Mass Spectrometry - Based Comprehensive Proteomic Analyses of Biological Samples: from Wood Frogs Liver Tissues to Cancer-Derived Extracellular Vesicles

Yingxi Li

Thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the Master's degree in Chemistry

Department of Chemistry and Biomolecular Sciences
Faculty of Science
University of Ottawa

Abstract

Proteomic studies have rapidly advanced in recent years, revolutionizing the field of biology and biomedical research. Proteomics is the large-scale study of proteins, and it provides an understanding of protein structures, compositions, functions, protein-protein interactions, and their cellular activities. In this thesis, MS-based proteomic approaches were applied to two different biological systems: liver tissues from wood frogs and breast cancer (BC)-derived extracellular vesicles. Chapter 2 focuses on exploring the molecular adaptations of freeze-tolerant wood frogs. Wood frogs have excellent freeze tolerance and can survive environmental challenges such as limited water availability, low oxygen levels, and extremely low temperatures that can lead to whole-body freezing. A comprehensive proteomic analysis was conducted on frog liver tissue exposed to anoxia, dehydration, or freezing using a label-free LC-MS/MS proteomic approach. The quantitative analysis identified 87, 118, and 86 proteins that were significantly upregulated in the dehydrated, anoxic, and frozen groups, respectively, indicating their potential protective roles. The study also confirmed the upregulation of three enzymes: GST, ALDOA, and SORD. These enzymes exhibited significantly higher specific activity in the livers of the frozen and anoxic groups compared to the controls. Chapter 2 suggests that GST, ALDOA, and SORD may be involved in the freeze tolerance mechanism by contributing to cellular detoxification and energy metabolism. Chapter 3 focuses on identifying potential biomarkers from BC-derived EVs. BC is one of the most diagnosed malignancies among women and the second leading cause of cancer-related death in North America. Small membrane-derived extracellular vesicles (sEVs) have considerable potential as diagnostic agents and can be detected at the earliest stage of cancers. We aim to study global proteome in small EVs to find biomarkers for BC diagnosis, and three BC cell lines were examined (two cancerous and one non-cancerous cell line). Three types of proteomics approaches were conducted: quantitative, phosphoproteomics, and proteomics acetylation analysis. Our quantitative proteomics analyses and previously reported data strongly suggest that 11 identified enzymes may be potential candidates as biomarkers for the diagnosis of BC. Among the phosphoproteins, we validated four enzymes associated with cancer and present only in sEVs isolated from cancerous cell lines: ACLY, PFKM, SIRT1, and SIRT6. The protein acetylation study revealed that ALDOA, PGK1 and ENO showed a significantly higher specific enzymatic activity in cancerous cell lines in comparison to the non-cancerous cell line. Thus, proteomic

studies reveal that sEVs contain enzymes that could be interesting potential candidates for early BC diagnostics. Future validation of enzymes using both cancer cell lines and blood from BC patients remains to be determined.

Statement of Contribution

Chapter 1: MS-based Proteomic Study

Yingxi Li wrote all contents in chapter 1.

Chapter 2: Proteomic Study of Wood Frogs (*Rana sylvatica*)

The manuscript was published in Scientific Reports.

Proteomic Analysis of *Rana sylvatica* Reveals Differentially Expressed Proteins in Liver in Response to Anoxia, Dehydration or Freezing Stress

Yingxi Li, Zoran Minic, Nico Hüttmann, Abdullah Khraibah, Kenneth B. Storey, Maxim V. Berezovski

Yingxi Li designed the study, conducted the experiments, wrote the manuscript and analyzed the data.

Nico Hüttmann and Abdullah Khraibah performed bioinformatics analysis.

Dr. Maxim V. Berezovski designed the study, discussed experimental design, offered critical input, and funded the study.

Dr. Zoran Minic discussed experimental design, performed LC-MS/MS analysis, and offered critical input as part of the John L. Holmes Mass Spectrometry Facility.

Dr. Kenneth B. Storey discussed experimental design, provided biological samples, and offered critical input.

Chapter 3: Proteomic Study of Breast Cancer-derived Extracellular Vesicles

The manuscripts in this chapter have been published in Biomedicines.

Phosphoproteomic Analysis of Breast Cancer-Derived Small Extracellular Vesicles Reveals Disease-Specific Phosphorylated Enzymes

Zoran Minic, Nico Hüttmann, Suttinee Poolsup, Yingxi Li, Vanessa Susevski, Emil Zaripov and Maxim V. Berezovski

Yingxi Li performed small extracellular vesicles isolation, proteomic sample preparation and validation.

Dr. Zoran Minic designed the study, performed LC-MS/MS analysis, analyzed the data, and wrote the manuscript.

Nico Hüttmann performed bioinformatics analysis.

Suttinee Poolsup and Vanessa Susevski performed cell culture.

Dr. Maxim V. Berezovski discussed experimental design, offered critical input, and funded the study.

Lysine Acetylome of Breast Cancer-Derived Small Extracellular Vesicles Reveals Specific Acetylation Patterns for Metabolic Enzymes

Zoran Minic, Yingxi Li, Nico Hüttmann, Gurcharan K. Uppal, Rochelle D'Mello and Maxim V. Berezovski

Yingxi Li discussed experimental design, performed cell culture, small extracellular vesicles isolation, proteomic sample preparation and validation.

Dr. Zoran Minic designed the study, performed LC-MS/MS analysis, analyzed the data, and wrote the manuscript.

Nico Hüttmann performed bioinformatics analysis.

Gurcharan K. Uppal and Rochelle D'Mello performed cell culture and small extracellular vesicles isolation.

Dr. Maxim V. Berezovski discussed experimental design, offered critical input, and funded the study.

Proteomics Approaches for the Discovery of Potential Enzymatic Biomarkers for Early Diagnosis of Breast Cancer

Yingxi Li, Nico Hüttmann, Zoran Minic and Maxim V. Berezovski

Yingxi Li designed the study, analyzed the data and wrote the manuscript.

Nico Hüttmann performed bioinformatics analysis.

Dr. Zoran Minic discussed experimental design and offered critical input.

Dr. Maxim V. Berezovski discussed experimental design and funded the study.

Acknowledgements

I would like to express my deepest gratitude to my thesis advisor, Dr. Maxim V. Berezovski, for his unwavering support throughout this journey. Your guidance, encouragement, and insightful feedback were crucial in shaping this work. Thank you for always being available for discussions, providing thoughtful critiques, and pushing me to refine my ideas. Your belief in my abilities has inspired me to reach beyond what I thought was possible, and for that, I am profoundly grateful.

A special thanks also goes to Dr. Zoran Minic, whose continuous support and expertise in MS have been invaluable. I appreciate the time you invested in helping me, always offering practical solutions whenever I faced challenges. Your thoughtful advice and encouragement have guided me through difficult moments, making this journey smoother and more rewarding.

Lastly, I would like to extend my deepest gratitude to my parents. Supporting your only child to study abroad is never easy, but you have always provided me with love, encouragement, and everything I needed to pursue my dreams. Your sacrifices, constant reassurance, and belief in my abilities have given me the strength to overcome every challenge I encountered. I am endlessly grateful for your dedication and for always standing by my side, even from afar. This achievement is as much yours as it is mine.

Table of Contents

List of Abbreviations

ALDOA	Aldolase
BC	Breast cancer
CFE	Cell-free extract
CID	Collision-induced dissociation
ESI	Electrospray ionization
ENO	Enolase
EH	Epoxide hydrolase
FDR	False discovery rate
FASP	Filter aided sample preparation
FC	Fold change
GO	Gene ontology
GST	Glutathione S-transferase
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GSK3β	Glycogen synthase kinase-3beta
HO	Heme oxygenase
HCD	Higher-energy collision dissociation
HPLC	High-performance liquid chromatography
LFQ	Label free quantification
LDH	Lactate dehydrogenase
LC-MS/MS	Liquid chromatography coupled to tandem mass spectrometry
MS	Mass spectrometry
MALDI	Matrix-assisted laser desorption/ionization
NTA	Nanoparticle tracking analysis
DDM	n-Dodecyl-B-D-Maltoside
PGK1	Phosphoglycerate kinase
PTM	Post-translational modification
PCA	Principal component analysis
PKM	Pyruvate kinase M1/2
ROS	Reactive oxygen species
sEVs	Small extracellular vesicles
SORD	Sorbitol dehydrogenase
SULT	Sulfotransferases
TOF	Time of flight

List of Figures

List of Tables

Chapter 1: MS-Based Proteomics

1.1 Introduction of Proteins

Proteins are large, complex molecules that play many critical roles in the body. They are essential for the structure, function, and regulation of the body's tissues and organs. Composed of long chains of amino acids, proteins are involved in everything from structural support to catalyzing metabolic reactions. They are encoded by genes and are synthesized within cells through the process of translation, where the sequence of nucleotides in messenger RNA (mRNA) is decoded to build a specific sequence of amino acids, ultimately folding into a functional three-dimensional structure. With 20 different standard amino acids available, the diversity of protein functions is achieved through variations in the sequence and length of these chains [1].

The structure of proteins is hierarchical and is classified into four levels: primary, secondary, tertiary, and quaternary. The primary structure refers to the linear sequence of amino acids in the protein chain, which determines how the protein will fold and function. Secondary structures, such as alpha helices and beta sheets, are local folded structures that form within a polypeptide due to interactions between the backbone atoms. The tertiary structure is the overall three-dimensional shape of a single polypeptide chain, driven by interactions between side chains of the amino acids. Some proteins consist of multiple polypeptide chains, and their arrangement in a multi-subunit complex defines the quaternary structure. Each level of structure is crucial in determining a protein's specific function [2].

Post-translational modifications (PTMs) are chemical changes that can occur to proteins after they have been synthesized. These modifications are crucial for the regulation of protein activity, stability, localization, and interaction with other cellular molecules. PTMs can include the addition of functional groups, such as phosphate (phosphorylation), acetyl (acetylation), or methyl groups (methylation), which can alter protein function or activity. Other modifications, such as ubiquitination, target proteins for degradation, while glycosylation, the addition of sugar molecules, can affect protein folding, stability, and cell signaling. PTMs provide a mechanism for cells to dynamically and reversibly regulate protein function in response to various cellular signals and environmental conditions [3].

The diversity of protein function is further enhanced by their ability to form complexes and interact with other biomolecules, such as DNA, RNA, lipids, and carbohydrates. Proteins can

function as enzymes, catalyzing biochemical reactions; as structural components, providing support and shape to cells and tissues; as transporters, moving molecules across cell membranes; and as receptors, detecting and responding to chemical signals. Proteins are also essential for the immune response, with antibodies being specialized proteins that recognize and neutralize foreign invaders such as bacteria and viruses. Additionally, proteins play a role in cell signaling pathways, where they act as messengers that transmit signals from the cell surface to the interior, initiating specific cellular responses.

Understanding the structure and function of proteins, as well as the effects of post-translational modifications, is crucial in many fields of biology and medicine. Misfolding of proteins or aberrant PTMs can lead to diseases such as Alzheimer's, Parkinson's, and cancer [3]. Consequently, studying proteins and their modifications can provide insights into the molecular basis of diseases and possibly lead to the development of targeted therapies. The study of proteins is a central focus in biochemistry and molecular biology, and advances in proteomics, which involves the large-scale study of proteins, continue to enhance our understanding of cellular mechanisms and the complexity of life.

1.2 Non-MS-based Methods to Study Proteins

To understand their complex functions and mechanisms, scientists have developed a variety of techniques to study proteins. These methods encompass structural analysis, functional assays, and interaction studies, each providing unique insights into protein biology.

X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and cryo-electron microscopy (cryo-EM) are methods that are commonly used in understanding the structure of proteins. X-ray crystallography is a powerful technique for determining the three-dimensional structures of proteins at atomic resolution. This method involves crystallizing a purified protein and exposing the crystal to X-ray beams. The diffraction pattern produced by the X-rays is analyzed to generate an electron density map, which reveals the arrangement of atoms within the protein. X-ray crystallography has been instrumental in elucidating the structures of numerous proteins, including enzymes, receptors, and large macromolecular complexes. However, the requirement for crystallization can be a significant limitation, as not all proteins readily form crystals [4].

NMR spectroscopy is another important method for studying protein structure and dynamics. Unlike X-ray crystallography, NMR does not require crystallization, allowing proteins to be examined in solution, which more closely resembles their natural state. NMR exploits the magnetic properties of certain atomic nuclei to provide information about the local environment and spatial arrangement of atoms within a protein. This technique is particularly useful for studying small to medium-sized proteins and for investigating protein dynamics, folding, and interactions. However, the complexity of data analysis and size limitations can pose challenges in NMR studies [5].

Cryo-electron microscopy (cryo-EM) has revolutionized structural biology by enabling the visualization of large protein complexes and membrane proteins at near-atomic resolution. In cryo-EM, proteins are rapidly frozen in a thin layer of vitreous ice, preserving their native structure without the need for crystallization. High-resolution images are captured using an electron microscope, and computational algorithms reconstruct three-dimensional models from these images. Cryo-EM is particularly valuable for studying proteins that are difficult to crystallize or are too large for NMR analysis. Recent advances in cryo-EM technology have significantly improved its resolution and broadened its applications [6].

Functional analysis is another important aspect of proteins study. Enzyme assays are fundamental tools for studying the catalytic activity of proteins. These assays measure the rate at which an enzyme catalyzes a specific biochemical reaction, providing insights into enzyme kinetics, substrate specificity, and regulatory mechanisms. Enzyme assays can be conducted using various detection methods, including spectrophotometry, fluorometry, and radiometry, depending on the nature of the reaction and available substrates. Kinetic parameters such as the Michaelis-Menten constant (Km) and the maximum reaction velocity (Vmax) can be determined, helping to characterize the enzyme's efficiency [7].

Fluorescence spectroscopy is a versatile technique used to study protein structure, dynamics, and interactions. Proteins can be intrinsically fluorescent or labeled with fluorescent probes to monitor changes in conformation, folding, and binding events. Techniques such as Förster resonance energy transfer (FRET) and fluorescence anisotropy provide information on protein-protein interactions and conformational changes. Fluorescence spectroscopy is widely used in real-time assays to investigate protein dynamics in living cells, offering insights into biological processes such as signaling and trafficking [8].

Circular dichroism (CD) spectroscopy is a technique that measures the differential absorption of left- and right-handed circularly polarized light by chiral molecules, such as proteins. CD spectroscopy provides information about the secondary structure content of proteins, including alpha helices, beta sheets, and random coils. This method is useful for studying protein folding, conformational changes, and interactions with ligands or other proteins. CD spectroscopy is particularly valuable for characterizing structural changes in response to environmental conditions such as temperature, pH, and ionic strength [9].

In addition to structural and functional analysis, protein-protein interaction studies are essential for understanding the complex networks and pathways that underpin cellular functions and biological processes. Proteins rarely act in isolation; instead, they interact with other proteins to form complexes that carry out specific tasks such as signal transduction, metabolic pathways, and structural assembly. By studying these interactions, researchers can elucidate the mechanisms by which proteins work together to regulate cellular activities, identify potential points of dysregulation in diseases, and discover new targets for therapeutic intervention. Understanding protein-protein interactions also provides insights into the evolution of protein functions and the development of novel biomolecular tools and drugs, thereby advancing both basic biological research and applied biomedical sciences [10].

The yeast two-hybrid (Y2H) system is a genetic method used to identify protein-protein interactions. It relies on the reconstitution of a transcription factor when two proteins of interest interact in the yeast nucleus, leading to the activation of a reporter gene. The Y2H system is widely used for mapping interaction networks and identifying novel protein partners. However, the method can produce false positives and negatives, and interactions detected in yeast may not always reflect physiological conditions in higher organisms [11].

Co-immunoprecipitation (Co-IP) is a biochemical technique used to detect physical interactions between proteins in a complex. It involves using an antibody specific to a target protein to precipitate the protein and its interacting partners from a cell lysate. The precipitated proteins are then identified using techniques such as Western blotting or mass spectrometry. Co-IP is a powerful method for confirming interactions observed in other assays and for studying post-translational modifications. The technique is limited by the availability and specificity of antibodies and may not detect transient or weak interactions [12].

Surface plasmon resonance (SPR) is a label-free technique for measuring real-time interactions between biomolecules. It detects changes in the refractive index near a sensor surface to quantify binding kinetics and affinity. SPR is used to study a wide range of interactions, including protein-protein, protein-DNA, and protein-ligand interactions. The technique provides detailed information on binding kinetics, such as association and dissociation rates, and can be used to screen potential drug candidates. However, the requirement for immobilization on a sensor chip can affect interaction dynamics, and SPR may be less effective for low-affinity interactions [12].

Other advanced techniques, such as protein microarrays and single-molecule techniques are also used in protein studies. Protein microarrays are high-throughput platforms used to study protein expression, interactions, and functions on a large scale. These arrays consist of thousands of proteins immobilized on a solid surface, allowing simultaneous analysis of multiple protein features. Protein microarrays are used in proteomics research to identify disease biomarkers, study protein-DNA interactions, and profile immune responses. They offer significant advantages in terms of speed and scalability but require careful optimization and validation to ensure reliable results.

Single-molecule techniques, such as atomic force microscopy (AFM) and single-molecule fluorescence spectroscopy, provide detailed insights into protein behavior at the individual molecule level. These techniques allow researchers to study protein folding, dynamics, and interactions in real-time, revealing information that is often obscured in ensemble measurements. Single-molecule studies have advanced our understanding of complex biological processes, such as enzyme catalysis and molecular motors, by uncovering heterogeneity and transient states that are crucial for function [13].

1.3 MS-based Proteomics Overview

The field of "omics" represents a collective group of disciplines dedicated to the comprehensive study of various biological molecules and their roles within living organisms. These disciplines include genomics (the study of genomes), transcriptomics (the study of RNA transcripts), metabolomics (the study of metabolites), and proteomics (the study of proteins), among others [14]. Each "omics" field provides a different perspective on the complex biological processes that govern life. By integrating data from multiple omics studies, researchers can obtain

a holistic understanding of biological systems, uncovering how different molecules interact and contribute to health and disease [15].

The concept of proteomics began to take shape with the development of two-dimensional gel electrophoresis in the 1970s. In 1975, O'Farrell, Klose, and Scheele introduced the two-dimensional gel technique to map proteins from Escherichia coli, mouse, and guinea pig, respectively [16-18]. This technique allowed scientists to separate and visualize proteins based on their isoelectric point and molecular weight, providing a way to analyze complex protein mixtures. However, protein identification remains unclear [19]. Their functions and interactions were also undiscovered [19]. Starting from the 1990s, the advances of mass spectrometry and bioinformatics have propelled the development of proteomics, enabling the detailed analysis of protein expression, protein functions, post-translational modifications, and protein-protein interactions [20].

Mass spectrometry (MS) is an analytical technique used to measure the mass-to-charge ratio of ions, enabling the identification and quantification of molecules in a sample. The analysis is achieved with three primary steps: ionization, mass analysis, and detection [21]. During ionization, the sample is ionized by an ion source to produce charged molecules. Common ionization methods include electrospray ionization (ESI), electron impact (EI), matrix assisted laser desorption ionization (MALDI), and etc [21]. After ionization, the charged molecules will enter the mass analyzer for mass analysis. In this step, the ions are separated based on their mass-to-charge ratio (m/z). Common types of mass analyzers include orbitrap, time-of-flight (TOF), quadrupole and etc [21]. Once the molecules are separated, they are sent to the detector, and the resulting data is used to generate a mass spectrum for further analysis [21].

The MS instrument that we used in our studies was the orbitrap fusion tribird mass spectrometer (Thermo Fisher Scientific) and ESI was the ion source. It applies a high voltage to a liquid sample as it flows through a fine capillary, forming a mist of charged droplets. As the solvent evaporates, the droplets become smaller, increasing the charge density and resulting in the formation of gas-phase ions [22]. Compared to other ionization sources, ESI can produce multiple charged ions which enables the analysis of high molecular weight proteins. It can efficiently ionize large, polar molecules such as peptides and proteins without extensive fragmentation [22]. This capability is essential for proteomics, where maintaining the integrity of large biomolecules is important for accurate analysis. In addition, ESI can couple with liquid chromatography (LC) easily. LC is an analytical technique that is used for the separation and analysis of complex

mixtures. In proteomics, the peptide mixture passes through a column packed with a stationary phase under high pressure. As peptides travel through the column, they interact differently with the stationary phase based on their chemical properties, then they are eluted from the column at different retention time, and thus separated. In an LC-MS system, the separated peptides would enter ESI and continued for MS analysis. In our studies, we utilized a nano LC-MS system for proteomic analysis. Nano LC works similar to conventional LC but it increases the sensitivity, because nano LC uses narrower columns and lower flow rate, which increases the concentration of peptides when entering ESI. Increase of sensitivity allows a better detection of low abundance proteins [23]. Also, nano LC can minimize the injected sample volume [23]. This is very important for proteomics especially when we work with limited biological samples. The nano LC-MS system that was used in our studies offers numerous advantages in proteomics, including enhanced sensitivity, superior separation efficiency, improved ionization efficiency, accurate quantification, and comprehensive proteome coverage [24]. These benefits make nano LC-MS an essential tool for detailed and accurate proteomic analysis, providing deeper insights into protein expression, function, and regulation.

In general, mass spectrometry is a powerful analytical technique widely used in proteomics to identify and quantify proteins in complex biological samples. Proteomic MS generates large datasets that require sophisticated statistical and bioinformatic analyses to extract meaningful biological insights. These analyses involve several key steps, including data preprocessing, statistical modeling, protein identification and quantification, and functional interpretation. By integrating statistical methods with bioinformatic tools, researchers can decipher complex proteomic data to understand cellular processes, identify biomarkers, and explore disease mechanisms.

Before statistical and bioinformatic analyses can begin, proteomic MS data must be preprocessed to ensure accuracy and reliability. Data preprocessing involves several critical steps, including peak detection, noise reduction, alignment, and normalization. Peak detection identifies the relevant ion signals corresponding to peptides, while noise reduction removes background signals that can interfere with accurate analysis. Alignment is crucial for comparing multiple MS runs, ensuring that peaks are consistently identified across samples. Normalization methods, such as total ion current normalization or variance stabilization, are applied to correct systematic biases and ensure comparability across samples.

The identification and quantification of proteins from MS data are fundamental steps in proteomic analysis. Protein identification typically involves matching experimental MS spectra to theoretical spectra generated from protein sequence databases. This process is facilitated by search algorithms such as SEQUEST, Mascot, and MaxQuant, which use scoring functions to evaluate the similarity between experimental and theoretical spectra. Accurate protein identification relies on comprehensive and up-to-date protein databases, as well as stringent criteria to minimize false positives.

Protein quantification can be achieved using label-free or labeled approaches. Label-free quantification methods, such as spectral counting and intensity-based approaches, directly measure the abundance of peptides in the sample. In contrast, labeled approaches, such as isotope-coded affinity tags (ICAT), tandem mass tags (TMT), and stable isotope labeling by amino acids in cell culture (SILAC), introduce isotopic labels into peptides to facilitate relative quantification [25]. Each quantification method has its advantages and limitations, and the choice of approach depends on the specific experimental design and research objectives.

Statistical analysis plays a crucial role in identifying differentially expressed proteins between experimental conditions. Common statistical methods include t-tests, analysis of variance (ANOVA), and linear models, which assess the significance of changes in protein abundance. These analyses must account for multiple hypothesis testing, as the large number of proteins analyzed can lead to an increased risk of false discoveries [26]. To address this issue, correction methods such as the Benjamini-Hochberg false discovery rate (FDR) procedure or Bonferroni correction are applied to control for false positives [27,28].

The ultimate goal of proteomic MS analyses is to derive meaningful biological insights from the data. Bioinformatic tools and databases, such as Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Reactome, are used to annotate proteins with functional information and identify enriched biological pathways. Pathway analysis helps to contextualize the findings by highlighting the cellular processes and networks affected by differential protein expression [29].

Functional enrichment analysis involves comparing the observed frequency of specific biological categories among differentially expressed proteins to the expected frequency based on a reference set. Statistical tests, such as Fisher's exact test or hypergeometric test, are used to

determine the significance of enrichment, identifying key pathways and biological processes that are altered in the experimental conditions [29].

Proteomic data can be further explored through network analysis, which examines the interactions and relationships between proteins. Protein-protein interaction networks provide insights into the functional organization of the proteome and help identify key regulators or hubs that play central roles in cellular processes [12]. Network-based approaches, such as cluster analysis and community detection, can reveal functional modules and pathways that are perturbed in disease states 30. Integration of proteomic data with other omics data, such as genomics, transcriptomics, and metabolomics, enhances the understanding of biological systems by providing a more comprehensive view of cellular processes. Multi-omics integration involves the use of bioinformatic frameworks and statistical models to combine data from different sources, uncovering synergistic interactions and regulatory mechanisms. This holistic approach is particularly valuable for elucidating complex diseases, where multiple molecular layers contribute to pathogenesis.

Despite the advancements in statistical and bioinformatic analyses of proteomic MS data, several challenges remain. The complexity and variability of biological samples, along with the technical limitations of MS, can introduce biases and uncertainties in data interpretation. Standardization of experimental protocols, data processing workflows, and reporting formats is essential to ensure reproducibility and comparability across studies. The continued development of advanced computational tools and algorithms will enhance the capacity to analyze large-scale proteomic datasets and integrate them with other omics data. Emerging technologies, such as single-cell proteomics and spatial proteomics, present new opportunities to explore protein expression and function with unprecedented resolution and context. As these technologies evolve, the integration of proteomic data with clinical and phenotypic information will be crucial for translating proteomic discoveries into clinical applications and personalized medicine.

1.4 Conclusions

Overall, mass spectrometry has emerged as the leading technology for proteomics due to its unparalleled ability to analyze complex biological samples with high sensitivity and specificity. MS can identify and quantify thousands of proteins in a single experiment, providing comprehensive coverage of the proteome. This capability is crucial for understanding the dynamic

nature of proteins, including their post-translational modifications and interactions. The versatility of MS allows for the analysis of various sample types, from simple mixtures to complex tissues, enabling researchers to study proteins in diverse biological contexts. Additionally, advancements in MS instrumentation and techniques, such as tandem mass spectrometry (MS/MS) and high-resolution mass analyzers, have significantly improved the accuracy and resolution of proteomic analyses, making it possible to detect even low-abundance proteins that are critical for understanding cellular processes and disease mechanisms.

Another reason MS is the leading technology in proteomics is its ability to provide quantitative data, which is essential for studying changes in protein expression and modification in response to biological stimuli or disease states. MS-based proteomics can employ both label-free and labeling techniques to achieve precise quantification of proteins across multiple samples or conditions. This quantitative capability is vital for biomarker discovery, drug development, and systems biology research, where understanding the relative abundance and dynamics of proteins is key. Moreover, the integration of MS data with bioinformatics tools and databases allows for the comprehensive interpretation of proteomic data, facilitating the identification of biological pathways and networks involved in various physiological and pathological processes. Together, these attributes make MS an indispensable tool in proteomics research, driving discoveries in biology and medicine.

1.5 References – Chapter 1

1 Rajendran, A., Endo, M. & Sugiyama, H. in *Advances in Protein Chemistry and Structural Biology* Vol. 87 (eds Christo Christov & Tatyana Karabencheva-Christova) 5-55 (Academic Press, 2012).

2 Sun, P. D., Foster, C. E. & Boyington, J. C. Overview of protein structural and functional folds. *Curr Protoc Protein Sci* **Chapter 17**, Unit 17.11, doi:10.1002/0471140864.ps1701s35 (2004).

3 Zhong, Q. *et al.* Protein posttranslational modifications in health and diseases: Functions, regulatory mechanisms, and therapeutic implications. *MedComm (2020)* **4**, e261, doi:10.1002/mco2.261 (2023).

4 Dessau, M. A. & Modis, Y. Protein crystallization for X-ray crystallography. *J Vis Exp*, doi:10.3791/2285 (2011).

5 Gronenborn, A. M. & Polenova, T. Introduction: Biomolecular NMR Spectroscopy. *Chem Rev* **122**, 9265-9266, doi:10.1021/acs.chemrev.2c00142 (2022).

6 Danev, R., Yanagisawa, H. & Kikkawa, M. Cryo-Electron Microscopy Methodology: Current Aspects and Future Directions. *Trends Biochem Sci* **44**, 837-848, doi:10.1016/j.tibs.2019.04.008 (2019).

7 Bisswanger, H. Enzyme assays. *Perspectives in Science* **1**, 41-55, doi:https://doi.org/10.1016/j.pisc.2014.02.005 (2014).

8 Basak, S. & Chattopadhyay, K. Studies of protein folding and dynamics using single molecule fluorescence spectroscopy. *Phys Chem Chem Phys* **16**, 11139-11149, doi:10.1039/c3cp55219e (2014).
9 Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochim Biophys Acta* **1751**, 119-139, doi:10.1016/j.bbapap.2005.06.005 (2005).
10 Lu, H. *et al.* Recent advances in the development of protein–protein interactions modulators: mechanisms and clinical trials. *Signal Transduction and Targeted Therapy* **5**, 213, doi:10.1038/s41392-020-00315-3 (2020).
11 Chen, Q. & Wei, T. Membrane and Nuclear Yeast Two-Hybrid Systems. *Methods Mol Biol* **2400**, 93-104, doi:10.1007/978-1-0716-1835-6_10 (2022).
12 Lin, J. S. & Lai, E. M. Protein-Protein Interactions: Co-Immunoprecipitation. *Methods Mol Biol* **1615**, 211-219, doi:10.1007/978-1-4939-7033-9_17 (2017).
13 Hill, F. R., Monachino, E. & van Oijen, A. M. The more the merrier: high-throughput single-molecule techniques. *Biochem Soc Trans* **45**, 759-769, doi:10.1042/bst20160137 (2017).
14 Dai, X. & Shen, L. Advances and Trends in Omics Technology Development. *Frontiers in Medicine* **9**, doi:10.3389/fmed.2022.911861 (2022).
15 Karczewski, K. J. & Snyder, M. P. Integrative omics for health and disease. *Nature Reviews Genetics* **19**, 299-310, doi:10.1038/nrg.2018.4 (2018).
16 Klose, J. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* **26**, 231-243, doi:10.1007/bf00281458 (1975).
17 O'Farrell, P. H. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* **250**, 4007-4021 (1975).
18 Scheele, G. A. Two-dimensional gel analysis of soluble proteins. Charaterization of guinea pig exocrine pancreatic proteins. *J Biol Chem* **250**, 5375-5385 (1975).
19 Graves, P. R. & Haystead, T. A. Molecular biologist's guide to proteomics. *Microbiol Mol Biol Rev* **66**, 39-63; table of contents, doi:10.1128/mmbr.66.1.39-63.2002 (2002).
20 Bradshaw, R. A. On the development of proteomics: a brief history. *Australian Journal of Chemistry* **76**, 418-428, doi:https://doi.org/10.1071/CH23012 (2023).
21 Vargas Medina, D. A., Maciel, E. V. S. & Lanças, F. M. in *Liquid Chromatography (Third Edition)* Vol. 1 (eds Salvatore Fanali *et al.*) 679-706 (Elsevier, 2023).
22 Ho, C. S. *et al.* Electrospray ionisation mass spectrometry: principles and clinical applications. *Clin Biochem Rev* **24**, 3-12 (2003).
23 Wilson, S. R., Vehus, T., Berg, H. S. & Lundanes, E. Nano-LC in proteomics: recent advances and approaches. *Bioanalysis* **7**, 1799-1815, doi:10.4155/bio.15.92 (2015).
24 Gaspari, M. & Cuda, G. Nano LC-MS/MS: A robust setup for proteomic analysis. *Methods in molecular biology (Clifton, N.J.)* **790**, 115-126, doi:10.1007/978-1-61779-319-6_9 (2011).
25 Anand, S., Samuel, M., Ang, C. S., Keerthikumar, S. & Mathivanan, S. Label-Based and Label-Free Strategies for Protein Quantitation. *Methods Mol Biol* **1549**, 31-43, doi:10.1007/978-1-4939-6740-7_4 (2017).
26 Lualdi, M. & Fasano, M. Statistical analysis of proteomics data: A review on feature selection. *Journal of Proteomics* **198**, 18-26, doi:https://doi.org/10.1016/j.jprot.2018.12.004 (2019).

27 Haynes, W. in *Encyclopedia of Systems Biology* (eds Werner Dubitzky, Olaf Wolkenhauer, Kwang-Hyun Cho, & Hiroki Yokota) 78-78 (Springer New York, 2013).
28 Diz, A. P., Carvajal-Rodríguez, A. & Skibinski, D. O. Multiple hypothesis testing in proteomics: a strategy for experimental work. *Mol Cell Proteomics* **10**, M110.004374, doi:10.1074/mcp.M110.004374 (2011).
29 Schmidt, A., Forne, I. & Imhof, A. Bioinformatic analysis of proteomics data. *BMC Systems Biology* **8**, S3, doi:10.1186/1752-0509-8-S2-S3 (2014).
30 Meunier, B. *et al.* Assessment of hierarchical clustering methodologies for proteomic data mining. *J Proteome Res* **6**, 358-366, doi:10.1021/pr060343h (2007).

Chapter 2: Proteomic Study of Wood Frogs (*Rana sylvatica*)

This chapter has been published in a peer-reviewed journal, Scientific Reports.

Proteomic Analysis of *Rana sylvatica* Reveals Differentially Expressed Proteins in Liver in Response to Anoxia, Dehydration or Freezing Stress

Yingxi Li [1,2], Zoran Minic [2], Nico Hüttmann [2], Abdullah Khraibah [1], Kenneth B. Storey [3], Maxim V. Berezovski [1,2,*]

[1] Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, ON K1N 6N5, Canada

[2] John L. Holmes Mass Spectrometry Facility, Faculty of Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada

[3] Department of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada

* Correspondence: Dr. Maxim V. Berezovski (M.V.B.)

2.1 Abstract

Ectothermic animals that live in seasonally cold regions must adapt to seasonal variation and specific environmental conditions. During the winter, some amphibians hibernate on land and encounter limited environmental water, deficient oxygen, and extremely low temperatures that can cause the whole body freezing. These stresses trigger physiological and biochemical adaptations in amphibians that allow them to survive. *Rana sylvatica*, commonly known as the wood frog, shows excellent freeze tolerance. They can slow their metabolic activity to a near halt and endure freezing of 65-70% of their total body water as extracellular ice during hibernation, returning to normal when the temperatures rise again. To investigate the molecular adaptations of freeze-tolerant wood frogs, a comprehensive proteomic analysis was performed on frog liver tissue after anoxia, dehydration, or freezing exposures using a label-free LC-MS/MS proteomic approach. Quantitative proteomic analysis revealed that 87, 118, and 86 proteins were significantly upregulated in dehydrated, anoxic, and frozen groups, suggesting potential protective functions. The presence of three upregulated enzymes, glutathione S-transferase (GST), aldolase (ALDOA),

and sorbitol dehydrogenase (SORD), was also validated. For all enzymes, the specific enzymatic activity was significantly higher in the livers of frozen and anoxic groups than in the controls. This study reveals that GST, ALDOA, and SORD might participate in the freeze tolerance mechanism by contributing to regulating cellular detoxification and energy metabolism.

