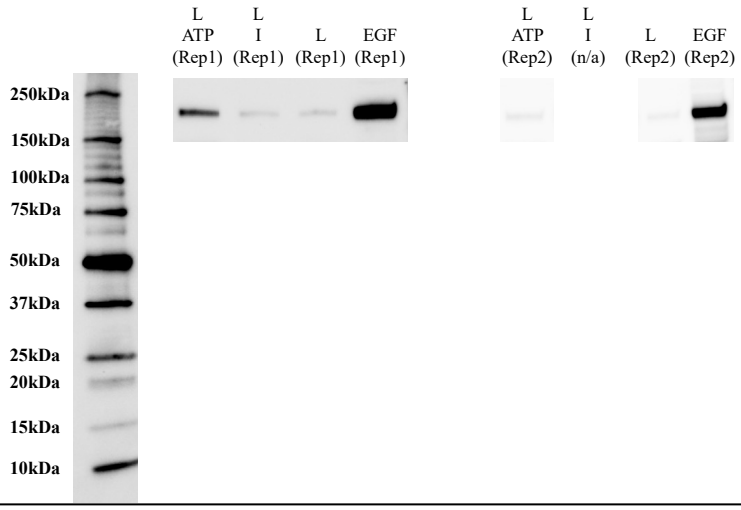
Appendix J. Transactivation testing. Schematic of the proposed transactivation activity and the proteins that were tested by western blot for their expression levels upon stimulation with EGF and lapatinib. Transactivation would be expected if ATP addition led to the expression of ERBB2 and ERK1/2 proteins. The experimental procedure was described in **Section 3.15**.



Protein: ERBB2/ P04626, MW = 185 kDa, upregulated in EGF (15 m full cell lysate fraction)
Conditions:

- Primary Ab – p-ERBB2 (Y1248) (CST-2247) 1:1000
- Secondary Ab – Anti Rabbit HRP conjugate (CST- 7074S) 1:2000

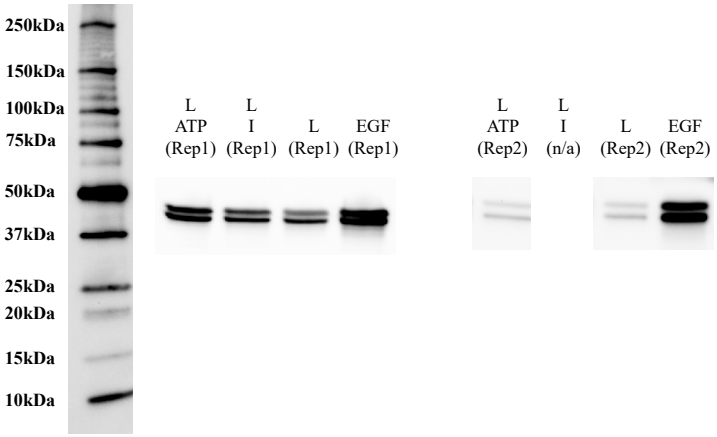
Note: L-lapatinib, I-ipatasertib



Protein: ERK1,2/ P28482, P27361 MW = 42, 44 kDa, upregulated in EGF (15 m full cell lysate fraction)
Conditions:

- Primary Ab –p-ERK (T202/Y204) (CST-9101) 1:1000
- Secondary Ab – Anti Rabbit HRP conjugate (CST- 7074S) 1:2000

Note: L-lapatinib, I-ipatasertib



Appendix K. Protein List (UniProt description)

ABCC1 - Multidrug resistance-associated protein 1
ABCC10 - ATP-binding cassette sub-family C member 10
ABCC2 - ATP-binding cassette sub-family C member 2
ABCC3 - ATP-binding cassette sub-family C member 3
ABCC5 - ATP-binding cassette sub-family C member 5
ACADSB - Short/branched chain specific acyl-CoA dehydrogenase, mitochondrial
ACVR1 - Activin receptor type-1
ACVR1B - Activin receptor type-1B
ADAM10 - Disintegrin and metalloproteinase domain-containing protein 10
ADAM15 - Disintegrin and metalloproteinase domain-containing protein 15
ADAM17 - Disintegrin and metalloproteinase domain-containing protein 17
ADGRE5/ CD97 - Adhesion G protein-coupled receptor E5
ADGRG1/ GPR56 - Adhesion G-protein coupled receptor G1
AGRN - Agrin
ALCAM/ CD166 - CD166 antigen
ALDH1A3 - Aldehyde dehydrogenase family 1 member A3
ANLN - Anillin
ANO1 - Anoctamin-1
ANXA1 - Annexin A1
ANXA5 - Annexin A5
ATM - Serine-protein kinase ATM
ATP1A1 - Sodium/potassium-transporting ATPase subunit alpha-1
ATP1A3 - Sodium/potassium-transporting ATPase subunit alpha-3
ATP5F1B - ATP synthase subunit beta, mitochondrial
ATP5A1/ATP5F1A - ATP synthase subunit alpha, mitochondrial
ATP8A1 - Phospholipid-transporting ATPase IA
ATRX - Transcriptional regulator ATRX
AURKA - Aurora kinase A
AURKB - Aurora kinase B
BAG3 - BAG family molecular chaperone regulator 3
BCAM/ CD239 - Basal cell adhesion molecule
BMPR1A - Bone morphogenetic protein receptor type-1A
BMPR2 - Bone morphogenetic protein receptor type-2
BZW1 - eIF5-mimic protein 2
CA12 - Carbonic anhydrase 12
CALM3 - Calmodulin-3
CALR - Calreticulin
CCNB1 - G2/mitotic-specific cyclin-B1
CD276 - CD276 antigen
CD44 - CD44 antigen
CD9 - CD9 antigen
CDH5 - Cadherin-5
CDK1 - Cyclin-dependent kinase 1
CDK2 - Cyclin-dependent kinase 2

CEACAM6 - Carcinoembryonic antigen-related cell adhesion molecule 6
CHMP4B - Charged multivesicular body protein 4b
CNN2 - Calponin-2
CNN3 - Calponin-3
COTL1 - Coactosin-like protein
COX5A - Cytochrome c oxidase subunit 5A, mitochondrial
CSF1R - Macrophage colony-stimulating factor 1 receptor
CSK - Tyrosine-protein kinase CSK
CTSL - CTD small phosphatase-like protein 2
CXADR - Coxsackievirus and adenovirus receptor
DAP3/MRPS29 - 28S ribosomal protein S29, mitochondrial
DAPK1 - Death-associated protein kinase 1
DBN1 - Drebrin
DCK - Deoxycytidine kinase
DIAPH2 - Protein diaphanous homolog 2
DSG2 - Desmoglein-2
DUSP3 - Dual specificity protein phosphatase 3
ECT2 - Protein ECT2
EIF4G2 - Eukaryotic translation initiation factor 4 gamma 2
EIF4G3 - Eukaryotic translation initiation factor 4 gamma 3
EPCAM/ CD326 - Epithelial cell adhesion molecule
ERBB1/ EGFR - Epidermal growth factor receptor
ERBB2/ HER2 - Receptor tyrosine-protein kinase erbB-2
ERBB3 - Receptor tyrosine-protein kinase erbB-3
ERBB4 - Receptor tyrosine-protein kinase erbB-4
F11R - Junctional adhesion molecule A
FANCM - Fanconi anemia group M protein
FGFR1 - Fibroblast growth factor receptor 1
FGFR2 - Fibroblast growth factor receptor 2
FGFR3 - Fibroblast growth factor receptor 3
FGFR4 - Fibroblast growth factor receptor 4
FLNA - Filamin-A
FN1 - Fibronectin
FOXA1 - Hepatocyte nuclear factor 3-alpha
FZD1 - Frizzled-1
GAA - Lysosomal alpha-glucosidase
GART - Trifunctional purine biosynthetic protein adenosine-3
GDI1 - Rab GDP dissociation inhibitor alpha
GNAS - Neuroendocrine secretory protein 55
GRIA3 - Glutamate receptor 3
GTF3C2 - General transcription factor 3C polypeptide 2
H1F0 - Histone H1.0
HNRNPH2 - Heterogeneous nuclear ribonucleoprotein H2
HSPB1 - Heat shock protein beta-1
HSPH1 - Heat shock protein 105 kDa
IGF1R - Insulin-like growth factor 1 receptor

ITGA2 - Integrin alpha-2
ITGAV/ CD51 - Integrin alpha-V
ITGB1/ CD29 - Integrin beta-1
ITGB2/ CD18 - Integrin beta-2
ITGB5 - Integrin beta-5
IVL - Involucrin
Ki67 - Proliferation marker protein Ki-67
KIF22 - Kinesin-like protein KIF22
KIF23 - Kinesin-like protein KIF23
KIF2C - Kinesin-like protein KIF2C
KIFC1 - Kinesin-like protein KIFC1
L1CAM/ CD171 - Neural cell adhesion molecule L1
LAMA1 - Laminin subunit alpha-1
LGR4 - Leucine-rich repeat-containing G-protein coupled receptor 4
LMO7 - LIM domain only protein 7
LRP2 - Low-density lipoprotein receptor-related protein 2
LRRC15 - Leucine-rich repeat-containing protein 15
MAFK - Transcription factor MafK
MAPK1/ ERK2 - Mitogen-activated protein kinase 1
MAPK3/ ERK1 - Mitogen-activated protein kinase 3
MARK2 - Serine/threonine-protein kinase MARK2
MET - Hepatocyte growth factor receptor
MIS12 - Protein MIS12 homolog
MRPL11 - 39S ribosomal protein L11, mitochondrial
MRPS30/mL65 - 39S ribosomal protein S30, mitochondrial
MTOR - Serine/threonine-protein kinase mTOR
MUC1/ CD227 - Mucin-1
MYH9 - Myosin-9
NAMPT - Nicotinamide phosphoribosyltransferase
NDRG1 - Protein NDRG1
NDUFV1 - NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial
NFKB1 - Nuclear factor NF-kappa-B p105 subunit
NOTCH1 - Neurogenic locus notch homolog protein 1
NOTCH2 - Neurogenic locus notch homolog protein 2
NRP1 - Neuropilin-1
NUMA1 - Nuclear mitotic apparatus protein 1
NUSAP1 - Nucleolar and spindle-associated protein 1
OGDH - 2-oxoglutarate dehydrogenase complex component E1
ORAI1 - Calcium release-activated calcium channel protein 1
P2RY2 - P2Y purinoceptor 2
P2RY6 - P2Y purinoceptor 6
PABPC3 - Polyadenylate-binding protein 3
PANX1 - Pannexin-1
PCDH1 - Protocadherin-1
PDCD4 - Programmed cell death protein 4
PDIA3 - Protein disulfide-isomerase A3

PDIA4 - Protein disulfide-isomerase A4
PDIA6 - Protein disulfide-isomerase A6
PFKFB3 - 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3
PIK3CA - Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform
PKD1 - Polycystin-1
PLIN3 - Perilipin-3
PNP - Purine nucleoside phosphorylase
POM121C - Nuclear envelope pore membrane protein POM 121C
PRC1 - Protein regulator of cytokinesis 1
PRMT5 - Protein arginine N-methyltransferase 5
PTPRD - Receptor-type tyrosine-protein phosphatase delta
PVR/ CD155 - Poliovirus receptor
PXN - Paxillin
RAB11A - Ras-related protein Rab-11A
RAB35 - Ras-related protein Rab-35
RAB5B - Ras-related protein Rab-5B
RACGAP1 - Rac GTPase-activating protein 1
RELA - Transcription factor p65
RGPD5 - RANBP2-like and GRIP domain-containing protein 5/6
RPTOR - Regulatory-associated protein of mTOR
RRM2 - Ribonucleoside-diphosphate reductase subunit M2
RSF1 - Remodeling and spacing factor 1
S100A9 - Protein S100-A9
SCYL1 - N-terminal kinase-like protein
SDHA - Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial
SFN - 14-3-3 protein sigma
SH3GL1 - Endophilin-A2
SIGLEC6 - Sialic acid-binding Ig-like lectin 6
SLC16A3 - Monocarboxylate transporter 4
SLC23A2 - Solute carrier family 23 member 2
SLC2A1 - Solute carrier family 2, facilitated glucose transporter member 1
SLC2A3 - Solute carrier family 2, facilitated glucose transporter member 3
SLC39A14 - Metal cation symporter ZIP14
SLC3A2 - 4F2 cell-surface antigen heavy chain
SLC7A5 - Large neutral amino acids transporter small subunit 1
SORL1 - Sortilin-related receptor
SPEN - Msx2-interacting protein
SRC - Proto-oncogene tyrosine-protein kinase Src
STAT3 - Signal transducer and activator of transcription 3
SUGT1 - Protein SGT1 homolog
SUSD2 - Sushi domain-containing protein 2
TAF2 - Transcription initiation factor TFIID subunit 2
TAF5 - Transcription initiation factor TFIID subunit 5
TBK1 - Serine/threonine-protein kinase TBK1
TFRC/ CD71 - Transferrin receptor protein 1
TGFBR1 - TGF-beta receptor type-1

TK1 - Thymidine kinase, cytosolic
TOP2A - DNA topoisomerase 2-alpha
TP53 - Cellular tumor antigen p53
TPM1 - Tropomyosin alpha-1 chain
TPM4 - Tropomyosin alpha-4 chain
TPX2 - Targeting protein for Xklp2
TRPV5 - Transient receptor potential cation channel subfamily V member 5
TRRAP - Transformation/transcription domain-associated protein
TSC2 - Tuberin
TUBB3 - Tubulin beta-3 chain
TUBB6 - Tubulin beta-6 chain
UBE2C - Ubiquitin-conjugating enzyme E2 C
UHRF1 - E3 ubiquitin-protein ligase UHRF1
ULK1 - Serine/threonine-protein kinase ULK1
UQCRC1 - Cytochrome b-c1 complex subunit 1, mitochondrial
UQCRC2 - Cytochrome b-c1 complex subunit 2, mitochondrial
VDAC1 - Voltage-dependent anion-selective channel protein 1
VDAC3 - Voltage-dependent anion-selective channel protein 3
VIM - Vimentin
VTCN1 - V-set domain-containing T-cell activation inhibitor 1
XPC - DNA repair protein complementing XP-C cells
ZBTB21 - Zinc finger and BTB domain-containing protein 21
ZNF512 - Zinc finger protein 512
ZPR1 - Zinc finger protein ZPR1

scientific reports

Check for updates

OPEN

The SKBR3 cell-membrane proteome reveals telltales of aberrant cancer cell proliferation and targets for precision medicine applications

Arba Karcini[1] & Iulia M. Lazar[1,2,3]✉

The plasma membrane proteome resides at the interface between the extra- and intra-cellular environment and through its various roles in signal transduction, immune recognition, nutrient transport, and cell–cell/cell–matrix interactions plays an absolutely critical role in determining the fate of a cell. Our work was aimed at exploring the cell-membrane proteome of a HER2+ breast-cancer cell line (SKBR3) to identify triggers responsible for uncontrolled cell proliferation and intrinsic resources that enable detection and therapeutic interventions. To mimic environmental conditions that enable cancer cells to evolve adaptation/survival traits, cell culture was performed under serum-rich and serum-deprived conditions. Proteomic analysis enabled the identification of ~ 2000 cell-membrane proteins. Classification into proteins with receptor/enzymatic activity, CD antigens, transporters, and cell adhesion/junction proteins uncovered overlapping roles in processes that drive cell growth, apoptosis, differentiation, immune response, adhesion and migration, as well as alternate pathways for proliferation. The large number of tumor markers (> 50) and putative drug targets (> 100) exposed a vast potential for yet unexplored detection and targeting opportunities, whereas the presence of 15 antigen immunological markers enabled an assessment of epithelial, mesenchymal or stemness characteristics. Serum-starved cells displayed altered processes related to mitochondrial OXPHOS/ATP synthesis, protein folding and localization, while serum-treated cells exhibited attributes that support tissue invasion and metastasis. Altogether, our findings advance the understanding of the biological triggers that sustain aberrant cancer cell proliferation, survival and development of resistance to therapeutic drugs, and reveal vast innate opportunities for guiding immunological profiling and precision medicine applications aimed at target selection or drug discovery.

Breast cancer is a common form of cancer that continues to lead, even in the present day, to a large number of deaths among women worldwide[1]. The different breast cancer subtypes are defined based on the presence of ER, HER2 or PR receptors, whether alone or in combination. HER2+ and triple negative breast cancers have the worst prognosis due to the fact that some HER2+ tumors are either non-responsive or develop resistance to anti-HER2 therapies, while triple negative cancers are non-responsive to hormonal therapies or drugs that target HER2 receptors[1,2]. As a result, focus has been placed on the development of novel therapeutic approaches that rely either on the use of various drug cocktails and treatment regimens that target multiple receptors or compensatory and downstream crosstalk signaling pathways of HER2, or, more recently, on triggering immune system responses that attack the cancer cells[2].