2.2 Introduction

Many animals show extraordinary resilience in the face of harsh environmental conditions, showcasing their remarkable ability to adapt and survive under diverse conditions [1]. Through the process of natural selection, species have developed an amazing array of adaptative strategies that enable them to thrive in challenging habitats, ranging from the Arctic tundra to scorching deserts. Under extreme cold conditions, these can include adaptations such as thick fur or insulating blubber to conserve body heat among mammals or, oppositely, the use of heat-dissipating mechanisms in order to regulate body temperatures within a narrow range [2-5]. Strong metabolic rate depression is also widespread to minimize rates of energy production and consumption in order to cope with limited food resources or extreme temperatures [6,7]. Behavioral adaptations, such as migration, hibernation, or burrowing, also aid various species in avoiding unfavorable conditions [8-10]. Furthermore, the ability to enter states of dormancy or suspended animation allows many species to survive through periods of environmental extremes (e.g. heat, cold, drought, etc.) [11,12].

The present study focuses on the wood frog, *Rana sylvatica* (also known as *Lithobates sylvaticus),* that stands out as one of Canada's most intriguing amphibians due to its amazing capacity for freeze tolerance. Wood frogs are widely distributed across diverse habitats, from the southern Appalachian Mountains to the boreal forests as far north as the Arctic Circle [13]. These adaptable frogs have mastered the art of survival under Canada's varying climates. One of the most astonishing aspects of their resilience is their remarkable ability to tolerate the freezing of up to 65-70% of their total body water as extracellular ice [14]. When they freeze, their cells face challenges of severe dehydration caused by water loss into extracellular ice masses and extended periods of oxygen deprivation due to freezing of blood plasma, that interrupts oxygen supply to organs [15]. Consequently, these frogs have developed tolerances for both cell/tissue dehydration and anoxia as adaptations to cope with harsh winter conditions. Most essential physiological

activities, such as respiration, heartbeat, muscle function, and brain activity, are strongly slowed or entirely halted during freezing [16].

In addition to physiological adaptations, wood frogs exhibit multiple biochemical adaptations for winter survival. They achieve their remarkable freeze tolerance via two main strategies: (1) the production of high concentrations of cryoprotectants and (2) the regulation of ice recrystallization [15]. Under the subzero temperatures of winter, extracellular water freezes causing a high proportion of intracellular water to be drawn out of cells to comply with osmosis. As a result, cells lose a high percentage of their water, but to prevent cells and organs from being damaged by dehydration, wood frogs and other freeze-tolerant species produce high concentrations of cryoprotectants that are packed into their cells (e.g., glucose, glycerol, sorbitol, or urea) that act as natural antifreeze. In addition, cells and tissues can be physically damaged by ice crystals penetrating cell membranes. To protect their intracellular environment, wood frogs can also produce ice-binding proteins to restrict the formation and maintenance of ice crystals to extracellular and extra-organ compartments only [15].

In addition to the production of high concentrations of cryoprotectants and ice-binding proteins, metabolic modifications also contribute to freeze tolerance. Wood frogs generally adapt their metabolism for fuel and energy preservation during the winter [17]. During this period, the overall metabolic activities of these ectotherms decrease strongly as temperature falls, leading to a substantial reduction in energy expenditure [18]. This slowdown in metabolism helps conserve fuel reserves, allowing frogs to survive the winter without food intake. In addition, some metabolic processes, such as detoxification, can be preserved as part of the stress response [19].

With their extraordinary freeze tolerance, wood frogs can survive being frozen solid for weeks or months. Then, with the arrival of spring, they thaw out and resume their normal activities with little or no tissue damage [16]. This unique survival strategy has captured the attention of many scientists and researchers, leading to extensive studies aimed at unraveling the biological mechanisms behind this incredible feat. However, unlike commonly used model organisms such as humans and rats with well-annotated genomes, non-model organisms, such as wood frogs, lack extensive genetic information, making traditional genetic studies challenging. Mass spectrometry-based proteomics has emerged as a powerful tool in studying non-model organisms, expanding our understanding of their biology and adaptation to diverse environments [20]. By analyzing the proteome, researchers can gain insights into protein expression patterns, post-translational

modifications, protein-protein interactions, and signaling pathways, providing valuable information about the organism's responses and adaptations to environmental stress.

In the present study, we undertook protein profiling of wood frog liver using MS-based proteomics approaches to understand further metabolic processes that contribute to freeze tolerance. Liver tissue was harvested from wood frogs after exposure to anoxia, dehydration, or freezing stress conditions. Tissue samples were analyzed by nano-liquid chromatography-tandem mass spectrometry (nLC-MS/MS), and protein profiles were compared with those of 5°C control frogs. The results revealed multiple differentially expressed proteins. Three significantly upregulated proteins were chosen for validation and deeper analysis.

2.3 Materials and Methods

2.3.1. Biological sample preparation

Adult wood frogs (~5-7 g) were collected from early spring breeding ponds in the Ottawa, Ontario region and transferred to the lab in coolers containing snow or crushed ice. In the lab, all frogs were given a bath in cold water containing tetracycline and were then housed in opaque plastic boxes (12 x 10 x 3 inches) lined with damp sphagnum moss and transferred into an incubator at 5°C for a two-week acclimation period. Frogs were checked daily. The control group was sampled randomly from this condition after the acclimation period [21-23].

Dehydration exposure was conducted as previously described [21]. Briefly, acclimated frogs were individually weighed, and then 5-6 frogs of diverse initial weights were placed in each of several tall, opaque, and dry plastic buckets in a 5°C cold incubator. Under these conditions, frogs slowly lost body water via evaporation across the skin. Frog weights were initially recorded twice a day and more frequently as the target water loss was approached. Calculation of the amount of body water lost was performed using the following equation:

$$\% \text{ Dehydration} = \frac{(M_i - M_d)}{(M_i - BWC_i)} \times 100\%$$

where M_i is the initial body mass of the frog, M_d is the mass after experimental dehydration, and BWC_i is the initial body water content of frogs before dehydration treatment. The BWC_i was previously determined to be 80.8 ± 1.2% g of H_2O per g body mass [21]. Water loss was discontinued when frogs reached approximately 40% dehydration.

Anoxia exposure was conducted following a previously published protocol [22]. Briefly, acclimated frogs were placed in cold containers (held in crushed ice) that had been pre-flushed

with nitrogen gas via ports in the lid for 30 minutes. Frogs were then quickly added (typically 5 per container), and then the container was flushed again with nitrogen gas for 30 minutes. Ports were sealed, and containers were transferred back to the 5°C incubator for 24 hours. After 24-hour exposure under a nitrogen gas atmosphere, containers were returned to a crushed ice container, the nitrogen gas line was reconnected, and frogs were rapidly sampled.

For freezing exposure, small groups of frogs were placed in plastic containers lined on the bottom with damp paper towels. Animals were held at -4°C in an incubator for 45 minutes to trigger ice nucleation in their bodies due to contact with ice on the paper towel. Temperature was then raised to -2.5 °C, and frogs were held for 24 hours at -2.5 °C followed by quick sampling [23].

Frogs in all groups were humanely euthanized using approved methods, followed by prompt dissection of the liver first (samples of other tissues then followed). Liver samples were immediately frozen in liquid nitrogen and then preserved at -80 °C until required for analysis. Groups of 18 frogs were prepared for each treatment group. Ethical procedures were strictly adhered to, with all animal protocols having received prior approval from the Carleton University Animal Care Committee (Protocol #13683), and following the guidelines set by the Canadian Council on Animal Care. The collection of wood frogs was authorized under Wildlife Scientific Collector's Authorization #1085726 granted by the Ministry of Natural Resources of Ontario [23].

2.3.2. Tissue homogenization

Frozen liver tissue samples were each suspended in 0.5 mL of ice-cold lysis buffer and homogenized for 3 minutes. Lysis buffer consisted of 1.5 M urea, 5% (v/v) glycerol, 1 mM DTT, 1:200 (v/v) protease inhibitor cocktail (Thermo Scientific, #78430), and 25 mM HEPES at pH 8.0. The suspension was centrifuged at 12,000 g for 10 min at 4°C, and soluble protein supernatants were separated from insoluble debris. Samples of 50 µL of the resulting supernatant were used for protein quantification, and 150 µL of the supernatant was used for proteomic sample preparation.

2.3.3 Protein Quantification by Bradford Assay

The protein concentration and quantity of the resulting supernatant were determined by the Bradford assay [24] using a protein assay kit (#23200, Thermo Fisher Scientific) following the manufacturer's protocol. The assay absorbance was measured with a spectrophotometer (Novaspec III, Biochrom) and semi-microvolume disposable polystyrene cuvettes (#2239955, Bio-Rad) at 595 nm.

2.3.4 Protein Reduction, Alkylation and Enzymatic Digestion

Aliquots of 50 µg of protein were further processed with a modified filter-aided sample preparation (FASP) protocol for proteomic analysis. The resulting supernatant samples containing 50 µg of protein were diluted to a total volume of 200 µL with a denaturation buffer (8 M urea, 25 mM HEPES at pH = 8.0). The samples were vortexed briefly and transferred to a 10 kDa MWCO filter (MRCPRT010, Millipore). The sample volume was reduced to about 20 µL by centrifugation for 15 minutes at 14,000 g, and proteins were reduced by the addition of 4 mM Tris(2-carboxyethyl) phosphine (TCEP) in 100 µL denaturation buffer. The samples were then incubated at 25°C for 30 minutes, followed by a 15-minute centrifugation at 14,000 g. Proteins were then alkylated with 20 mM of iodoacetamide (IAA) in 100 µL denaturation buffer. The samples were incubated at 25°C for 40 minutes, followed by a 15-minute centrifugation at 14,000 g. Next, 100 µL of digestion buffer (0.6% v/v glycerol, 25 mM HEPES, pH = 8.0) was added into the filter, followed by a 15-minute centrifugation at 14,000 g. After buffer exchange, the filter was transferred to a clean collection tube. Proteolytic digestion was performed by the addition of MS-grade trypsin/Lys-C mix (#V5072, Promega), 1:150 enzyme to protein ratio, and incubated in the dark with shaking at 600 rpm at 37 °C for 12 hours. Peptides were eluted as filtrate by centrifugation at 14,000 g for 15 minutes, and 2% (v/v) of formic acid was added to the collected filtrates to stop digestion. The collected filtrate containing about 50 µg of protein was desalted on disposable TopTip C-18 micro-spin columns (#TT2C18.96, Glygen, Ellicott City, MD, USA) and dried by vacuum centrifugation (Savant SPD111V SpeedVac Concentrator, Thermo Scientific).

2.3.5 Nano-LC-MS/MS

Protein samples were analyzed by an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific) coupled to an UltiMate 3000 nanoRSLC (Dionex, Thermo Fisher Scientific). One microliter of peptides (equivalent to 2.0 µg of protein) was loaded onto an in-house packed column (Polymicro Technology), 15 cm × 70 µm ID, Luna C18(2), 3 µm, 100 Å (Phenomenex) employing a water/acetonitrile/0.1% formic acid gradient. Peptides were separated for 105 min at a flow rate of 0.30 µL/min and the following steps: 0–7 min, 2–2% ACN; 7–77 min, 2–38% ACN; 77-82 min, 38–98% ACN; 82–92 min, 98–98% ACN; 92–95 min, 98–2% ACN; 95-105 min, 2–2% ACN. Eluted peptides were directly sprayed into a mass spectrometer using positive electrospray ionization (ESI) at an ion source temperature of 250 °C and an ion spray voltage of 2.1 kV. The Orbitrap Fusion mass spectrometer was run in data-dependent MS/MS acquisition following a full MS survey scan in top-speed mode. Full-scan MS spectra (m/z 350–2000) were acquired at a

resolution of 60,000. Precursor ions were filtered according to monoisotopic precursor selection, charge state (+ 2 to + 7), and dynamic exclusion (30 s with a ± 10 ppm window). The automatic gain control settings were 5×10^5 for full FTMS scans and 1×10^4 for MS/MS scans. Fragmentation was performed with collision-induced dissociation (CID) in the linear ion trap. Precursors were isolated using a 2 m/z isolation window and fragmented with a normalized collision energy of 35%.

2.3.6. MS Spectra Processing

Mass spectrometry raw data files of 72 samples were analyzed using MaxQuant, version 1.6.4.0 [25]. Peptides were searched against the *Lithobates* genus (NCBI taxonomy ID: 192752) FASTA file containing 30,829 reviewed and unreviewed entries (09.04.2021) and a default contaminants database using the Andromeda search engine, integrated into MaxQuant [26]. Default parameters were used unless stated otherwise. Methionine oxidation and N-terminal acetylation were set as variable modifications, while cysteine carbamidomethylation was set as a fixed modification. Trypsin and LysC proteases were chosen as the digestion enzyme-generating peptides of at least 7 amino acids with a maximum of 2 missed cleavages. A false discovery rate (FDR) was set to 0.01 for peptides and proteins, which were determined using a reverse sequence decoy database. A contaminant database provided by the Andromeda search engine was used. Peptides were identified with an initial precursor mass deviation and a fragment deviation of up to 10 ppm, and 0.5 Da, respectively. To increase the peptide identification rate, the 'Match between runs' algorithm in MaxQuant was performed between all samples [27]. Proteins and peptides matching to the reverse database or identified as contaminant were filtered out. A minimum ratio count of 2 was required for label-free quantification. Label-free quantification (LFQ) values were obtained through MaxQuant quantitative label-free analysis.

2.3.7. Data Filtering

The MaxQuant output tables *proteinGroups.txt* were loaded in R and analyzed using an in-house pOmics R package (github.com/nicohuttmann/pOmics). Potential contaminants and reverse proteins annotated by MaxQuant were excluded from the analysis. Principle component analysis (PCA) was generated using the prcomp R function based on the imputed label-free quantification (LFQ) intensity values. Proteins with detected LFQ intensities in 50% of the samples were used for imputation, and imputation of missing LFQ intensities was based on a down-shifted Gaussian distribution of log-transformed protein LFQ intensities [shift = 1.8 standard deviations (sd), width

= 0.3 sd], stimulating low abundance profiles [28]. The $\log_2$ (fold-change) values and p-values by unpaired t-test for proteins were computed using the imputed LFQ intensities. The GO enrichment analysis was made by enriching the detected proteins against the *Lithobates catesbeianus* (NCBI taxonomy ID: 8400) database.

2.3.8 Data Availability

All MS raw data were submitted to the PRIDE repository (Accession: PXD050540) at the European Bioinformatics Institute.

2.3.9 Glutathione S-transferase (GST) Activity Assay

Glutathione S-transferase activity was measured using a glutathione S-transferase activity assay kit (ab65326, Abcam, Toronto, ON, Canada) according to the manufacturer's procedures. Frozen liver tissue samples from four wood frog groups were resuspended in 0.5 mL of ice-cold 1x phosphate buffer saline and homogenized for 3 minutes. The suspension was centrifuged at 12,000 g for 10 min at 4°C, and supernatants were collected and used for measurement of enzymatic activities. Assays were performed at 22 °C, and absorbances were measured at 450 nm in a kinetic mode according to the manufacturer's instructions. Enzyme activities were calculated from the assay time of 40 minutes using the extinction coefficient factor provided by the standard operating procedure.

2.3.10 Aldolase (ALDOA) Activity Assay

Aldolase activity was measured using an aldolase activity assay kit (ab196994, Abcam, Toronto, ON, Canada) according to the manufacturer's procedures. Supernatant samples collected from liver tissue homogenates prepared initially for GST activity measurement were also used for the aldolase activity assay. Assays were performed at 25 °C, and absorbances were measured at 340 nm in a kinetic mode according to the manufacturer's instructions. The activity of enzymes was calculated from the assay time of 40 minutes via the calibration curve of the NADH standard at known concentrations (Figure S1A).

2.3.11 Sorbitol Dehydrogenase (SORD) Activity Assay

Sorbitol dehydrogenase activity was measured using a sorbitol dehydrogenase activity assay kit (ab252902, Abcam, Toronto, ON, Canada) according to the manufacturer's procedures. Supernatants collected from liver tissue homogenates for GST activity measurement were used for the sorbitol dehydrogenase activity assay. Assays were performed at 25 °C, and absorbances were measured at 450 nm in a kinetic mode according to the manufacturer's instructions. Enzyme

activities were calculated from the assay time of 40 minutes via the calibration curve of the NADH standard at known concentrations (Figure S1B).

2.4 Results

2.4.1 Liver Protein Profiling Upon External Stressors

Wood frog liver tissue samples containing approximately 50 μg of proteins were prepared from control, anoxic, dehydrated, and frozen groups. Eighteen biological replicates were obtained per group and examined independently. Proteins were digested using a FASP method that efficiently removes impurities from metabolites and consequently can increase the number of proteins identified [29]. A schematic overview of our experimental strategy is presented in Figure 2.1.

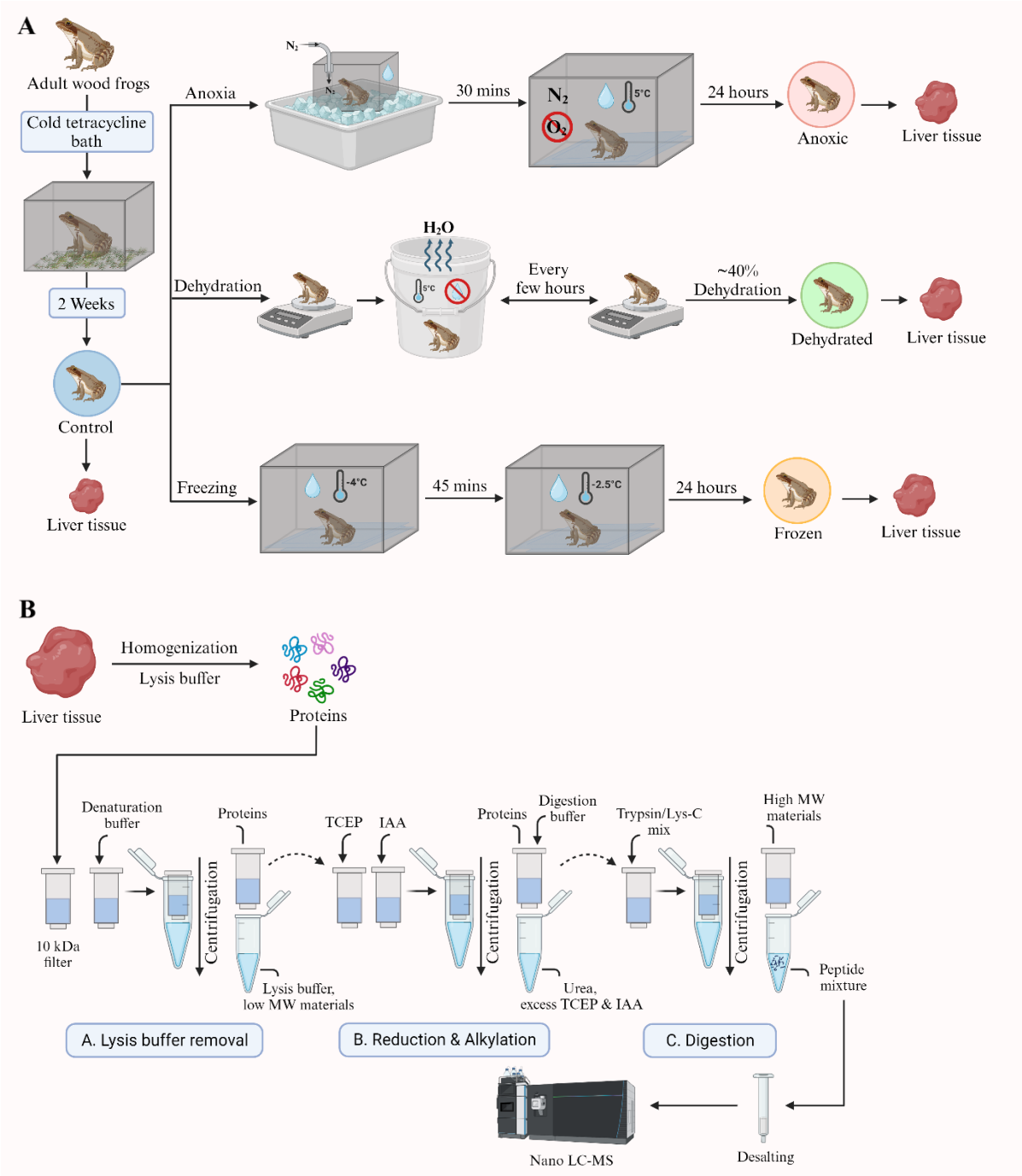

Figure 2.1. Workflow of proteomic sample preparation. (A) Frogs were treated under different conditions: control, anoxic, dehydrated, and frozen exposures. Liver tissues were sampled after the treatment and used for proteomic analysis. (B) Proteomic sample preparation was performed using the FASP method. Extracted proteins from frog liver tissues were reduced, alkylated, and digested into peptides and then analyzed with nLC-MS/MS.

Peptides obtained after digestion were separated and analyzed using nano-liquid chromatography-tandem mass spectrometry (nLC-MS/MS). Rigorous bioinformatics analysis and protein identification were applied to the proteomic mass spectrometry data. Proteins were identified and quantified using MaxQuant search engine using the *Lithobates* genus (NCBI taxonomy ID: 192752) database. Proteins were analyzed with strict criteria to ensure high confidence with a false discovery rate of < 1%. After the removal of reverse sequences and potential contaminants, a list of all identified proteins along with relevant parameters for protein identification is presented in Table S1. The total number of identified and quantified proteins for each group is also presented in Figure 2.2A and Table S2. In total, 1245, 1301, 1308, and 1253 proteins were identified, and 817, 883, 891, and 833 proteins have been quantified in control, anoxic, dehydrated, and frozen groups, respectively (Figure 2.2A). The Venn diagram of proteins revealed that 86% of the identified proteins were found in all four groups, and about 1% of proteins were found as unique proteins for each group (Figure 2.2B).

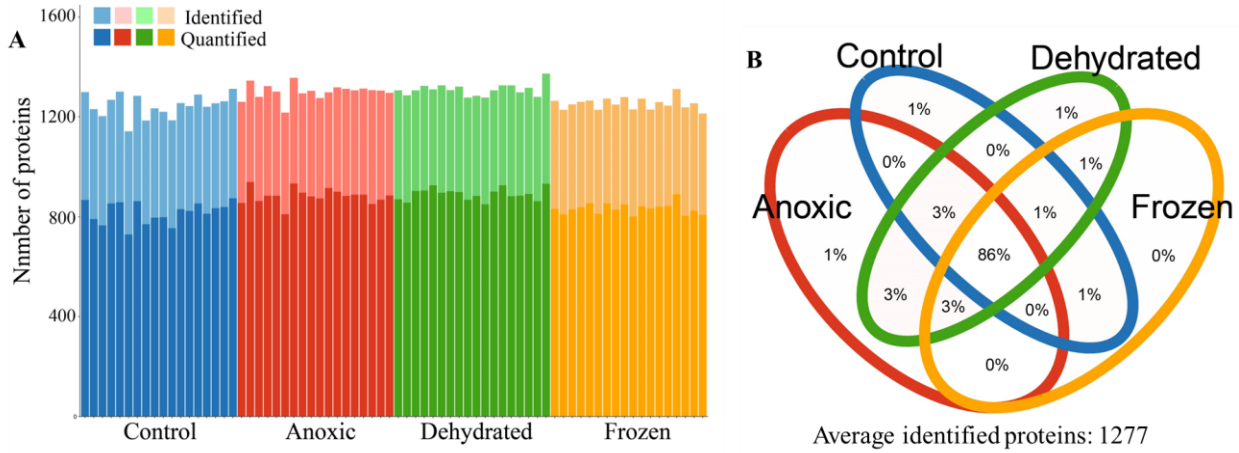

Figure 2.2. Overview of the identified and quantified proteins. (A) The total number of identified and quantified proteins from control, anoxic, dehydrated, and frozen groups. (B) A Venn diagram showing the number of overlapped proteins from the four groups.

2.4.2 Clustering Analysis of Proteins

A principal component analysis (PCA) and a hierarchical clustering analysis were performed to determine group similarity. The generation of principal component analysis (PCA) and hierarchical clustering analysis was based on imputed LFQ intensities for which proteins needed to be present in at least 50% of the samples in one group. The PCA and the hierarchical clustering

analysis showed that there were clear differences between the control group and the three other groups (Figure 2.3). Furthermore, the anoxic and dehydrated groups appear to have a high degree of similarity in their protein profile as their clusters were partially overlapped.

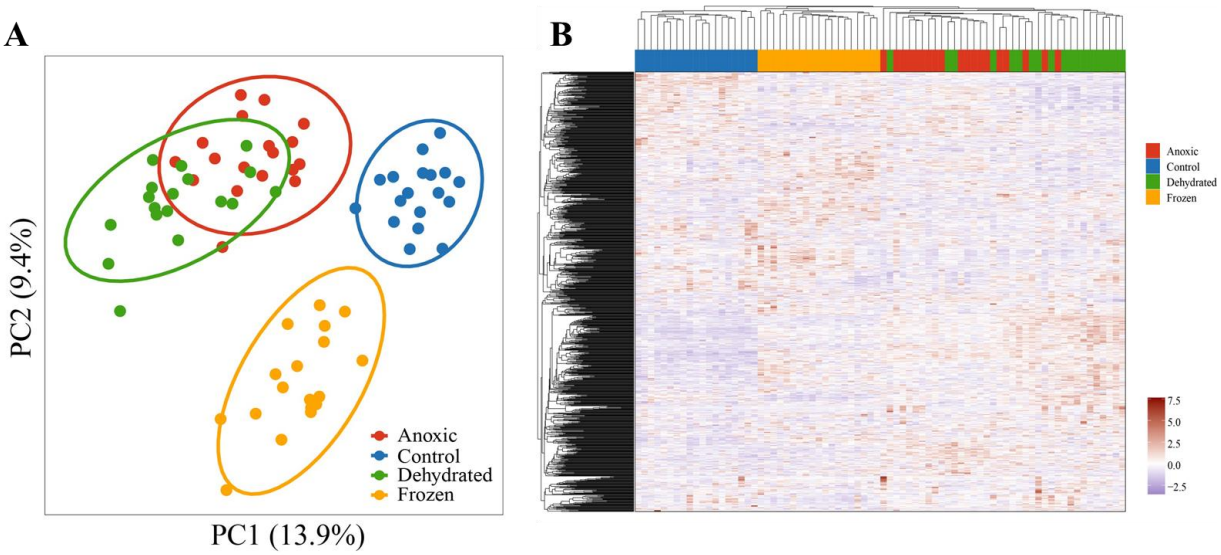

Figure 2.3. Clustering analysis of identified proteins. (A) Principal component analysis of samples from control, anoxic, dehydrated, and frozen groups. (B) Heatmap analysis of samples from the four groups.

2.4.3 Visualization of Protein Abundance

Volcano plots were produced to visualize the differentially expressed proteins in each treatment group (Figure 2.4). Fold changes (FC) calculated by LFQ protein intensity and t-tests for p-values were implemented. $FC > 1.5$ and p-values < 0.05 were considered thresholds to identify differentially expressed proteins (DEPs). Overall, 87, 118, and 86 proteins were identified as significantly upregulated in anoxic, dehydrated, and frozen groups, respectively, as compared to the control group. The complete list of upregulated, downregulated, and non-significantly differentially expressed proteins is presented in Table S3. The complete list of upregulated and downregulated proteins is presented in Supplementary Table S2. In this study, we focus on upregulated proteins to identify molecular markers that are associated with particular stress conditions.

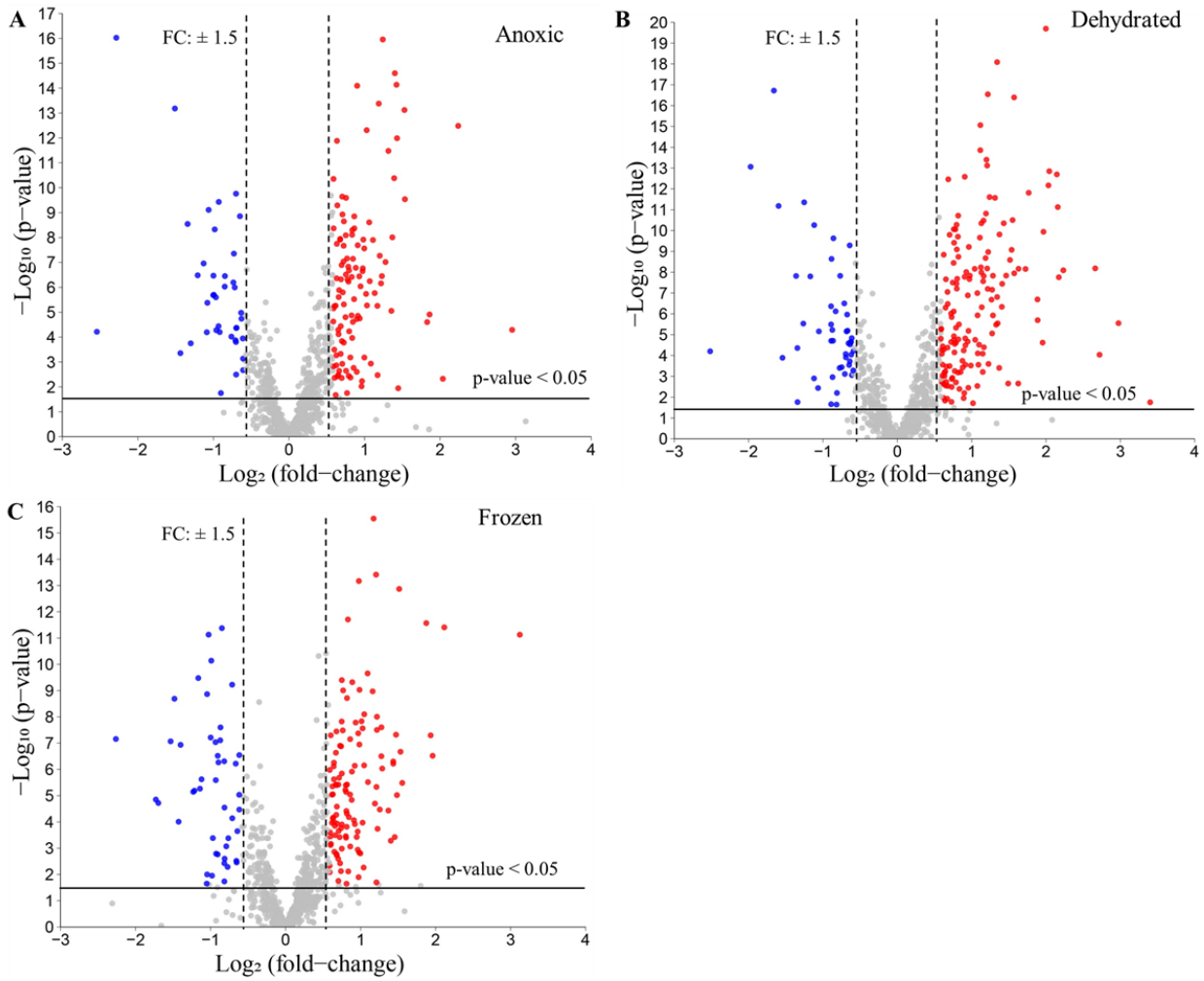

Figure 2.4. Volcano plots of quantified proteins from liver. Protein quantities between the three conditions and the control group were compared. Red dots indicate proteins that were upregulated under the stress condition; blue dots indicate proteins that were downregulated under the stress condition; grey dots show proteins that did not change significantly between control and stress conditions. The analysis threshold was p-value < 0.05 and the fold change (FC) > ±1.5; anything below these thresholds was considered not significant. (A) Volcano plot of anoxic vs. control; (B) Volcano plot of dehydrated vs. control; (C) Volcano plot of frozen vs. control.

Figure S2 shows Venn diagram analysis of up- and down-regulated proteins extracted from the results of Volcano plot (Figure 2.4, Table S3). For the up-regulated proteins, 18, 34 and 47 proteins were unique for anoxic, dehydrated and frozen conditions, respectively. An important number of up-regulated proteins overlap between these three groups (69 proteins). For the down-regulated proteins, 17, 29 and 38 proteins were unique for anoxic, dehydrated and frozen conditions,

respectively. There was not a large overlap between these three conditions, but a significant amount of overlap is observed between dehydrated and anoxic groups.

2.4.4 Functional Analysis of Significantly Upregulated Proteins

Gene Ontology (GO) analysis was used to annotate the significantly upregulated proteins in anoxic, dehydrated, and frozen groups as compared to the control group (Figure 2.5). The GO enrichment analysis of up-regulated proteins used cellular components, biological processes, and molecular function terms. Within the GO cellular component terms, the upregulated proteins were mainly localized to the cytoplasm, nucleus, ribosomes, and integral component of membrane categories. The GO biological process was enriched mainly in the term Translation, but several other terms related to transport were also enriched, including intracellular protein transport, vesicle-mediated transport, and protein transport. Interestingly, the dehydrated group had more enriched proteins related to protein transport as compared to the other treatment groups. Within the GO molecular function terms, the upregulated proteins were mainly annotated as structural constituents of ribosome, ATP binding, RNA binding, and ATP hydrolysis activity. The results for these three stress conditions show an overlap between the GO of cellular components, biological processes, and molecular function terms. However, the number of proteins for each condition and term differed. Enrichment analysis of GO terms for diverse stress and response conditions revealed an unexpectedly low number of proteins (Table 2.1).

Upregulated Enzymes	**GO for stress-responsive terms**	**Conditions**
Eosinophil peroxidase	Defense response [GO:0006952]	Anoxic, Dehydrated
Epoxide hydrolase	Response to toxic substance [GO:0009636]	Dehydrated, Frozen
Aconitate hydratase	Response to iron (II) ion [GO:0010040]	Frozen

Table 2.1. GO enrichment analysis of upregulated proteins in anoxic, dehydrated, and frozen groups in response to diverse stress and response conditions.

Gene Ontology (GO) enrichment analysis of downregulated proteins in anoxic, dehydrated, and frozen groups as compared to the control group is presented in Figure S3. Although this analysis was performed on a small number of proteins (39, 49 and 46 proteins in the anoxic, dehydrated, and frozen groups, respectively), it appears that there are substantial differences in the terms for GO enrichments in comparison to upregulated protein. It is interesting to note that

tricarboxylic acid cycle for all three conditions were affected. Therefore, two proteins of tricarboxylic acid cycle were found to be downregulated in frozen, dehydrated and anoxic conditions. These proteins include malate dehydrogenase and oxoglutarate dehydrogenase.