The heavy interest in the study of cancer cell-membrane receptors has been fueled by their central role in initiating cellular signaling cascades that lead to aberrant cell proliferation, as well as by their potential as cancer markers or drug targets. Cell-membrane receptors include three traditional protein categories, i.e., G-protein-coupled receptors (GPCRs), ion channels, and enzyme-linked receptors-mostly represented by receptor tyrosine kinases (RTKs)[3]. GPCRs represent the largest class of receptors[4], while the enzyme-linked receptors the most

[1]Department of Biological Sciences, Virginia Tech, Blacksburg, VA 24061, USA. [2]Fralin Life Sciences Institute, Virginia Tech, Blacksburg, VA 24061, USA. [3]Virginia Tech Carilion School of Medicine, Roanoke, VA 24016, USA. ✉email: malazar@vt.edu

studied one[5], and together they comprise the majority of drug targets. The aberrant activity of these receptors was linked to many diseases including inflammation, metabolic disorders, and cancer[6]. Targeting, for example, HER2 receptors has been at the core of targeting HER2+ tumors. Proteomic analysis of cell-surface (CS) proteins has revealed, however, many important, additional roles for other CS proteins in cancer proliferation[7]. The detection and characterization of these cell-surface targets has been, nevertheless, challenging due to compounding factors such as low abundance, hydrophobicity, presence of post-translational modifications (PTMs), and heterogeneity[8,9].

Several methods have been developed for the isolation of cell-surface proteins relying mainly on ultra-centrifugation, coating of the plasma membrane with silica-beads, and chemical labeling of N-linked glycosylated proteins or of protein amine, sulfhydryls or aldehyde groups, followed by affinity pulldown[8–12]. After isolation, the state-of-the art for detecting the enriched CS protein fractions involves mass spectrometry (MS) analysis. The advanced capabilities of the MS technology (i.e., high sensitivity, high mass accuracy and quantification capability) enabled the detection of thousands of proteins per cell line, the compilation of comprehensive cell-surface protein data into interactive databases such as The Cell-Surface Protein Atlas (~ 1500 human proteins[13]), and the development of even more comprehensive lists constructed with machine learning based predictor tools (~ 2900 human proteins[14]). Altogether, these studies have contributed to the overall knowledge of what has been named the "surfaceome" and its associated signaling networks in humans[14,15], largely captured in comprehensive public repositories[16–19].

To capitalize on the wealth of information that can be generated through mass spectrometric analysis, this study was aimed at characterizing the cell-surface proteome of SKBR3/HER2+ breast cancer cells by using orthogonal methods for cell-surface protein enrichment and isolation, categorizing these proteins based on their functional role and relevance to cancer, identifying key drivers of aberrant proliferation, and exploring the opportunities presented by such cells for the development of effective diagnostic and therapeutic approaches. We also report on the remodeling of the cell-membrane proteome under serum-starved and serum-supplemented conditions, and, lastly, we draw insights into the signaling cascades initiated at the plasma membrane and the potential crosstalk activities that fuel the development of resistance to treatment with therapeutic drugs.

Methods

Reagents and materials. SKBR3 cells, trypsin (0.25%)/EDTA (0.53 mM) and PenStrep solution were purchased from ATCC (Manassas, VA), and fetal bovine serum (FBS) from Gemini Bio-Products (West Sacramento, CA). McCoy's 5A (Modified) medium, Dulbecco's Phosphate Buffered Saline solution (DPBS), DPBS with calcium and magnesium (+Ca^{2+}/Mg^{2+}), and TrypLE Select Enzyme solutions were purchased from Gibco (Carlsbad, CA). Sample processing reagents such as NaF, Na_3VO_4, dithiothreitol (DTT), urea, ammonium bicarbonate (NH_4HCO_3), acetic acid, trifluoroacetic acid (TFA), and Triton-X were from Sigma (St. Louis, MO). Aniline was from BeanTown Chemical Corporation (Hudson, NH). Sequencing grade trypsin and trypsin/LysC were purchased from Promega (Madison, WI). Protease inhibitors cocktail (HALT), EZ-Link Sulfo-NHS-SS-Biotin Pierce Cell-surface Biotinylation and Isolation Kit, EZ-Link Alkoxyamine-PEG4-Biotin, Pierce Sodium meta-Periodate, EasyPep Mini MS Sample Prep Kit, and Streptavidin Alexa Fluor 488 conjugate were purchased from Thermo Scientific (Rockford, IL). Primary polyclonal rabbit ATP5A and P2Y2 antibodies, as well as goat anti-rabbit IgG (H + L) secondary antibody [DyLight 488], were purchased from Novus Biologicals (Centennial, CO). DAPI powder was obtained from Cell Signaling Technology (Danvers, MA), and ProLong™ Diamond Antifade Mountant with DAPI solution from Life Technologies Corp. (Carlsbad, CA). SPEC-PTC18, SPEC-PTSCX sample cleanup pipette tips and Bond Elut C18/3 mL cleanup cartridges were from Agilent (Santa Clara, CA), cell culture slides (8-chamber) for fluorescent visualization of cells from MatTek (Ashland, MA), and Nunc cell culture flasks from Thermo Scientific. HPLC-grade solvents such as methanol and acetonitrile were purchased from Fisher Scientific (Fair Lawn, NJ). Water for the preparation of sample solutions and LC eluents was either produced by a MilliQ Ultrapure water system (Millipore, Bedford, MA) or was distilled from DI water.

Cell culture. The SKBR3 cells were cultured in McCoy's 5A medium and FBS (10%) in T175 Nunc flasks, at 37 °C and in the presence of CO_2 (5%). After reaching ~ 70–80% confluence, for the first set of culture conditions, the cells were washed twice with serum-free medium and incubated in McCoy 5A for 48 h without any supplements. For the second set of culture conditions, after 48 h serum starvation, the cells were incubated for 24 h in McCoy 5A supplemented with FBS (10%). Penstrep (0.5%) was added to all culture media to prevent bacterial contamination. Two T175 flasks of serum-free (SF) or serum-treated (ST) cells (~ 90% confluence, 15–20 million cells/flask) were prepared for each cell-surface protein harvesting procedure, by either chemical labelling or proteolytic cleavage methods, as described below. Three distinct biological replicates (n = 3) of each condition were generated for analysis.

Microscopy. Several cell surface proteins were visualized by immunolabeling with primary rabbit antibodies against ATP5F1A or P2RY2 and secondary antibody conjugated to DyLight488 (NovusBio). The cells were fixed in cold methanol (− 20 °C, 5 min), blocked with BSA (5% in PBS, 1 h, room temperature-RT), and incubated with the primary antibody (1:100 dilution, 4 °C, overnight) in BSA (1% in PBS). The following day, the cells were incubated with the secondary antibody (1:2500 dilution, RT, 1 h, dark) in BSA (1% in PBS), and cured with ProLong Diamond antifade mountant with DAPI (RT, 24 h, dark) or DAPI solution (1 µg/mL, RT, 5 min, dark). Alternatively, the cells were fixed with paraformaldehyde (PFA) solution (2% in PBS, RT, 15 min) and permeabilized with Triton X-100 (0.5% in PBS, RT, 5 min). The blocking and incubation with antibody steps were the same as above. Images were acquired either with an inverted epi-fluorescence Eclipse TE2000-U microscope (Nikon Instruments Inc, Melville, NY) with a 20X air objective, or by confocal scanning with SoRa mode with

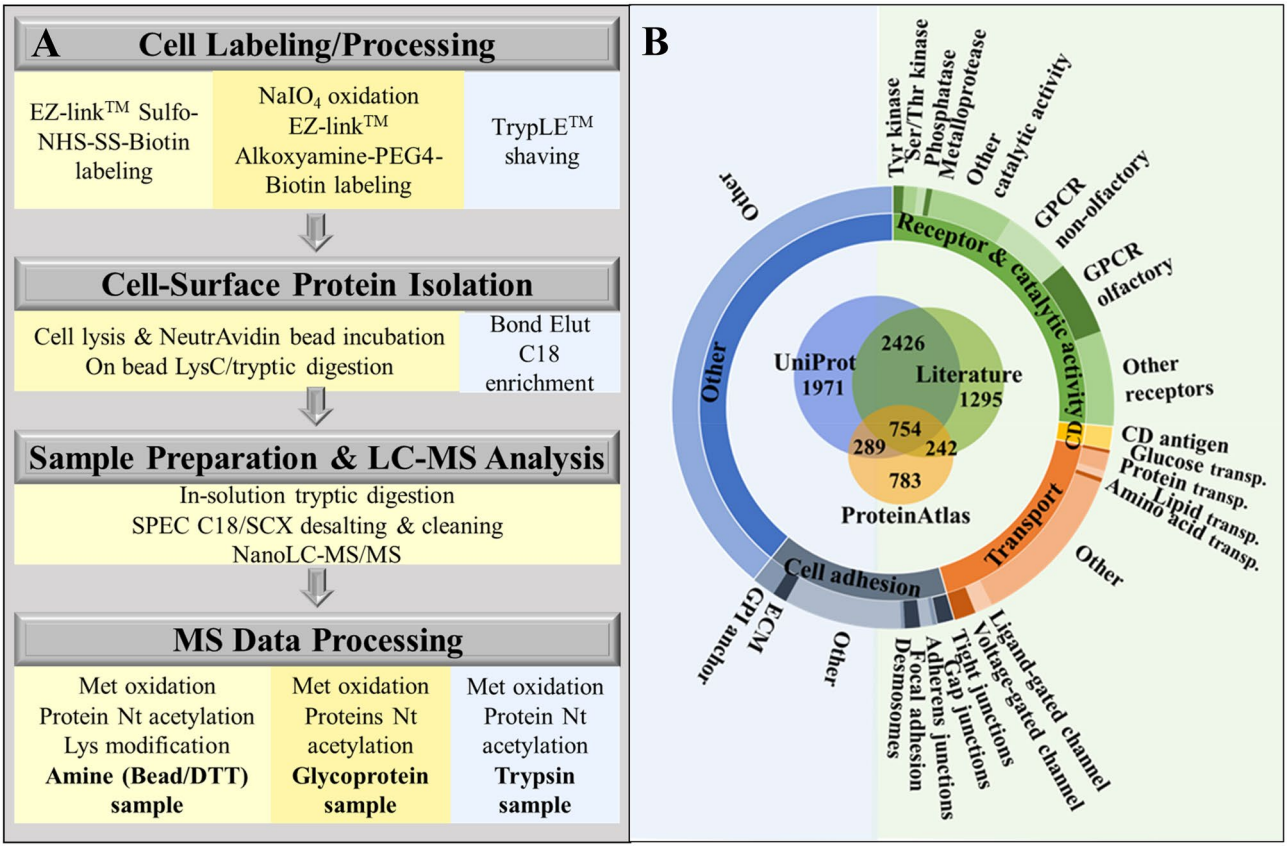


Figure 1. Cell-membrane protein isolation flowchart and database classification. (**A**) SKBR3 cell-membrane protein isolation and processing workflow via three distinct methods: biotin labeling of protein primary amine groups, biotin labeling of glycoproteins, and enzymatic shaving. (**B**) In-house built database of 7760 cell-membrane proteins classified based on GO controlled vocabulary terms.

a Nikon Eclipse Ti2 with a 40X water objective. The images were processed with Nikon software Denoise.ai and NIS-Elements AR Analysis 5.11.01.

Cell-membrane protein labeling and harvesting. To isolate the cell-membrane fraction of SKBR3 cells, a combination of chemical labeling and enzymatic approaches was followed (Fig. 1A). Based on reported yields and processing times[8–13], three methods, relying on protein isolation by biotinylation of amino groups and of glycan posttranslational modifications and affinity pulldown, as well as on tryptic shaving of receptors in cell culture, were chosen. Sulfo-NHS-SS-biotin based isolation of proteins enabled the labeling of primary amino groups at the protein N-terminal (α-amino) and Lys (ε-amino) residues, while alkoxamine-PEG4-biotin based isolation, enabled the labeling of carbohydrate moieties that are commonly encountered on the cell-surface proteins. Trypsinization of cells in culture was the least time-consuming method due to minimal processing prior and after sample collection. All reagent solutions were prepared fresh before use, and the reagent and rinse solutions that were used for biotin labeling were cooled to 4 °C before adding to the cells. The first labeling procedure involved the use of EZ-Link Sulfo-NHS-SS-Biotin (0.5 mg/mL) for labeling the protein N-terminal and Lys side-chain amino groups. Cells were rinsed twice with DPBS (+ Ca^{2+}/Mg^{2+}) and then incubated at 4 °C for 30 min, in the dark, with the biotin reagent. After incubation, the biotin reagent was removed, and each flask was washed twice with 20 mL Tris quenching buffer solution (0.1 M) provided in the kit. The cells were collected by scraping in Tris-buffer (10 mL per flask), and centrifuged for 5 min at 800×*g* and 4 °C. The second approach involved the labeling of cell-surface glycoproteins with EZ-Link Alkoxyamine-PEG4-Biotin (0.5 mg/mL) following protocols described by the manufacturer and in previous manuscripts[12], with some modifications. Briefly, the cells were rinsed twice with DPBS (+ Ca^{2+}/Mg^{2+}) and incubated at 4 °C for 30 min, in the dark, with 20 mL sodium meta-periodate solution (1 mM, pH 6.5) to oxidize the glycan moieties of cell-surface proteins. The cells were rinsed again, twice, with DPBS (+ Ca^{2+}/Mg^{2+}), and incubated with 12 mL biotin reagent solution in the presence of 10 mM aniline at 4 °C for 30 min, in the dark. After the completion of the labeling reaction, the biotin reagent was removed, and each flask was washed twice with 20 mL DPBS (+ Ca^{2+}/Mg^{2+}). Cell-surface protein biotinylation of cells was visualized with an inverted Eclipse TE2000-U epi-fluorescence microscope (Nikon, Melville, NY), after staining the cells with Streptavidin Alexa Fluor™ 488 (4 µg/mL). The cells were collected by scraping in 10 mL DPBS and centrifuged for 5 min at 800×*g* and 4 °C. The labeled cell pellets generated by either procedure were frozen at − 80 °C for further processing or subjected to immediate lysis. The third approach consisted of shaving the cell-surface protein ectodomains with TrypLE, a reagent that contains recombinant enzymes for

cell dissociation that are free of animal origin trypsin. For this procedure, the SKBR3 cells were washed twice with serum-free medium, and incubated with 10 mL TrypLE solution at 37 °C, with 5% CO_2, for 2–4 min. The incubation time was short, to prevent cell detachment. The cell supernatant containing the cell-surface protein ectodomains was then collected, centrifuged for 5 min at 500×*g* and 4 °C for the removal of floating cells, and frozen at − 80 °C. The samples generated through the three enrichment procedures will be referred from now on as the amine, glyco, and trypsin samples.