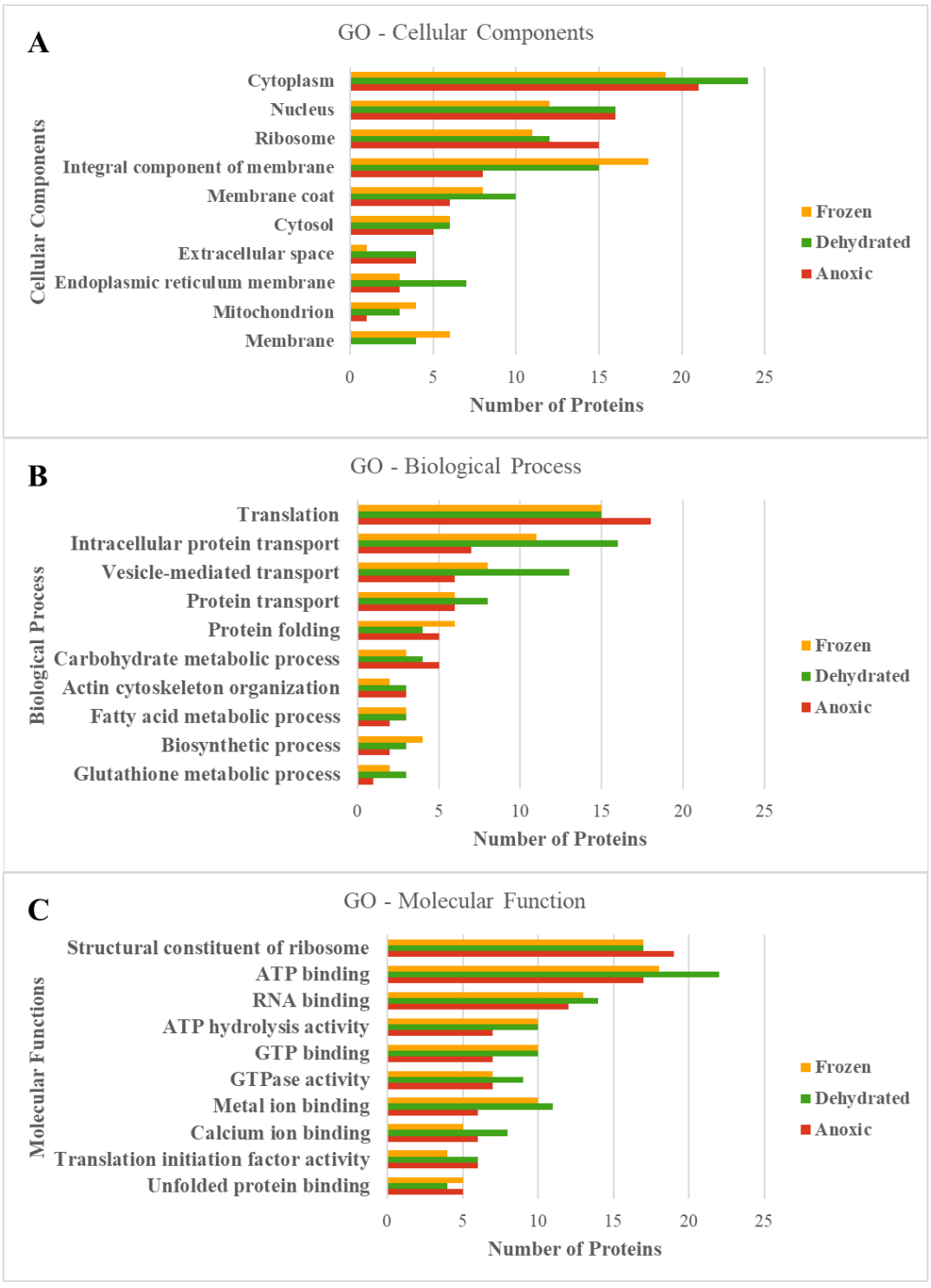

Figure 2.5. Top 10 Gene Ontology (GO) functional annotation labels for upregulated proteins in anoxic, dehydrated, and frozen groups. (A) Cellular Components, (B) Biological processes, and (C) Molecular functions.

2.4.5 Significant Differentially Expressed Enzymes

Since metabolic changes play a pivotal role in the adaptation of wood frogs to extreme environments [30] and this change can be easily monitored by enzymatic assay to assess these metabolic alterations, the significantly upregulated protein list was further narrowed by selecting only enzymes. In the upregulated protein list, 16, 29, and 19 proteins are enzymes in the anoxic, dehydrated, and frozen groups, respectively. The list of significantly upregulated enzymes can be found in Supplementary Table S4. Some of these enzymes have a molecular function that can be linked to the wood frog's freeze tolerance mechanism (Table 2.2). In addition to the enzymes shown in Table 2.2, proteomic analysis identified aldolase as a unique protein present only in the anoxic group.

Upregulated Enzymes	**Molecular Function**	**Fold Change**		
		Anoxic	**Dehydrated**	**Frozen**
Glutathione S-transferase (GST)	Detoxification of ROS	3.6	4.5	8.7
Sorbitol dehydrogenase (SORD)	Carbohydrate metabolism	1.8	1.7	-
Epoxide hydrolase (EH)	Metabolism of xenobiotics	-	2.1	3.8
Sulfotransferases (SULT)	Detoxification	3.6	3.9	-
Lactate dehydrogenase (LDH)	Anaerobic metabolism	2.3	2.5	-
Heme oxygenase (HO)	Oxidative stress response	-	-	1.9

Table 2.2. Significantly upregulated enzymes in anoxic, dehydrated, and frozen groups compared to the control group, and their molecular functions. "-" represents that enzymes were not significantly upregulated in such groups.

2.4.6 Validation of Specific Activity of GST, ALDOA, and SORD

Three significantly upregulated enzymes identified by the proteomics approach were selected for further analysis: glutathione S-transferase (GST), aldolase (ALDOA), and sorbitol dehydrogenase (SORD). Enzymatic kits for other enzymes presented in Table 2.1 were not commercially available and consequently further analyses of these enzymes were not conducted. Enzymatic activities were tested in frog liver tissues for each enzyme using five biological replicates. The specific activities were measured in nmol per minute per milligram of proteins, and the mean specific activities of biological replicates are presented in Figure 2.6. The difference in specific activity was considered statistically significant if the p-value < 0.05. The specific activity

levels of all tested enzymes were significantly higher in the frozen and anoxic groups using t-test analysis as compared with the control group. Similarly, the specific activity of GST was significantly higher in the dehydrated group as compared with the control group, but ALDOA and SORD did not show significantly different specific activities among the two groups. To provide more information regarding multiple comparisons between groups, the results of one-way analysis of variance (ANOVA) are performed (Figure S4). The results confirmed significantly different specific activities among the four groups.

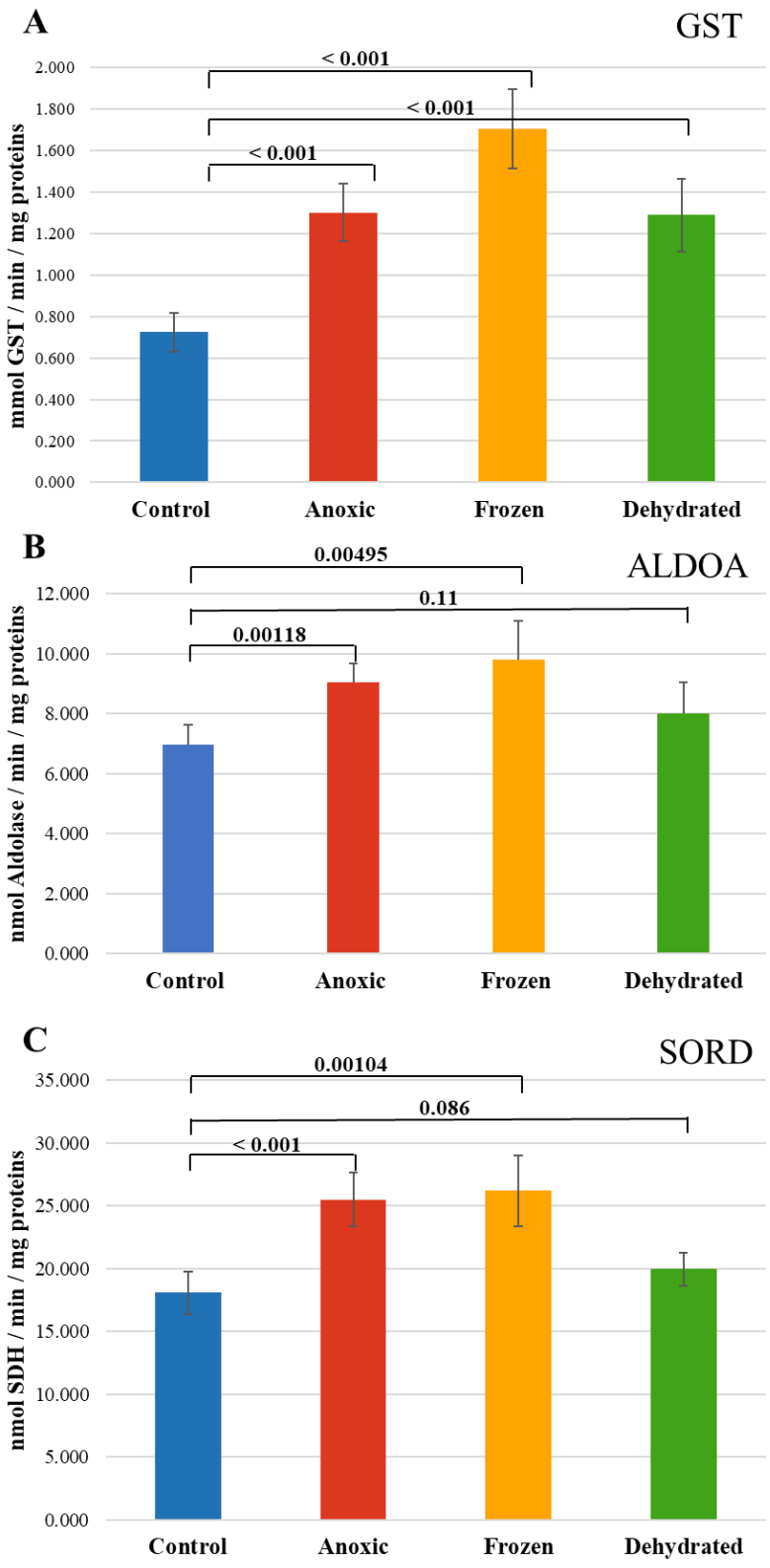

Figure 2.6. Specific enzymatic activity of (A) Glutathione S-transferase (GST), (B) Aldolase (ALDOA), and (C) Sorbitol dehydrogenase (SORD) from frog liver. The bar graph represents the

mean value of the biological replicates, whereas error bars indicate the standard deviation of the replicates, and p-values obtained from a Student's t-test are indicated on the top. Data were considered statistically significant if $p < 0.05$.

2.5 Discussion

The field of freeze tolerance among animals has been explored since the 1960s [31,32]. In its early stages, the field was limited in scope, primarily concentrating on essential data collection about freeze-tolerant organisms, including information on species variety, freeze survival parameters, ice nucleators, and identification of cryoprotectants [33]. The field has now grown significantly, and studies from different aspects have been conducted among many animal groups [15]. The wood frog, *Rana sylvatica,* is the best-researched freeze-tolerant vertebrate and serves as the primary model for investigating the molecular, biochemical, and physiological mechanisms underlying freezing survival [16]. MS-based proteomics is considered a powerful tool for investigating the molecular adaptations of organisms for survival under environmental stresses of many kinds. However, there have been limited proteomic studies conducted on wood frogs. Kiss *et al.* performed a proteomic analysis on wood frogs collected during both summer and winter, identifying 33 proteins with significant seasonal variations. Their results revealed that winter frogs exhibited elevated expression of proteins associated with cryoprotection, whereas proteins involved in cell proliferation, protein synthesis, and mitochondrial function showed reduced abundance [34]. Hawkins *et al.* employed phosphoproteome analysis to investigate the phosphorylation of metabolic enzymes in wood frogs. Their study focused on enzymes related to glucose and urea production, specifically in response to freezing, anoxia, or dehydration exposures. The study revealed distinct stress-specific variations in phosphopeptide abundance of 9 glycolytic enzymes and 3 urea cycle enzymes in the liver of wood frogs [23].

A shotgun label-free quantitative proteomic approach was performed in our study to understand the molecular basis of freeze tolerance in wood frogs. The results of this analysis showed unique proteome patterns and many differentially expressed proteins under different stress conditions, including anoxia, dehydration, and frozen groups. It is important to note that alteration in protein abundance is not a necessary consequence in gene expression and can be regulated at multiple levels including post-transcriptional, translational and protein degradation regulation. As ectotherms, the metabolic activities of wood frogs dramatically slow in the winter [30], although a

few activities rise as part of the cold stress response. Previous studies showed that winter wood frogs (when triggered by ice formation on their skin) immediately activate liver glycogenolysis to produce copious amounts of glucose that is then exported as a cryoprotectant and rapidly taken up by all tissues [35,36]. These high sugar concentrations mitigate cell/tissue damage from excessive water loss into extracellular ice masses. Antioxidant defenses and pro-survival pathways are also enhanced [37,38]. Since ATP provides the energy needed for many essential processes in cells and organisms [30,39,40], it is not surprising that the GO molecular function term "ATP hydrolysis activity" was significantly enriched among upregulated proteins. Interestingly, the dehydration group was mostly enriched in the GO biological process terms "intracellular protein transport", "vesicle-mediated transport", and "protein transport". Previous studies have demonstrated that the production of cryoprotectants such as glucose and urea was also increased as part of the response to dehydration stress in wood frogs [23]. Our finding suggests that proteins related to the transport biological process were upregulated during dehydration to aid cryoprotectant transport and distribution from liver to other organs in order to provide osmotic resistance against cell shrinkage below a critical minimum volume.

In order to have a better understanding of the metabolic changes supporting freeze tolerance, we focused on the significantly upregulated enzymes in treatment groups. Some of these enzymes were closely related to the freeze tolerance mechanism. For example, the sulfotransferases (SULT) are a group of transferase enzymes responsible for transferring sulfate groups from a common donor molecule 3'-phosphoadenosine 5'-phosphosulfate (PAPS) to an acceptor group of numerous substrates [41]. The anoxic condition promotes an increase in reactive oxygen species (ROS), leading to oxidative damage [42]. Dehydration can also cause oxidative stress either by elevating the production of ROS or by deactivating antioxidant enzymes [43]. Hossain *et al.* observed that the depletion of SULT results in the accumulation of ROS [44]. In our work, we found upregulation of the SULT enzyme in anoxic and dehydrated groups, suggesting that SULTs may be involved in the oxidative stress response.

Epoxide hydrolases (EH) catalyze the hydration of epoxides to trans-dihydrodiols and are very important enzymes in toxification–detoxification processes [45]. EH is a cytosolic enzyme that is highly expressed in liver [46]. It is known that glycogen synthase kinase-3beta (GSK3β) shows a strong increase in protein levels in liver of glucose-injected wood frogs [47]. Li *et al.* proved that soluble EH has a protective effect on oxidative damage by reducing ROS levels in rats and

activating the PI3K/Akt/GSK3β signaling pathway [48,49]. In turn, activation of the GSK3β signaling pathway can lead to the overexpression of EH as a component of the stress response. Based on these potential roles of EH, it is not surprising that significant upregulation of EH in both dehydrated and frozen groups was identified in this study.

Other than SULT and EH, heme oxygenase (HO) was also identified as being involved in the cold stress response. Katori *et al.* showed that overexpression of HO-1 protected rat livers against cold ischemia-reperfusion injury (IRI) [50]. In addition, Venkatachalam *et al.* found that HO-1 can protect liver cells from cold ischemia in both *in vitro* and *ex vivo* models [51]. Ischemia occurs when blood supply falls below the body's normal demand, causing a shortage of oxygen. When the affected tissue is reperfused, a rapid and excessive accumulation of ROS occurs, potentially leading to IRI tissue damage [52]. Wood frogs remain frozen for weeks at a time and can also undergo multiple freeze-thaw cycles showing little or no tissue damage after thawing [16]. Our findings show HO upregulation in the frozen group, which suggests that HO potentially plays a key role in protecting tissues from damage under cold stress.

GO enrichment analysis of upregulated proteins in anoxic, dehydrated, and frozen groups using diverse stress and response terms revealed only three proteins: eosinophil peroxidase, epoxide hydrolase, and aconitate hydratase. The frog's liver appears to activate different adaptive molecular mechanisms under these three conditions but not many proteins are involved in the response to stress. In general, stress conditions can damage the structure and function of macromolecules and prolonged stress may cause death [53]. Therefore, *R. sylvatica* has developed a molecular adaptive mechanism to prevent potential damage to liver cells.

Antibodies specific to frog proteins are not commonly available commercially, so we chose an alternate method to validate the proteomic results. We chose enzymatic assays that are generally very sensitive and selective, for our validation method. Enzyme assays were used to assess the activities of three targets related to the freeze tolerance mechanism: glutathione S-transferases (GSTs), sorbitol dehydrogenase (SORD), and aldolase (ALDOA).

GSTs are a class of enzymes that catalyze the conjugation of glutathione, a major cellular antioxidant, to a wide variety of electrophilic compounds [54]. Under anoxic conditions, cells experience significant stress due to the lack of oxygen, which can lead to the accumulation of ROS and other toxic compounds [55]. Research has shown that GSTs are upregulated in response to cold stress in various organisms, including fungi [56], plants [57,58], and reptiles [7]. Interestingly, Willmore

et al. showed that anoxia exposure caused significant reduction in the specific activity of liver GST in turtles [59,60]. Our findings from proteomic analysis showed significant upregulation of GSTs in anoxic, dehydrated, and frozen frogs. GST enzymatic activity was significantly higher in all three treatment groups as well. In the frozen condition, proteomics analysis revealed that GST was found to be 8.7-fold more expressed when compared to the control group. However, specific enzymatic activity increased only by about 1.8-fold. This discrepancy could be explained by the fact that enzymatic activity depends on the post-translational modifications as well as protein-protein interactions [61]. Notably, freeze-exposed wood frogs showed the highest protein expression and specific enzymatic activity of GST. Our results, along with previous studies, suggest that GSTs play a role in cellular detoxification by aiding the clearance of stress-induced ROS and consequently contributing to prevention of cellular damage.

Sorbitol dehydrogenase (SORD) is an enzyme involved in carbohydrate metabolism, transforming sorbitol (a sugar alcohol derived from glucose) into fructose, and providing an alternate source of energy [62]. Our proteomic results showed significant upregulation of SORD in anoxic and dehydrated groups, but not in the frozen group. However, SORD specific activity was significantly higher only in anoxic and frozen groups. This discrepancy between proteomics analysis and enzymatic activity assay could be due to the regulation of enzymatic activity by various post-translation modifications of SORD. The results of this study might suggest that this enzyme, as part of energy metabolism, can have a role in the adaptation of frogs to anoxic and dehydrated groups.

Aldolase (ALDOA) is a glycolytic enzyme that splits fructose 1,6-bisphosphate into two triose phosphate moieties that are further processed to pyruvate and then to either an anaerobic end point (e.g. lactate) or converted to acetyl-CoA that enters the Krebs cycle for aerobic catabolism. Both processes lead to ATP production but the aerobic route produces much more ATP [63]. Our findings from proteomic analysis showed that fructose-bisphosphate aldolase B was a unique protein in the anoxic group. It is reported that aldolase B is predominantly located in the liver and plays a role in the glycolysis pathway [64]. Our results using an enzymatic assay showed that ALDOA was detected in all groups, but its specific activity was significantly higher in anoxic and frozen groups. This result is also in agreement with previous studies [65,66]. Michaelidis *et al.* showed that aldolase activity in frogs initially increased within the first month of hibernation [65], and Niu *et al.* revealed that the mRNA level of fructose-bisphosphate aldolase was significantly increased in hibernating

frogs, *Nanorana parkeri* [66]. When considering the fact that regular glycolysis is inhibited in wood frogs during freezing [67], these findings indicate that anaerobic glycolysis continues to function during hibernation. It appears that it is capable of supplying ATP for energy-requiring activities, but its efficiency is considerably lower compared to aerobic metabolic pathways. Therefore, significantly higher specific activity of aldolase in anoxic and frozen conditions could have roles in maintaining carbohydrate utilization during hibernation and retaining the functional ability to respond to any environmental changes.

Identified down-regulated proteins are related to different biological functions, but it was found that two proteins, malate dehydrogenase and oxoglutarate dehydrogenase, of the tricarboxylic acid cycle were affected for all three conditions. These results are in agreement with a previously reported investigation whereby pyruvate dehydrogenase from *Rana sylvatica* was found to be inhibited during 24 h freezing and 24 h anoxia. It has been suggested that the inhibition of this enzyme could reduce glycolytic flux and carbon entry into the tricarboxylic acid cycle as part of metabolic rate depression [68].

Most enzyme-catalyzed reactions are temperature dependent, and most enzymes show reduced activity or are even inactivated at low temperatures [69,70] The enzymatic assay results revealed that the frozen group displayed the relatively highest specific activity among the three enzymes tested: GST, ALDOA, and SORD. This increased enzymatic activity in the frozen group suggests the potential involvement of these enzymes in maintaining basic life functions of wood frogs under stress. In addition, these enzymes likely play a role in safeguarding tissues against damage under frozen conditions by maintaining the functionality of the most important metabolic pathways. The main limitation of this study is the lack of whole-genome sequencing of *Rana sylvatica*. The identification of proteins was aided by the genus rank taxonomic representation (*Lithobates*) in the Uniprot database instead of the species rank. This may cause potential misidentification of some proteins. Similar situations were applied to the gene ontology enrichment analysis, where the analysis was made using the database of a different species, American bullfrogs (*Rana catesbeianus*), that has the highest identified proteins under all conditions. Once the genome sequence of *Rana sylvatica* becomes available, it will be possible to obtain a more detailed image of proteome profiles using our reported data.

2.6 Conclusion

The aim of this study was to analyze liver tissue from wood frogs under normal conditions and after anoxia, dehydration, or freezing exposures. The proteomic analysis confirmed differences in protein expression between controls and the three experimental groups. These results indicated that frogs developed many aspects of the molecular and cellular responses to these three stress conditions. Among differentially regulated proteins, three enzymes: GST, ALDOA, and SORD, were validated. The enzymatic assay results revealed that these enzymes showed a significantly higher specific enzymatic activity under anoxia and frozen treatments than the control group. Thus, our results strongly suggest that these enzymes might have important roles in maintaining normal metabolic function under extreme stress conditions. Overall, our data can provide a rich source of protein markers that may be used for further exploration of animal freeze tolerance.

2.7 Acknowledgements

The research was supported by the Natural Sciences and Engineering Research Council of Canada (grant #RGPIN-2020-04733 to KBS and #RGPIN-2020-05775 to MVB).

2.8 References – Chapter 2

1 Malhi, Y. *et al.* The role of large wild animals in climate change mitigation and adaptation. *Current Biology* **32**, R181-R196, doi:https://doi.org/10.1016/j.cub.2022.01.041 (2022).

2 Gaughan, J. B., Sejian, V., Mader, T. L. & Dunshea, F. R. Adaptation strategies: ruminants. *Animal Frontiers* **9**, 47-53, doi:10.1093/af/vfy029 (2019).

3 Iverson, S. J. in *Encyclopedia of Marine Mammals (Second Edition)* (eds William F. Perrin, Bernd Würsig, & J. G. M. Thewissen) 115-120 (Academic Press, 2009).

4 Besch, E. L. & Woods, J. E. Heat dissipation biorhythms of laboratory animals. *Lab Anim Sci* **27**, 54-59 (1977).

5 Vejmělka, F. *et al.* Heat dissipation in subterranean rodents: the role of body region and social organisation. *Scientific Reports* **11**, 2029, doi:10.1038/s41598-021-81404-3 (2021).

6 Storey, K. B., Storey, J. M., Brooks, S. P., Churchill, T. A. & Brooks, R. J. Hatchling turtles survive freezing during winter hibernation. *Proc Natl Acad Sci USA* **85**, 8350-8354, doi:10.1073/pnas.85.21.8350 (1988).

7 Hermes-Lima, M. & Storey, K. B. Antioxidant defenses in the tolerance of freezing and anoxia by garter snakes. *Am J Physiol* **265**, R646-652, doi:10.1152/ajpregu.1993.265.3.R646 (1993).

8 Black, R., Bennett, S. R. G., Thomas, S. M. & Beddington, J. R. Migration as adaptation. *Nature* **478**, 447-449, doi:10.1038/478477a (2011).

9 Constant, T. *et al.* Integrating mortality risk and the adaptiveness of hibernation. *Front Physiol* **11**, 706, doi:10.3389/fphys.2020.00706 (2020).

10 Li, M. *et al.* The hypoxia adaptation of small mammals to plateau and underground burrow conditions. *Animal Model Exp Med* **4**, 319-328, doi:10.1002/ame2.12183 (2021).
11 Padilla, P. A. & Ladage, M. L. Suspended animation, diapause and quiescence: arresting the cell cycle in *C. elegans*. *Cell Cycle* **11**, 1672-1679, doi:10.4161/cc.19444 (2012).
12 Ragland, G., Denlinger, D. & Hahn, D. Mechanisms of suspended animation are revealed by transcript profiling in the flesh fly. *Proc Nat Acad Sci USA* **107**, 14909-14914, doi:10.1073/pnas.1007075107 (2010).
13 Larson, D. J. *et al.* Wood frog adaptations to overwintering in Alaska: new limits to freezing tolerance. *J Exp Biol* **217**, 2193-2200, doi:10.1242/jeb.101931 (2014).
14 Storey, K. B. Life in a frozen state: adaptive strategies for natural freeze tolerance in amphibians and reptiles. *Am J Physiol* **258**, R559-568, doi:10.1152/ajpregu.1990.258.3.R559 (1990).
15 Storey, K. B. & Storey, J. M. Molecular physiology of freeze tolerance in vertebrates. *Physiological Reviews* **97**, 623-665, doi:10.1152/physrev.00016.2016 (2017).
16 Storey, J. M., Wu, S. & Storey, K. B. Mitochondria and the frozen frog. *Antioxidants (Basel)* **10**, doi:10.3390/antiox10040543 (2021).
17 Storey, K. B. & Storey, J. M. Mitochondria, metabolic control and microRNA: Advances in understanding amphibian freeze tolerance. *Biofactors* **46**, 220-228, doi:10.1002/biof.1511 (2020).
18 Tsiouris, J. A. Metabolic depression in hibernation and major depression: an explanatory theory and an animal model of depression. *Med Hypotheses* **65**, 829-840, doi:10.1016/j.mehy.2005.05.044 (2005).
19 Cowan, K. J. & Storey, K. B. Freeze-thaw effects on metabolic enzymes in wood frog organs. *Cryobiology* **43**, 32-45, doi:10.1006/cryo.2001.2338 (2001).
20 Russeth, K. P., Higgins, L. & Andrews, M. T. Identification of proteins from non-model organisms using mass spectrometry: application to a hibernating mammal. *J Proteome Res* **5**, 829-839, doi:10.1021/pr050306a (2006).
21 Churchill, T. & Storey, K. Metabolic responses to dehydration by liver of the wood frog, *Rana sylvatica*. *Canadian Journal of Zoology* **72**, 1420-1425, doi:10.1139/z94-188 (2011).
22 Al-Attar, R. & Storey, K. B. Effects of anoxic exposure on the nuclear factor of activated T cell (NFAT) transcription factors in the stress-tolerant wood frog. *Cell Biochem Funct* **36**, 420-430, doi:10.1002/cbf.3362 (2018).
23 Hawkins, L. J. *et al.* Glucose and urea metabolic enzymes are differentially phosphorylated during freezing, anoxia, and dehydration exposures in a freeze tolerant frog. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* **30**, 1-13, doi:https://doi.org/10.1016/j.cbd.2019.01.009 (2019).
24 Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254, doi:https://doi.org/10.1016/0003-2697(76)90527-3 (1976).
25 Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* **26**, 1367-1372, doi:10.1038/nbt.1511 (2008).
26 Inui, H. *et al.* Cucurbita pepo contains characteristic proteins without a signal peptide in the xylem sap. *Journal of Plant Physiology* **287**, 154038, doi:https://doi.org/10.1016/j.jplph.2023.154038 (2023).

27 Nagaraj, N. *et al.* System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap*. *Molecular & Cellular Proteomics* **11**, M111.013722, doi:https://doi.org/10.1074/mcp.M111.013722 (2012).
28 Lazar, C., Gatto, L., Ferro, M., Bruley, C. & Burger, T. Accounting for the multiple natures of issing values in label-free quantitative proteomics data sets to compare imputation strategies. *Journal of Proteome Research* **15**, 1116-1125, doi:10.1021/acs.jproteome.5b00981 (2016).
29 Wiśniewski, J. R. in *Microbial Proteomics: Methods and Protocols* (ed Dörte Becher) 3-10 (Springer New York, 2018).
30 Muir, T. J., Costanzo, J. P. & Lee Jr., R. E. Metabolic depression induced by urea in organs of the wood frog, *Rana sylvatica*: effects of season and temperature. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* **309A**, 111-116, doi:https://doi.org/10.1002/jez.436 (2008).
31 Miller, L. K. Freezing tolerance in an adult insect. *Science* **166**, 105-106, doi:10.1126/science.166.3901.105 (1969).
32 Somero, G. N. & DeVries, A. L. Temperature tolerance of some Antarctic fishes. *Science* **156**, 257-258, doi:10.1126/science.156.3772.257 (1967).
33 Storey, K. B. & Storey, J. M. Freeze tolerance in animals. *Physiol Rev* **68**, 27-84, doi:10.1152/physrev.1988.68.1.27 (1988).
34 Kiss, A. J., Muir, T. J., Lee, R. E., Jr. & Costanzo, J. P. Seasonal variation in the hepatoproteome of the dehydration and freeze-tolerant wood frog, *Rana sylvatica*. *Int J Mol Sci* **12**, 8406-8414, doi:10.3390/ijms12128406 (2011).
35 Costanzo, J. P. & Jr., R. E. L. Cryoprotectant production capacity of the freeze-tolerant wood frog, *Rana sylvatica*. *Canadian Journal of Zoology* **71**, 71-75, doi:10.1139/z93-011 (1993).
36 do Amaral, M. C., Lee, R. E., Jr. & Costanzo, J. P. Enzymatic regulation of seasonal glycogen cycling in the freeze-tolerant wood frog, *Rana sylvatica*. *J Comp Physiol B* **186**, 1045-1058, doi:10.1007/s00360-016-1012-2 (2016).
37 Gupta, A. & Storey, K. B. Regulation of antioxidant systems in response to anoxia and reoxygenation in *Rana sylvatica*. *Comp Biochem Physiol B Biochem Mol Biol* **243-244**, 110436, doi:10.1016/j.cbpb.2020.110436 (2020).
38 Zhang, J., Gupta, A. & Storey, K. B. Freezing stress adaptations: Critical elements to activate Nrf2 related antioxidant defense in liver and skeletal muscle of the freeze tolerant wood frogs. *Comp Biochem Physiol B Biochem Mol Biol* **254**, 110573, doi:10.1016/j.cbpb.2021.110573 (2021).
39 Dunn, J. & Grider, M. H. in *StatPearls*, StatPearls Publishing, Copyright © 2023, StatPearls Publishing LLC., 2023).
40 Shekhovtsov, S. V. *et al.* Metabolic response of the Siberian wood frog *Rana amurensis* to extreme hypoxia. *Sci Rep* **10**, 14604, doi:10.1038/s41598-020-71616-4 (2020).
41 Negishi, M. *et al.* Structure and function of sulfotransferases. *Arch Biochem Biophys* **390**, 149-157, doi:10.1006/abbi.2001.2368 (2001).
42 McGarry, T., Biniecka, M., Veale, D. J. & Fearon, U. Hypoxia, oxidative stress and inflammation. *Free Radical Biology and Medicine* **125**, 15-24, doi:https://doi.org/10.1016/j.freeradbiomed.2018.03.042 (2018).

43 França, M. B., Panek, A. D. & Eleutherio, E. C. A. Oxidative stress and its effects during dehydration. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **146**, 621-631, doi:https://doi.org/10.1016/j.cbpa.2006.02.030 (2007).
44 Hossain, M. I. *et al.* SULT4A1 Protects against oxidative-stress induced mitochondrial dysfunction in neuronal cells. *Drug Metab Dispos* **47**, 949-953, doi:10.1124/dmd.119.088047 (2019).
45 Gautheron, J. & Jéru, I. The multifaceted role of spoxide hydrolases in human health and disease. *Int J Mol Sci* **22**, doi:10.3390/ijms22010013 (2020).
46 Mello, A. *et al.* Soluble epoxide hydrolase hepatic deficiency ameliorates alcohol-associated liver disease. *Cell Mol Gastroenterol Hepatol* **11**, 815-830, doi:10.1016/j.jcmgh.2020.10.002 (2021).
47 Zhang, J., Gupta, A. & Storey, K. B. Freezing stress adaptations: Critical elements to activate Nrf2 related antioxidant defense in liver and skeletal muscle of the freeze tolerant wood frogs. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **254**, 110573, doi:https://doi.org/10.1016/j.cbpb.2021.110573 (2021).
48 Li, J. *et al.* Silencing of soluble epoxide hydrolase 2 gene reduces H(2)O(2)-induced oxidative damage in rat intestinal epithelial IEC-6 cells via activating PI3K/Akt/GSK3β signaling pathway. *Cytotechnology* **72**, 23-36, doi:10.1007/s10616-019-00354-x (2020).
49 Li, X. & Wu, X. Soluble epoxide hydrolase (Ephx2) silencing attenuates the hydrogen peroxide-induced oxidative damage in IEC-6 cells. *Arch Med Sci* **17**, 1075-1086, doi:10.5114/aoms.2019.87137 (2021).
50 Katori, M. *et al.* Heme oxygenase-1 overexpression protects rat hearts from cold ischemia/reperfusion injury via an antiapoptotic pathway. *Transplantation* **73**, 287-292, doi:10.1097/00007890-200201270-00023 (2002).
51 Venkatachalam, A. B. *et al.* Delivery of soluble heme oxygenase 1 cell-penetrating peptide into liver cells in *in vitro* and *ex vivo* models of cold ischemia. *Eur Surg Res* **58**, 51-68, doi:10.1159/000451079 (2017).
52 Fitridge, R. & Thompson, M., eds *Mechanisms of Vascular Disease: A Reference Book for Vascular Specialists* (eds Fitridge, R. & Thompson, M.) (University of Adelaide Press, 2011).
53 Poljšak, B. & Milisav, I. Clinical implications of cellular stress responses. *Biomolecules and Biomedicine* **12**, 122-126, doi:10.17305/bjbms.2012.2510 (2012).
54 Sheehan, D., Meade, G., Foley, V. M. & Dowd, C. A. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochem J* **360**, 1-16, doi:10.1042/0264-6021:3600001 (2001).
55 Jakubczyk, K. *et al.* Reactive oxygen species - sources, functions, oxidative damage. *Pol Merkur Lekarski* **48**, 124-127 (2020).
56 Gurdo, N., Novelli Poisson, G. F., Juárez Á, B., Rios de Molina, M. C. & Galvagno, M. A. Improved robustness of an ethanologenic yeast strain through adaptive evolution in acetic acid is associated with its enzymatic antioxidant ability. *J Appl Microbiol* **125**, 766-776, doi:10.1111/jam.13917 (2018).
57 Zhang, Y. *et al.* ERF9 of *Poncirus trifoliata* (L.) Raf. undergoes feedback regulation by ethylene and modulates cold tolerance via regulating a glutathione S-transferase U17 gene. *Plant Biotechnol J* **20**, 183-200, doi:10.1111/pbi.13705 (2022).

58 Seppänen, M. M., Cardi, T., Borg Hyökki, M. & Pehu, E. Characterization and expression of cold-induced glutathione S-transferase in freezing tolerant *Solanum commersonii*, sensitive *S. tuberosum* and their interspecific somatic hybrids. *Plant Sci.* **153**, 125-133, doi:10.1016/s0168-9452(99)00252-6 (2000).
59 Willmore, W. G. & Storey, K. B. Glutathione systems and anoxia tolerance in turtles. *Am J Physiol* **273**, R219-225, doi:10.1152/ajpregu.1997.273.1.R219 (1997).
60 Willmore, W. G. & Storey, K. B. Purification and properties of the glutathione S-transferases from the anoxia-tolerant turtle, *Trachemys scripta elegans*. *Febs J.j* **272**, 3602-3614, doi:10.1111/j.1742-4658.2005.04783.x (2005).
61 Minic, Z., Dahms, T. E. S. & Babu, M. Chromatographic separation strategies for precision mass spectrometry to study protein-protein interactions and protein phosphorylation. *J Chromatogr B Analyt Technol Biomed Life Sci* **1102-1103**, 96-108, doi:10.1016/j.jchromb.2018.10.022 (2018).
62 El-Kabbani, O., Darmanin, C. & Chung, R. P. Sorbitol dehydrogenase: structure, function and ligand design. *Curr Med Chem* **11**, 465-476, doi:10.2174/0929867043455927 (2004).
63 Arakaki, T. L. *et al.* Structure of human brain fructose 1,6-(bis)phosphate aldolase: linking isozyme structure with function. *Protein Sci.* **13**, 3077-3084, doi:10.1110/ps.04915904 (2004).
64 Campbell, E., Schlappal, A., Geller, E. & Castonguay, T. W. in *Nutrition in the Prevention and Treatment of Abdominal Obesity* (ed. Ronald Ross Watson) 197-205 (Academic Press, 2014).
65 Michaelidis, B., Kyriakopoulou-Sklavounou, P., Staikou, A., Papathanasiou, I. & Konstantinou, K. Glycolytic adjustments in tissues of frog *Rana ridibunda* and land snail *Helix lucorum* during seasonal hibernation. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **151**, 582-589, doi:https://doi.org/10.1016/j.cbpa.2008.07.017 (2008).
66 Niu, Y., Zhang, X., Men, S., Storey, K. B. & Chen, Q. Integrated analysis of transcriptome and metabolome data reveals insights for molecular mechanisms in overwintering Tibetan frogs, *Nanorana parkeri*. *Front. Physiol.* **13**, 1104476, doi:10.3389/fphys.2022.1104476 (2022).
67 Varma, A. & Storey, K. B. Freeze-induced suppression of pyruvate kinase in liver of the wood frog (*Rana sylvatica*). *Adv Biol Regul* **88**, 100944, doi:10.1016/j.jbior.2022.100944 (2023).
68 Al-Attar, R., Wijenayake, S. & Storey, K. B. Metabolic reorganization in winter: Regulation of pyruvate dehydrogenase (PDH) during long-term freezing and anoxia. *Cryobiology* **86**, 10-18, doi:10.1016/j.cryobiol.2019.01.006 (2019).
69 Wolfenden, R., Snider, M., Ridgway, C. & Miller, B. The temperature dependence of enzyme rate enhancements. *J. Am. Chem. Soc.* **121**, 7419-7420, doi:10.1021/ja991280p (1999).
70 More, N., Daniel, R. M. & Petach, H. H. The effect of low temperatures on enzyme activity. *Biochem. J.* **305 (Pt 1)**, 17-20, doi:10.1042/bj3050017 (1995).
1 Malhi, Y. *et al.* The role of large wild animals in climate change mitigation and adaptation. *Current Biology* **32**, R181-R196, doi:https://doi.org/10.1016/j.cub.2022.01.041 (2022).
2 Gaughan, J. B., Sejian, V., Mader, T. L. & Dunshea, F. R. Adaptation strategies: ruminants. *Animal Frontiers* **9**, 47-53, doi:10.1093/af/vfy029 (2019).

3 Iverson, S. J. in *Encyclopedia of Marine Mammals (Second Edition)* (eds William F. Perrin, Bernd Würsig, & J. G. M. Thewissen) 115-120 (Academic Press, 2009).
4 Besch, E. L. & Woods, J. E. Heat dissipation biorhythms of laboratory animals. *Lab Anim Sci* **27**, 54-59 (1977).
5 Vejmělka, F. *et al.* Heat dissipation in subterranean rodents: the role of body region and social organisation. *Scientific Reports* **11**, 2029, doi:10.1038/s41598-021-81404-3 (2021).
6 Storey, K. B., Storey, J. M., Brooks, S. P., Churchill, T. A. & Brooks, R. J. Hatchling turtles survive freezing during winter hibernation. *Proc Natl Acad Sci U S A* **85**, 8350-8354, doi:10.1073/pnas.85.21.8350 (1988).
7 Hermes-Lima, M. & Storey, K. B. Antioxidant defenses in the tolerance of freezing and anoxia by garter snakes. *Am J Physiol* **265**, R646-652, doi:10.1152/ajpregu.1993.265.3.R646 (1993).
8 Black, R., Bennett, S. R. G., Thomas, S. M. & Beddington, J. R. Migration as adaptation. *Nature* **478**, 447-449, doi:10.1038/478477a (2011).
9 Constant, T. *et al.* Integrating Mortality Risk and the Adaptiveness of Hibernation. *Front Physiol* **11**, 706, doi:10.3389/fphys.2020.00706 (2020).
10 Li, M. *et al.* The hypoxia adaptation of small mammals to plateau and underground burrow conditions. *Animal Model Exp Med* **4**, 319-328, doi:10.1002/ame2.12183 (2021).
11 Padilla, P. A. & Ladage, M. L. Suspended animation, diapause and quiescence: arresting the cell cycle in C. elegans. *Cell Cycle* **11**, 1672-1679, doi:10.4161/cc.19444 (2012).
12 Ragland, G., Denlinger, D. & Hahn, D. Mechanisms of suspended animation are revealed by transcript profiling in the flesh fly. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 14909-14914, doi:10.1073/pnas.1007075107 (2010).
13 Larson, D. J. *et al.* Wood frog adaptations to overwintering in Alaska: new limits to freezing tolerance. *J Exp Biol* **217**, 2193-2200, doi:10.1242/jeb.101931 (2014).
14 Storey, J. M., Wu, S. & Storey, K. B. Mitochondria and the Frozen Frog. *Antioxidants (Basel)* **10**, doi:10.3390/antiox10040543 (2021).
15 Storey, K. B. & Storey, J. M. Molecular Physiology of Freeze Tolerance in Vertebrates. *Physiological Reviews* **97**, 623-665, doi:10.1152/physrev.00016.2016 (2017).
16 Storey, J. M., Wu, S. & Storey, K. B. Mitochondria and the Frozen Frog. *Antioxidants (Basel)* **10**, doi:10.3390/antiox10040543 (2021).
17 Storey, K. B. & Storey, J. M. Mitochondria, metabolic control and microRNA: Advances in understanding amphibian freeze tolerance. *Biofactors* **46**, 220-228, doi:10.1002/biof.1511 (2020).
18 Tsiouris, J. A. Metabolic depression in hibernation and major depression: an explanatory theory and an animal model of depression. *Med Hypotheses* **65**, 829-840, doi:10.1016/j.mehy.2005.05.044 (2005).
19 Cowan, K. J. & Storey, K. B. Freeze-thaw effects on metabolic enzymes in wood frog organs. *Cryobiology* **43**, 32-45, doi:10.1006/cryo.2001.2338 (2001).
20 Russeth, K. P., Higgins, L. & Andrews, M. T. Identification of proteins from non-model organisms using mass spectrometry: application to a hibernating mammal. *J Proteome Res* **5**, 829-839, doi:10.1021/pr050306a (2006).
21 Churchill, T. & Storey, K. Metabolic responses to dehydration by liver of the wood frog, Rana sylvatica. *Canadian Journal of Zoology* **72**, 1420-1425, doi:10.1139/z94-188 (2011).

22 Al-Attar, R. & Storey, K. B. Effects of anoxic exposure on the nuclear factor of activated T cell (NFAT) transcription factors in the stress-tolerant wood frog. *Cell Biochem Funct* **36**, 420-430, doi:10.1002/cbf.3362 (2018).
23 Hawkins, L. J. *et al.* Glucose and urea metabolic enzymes are differentially phosphorylated during freezing, anoxia, and dehydration exposures in a freeze tolerant frog. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* **30**, 1-13, doi:https://doi.org/10.1016/j.cbd.2019.01.009 (2019).
24 Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254, doi:https://doi.org/10.1016/0003-2697(76)90527-3 (1976).
25 Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* **26**, 1367-1372, doi:10.1038/nbt.1511 (2008).
26 Inui, H. *et al.* Cucurbita pepo contains characteristic proteins without a signal peptide in the xylem sap. *Journal of Plant Physiology* **287**, 154038, doi:https://doi.org/10.1016/j.jplph.2023.154038 (2023).
27 Nagaraj, N. *et al.* System-wide Perturbation Analysis with Nearly Complete Coverage of the Yeast Proteome by Single-shot Ultra HPLC Runs on a Bench Top Orbitrap*. *Molecular & Cellular Proteomics* **11**, M111.013722, doi:https://doi.org/10.1074/mcp.M111.013722 (2012).
28 Lazar, C., Gatto, L., Ferro, M., Bruley, C. & Burger, T. Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies. *Journal of Proteome Research* **15**, 1116-1125, doi:10.1021/acs.jproteome.5b00981 (2016).
29 Wiśniewski, J. R. in *Microbial Proteomics: Methods and Protocols* (ed Dörte Becher) 3-10 (Springer New York, 2018).
30 Muir, T. J., Costanzo, J. P. & Lee Jr., R. E. Metabolic depression induced by urea in organs of the wood frog, *Rana sylvatica*: effects of season and temperature. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* **309A**, 111-116, doi:https://doi.org/10.1002/jez.436 (2008).
31 Miller, L. K. Freezing tolerance in an adult insect. *Science* **166**, 105-106, doi:10.1126/science.166.3901.105 (1969).
32 Somero, G. N. & DeVries, A. L. Temperature tolerance of some Antarctic fishes. *Science* **156**, 257-258, doi:10.1126/science.156.3772.257 (1967).
33 Storey, K. B. & Storey, J. M. Freeze tolerance in animals. *Physiol Rev* **68**, 27-84, doi:10.1152/physrev.1988.68.1.27 (1988).
34 Kiss, A. J., Muir, T. J., Lee, R. E., Jr. & Costanzo, J. P. Seasonal variation in the hepatoproteome of the dehydration and freeze-tolerant wood frog, *Rana sylvatica*. *Int J Mol Sci* **12**, 8406-8414, doi:10.3390/ijms12128406 (2011).
35 Costanzo, J. P. & Jr., R. E. L. Cryoprotectant production capacity of the freeze-tolerant wood frog, *Rana sylvatica*. *Canadian Journal of Zoology* **71**, 71-75, doi:10.1139/z93-011 (1993).
36 do Amaral, M. C., Lee, R. E., Jr. & Costanzo, J. P. Enzymatic regulation of seasonal glycogen cycling in the freeze-tolerant wood frog, *Rana sylvatica*. *J Comp Physiol B* **186**, 1045-1058, doi:10.1007/s00360-016-1012-2 (2016).

37 Gupta, A. & Storey, K. B. Regulation of antioxidant systems in response to anoxia and reoxygenation in *Rana sylvatica*. *Comp Biochem Physiol B Biochem Mol Biol* **243-244**, 110436, doi:10.1016/j.cbpb.2020.110436 (2020).
38 Zhang, J., Gupta, A. & Storey, K. B. Freezing stress adaptations: Critical elements to activate Nrf2 related antioxidant defense in liver and skeletal muscle of the freeze tolerant wood frogs. *Comp Biochem Physiol B Biochem Mol Biol* **254**, 110573, doi:10.1016/j.cbpb.2021.110573 (2021).
39 Dunn, J. & Grider, M. H. in *StatPearls* (StatPearls Publishing Copyright © 2023, StatPearls Publishing LLC., 2023).
40 Shekhovtsov, S. V. *et al.* Metabolic response of the Siberian wood frog *Rana amurensis* to extreme hypoxia. *Sci Rep* **10**, 14604, doi:10.1038/s41598-020-71616-4 (2020).
41 Negishi, M. *et al.* Structure and function of sulfotransferases. *Arch Biochem Biophys* **390**, 149-157, doi:10.1006/abbi.2001.2368 (2001).
42 McGarry, T., Biniecka, M., Veale, D. J. & Fearon, U. Hypoxia, oxidative stress and inflammation. *Free Radical Biology and Medicine* **125**, 15-24, doi:https://doi.org/10.1016/j.freeradbiomed.2018.03.042 (2018).
43 França, M. B., Panek, A. D. & Eleutherio, E. C. A. Oxidative stress and its effects during dehydration. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **146**, 621-631, doi:https://doi.org/10.1016/j.cbpa.2006.02.030 (2007).
44 Hossain, M. I. *et al.* SULT4A1 Protects Against Oxidative-Stress Induced Mitochondrial Dysfunction in Neuronal Cells. *Drug Metab Dispos* **47**, 949-953, doi:10.1124/dmd.119.088047 (2019).
45 Gautheron, J. & Jéru, I. The Multifaceted Role of Epoxide Hydrolases in Human Health and Disease. *Int J Mol Sci* **22**, doi:10.3390/ijms22010013 (2020).
46 Mello, A. *et al.* Soluble Epoxide Hydrolase Hepatic Deficiency Ameliorates Alcohol-Associated Liver Disease. *Cell Mol Gastroenterol Hepatol* **11**, 815-830, doi:10.1016/j.jcmgh.2020.10.002 (2021).
47 Zhang, J., Gupta, A. & Storey, K. B. Freezing stress adaptations: Critical elements to activate Nrf2 related antioxidant defense in liver and skeletal muscle of the freeze tolerant wood frogs. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **254**, 110573, doi:https://doi.org/10.1016/j.cbpb.2021.110573 (2021).
48 Li, J. *et al.* Silencing of soluble epoxide hydrolase 2 gene reduces H(2)O(2)-induced oxidative damage in rat intestinal epithelial IEC-6 cells via activating PI3K/Akt/GSK3β signaling pathway. *Cytotechnology* **72**, 23-36, doi:10.1007/s10616-019-00354-x (2020).
49 Li, X. & Wu, X. Soluble epoxide hydrolase (Ephx2) silencing attenuates the hydrogen peroxide-induced oxidative damage in IEC-6 cells. *Arch Med Sci* **17**, 1075-1086, doi:10.5114/aoms.2019.87137 (2021).
50 Katori, M. *et al.* Heme oxygenase-1 overexpression protects rat hearts from cold ischemia/reperfusion injury via an antiapoptotic pathway. *Transplantation* **73**, 287-292, doi:10.1097/00007890-200201270-00023 (2002).
51 Venkatachalam, A. B. *et al.* Delivery of Soluble Heme Oxygenase 1 Cell-Penetrating Peptide into Liver Cells in in vitro and ex vivo Models of Cold Ischemia. *Eur Surg Res* **58**, 51-68, doi:10.1159/000451079 (2017).
52 in *Mechanisms of Vascular Disease: A Reference Book for Vascular Specialists* (eds R. Fitridge & M. Thompson) (University of Adelaide Press

53 Poljšak, B. & Milisav, I. Clinical implications of cellular stress responses. *Biomolecules and Biomedicine* **12**, 122-126, doi:10.17305/bjbms.2012.2510 (2012).
54 Sheehan, D., Meade, G., Foley, V. M. & Dowd, C. A. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochem J* **360**, 1-16, doi:10.1042/0264-6021:3600001 (2001).
55 Jakubczyk, K. *et al.* Reactive oxygen species - sources, functions, oxidative damage. *Pol Merkur Lekarski* **48**, 124-127 (2020).
56 Gurdo, N., Novelli Poisson, G. F., Juárez Á, B., Rios de Molina, M. C. & Galvagno, M. A. Improved robustness of an ethanologenic yeast strain through adaptive evolution in acetic acid is associated with its enzymatic antioxidant ability. *J Appl Microbiol* **125**, 766-776, doi:10.1111/jam.13917 (2018).
57 Zhang, Y. *et al.* ERF9 of Poncirus trifoliata (L.) Raf. undergoes feedback regulation by ethylene and modulates cold tolerance via regulating a glutathione S-transferase U17 gene. *Plant Biotechnol J* **20**, 183-200, doi:10.1111/pbi.13705 (2022).
58 Seppänen, M. M., Cardi, T., Borg Hyökki, M. & Pehu, E. Characterization and expression of cold-induced glutathione S-transferase in freezing tolerant *Solanum commersonii*, sensitive *S. tuberosum* and their interspecific somatic hybrids. *Plant Sci* **153**, 125-133, doi:10.1016/s0168-9452(99)00252-6 (2000).
59 Willmore, W. G. & Storey, K. B. Glutathione systems and anoxia tolerance in turtles. *Am J Physiol* **273**, R219-225, doi:10.1152/ajpregu.1997.273.1.R219 (1997).
60 Willmore, W. G. & Storey, K. B. Purification and properties of the glutathione S-transferases from the anoxia-tolerant turtle, *Trachemys scripta elegans*. *Febs j* **272**, 3602-3614, doi:10.1111/j.1742-4658.2005.04783.x (2005).

Chapter 3: Proteomic Study of Breast Cancer-derived Extracellular Vesicles

This chapter has been published in a peer-reviewed journal, Biomedicines.

Phosphoproteomic Analysis of Breast Cancer-Derived Small Extracellular Vesicles Reveals Disease-Specific Phosphorylated Enzymes

Zoran Minic [1,*], Nico Hüttmann [2], Suttinee Poolsup [2], Yingxi Li [2], Vanessa Susevski [2], Emil Zaripov [2] and Maxim V. Berezovski [1,2,*]

[1] John L. Holmes Mass Spectrometry Facility, Faculty of Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada

[2] Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, ON K1N 6N5, Canada

[*] Correspondence: Dr. Zoran Minic (Z.M.); Dr. Maxim V. Berezovski (M.V.B.)

Lysine Acetylome of Breast Cancer-Derived Small Extracellular Vesicles Reveals Specific Acetylation Patterns for Metabolic Enzymes

Zoran Minic [1,*], Yingxi Li [2] , Nico Hüttmann [1] , Gurcharan K. Uppal [2] , Rochelle D'Mello [2] and Maxim V. Berezovski [1,2,*]

[1] John L. Holmes Mass Spectrometry Facility, Faculty of Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada

[2] Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, ON K1N 6N5, Canada

[*] Correspondence: Dr. Zoran Minic (Z.M.); Dr. Maxim V. Berezovski (M.V.B.)

Proteomics Approaches for the Discovery of Potential Enzymatic Biomarkers for Early Diagnosis of Breast Cancer

Yingxi Li [1], Nico Hüttmann [2], Zoran Minic [2,*] and Maxim V. Berezovski [1,2,*]

[1] Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, ON K1N 6N5, Canada

[2] John L. Holmes Mass Spectrometry Facility, Faculty of Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada

[*] Correspondence: Dr. Zoran Minic (Z.M.); Dr. Maxim V. Berezovski (M.V.B.)

3.1 General Introduction

Breast cancer, one of the most common cancers affecting women worldwide, poses significant challenges in early detection, prognosis, and treatment. Traditional diagnostic methods, such as mammography and biopsy, while effective, have limitations in sensitivity and specificity, particularly in early-stage detection. Statistic results indicate that early detection of BC greatly enhances patient outcomes and can prevent the progression of symptoms [1]. Consequently, there is a critical need for more precise and non-invasive biomarkers that can improve early diagnosis, monitor disease progression, and possibly predict treatment responses.

Small extracellular vesicles are membrane-bound particles released by cells into the extracellular environment. They play a crucial role in cell-to-cell communication, as they carry a variety of molecules, including proteins, lipids and nucleic acids [2]. Through the transfer of these molecules, sEVs influence various aspects of tumor development, metastasis, and resistance to therapy [2]. Recent studies show that proteomics offers a promising avenue for discovering such biomarkers from small extracellular vesicles (sEVs) by providing a comprehensive analysis of protein expression, post-translational modifications (PTMs), and interactions [3-6]. Many PTMs, including phosphorylation and acetylation, have been linked to tumor progression, growth, and survival by altering the normal functions of proteins in tumor cells [7].

This chapter provides proteomic analysis from three approaches, quantitative, phosphoproteomics, and proteomics acetylation analysis, in analyzing sEVs proteins from BC cell lines. It also reports the distinct enzymes found only in cancerous cell lines, and some of them were validated by enzymatic assays.

3.1.1 References

1. DeSantis, C. E. *et al.* Cancer treatment and survivorship statistics, 2014. *CA Cancer J Clin* **64**, 252-271, doi:10.3322/caac.21235 (2014).
2. Wang, Y. *et al.* Small Extracellular Vesicles: Functions and Potential Clinical Applications as Cancer Biomarkers. *Life (Basel)* **11**, doi:10.3390/life11101044 (2021).
3. Lane, R. E., Korbie, D., Hill, M. M. & Trau, M. Extracellular vesicles as circulating cancer biomarkers: opportunities and challenges. *Clin Transl Med* **7**, 14, doi:10.1186/s40169-018-0192-7 (2018).
4. Urabe, F. *et al.* Extracellular vesicles as biomarkers and therapeutic targets for cancer. *American Journal of Physiology-Cell Physiology* **318**, C29-C39, doi:10.1152/ajpcell.00280.2019 (2020).

5. Xu, F. *et al.* Tumor-derived extracellular vesicles as a biomarker for breast cancer diagnosis and metastasis monitoring. *iScience* **27**, 109506, doi:https://doi.org/10.1016/j.isci.2024.109506 (2024).
6. Lee, Y. *et al.* Recent advances of small extracellular vesicle biomarkers in breast cancer diagnosis and prognosis. *Molecular Cancer* **22**, 33, doi:10.1186/s12943-023-01741-x (2023).
7. Srivastava, A. K., Guadagnin, G., Cappello, P. & Novelli, F. Post-Translational Modifications in Tumor-Associated Antigens as a Platform for Novel Immuno-Oncology Therapies. *Cancers (Basel)* **15**, doi:10.3390/cancers15010138 (2022).

3.2 Phosphoproteomic Analysis of Breast Cancer-Derived Small Extracellular Vesicles Reveals Disease-Specific Phosphorylated Enzymes

3.2.1 Abstract

Small membrane-derived extracellular vesicles have been proposed as participating in several cancer diseases, including breast cancer (BC). We performed a phosphoproteomic analysis of breast cancer-derived small extracellular vesicles (sEVs) to provide insight into the molecular and cellular regulatory mechanisms important for breast cancer tumor progression and metastasis. We examined three cell line models for breast cancer: MCF10A (non-malignant), MCF7 (estrogen and progesterone receptor-positive, metastatic), and MDA-MB-231 (triple-negative, highly metastatic). To obtain a comprehensive overview of the sEV phosphoproteome derived from each cell line, effective phosphopeptide enrichment techniques IMAC and TiO_2, followed by LC-MS/MS, were performed. The phosphoproteome was profiled to a depth of 2003 phosphopeptides, of which 207, 854, and 1335 were identified in MCF10A, MCF7, and MDA-MB-231 cell lines, respectively. Furthermore, 2450 phosphorylation sites were mapped to 855 distinct proteins, covering a wide range of functions. The identified proteins are associated with several diseases, mostly related to cancer. Among the phosphoproteins, we validated four enzymes associated with cancer and present only in sEVs isolated from MCF7 and MDA-MB-231 cell lines: ATP citrate lyase (ACLY), phosphofructokinase-M (PFKM), sirtuin-1 (SIRT1), and sirtuin-6 (SIRT6). With the exception of PFKM, the specific activity of these enzymes was significantly higher in MDA-MB-231 when compared with MCF10A-derived sEVs. This study demonstrates that sEVs contain functional metabolic enzymes that could be further explored for their potential use in early BC diagnostic and therapeutic applications.

Keywords: ATP citrate lyase (ACLY); breast cancer; phosphofructokinase-M (PFKM); phosphoproteomics; sirtuin-1 (SIRT1); sirtuin-6 (SIRT6); small extracellular vesicles.

3.2.2 Introduction

Extracellular vesicles (EVs) derived from human cancer cell lines are involved in multiple biological processes in tumor biology, such as modulation of the microenvironment, angiogenesis, sustained growth, and tissue invasion [1-7]. EVs act as transport vectors or signal transducers that can deliver specific biological information by transferring bioactive content (nucleic acids,

proteins, and metabolites) from donor to nearby or distant cells [8-10]. Therefore, EVs emerged as key regulators of cell–cell communication within multicellular organisms, in health and disease. EVs include several subtypes of membrane-bound vesicles, including exosomes and microvesicles, distinguished by their biogenesis pathway. These subtypes may be roughly categorized by their measured diameter, with small extracellular vesicles (sEVs) or exosomes in the range of 30 to 150 nm [11,12] and medium extracellular vesicles (mEVs) in the range of 50 nm to 1000 nm in diameter [13], this study's vesicles with a measured diameter in the range of 50 to 300 nm are referred to as sEVs.

In order to elucidate the molecular changes that coincide with cancer or discover new biomarkers, it is important to measure protein expression in EVs, followed by the identification of post-translationally modified proteins such as phosphorylated proteins. Protein phosphorylation is the most important post-translational modification that dominates signaling transduction, which plays an essential role for almost all cellular functions and is involved in major regulatory mechanisms of cell signaling networks [14]. Recent studies have indicated that certain phosphoproteins encapsulated within EVs function as key regulators in the tumor microenvironment [15,16]. A recent study found high expression levels of tyrosine kinases (RTKs), such as phosphorylated EGFR and HER2, in MCF7 cell line-derived exosomes; these sEVs were capable of stimulating the MAPK pathway in monocytes through the transport of functional RTKs, leading to the inactivation of apoptosis-related caspases [17]. It has been reported that human colorectal cancer exosomes, derived from the SW620 cell line, were found to contain 313 phosphoproteins with 1091 phosphosites, of which 202 were newly discovered [16]. These exosome-derived phosphoproteins had a remarkably high level of tyrosine-phosphorylated proteins (6.4%) which were functionally relevant to ephrin signalling pathway-directed cytoskeleton remodeling.

Some studies reported the detection of phosphoproteins in biofluids for disease diagnosis [18-20]. For instance, an in-depth analysis of phosphoproteomes in plasma from both microvesicles and exosomes measured phosphorylation changes across patients with breast cancer (BC) and healthy individuals [18]. Using label-free quantitative phosphoproteomics, 144 phosphoproteins were identified in plasma EVs, which were more abundantly expressed in patients diagnosed with breast cancer compared with healthy controls [18]. In another quantitative EV phosphoproteomic study, plasma samples from patients diagnosed with kidney cancer were analyzed in order to identify direct markers of cellular signaling and disease progression [19]. The results of these analyses

revealed 28 proteins present in kidney cancer samples that were not detected in the control samples. Several EV isolation methods have been assessed for highly efficient capture of EVs from human urine samples [20]. For example, using the EVTRAP isolation method, close to 2000 unique phosphopeptides were identified from 10 mL of urine. Data obtained for protein phosphorylation in EVs presents a potential opportunity for understanding cancer signaling and early-stage cancer diagnosis.

Cell lines are commonly used as models in cancer biology since they are easily grown and relatively inexpensive, making them suitable for high-throughput testing and omics studies [21]. Blood is a systemic source of EVs from all tissue types in the body, not only breast tumors; therefore, it is difficult to select appropriate biomarkers relevant to breast cancer (BC). Comparative analyses of body fluids and cell lines may help find more tissue specific biomarkers. In recently published work, the difference in proteomes between cancerous (MCF7 and MDA-MB-231) and noncancerous (MCF10A) cell line-derived EVs has been investigated [6,7]. The results revealed that 87 sEV [6] and 112 mEV [7] proteins relevant to BC in which MDA-MB-231 cell line-derived sEV proteins were proposed as potential breast cancer biomarkers for disease diagnosis and prognosis as well as for potential therapeutic targets and resistance against chemotherapy agents. Moreover, the study also validated three enzymes—ornithine aminotransferase (OAT), transaldolase (TALDO1) and bleomycin hydrolase (BLMH)—using standard enzymatic assays. Two of the three enzymes, OAT and TALDO1, had significantly higher specific enzymatic activities in MDA-MB-231-derived sEVs than MCF10A. BLMH was found to be highly expressed in MDA-MB-231 microvesicles (MVs) when compared with MCF10A-derived MVs. EV enzymes have not been previously investigated in a comprehensive manner. Therefore, this study aimed to provide support for the investigation of enzymes derived from EVs for BC diagnosis and therapeutic applications.

The aim of this study was to gain insight into the phosphoproteome of sEVs derived from the metastatic BC cell lines MDA-MB-231 and MCF7 and from the non-cancerous cell line MCF10A. Differences in phosphoproteomes were observed between the control cell line, MCF10A, and BC cell lines, MDA-MB-231 and MCF7. We selected and validated four phosphoenzymes: ATP citrate lyase (ACLY), phosphofructokinase-M (PFKM), sirtuin-1 (SIRT1), and sirtuin-6 (SIRT6) identified in both MDA-MB-231 and MCF7 that might be potentially relevant to BC diagnosis.

3.2.3 Materials and Methods

3.2.3.1 Cell Culturing and sEV Isolation

Breast cancer epithelial cell lines MDA-MB-231 (ATCC HTB-26) and MCF7 (ATCC HTB-22) and the non-tumorigenic breast epithelial cell line MCF10A (ATCC CRL-10317) were obtained from the American Type Culture Collection (ATCC) and used in this study. MDA-MB-231 and MCF10A cells were cultured as described in our previous work [6]. MCF7 cells were cultured in EV-depleted DMEM/Ham's F12 (GIBCO-Invitrogen) supplemented with 10% fetal bovine serum (Sigma Life Science, St. Louis, USA). The three cell lines were plated in increased attachment cell culture dishes (VWR) and grown for 7 days. Each liter of cell culture supernatant was harvested from approximately 4×10^8 cells.

3.2.3.2 Differential Ultracentrifugation (UC)

Small EVs were isolated from cell culture supernatant by UC, as described previously [6]. Conditioned medium (240 mL) was harvested after 7 days of incubation and immediately centrifuged at 300× *g* for 10 min. Apoptotic bodies were removed by centrifugation at 2000× *g* using a Sigma13190 rotor (MBI) for 20 min. Samples were then spun at 16,500× *g* for 1 h using an SW28 Ti rotor (Beckman Coulter, Indianapolis, USA) to deplete microvesicles. To pellet sEVs, the same rotor was used to centrifuge samples for 3 h at 100,000× *g*. Collected sEVs were washed with PBS and centrifuged at 100,000× *g* for 1 h and finally resuspended in 200 μL PBS and stored at −80 °C. A Bradford protein assay kit (Thermo Scientific, Rockford, USA, Cat No. 23236) was used to assess sample protein concentrations. The average amount of EV protein from three replicate samples was roughly 22 ± 5, 27 ± 6 and 34 ± 6 μg for MCF10A, MCF7 and MDA-MB-231 cell lines, respectively. These samples were used for proteomic analysis.

3.2.3.3 Quantification of sEVs by Nanoparticle Tracking Analysis (NTA)

The average particle diameter and concentration of isolated EV samples were measured by the ZetaView PMX-110 (Particle Metrix, Meerbusch, Germany). The camera shutter speeds were adjusted to 85 and 40. The instrument was calibrated and focused with 102 nm polystyrene beads (Microtrac, Cat No. 900383).

3.2.3.4 Sample Preparation for Phosphoproteomics

Isolated sEVs were resuspended in the lysis buffer with a volume ratio of 4:1 (fraction/buffer) with a final concentration of 20 mM HEPES, pH 8.0, 5% glycerol, 0.1% n-Dodecyl-β-D-Maltoside, 0.2 mM DTT, 1.6 M urea, and 1/250 (v:v) phosphatase inhibitors (phosphatase inhibitor 2 and 3;

Sigma: Cat No. P0044 and P5726) and gently vortexed for 3 min. The suspension was centrifuged for 1 min at 10,000× *g*, and the supernatant was then collected.

3.2.3.5 Phosphopeptide Enrichment by IMAC and TiO_2

Protein samples obtained from the isolated sEVs were reduced, alkylated, digested, and desalted, as described previously [6]. Phosphopeptide enrichment by IMAC was performed according to the protocol of Pierce Fe-NTA phosphopeptide enrichment kit (Thermo Scientific, Cat No. 88300). TiO_2-phosphopeptide enrichment was carried out according to the manufacturer's protocol (Pierce TiO_2 phosphopeptide enrichment and clean-up kit, Cat No. 88301). All enriched phosphopeptide samples were acidified by adding TFA (1% final concentration), subsequently desalted (TopTip C-18 columns; Glygen Corp.), and dried in a vacuum evaporator.

3.2.3.6 Nano-LC-MS/MS

To process samples, the Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific, Mississauga, ON, Canada) coupled to an UltiMate 3000 nanoRSLC (Thermo Fisher Scientific, Mississauga, ON, Canada) was utilized, as previously described with modified instrument parameters [6,7]. Following the reconstitution of digested peptides with 20 µL MS-grade H_2O/1% formic acid (*v*/*v*), three microliters of sample was injected onto an in-house packed column and eluted for 65 min at a flow rate of 300 nL/min (0–10 min, 2–2% ACN; 10–40 min, 2–38% ACN; 40–45 min, 38–98% ACN; 45–50 min, 98–98% ACN; 50–55 min, 98–2% ACN; 55–65 min, 2–2% ACN). The ESI+ parameters were set as follows: top speed mode, ion source temperature 250 °C, ion spray voltage 2.1 kV, and a full-scan MS (m/z 350–2000) resolution of 60,000. For collision-induced dissociation (CID), the automatic gain control (AGC) target was set to 5×10^5 for full scans and 1×10^4 for MS/MS scans. Precursor ions were filtered from +2 to +7 charge states within 2 m/z isolation windows. CID in the linear ion trap was performed at a normalized collision energy of 35%. For the higher-energy collision dissociation (HCD) method, the AGC target for precursor ions was set to 5×10^5 with an ion filling time of 150 ms. The highest intensity ions were isolated and fragmented with a normalized collision energy of 32% and detected at a mass resolution of 15,000. The AGC target for MS/MS was set to 5×10^4 with a maximum injection time of 150 ms and a dynamic exclusion of 30 s.

3.2.3.7 MS Spectra Processing

MS raw files were analyzed with MaxQuant (version 2.0.1.0) [22] and the built-in Andromeda search engine [23], as previously described with slight modifications [6]. Peptides were searched

against the human UniProt FASTA file, containing 20,396 entries (21 April 2021), and a default contaminants database. MaxQuant's default parameters were used unless otherwise stated. In addition to N-terminal acetylation and methionine oxidation, phosphorylation of serine, threonine, and tyrosine were set as variable modifications. Meanwhile, carbamidomethylation of cysteine was set as a fixed modification. Tryptic peptides with a minimum of 6 amino acids and a molecular weight maximum of 4600 Da were searched with a maximum of two missed cleavages. The initial precursor mass deviation was set to 10 ppm, while a fragment mass deviation of 0.5 Da was used. The false discovery rate (FDR) was set to 0.01 for peptide spectrum matches (PSM) and protein identification, using a reverse sequence decoy database.