Cell-membrane protein recovery and processing. To isolate the cell-surface proteins of the amine-biotinylated samples, the cells were lysed with 500 uL Lysis Buffer (Pierce) supplemented with HALT protease inhibitor cocktail (5 uL), for 30 min, on ice, with intermittent vortexing and sonication. The lysate (~ 500 uL) was collected by centrifugation (15,000×*g*, 5 min, 4 °C) and incubated with 250 uL NeutrAvidin beads at room temperature for 2 h, followed by 4 washes with Wash Buffer (Pierce) and 3 washes with NH_4HCO_3 (100 mM). After each wash, the beads were isolated by centrifugation (1000×*g*, 1 min). Protein recovery from the beads was performed by proteolytic digestion, on the bead, overnight, RT, in 200 uL NH_4HCO_3 (100 mM) supplemented with 25 uL trypsin/Lys C solution (10–12 µg enzyme). After centrifugation (1000×*g*), the beads were further treated with 200 uL DTT (10 mM) for 1 h at RT to recover the di-thiol, covalently bound remaining protein fragments. Both on-bead protein digest and DTT-released fractions were collected and denatured with urea (8 M) for 1 h at 57 °C (the on-bead digest solution was also added DTT, 5 mM). After dilution with NH_4HCO_3 (100 mM) to reduce the urea concentration to < 1 M, the samples were subjected to a second digestion in solution with 10 uL trypsin (~ 5 µg enzyme) for 4 h at 37 °C. After quenching the enzymatic reaction with TFA, the cell-surface peptide extracts were processed for salt and detergent disposal with SPEC-PTC18 and SPEC-SCX cartridges. Isolation of the cell-surface proteins of the biotinylated glyco samples was performed by following a similar procedure to the one that was used for the amine-labeled samples. The cell lysate was incubated with NeutrAvidin beads, the beads were treated with 200 uL DTT (45 mM, 1 h, RT, dark), and after the removal of the DTT solution by centrifugation (1000×*g*, 1 min), on-bead proteolytic digestion was performed overnight, RT, in 200 uL solution of NH_4HCO_3 (100 mM) with 25 uL trypsin/Lys C (10–12 µg enzyme) in the presence of urea (1 M). An additional 4 h digestion at RT was performed by adding to the beads 100 uL NH_4HCO_3 (50 mM) and 10 uL trypsin solution (~ 5 µg enzyme). The collected glycoprotein fraction was then processed with SPEC-PTC18 and SPEC-SCX cartridges. Control samples were prepared from unlabeled cells in the same exact manner. To isolate the cell-surface proteins of the trypsinized samples, the collected solution (~ 10 mL) was digested in a preliminary stage, overnight, with 20 µg trypsin at 37 °C, and concentrated then on a Bond Elut C18 column to remove the large volume of TrypLE solution. The sample was then reconstituted in Tris-buffer (120 uL, 50 mM, pH = 8) and denatured with urea (8 M) and DTT (5 mM) for 1 h at 57 °C. After reducing the urea concentration to < 1 M with NH_4HCO_3 (100 mM), the sample was subjected to a second digestion with trypsin (~ 5 µg) for 4 h at 37 °C. After cleanup, all peptide samples were dissolved in 30 uL CH_3CN/H_2O/TFA (95–98):(2–5):0.01 v/v for LC–MS analysis. Protein concentration measurements for either of these samples, prior to processing and proteolytic digestion, could not be performed due to limited sample availability and low abundance of the cell-surface proteins in solution.

LC–MS analysis. The peptide samples were analyzed with an EASY-nLC 1200 UHPLC system (ThermoFisher Scientific) by using a heated nano-electrospray ionization (ESI) source (2 kV) and a Q Exactive hybrid quadrupole-Orbitrap mass spectrometer (ThermoFisher Scientific). An EASY-Spray column ES802A (150 mm long, 75 µm i.d., 3 µm C18/silica particles, ThermoFisher Scientific) was used at 45 °C and flow rates of 250 nL/min. The mobile phases were prepared from H_2O:CH_3CN:TFA, and mixed in ratios of 96:4:0.01 v/v for mobile phase A and 10:90:0.01 v/v for B. During a separation gradient of 85 min, the eluent B concentration was increased from 3 to 30% (5–65 min), 45% (65–72 min), 60% (72–73 min), and 90% (73–74 min), where it was kept for 5 min, and then decreased to a final concentration of 3%. LC separation stability was monitored via the output pressure which was maintained at 90–93 bar at 5% B. Control samples were separated on a 250 mm nano-LC column, with a 2 h long gradient, to maximize the detection of proteins retained on NeutrAvidin beads through non-specific interactions. The MS data were acquired over a range of 400–1600 m/z with resolution set to 70,000, AGC target to 3E6, and maximum IT to 100 ms. Data-dependent MS2 acquisition (dd-MS2) was enabled by using higher-energy collisional dissociation (HCD), isolating the precursor ions with a width of 2.4 m/z, and fragmenting them at 30% normalized collision energy (NCE). dd-MS2 acquisition parameters were set to resolution 17,500, AGC target 1E5 (minimum AGC target 2E3 and intensity threshold 4E4), maximum IT 50 ms, and loop count 20. Charge exclusion was enabled for unassigned and + 1 charges, apex trigger was set to 1 to 2 s, dynamic exclusion lasted for 10 s for chromatographic peak widths of 8 s, and the features of isotope exclusion and preferred peptide match were turned on. For parallel reaction monitoring (PRM) validation, the peptides of interest were searched within a time-window of +/− 10 min of the precursor ion retention time, following a separation gradient of 2 h on a 250 mm Easy-Spray LC column (ES902 PepMap™ RSLC C18, 75 µm i.d., 2 µm particles, 100 Å). The precursor ions were isolated with a width of 2.0 m/z, and fragmented at 30% normalized collision energy with PRM parameters set as follows: resolution 35,000, AGC target 2E5, and maximum IT 110 ms.

MS raw data processing. The MS data were processed by the Proteome Discoverer 2.4 package (Thermo Fisher Scientific, Waltham, MA) and searched with Sequest HT against a *Homo sapiens* database (DB) of 20,433 reviewed, non-redundant protein sequences downloaded from the UniProtKB/Swiss-Prot public repository (March 2019 download). The processing workflow spectrum filter was set for a peptide precursor mass range of 400–5000 Da, and the Sequest HT node parameters allowed for the selection of fully tryptic peptides (6–144 aa

length) with maximum two missed cleavages, 15 ppm precursor ion tolerance, b/y/a ion fragments with 0.02 Da tolerance, and dynamic modifications (maximum 4 per peptide) on Met (15.995 Da/oxidation) and the protein N-terminal amino acids (42.011 Da/acetyl). Carbohydrate group labeling and trypsinization do not alter the chemical structure of the cell-surface proteins, but labeling of amine groups with the biotinylation reagent forms a 3-mercapto-propanamide derivative, for which a dynamic modification of 87.998 on the Lys residues was also enabled. The raw files were processed independently for the fraction of proteins generated by direct on-bead digestion and DTT reduction, but for reporting, the results were merged. Overwhelmingly, though, the majority of protein identifications were enabled by the on-bead digestion step, rendering the additional DTT recovery step unnecessary. The peptide spectrum match (PSM) validator node used a target/decoy concatenated database strategy to calculate the *FDR* targets of 0.01 (strict) and 0.03 (relaxed) based on search engine Xcorr scores (input data of maximum DeltaCn 0.05 and maximum rank 1). Additional parameters were set in the consensus workflow for both peptide and protein levels. The peptide group modification site probability threshold was set to 75. Peptide confidences were represented by the corresponding best PSM confidences. Only peptides of at least medium confidence and proteins matched by only rank 1 peptides were retained in the peptide/protein filter node. The peptides were counted only for top scoring proteins. The protein *FDR* validator node used the protein scores from the target and decoy searches to calculate the *FDR*s and rank the proteins, and then calculate the q-values from the *FDR*s at each score threshold. The *FDR*s were set to 0.01 (high) and 0.03 (medium) for PSMs, peptides, and proteins, and the strict parsimony principle was enabled for protein grouping. The PRM data were processed by Skyline 20.2[20] by using a mass spectral library generated from cell-surface protein samples produced by the glycoprotein enrichment method. The *b* and *y* ions were selected from "ion 2" to "last ion" with precursor charges of 2 and 3, and fragment charges of 1 and 2. The library ion match tolerance was 0.02 m/z, and the 5 or 10 most intense product ions were picked from the filtered product ions. The presence of a peptide was considered validated when the peptide displayed a minimum of 5 transitions and when the dot product (*dotp*) score was roughly > 0.8.

Bioinformatics data interpretation and visualization. An in-house database of cell-membrane proteins was built by extracting relevant entries from the UniProtKB/Swiss-Prot database based on controlled vocabulary terms[16], from the Human Protein Atlas (HPA) Cellular and Organelle Proteome[17,18], and from the scientific literature[12–15]. GeneCards[21] and UniProt[16] were used to assess protein functionality. STRING 11.5 was used to build protein–protein interactions (PPI) networks and assess GO enrichment in biological processes[22], with interaction score confidences set to medium/high and enrichment *FDR* < 0.05. Cytoscape 3.8.2[23] was utilized to visualize protein networks based on interactomics data exported from STRING, RAWGraphs-an open source data visualization framework[24]-was used for building the dendograms, InteractiVenn.net for building Venn diagrams, and Protter was used for visualizing the location of a protein relative to the cell-membrane bilayer[25].

Statistical analysis of changes in protein abundance. For each of the three biological replicates ($n = 3$), three LC–MS/MS technical replicates were performed, and the results of the three technical replicates were combined in one multiconsensus protein and peptide report. Protein detection reproducibility and quantitation was performed based on spectral counting. The strength of the bivariate (linear) relationship between any two sets of biological replicates was evaluated with the Pearson correlation coefficient "*r*". For evaluating changes in protein abundance, missing values were handled by adding one spectral count to each protein from the dataset. Data normalization was performed based on spectral counting, in two steps. In the first step, normalization was performed at the global level by averaging the total spectral counts (SC) of the six samples taken into consideration (i.e., three SF and three ST biological replicates), and using the resulting average as a correction factor (CF1) for adjusting the counts of individual proteins in each sample. In the second step, normalization was performed at the cell-surface protein level by calculating a second correction factor (CF2) based on the spectral counts of only a short list of 10 endogenous cell-surface proteins that were already corrected by CF1 [see Eqs. (1) and (2) below]. Proteins that changed abundance in the cell-surface proteome were selected by calculating the Log2 values of the ST/SF spectral count ratios and using a two-tailed *t*-test for assessing significance. Proteins matched by two unique peptides with fold change (FC) ≥ 2 in normalized spectral counts and p-value < 0.05 were considered for discussion.

$$SC_{jiN} = SC_{ji} \times CF1_j \times CF2_j \tag{1}$$

$$SC_{jiN} = SC_{ji} * \frac{\frac{1}{6}\sum_{j=1}^{6}\left(\sum_{i=1}^{x} SC_{ji}\right)}{\sum_{i=1}^{x} SC_{ji}} \times \frac{\frac{1}{6}\sum_{j=1}^{6}\left[\sum_{i=1}^{10}\left(SC_{jiE} \times CF1_j\right)\right]}{\sum_{i=1}^{10}\left(SC_{jiE} \times CF1_j\right)} \tag{2}$$

SC_{ji} = spectral count of protein "*i*" in data set "*j*" ($j = 1$–6); SC_{jiN} = normalized SC_{ji}; SC_{jiE} = spectral count of endogenous membrane protein "*i*" in data set "*j*" (10 endogenous proteins were considered); x = total number of proteins identified in the 6 sample sets taken for comparison (3 × ST vs. 3 × SF).

Results

Cell-membrane protein database. To create a theoretical framework for mapping the cell-membrane proteome, an in-house database containing 7760 proteins was assembled by using information from the literature and two public resources, i.e., UniProt[16] and the Human Protein Atlas (HPA)[17,18] (Supplementary Data S1). UniProt proteins were derived by using the advanced search interface that returned 5440 protein IDs. Controlled

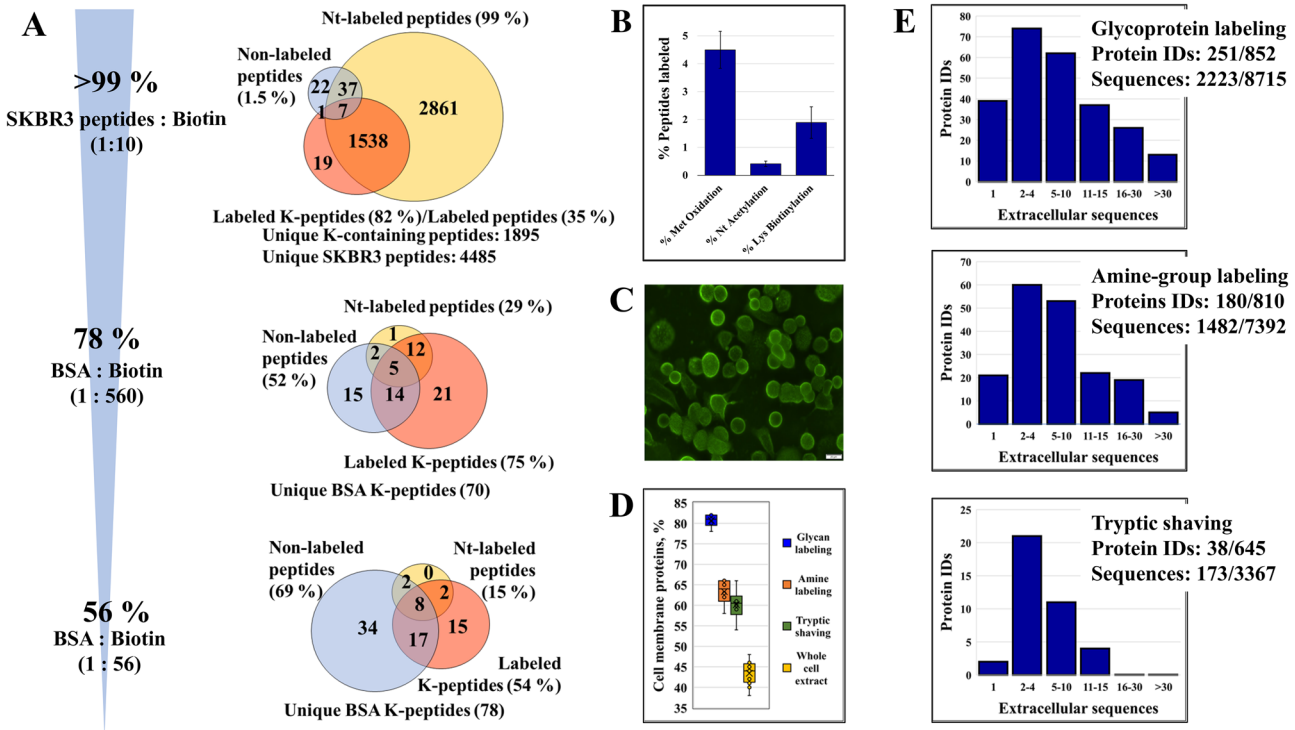


Figure 2. Cell membrane protein labeling efficiency and enrichment in extracellular sequences. (**A**) Percent peptides carrying a biotinylation-induced label in an SKBR3 cell extract and in BSA tryptic digests using various peptide/protein:biotin molar ratios (applicable to the amine group labeling method); the Venn Diagrams represent the number of labeled Lys-containing peptides (+87.998 Da), peptides labeled at the N-terminus (+87.998 Da), and the number of non-labeled peptides. (**B**) Percent peptides carrying a PTM: Met oxidation, peptide Nt acetylation, and Lys biotinylation (case of the amine group labeling method with on-bead proteolytic digestion); the error bars represent the SD of biological replicates. (**C**) SKBR3 cells labeled by alkoxyamine-PEG4-biotin and conjugated with streptavidin antibody-Alexa Fluor™ 488. (**D**) Cell membrane protein enrichment effectiveness represented by the number of cell membrane proteins in the top 100 most abundant proteins (abundance determined by the number of matching unique peptides). (**E**) Histograms of protein IDs matched by different numbers of extracellular peptide sequences (with extracellular sequences detected/total); extracellular sequence assignments were made based on topological domain information extracted from UniProt.

vocabulary terms were used for searching the cellular compartment (CC), Gene Ontology (GO) and the Keyword (KW) fields, all filtered for cell membrane localization. Lists of proteins localized to the cell-membrane, cell-surface, cell junction, cell projection, and peripheral proteins, with roles in signaling (i.e., receptor and catalytic activity), immune response (e.g., CD antigens), adhesion, and transport, were extracted. Plasma membrane proteins from the HPA were retrieved in bulk (2068 IDs), and additional cell-surface proteins detected *in-vitro* by using various experimental enrichment techniques[12,13] or predicted via *in-silico* studies[14,15] were added to the list (4717 IDs). Complementary information about GPCR families was acquired through IUPHAR[19]. A classification of the cell-membrane proteins included in the database is presented in Fig. 1B. The overlap between the various protein categories was rather minimal (~10–15%), but unavoidable, due to the complex roles that the cell-membrane proteins play in several biological processes. With improvements in sample preparation technologies, MS detection sensitivity, and search engine machine learning capabilities, a more consistent consensus between the various information sources is also expected.

Effectiveness of cell-membrane protein isolation. The efficiency of the biotin labeling reaction was evaluated by using two BSA protein samples and SKBR3 tryptic peptides. The sample:Sulfo-NHS-SS-biotin molar ratios for the BSA protein samples were 1:56 and 1:560, respectively, and for the SKBR3 tryptic peptides was 1:10 (Fig. 2A). The labeling efficiency of the tryptic peptides was very high, reaching ~99% at the N-terminal (Nt) and ~82% for all Lys (K)-containing peptides. Only ~1.5% of the peptides were non- or partially labeled. However, when labeling was performed at the protein level and followed by proteolytic digestion, the labeling efficiency of Lys-containing BSA peptides dropped progressively to ~75% and ~54%, respectively, with the decrease in the molar ratio of the added biotinylation reagent. This was matched by a concomitant increase in the non-labeled or partially labeled peptides to 52% and 69%. N-terminal labeling of multiple BSA peptides was observed, as well, presumably due to incomplete quenching of the labeling reagent prior to proteolytic digestion. Nonetheless, the labeling of the BSA protein N-terminal amino acid could not be detected. The results underscore the impact of limited reagent accessibility to hindered Lys sites in a protein, which becomes a much

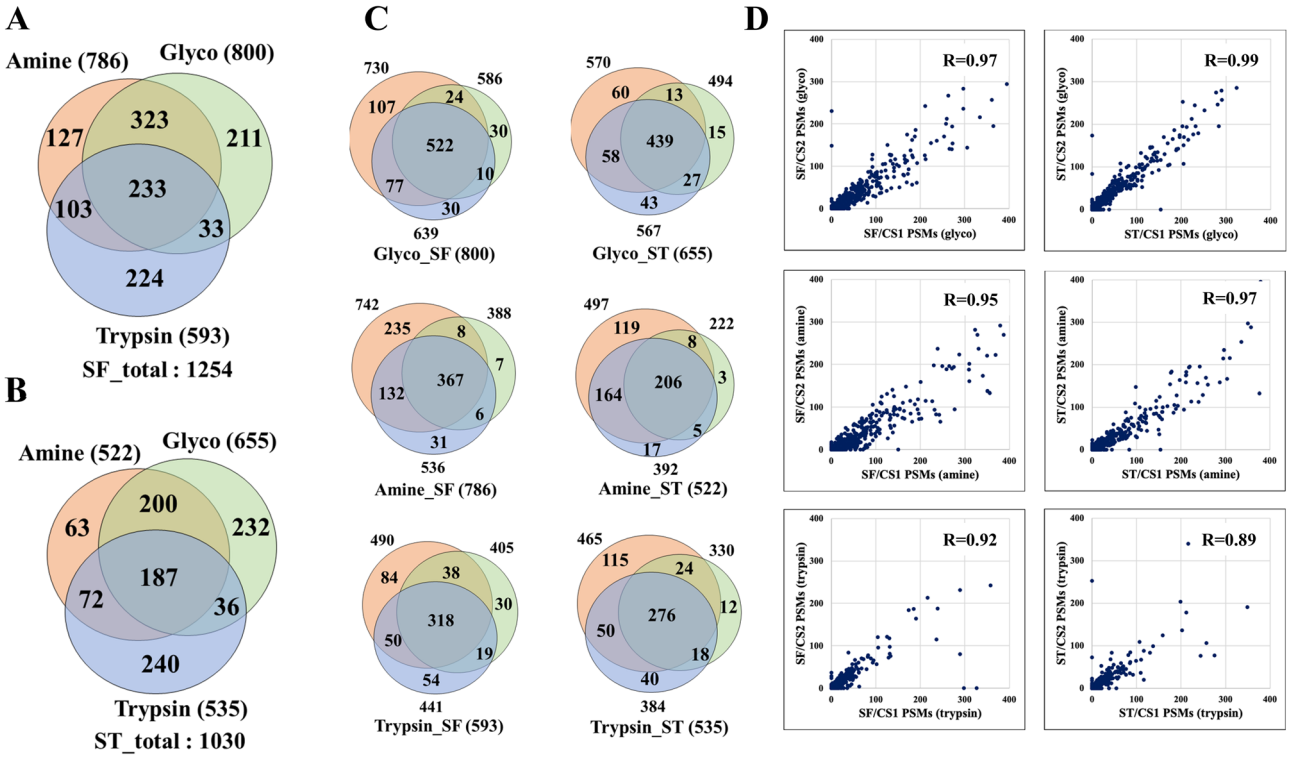


Figure 3. Protein ID Venn and PSM correlation diagrams representing the complementarity and reproducibility of the three labeling methods in detecting cell-membrane proteins matched by at least 2 unique peptides (*FDR* < 3%). (**A**) Serum-free cultured cells. (**B**) Serum-treated cells. (**C**) Reproducibility of protein detection between three biological replicates for each labeling method and cell treatment condition (SF and ST). (**D**) PSM correlations between any two biological replicates for each of the three cell-membrane protein enrichment methods; the correlations are shown for the 0–400 PSM range in which the vast majority of proteins could be found (R = Pearson correlation coefficient).

more challenging factor in the case of live cells when the extracellular domain of intact membrane proteins often displays a heavily modified and entangled structure[26]. Accordingly, Lys biotinylation of cell-surface proteins on live cells was observed at a substantially reduced level (~ 1.9%), even less than that of Met oxidation (~ 4.5%). Protein N-terminal acetylation was also detected at a low level (~ 0.4%), but biotinylation was not observable (Fig. 2B). As a result, in the final analysis, the biotinylation-induced modification on the N-teminus of proteins was not included in the list of enabled DB search modifications. In the case of alkoxamine-PEG4-biotin-based labeling and isolation of glycosylated proteins, the labeling efficiency could not be evaluated by MS as there was no change in mass involved, however, the attachment of the labeling reagent to the cell-surface proteins could be visualized by microscopy and indicated a uniform coverage (Fig. 2C).

The enrichment efficiency in cell-membrane proteins was assessed by MS, by calculating the proportion of membrane proteins in the top 100 most abundant proteins, with abundance defined by the counts of unique peptides per protein. According to the controlled vocabulary annotations in the database that was described above, the percentage of cell-membrane proteins in typical whole cell extracts was ~ 43%. Upon enrichment, this percentage increased to ~ 60%, ~ 64% and ~ 81% for tryptic shaving, amino group, and glycan labeling, respectively (Fig. 2D). Cell-surface protein enrichment based on glycan labeling provided the highest yield, most likely due to the heavy glycosylation of the extracellular protein domains that could be more efficiently labeled than the protein N-termini and Lys residues in the case of the amine labeling method[27]. Poor penetration of trypsin through the cell-surface protein coat, slow tryptic activity, and the possible contribution of lysed cell content to the pool of identified proteins may have led, on the other hand, to the lowest enrichment yield for the tryptic shaving method. For proteins for which topological information was available in UniProt (i.e., for 2923 proteins from the in-house built DB, matched by 7546 extracellular sequences), the topological domain assignments validated the presence of numerous cell-membrane proteins from the dataset. The histograms from Fig. 2E indicate that many of the detected proteins were identified by multiple extracellular sequences, confirming thus their presence in the cell-membrane or on the cell-surface, and also that in comparison to trypsinization the chemical labeling methods were more effective for capturing the cell-membrane proteome.

Cell-membrane proteome data analysis. The combination of orthogonal enrichment approaches led to the identification of a total of 2054 cell-membrane proteins in the combined SF and ST samples, of which 1921 were present in the SF and 1435 in the ST cell states. The number of proteins identified by at least two unique medium or high confidence peptides was 1316, 1254, and 1030, respectively (Fig. 3A,B; Supplementary

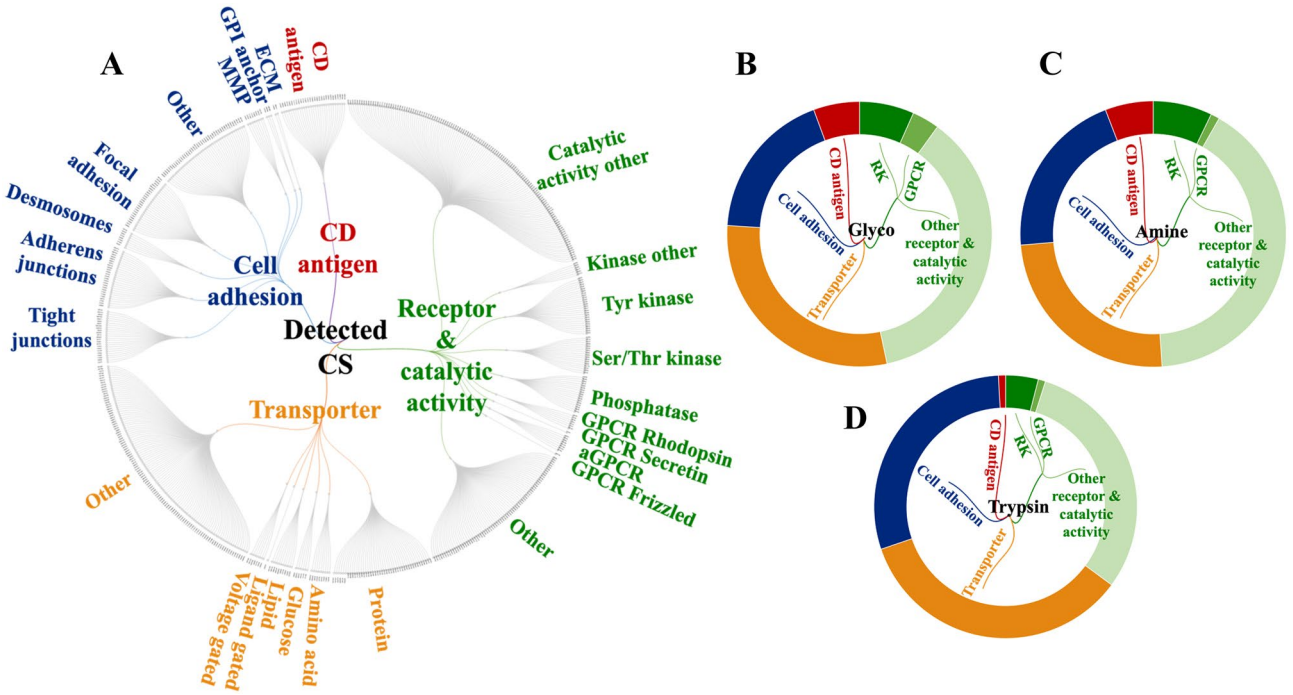


Figure 4. Functional categorization of the detected cell-membrane proteins based on GO controlled vocabulary terms (proteins detected by at least 2 unique peptides, *FDR* < 3%). (**A**) Dendrogram of cell-membrane proteins detected by all labeling methods and conditions. (**B**) (C), and (**D**) Doughnut charts of detected cell-membrane proteins enriched via glycoprotein labeling, amino group labeling, and tryptic shaving methods, respectively.

Data S2). The three methods were complementary to each other, however, as shown in the Venn diagrams from Fig. 3C, cell-membrane protein enrichment based on the labeling of glycoproteins enabled the identification of the largest number of proteins and with the best reproducibility, i.e., 65–67% overlap between three biological replicates. In terms of peptide spectrum matches, the quality of protein identification was high and consistent across all three methods with correlation factors ranging from 0.95 to 0.99 (~0.9 for some tryptic samples) (Fig. 3D). Non-biotinylated control cells processed with the glycoprotein enrichment method yielded 28 high-confidence proteins, of which only 11 were matched by two unique peptides and known to be associated with the cell-membrane (i.e., mainly abundant cytoskeletal/cytosol proteins). This indicated that the experimental procedure used in conjunction with the cell-membrane protein database worked well together to reduce the impact of contaminants retained by non-specific interactions on the NeutrAvidin beads (Supplementary Data S2).

The SKBR3 surfaceome. The remarkable ability of cancer cells to enact aberrant proliferation programs and metastasize to distant sites is mediated via an altered cell-surface proteome that facilitates in-and-out cell signaling processes as well as adhesion and migratory functions. To gain a better insight into these processes, a functional characterization of the SKBR3 surfaceome was performed by categorizing the set of 1316 proteins into four major groups: receptors and proteins with catalytic activity, transporters, cell adhesion/junction proteins, and proteins with immune functions such as CDs. Less abundant categories included receptor substrates, cell-surface binding or associated proteins, or proteins that are not typically recognized as membrane proteins (e.g., peripheral membrane proteins, cell projection, GPI anchors, matrix metalloproteinases/MMPs, and ECM molecules). The dendogram from Fig. 4A provides an overview of a representative subset of 525 cell-membrane proteins that could be placed in specific compartments, with protein IDs being included in Supplementary Data S3. Cell-surface protein enrichment by glycan or amino group labeling yielded the largest number of receptor/catalytic proteins and CDs (Fig. 4B,C), while trypsinization enabled the identification of a more abundant fraction in cell adhesion and transport proteins (Fig. 4D).

To further assess biological utility, the combined results of the three enrichment methods were compared to the output of three additional independent experiments, one including cell-surface protein enrichment from proliferative cells grown in serum-rich culture media, never exposed to serum starvation, and two including whole cell analysis of SF and ST cells without enrichment in cell-surface proteins (Table 1). As expected, when enrichment was performed, a larger number of cell-surface proteins were identified. A clear advantage, however, was observable only when proteins matched by at least two unique peptides were counted. In particular, the enrichment process enabled the high confidence detection of a much larger number of signaling receptors (GPCRs, Tyr receptor kinases), GPI anchors and CD antigens (columns 5, 6 vs. 7, 8). Notably, the glyco enrichment method alone enabled the identification of the vast majority of kinase/GPCR receptors and CD antigens, rendering it, therefore, the method of choice for profiling valuable targets for therapeutic treatment and immunophenotyping (Fig. 5). In contrast, cell-membrane Ser/Thr kinase receptors were detectable in higher numbers

	G1 + S (CS)	Proliferating (CS)	G1 + S (WC*)	G1 + S (WC#)	G1 + S (CS/2pep)	Proliferating (CS/2pep)	G1 + S (WC*/2pep)	G1 + S (WC#/2pep)
	1	2	3	4	5	6	7	8
Matches to CSDB	**2054**	**1359**	2265	2821	**1316**	**1175**	1295	1108
Matches to CS/Swiss-Prot	**1339**	**874**	1427	1831	**861**	**754**	772	667
Receptors	**168**	**112**	114	211	***117***	***106***	*24*	*21*
GPCRs	**36**	**15**	36	65	***15***	***11***	*1*	*1*
Tyr kinases	**29**	**20**	16	30	***26***	***20***	*8*	*4*
Ser/Thr kinases	**29**	**17**	40	44	**17**	**15**	25	19
Transport (ers)	**381**	**275**	348	425	**279**	**247**	198	163
Cell junction/cell adhesion	**348**	**251**	347	421	**255**	**231**	187	158
GPI anchors	**25**	**21**	16	28	***17***	***20***	*4*	*2*
Signal anchor	**24**	**21**	28	35	***20***	***19***	*12*	*11*
CD antigens	**105**	**84**	58	84	***89***	***80***	*19*	*13*
Secreted	**643**	**470**	718	897	**454**	**401**	396	366

Table 1. Cell-surface protein identification effectiveness with and without enrichment in cell-surface proteins, by considering the whole protein set or only proteins matched by two unique peptides. *CS* cell surface, *WC* whole cell, *2pep-2* unique peptides. *Data acquired with a QExactive Plus Orbitrap mass spectrometer. #The analysis included many replicates; as a result, a larger number of proteins were identified. Bold values represents protein count data from SKBR3 cells subjected to enrichment in cell-surface proteins. Underlined/italic values indicate conditions for which a substantial improvement in cell-surface protein enrichment was observed (vs. whole cell, non-enriched samples).

and with a larger number of unique peptides without performing cell-surface protein enrichment, likely due to the prevalently longer cytoplasmic tails in comparison to the shorter extracellular N-terminal domains.

Altogether, based on controlled vocabulary terms, the enrichment process enabled the classification of about 275 proteins with catalytic and receptor activity (including 15 GPCRs, 26 Tyr kinases, 17 Ser/Thr kinases), 89 CD antigens, 255 cell adhesion/junction molecules, and 279 transport proteins (Figs. 6A, 7). An interrogation of the biological processes (Fig. 6B) and associated pathways (Fig. 6C) represented by these proteins revealed for each category not just one, but multiple and complex roles with broad impact on essential cellular processes such as cell communication/signaling, biological adhesion and migration, transport, immune response, cell growth, death, and differentiation (Supplementary Data S2). Important to note is the additional impact imparted by proteins that are just temporarily associated with the cell membrane (e.g., peripheral membrane, ECM, secreted or exosome-associated proteins).

Proteins with receptor and catalytic activity. The category of cell-membrane proteins with catalytic and receptor activity included, in addition to a large group of kinase receptors, non-receptor kinases, phosphatases, MMPs, GTPase molecular switches, and proteins with ATPase activity. Together, these proteins are engaged in extensive cell-to-cell signaling and intracellular signal transduction, cell growth, apoptosis, cell locomotion and migration, trafficking of various cellular components, transport (ions, lipids, amino acid, metabolites), and regulation of actin cytoskeleton organization and cell polarization. Many of these biological processes are altered in cancer cells due to the presence of mutations[28].