3.2.3.8 Data Filtering and Phosphorylation Site Localization

MaxQuant output tables for protein groups, peptides, and phosphosites were used for all analyses in R [24]. We only retained phosphopeptides, which were measured in at least two out of the three replicates. Remaining phosphopeptides were used to classify phosphosites based on their combined localization probability into class I, II and III, and IV [25]. Known phosphosites and kinase-substrate interactions were downloaded from PhosphoSitePlus (30 July 2021) [26]. In addition, kinases were predicted for remaining phosphosites using GPS 5.0, with the threshold parameter set to high [27]. Predicted kinases were represented as kinase groups and related to previously identified kinases in MDA-MB-231 sEVs [6].

3.2.3.9 Disease and Functional Annotation Analysis

DisGeNET protein–disease annotations were retrieved with the disgenet2r R package [28]. Cancer- and breast cancer-related proteins were identified by semantically related terms such as "mammary carcinoma" or "malignant neoplasm of breast". Gene ontology functional annotations were obtained with the org.Hs.eg.db annotation package [29-31]. Absolute number of proteins for each biological term was compared between cell lines.

3.2.3.10 Data Availability

All MS raw data were submitted to the PRIDE repository (Accession: PXD030424) at the European Bioinformatics Institute.

3.2.3.11 ATP-Citrate Synthase Activity Assay

Cells and sEVs were lysed by gentle vortexing in extraction buffer provided by the ACLY Activity Assay Kit (AMSBIO, Cat No. 79904). Supernatants were collected and used for the

analyses of enzymatic activities. ACLY Activity Assay Kit was used according to the manufacturer's instructions. The assays were performed at room temperature for 60 min.

3.2.3.12 Phosphofructokinase Activity Assay

The phosphofructokinase activity was measured using kits from Sigma-Aldrich (Phosphofructokinase Activity Colorimetric Assay Kit MAK093) according to manufacturer's instructions. The assays were performed at room temperature in 100 µL of reaction mixture. The phosphofructokinase activity was calculated via the standard curve of the NADH standard at known concentrations. The samples were mixed and incubated for 30 min, and the absorbance was monitored at 450 nm.

3.2.3.13 SIRT1 and SIRT6 Activity Assay

The enzymatic activity of SITR1 and SIRT6 was assessed by using the Fluorescent Screening Assay Kit (ab156065, ab156068; Abcam, Toronto, Canada) following the manufacturer's instructions. The protein extract from cells and EVs was obtained by treatment with Triton X-100 (final concentration 0.5%). The assay was performed in 96-well black microplate (Greiner Bio-One 655209; Fischer Scientific, Monroe, USA) with a reaction volume of 50 µL per well. Briefly, the reaction was started by incubating the protein extract with the reaction mixture containing an acetylated peptide substrate. Samples were incubated for 30 min at 23 °C. Control samples were prepared in absence of NAD^+. Fluorescence intensities of SIRT1 (λ_{ex} = 340 nm, λ_{em} = 450 nm) and SIRT6 (λ_{ex} = 488 nm, λ_{em} = 530 nm) were measured using a microplate reader (FilterMax F3 and F5 Multi-Mode Microplate Readers from Molecular Devices). The activity of enzymes was calculated from the assay time between 5 and 10 min.

3.2.3.14 Western Blot

SDS electrophoresis and Western blot experiments were carried out as previously described [6].

3.2.4 Results

3.2.4.1 Isolation of sEV

In this work, we carried out a systematic analysis of the phosphoproteome of sEVs derived from MCF10A, MCF7, and MDA-MB-231 cell lines to provide insight into the molecular mechanism of breast cancer. To evaluate the range of measured diameters of sEVs isolated by differential ultracentrifugation, nanoparticle tracking analysis was performed. Isolated vesicles, from the three cell lines, measured 50 to 300 nm in diameter, with an average size of about 125

nm (Figure 3.2.1A). Additionally, Western blots were performed to validate the presence of EV markers CD9, CD63, and CD81. The results revealed the presence of markers in all free sEV fractions derived from three cell lines: MCF10A, MCF7, and MDA-MB-231 (Figure 3.2.1B and Figure S1).

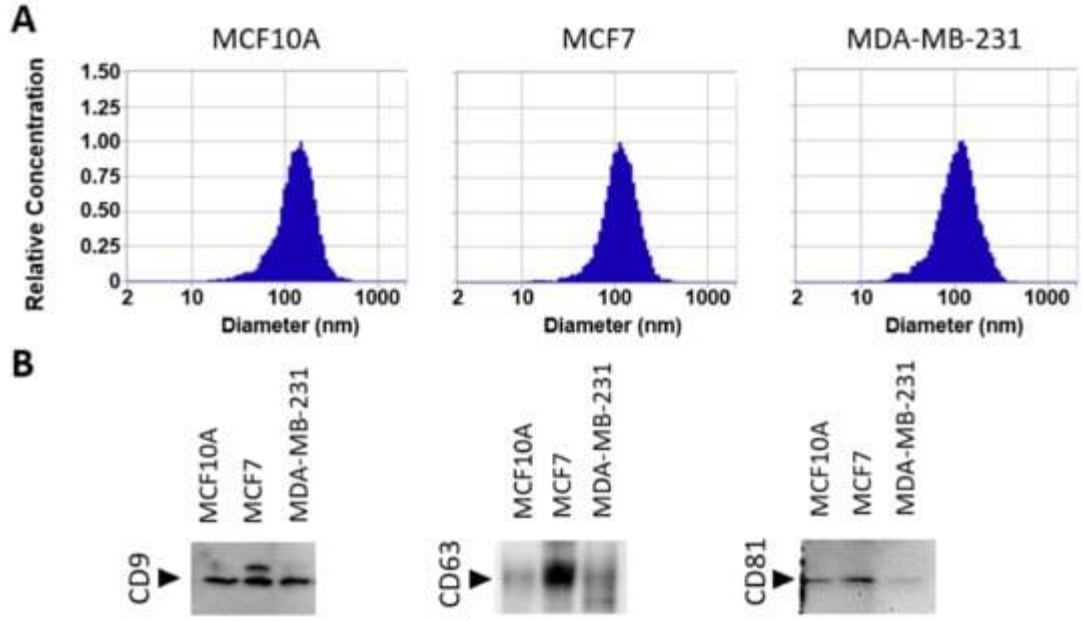

Figure 3.2.1. Characterization of extracellular vesicles by nanoparticle tracking analysis (NTA) and Western blot. (A) NTA characterization showing size distributions of isolated sEVs from MCF10A, MCF7, and MDA-MB-231 cell lines. (B) Western blot analysis of sEV marker proteins, CD9, CD63, and CD81.

3.2.4.2 Overall Phosphoproteome Profiling

A total of three biological replicates were prepared from MCF10A-, MCF7-, and MDA-MB-231-derived sEVs and examined independently. EVs were buffered with phosphatase inhibitors, digested, and the phosphopeptides were enriched to enhance the identification of low-abundance phosphorylated proteins. In addition, we combined two techniques for the enrichments of phosphopeptides: Fe-IMAC and TiO_2 affinity chromatography. An overview of our experimental strategy is presented in Figure 3.2.2A. The enriched phosphopeptides were analyzed by ultra-high-performance nLC-nESI-MS/MS using two fragmentation modes (CID and HCD), and the obtained data were subjected to rigorous assessment and peptide identification. Each sample represents the data obtained from peptides of three independent experiments after filtering as described in

Materials and Methods (Section 3.2.3.8, Data Filtering). The reliable phosphopeptides were considered if they were identified in at least two biological replicates. After removal of potential contaminants, the total number of identified phosphopeptides for each sEV sample, using two enrichment and fragmentation methods, are presented in Figure 3.2.2B and Table S1. Significant differences were observed using CID and HCD fragmentation for both enrichment methods and Fe-IMAC and TiO_2 affinity chromatography from all sEVs (Figure S2). In addition, a higher number of phosphopeptides was obtained using IMAC in comparison with TiO_2. In total, from 2003 phosphopeptides, 207, 854, and 1335 were identified in MCF10A, MCF7, and MDA-MB-231 sEVs, respectively. Among all identified phosphosites (2450), the probability of correct site analysis identified a total of 1613 class I (>75% confidence), 774 class II and III (25–75% confidence), and 63 class IV (<25% confidence) phosphosites, respectively [25]. Among Class I phosphosites, 60 are novel phosphorylated sites, not previously reported (Table S2. Identified phosphopeptides correspond to 145, 462, and 587 phosphoproteins in MCF10A-, MCF7-, and MDA-MB-231-derived sEVs, respectively. The Venn diagram of combined phosphoproteins from three cell lines revealed a total of 855 distinct proteins (Figure 3.2.2D).

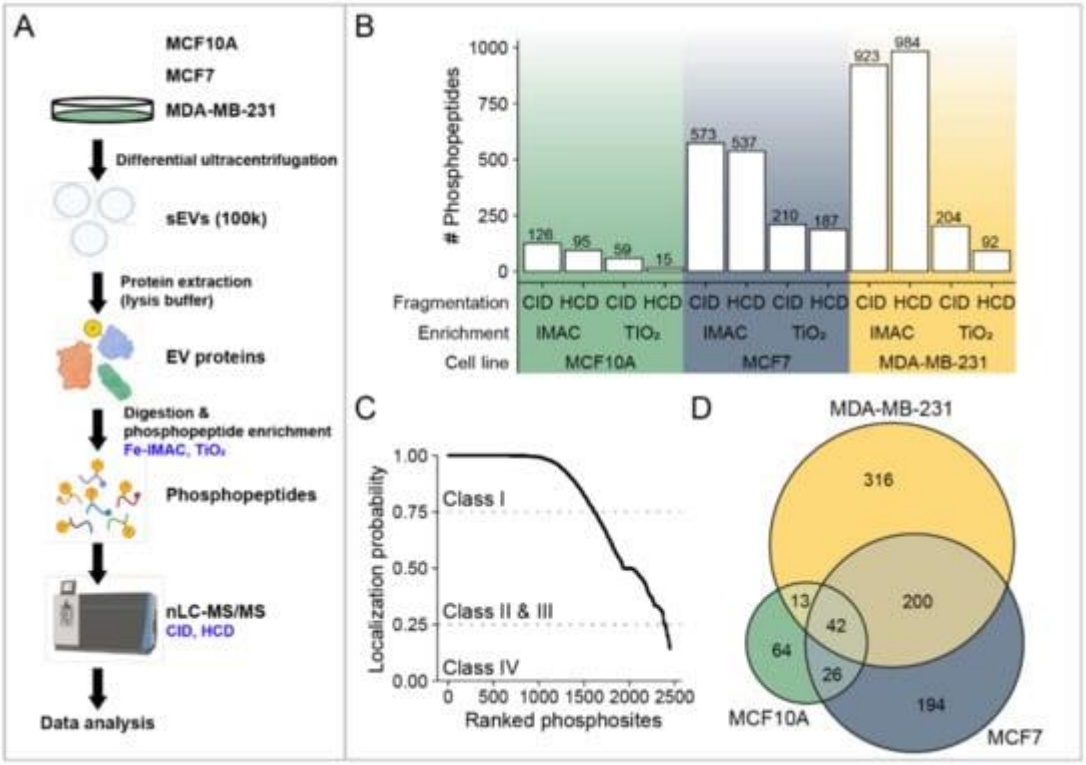

Figure 3.2.2. Overview of the identified phosphopeptides and phosphoproteins. (A) Workflow illustration of phosphoproteome analysis of sEVs from MCF10A, MCF7, and MDA-MB-231 cells.

(B) Total number of phosphopeptides identified using Fe-IMAC and TiO_2 enrichment methods, as well as two fragmentation methods: CID and HCD. (C) The probability of correct site identification in peptide sequences. (D) A Venn diagram showing the number of EV phosphoproteins identified from the sEVs of three cell lines.

3.2.4.3 Phosphorylation Site Distributions, Phosphorylation Motifs and Predicted Potential Kinases for Identified Phosphorylation Sites

Figure 3.2.3A presents the distribution of phosphosite residues for different cell line-derived sEVs. In total, 1987 phosphoserine (pS), 433 phosphothreonine (pT), and 30 phosphotyrosine (pY) sites were identified (Table S2). Similar distributions of pS, pT, and pY sites were observed among sEVs from the three cell lines. The result revealed that pS was most strongly represented, followed by pT and pY sites. The distribution of the number of phosphorylation sites localized on each protein from the three cell lines is given in Figure 3.2.3B. A single phosphorylation site was localized on most of the identified phosphoproteins in all three sEV fractions. An important number of the identified phosphoproteins was found to be phosphorylated more than once. Small EVs derived from MCF7 (275 phosphoproteins; 60.0%) and MDA-MB-231 (355 phosphoproteins; 62.3%) cell lines contained a higher number of multiply phosphorylated proteins than MCF10A (78 phosphoproteins; 54.5%) cells (Figure 3.2.3B). Some proteins contained a high number of phosphorylated sites, such as serine/arginine repetitive matrix protein 2 (SRRM2) (72 in MDA-MB-231 EVs, 81 MCF7 EVs, and only 3 in MCF10A). Other proteins that phosphorylated multiple times were mainly present in MDA-MB-231 sEVs, and MCF7 sEVs include Bcl-2-associated transcription factor 1 (BCLAF1), Serine/threonine-protein kinase PRP4 homolog (PRPF4B) and Thyroid hormone receptor-associated protein 3 (THRAP3) (Table S3).

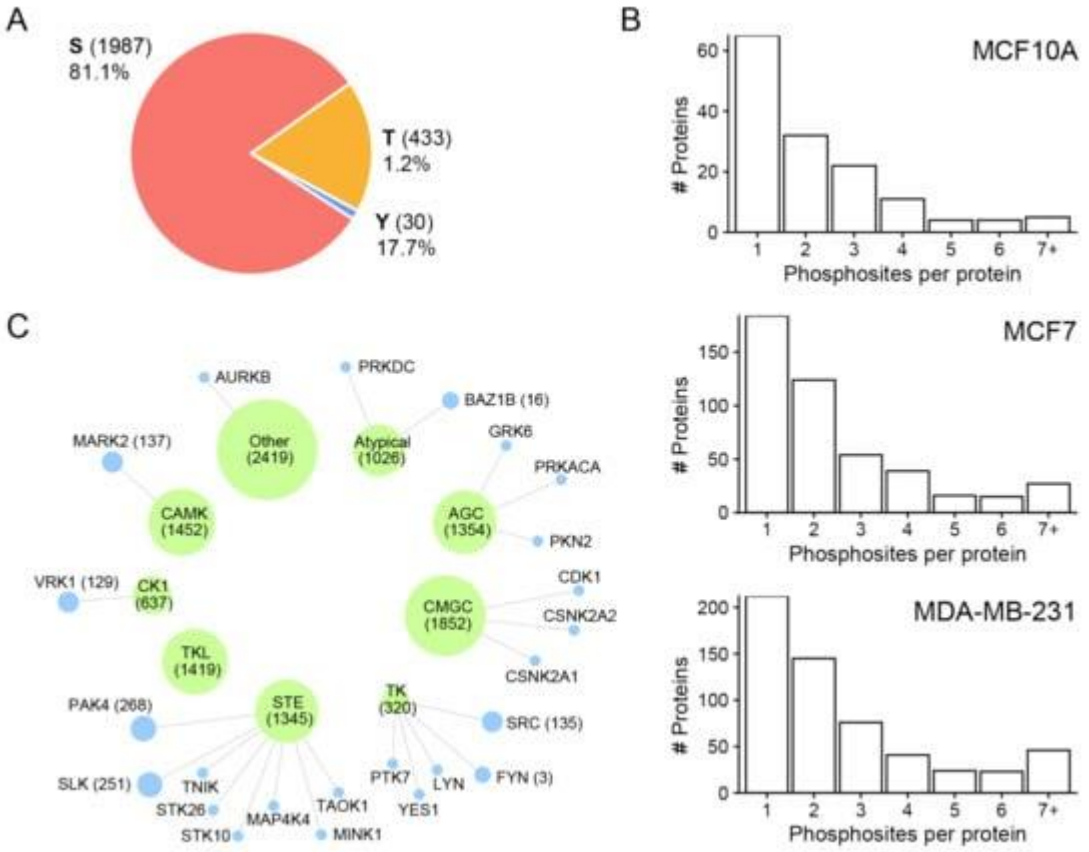

Figure 3.2.3. Overview of the identified phosphosites and frequency distribution of phosphorylated amino acid and phosphorylation motifs. (A) Frequency distribution of phosphorylated amino acid: pS, pY, and pY. (B) Numbers of sites observed per protein in all three sEV fractions. (C) Pie chart representation of predicted kinases groups (green) and their link to identified kinases (blue) in MDA-MB-231 sEVs [6]. Node size represents the scaled number of predicted phosphorylation sites.

We visualized sequence motifs for the phosphorylated amino acids for each cell line fraction (Figure S3). The majority of phosphorylation sites that contained serine and threonine residues were followed by proline motifs (P at +1). These motifs for pS and pT are well known to be targeted mitogen-activated protein kinases (MAPKs) and cyclin-dependent kinases (CDKs) [32]. Significantly enriched S-based motifs were (D,E+1). The substrate of peptides containing D/E-rich motifs belonged to the casein kinase 1 (CK1) and 2 (CK2) families and some other kinases [33].

In addition, we used the GPS 5.0 software [27] to predict which kinases are responsible for phosphorylation of proteins at the identified sites. Among the phosphorylation sites identified in this analysis, most of them were predicted to be phosphorylated by groups of kinase family such

as CMGC, CAMKs, STE, AGC, TKL, CK1, and TK (Figure 3.2.3C). These results were consistent with those derived from the phosphosite sequence motifs. Kinases previously identified in MDA-MB-231 EVs [6] mainly belonged to the STE and TK kinase families that largely include subfamily members p21-activated kinase 4 (PAK4), STE20-like serine/threonine-protein kinase (SLK), tyrosine-protein kinase Fyn (FYN), and proto-oncogene tyrosine-protein kinase Src (SRC).

3.2.4.4. Identified Phosphoproteins in the Context of Cancer/Breast Cancer

To determine whether identified phosphoproteins are linked to cancer diseases, we annotated proteins using the recent set of disease annotations from the DisGeNET database [28]. Phosphoproteins are associated with several diseases, but they are most commonly related to cancer (Figure S4). In total, 137 and 161 phosphoproteins were associated with cancer diseases in sEVs derived from MCF7 and MDA-MB-231 cell lines, respectively (Figure 3.2.4A). Among them, 40 and 50 sEV phosphoproteins derived from MCF7 and MDA-MB-231 are related to breast cancer, respectively (Figure 3.2.4B). The commonly associated breast cancer phosphoproteins unique to both MCF7 and MDA-MB-231 sEVs include 23 phosphoproteins (Figure 3.2.4C, Table S3). Furthermore, phosphosites unique for MCF7 and MDA-MB-231 sEVs were compared with previously identified phosphosites in plasma EVs in patients diagnosed with breast cancer and healthy controls [18]. We found in our study nine phosphosites that were common with previously reported studies (Table S4).

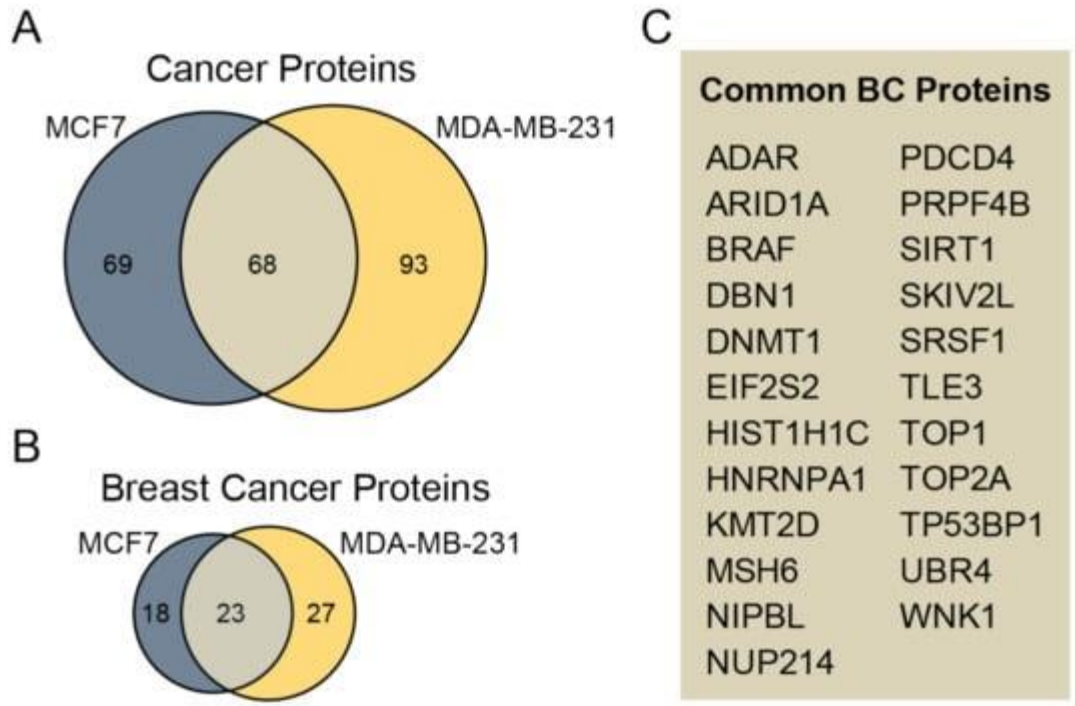

Figure 3.2.4. Assessment of identified phosphoproteins related to cancer. (A) The Venn diagram compares cancer-related phosphoproteins unique for MCF7 and MDA-MB-231 sEVs. (B) The Venn diagram compares the number of phosphoproteins relevant to BC unique for sEVs from cancerous cell lines. (C) A list of 23 common BC proteins from MCF7 and MDA-MB-231 sEVs is presented.

3.2.4.5. Functional and Pathway Analysis of Identified Phosphoproteins

We annotated and classified proteins using the KEGG pathway database. Pathways terms such as metabolic pathway, spliceosome, cell cycle, and viral carcinogenesis were highly abundant in MCF7- and MDA-MB-231-derived sEVs, compared with MCF10A-derived sEVs (Figure 3.2.5) (Table S5). Interestingly, 13 and 10 proteins identified in the MDA-MB-231 and MCF7 sEVs, respectively, were related to cell cycle. Among these cell cycle proteins, five proteins belonged to kinases: cyclin-dependent kinase 1 (CDK1) and 2 (CDK2), glycogen synthase kinase 3 beta (GSK3B), polo-like kinase 1 (PLK1), and protein kinase/DNA-activated, catalytic subunit (PRKDC) (Table S5). We previously studied the presence of three functional metabolic enzymes in BC-derived MVs that might be used as potential biomarkers in BC therapy [7]. This investigation indicated that sEV enzymes can effectively be incorporated into accurate, quick, and sensitive

early diagnostic assays that work by measuring their enzymatic activities. Therefore, we focused on sEV enzymes annotated by the term “metabolic pathways” by the KEGG database. After data analysis, six enzymes were identified only in the MCF7- or MDA-MB-231-derived sEVs at least once in three biological replicates. These enzymes are involved in the biosynthesis of cofactors (ACLY, ACSS2), glycolysis (PKFM), deacetylase activity (SIRT1 and SIRT6), and pyrimidine biosynthesis (CTPS1) (Table 3.2.1). Two enzymes, ACLY and PKFM, contained the highest number of identifications of phosphopeptides among the list of proteins presented in Table 3.2.1. The SIRT1 and SITR6 enzymes were identified in all three biological replicates of MCF7 and/or MDA-MB-231 sEVs with the IMAC phosphopeptide enrichment technique using either CID or HCD fragmentation methods. Therefore, four enzymes that have at least six identifications of phosphopeptides were further investigated: ACLY, PKFM, SIRT1, and SIRT6. Representative MS/MS spectra for these four enzymes are presented in Figure S5. Measuring the enzymatic assay for CTPS1 requires specific instrumentation, is time consuming, and is difficult to implement for routine use [34]. For this reason, no further validation of this enzyme was performed.

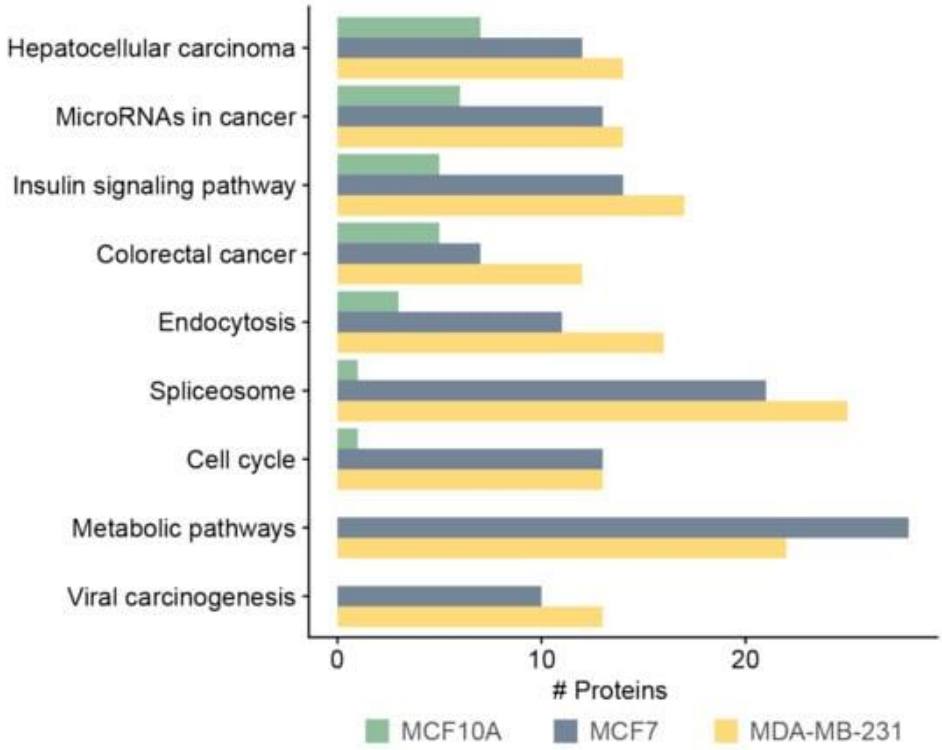

Figure 3.2.5. KEGG pathway analysis of phosphoproteins derived from MCF10A, MCF7, and MDA-MB-231 sEVs.

Table 3.2.1. Summary of the identifications of phosphosites for enzymes in three biological replicates of MCF7 and MDA-MB-231 sEVs. Six proteins were identified using IMAC and either CID or HCD fragmentation method in all three replicate samples in at least one cell line. In addition, PKFM has also been identified from TiO_2 enrichment in two replicates using HCD fragmentation in contrast with the rest.

Enzyme	Gene	Number of Phosphorpeptide Identifications	MCF7 sEV		MDA-MB-231 sEV	
			CID	HCD	CID	HCD
			Found in sample replicates (Number/3)			
ATP citrate lyase	ACLY	8	3	3	1	1
6-Phosphofructokinase	PFKM	8	1	0	3	3
Sirtuin 1	SIRT1	7	3	1	0	3
CTP synthetase 1	CTPS1	7	1	0	3	3
Sirtuin 6	SIRT6	6	0	3	0	3
Acetyl-CoA synthetase 2	ACSS2	5	2	3	0	0

3.2.4.6. Analysis of ACLY, PKFM, SIRT1, and SIRT6 in Cells and Their sEVs

ACLY, PKFM, SIRT1, and SIRT6 activities were detected in all extracted protein fractions from the three cell lines, as well as their sEVs (Figure 3.2.6, Table S6). The specific activity of ACLY was significantly higher in MDA-MB-231 when compared with MCF7- and MCF10A-derived sEVs. The specific activity of PFKM in sEV fractions, although slightly higher in MDA-MB-231 in comparison with the other two cell lines, was relatively similar in all sEV fractions. The specific activity of SIRT1 and SIRT6 in MDA-MB-231 and MCF7 was substantially higher in comparison with MCF10A in both cell-free extract (CFE) and sEV fractions, respectively (Figure 3.2.6). However, a more significant difference in specific activity between sEVs derived from non-cancerous and cancerous cell lines was observed for SIRT1.

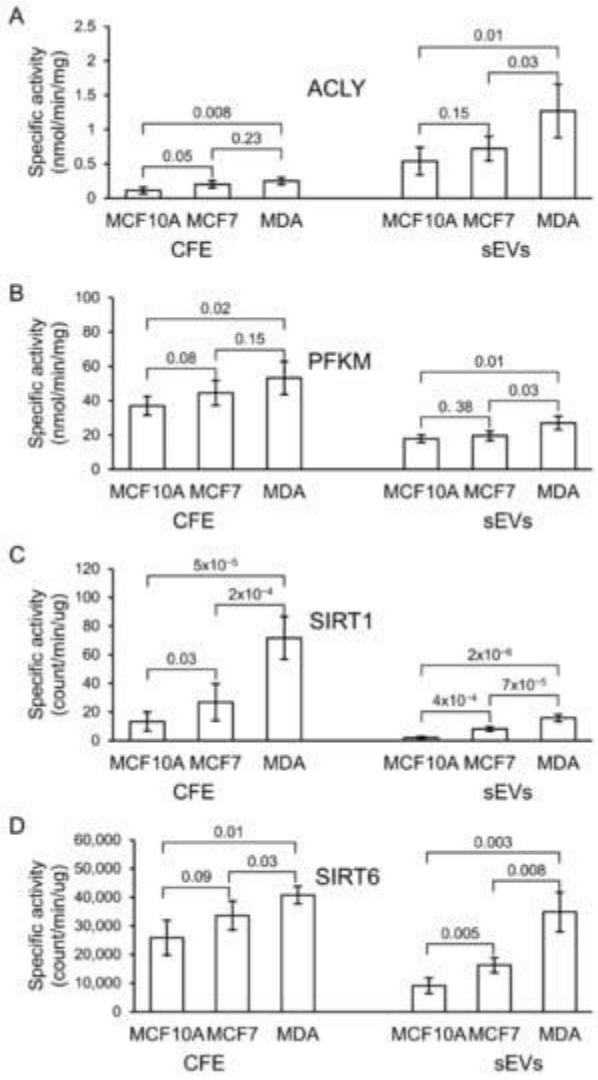

Figure 3.2.6. Specific enzymatic activity of (**A**) ATP citrate lyase (ACLY), (**B**) phosphofructokinase-M (PFKM), (**C**) sirtuin-1 (SIRT1), and (**D**) sirtuin-6 (SIRT6) in cell-free extract (CFE) and their corresponding sEV fractions. The bar graph represents mean values, while error bars indicate the standard deviation (SD) of four replicates and *p*-values obtained from a Student's *t*-test.

3.2.5 Discussion

To understand the molecular and cellular regulatory mechanisms important for BC tumor progression and metastasis, we performed a phosphoproteomic analysis of BC-derived sEVs. In our study, we used sEVs from two breast cancer cell lines, MCF7 and MDA-MB-231, and a non-tumorigenic breast cell line, MCF10A. Two enrichment techniques, IMAC and TiO_2, and

fragmentation methods, CID and HCD, were employed to maximize the purification and identification of phosphopeptides. We identified 2450 phosphorylation sites with 60 novel ones. These phosphorylation sites were mapped to 870 distinct proteins, covering a broad range of functions. Many identified phosphoproteins were unique to BC EVs encompassing a variety of signaling, metabolic, and regulatory pathways and cellular processes. The significant difference in the number of phosphoproteins identified from sEVs isolated from MCF10A and either MDA-MB-231 or MCF7 could be due to MCF10A producing a lower number of sEVs. This is supported by studies showing that MCF10A cells produce significantly fewer sEVs and less protein content than MDA-MB-231 [6,35]. Moreover, differences in the phosphoproteins identified were also observed between MCF7 and MDA-MB-231 EVs (Figure 3.2.2D). Furthermore, the intensity of EV markers observed by Western blot analysis varied for these cell lines, particularly for CD63 and CD81 markers (Figure 3.2.1B). The difference in phosphoproteome patterns and abundance of EV markers may be attributed to natural cell-to-cell variation in protein expression and sEV biogenesis. Distinct cell lines may release different ratios of EV subtypes that consequently change the overall number of identified proteins, as well as their expression.

In this study, among all groups of kinases, CMGC, CAMKs, STE, AGC, and TKL were predicted to regulate the largest number of phosphorylation events in sEVs (Figure 3.2.3C). Kinases previously identified in MDA-MB-231 EVs [6] mainly belong to the STE and TK kinase families including subfamily members such as p21-activated kinase 4 (PAK4), STE20-like serine/threonine-protein kinase (SLK), tyrosine-protein kinase Fyn (FYN), and proto-oncogene tyrosine-protein kinase Src (SRC). The PAK4 protein kinase is often highly expressed in TNBC cells and plays an important role in cell growth, survival, and migration [36]. The Ste20-like kinase, SLK, is involved in the control of BC cell motility [37]. It has been reported that FYN, a member of the SRC family kinases, is required for the maintenance of the basal breast cancer subtype [38]. For instance, c-Src proto-oncogene tyrosine kinase has been shown to support cancer cell migration and proliferation [39]. The presence of these kinases in EVs may be important for the regulation of protein phosphorylation in recipient cells, and consequently, the promotion of cancer progression.

In total, 266 phosphoproteins were associated with cancer in sEVs derived from MCF7 and MDA-MB-231 cell lines. Among them, 78 are related to breast cancer, suggesting that the cargo of phosphoproteins in sEVs may play an important role in cancer progression and metastasis. Moreover, these identified phosphoproteins could serve as biomarkers for BC diagnosis. Therefore,

this study identified nine proteins (Table S4) previously found in plasma EVs of patients diagnosed with BC and healthy controls [18]. These nine phosphoproteins may be further explored in BC biomarker development and clinical use.

Advancements in proteomics facilitate discovery of enzymes that can provide an attractive source for cancer biomarker tests. Due to their unique enzyme specificity and selectivity, enzymatic assays are a reliable, simple, and rapid diagnostic method [40]. A phosphoproteomic approach is particularly important because the phosphorylation of enzymes may affect enzymatic activities [14,41-43]. Therefore, for the selection of potential BC biomarkers for detection and prognosis as well as for pharmacological purposes, the focus in this work was on the identification of phosphorylated sEV enzymes identified only in BC cell lines. This led to the selection of phosphorylated enzymes present only in cancerous cell lines. These enzymes include ACLY, PKFM, SIRT1, and SIRT6.