Enzyme-linked receptors display extracellular domains for binding growth factors, cytokines or hormones, and initiate the transmission of chemical signals via their intracellular cytoplasmic domains that either have, or interact with proteins that have, catalytic activity. Among the detected catalytic receptors, the most relevant to cancer growth, proliferation and differentiation, and breast cancer specifically, were the Tyr protein kinases of the EGFR/ERBB, FGFR and IGFR families of growth factor and hormone binding receptors (Fig. 6B, Ca). These also included new RTK drug targets for breast carcinoma, such as the discoidin domain receptor (DDR1), which is involved in the activation of cell proliferation, survival, ECM remodeling, migration and invasion pathways[29]. In addition, Ser/Thr kinase receptors for a number of TGF-β superfamily of ligands (BMPR1A/BMPR2 bone morphogenic proteins, ACVR1/ACVR2 activin receptors)[30], as well as members of the TGFBR complex (ENG), were present. Modulation of TGFBR signaling is accomplished by interactions with a broad range of cell surface receptors and non-receptors, which were all detected in the cell membrane fraction of the SKBR3 cells (e.g., with ENG, NRP1, PDGFRβ, CD44, and integrins). Moreover, a fairly large collection of Tyr kinase ephrin receptors, semaphorin plexin receptors, neuropilin (NRP1), and ROBO1 formed a group with multiple important functions in developmental processes, differentiation, cytoskeleton remodeling, chemotaxis, and migration[31–35]. Ephrins are known to be involved in the regulation of multiple signaling pathways (e.g., MAPK, ERK, RAS, estrogen), and together with ROBO1 and NRP1 (a VEGF receptor) play major roles in angiogenesis and vascular development[16,31]. Plexin receptors, on the other hand, have been shown to be involved in invasive cell growth and ERBB signaling. Together with FGFR1, EGFR/ERBB, HMGB1, BMPR1A/BMPR2, and the non-catalytic NOTCH1/2/3 group, the plexin/ephrin receptors are further massively implicated in cell differentiation processes.

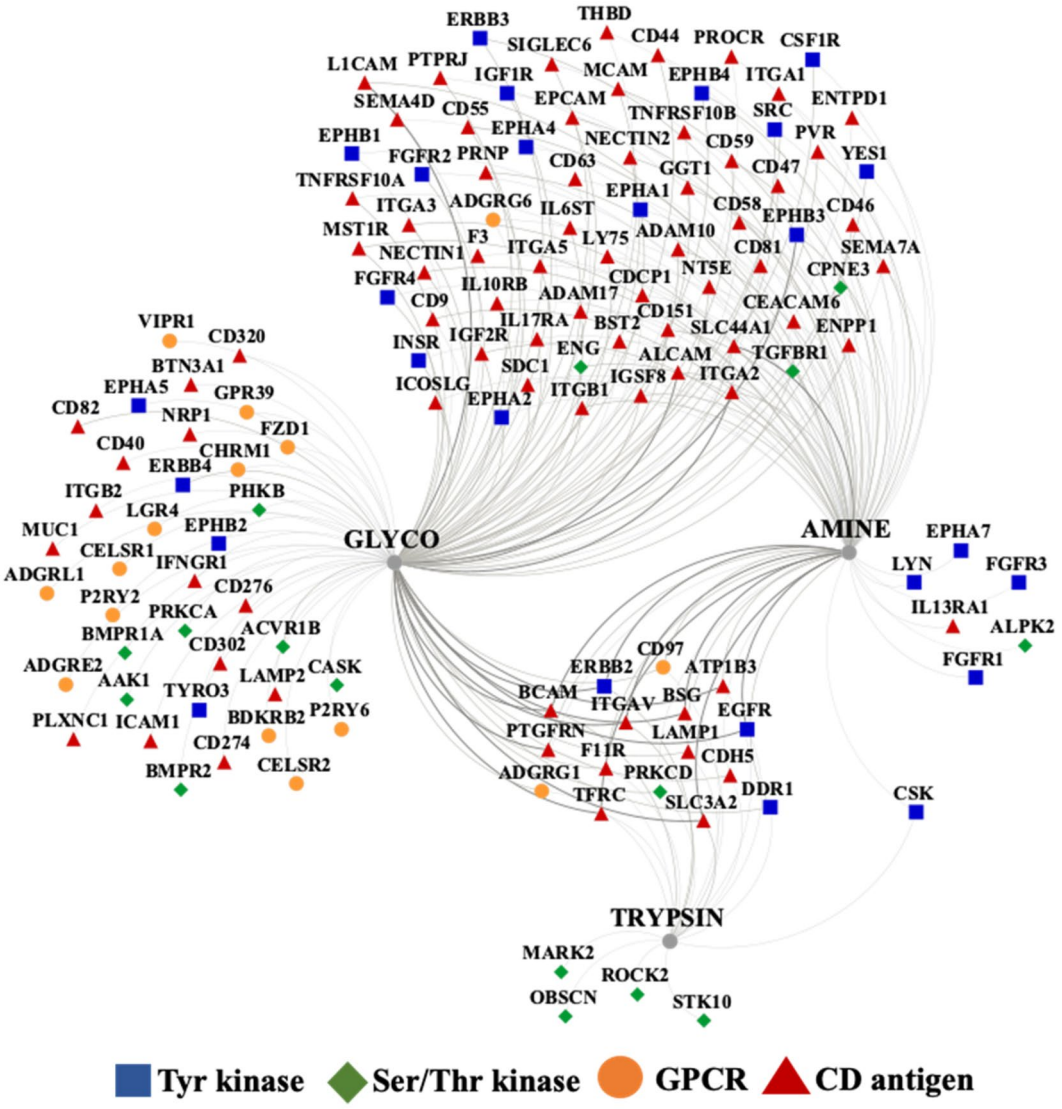


Figure 5. Detectability of cell-membrane receptors and CD proteins. The diagram represents the overlap between the detected proteins produced by the three enrichment methods (visualization performed with Cytoscape); the edge thickness reflects the protein abundance as evidenced by PSMs.

In contrast to RTKs, the GPCRs transduce extracellular signals by changing their conformation upon binding of a ligand and transmiting the signal through G-protein modulation. The GPCRs encompass ~ 800 protein members and are the largest family of cell-surface receptors[36]. The detected GPCRs were part of the Rhodopsin (class A) and Secretin (Class B) families (Fig. 7B), and were yielded mainly by the glycan labeling method, likely due to their rather low abundance and high glycosylation rate at the N-terminal sequences[37]. Supplementary Data S4 provides the PRM/MS data that validated the presence of the detected GPCRs. These GPCRs included adhesion (aGPCRs), Wnt signaling, and neuroactive ligand-receptors, several with EGF-like (CELSR1/2, CD97, EMR2) and hormone-receptor (VIPR1, CELSR1, LPHN1) Pfam domains. Most Secretin GPCRs were also adhesion GPCRs that contained GPCR proteolysis sites. The aGPCRs are evolutionarily conserved[37], and, in addition to their involvement in cell adhesion and migration processes, have emerged for their role in tumorigenesis and metastasis[38]. For example, ADGRE5 (CD97) and ADGRG1 (GPR56) are two aGPCRs that have been detected by the largest peptide counts and by all enrichment methods, and have been reported for increased expression in various cancers[38,39]. The Rhodopsin class A GPCRs included a cluster involved in neuroligand-receptor ligand interactions that also comprised two G protein-coupled purinergic nucleotide receptors. One of them, the P2RY6 receptor, is a known target for colorectal cancer due to its role in protecting cancer cells from apoptotic processes, possibly via AKT and/or ERK1/2 signaling[40]. While less is known about its role in breast cancer, GPCRs have been frequently found to be involved in modulating signaling pathways via cross-talk with other receptors[41,42].

Immune system receptors, CDs, and antigen characteristics. Cancer cell receptors that trigger cytotoxic innate and adaptive immune system responses are critical to the path of tumor development, and

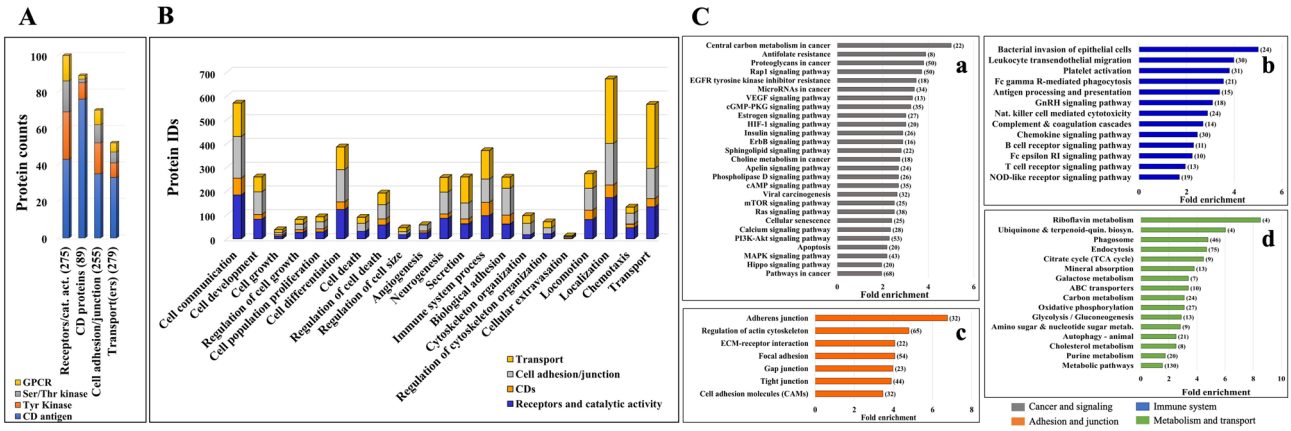


Figure 6. Bar charts of selected functional categories and pathways associated with the detected cell-membrane proteins. (**A**) Categorization of the detected receptors/enzymes, antigens, transporters and cell adhesion/junction proteins into Tyr kinase, Ser/Thr kinase, GPCR and CD groups. (**B**) Cancer-relevant enriched biological processes represented by the cell-membrane proteins. (**C**) Enriched KEGG pathways represented by cell-membrane proteins involved in: (Ca) signaling and cancer, (Cb) immune response, (Cc) adhesion/junction, and (Cd) metabolism and transport. Notes: Numbers in parentheses represent the number of proteins matched to each process; full lists, fold-enrichment and *FDR* values are provided in Supplemental file 2; the background gene sets were the full set of corresponding proteins in the human proteome. Examples of cell-membrane or membrane-associated proteins: Growth factor receptors (EGFRs, FGFRs, IGFRs, MET, OGFR), hormone receptors (INSR, IGF1R, ADIPOR1, VIPR), receptor type tyrosine protein phosphatases (PTPRA, PTPRJ), non-receptor Tyr kinases (LYN), Tyr kinase ephrins (EPHAs, EPHBs, EFNAs, EFNBs), plexins (PLXNs), nectins (PVRL1), GPCRs (rhodopsin, secretin, adhesion, frizzled), small GTPases (RHOA, RHOG, HRAS, RAC1, RAB13), proteins with roles in immune response (IL/IFN receptors, chemokine receptors and ligands, HLA class 1 antigens, HMGB1), adaptor proteins (SHC1), integrin receptors (ITGAs, ITGBs, ITGAV), cell–matrix adhesion (focal adhesion integrins, DAG1), cell–cell adhesion (adherens junction cadherins/protocodherins, desmosomal desmocollin/desmoglein/plakins, gap junction connexins, tight junction claudins, occludins, JAMs and ZO proteins), CAMs (EPCAM, BCAM, MCAM, CDH5/CD144, NEO1, mucins), immunoglobulin-like CAMs (ICAM1, L1CAM, ALCAM, nectins), MMPs (MMP15/24), disintegrins (ADAM 9/10), Ca-binding proteins (S100A8/9, PRKCA, spectrins), transport (ABC transporters, SLC carriers), ion channels (ligand/voltage gated, TRPM cation channels), and cytoskeleton reorganization proteins (ENAH, VASP, MYH10, EZR, CORO1A/1B).

are key determinants of the biological processes that help cancer cells evade destruction by immune attack. The presence of a group of interleukin and interferon cytokine receptors (ILs, IFNs), C-lectin/Fc/scavenger receptors, macrophage stimulating protein receptors (MST1R, CSF1R), HLA class I histocompatibility antigens (HLA-E, HLA-G), and B-cell/T-cell activating proteins (LYN, CD40, CD81) were indicative of SKBR3 triggers capable of eliciting a spectrum of innate, adaptive and inflammatory reactions that included among others cytokine production, positive regulation of innate immunity and defense responses, and elicitation of B-cell proliferation and T-cell killer cytotoxic effects (Fig. 6B, Cb). Many of these proteins were part of a group of 89 cluster of differentiation (CD) antigens with multiple roles not just in immune system processes but also in cell communication, signal transduction, adhesion, cell locomotion and transport (Fig. 7C). The detected CD antigens encompassed classical receptors, integrins and integrin binding proteins, cell adhesion molecules (CAMs), and disintegrin metalloproteinase domain-containing proteins, several of them being used in immunological profiling[43–49]. The relevance of this cell-surface category was underscored by participation in- or regulation of pathways such as MAPK, PI3K-AKT, ERK, JAK/STAT, ECM-receptor interactions, and B-cell/T-cell activation. The most abundant CDs included members of all protein categories, with high degree centrality nodes being represented primarily by adhesion proteins (ITGB1, ITGAV, ICAM1, CD9, and CD44) (Fig. 7C).

Cell adhesion and junction proteins. These molecules are often in a gray area of categorization because they participate not just in cell adhesion and locomotion, but also in a broad range of cancer-relevant processes including cell–cell and intracellular signaling, cell growth/proliferation/differentiation and death, secretion, angiogenesis, endocytosis, and chemotaxis, just to name a few (Fig. 6B, Cc)[50]. The CAM receptors that are involved in signaling belong to several families that include Ca-dependent cadherins, integrins, selectins, and Ca-independent immunoglobulin-like proteins. CAMs do not have catalytic domains, but engage in signaling by association with signaling adaptors and nonreceptor tyrosine kinases. Adhesion molecules that use non-enzymatic mechanisms for signal transduction have been, however, much less studied with respect to the details of signal recognition and transfer. All classical cell–matrix and cell–cell adhesion categories, as well as a number of additional CAMs and immunoglobulin-like CAMs, disintegrins, and MMPs were represented in the dataset. The integrins and selectins have been shown to be involved in various aspects of the metastatic process[51–53]. Integrins are transmembrane adhesion receptors that recognize a variety of cell-surface or extracellular matrix