ACLY is a cytoplasmic homologous tetramer composed of four polypeptide chains and acts as a metabolic enzyme involved in fatty acid synthesis in rapidly proliferating cancer cells [44-46]. ACLY plays a role in modulating proliferation, growth, migration, and apoptosis that has been reported in many cancer cells [44-46]. The expression of ACLY in human lung adenocarcinoma has been investigated to be higher compared with non-carcinogenic lung control tissue [47] and contributes to increased lipogenesis and tumor growth [48]. It has been shown that miR-22 inhibits the growth and metastasis of MCF7 cells by decreasing ACLY expression [49]. BMS-303141, an ACLY inhibitor, has been suggested for HCC treatment [50]. It could induce endoplasmatic reticulum stress and activate the p-eIF2α/ATF4/CHOP axis, promoting HCC cell apoptosis. Recent studies highlight ACLY as a potential biomarker for predicting breast cancer recurrence in patients [51]. In this study, we found that ACLY enzymatic activity was increased in breast cancer-derived sEVs isolated from the MDA-MB-231 cell line, in comparison with sEVs isolated from MCF7 and MCF10A cell lines. Further studies are needed to investigate the potential of this protein as a clinical indicator for prognosis in breast cancer.

Expression levels of phosphofructokinase-M (PFKM) are closely related to the occurrence and development of malignant tumors. To meet the metabolic demands of tumor cells in energy, the activity of PFKM in cancer cells is increased [52-56]. The inhibition of 2,6-2-fructose production decreases PFKM activity, which results in the inhibition of the growth of tumor cells [57]. The gene-based analysis of early-onset BC has identified a region containing the key glycolysis regulation

gene, PFKM, which is proposed as a potential target for BC prevention and treatment [58]. Our study indicates that the specific activity of PFKM in sEV fractions of MDA-MB-231 was slightly higher than its activity in sEVs from MCF7 and non-cancerous MCF10A cell lines, suggesting that the phosphorylation of PFKM in MDA-MB-231 sEVs does not significantly affect the enzyme's activity. Therefore, this enzyme from sEVs might not effectively serve as a prognostic biomarker for BC.

Sirtuin-1 (SIRT1) is a class-III histone deacetylase (HDAC) enzyme involved in gene regulation, genome stability, apoptosis, autophagy, senescence, proliferation, aging, and tumorigenesis [59]. This enzyme deacetylates histones and non-histone proteins important in cancer biology such as p53, p73, Rb, and NF-κB [60,61]. SIRT1 is proposed as a prognostic indicator as well as a novel therapeutic candidate in triple-negative breast cancer (TNBC) [62]. Studies looking at the expression of SIRT1 in BC indicate its contradictory roles as a tumor suppressor or promoter [63-65]. Specific activity of SIRT1 from whole cell and sEV protein fractions is higher in MCF7 and MDA-MB-231 than its corresponding MCF10A fractions. This suggests that SIRT1 could be useful to investigate as a potential prognostic and therapeutic target for BC.

Sirtuin-6 (SIRT6) is involved in multiple molecular pathways related to DNA repair, glycolysis, gluconeogenesis, tumorigenesis, neurodegeneration, and cardiac hypertrophic responses [66]. SIRT6 may be linked to cancer progression and tumor growth. It was identified as a tumor suppressor that regulates aerobic glycolysis in cancer cells [67,68]. In vivo studies revealed the critical role of high SIRT6 levels in slowing down hepatic cancer at an early stage of its progression [66]. 4H-Chromen, an activator of SIRT6, has been studied in various breast cancer cells and demonstrated to decrease cell proliferation in TNBC cells [69]. In this work, like SIRT1, cancerous MCF7 and MDA-MB-231 cell line-derived sEVs exhibited higher enzymatic activity of SIRT6 than sEVs of the non-cancerous MCF10A cell line. This finding suggests that this enzyme may be useful to investigate as a potential biomarker for BC diagnosis.

Previous studies have investigated the phosphoproteome of BC cell lines, as well as their EVs [70,71]. However, our methodological approach for studying phosphoproteins of BC cell lines was different from these reports [71,72]. Phosphoproteomic analysis of EVs derived from several BC cell lines including MCF7, MDA-MB-231, and MCF10A was performed using only IMAC enrichment and HCD fragmentation methods; ACLY and SIRT1 were identified, but not PFKM and SIRT6. In another report that studied the phosphoproteomic characterization of KAIMRC1, MCF-7, and

MDA-MB-231 BC cell lines, TiO_2 enrichment and CID fragmentation methods were employed for peptide enrichment and fragmentation; the study was based on cell culture lysates and did not identify the four enzymes presented in our work.

The objective of this study, along with previously published works [6,7], was to investigate proteins or enzymes that may be later explored for their efficacy as biomarkers for BC early detection. To achieve this objective, blood sEVs from healthy and BC patients should be studied for the presence or activity of these potential biomarkers. This study was limited by the isolation method of sEVs since extracellular vesicle subtypes are diverse in their diameter and density. Therefore, ultracentrifugation does not differentiate sEVs that originate from different biogenesis pathways. Due to the heterogeneity of EV populations that come from different intracellular origins, our study did not examine a specific EV subtype, but instead considered a diverse population of sEVs. It has been reported that the EV isolation method can significantly impact EV yield and purity from human serum [72]. For this reason, it will be most challenging to determine the most suitable method for the isolation of EVs from blood in the future. In this study, we found that phosphorylated enzymes identified from sEV fractions were already proposed to play a role in cancer therapy. Our findings support further efforts to investigate these enzymes, especially ACLY, SIRT1, and SIRT6 for breast cancer diagnosis and therapy.

3.2.6. Conclusions

In this study, we conducted a phosphoproteomic analysis of breast cancer-derived extracellular vesicles from MCF10A, MCF7, and MDA-MB-231 cell lines. In total, 855 distinct phosphoproteins were identified collectively among the cell lines, covering a wide range of functions, most of which are related to cancer. Among these phosphoproteins, we validated four enzymes: ACLY, PFKM, SIRT1, SIRT6. The results demonstrate that the specific activity of PFKM in BC cancer cell lines was not statistically different from the non-cancerous cell line. In contrast, the three phosphorylated enzymes ACLY, SIRT1, and SIRT6 showed a significantly higher specific enzymatic activity in MDA-MB-231 in comparison to MCF10A-derived sEVs. These three enzymes might serve as a potential prognostic biomarker for BC. They have been previously proposed as therapeutic targets for cancer therapy. Thus, our findings justify further investigation of these enzymes as promising drug targets for BC treatment.

3.2.7 Supplementary Materials

The following supporting information can be downloaded at: **https://www.mdpi.com/article/10.3390/biomedicines10020408/s1**, Figure S1: Full-length images of Western blot gels, Figure S2: Venn diagrams presenting the overlap of phosphopeptides identified from either IMAC or TiO2 enrichment method using two fragmentation methods, CID and HCD, Figure S3: Phosphorylation motifs of EVs from MCF10A, MCF7, and MDA-MB-231 cells for pS, pY, and pY, Figure S4: Fractions of disease annotations for selected disease categories using phosphoproteins unique for MCF7 and MDA-MB-231 sEVs, Figure S5: Representative MS/MS spectra of phosphopeptides from ACLY, PFKM, SIRT1, and SIRT6, Table S1: Total number of identified phosphopeptides for each sEV sample, using two enrichment and fragmentation methods, Table S2: The probability of correct site analysis for phosphoserine, phosphothreonine, and phosphotyrosine sites identified from three cell line-derived sEVs, Table S3: The list of identified phosphoproteins indicating breast cancer phosphoproteins common to MDA-MB-231 and MCF7 sEVs and number of identified phosphosites per protein per cell line, Table S4: Phosphosites found in MCF7 and/or MDA-MB-231 EVs common to previously identified phosphosites in plasma EV of healthy patients and diagnosed with breast cancer [18], Table S5: KEGG pathway functional annotation terms for sEV MCF10A of MCF7 and MDA-MB-231, Table S6: Activities of ACLY, PFKM, SIRT1, and SIRT6 in the different protein extracts from cells and their derived sEV fractions. The numbers in parentheses indicate the number of biological replicates.

3.2.8 Funding

The study was supported by the John L. Holmes Mass Spectrometry Facility for the sample preparation and MS analysis and funded by Natural Sciences and Engineering Research Council of Canada (grant # RGPIN-2020-05775 for M.V.B.).

3.2.9 Data Availability Statement

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD030424.

3.2.10 Acknowledgments

We would like to acknowledge Ryan Reshke at Gibbings Laboratory for excellent technical assistance in nanoparticle tracking analysis and Dylan Burger at the Kidney Research Centre, the Ottawa Research Institute, for the NTA technical training.

3.2.11 References

1. Kruger, S.; Abd Elmageed, Z.Y.; Hawke, D.H.; Wörner, P.M.; Jansen, D.A.; Abdel-Mageed, A.B.; Alt, E.U.; Izadpanah, R. Molecular Characterization of Exosome-Like Vesicles from Breast Cancer Cells. *BMC Cancer* **2014**, *14*, 44.
2. Palazzolo, G.; Albanese, N.N.; DI Cara, G.; Gygax, D.; Vittorelli, M.L.; Pucci-Minafra, I. Proteomic Analysis of Exosome-Like Vesicles Derived from Breast Cancer Cells. *Anticancer Res.* **2012**, *32*, 847–860.
3. Hurwitz, S.N.; Rider, M.A.; Bundy, J.L.; Liu, X.; Singh, R.K.; Meckes, D.G., Jr. Proteomic Profiling of NCI-60 Extracellular Vesicles Uncovers Common Protein Cargo and Cancer Type-Specific Biomarkers. *Oncotarget* **2016**, *7*, 86999–87015.
4. Demory Beckler, M.; Higginbotham, J.N.; Franklin, J.L.; Ham, A.J.; Halvey, P.J.; Imasuen, I.E.; Whitwell, C.; Li, M.; Liebler, D.C.; Coffey, R.J. Proteomic Analysis of Exosomes from Mutant KRAS Colon Cancer Cells Identifies Intercellular Transfer of Mutant KRAS. *Mol. Cell Proteom.* **2013**, *12*, 343–355.
5. Liang, B.; Peng, P.; Chen, S.; Li, L.; Zhang, M.; Cao, D.; Yang, J.; Li, H.; Gui, T.; Li, X.; et al. Characterization and Proteomic Analysis of Ovarian Cancer-Derived Exosomes. *J. Proteom.* **2013**, *80*, 171–182.
6. Risha, Y.; Minic, Z.; Ghobadloo, S.M.; Berezovski, M.V. The Proteomic Analysis of Breast Cell Line Exosomes Reveals Disease Patterns and Potential Biomarkers. *Sci. Rep.* **2020**, *10*, 13572.
7. Risha, Y.; Susevski, V.; Hüttmann, N.; Poolsup, S.; Minic, Z.; Berezovski, M.V. Breast Cancer-Derived Microvesicles Are the Source of Functional Metabolic Enzymes as Potential Targets for Cancer Therapy. *Biomedicines.* **2021**, *9*, 107.
8. Huang, X.; Yuan, T.; Tschannen, M.; Sun, Z.; Jacob, H.; Du, M.; Liang, M.; Dittmar, R.L.; Liu, Y.; Liang, M.; et al. Characterization of Human Plasma-Derived Exosomal RNAs by Deep Sequencing. *BMC Genom.* **2013**, *14*, 319.
9. Carvalho, A.S.; Henrique Baeta, H.; Moraes, M.C.S.; Beck, H.C.; Rodriguez, M.S.; Saraswat, M.; Pandey, A.; Matthiesen, R. Extra-Cellular Vesicles Carry Proteome of Cancer Hallmarks. *Front. Biosci.* **2020**, *25*, 398–436.
10. Brzozowski, J.S.; Jankowski, H.; Bond, D.R.; Mccague, S.B.; Munro, B.R.; Predebon, M.J.; Scarlett, C.J.; Skelding, K.A.; Weidenhofer, J. Lipidomic Profiling of Extracellular Vesicles Derived from Prostate and Prostate Cancer Cell Lines. *Lipids Health Dis.* **2018**, *17*, 211.
11. Xu, R.; Greening, D.W.; Zhu, H.-J.; Takahashi, N.; Simpson, R.J. Extracellular Vesicle Isolation and Characterization: Toward Clinical Application. *J. Clin. Investig.* **2016**, *126*, 1152–1162.
12. Minciacchi, V.R.; Freeman, M.R.; Vizio, D.D. Extracellular Vesicles in Cancer: Exosomes, Microvesicles and the Emerging Role of Large Oncosomes. Semin. *Cell Dev. Biol.* **2015**, *40*, 41–51.

13. Vizio, D.D.; Morello, M.; Dudley, A.C.; Schow, P.W.; Adam, R.M.; Morley, S.; Mulholland, D.; Rotinen, M.; Hager, M.H.; Insabato, L.; et al. Large Oncosomes in Human Prostate Cancer Tissues and in the Circulation of Mice with Metastatic Disease. *Am. J. Pathol.* **2012**, *181*, 1573–1584.
14. Minic, Z.; Dahms, T.E.S.; Babu, M. Chromatographic Separation Strategies for Precision Mass Spectrometry to Study Protein-Protein Interactions and Protein Phosphorylation. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2018**, *1102-1103*, 96–108.
15. Yu, S.; Cao, H.; Shen, B.; Feng, J. Tumor-Derived Exosomes in Cancer Progression and Treatment Failure. *Oncotarget* **2015**, *6*, 37151–37168.
16. Guo, J.; Cui, Y.; Yan, Z.; Luo, Y.; Zhang, W.; Deng, S.; Tang, S.; Zhang, G.; He, Q.Y.; Wang, T. Phosphoproteome Characterization of Human Colorectal Cancer SW620 Cell-Derived Exosomes and New Phosphosite Discovery for C-HPP. *J. Proteome Res.* **2016**, *15*, 4060–4072.
17. Song, X.; Ding, Y.; Liu, G.; Yang, X.; Zhao, R.; Zhang, Y.; Zhao, X.; Anderson, G.J.; Nie, G. Cancer Cell-derived Exosomes Induce Mitogen-activated Protein Kinase-dependent Monocyte Survival by Transport of Functional Receptor Tyrosine Kinases. *J. Biol. Chem.* **2016**, *291*, 8453–8464.
18. Chen, I.H.; Xue, L.; Hsu, C.C.; Paez, J.S.; Pan, L.; Andaluz, H.; Wendt, M.K.; Iliuk, A.B.; Zhu, J.K.; Tao, W.A. Phosphoproteins in Extracellular Vesicles as Candidate Markers for Breast Cancer. *Proc Natl Acad Sci U S A.* **2017**, *114*, 3175–3180.
19. Iliuk, A.; Wu, X.; Li, L.; Sun, J.; Hadisurya, M.; Boris, R.S.; Tao, W.A. Plasma-Derived Extracellular Vesicle Phosphoproteomics through Chemical Affinity Purification. *J. Proteome Res.* **2020**, *19*, 2563–2574.
20. Wu, X.; Li, L.; Iliuk, A.; Tao, W.A. Highly Efficient Phosphoproteome Capture and Analysis from Urinary Extracellular Vesicles. *J. Proteome Res.* **2018**, *17*, 3308–3316.
21. Goodspeed, A.; Heiser, L.M.; Gray, J.W.; Costello, J.C. Tumor-Derived Cell Lines as Molecular Models of Cancer Pharmacogenomics. *Mol. Cancer Res.* **2016**, *14*, 3–13.
22. Cox, J.; Mann, M. MaxQuant Enables High Peptide Identification Rates, Individualized p.p.b.-Range Mass Accuracies and Proteome-Wide Protein Quantification. *Nat. Biotechnol.* **2008**, *26*, 1367–1372.
23. Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R.A.; Olsen, J.V.; Mann, M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J. Proteome Res.* **2011**, *10*, 1794–1805.
24. R. Core Team. R: A Language and Environment for Statistical Computing. R. Core Team: Vienna, Austria, 2021.
25. Olsen, J.V.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M. Global, in Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* **2006**, *127*, 635–648.
26. Hornbeck, P.V.; Zhang, B.; Murray, B.; Kornhauser, J.M.; Latham, V.; Skrzypek, E. PhosphoSitePlus, 2014: Mutations, PTMs and Recalibrations. *Nucleic Acids Res.* **2015**, *43*, D512-520.
27. Wang, C.; Xu, H.; Lin, S.; Deng, W.; Zhou, J.; Zhang, Y.; Shi, Y.; Peng, D.; Xue, Y. GPS 5.0: An Update on the Prediction of Kinase-Specific Phosphorylation Sites in Proteins. *Genomics Proteomics Bioinformatics* **2020**, *18*, 72–80.

28. Piñero, J.; Ramírez-Anguita, J.M.; Saüch-Pitarch, J.; Ronzano, F.; Centeno, E.; Sanz, F.; Furlong, L.I. The DisGeNET Knowledge Platform for Disease Genomics: 2019 Update. *Nucleic Acids Res.* **2020**, *48 (D1)*, D845–D855.
29. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene Ontology: Tool for the Unification of Biology. *Nat. Genet.* **2000**, *25*, 25–29.
30. The Gene Ontology Consortium. The Gene Ontology Resource: Enriching a GOld Mine. *Nucleic Acids Res.* **2021**, *49 (D1)*, D325–D334.
31. Carlson, M. *Org.Hs.Eg.Db: Genome Wide Annotation for Human*, R Package Version 3.8.2. 2019.
32. Adams, J.A. Kinetic and Catalytic Mechanisms of Protein Kinases. *Chem. Rev.* **2001**, *101*, 2271–2290.
33. Sugiyama, N.; Imamura, H.; Ishihama, Y. Large-scale Discovery of Substrates of the Human Kinome. *Sci Rep.* **2019**, *9*, 10503.
34. Boschat, A.C.; Minet, N.; Martin, E.; Barouki, R.; Latour, S.; Sanquer, S. CTP Synthetase Activity Assay by Liquid Chromatography Tandem Mass Spectrometry in the Multiple Reaction Monitoring Mode. *J. Mass Spectrum.* **2019**, *54*, 885–893.
35. Im, E.J.; Lee, C.H.; Moon, P.G.; Rangaswamy, G.G.; Lee, B.; Lee, J.M.; Lee, J.C.; Jee, J.G.; Bae, J.S.; Kwon, T.K.; et al. Sulfisoxazole Inhibits the Secretion of Small Extracellular Vesicles by Targeting the Endothelin Receptor A. *Nat. Commun.* **2019**, *10*, 1387.
36. Cordover, E.; Wei, J.; Patel, C.; Shan, N.L.; Gionco, J.; Sargsyan, D.; Wu, R.; Cai, L.; Kong, A.N.; Jacinto, E.; et al. KPT-9274, an Inhibitor of PAK4 and NAMPT, Leads to Downregulation of mTORC2 in Triple Negative Breast Cancer Cells. *Chem. Res. Toxicol.* **2020**, *33*, 482–491.
37. Roovers, K.; Wagner, S.; Storbeck, C.J.; O'Reilly, P.; Lo, V.; Northey, J.J.; Chmielecki, J.; Muller, W.J.; Siegel, P.M.; Sabourin, L.A. The Ste20-like Kinase SLK is Required for ErbB2-Driven Breast Cancer Cell Motility. *Oncogene* **2009**, *28*, 2839–4288.
38. Lee, G.H.; Yoo, K.C.; An, Y.; Lee, H.J.; Lee, M.; Uddin, N.; Kim, M.J.; Kim, I.G.; Suh, Y.; Lee, S.J. FYN Promotes Mesenchymal Phenotypes of Basal Type Breast Cancer Cells Through STAT5/NOTCH2 Signaling Node. *Oncogene* **2018**, *37*, 1857–1868.
39. Kim, L.C.; Song, L.; Haura, E.B. Src Kinases as Therapeutic Targets for Cancer. *Nat. Rev. Clin. Oncol.* **2009**, *6*, 587–595.
40. Singh, R.S.; Singh, T.; Singh, A.K. Enzymes as Diagnostic Tools. *Adv. Enzym. Technol* **2019**, 225–271.
41. Catalano, S.; Barone, I.; Marsico, S.; Bruno, R.; Andò, S. Phosphorylation Processes Controlling Aromatase Activity in Br east Cancer: An Update. *Mini Rev. Med. Chem.* **2016**, *16*, 691–698.
42. Seshacharyulu, P.; Pandey, P.; Datta, K.; Batra, S.K. Phosphatase: PP2A Structural Importance, Regulation and Its Aberrant Expression in Cancer. *Cancer Lett.* **2013**, *335*, 9–18.
43. Longati, P.; Comoglio, P.M.; Bardelli, A. Receptor Tyrosine Kinases as Therapeutic Targets: The Model of the MET Oncogene. *Curr. Drug Targets* **2001**, *2*, 41–55.
44. Singh, M.; Richards, E.G.; Mukherjee, A.; Srere, P.A. Structure of ATP Citrate Lyase from Rat Liver. Physicochemical Studies and Proteolytic Modification. *J. Biol. Chem.* **1976**, *251*, 5242–5250.

45. Granchi, C. ATP Citrate Lyase (ACLY) Inhibitors: An Anti-Cancer Strategy at the Crossroads of Glucose and Lipid Metabolism. Eur. *J. Med. Chem.* **2018**, *157*, 1276–1291.
46. Chypre, M.; Zaidi, N.; Smans, K. ATP-Citrate Lyase: A Mini-Review. *Biochem. Biophys. Res. Commun.* **2012**, *422*, 1–4.
47. Migita, T.; Narita, T.; Nomura, K.; Miyagi, E.; Inazuka, F.; Matsuura, M.; Ushijima, M.; Mashima, T.; Seimiya, H.; Satoh, Y.; et al. ATP Citrate Lyase: Activation and Therapeutic Implications in Non-Small Cell Lung Cancer. *Cancer Res.* **2008**, *68*, 8547–8554.
48. Lin, R.; Tao, R.; Gao, X.; Li, T.; Zhou, X.; Guan, K.L.; Xiong, Y.; Lei, Q.Y. Acetylation Stabilizes ATP-Citrate Lyase to Promote Lipid Biosynthesis and Tumor Growth. *Mol. Cell.* **2013**, *51*, 506–518.
49. Liu, H.; Huang, X.; Ye, T. MiR-22 Down-Regulates the Proto-Oncogene ATP Citrate Lyase to Inhibit the Growth and Metastasis of Breast Cancer. *Am. J. Transl. Res.* **2018**, *10*, 659–669.
50. Zheng, Y.; Zhou, Q.; Zhao, C.; Li, J.; Yu, Z.; Zhu, Q. ATP Citrate Lyase Inhibitor Triggers Endoplasmic Reticulum Stress to Induce Hepatocellular Carcinoma Cell Apoptosis via p-eIF2α/ATF4/CHOP Axis. *J. Cell Mol. Med.* **2021**, *25*, 1468–1479.
51. Chen, Y.; Li, K.; Gong, D.; Zhang, J.; Li, Q.; Zhao, G.; Lin, P. ACLY: A Biomarker of Recurrence in Breast Cancer. *Pathol. Res. Pract.* **2020**, *216*, 153076.
52. Musumeci, O.; Bruno, C.; Mongini, T.; Rodolico, C.; Aguennouz, M.; Barca, E.; Amati, A.; Cassandrini, D.; Serlenga, L.; Vita, G.; et al. Clinical Features and New Molecular Findings in Muscle Phosphofructokinase Deficiency (GSD Type VII). *Neuromuscul. Disord.* **2012**, *22*, 325–330.
53. Mor, I.; Cheung, E.C.; Vousden, K.H. Control of Glycolysis Through Regulation of PFK1: Old Friends and Recent Additions. Cold Spring Harb Symp. *Quant. Biol.* **2011**, *76*, 211–216.
54. Webb, B.A.; Forouhar, F.; Szu, F.E.; Seetharaman, J.; Tong, L.; Barber, D.L. Structures of Human Phosphofructokinase-1 and Atomic Basis of Cancer-Associated Mutations. *Nature* **2015**, *523*, 111–114.
55. Ismail, R.; Ul Hussain, M. The Up Regulation of Phosphofructokinase1 (PFK1) Protein During Chemically Induced Hypoxia is Mediated by the Hypoxia-Responsive Internal Ribosome Entry Site (IRES) Element, Present in Its 5'Untranslated Region. *Biochimie* **2017**, *139*, 38–45.
56. Lee, J.H.; Liu, R.; Li, J.; Zhang, C.; Wang, Y.; Cai, Q.; Qian, X.; Xia, Y.; Zheng, Y.; Piao, Y.; et al. Stabilization of Phosphofructokinase 1 Platelet Isoform by AKT Promotes Tumorigenesis. *Nat. Commun.* **2017**, *8*, 949.
57. Clem, B.; Telang, S.; Clem, A.; Yalcin, A.; Meier, J.; Simmons, A.; Rasku, M.A.; Arumugam, S.; Dean, W.L.; Eaton, J.; et al. Small-Molecule Inhibition of 6-Phosphofructo-2-Kinase Activity Suppresses Glycolytic Flux and Tumor Growth. *Mol. Cancer Ther.* **2008**, *7*, 110–120.
58. Ahsan, H.; Halpern, J.; Kibriya, M.G.; Pierce, B.L.; Tong, L.; Gamazon, E.; McGuire, V.; Felberg, A.; Shi, J.; Jasmine, F.; et al. A Genome-Wide Association Study of Early-Onset Breast Cancer Identifies PFKM as a Novel Breast Cancer Gene and Supports a Common Genetic Spectrum for Breast Cancer at Any Age. Cancer Epidemiol. *Biomarkers Prev.* **2014**, *23*, 658–669.
59. Alves-Fernandes, D.K.; Jasiulionis, M.G. The Role of SIRT1 on DNA Damage Response and Epigenetic Alterations in Cancer. *Int. J. Mol. Sci.* **2019**, *20*, 3153.

60. Yeung, F.; Hoberg, J.E.; Ramsey, C.S.; Keller, M.D.; Jones, D.R.; Frye, R.A.; Mayo, M.W. Modulation of NF-kappaB-Dependent Transcription and Cell Survival by the SIRT1 Deacetylase. *EMBO J.* **2004**, *23*, 2369–2380.
61. Yi, Y.W.; Kang, H.J.; Kim, H.J.; Kong, Y.; Brown, M.L.; Bae, I. Targeting Mutant p53 by a SIRT1 Activator YK-3-237 Inhibits the Proliferation of Triple-Negative Breast Cancer Cells. *Oncotarget* **2013**, *4*, 984–994.
62. Jin, M.S.; Hyun, C.L.; Park, I.A.; Kim, J.Y.; Chung, Y.R.; Im, S.A.; Lee, K.H.; Moon, H.G.; Ryu, H.S. SIRT1 Induces Tumor Invasion by Targeting Epithelial Mesenchymal Transition-Related Pathway and is a Prognostic Marker in Triple Negative Breast Cancer. *Tumour Biol.* **2016**, *37*, 4743–4753.
63. Wang, C.; Yang, W.; Dong, F.; Guo, Y.; Tan, J.; Ruan, S.; Huang, T. The Prognostic Role of Sirt1 Expression in Solid Malignancies: A Meta-Analysis. *Oncotarget* **2017**, *8*, 66343–66351.
64. Cao, Y.-W.; Li, W.-Q.; Wan, G.-X.; Li, Y.-X.; Du, X.-M.; Li, Y.-C.; Feng, L. Correlation and Prognostic Value of SIRT1 and Notch1 Signaling in Breast Cancer. *J. Exp. Clin. Cancer Res CR* **2014**, *33*, 97.
65. Chung, Y.R.; Kim, H.; Park, S.Y.; Park, I.A.; Jang, J.J.; Choe, J.-Y.; Jung, Y.Y.; Im, S.A.; Moon, H.G.; Lee, K.H.; et al. Distinctive role of SIRT1 expression on tumor invasion and metastasis in breast cancer by molecular subtype. *Hum. Pathol.* **2015**, *46*, 1027–1035.
66. Khan, R.I.; Nirzhor, S.S.R.; Akter, R. A Review of the Recent Advances Made with SIRT6 and its Implications on Aging Related Processes, Major Human Diseases, and Possible Therapeutic Targets. *Biomolecules* **2018**, *8*, 44.
67. Zhong, L.; D'Urso, A.; Toiber, D.; Sebastian, C.; Henry, R.E.; Vadysirisack, D.D.; Guimaraes, A.; Marinelli, B.; Wikstrom, J.D.; Nir, T.; et al. The Histone Deacetylase Sirt6 Regulates Glucose Homeostasis via Hif1α. *Cell* **2010**, *140*, 280–293.
68. Sebastián, C.; Zwaans, B.M.M.; Silberman, D.M.; Gymrek, M.; Goren, A.; Zhong, L.; Ram, O.; Truelove, J.; Guimaraes, A.R.; Toiber, D.; et al. The Histone Deacetylase SIRT6 Is a Tumor Suppressor that Controls Cancer Metabolism. *Cell* **2012**, *151*, 1185–1199.
69. Tenhunen, J.; Kučera, T.; Huovinen, M.; Küblbeck, J.; Bisenieks, E.; Vigante, B.; Ogle, Z.; Duburs, G.; Doležal, M.; Moaddel, R.; et al. Screening of SIRT6 Inhibitors and Activators: A Novel Activator has an Impact on Breast Cancer Cells. *Biomed. Pharmacother.* **2021**, *138*, 111452.
70. Rontogianni, S.; Synadaki, E.; Li, B.; Liefaard, M.C.; Lips, E.H.; Wesseling, J.; Wu, W.; Atelaar, M. Proteomic profiling of extracellular vesicles allows for human breast cancer subtyping. *Commun Biol.* **2019**, *2*, 325.
71. Alghanem, B.; Ali, R.; Nehdi, A.; Al Zahrani, H.; Altolayyan, A.; Shaibah, H.; Baz, O.; Alhallaj, A.; Moresco, J. Proteomics profiling of KAIMRC1 in comparison to MDA-MB231 and MCF-7. *Int. J. Mol. Sci.* **2020**, *21*, 4328.
72. Brennan, K.; Martin, K.; FitzGerald, S.P.; O'Sullivan, J.; Wu, Y.; Blanco, A.; Richardson, C.; Mc Gee, M.M. A Comparison of Methods for the Isolation and Separation of Extracellular Vesicles from Protein and Lipid Particles in Human Serum. *Sci. Rep.* **2020**, *10*, 1039.

3.3 Lysine Acetylome of Breast Cancer-Derived Small Extracellular Vesicles Reveals Specific Acetylation Patterns for Metabolic Enzymes

3.3.1 Abstract

Cancer-derived small extracellular vesicles have been proposed as promising potential biomarkers for diagnosis and prognosis of breast cancer (BC). We performed a proteomic study of lysine acetylation of breast cancer-derived small extracellular vesicles (sEVs) to understand the potential role of the aberrant acetylated proteins in the biology of invasive ductal carcinoma and triple-negative BC. Three cell lines were used as models for this study: MCF10A (non-metastatic), MCF7 (estrogen and progesterone receptor-positive, metastatic) and MDA-MB-231 (triple-negative, highly metastatic). For a comprehensive protein acetylation analysis of the sEVs derived from each cell line, acetylated peptides were enriched using the anti-acetyl-lysine antibody, followed by LC-MS/MS analysis. In total, there were 118 lysine-acetylated peptides, of which 22, 58 and 82 have been identified in MCF10A, MCF7 and MDA-MB-231 cell lines, respectively. These acetylated peptides were mapped to 60 distinct proteins and mainly identified proteins involved in metabolic pathways. Among the acetylated proteins identified in cancer-derived sEVs from MCF7 and MDA-MB-231 cell lines are proteins associated with the glycolysis pathway, annexins and histones. Five acetylated enzymes from the glycolytic pathway, present only in cancer-derived sEVs, were validated. These include aldolase (ALDOA), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate kinase (PGK1), enolase (ENO) and pyruvate kinase M1/2 (PKM). For three of these enzymes (ALDOA, PGK1 and ENO) the specific enzymatic activity was significantly higher in MDA-MB-231 when compared with MCF10A-derived sEVs. This study reveals that sEVs contain acetylated glycolytic metabolic enzymes that could be interesting potential candidates for early BC diagnostics.

Keywords: breast cancer, protein acetylation, small extracellular vesicles, aldolase (ALDOA), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate kinase (PGK 1), enolase (ENO), pyruvate kinase M1/2 (PKM)

3.3.2 Introduction

Extracellular vesicles (EVs), derived from human cancer cells, play a role in driving various cellular processes related to cancer biology, such as the modulation of tumor microenvironment,

angiogenesis, sustained growth and tissue invasion and metastasis [1-8]. EVs carry a variety of molecules (nucleic acids, proteins and metabolites) that can be released from the donor to nearby or distant cells [8-11]. Consequently, EV exchange between cells appears as a crucial modulator of cell–cell communication and could be an important player in health and disease states. Membrane-bound EVs are classified, based on their distinct biogenesis pathways, into three main subtypes: exosomes, microvesicles and apoptotic bodies. Additionally, EVs can also be classified based on their substantially different sizes. Small extracellular vesicles (sEVs) are in the range of 30 to 150 nm [12,13], while medium extracellular vesicles (mEVs) are in the range of 100 nm to 1000 nm in diameter [14,15]. In our study, we isolated vesicles in the range of 50–300 nm in diameter, referred to as sEVs.

EVs are of great interest for molecular biomarker discovery that enable monitoring cancer progression as well as early diagnosis and accurate prognosis of BC. In our previous investigations, using proteomics approaches, numerous proteins have already been identified from EVs as potential biomarkers for early diagnosis or as therapeutic targets [7,8]. Moreover, using a phosphoproteomics approach, we previously assessed the enzymatic activity of three enzymes associated with cancer and only present in sEVs derived from the cancerous cell lines MCF7 and MDA-MB-231: ATP citrate lyase (ACLY), sirtuin-1 (SIRT1) and sirtuin-6 (SIRT6) [16]. The specific activity of these enzymes was significantly higher in MDA-MB-231 sEV fractions when compared with similar MCF10A fractions. Our previous work demonstrated the presence of functional metabolic enzymes in sEVs specific to BC cells that may be further explored for early BC diagnostics. Furthermore, this investigation also indicates that post-translational modifications (PTMs) of EV proteins can be a valuable resource for biomarker candidates. Enzymes related to the NAD+ deacetylase group, such as sirtuin-1 (SIRT1) and sirtuin-6 (SIRT6), can have several different functions [17-26]. SIRT1 has been shown to participate in apoptosis, autophagy, senescence, proliferation, aging, tumorigenesis, genome regulation, stability and maintenance [17-23]. SIRT6 also affect several processes and is involved in DNA repair, glycolysis, gluconeogenesis, tumorigenesis, neurodegeneration and cardiac hypertrophic responses [24-26]. The presence of these enzymes in sEVs questions the role of protein acetylation in the regulation of BC progression and tumor growth.

Protein acetylation is a reversible PTM in which the acetyl group from acetyl coenzyme A (Ac-CoA) is transferred to either the α-amino group of the protein's N-terminus or to the ε-amino group

of lysine residues [27]. The lysine acetylation of proteins is catalyzed by lysine acetyltransferases (KATs), and this PTM is reversible and regulated by enzymes, namely deacetylases (KDACs), that can remove the acetyl group [28,29]. Acetylation results in the neutralization of the positive charge on the lysine residue and consequently alters protein function [28,29]. Interestingly, some studies have reported that lysine acetylation in mitochondria, where a high concentration of acetyl-CoA and an elevated pH are present, can also occur in a non-enzymatic manner [30,31].