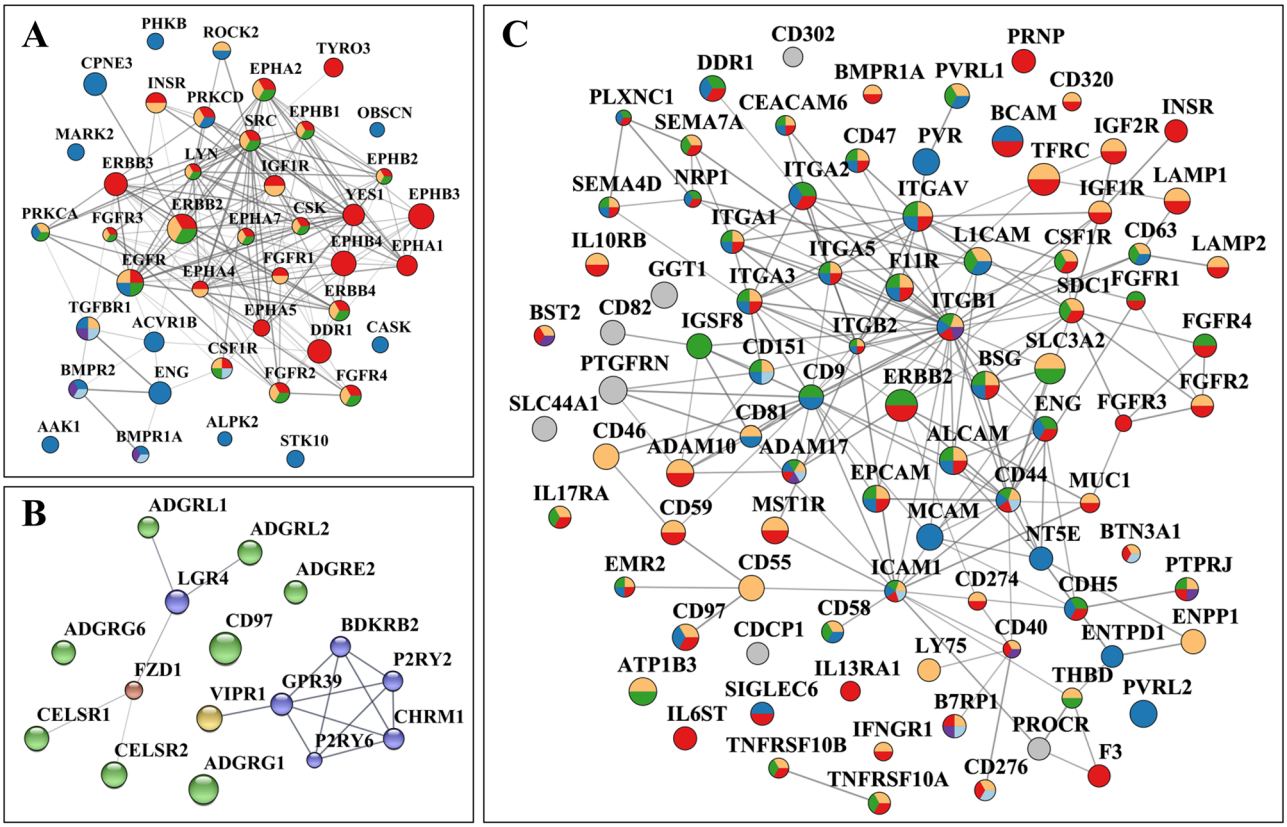


Figure 7. Protein–protein interaction networks of selected cell-membrane protein categories. (**A**) Receptor kinases: Red-Tyr kinase receptors (EGFRS, FGFRs, ephrins, DDR1), Blue-Ser/Thr kinase receptors (TGFBR1, BMPRs), Yellow-MAPK regulation, Green-ERK1/ERK2 regulation, Light blue-Cytokine-cytokine receptor signaling, Magenta-TGFB signaling. (**B**) G-protein coupled receptors: Blue-Class A Rhodopsin (P2RY2, P2RY6, CHRM1, BDKRB2, GPR39), Yellow-Class B1 Secretin (V1PR1), Green-Class B2 Adhesion (ADGRL1/2, ADGRG1/6, ADGRE2, CELSR1/2, CD97), Red-Class F Frizzled (FZD1). (**C**) CD antigens: Classical receptors (ERBB2, FGFRs, IGFRs, TNFRs, TFRC, ILs, IFNs), integrins and integrin binding proteins (ITGAV, ITGAs, ITGBs, semaphorins), CAMs (EPCAM, BCAM, ICAM1, L1CAM, MCAM, CDH5/CD144, mucins, nectins), and Disintegrin metalloproteinase domain-containing proteins (ADAM10/17); Yellow-Immune system process, Blue-Biological adhesion, Red-Cell communication, Green-Locomotion, Magenta-B cell activation, Light blue-T-cell activation. Notes: The PPI networks were generated with STRING and visualized with Cytoscape; node size is proportional to the total spectral counts that matched a protein, from < 10 (small) to > 10,000 (large).

(ECM) ligands (e.g., fibronectin, vitronectin, laminin, and collagen). The binding is mediated by the 24 α- and 9 β glycoprotein subunits that form noncovalent heterodimers (ITGA/ITGB) with binding activity modulated by various extracellular (e.g., Ca^{2+}/Mg^{2+}) or cell-type specific factors, and affinity for either cell–matrix or cell–cell interactions. Selectins, on the other hand, are adhesion molecules found on the surface of leukocytes, platelets and endothelial cells that through interactions with ligands expressed on the surface of cancer cells (mucins, glycosaminoglycans or sulfated glycolipids) facilitate metastatic spread within blood vessels[51]. Along with other receptors, many adhesion proteins that are tumor markers used in diagnostics and therapeutic decisions were detected[47], among which, IDH2, MUC16 and ITGAV/CD5 in high abundance (Fig. 8A,C). The cell–cell anchoring junctions were represented by adherence cadherin molecules associated with the actin filaments through cytoplasmic proteins such as catenins (e.g., CTNNA1), and desmosomal desmocollins (DSC2) and desmogleins (DSG2) bound to keratin intermediate filaments via plakin linkers. These types of junctions have roles in tissue morphogenesis, in maintaining tissue architecture and epithelial homeostasis, in cell proliferation and differentiation, and in facilitating cell movement[54–58].

Transporters and ion channels. This category included members of the entire range of transport (i.e., ABCs-ATP binding cassette transporters and SLCs-solute carriers superfamilies) and ion channel proteins (i.e., ligand and voltage gated), as well as other receptor/signaling, adhesion, MMP or peripheral proteins that either have transporter activity or are adaptors or accessory components of the transport complexes[59–63]. Altogether, the pool of membrane transport comprised 95 proteins with transporter activity and 22 with ion channel activity. Defective transport has been correlated with a variety of metabolic diseases, and also with cancer[59]. The group carries out, however, functions that are related not just to cellular transport and localization, but also to signaling/communication, development/differentiation, immune response, secretion, and adhesion (Fig. 6B,

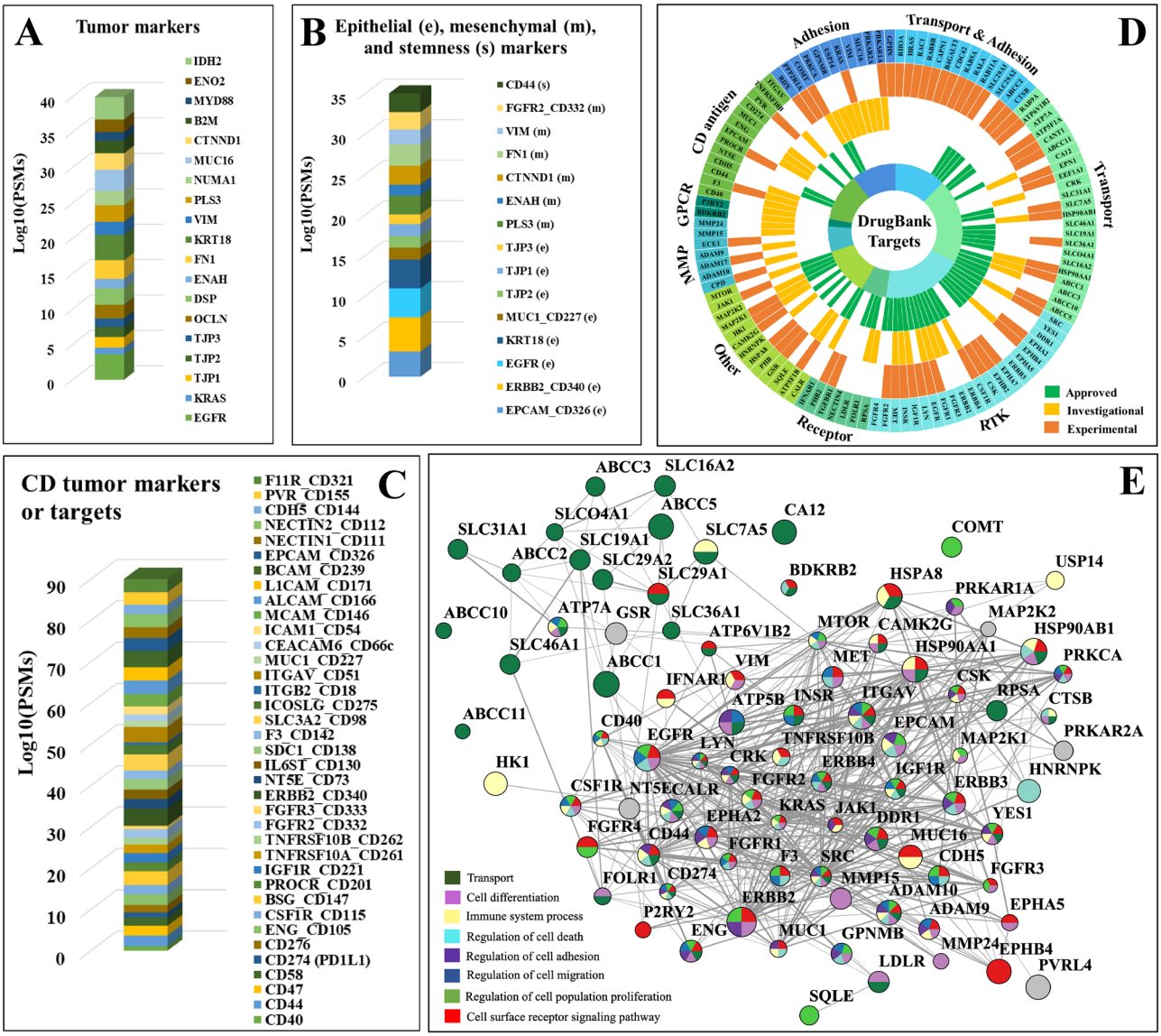


Figure 8. Cancer markers and drug targets detected in the SKBR3 cell-membrane proteome. (**A**) Tumor markers. (**B**) Epithelial (e), mesenchymal (m), and stemness (s) cancer markers. (**C**) CD Tumor markers and drug targets. (**D**) Sunburst chart representing receptors by functional categories identified in DrugBank as potential cancer therapeutic drug targets based on approved (green), investigational (yellow), and experimental (orange) status levels. (**E**) PPI network of the approved and investigational cancer drug targets. Note: Node size is proportional to the log10(SC) and edge thickness reflects the STRING interaction score.

Cd). The transport/adhesion proteins could be associated with functions related to cytoskeletal organization, the vesicle mediated transport with endocytosis, and the secretion proteins with immune responses (SLCs, ATPases, GTPases, TMEMs, TRPMs, MMPs, integrins, ORAI1, VAMP).

Drug targeting potential. Given the immense therapeutic opportunities offered by cell-membrane proteins[64–66], the detected RTKs, GPCRs, CD antigens, MMPs, adhesion and transport proteins were searched in the DrugBank database to identify prospects for targeting HER2+ breast cancer cells[64]. A total of 113 proteins were found in this pool, 56 for which approved and 48 for which investigational cancer drug targeting data existed (Fig. 8D and Supplementary Data S5). An additional category of 62 detected cell-surface proteins, not necessarily cancer-relevant, targeted by experimental drugs, was added to the list. The approved list included small molecule or monoclonal antibody drugs for various solid or liquid, early or advanced/metastatic cancers, and administered via chemo, targeted, combination, MDR, or topical therapeutic regimes. The RTKs, transporters and adhesion proteins represented the largest class of targets, followed by CD antigens, MMPs and GPCRs. In consensus, the targets of approved and investigational drugs clustered into two main categories, highlighting novel prospects for the development of anticancer drug cocktails (Fig. 8E). One category encompassed regula-

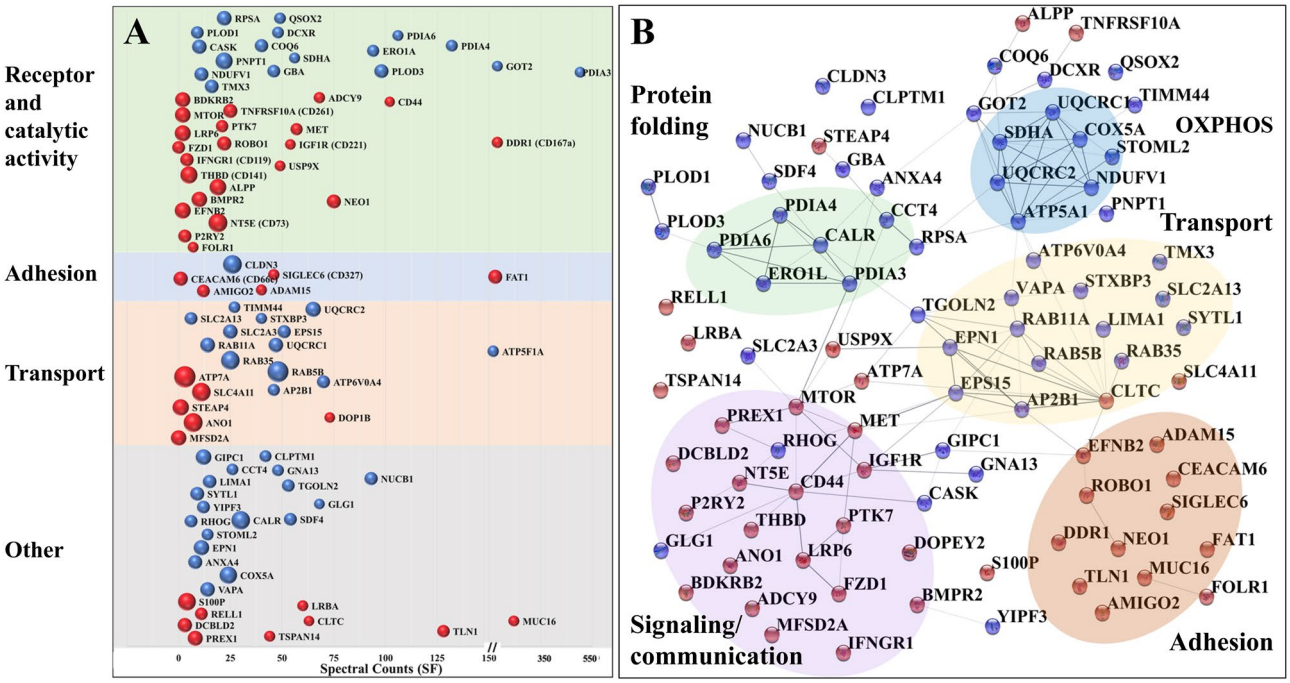


Figure 9. Proteins with changed abundance in the cell-membrane proteome. (**A**) Proteins with increased (red) and decreased (blue) abundance in ST vs. SF cells, categorized by function (y-axis) and spectral counts in SF cells (x-axis); the sphere size is proportional to the log2(FC) in protein spectral counts. (**B**) Cytoscape visualization of the STRING PPI network created with the proteins that displayed a change in abundance (same color scheme as in **A**).

tors of signal transduction, cell proliferation, differentiation, adhesion/migration, immune response and death, while the other solute transporters through the cell-membrane.

Changes in cell-surface protein abundance. The regulation of plasma membrane protein abundances represents a key biological process through which the cells mediate intercellular communication, preserve cellular homeostasis, or exert their function in response to environmental stimuli. This regulation can be slow when it involves protein de novo synthesis or degradation, or fast when it relies on rapid removal or insertion of proteins from and into the plasma membrane by making use of proteins stored in endosomal compartments or exocytic vesicles[67]. Given that harvesting of the cell-membrane proteome in this study occurred after prolonged exposure to SF or ST culture conditions, observable changes were expected to be representative of homeostatic processes rather than fast, transitory or cyclic events. Changes in cell-membrane protein expression between SF and ST cells were investigated after performing a two-step data normalization process, with the first, global normalization step being intended to account for data variability induced by sample processing, and the second step for the possible contamination of the cell-surface proteins by proteins from other cell compartments. Correction factors for global normalization were calculated by using the spectral counts of all proteins identified in each of the six samples under consideration (see Methods section), while for the second step, by using the spectral counts of only 10 endogenous cell-membrane proteins that met the following criteria: (a) the proteins were detected in every biological replicate of every isolation method, (b) the proteins were primarily associated with the cell-membrane but not with other cell fractions (i.e., nucleus, cytoplasm, ECM or secretome), and (c) the proteins had to have a transmembrane domain. The endogenous normalization proteins included solute carriers (SLC2A1, SLC3A2, SLC16A3), adhesion proteins (CDH5, PCDH1, DSG2, F11R), and receptors or proteins with catalytic activity (ITGB5, GNAS, SUSD2). The correction factors calculated by this approach ranged from 0.8 to 1.3 for the 1st set, and from 0.94 to 1.06 for the 2nd set. The method was applied to the dataset generated by the glycoprotein enrichment method that returned the largest number of total protein IDs, with the highest reproducibility and enrichment effectiveness (i.e., 852 proteins with 2 unique peptides/protein). Changes in abundance were observed for members of all protein categories (Fig. 9A), however, as the cell-surface protein labeling procedure induced the detachment and lysis of a small fraction of fragile serum-starved cells, only proteins for which a GO annotation of cell-membrane, cell-surface, peripheral cell-membrane, cell junction, or cell projection could be found were considered for the comparative analysis of the cell-surface glycoprotein ST vs. SF cells (i.e., 581 proteins). Supplementary Data S6 includes the list of analyzed proteins and the biological processes and pathways associated with the quantitative comparisons.

The biological processes that were up-regulated in the serum-deprived cells were represented by proteins involved in (a) OXPHOS and mitochondrial ATP synthesis coupled electron transport, (b) ER protein folding, negative regulation of unfolded protein response (UPR) and Ca^{2+} homeostasis, and (c) intracellular molecular and vesicle-mediated transport, localization, secretion, and cellular homeostasis (Fig. 9B). Intensified transport

and localization processes were presumably provoked by adaptation and survival responses to serum-deprived stress. Vesicle trafficking and membrane fusion included Rab11-mediated endocytic recycling, clathrin mediated endocytosis, and exocytosis, and appeared to target members of the ER protein folding machinery, Golgi apparatus and mitochondrial oxidative phosphorylation (OXPHOS). Calreticulin (CALR) and members of the protein disulphide isomerase (PDI) family are involved in maintaining cellular homeostasis by acting as chaperones that aid the folding of proteins destined for secretion in the ER. Their over expression has been observed in ER stress, but was also correlated with various cancerous cell states, while their translocation to the cell surface was associated with cancer progression and invasion[68–70]. The serum-stimulated cells, as expected, displayed two major categories of up-regulated processes represented by proteins involved on one hand in cell communication and cell-surface receptor signaling, and, on the other hand, in cell–cell adhesion and cell–matrix interactions, locomotion and migration.

PRM-MS was used to validate the observed changes in the abundance of a selected set of proteins identified in the cell membrane proteome. Supplementary Data S7 provides the PRM comparisons of SF to ST cells, for two biological sample replicates. Elevated proteins that were selected for validation were part of both SF (ATP5F1A, ATP5F1B, COX5A, UQCRC1, UQCRC2, SDHA, STOML2) and ST samples (MET, IGFR1, CD44, P2RY2). In addition, immunofluorescence (IF) microscopy was used to further explore the changes in the cell-surface abundance of ATP5F1A (elevated in SF) and P2RY2 (elevated in ST). The PRM data re-enforced the results obtained by DDA-MS analysis. ATP5F1A and P2RY2 were just above the FC ≥ 2 threshold in PSMs to be included in the list of proteins with changed abundance, and IF microscopy revealed similar trends. Based on the fluorescence intensity profiles taken across the cells, however, the changes were small, and observable only in some cells, but not in all (Supplementary Data S8). Given the importance of clarifying the mechanisms of altered metabolism in cells, further studies targeted to the study of mitochondrial protein re-localization to the cell surface and P2RY2/ATP activated pathways will be necessary.

Discussion

Overall, a comprehensive landscape of the cell-membrane proteome underscored both, the boundless opportunities for biological research, diagnostics and therapeutics (Figs. 8 and 9), and also the challenges posed by its dynamic and often transient composition that can be induced by protein shuffling between various cellular compartments. The RTKs formed highly interconnected PPI networks through which they control essential cellular functions (Fig. 7A). Aberrant signaling initiated by these receptors, due to changes in expression level or the presence of mutations, was linked to many diseases including not just to cancer, but also inflammation and metabolic disorders. The RTKs are highly mutated in many cancers[28], and represent promising tumor markers and/or drug targets (Fig. 8). Mutations that lead to gain of function, genomic amplification or chromosomal rearrangements are often responsible for abnormal activation, signaling, and uncontrolled cell proliferation. As a result, extensive efforts have been invested into the discovery of novel RTK drug targets. Not surprisingly, focus has been also recently placed on investigating the targeting potential of ephrin receptors, which are the largest sub-family of RTKs. In the SKBR3 dataset, the complex role of ephrins in cell communication, development and migration emerged from their PPIs with the RTKs (FGFRs, EGFRs, HMGBs, BMPRs), non-catalytic NOTCH, and the plexin receptors. The compounding impact of these receptors on abnormal cell behavior clearly offers further opportunities for effective drug targeting. PLXNB2, for example, promotes the phosphorylation of ERBB2 at Tyr 1248[16], a phosphorylation site that downstream affects biological processes related to cell cycle and growth, cytoskeleton organization, motility, apoptosis, and carcinogenesis[32]. Aberrant activation of NOTCH signaling in breast cancers, on the other hand, by either receptor overexpression or mutations, leads to uncontrolled cell proliferation and survival[33]. Downstream cross-talk between pathways contributes to an even greater extent to altered signaling processes. In the case of TGF-β receptors, for example, activation is controlled by interactions with other proteins and by various posttranslational modifications (e.g., phosphorylation, ubiquitylation, sumoylation, and neddylation). As a result, these receptors are capable of triggering downstream signaling processes via multiple pathways, including SMAD, ERK, JNK and p38MAPK[30].

It is worth emphasizing some additional cancer-supportive capabilities enabled by the ephrin/plexin receptors. The ephrin/plexin group was part of a larger functional cluster with roles in chemokine signaling and increased chemotaxis[34,35] that included not just well-known receptor/non-receptor signaling kinases (e.g., ERBB2, CSF1R, SRC, PRKCD), but also proteins with functionally diverse activities (i.e., proteins with roles in immune response, cell adhesion cell-ECM binding molecules, and transport). In tumors, an altered expression of chemokines is responsible for the recruitment of immune cells and for cellular processes that support angiogenesis, proliferation, cancer dissemination and metastasis[34]. As a result, chemokines and their cognate chemokine receptors evolved as valuable drug targets for the development of novel immunotherapeutic interventions[35].

An even larger family of drug targets is constituted by the GPCRs, due to their comprehensive role in a wide range of signaling processes and physiological conditions. However, despite the GPCRs being the largest family of cell-surface and also druggable receptors, only few cancer therapies target these GPCRs. Generally, the discovery of novel drugs for GPCR targets has been limited by a high degree of sequence homology between many GPCRs at the binding site of ligands, and the lack of a clarified structure for GPCRs that are hard to isolate, purify or crystallize[65]. In addition, the majority of aGPCRs lack an endogenous ligand and their mechanism of action is not fully understood. Nonetheless, emerging crosstalk activity between GPCRs and catalytic receptors (e.g., RTKs such as EGFR) has revealed novel signaling mechanisms with roles in cell proliferation and differentiation[41], the GPCRs being capable of initiating distinct MAPK signaling pathways via stimulation of ERK, JNK and p38MAPK[42]. Depending on the transduction mechanism, however, the GPCRs can have either an inhibiting or stimulating role on the downstream pathways, and the mechanistic details of RTK transactivation are yet to be explored[41]. To boost the discovery of novel therapeutic targets, the challenge of transactivation

studies is twofold, i.e., to demonstrate the presence of existing crosstalk interactions and to clarify the relevance of such crosstalk to disease.

Yet another documented category of tumor biomarkers or drug targets was represented by the CDs (Fig. 8A–C)[43–47]. Of particular interest was the presence of antigen immunological markers that define the epithelial, mesenchymal or stemness characteristics of cells (Fig. 8B). A group of 8 epithelial, 6 putative mesenchymal, and one stemness marker were present, with the ERBB2, EGFR, KRT and EpCAM epithelial markers being highly abundant on the cell-surface[43]. The presence of non-epithelial markers, however, indicated that the SKBR3 cells were a mixed population of differentiated epithelial cells and cells undergoing EMT with stemness characteristics[43]. It must be noted, though, that from the detectable peptide sequences it was not clear whether the mesenchymal or epithelial splice variants and protein isoforms of ENAH and FGFR2 were identified[43]. The presence of the CD44 stemness marker was also reflective of the metastatic characteristics of SKBR3 cells, while that of PDL1/CD274 (programmed cell death 1 ligand), a receptor ligand that blocks T-cell activation and that is upregulated by many tumor cells, of the ability to escape immune surveillance[48]. The PD1/PD1L1 receptor/ligand pair is the target of thousands of clinical trials that test immune checkpoint inhibitors[49]. Possible similar roles have been attributed to CD276, as well.

Cancer progression is also supported by adhesion, adhesion/receptor and junction molecules. Integrins have multiple and complex roles in this process, and distinct integrin expression patterns were used to predict survival and organ-specific metastases via tumor-derived exosome uptake[52]. When activated by the binding of matrix components, the integrins can engage catalytic receptors and co-operate in triggering intracellular signaling pathways that regulate cell growth, survival, proliferation and differentiation. Vice versa, signaling processes initiated by conventional receptors can alter the expression and ligand-binding properties of integrins. By mediating the interactions between the ECM and the actin cytoskeleton, via binding intracellular anchor proteins (α-actinin, talin, filamin, vinculin) and recruiting downstream focal adhesion kinase (FAK) and Src kinases, the integrins regulate cell shape, motility, and further, cell migration and invasion[53]. The central role of integrins and CAMs in signaling, immune recognition and cell migration was underscored, as also highlighted above, by the high degree centrality of their nodes in the PPI network of CD antigens (Fig. 7C). On the other hand, altered expression of gap and tight junction proteins, and damaged junction integrity or functionality, have been associated with inflammatory conditions, anchorage-independent growth, cancer invasion and survival, and growth at the metastatic site[57,58]. The loss of cell–cell anchoring junctions is considered, in fact, a prerequisite to epithelial-to-mesenchymal transition (EMT), migration and cancer invasion[55,56].

As the ABC transporters mediate the efflux of drugs, they have been associated with the development of multidrug-resistance (MDR) and failure of chemotherapies. The SLCs represent the 2nd largest family of cell-membrane proteins, after GPCRs, that facilitate the transport of amino acids, peptides, sugars, neurotransmitters, vitamins, metals, inorganic/organic ions and electrolytes[59,60]. Six ABC and 42 SLC transporters with symporter or antiporter activity were identified in the SKBR3 membrane proteome, mostly involved in ion transport, of which three were MDR-associated proteins (ABCC1/2/5/10) and two were relevant to drug uptake (SLCO4A1 and SLC22A18)[60,61]. Elevated ABCC1 levels have been found in many cancers and were associated with unfavorable outcomes. Moreover, among the 22 proteins with ion channel activity (seven voltage-gated and one ligand-gated), ORAI1, ANO1, STIM1, and PANX1 have been found to be involved in evasion of cancer cells from the primary tumor[62]. EMT was found to be associated with a remodeling of the Ca^{2+} signalosome[62], and voltage gated Na^{+} channels were found to be upregulated in breast cancer and to promote tissue invasion[63]. Several recent reviews have summarized the therapeutic potential of cell-membrane transport proteins, and highlighted the anti-tumor or anti-metastatic potential of channel inhibitors[62,63].

Altogether, the data exemplifies the framework that can inform the development of precision medicine therapeutic approaches. The broad context of the surfaceome and PPI maps (Fig. 8E) can enable the identification of aberrantly behaving signature proteins, with diverse functional role, laying the basis for combinatorial or network targeting approaches that can act synergistically, more effectively, and with less side effects and toxicity[66].

Metabolic rewiring of growing and proliferating cancer cells to use aerobic glycolysis for ATP production, instead of OXPHOS, is an established and extensively studied mechanism of energy production (the Warburg effect[71]). It has been recognized, however, that different types of cancer cells can use both OXPHOS and aerobic glycolysis for ATP production[72], and, as a result, OXPHOS inhibitors have been suggested for targeting metabolic processes in high OXPHOS cancers[72]. However, the behavior of cancer cells under nutrient-deprived conditions is not well understood, certain studies pointing toward reinforced OXPHOS activity in serum-deficient cells[73,74]. In our study, the increased abundance of a number of mitochondrial inner membrane respiratory chain proteins in the serum-deprived cells was also indicative of cells relying more heavily on OXPHOS for producing ATP (i.e., complex I-NDUFV1, complex II-SDHA, complex III-UQCRC1/UQCRC2, complex IV-COX5A, and complex V-ATP5A1). It was not clear, however, whether elevated OXPHOS and/or stress-induced protein re-localization to the plasma membrane was responsible for the increased abundance of these proteins in the cell-membrane fraction of the serum-starved cells. In whole cell extracts, differences in the abundance of mitochondrial proteins between SF and ST cells were not observed. Nonetheless, the presence and activity of the ATP synthase complex components and of other mitochondrial matrix proteins and OXPHOS complexes in the cell-membrane and lipid rafts has been described before[75,76]. The components of the F_1F_0ATP synthase complex on the surface of certain tumors has been associated with more aggressive, late stage metastatic cancers[77]-as was the case of the SKBR3 cells that were collected from a pleural effusion metastatic site. Cell-surface ATP synthase activity has been also associated with the synthesis of extracellular ATP, binding of various ligands, and purinergic signaling[76]. Extracellular ATP acts as an intercellular messenger that can interact with various cell-surface receptors, triggering, depending on conditions, cell death, proliferation, or various immune responses. When acting on P2RY receptors, such as the P2RY2 and P2RY6 detected on the surface of SKBR3 cells, ATP can activate ERK-MAPK, PI3K-AKT and survival pathways, or support EMT, invasiveness and metastatic spreading[78]. ATP targeting in

the tumor microenvironment has been attempted, therefore, as a cancer therapy in several clinical trials[78]. In a similar manner, it has been suggested that OXPHOS complexes in the plasma membrane represent a source of extracellular superoxide which can exert various regulating roles in cellular function[76]. Altogether, in an effort to develop targeted therapies, the altered mechanisms of glucose metabolism in cancer cells have come recently under much scrutiny[79].

MTOR signaling (represented by MTOR, IGF1R, FZD1, LRP6), which plays a central role in regulating cell growth and anabolic/catabolic metabolism, appeared to be activated by the presence of growth factors, hormones and nutrients from serum[80]. MTOR hyperactivation is integral to several oncogenic pathways (e.g., PI3K/AKT and MAPK)[80], and, similarly, IGF1R overexpression and signaling has been shown to be implicated in the regulation of survival and proliferation of many cancers[81]. Therefore, co-targeting PI3K, mTOR, and IGF1R proved to be effective in reducing tumor growth and decreasing cell migration and invasion[82]. In addition, targeting of the mesenchymal epithelial transition (MET) receptor Tyr kinase that is coded by the proto-oncogene *MET*, has also gained momentum, as its aberrant activation has been associated with a number of signaling pathways that promote cell survival, growth, proliferation, morphogenetic effects, and migration (e.g., PI3K/AKT, Ras/MAPK, JAK/STAT, SRC, Wnt/β-catenin)[83]. Relevant to note is that MTOR, IGF1R and MET have been also implicated in the activation of alternative pathways that drive resistance to therapeutic treatment with EGFR tyrosine kinase inhibitors (TKIs), resulting frequently in the recurrence of tumors[84]. Processes related to cell differentiation and immune system responses were also elicited in the stimulated cells by many of the same proteins. In addition, CD44, an adhesion receptor expressed on the surface of many cancer cells that mediates cell–cell interactions and cell migration, has been recognized for its multifunctional roles in survival, angiogenesis, metastasis and activation of immune responses and inflammation[21,85]. Altogether, in the presence of nutrients, these proteins were representative of key biological processes that support cancer cell proliferation, adhesion, migration and propensity for tissue invasion and metastasis.

Conclusions

The metastatic SKBR3 cell-membrane proteome revealed a broad and rich map of receptors, immune response, adhesion and transporter proteins that sustain cancer-cell interactions with the native or drug-altered tumor microenvironment. As evidenced by PPI networks, the concerted action of cell-membrane proteins exposed synergistic capabilities that nourish aberrant cell proliferation and metastatic potential not just through well-known signaling mechanisms, but through all functional roles of the surfaceome. Various aspects of cell growth, proliferation and differentiation were mainly orchestrated by proteins with catalytic activity, kinase receptors, plexins, and some CDs via growth-factor initiated signaling pathways. Several important groups of proteins with newly identified activities in cancer development included the families of metalloproteinases, nectins, ephrins, and bone morphogenetic proteins. Alterations in the abundance of certain cell-membrane proteins in response to serum withdrawal provided novel insights into how cancer cells may exploit metabolic mechanisms of energy production to sustain their proliferation. Further studies will be needed to confirm these findings. Cell migration, invasion and metastatic propensity were facilitated by proteins with roles in sustaining or regulating angiogenesis, cell–cell or cell-ECM interactions and EMT processes. Essentially, all cell-membrane protein categories, not just the CDs, contributed to mounting innate/adaptive or inflammatory immune responses, with ephrin and plexin receptors supporting chemokine signaling and chemotactic processes. The presence of receptor ligands with T-cell inhibitory functions, such as PDL1, pointed to abilities to stage immune escape. Propensity for altered drug uptake mechanisms and development of multi-drug resistance, mediated by solute carriers or ABC transporters, respectively, were also evident. The availability of a vast range of multi-functional cell-membrane proteins underscored, however, encouraging prospects for the development of more effective combination therapies that co-target proliferative, autocrine/survival, apoptotic, angiogenesis, and cell migratory pathways, as well as more subtle cancer checkpoint immunotherapies. Reassuring was the presence of a large number of immunological markers that reflected yet unexplored opportunities for cancer diagnosis, prognosis and assessment of recurrence after therapy.