By using proteomic approaches, 181 and 244 acetylation sites have been identified in human BC MDA-MB-231 cells from enrichment with the monoclonal antibody cocktail and the polyclonal antibody, respectively [32]. The overall acetylome revealed that the acetylation levels of the majority of proteins in BC tissue were significantly higher than those in normal tissue [33]. This study revealed that highly acetylated proteins were significantly enriched in histone H2A.X (H2A.X) complexes and nucleophosmin (NPM1). Atypically acetylated proteins have been shown to promote breast cancer metastasis and the proliferation of breast cancer [34-40]. Moreover, acetylation can inhibit the sensitivity of tumor cells to anti-tumor therapy oncogenesis and progression of BC [41,42]. The regulation of acetylated proteins such as CREB binding protein (CBP), ALDH1A1, α-tubulin, cortactin and Forkhead Box O3 (FOXO3) show tumor-suppressing effects in BC [43,44]. It has been demonstrated that inhibitors targeting protein acetylation can be used as potential drug candidates [34].

Reversible lysine acetylation acts as an indispensable regulator in multiple cellular pathways, oncogenesis and progression of BC. To date, the potential role of this PTM in extracellular vesicles related to cancer and more specifically to BC is still unknown. The fact that protein acetylation can affect protein function will elicit interest in exploring and evaluating this PTM for early diagnosis of BC. This study explored the lysine acetylome of sEVs derived from metastatic BC cell lines, MDA-MB-231 and MCF7, and the non-cancerous breast tissue cell line MCF10A. Differences in EV acetylomes were observed between the non-cancerous and cancerous breast cell lines. We selected and tested the enzymatic activity of five acetylated enzymes: aldolase (ALDOA), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate kinase (PGK 1), enolase (ENO) and pyruvate kinase M1/2 (PKM). ALDOA, PGK 1 and ENO had significantly higher specific enzymatic activities in MDA-MB-231 compared to MCF10A-derived sEVs.

3.3.3 Materials and Methods

3.3.3.1 Cell Culturing and sEV Isolation

Culturing of epithelial breast cancer cell line MDA-MB-231 (ATCC HTB-26), MCF7 (ATCC HTB-22) and MCF10A non-tumorigenic epithelial breast cancer cell line (ATCC CRL-10317) was performed as described previously [7,8,16]. sEVs were isolated by using differential ultracentrifugation as described in our previous investigation [7,16].

3.3.3.2 Quantification of sEVs by Nanoparticle Tracking Analysis (NTA)

The concentration and size distribution of sEVs were determined using The ZetaView nanoparticle tracking microscope PMX-110 (Particle Metrix, Meerbusch, Germany) at 85 and 40 camera shutter speeds [16].

3.3.3.3 Characterization of EVs Protein Markers

Marker-based assessment of EVs isolated from MCF10A, MCF7 and MDA-MB-231 cell lines was performed using the commercially available Exo-Check Exosome Antibody Array kit (System Biosciences, Palo Alto, CA, USA) according to the manufacturer's protocol.

3.3.3.4 Sample Preparation for Acetylomics

Isolated sEVs were added in lysis buffer with a volume ratio of 4:1 (fraction/buffer) consisting of a final concentration of 20 mM HEPES, pH 8.0, 0.1% NP-40, 1 mM DTT, 1.6 M urea, 1/1000 (*v*:*v*) protease inhibitor cocktail (Cat No. 78430, Thermo Fischer Scientific, Mississauga, ON, Canada), 3 µM Trichostatin A (TSA) and 10 mM Nicotinamide (NAM) and were gently vortexed for 2 min. The suspension was centrifuged for 3 min at 10,000x *g* and the supernatant was then collected. The protein concentration was determined using a Bradford protein assay kit (Thermo Scientific, Cat No. 23236, Waltham, MA, USA).

3.3.3.5 Protein Reduction, Alkylation and Enzymatic Digestion

Protein samples obtained from the isolated sEVs were reduced, alkylated and digested using FASEB method as described previously [45]. Proteolytic digestion was performed by addition of sequencing grade modified trypsin (Promega, #V5111, Madison, WI, USA), 1:300 enzyme to protein ratio, and incubated under shaking at 500 rpm at ambient temperature overnight. The digestion was stopped by addition of formic acid (1% final concentration) and centrifuged at 15,000× *g* for 3 min. The supernatant containing about 100 µg of digested protein was desalted on disposable TopTip C-18 columns (Glygen, #TT2C18.96, Ellicott City, MD, USA) and dried by vacuum centrifugation.

3.3.3.6 Enrichment of Acetylated Peptides

The peptides were dissolved in 100 µL of the immunoprecipitation buffer solution containing 50 mM HEPES (pH 8.0), 100 mM NaCl, 1 mM EDTA, 0.1% NP-40 and incubated with 30 µL pre-washed antibody beads (catalog no. PTM-104 for Kac, PTM Biolabs, Inc., Hangzhou, China) at 4 °C overnight with gentle shaking. The bound lysine-acetylated peptides were then processed according to the protocol of PTM Biolabs. The eluted peptides were collected and vacuum-dried followed by resuspension in 20 µL of 0.1% FA and analyzed by LC-MS/MS.

3.3.3.7 Nano-LC-MS/MS

An Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific, Mississauga, ON, Canada) equipped with an UltiMate 3000 nanoRSLC (Thermo Fisher Scientific, Mississauga, ON, Canada) was used for nanoLC-MS/MS analysis. Three microliters of enriched acetylated peptides were loaded onto the column for 65 min at a flow rate of 0.30 µL/min and separated on an in-house packed column (Polymicro Technology, Phoenix, AZ, USA), 15 cm × 70 µm ID, Luna C18(2), 3 µm, 100 Å (Phenomenex) employing a water/ACN/0.1% formic acid gradient. The following steps were employed: 0–10 min, 2–2% ACN; 10–40 min, 2–38% ACN; 40–45 min, 38–98% ACN; 45–50 min, 98–98% ACN; 50–55 min, 98–2% ACN; 55–65 min, 2–2% ACN. Data-dependent MS/MS acquisition was performed following a full MS survey scan. The Orbitrap parameters in ESI+ were set up as follows: ion spray voltage 2.1 kV, ion source temperature 250 °C, top speed mode over the *m/z* range (*m/z* 350–2000) at a resolution of 60,000. Precursor ions were filtered according to monoisotopic precursor selection and charge state (+2 to +7), and dynamic exclusion was enabled for 30 s. The automatic gain control settings were 5×10^5 for full scan and 1×10^4 for MS/MS scans. Fragmentation was performed with collision-induced dissociation (CID) in the linear ion trap. Precursors were isolated using a 2 *m/z* isolation window and fragmented with a normalized collision energy of 35%.

3.3.3.8 MS Spectra Processing

Proteome Discoverer 2.1 (Thermo Fisher Scientific, Mississauga, ON, Canada) was used for protein identification. The precursor mass tolerance was set at 10 ppm and 0.6 Da mass tolerance for fragment ions. Search engine SEQUEST-HT implemented in Proteome Discoverer was applied for all MS raw files. Search parameters were set to allow for dynamic modification of methionine oxidation, lysine (K) acetylation, N-terminus acetylation and static modification of cysteine carbamidomethylation. Peptides were searched against a human UniProt FASTA file containing

20,396 entries (21 April 2021) and a default contaminants database. The false discovery rate (FDR) was set to 0.01 for both the protein and peptide level.

3.3.3.9 Data Filtering and Acetylation Site Localization

The SEQUEST-HT output table containing information on PSMs was used for all analysis in R [46]. Each cell line was cultured and analyzed in triplicate. Acetylated peptides which were found in at least two out of the three replicates were used for further analysis. Acetylated peptides were then collapsed to acetylation sites on unique proteins. Enriched motifs for acetylation sites per cell line were identified with motif-x. The biological function of acetylated proteins in each cell line were annotated with Gene Ontology (GO) and KEGG terms. Identified acetylation sites were compared with known acetylation sites downloaded from PhosphoSitePlus (20 October 2022) [47].

3.3.3.10 Disease and Functional Annotation Analysis

Gene ontology and KEGG functional annotations were obtained through the clusterProfiler R package [48-50]. Absolute numbers of proteins belonging to presented biological themes were compared between cell lines.

3.3.3.11 Data Availability

All MS raw data were submitted to the PRIDE repository (Accession: PXD040413) at the European Bioinformatics Institute.

3.3.3.12 Aldolase Activity Assay

Cells and isolated extracellular vesicles were resuspended in ice-cold Aldolase Assay Buffer from the kits (ab196994, Abcam, Toronto, ON, Canada) and gently vortexed. After centrifugation at 10,000× *g* for 1 min, supernatants were collected and used for enzymatic activities. The assays were performed at 37 °C, and absorbances were measured at 450 nm in a kinetic mode according to the manufacturer's instructions. The activity of enzymes was calculated from the assay time of 30 min.

3.3.3.13 Glyceraldehyde 3 Phosphate Dehydrogenase Activity Assay

The glyceraldehyde 3 phosphate dehydrogenase activity was measured using a Glyceraldehyde 3 phosphate dehydrogenase Activity kit (ab204732, Abcam, Toronto, ON, Canada) according to the manufacturer's procedures. The assays were performed in 100 µL of the reaction mixture on a microplate reader at 37 °C, and absorbances were measured at 450 nm in a kinetic mode based on the manufacturer's instructions. The activity of enzymes was calculated from the assay time at 10 (cells) and 50 min (EVs).

3.3.3.14 Phosphoglycerate Kinase Activity Assay

The Phosphoglycerate Kinase Activity was assessed by using a colorimetric assay (The Phosphoglycerate Kinase Activity Assay Kit, ab252890, Abcam, Toronto, ON, Canada) following the manufacturer's procedures. The assays were performed on a microplate reader (Greiner Bio-One 655209; Fischer Scientific, Toronto, ON, Canada) at 37 °C in a kinetic mode at 340 nm as described in the manufacturer's instructions. The activity of enzymes was calculated from the assay time of 20 min.

3.3.3.15 Enolase Activity Assay

The protein extracts from cells and EVs were treated with enolase assay buffer (Abcam Enolase assay kit, ab241024, Abcam, Toronto, ON, Canada), and supernatants were collected after centrifugation at 10,000× *g* for 1 min. The assay was performed as indicated in the manufacturer's instructions at 37 °C in a kinetic mode using a fluorometric method (Ex/Em = 535/587). The activity of enzymes was calculated from the assay time between 10 and 30 min.

3.3.3.16 Pyruvate Kinase Activity Assay

Briefly, the reaction was started by incubating the protein extract with the reaction mixture following the manufacturer's instructions (Pyruvate Kinase Assay Kit, ab83432, Abcam, Toronto, ON, Canada). Samples were incubated at 37 °C in a kinetic mode, and the fluorescence was measured between 0 and 10 min using a microplate reader (Ex/Em = 535/587).

3.3.3.17 Statistical Analysis

Enzymatic assays were conducted at least in quadruplicates. Replicates of each cell line within a sample origin group, cell-free extract or sEVs, were compared for statistical significance of their means by a Student's *t*-test. All *p*-values were annotated and a *p*-value of <0.05 was considered significant.

3.3.4 Results

3.3.4.1 Isolation of sEV

To explore the potential role of acetylated proteins for early BC diagnostics, we performed a systematic acetylome analysis of sEVs derived from MCF10A, MCF7 and MDA-MB-231 cell lines. Nanoparticle tracking analysis (NTA) was used to evaluate the size distribution of sEVs isolated by differential ultracentrifugation. The diameter of isolated sEVs ranged from 50 to 300 nm, with an average size of about 125 nm (Figure 3.3.1). Additionally, the ExoCheck kit validated

the presence of several external (CD63, EpCAM, ANXA5, CD81 and ICAM) and internal (TSG101, ALIX and FLOT1) EV markers in isolates from all three cell lines (Figure S1).

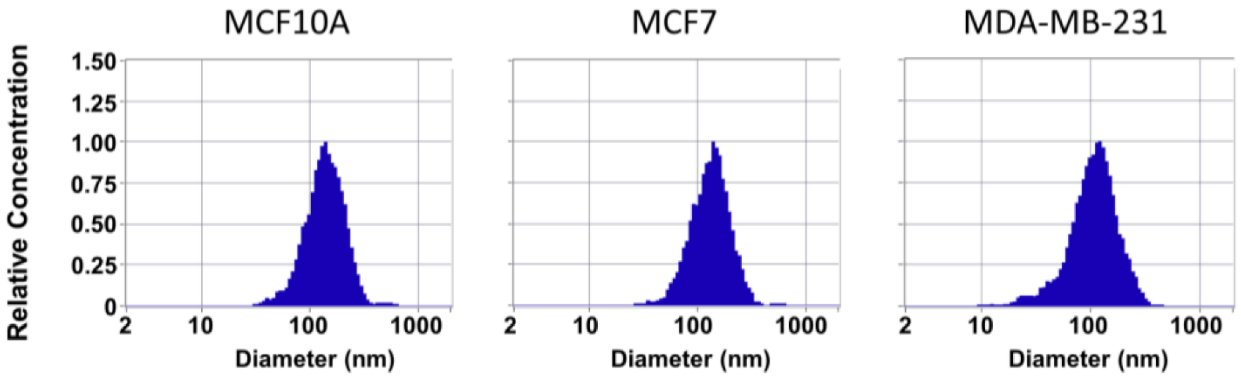

Figure 3.3.1. Nanoparticle tracking analysis (NTA) of extracellular vesicles. NTA plots show the size distribution profiles of isolated sEVs from MCF10A, MCF7 and MDA-MB-231 cell lines.

3.3.4.2 Overall Acetylome Profiling

Samples of about 100 µg from MCF10A, MCF7 and MDA-MB-231 derived sEVs collected from three biological replicates were prepared and examined independently. Proteins were digested and the acetylated peptides were enriched using anti-acetyl-lysine (KAC) immunoaffinity chromatography to enhance the identification of low-abundance acetylated proteins. A schematic overview of our experimental strategy is presented in Figure 3.3.2A. After harvesting and isolating sEVs using ultra-centrifugation, acetylated peptides were enriched then analyzed by nano liquid chromatography-tandem mass-spectrometry (nLC-MS/MS) using CID fragmentation mode. The proteomic mass spectrometry data were subjected to rigorous bioinformatics assessment and protein identification related to lysine acetylation. Acetylated peptides were considered reliable if they were identified in at least two biological replicates. After removal of potential contaminants like keratins, the total number of identified acetylated peptides for each cell line is presented in Figure 3.3.2B and Table S1. In total, from 118 acetylated peptides 22, 58 and 82 have been identified in MCF10A, MCF7 and MDA-MB-231sEVs, respectively (Figure 3.3.2B). The Venn diagram of acetylated proteins revealed in total 60 identified proteins from three cell lines including 13, 41 and 44 acetylated proteins from MCF10A, MCF7 and MDA-MB-231 sEVs, respectively (Figure 3.3.2C and Table S2).

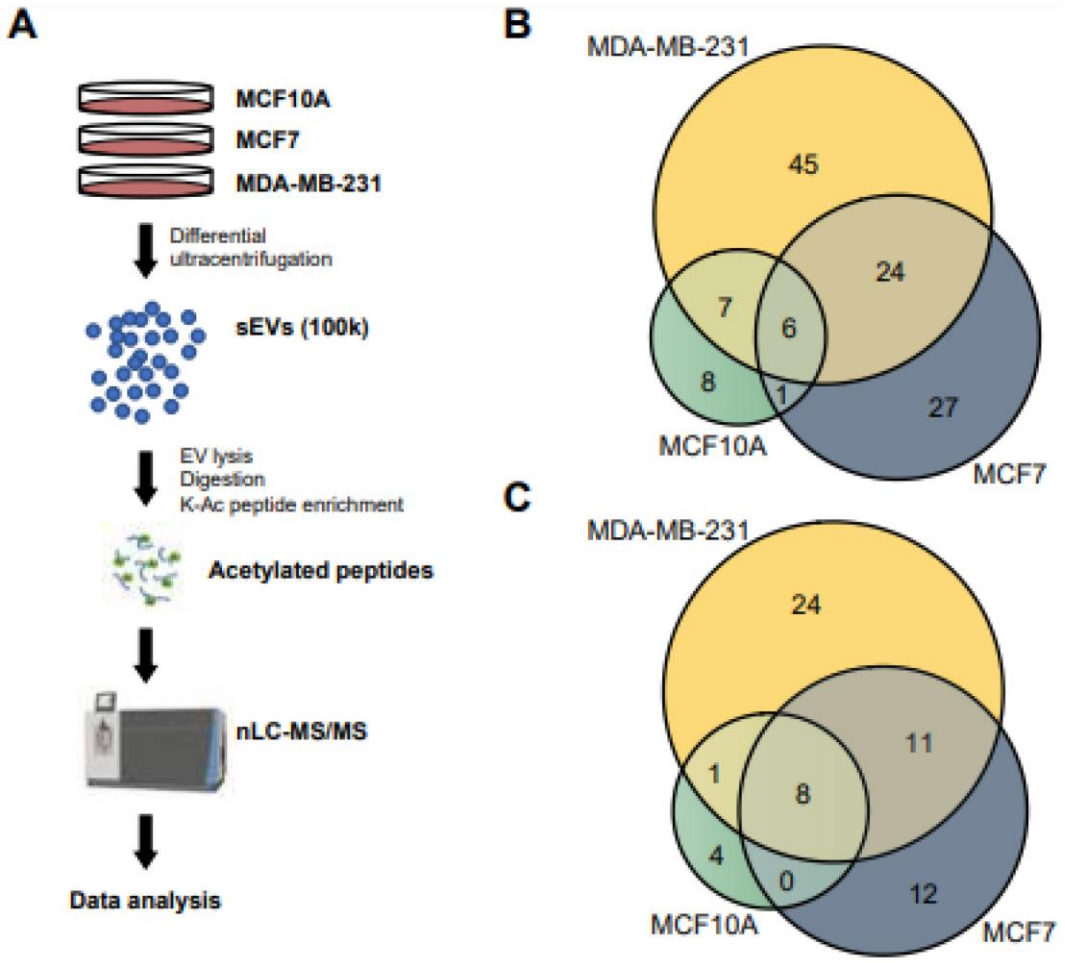

Figure 3.3.2. Overview of the identified acetylated peptides/proteins. (**A**) Workflow illustration of acetylome analysis of sEVs from MCF10A, MCF7 and MDA-MB-231 cells. (**B**) Total number of acetylated peptides identified from three independent replicates of each cell line (**C**) A Venn diagram showing the number of EV acetylated proteins identified from the three cell lines.

3.3.4.3 Acetylation Site Distributions and Motifs for the Identified Acetylation Sites

The distribution of acetylated residues per protein in sEVs derived from three breast cell lines is presented in Figure 3.3.3A. Among the 97 identified sites, 25 are not previously reported (Table S3). A single acetylation site was localized on most of the identified acetylated proteins in all three sEV fractions. A notable number of the identified proteins were found to be lysine acetylated at two or more sites. MCF7 (eight acetylated proteins) and MDA-MB-231 (16 acetylated proteins) derived sEVs contained a higher number of multiply acetylated proteins than MCF10A (five

acetylated proteins) derived particles (Figure 3.3.3A). Some proteins contained five or more acetylated sites such as fatty acid synthase (FASN) and H4 clustered histone 9 (H4C9).

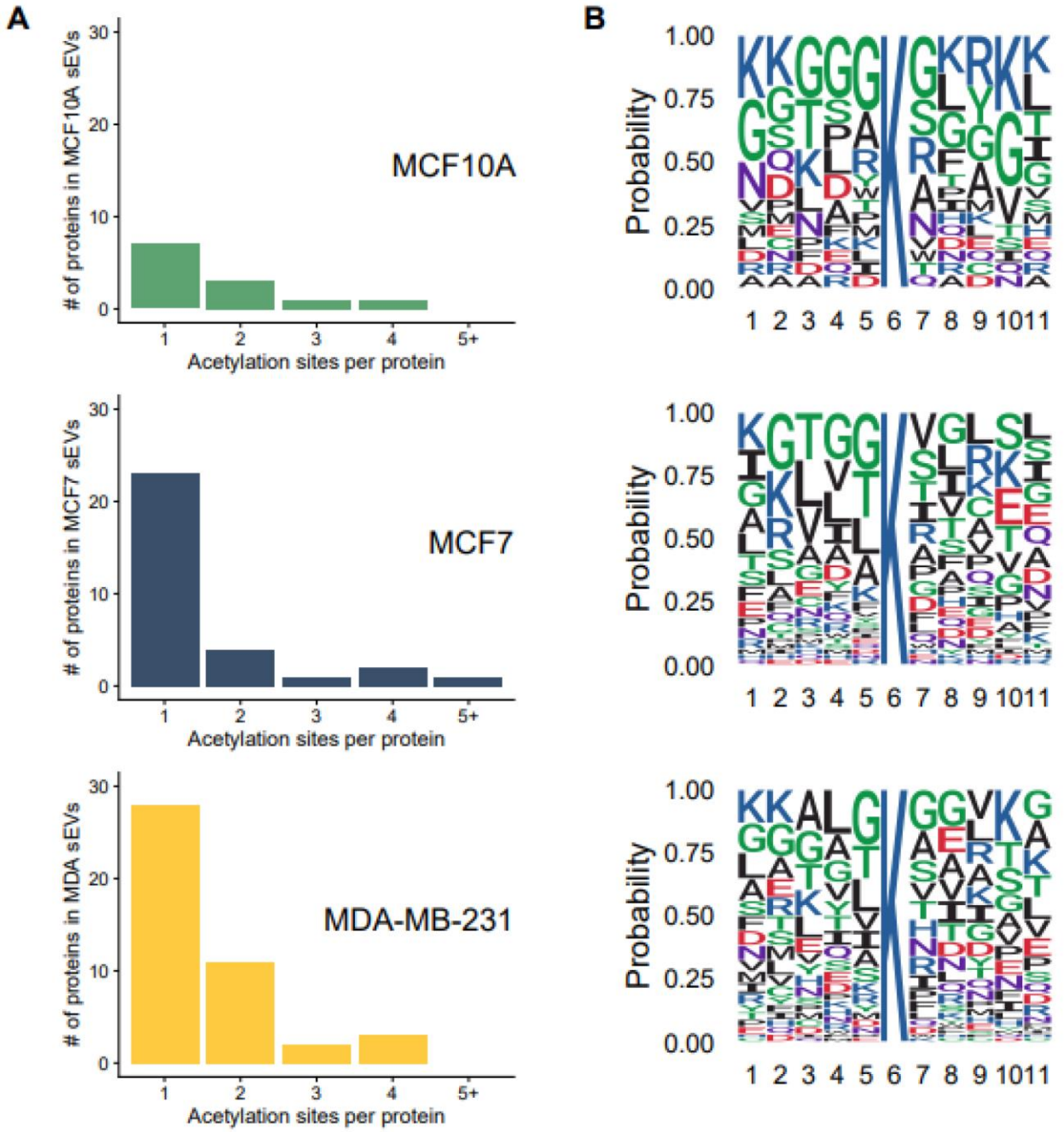

Figure 3.3.3. Overview of the acetylated sites and acetylation motifs. (A) Barplot depicting the distribution of numbers of acetylation sites observed per protein in all three EV-derived cell lines. (B) Conservation of acetylation motif for ±10 amino acids around the lysine acetylation sites. The central K refers to the acetylated lysine and the size of each letter corresponds to the frequency of the amino acid residue in that position.

Visualized sequence motifs from the acetylation for each cell lines derived sEV are presented in Figure 3.3.3B. Acetylation motives have not revealed substantially different motifs for lysine acetylation from all EV fractions. It appears that G at the −1 and −2 positions as well as K and G at the +4 position are overrepresented in MCF10A in comparison to MCF7 and MDA EVs. Some variations of amino acids were observed at the position from +1 to +5 in all three EV fractions. Therefore, these results may suggest different acetylation dynamics between the three cell lines.

3.3.4.4 Functional and Pathway Analysis of Identified Acetylated Proteins

Identified acetylated proteins were annotated and classified using the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. Within the GO subcellular localization terms, the identified lysine-acetylated proteins were mainly localized to the membrane-bounded organelle, cytoplasm, extracellular vesicle, nucleus and membrane (Figure 3.3.4A). KEGG enrichment demonstrated that many proteins were associated with the terms of metabolic pathways, glycolysis/gluconeogenesis, biosynthesis of amino acids, HIF-1 signaling pathway and pyruvate metabolism (Figure 3.3.4B). For all these terms, a larger number of proteins was presented in MCF7 and MDA-MB-231 derived sEVs, in comparison to MCF10A derived sEVs. Remarkably, the majority of acetylated proteins were associated to the glycolysis pathway (Figure 3.3.5).

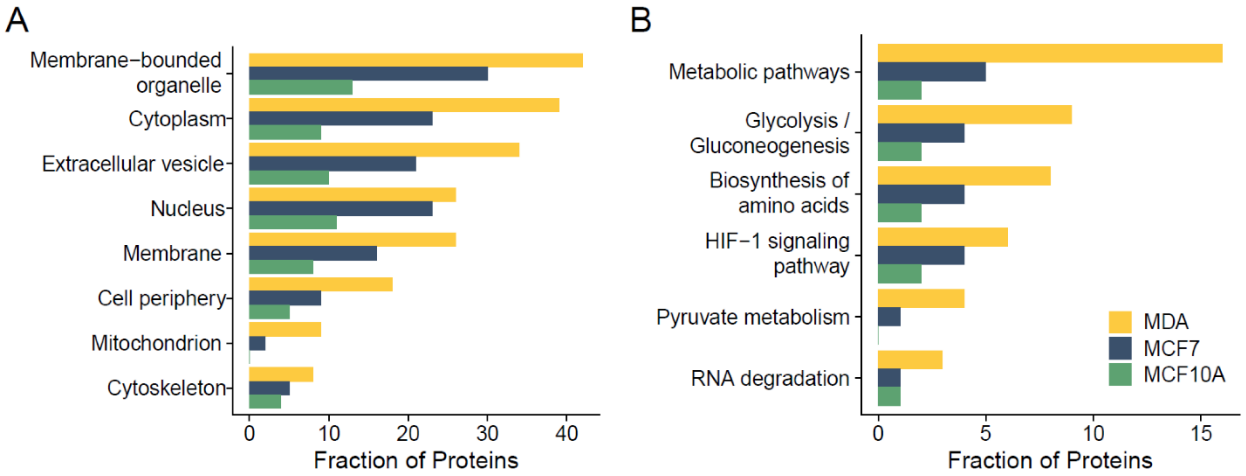

Figure 3.3.4. GO cellular localization (**A**) and KEGG pathways analysis (**B**) of acetylated proteins derived from MCF10A, MCF7 and MDA-MB-231 sEVs.

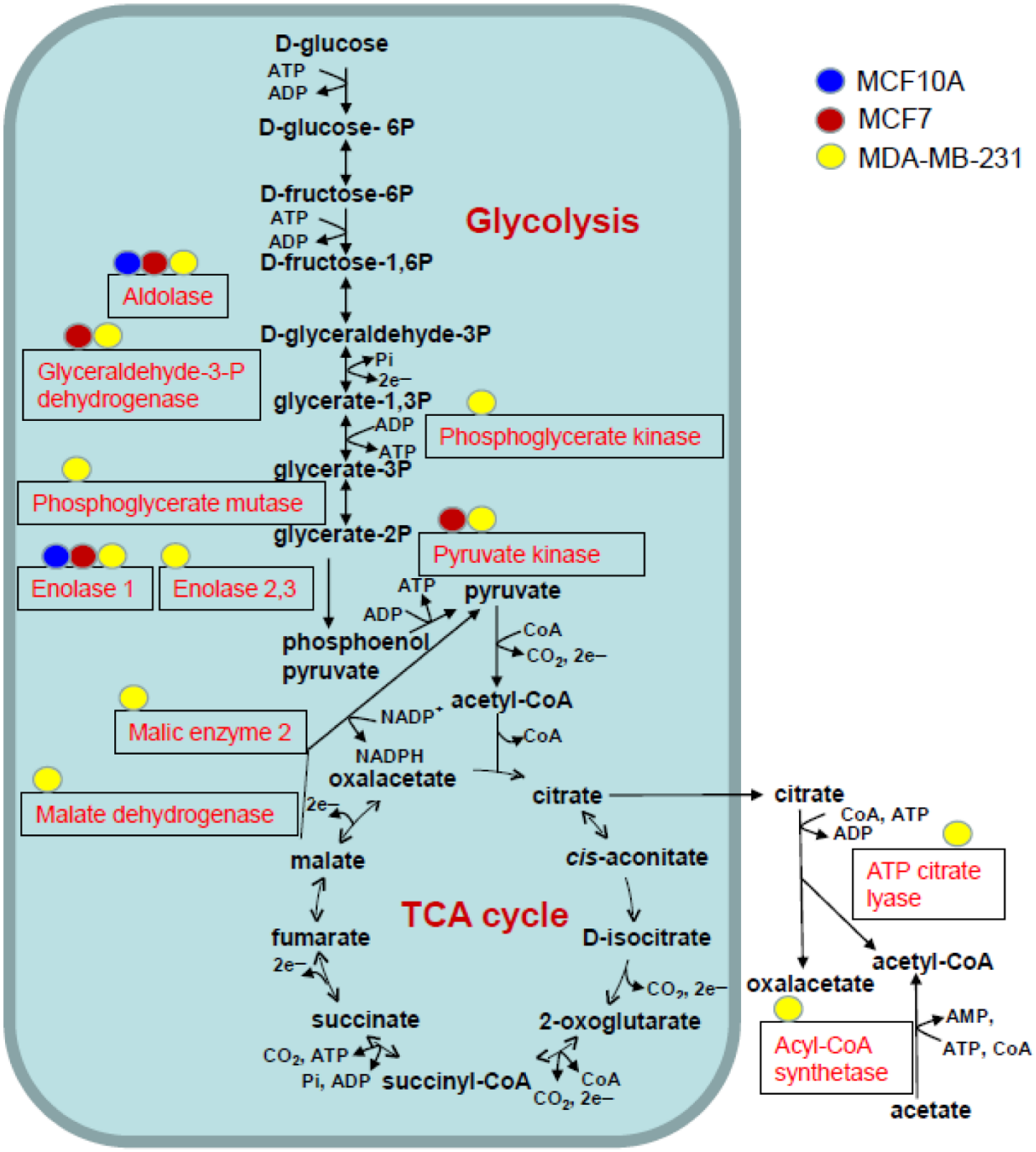

Figure 3.3.5. The enrichment of acetylated enzymes in the glycolysis pathway and TCA cycle in breast cancer cell lines. The acetylated enzymes from MCF10A, MCF7 and MDA-MB-231 are marked in blue, red and yellow respectively.

A noteworthy number of lysine-acetylated proteins from MCF7 and MDA-MB-231 derived sEVs were observed for annexins (Figure S2). Numerous acetylated histones were identified

among different cell line derived sEVs including H4 clustered histone 9, H3 clustered histone 1, H3.4 histone, H2B clustered histone 18, H2B clustered histone 20 and H1.5 linker histone (Table S2). Interestingly, KAT8 regulatory NSL complex subunit 1, which was identified in this study, is known to be involved in the acetylation of nucleosomal histone H4 in various lysed residues [51].

Our previous work focused on identifying metabolic enzymes in BC-derived EVs and MVs that could be explored as diagnostic BC biomarker [8,16]. We chose to focus on enzymes as biomarkers since they can be easily incorporated into quick, accurate and sensitive diagnostic assays that can assess specific enzymatic activity. sEV enzymes annotated by the KEGG database with the term "glycolysis/metabolic pathways" were selected for downstream experiments. We investigated five enzymes (Table 3.3.1, Figure S3) mainly acetylated in MCF7 and/or MDA-MB-231 sEVs: aldolase (ALDOA), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate kinase (PGK1), enolase (ENO) and pyruvate kinase M1/2 (PKM). We did not perform further validation for phosphoglycerate mutase 1 (PGAM1) because of the unavailability of a kit capable of measuring its enzymatic activity.

Table 3.3.1. Summary of the acetylation sites for enzymes related to the glycolysis pathway, TCA cycle, and acetyl-CoA metabolism from MCF10A, MCF7 and MDA-MB-231 sEVs. These acetylated enzymes were identified at least in two replicate samples for each cell line.

Enzyme	**Gene**	**Acetylation Sites**	**MCF10A sEV**	**MCF7 sEV**	**MDA-MB-231 sEV**
			Position of Acetylation		
Aldolase	ALDOA	3	K-147	K-42, K-147, K-230	K-147, K-230
Glyceraldehyde-3-P-dehydrogenase	GAPDH	3		K-219	K-61, K-194, K-219
Phosphoglycerate kinase 1	PGK1	1			K-131
phosphoglycerate mutase 1	PGAM1	1			K-100

Enzyme	Gene	Acetylation Sites	MCF10A sEV	MCF7 sEV	MDA-MB-231 sEV
			Position of Acetylation		
Enolase 1	ENO1	4	K-343	K-343	K-60, K-193, K-203, K-343
Enolase 2	ENO2	1			K-343
Enolase 3	ENO3	2			K-60, K394
pyruvate kinase M1/2	PKM	1		K-433	K-433

3.3.4.5 Analysis of ALDOA, GAPDH, PGK 1, ENO and PKM in Cells and Their sEVs

The enzymatic activity was detected in all extracted protein fractions from the three cell lines (Figure 3.3.6, Table S5). The specific activity levels of all tested enzymes are significantly higher in MDA-MB-231 when compared with MCF10A cell line. Similarly, the specific activities of ALDOA and ENO were significantly higher in MDA-MB-231 and MCF7 when compared with MCF10A cell derived sEVs. The specific activity of PGK1 was significantly higher in MDA-MB-231 in comparison to MCF10A and MCF7 sEV fractions. Very low or absence of activity was observed for GAPDH and PKM in all EV fractions. Although some activity of GAPDH was observed in MDA-MB-231 and MCF7 from sEV in comparison to MCF10A, the results cannot be conclusive because of very low specific activity. Therefore, ALDOA and ENO reveal a more significant difference in specific activity between sEV derived from MCF10A and MCF7 or MDA-MB-231.