Data availability

The mass spectrometry raw files were deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the following dataset identifiers: PXD028976, PXD028977, and PXD028978.

Received: 17 December 2021; Accepted: 7 June 2022
Published online: 27 June 2022

References

1. Watkins, E. J. Overview of breast cancer. *J. Am. Acad. PAs* **32**, 13–17 (2019).
2. Nahta, R. Novel therapies to overcome HER2 therapy resistance in breast cancer. In *Current Applications for Overcoming Resistance to Targeted Therapies* (ed. Szewczuk, M.R, Qorri, B., Sambi, M.) 191–221 (Springer, 2019).
3. Cooper, G. M. *The Cell: A Molecular Approach* 2nd edn. (Sinauer Associates, 2000).
4. Paul, A. I. *et al.* GPCRomics: GPCR expression in cancer cells and tumors identifies new, potential biomarkers and therapeutic targets. *Front. Pharmacol.* **9**, 1–11 (2018).
5. Mark, A. L. & Joseph, S. Cell signaling by receptor tyrosine kinases. *Cell* **141**, 1117–1134 (2010).
6. Gutierrez, A. N. & McDonald, P. H. GPCRs: Emerging anti-cancer drug targets. *Cell. Signal.* **41**, 65–74 (2018).
7. Ziegler, Y. S., Moresco, J. J., Tu, P. G., Yates, J. R. III. & Nardulli, A. M. Plasma membrane proteomics of human breast cancer cell lines identifies potential targets for breast cancer diagnosis and treatment. *PLoS ONE* **9**, 1–18 (2014).
8. Elschenbroich, S., Kim, Y., Medin, J. A. & Kislinger, T. Isolation of cell surface proteins for mass spectrometry-based proteomics. *Expert Rev. Proteomics* **7**, 141–154 (2010).

9. Kuhlmann, L., Cummins, E., Samudio, I. & Kislinger, T. Cell-surface proteomics for the identification of novel therapeutic targets in cancer. *Expert Rev. Proteomics* **15**, 259–275 (2018).
10. Li, Y. *et al.* Sensitive profiling of cell surface proteome by using an optimized biotinylation method. *J. Proteomics* **196**, 33–41 (2019).
11. Wollscheid, B. *et al.* Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat. Biotechnol.* **27**, 378–386 (2009).
12. Kalxdorf, M., Gade, S., Eberl, H. C. & Bantscheff, M. Monitoring cell-surface n-glycoproteome dynamics by quantitative proteomics reveals mechanistic insights into macrophage differentiation. *Mol. Cell. Proteomics* **16**, 770–785 (2017).
13. Bausch-Fluck, D. *et al.* A mass spectrometric-derived cell surface protein atlas. *PLoS ONE* **10**, 1–22 (2015).
14. Bausch-Fluck, D. *et al.* The in silico human surfaceome. *Proc. Natl. Acad. Sci. USA* **115**, E10988–E10997 (2018).
15. Ramilowski, J. A. *et al.* A draft network of ligand–receptor-mediated multicellular signalling in human. *Nat. Commun.* **6**, 1–12 (2015).
16. UniProt Consortium. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).
17. Thul, P. J. *et al.* A subcellular map of the human proteome. *Science* **356**, 1–12 (2017).
18. Uhlén, M. et al. Proteomics. Tissue-based map of the human proteome. *Science* **347,** 1260419:1–9 (2015).
19. Alexander, S. P. H. et al. The concise guide to PHARMACOLOGY 2019/20 *Br J Pharmacol.* **176,** S21-S141 (2019).
20. MacLean, B. *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966–968 (2010).
21. Stelzer, G. *et al.* The GeneCards suite: From gene data mining to disease genome sequence analyses. *Curr. Protoc. Bioinf.* **54**, 1–33 (2016).
22. Szklarczyk, D. *et al.* STRING v11: Protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).
23. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
24. Mauri, M., Elli, T., Caviglia, G., Uboldi, G., Azzi, M. RAWGraphs: A visualisation platform to create open outputs. in *CHItaly'17, Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter* 28:1–28:5 (Association for Computing Machinery, 2017).
25. Omasits, U., Ahrens, C. H., Müller, S. & Wollscheid, B. Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* **30**, 884–886 (2014).
26. Hörmann, K. *et al.* A surface biotinylation strategy for reproducible plasma membrane protein purification and tracking of genetic and drug-induced alterations. *J. Proteome Res.* **15**, 647–658 (2016).
27. Lazar, I. M., Deng, J., Ikenishi, F. & Lazar, A. C. Exploring the glycoproteomics landscape with advanced MS technologies. *Electrophoresis* **36**, 225–237 (2015).
28. Yang, X. & Lazar, I. M. XMAn: A Homo sapiens mutated-peptide database for MS analysis of cancerous cell states. *J. Proteome Res.* **13**, 5486–5495 (2014).
29. Jing, H., Song, J. & Zheng, J. Discoidin domain receptor 1: New star in cancer-targeted therapy and its complex role in breast carcinoma. *Oncol. Lett.* **15**, 3403–3408 (2018).
30. Heldin, C. H. & Moustakas, A. Signaling receptors for TGF-β family members. *Cold Spring Harbor Perspect. Biol.* **8**, 1–33 (2016).
31. Mosch, B., Reissenweber, B., Neuber, C. & Pietzsch, J. Eph receptors and ephrin ligands: important players in angiogenesis and tumor angiogenesis. *J. Oncol.* **2010**, 1–12 (2010).
32. Hornbeck, P. V. *et al.* PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43**, D512–D520 (2015).
33. Masuko, K. & Masaru, K. Precision medicine for human cancers with Notch signaling dysregulation. *Int. J. Mol. Med.* **45**, 279–297 (2020).
34. Jin, T., Xu, X. & Hereld, D. Chemotaxis, chemokine receptors and human disease. *Cytokine* **44**, 1–8 (2008).
35. Poeta, V. M., Massara, M., Capucetti, A. & Bonecchi, R. Chemokines and chemokine receptors: New targets for cancer immunotherapy. *Front. Immunol.* **10**, 1–10 (2019).
36. Baker, M. S. *et al.* Accelerating the search for the missing proteins in the human proteome. *Nat. Commun.* **8**, 1–13 (2017).
37. Bassilana, F., Nash, M. & Ludwig, M. G. Adhesion G protein-coupled receptors: Opportunities for drug discovery. *Nat. Rev. Drug Discov.* **18**, 869–884 (2019).
38. Aust, G., Zhu, D., Van Meir, E. G. & Xu, L. Adhesion GPCRs in tumorigenesis. *Handb. Exp. Pharmacol.* **234**, 369–396 (2016).
39. Gad, A. A. & Balenga, N. The emerging role of adhesion GPCRs in cancer. *ACS Pharmacol. Transl. Sci.* **3**, 29–42 (2020).
40. Placeta, M. *et al.* The G protein-coupled P2Y6 receptor promotes colorectal cancer tumorigenesis by inhibiting apoptosis. *Biochim. Biophys. Acta Mol. Basis Dis.* **1864**, 1539–1551 (2018).
41. Cattaneo, F. *et al.* Cell-surface receptors transactivation mediated by g protein-coupled receptors. *Int. J. Mol. Sci.* **15**, 19700–19728 (2014).
42. Goldsmith, Z. & Dhanasekaran, D. G. Protein regulation of MAPK networks. *Oncogene* **26**, 3122–3142 (2007).
43. Barriere, G. *et al.* Circulating tumor cells and epithelial, mesenchymal and stemness markers: Characterization of cell subpopulations. *Ann. Transl. Med.* **2**, 1–8 (2014).
44. Krawczyk, N. *et al.* Expression of stem cell and epithelial-mesenchymal transition markers in circulating tumor cells of breast cancer patients. *BioMed Res. Int.* **2014**, 1–11 (2014).
45. Ribatti, D., Tamma, R. & Annese, T. Epithelial–mesenchymal transition in cancer: A historical overview. *Transl. Oncol.* **13**, 1–9 (2020).
46. Liao, T. T. & Yang, M. H. Hybrid epithelial/mesenchymal state in cancer metastasis: Clinical significance and regulatory mechanisms. *Cells* **9**, 1–13 (2020).
47. National Cancer Institute. https://www.cancer.gov/about-cancer/diagnosis-staging/diagnosis/tumor-markers-list/ (2021).
48. Kim, D. W., Uemura, M. & Diab, A. Comprehensive review of PD1/L1 inhibition in metastatic solid tumors: Safety, efficacy and resistance. *J. Biomed. Sci.* **6**, 1–9 (2017).
49. Tang, J., Shalabi, A. & Hubbard-Lucey, V. M. Comprehensive analysis of the clinical immuno-oncology landscape. *Ann. Oncol.* **29**, 84–91 (2018).
50. Santos A.K. et al. The Role of Cell Adhesion, Cell Junctions, and Extracellular Matrix in Development and Carcinogenesis. in *Trends in Stem Cell Proliferation and Cancer Research* (ed. Resende, R. & Ulrich, H.) 13–49 (Springer, 2013).
51. Bendas, G. & Borsig, L. Cancer cell adhesion and metastasis: selectins, integrins, and the inhibitory potential of heparins. *Int. J. Cell Biol.* **2012**, 1–10 (2012).
52. Hoshino, A. *et al.* Tumour exosome integrins determine organotropic metastasis. *Nature* **527**, 329–335 (2015).
53. Alberts, B. *et al. Molecular Biology of the Cell* 4th edn. (Garland Science, 2002).
54. Aberle, H., Schwartz, H. & Kemler, R. Cadherin-catenin complex: protein interactions and their implications for cadherin function. *J. Cell. Biochem.* **61**, 514–523 (1996).
55. Csidgey, M. & Dawson, C. Desmosomes: A role in cancer?. *Br. J. Cancer* **96**, 1783–1787 (2007).
56. Gloushankova, N. A., Rubtsova, S. N., Zhitnyak, I. Y. Cadherin-mediated cell-cell interactions in normal and cancer cells. *Tissue Barriers* **5,** e1356900-1-15 (2017).
57. Aasen, T., Mesnil, M., Naus, C. C., Lampe, P. D. & Laird, D. W. Gap junctions and cancer: communicating for 50 years. *Nat. Rev. Cancer* **16**, 775–788 (2016).

58. Bhat, A. A. *et al.* Tight junction proteins and signaling pathways in cancer and inflammation: A functional crosstalk. *Front. Physiol.* **9**, 1–19 (2019).
59. Sahoo, S., Aurich, M. K., Jonsson, J. J. & Thiele, I. Membrane transporters in a human genome-scale metabolic knowledge base and their implications for disease. *Front. Physiol.* **5**, 1–24 (2014).
60. Lin, L., Yee, S. W., Kim, R. B. & Giacomini, K. M. SLC transporters as therapeutic targets: Emerging opportunities. *Nat. Rev. Drug Discov.* **14**, 543–560 (2015).
61. Keogh, J., Hagenbuch, B., Rynn, C., Stieger, B., Nicholls, G. Chapter 1: Membrane Transporters: Fundamentals, Function and Their Role in ADME. In *Drug Transporters: Volume 1: Role and Importance in ADME and Drug Development* (ed. Nicholls, G. & Youdim, K.) 1–56 (The Royal Society of Chemistry, 2016).
62. Klumpp, L., Sezgin, E. C., Eckert, F. & Huber, S. M. Ion channels in brain metastasis. *Int. J. Mol. Sci.* **17**, 1–14 (2016).
63. Almasi, S. & El Hiani, Y. Exploring the therapeutic potential of membrane transport proteins: Focus on cancer and chemoresistance. *Cancers* **12**, 1–31 (2020).
64. Wishart, D. S. et al. DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Res.* **46,** D1074–D1082 (2017).
65. Schalop, L. & Allen, J. *GPCRs, Desirable Therapeutic Targets in Oncology*. *Drug Discovery and Development* https://www.drugdiscoverytrends.com/gpcrs-desirable-therapeutic-targets-in-oncology/. 2017.
66. Csermely, P., Agoston, V. & Pongor, S. The efficiency of multi-target drugs: the network approach might help drug design. *Trends Pharmacol. Sci.* **26**, 178–182 (2005).
67. Royle, S. J. & Murrell-Lagnado, R. D. Constitutive cycling: a general mechanism to regulate cell surface proteins. *BioEssays* **25**, 39–46 (2002).
68. Laurindo, F. R., Pescatore, L. A. & Fernandes, D. Protein disulfide isomerase in redox cell signaling and homeostasis. *Free Radical Biol. Med.* **52**, 1954–1969 (2012).
69. Fucikova, J., Spisek, R., Kroemer, G. & Galluzzi, L. Calreticulin and cancer. *Cell Res.* **31**, 5–16 (2021).
70. Parakh, S. & Atkin, J. D. Novel roles for protein disulphide isomerase in disease states: A double edged sword?. *Front. Cell Dev. Biol.* **3**, 1–11 (2015).
71. Liberti, M. V. & Locasale, J. W. The Warburg effect: How does it benefit cancer cells?. *Trends Biochem. Sci.* **41**, 211–218 (2016).
72. Ashton, T. M., McKenna, W. G., Kunz-Schughart, L. A. & Higgins, G. F. Oxidative phosphorylation as an emerging target in cancer therapy. *Clin. Cancer Res.* **24**, 2482–2490 (2018).
73. Liu, Z. *et al.* Nutrient deprivation-related OXPHOS/glycolysis interconversion via HIF-1α/C-MYC pathway in U251 cells. *Tumor Biol.* **37**, 6661–6671 (2016).
74. Antico Arciuch, V. G., Elguero, M. E., Poderoso, J. J. & Carreras, M. C. Mitochondrial regulation of cell cycle and proliferation. *Antioxid. Redox Signal.* **16**, 1150–1180 (2012).
75. Moser, T. L. *et al.* Endothelial cell surface F1-FO ATP synthase is active in ATP synthesis and is inhibited by angiostatin. *Proc. Natl. Acad. Sci. USA* **98**, 6656–6661 (2001).
76. Kim, B. W. *et al.* Lipid raft proteome reveals that oxidative phosphorylation system is associated with the plasma membrane. *Expert Rev. Proteomics* **7**, 849–866 (2010).
77. Speransky, S. *et al.* A novel RNA aptamer identifies plasma membrane ATP synthase beta subunit as an early marker and therapeutic target in aggressive cancer. *Breast Cancer Res. Treat.* **176**, 271–289 (2019).
78. Vultaggio-Poma, V., Sarti, A. C. & Di Virgilio, F. Extracellular ATP: A feasible target for cancer therapy. *Cells* **9**, 1–22 (2020).
79. Hay, N. Reprogramming glucose metabolism in cancer: Can it be exploited for cancer therapy?. *Nat. Rev. Cancer* **16**, 635–649 (2016).
80. Saxton, R. A. & Sabatini, D. M. mTOR signaling in growth, metabolism, and disease. *Cell* **168**, 960–976 (2017).
81. Larsson, O., Girnita, A. & Girnita, L. Role of insulin-like growth factor 1 receptor signalling in cancer. *Br. J. Cancer* **92**, 2097–2101 (2005).
82. May, C. D. *et al.* Co-targeting PI3K, mTOR, and IGF1R with small molecule inhibitors for treating undifferentiated pleomorphic sarcoma. *Cancer Biol. Ther.* **18**, 816–826 (2017).
83. Zhang, Y. *et al.* Function of the c-Met receptor tyrosine kinase in carcinogenesis and associated therapeutic opportunities. *Mol. Cancer* **17**, 1–14 (2018).
84. Huang, L. & Fu, L. Mechanisms of resistance to EGFR tyrosine kinase inhibitors. *Acta Pharm. Sin. B* **5**, 390–401 (2015).
85. Senbanjo, L. T. & Chellaiah, M. A. CD44: a multifunctional cell surface adhesion receptor is a regulator of progression and metastasis of cancer cells. *Front. Cell Dev. Biol.* **5**, 1–6 (2017).

Acknowledgements

This work was supported by an award from the National Institute of General Medical Sciences (Grant No. 1R01GM121920) to IML. We thank Dr. Joshua Nicklay and Dr. John Veneski from Thermo Scientific for analyzing a preliminary set of cell extracts on a QExactive Plus Orbitrap mass spectrometer. We also thank Dr. Samy Lamouille and Christina Wheeler from the Fralin Biomedical Research Institute at VTC for providing support with the acquisition of confocal microscopy data.

Author contributions

A.K. performed the experiments, processed the data, and prepared the figures; I.M.L. and A.K. analyzed the data and wrote the manuscript; I.M.L. coordinated the work. Both authors reviewed and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-14418-0.

Correspondence and requests for materials should be addressed to I.M.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.