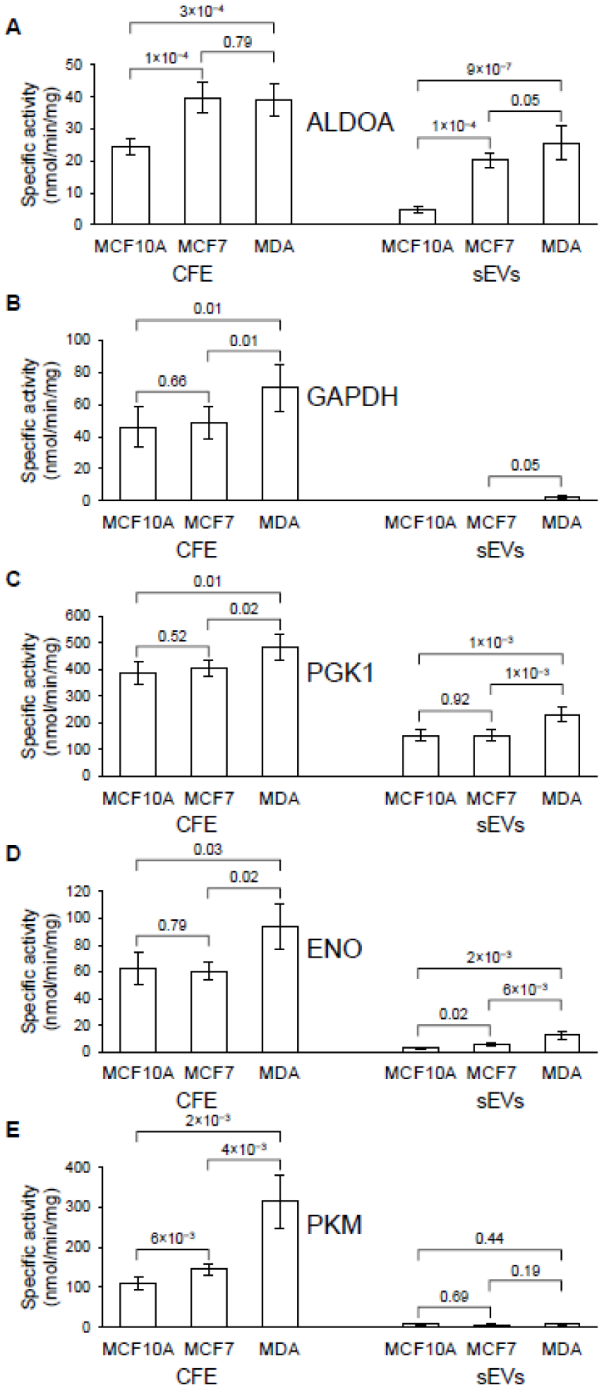

Figure 3.3.6. Specific enzymatic activity of (**A**) aldolase (ALDOA), (**B**) glyceraldehyde-3-phosphate dehydrogenase (GAPDH), (**C**) phosphoglycerate kinase (PGK1), (**D**) enolase (ENO) and (**E**) pyruvate kinase M1/2 (PKM) in cell-free extract (CFE) and their corresponding sEV fractions. The bar graph represents mean values, while error bars indicate the standard deviation (SD) of at least five replicates, and *p*-values obtained from a Student's t-test are indicated on the top.

3.3.5 Discussion

In recent years, proteomics has played a remarkable role in providing information about EVs and their function in breast cancer tumor progression and metastasis [2-8]. Post-translational modification of proteins intersects with cancer, playing a decisive role in regulating various cellular processes and being implicated in the major regulatory mechanisms of cell signaling networks. In our previous investigation, using proteomics analysis, we demonstrated that EVs contain functional, phosphorylated, metabolic enzymes that could be potential candidates in early BC diagnostic and therapeutic applications [16]. In this study, 25 novel acetylated sites were identified in sEVs from two breast cancer cell lines (MCF7 and MDA-MB-231) and the non-metastatic (MCF10A) cell line. These acetylated sites were mapped to 60 distinct proteins. The cancerous cell lines used in this work present a good model for studying invasive ductal carcinoma (MCF7) and triple-negative BC (MDA-MB-231).

Although enrichment analysis was performed using the same initial quantity of proteins, about 100 µg from all three sEVs, there was a significant difference in the number of identified acetylated proteins isolated from these three cell lines. Therefore, EVs from MCF10A produced a lower number of lysine acetylated proteins in comparison to MDA-MB-231 and MCF7. This finding suggests that lysine acetyltransferase enzymes are working more effectively in cancerous cell lines. Some of these enzymes have been overexpressed in various types of cancers such as colon, liver, glioma, bone, lung, breast and prostate cancers [52,53]. Accordingly, some studies have explored small-molecule inhibitors from the KAT protein family for cancer therapy [53]. In this study, numerous identified acetylated proteins in sEV-derived BC cell lines indicated that their modification could affect protein function and holds a biological role and prognostic value in breast cancer.

The majority of acetylated proteins were identified to participate in various metabolic processes such as glycolysis, biosynthesis of amino acids and pyruvate metabolism. It has been investigated that cancer cells are able to reprogram different metabolic processes to meet requirements for proliferation, invasion and survival in hostile environments [54,55]. Glucose is one of the primary fuel sources for malignant cells and plays an important role in tumorigenic metabolism and cancer progression [56]. Therefore, it is not surprising that several identified acetylated proteins are associated with the glycolysis pathway. During the process of oncogenesis,

an essential piece of metabolic reprogramming of cancerous cells is enhancing aerobic glycolysis and glucose uptake [57,58]. Five enzymes involved in glycolysis present only or mainly in sEVs isolated from cancer-derived cell lines were validated: aldolase (ALDOA), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate kinase (PGK1), enolase (ENO) and pyruvate kinase (PKM). However, only three of these enzymes, ALDOA, PGK1 and ENO, have a higher value for specific enzymatic activity in MDA-MB-231 derived sEVs when compared with MCF10A ones. Several investigations demonstrated increasing activity/expression of these enzymes in breast cancer cells or tissues and are proposed as promising breast tumor markers. The enzyme activities of ALDOA and ENO were found to significantly increase in cancerous breast tissue when compared to normal tissue [59]. The overexpression of phosphoglycerate kinase has been demonstrated in most cancer types, including BC [60]. This study validated that PGK1 can be used for BC prognosis. In the present work, we found that the activities of ALDOA, PGK1 and ENO were increased in BC derived sEVs isolated from the MDA-MB-231 cell line, in comparison to sEVs isolated from MCF10A cell lines. These findings suggest that sEV enzymes may be potential clinical indicators for the diagnosis of BC.

The presence of metabolite products from identified acetylated enzymes from the glycolysis pathway and malate dehydrogenase from the TCA cycle can serve as precursors for amino and fatty acid synthesis [61]. This finding suggests that pathways for amino and fatty acid synthesis were affected with lysine acetylation. This is not surprising because it is known that in addition to glucose, tumor cells metabolize amino and fatty acids at much higher rates than their nontumor equivalents [62].

In this study, we also identified four acetylated annexin proteins (annexin 1, 2, 5 and 6) present in cancerous cell line derived sEVs. Annexin family members can regulate various cellular functions including vesicle trafficking, vesicle fusion, plasma membrane repair, promotion of membrane segregation and actin cytoskeleton dynamic regulation [63-66]. Dysregulation of annexins has been associated with multiple cancers including BC, and these proteins have emerged as potential biomarkers and pharmacological targets for medical applications [67]. Quantitative mass spectrometry-based proteomic analyses revealed significantly higher ANXA1 protein expression in more invasive glioblastoma cells [68]. The abundance of ANXA1 was significantly higher in the high-grade glioblastoma EVs compared with low-grade glioblastoma EVs [69]. Therefore, this protein has been proposed as a potential EV biomarker for glioblastoma. In a model of human BC,

ANXA1 can promote tumor-formation, and it has been suggested that some basal-like TNBCs may require high endogenous tumor cell Annexin A1 expression for continued growth [70]. In addition, ANXA1 has been suggested to be a useful prognostic marker in HER-2+ BC patients [71]. The protein Annexin A2 has been explored as a prognostic marker because of the expression of this protein in various cancer cells [72] including oral squamous cell carcinoma [73] and prostate cancer [74-76]. It was found that Annexin A2 is highly expressed in breast tumor tissues [77] and that FOXD1-dependent RalA-ANXA2-Src complex promotes circulating tumor cell formation in BC [78]. Moreover, Annexin A5 (ANXA5) was found to promote prostate cancer stem cells, pancreatic adenocarcinoma, sarcoma and tumorigenesis and progression of BC [79]. This protein has been proposed to be a predictive biomarker for tumor development, metastasis and invasion and be of diagnostic, prognostic and therapeutic significance in cancer [79]. Moreover, in BC cells, ANXA5 up-regulation suppresses Raf-1 and MEK1/2 expressions, ERK1/2 phosphorylation and Ras activation in MCF-7 [80]. The protein ANXA6 has been implicated in various cellular functions including cell growth, differentiation and motility, which underlie tumor progression [81]. Consequently, reduced expression of ANXA6 is associated with decreased cell motility and rapid growth of xenograft TNBC tumors in vivo [81]. Our work demonstrates the high level of acetylation of several annexins in cancerous BC derived EVs suggesting that this post-translational modification may cause a change in protein structure and function, and consequently alter its function.

Aberrant acetylated histones identified mainly in cancerous cell line derived sEVs include H4 clustered histone 9, H3 clustered histone 1, H3.4 histone, H2B clustered histone 18, H2B clustered histone 20 and H1.5 linker histone. Epigenetic regulation via histone acetylation plays an important role in chromatin remodeling and in the regulation of gene transcription, which can be involved in tumorigenesis [53,82]. Several studies have suggested modifications to histone acetylation and proposed them as potential diagnostic or prognostic biomarkers in cancer [53]. Analysis of global histone modification in BC patients' analysis revealed that the loss of acetylated H4 at K16 may serve as an early indication of cancer, and low levels of H3 acetylated at K9 and K14, and H4 at K12 are prognostic of poor outcomes [83]. In our work we identified the acetylation of H2B clustered histone 18 in both MCF7 and MDA-MB-231 derived EVs. It has been shown that the expression profile of MAP3K4-deficient trophoblast stem cells includes an H2B acetylation-regulated gene signature that closely overlaps with that of human BC cells [84].

In our previously published works [8,16], we investigated potential enzyme biomarkers for BC early detection, exploring the proteomic content of sEVs from cancerous cell lines. The discovery of potential biomarkers based on enzymatic assays with reliable clinical significance would most likely require a panel of multiple enzymes. For this reason, future studies to find more enzyme candidates should be performed using proteomics approaches on different cell lines as well as isolated blood from healthy and BC patients. Our results, based on proteomics analysis of sEVs from cancerous cell lines, found that three enzymes, ALDOA, PGK1 and ENO, may be potential biomarkers for BC diagnosis. These enzymes were already investigated to significantly increase activity in BC cells and tissue [59,60] supporting future exploration of these enzymes for BC prognosis.

3.3.6 Conclusions

Here, we performed a protein acetylation analysis of breast cancer-derived extracellular vesicles from MCF10A, MCF7 and MDA-MB-231 cell lines. We found that 60 distinct acetylated proteins were identified collectively among cell lines, covering different metabolic functions and especially within the glycolytic pathway. Among these acetylated proteins, we validated five glycolytic enzymes, ALDOA, GAPDH, PGK1, ENO and PKM. The results revealed that the specific activity of GAPDH and PKM in cancerous cell lines was at the limit of detection and therefore cannot be taken in consideration as a viable biomarker. In contrast, acetylated enzymes ALDOA, PGK1 and ENO showed a significantly higher specific enzymatic activity in MDA-MB-231 in comparison to MCF10A-derived EVs. Our results, based on protein acetylation analysis of BC-derived sEVs as well as previously reported studies on BC cells or tissues, strongly suggest that these enzymes may be good candidates as potential prognostic biomarkers for BC.

3.3.7 Acknowledgements

We would like to acknowledge Yousef Risha for his technical help in Exo-Check exosome antibody array analysis and Ryan Reshke at the University of Ottawa for his excellent technical assistance in nanoparticle tracking analysis.

3.3.8 Supplementary Materials

The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/biomedicines11041076/s1, Figure S1: Exo-Check

exosome antibody array analysis for expression of well-known five external (CD63, EpCAM, ANXA5, CD81 and ICAM) and three internal (TSG101, ALIX and FLOT1) EV protein exosomal (or sEV) markers isolated from MCF10A, MCF7 and MDA-MB-231 cells. GM130 cis-Golgi marker is used to monitor any cellular contamination in sEV isolations, a labeled positive control for HRP detection, and a blank spot as a background control; Figure S2: The position of annexin acetylation sites from MCF10A, MCF7 and MDA-MB-231 marked with blue, red and yellow, respectively; Figure S3: The position of acetylated sites of enzymes in the glycolysis pathway from MCF10A, MCF7 and MDA-MB-231, respectively; Table S1: A total number of identified acetylated peptides for each sEV sample; Table S2: The list of identified acetylated proteins from three cell lines derived sEV; Table S3: The list of acetylated sites identified from three cell lines derived sEV; Table S4: GO functional annotation labels for cellular compartment and KEGG pathway analysis for sEVs from MCF10A, MCF7 and MDA-MB-231; Table S5: The specific activities of ALDOA, GAPDH, PGK1, ENO and PKM in the different cell free extracts and their derived sEV fractions. The numbers in parentheses indicate the number of biological replicates.

3.3.9 Funding Statement

The study was supported by the John L. Holmes Mass Spectrometry Facility for the sample preparation and MS analysis and granted by Natural Sciences and Engineering Research Council of Canada (M.V.B: grant # RGPIN-2020-05775).

3.3.10 Data Availability Statement

The raw MS data presented in this study is openly available at the PRIDE repository. This data can be found here: https://www.ebi.ac.uk/pride/archive/projects/PXD040413.

3.3.11 References

1. Kruger S., Abd Elmageed Z.Y., Hawke D.H., Wörner P.M., Jansen D.A., Abdel-Mageed A.B., Alt E.U., Izadpanah R. Molecular characterization of exosome-like vesicles from breast cancer cells. *BMC Cancer*. 2014;14:44. doi: 10.1186/1471-2407-14-44.
2. Palazzolo G., Albanese N.N., DI Cara G., Gygax D., Vittorelli M.L., Pucci-Minafra I. Proteomic analysis of exosome-like vesicles derived from breast cancer cells. *Anticancer Res*. 2012;32:847–860.
3. Hurwitz S.N., Rider M.A., Bundy J.L., Liu X., Singh R.K., Meckes D.G., Jr. Proteomic profiling of NCI-60 extracellular vesicles uncovers common protein cargo and cancer type-specific biomarkers. *Oncotarget*. 2016;7:86999–87015. doi: 10.18632/oncotarget.13569.

4. Demory Beckler M., Higginbotham J.N., Franklin J.L., Ham A.J., Halvey P.J., Imasuen I.E., Whitwell C., Li M., Liebler D.C., Coffey R.J. Proteomic analysis of exosomes from mutant KRAS colon cancer cells identifies intercellular transfer of mutant KRAS. *Mol. Cell Proteomics.* 2013;12:343–355. doi: 10.1074/mcp.M112.022806.
5. Liang B., Peng P., Chen S., Li L., Zhang M., Cao D., Yang J., Li H., Gui T., Li X., et al. Characterization and proteomic analysis of ovarian cancer-derived exosomes. *J. Proteomics.* 2013;80:171–182. doi: 10.1016/j.jprot.2012.12.029.
6. Maia J., Caja S., Strano Moraes M.C., Couto N., Costa-Silva B. Exosome-Based Cell-Cell Communication in the Tumor Microenvironment. *Front. Cell Dev. Biol.* 2018;6:18. doi: 10.3389/fcell.2018.00018.
7. Risha Y., Minic Z., Ghobadloo S.M., Berezovski M.V. The proteomic analysis of breast cell line exosomes reveals disease patterns and potential biomarkers. *Sci. Rep.* 2020;10:13572. doi: 10.1038/s41598-020-70393-4.
8. Risha Y., Susevski V., Hüttmann N., Poolsup S., Minic Z., Berezovski M.V. Breast Cancer-Derived Microvesicles Are the Source of Functional Metabolic Enzymes as Potential Targets for Cancer Therapy. *Biomedicines.* 2021;9:107. doi: 10.3390/biomedicines9020107.
9. Huang X., Yuan T., Tschannen M., Sun Z., Jacob H., Du M., Liang M., Dittmar R.L., Liu Y., Liang M., et al. Characterization of human plasma-derived exosomal RNAs by deep sequencing. *BMC Genom.* 2013;14:319. doi: 10.1186/1471-2164-14-319.
10. Carvalho A.S., Baeta H., Silva B.C., Moraes M.C.S., Bodo C., Beck H.C., Rodriguez M.S., Saraswat M., Pandey A., Matthiesen R. Extra-cellular vesicles carry proteome of cancer hallmarks. *Front. Biosci. (Landmark Ed.)* 2020;25:398–436. doi: 10.2741/4811.
11. Brzozowski J.S., Jankowski H., Bond D.R., McCague S.B., Munro B.R., Predebon M.J., Scarlett C.J., Skelding K.A., Weidenhofer J. Lipidomic profiling of extracellular vesicles derived from prostate and prostate cancer cell lines. *Lipids Health Dis.* 2018;17:211. doi: 10.1186/s12944-018-0854-x.
12. Xu R., Greening D.W., Zhu H.J., Takahashi N., Simpson R.J. Extracellular vesicle isolation and characterization: Toward clinical application. *J. Clin. Investig.* 2016;126:1152–1162. doi: 10.1172/JCI81129.
13. Minciacchi V.R., Freeman M.R., Di Vizio D. Extracellular vesicles in cancer: Exosomes, microvesicles and the emerging role of large oncosomes. *Semin. Cell Dev. Biol.* 2015;40:41–51. doi: 10.1016/j.semcdb.2015.02.010.
14. Di Vizio D., Morello M., Dudley A.C., Schow P.W., Adam R.M., Morley S., Mulholland D., Rotinen M., Hager M.H., Insabato L., et al. Large oncosomes in human prostate cancer tissues and in the circulation of mice with metastatic disease. *Am. J. Pathol.* 2012;181:1573–1584. doi: 10.1016/j.ajpath.2012.07.030.
15. Doyle L.M., Wang M.Z. Overview of Extracellular Vesicles, Their Origin, Composition, Purpose, and Methods for Exosome Isolation and Analysis. *Cells.* 2019;8:727. doi: 10.3390/cells8070727.
16. Minic Z., Hüttmann N., Poolsup S., Li Y., Susevski V., Zaripov E., Berezovski M.V. Phosphoproteomic Analysis of Breast Cancer-Derived Small Extracellular Vesicles Reveals Disease-Specific Phosphorylated Enzymes. *Biomedicines.* 2022;10:408. doi: 10.3390/biomedicines10020408.

17. Alves-Fernandes D.K., Jasiulionis M.G. The Role of SIRT1 on DNA Damage Response and Epigenetic Alterations in Cancer. *Int. J. Mol. Sci.* 2019;20:3153. doi: 10.3390/ijms20133153.
18. Yeung F., Hoberg J.E., Ramsey C.S., Keller M.D., Jones D.R., Frye R.A., Mayo M.W. Modulation of NF-kappaB-dependent transcription and cell survival by the SIRT1 deacetylase. *EMBO J.* 2004;23:2369–2380. doi: 10.1038/sj.emboj.7600244.
19. Yi Y.W., Kang H.J., Kim H.J., Kong Y., Brown M.L., Bae I. Targeting mutant p53 by a SIRT1 activator YK-3-237 inhibits the proliferation of triple-negative breast cancer cells. *Oncotarget.* 2013;4:984–994. doi: 10.18632/oncotarget.1070.
20. Jin M.S., Hyun C.L., Park I.A., Kim J.Y., Chung Y.R., Im S.A., Lee K.H., Moon H.G., Ryu H.S. SIRT1 induces tumor invasion by targeting epithelial mesenchymal transition-related pathway and is a prognostic marker in triple negative breast cancer. *Tumour Biol.* 2016;37:4743–4753. doi: 10.1007/s13277-015-4231-3.
21. Wang C., Yang W., Dong F., Guo Y., Tan J., Ruan S., Huang T. The prognostic role of Sirt1 expression in solid malignancies: A meta-analysis. *Oncotarget.* 2017;8:66343–66351. doi: 10.18632/oncotarget.18494.
22. Cao Y.W., Li W.Q., Wan G.X., Li Y.X., Du X.M., Li Y.C., Li F. Correlation and prognostic value of SIRT1 and Notch1 signaling in breast cancer. *J. Exp. Clin. Cancer Res.* 2014;33:97. doi: 10.1186/s13046-014-0097-2.
23. Chung Y.R., Kim H., Park S.Y., Park I.A., Jang J.J., Choe J.Y., Jung Y.Y., Im S.A., Moon H.G., Lee K.H., et al. Distinctive role of SIRT1 expression on tumor invasion and metastasis in breast cancer by molecular subtype. *Hum. Pathol.* 2015;46:1027–1035. doi: 10.1016/j.humpath.2015.03.015.
24. Khan R.I., Nirzhor S.S.R., Akter R. A Review of the Recent Advances Made with SIRT6 and its Implications on Aging Related Processes, Major Human Diseases, and Possible Therapeutic Targets. *Biomolecules.* 2018;8:44. doi: 10.3390/biom8030044.
25. Zhong L., D'Urso A., Toiber D., Sebastian C., Henry R.E., Vadysirisack D.D., Guimaraes A., Marinelli B., Wikstrom J.D., Nir T., et al. The histone deacetylase Sirt6 regulates glucose homeostasis via Hif1alpha. *Cell.* 2010;140:280–293. doi: 10.1016/j.cell.2009.12.041.
26. Sebastián C., Zwaans B.M., Silberman D.M., Gymrek M., Goren A., Zhong L., Ram O., Truelove J., Guimaraes A.R., Toiber D., et al. The histone deacetylase SIRT6 is a tumor suppressor that controls cancer metabolism. *Cell.* 2012;151:1185–1199. doi: 10.1016/j.cell.2012.10.047.
27. Drazic A., Myklebust L.M., Ree R., Arnesen T. The world of protein acetylation. *Biochim. Biophys. Acta.* 2016;1864:1372–1401. doi: 10.1016/j.bbapap.2016.06.007.
28. Choudhary C., Weinert B.T., Nishida Y., Verdin E., Mann M. The growing landscape of lysine acetylation links metabolism and cell signalling. *Nat. Rev. Mol. Cell Biol.* 2014;15:536–550. doi: 10.1038/nrm3841.
29. Verdin E., Ott M. 50 years of protein acetylation: From gene regulation to epigenetics, metabolism and beyond. *Nat. Rev. Mol. Cell Biol.* 2015;16:258–264. doi: 10.1038/nrm3931.
30. Evjenth R., Hole K., Karlsen O.A., Ziegler M., Arnesen T., Lillehaug J.R. Human Naa50p (Nat5/San) displays both protein N alpha- and N epsilon-acetyltransferase activity. *J. Biol. Chem.* 2009;284:31122–31129. doi: 10.1074/jbc.M109.001347.

31. Starheim K.K., Arnesen T., Gromyko D., Ryningen A., Varhaug J.E., Lillehaug J.R. Identification of the human N(alpha)-acetyltransferase complex B (hNatB): A complex important for cell-cycle progression. *Biochem. J.* 2008;415:325–331. doi: 10.1042/BJ20080658.
32. Shaw P.G., Chaerkady R., Zhang Z., Davidson N.E., Pandey A. Monoclonal antibody cocktail as an enrichment tool for acetylome analysis. *Anal Chem.* 2011;83:3623–3626. doi: 10.1021/ac1026176.
33. Gao X., Bao H., Liu L., Zhu W., Zhang L., Yue L. Systematic analysis of lysine acetylome and succinylome reveals the correlation between modification of H2A.X complexes and DNA damage response in breast cancer. *Oncol. Rep.* 2020;43:1819–1830. doi: 10.3892/or.2020.7554.
34. Liu J., Wang Q., Kang Y., Xu S., Pang D. Unconventional protein post-translational modifications: The helmsmen in breast cancer. *Cell Biosci.* 2022;12:22. doi: 10.1186/s13578-022-00756-z.
35. Rey M., Irondelle M., Waharte F., Lizarraga F., Chavrier P. HDAC6 is required for invadopodia activity and invasion by breast tumor cells. *Eur. J. Cell Biol.* 2011;90:128–135. doi: 10.1016/j.ejcb.2010.09.004.
36. Riolo M.T., Cooper Z.A., Holloway M.P., Cheng Y., Bianchi C., Yakirevich E., Ma L., Chin Y.E., Altura R.A. Histone deacetylase 6 (HDAC6) deacetylates survivin for its nuclear export in breast cancer. *J. Biol. Chem.* 2012;287:10885–10893. doi: 10.1074/jbc.M111.308791.
37. Malonia S.K., Yadav B., Sinha S., Lazennec G., Chattopadhyay S. Chromatin remodeling protein SMAR1 regulates NF-κB dependent Interleukin-8 transcription in breast cancer. *Int. J. Biochem. Cell Biol.* 2014;55:220–226. doi: 10.1016/j.biocel.2014.09.008.
38. Chang Y.W., Chen H.A., Tseng C.F., Hong C.C., Ma J.T., Hung M.C., Wu C.H., Huang M.T., Su J.L. De-acetylation and degradation of HSPA5 is critical for E1A metastasis suppression in breast cancer cells. *Oncotarget.* 2014;5:10558–10570. doi: 10.18632/oncotarget.2510.
39. You D., Zhao H., Wang Y., Jiao Y., Lu M., Yan S. Acetylation Enhances the Promoting Role of AIB1 in Breast Cancer Cell Proliferation. *Mol. Cells.* 2016;39:663–668. doi: 10.14348/molcells.2016.2267.
40. Kawai H., Li H., Avraham S., Jiang S., Avraham H.K. Overexpression of histone deacetylase HDAC1 modulates breast cancer progression by negative regulation of estrogen receptor alpha. *Int. J. Cancer.* 2003;107:353–358. doi: 10.1002/ijc.11403.
41. Liu B., Wang T., Wang H., Zhang L., Xu F., Fang R., Li L., Cai X., Wu Y., Zhang W., et al. Oncoprotein HBXIP enhances HOXB13 acetylation and co-activates HOXB13 to confer tamoxifen resistance in breast cancer. *J. Hematol. Oncol.* 2018;11:26. doi: 10.1186/s13045-018-0577-5.
42. Zhao D., Mo Y., Li M.T., Zou S.W., Cheng Z.L., Sun Y.P., Xiong Y., Guan K.L., Lei Q.Y. NOTCH-induced aldehyde dehydrogenase 1A1 deacetylation promotes breast cancer stem cells. *J. Clin. Investig.* 2014;124:5453–5465. doi: 10.1172/JCI76611.
43. Wang H., Holloway M.P., Ma L., Cooper Z.A., Riolo M., Samkari A., Elenitoba-Johnson K.S., Chin Y.E., Altura R.A. Acetylation directs survivin nuclear localization to repress STAT3 oncogenic activity. *J. Biol. Chem.* 2010;285:36129–36137. doi: 10.1074/jbc.M110.152777.

44. Sun Y., Sun J., Lungchukiet P., Quarni W., Yang S., Zhang X., Bai W. Fe65 Suppresses Breast Cancer Cell Migration and Invasion through Tip60 Mediated Cortactin Acetylation. *Sci. Rep.* 2015;5:11529. doi: 10.1038/srep11529.
45. Wiśniewski J.R., Zougman A., Nagaraj N., Mann M. Universal sample preparation method for proteome analysis. *Nat. Methods.* 2009;6:359–362. doi: 10.1038/nmeth.1322.
46. R Core Team . *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing; Vienna, Austria: 2022. [(accessed on 28 March 2023)]. Available online: https://www.R-project.org/
47. Hornbeck P.V., Zhang B., Murray B., Kornhauser J.M., Latham V., Skrzypek E. PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Res.* 2015;43:D512–D520. doi: 10.1093/nar/gku1267.
48. Wu T., Hu E., Xu S., Chen M., Guo P., Dai Z., Feng T., Zhou L., Tang W., Zhan L., et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation.* 2021;2:100141. doi: 10.1016/j.xinn.2021.100141.
49. The gene ontology consortium The Gene Ontology resource: Enriching a GOld mine. *Nucleic Acids Res.* 2021;49:D325–D334. doi: 10.1093/nar/gkaa1113.
50. Carlson M. *Org.Hs.Eg.Db: Genome Wide Annotation for Human.* 2019. R Package Version 3.8.2.
51. Li T., Lu D., Yao C., Li T., Dong H., Li Z., Xu G., Chen J., Zhang H., Yi X., et al. Kansl1 haploinsufficiency impairs autophagosome-lysosome fusion and links autophagic dysfunction with Koolen-de Vries syndrome in mice. *Nat. Commun.* 2022;13:931. doi: 10.1038/s41467-022-28613-0.
52. Kim S.M., Ha E., Kim J., Cho C., Shin S.J., Seo J.H. NAA10 as a New Prognostic Marker for Cancer Progression. *Int. J. Mol. Sci.* 2020;21:8010. doi: 10.3390/ijms21218010.
53. Di Martile M., Del Bufalo D., Trisciuoglio D. The multifaceted role of lysine acetylation in cancer: Prognostic biomarker and therapeutic target. *Oncotarget.* 2016;7:55789–55810. doi: 10.18632/oncotarget.10048.
54. Martínez-Reyes I., Chandel N.S. Cancer metabolism: Looking forward. *Nat. Rev. Cancer.* 2021;21:669–680. doi: 10.1038/s41568-021-00378-6.
55. Li L., Yang L., Fan Z., Xue W., Shen Z., Yuan Y., Sun X., Wang D., Lian J., Wang L., et al. Hypoxia-induced GBE1 expression promotes tumor progression through metabolic reprogramming in lung adenocarcinoma. *Signal Transduct. Target. Ther.* 2020;5:54. doi: 10.1038/s41392-020-0152-8.
56. Phan L.M., Yeung S.C., Lee M.H. Cancer metabolic reprogramming: Importance, main features, and potentials for precise targeted anti-cancer therapies. *Cancer Biol. Med.* 2014;11:1–19. doi: 10.7497/j.issn.2095-3941.2014.01.001.
57. Gwangwa M.V., Joubert A.M., Visagie M.H. Crosstalk between the Warburg effect, redox regulation and autophagy induction in tumourigenesis. *Cell Mol. Biol. Lett.* 2018;23:20. doi: 10.1186/s11658-018-0088-y.
58. Vinaik R., Barayan D., Auger C., Abdullahi A., Jeschke M.G. Regulation of glycolysis and the Warburg effect in wound healing. *JCI Insight.* 2020;5:e138949. doi: 10.1172/jci.insight.138949.
59. Hennipman A., Smits J., van Oirschot B., van Houwelingen J.C., Rijksen G., Neyt J.P., Van Unnik J.A., Staal G.E. Glycolytic enzymes in breast cancer, benign breast disease and normal breast tissue. *Tumour Biol.* 1987;8:251–263. doi: 10.1159/000217529.

60. Li L., Bai Y., Gao Y., Li D., Chen L., Zhou C., Feng M., Chen X., Jin W., Cao Y. Systematic Analysis Uncovers Associations of PGK1 with Prognosis and Immunological Characteristics in Breast Cancer. *Dis. Markers.* 2021;2021:7711151. doi: 10.1155/2021/7711151.
61. Minic Z. Proteomic Studies of the Effects of Different Stress Conditions on Central Carbon Metabolism in Microorganisms. *J. Proteom. Bioinform.* 2015;8:80–90.
62. Martinez-Outschoorn U.E., Peiris-Pagés M., Pestell R.G., Sotgia F., Lisanti M.P. Cancer metabolism: A therapeutic perspective. *Nat. Rev. Clin. Oncol.* 2017;14:11–31. doi: 10.1038/nrclinonc.2016.60.
63. Gerke V., Moss S.E. Annexins: From structure to function. *Physiol. Rev.* 2002;82:331–371. doi: 10.1152/physrev.00030.2001.
64. Miwa N., Uebi T., Kawamura S. S100-annexin complexes—Biology of conditional association. *FEBS J.* 2008;275:4945–4955. doi: 10.1111/j.1742-4658.2008.06653.x.
65. Zhao X.Q., Naka M., Muneyuki M., Tanaka T. Ca^{2+}-dependent inhibition of actin-activated myosin ATPase activity by S100C (S100A11), a novel member of the S100 protein family. *Biochem. Biophys. Res. Commun.* 2000;267:77–79. doi: 10.1006/bbrc.1999.1918.
66. Hayes M.J., Rescher U., Gerke V., Moss S.E. Annexin-actin interactions. *Traffic.* 2004;5:571–576. doi: 10.1111/j.1600-0854.2004.00210.x.
67. Prieto-Fernández L., Menéndez S.T., Otero-Rosales M., Montoro-Jiménez I., Hermida-Prado F., García-Pedrero J.M., Álvarez-Teijeiro S. Pathobiological functions and clinical implications of annexin dysregulation in human cancers. *Front. Cell Dev. Biol.* 2022;10:1009908. doi: 10.3389/fcell.2022.1009908.
68. Mallawaaratchy D.M., Buckland M.E., McDonald K.L., Li C.C., Ly L., Sykes E.K., Christopherson R.I., Kaufman K.L. Membrane proteome analysis of glioblastoma cell invasion. *J. Neuropathol. Exp. Neurol.* 2015;74:425–441. doi: 10.1097/NEN.0000000000000187.
69. Mallawaaratchy D.M., Hallal S., Russell B., Ly L., Ebrahimkhani S., Wei H., Christopherson R.I., Buckland M.E., Kaufman K.L. Comprehensive proteome profiling of glioblastoma-derived extracellular vesicles identifies markers for more aggressive disease. *J. Neurooncol.* 2017;131:233–244. doi: 10.1007/s11060-016-2298-3.
70. Johnstone C.N., Tu Y., Langenbach S., Baloyan D., Pattison A.D., Lock P., Britt K.L., Lehmann B.D., Beilharz T.H., Ernst M., et al. Annexin A1 Is Required for Efficient Tumor Initiation and Cancer Stem Cell Maintenance in a Model of Human Breast Cancer. *Cancers.* 2021;13:1154. doi: 10.3390/cancers13051154.
71. Silva-Oliveira R., Pereira F.F., Petronilho S., Martins A.T., Lameirinhas A., Constâncio V., Caldas-Ribeiro I., Salta S., Lopes P., Antunes L., et al. Clinical Significance of ARID1A and ANXA1 in HER-2 Positive Breast Cancer. *J. Clin. Med.* 2020;9:3911. doi: 10.3390/jcm9123911.
72. Christensen M.V., Høgdall C.K., Jochumsen K.M., Høgdall E.V.S. Annexin A2 and cancer: A systematic review. *Int. J. Oncol.* 2018;52:5–18. doi: 10.3892/ijo.2017.4197.
73. Rodrigo Tapia J.P., Pena Alonso E., García-Pedrero J.M., Florentino Fresno M., Suárez Nieto C., Owen Morgan R., Fernández M.P. Annexin A2 expression in head and neck squamous cell carcinoma. *Acta Otorrinolaringol. Esp.* 2007;58:257–262. doi: 10.1016/S0001-6519(07)74923-3.
74. Yee D.S., Narula N., Ramzy I., Boker J., Ahlering T.E., Skarecky D.W., Ornstein D.K. Reduced annexin II protein expression in high-grade prostatic intraepithelial neoplasia and

prostate cancer. *Arch. Pathol. Lab. Med.* 2007;131:902–908. doi: 10.5858/2007-131-902-RAIPEI.
75. Ding T., Yang L., Wang Y., Yuan J., Chen T., Cai X. Down-regulation of annexin II in prostate cancer is associated with Gleason score, recurrence, metastasis and poor prognosis. *Mol. Med. Rep.* 2010;3:781–787. doi: 10.3892/mmr.2010.332.
76. Smitherman A.B., Mohler J.L., Maygarden S.J., Ornstein D.K. Expression of annexin I, II and VII proteins in androgen stimulated and recurrent prostate cancer. *J. Urol.* 2004;171:916–920. doi: 10.1097/01.ju.0000104674.70170.cd.
77. Gibbs L.D., Mansheim K., Maji S., Nandy R., Lewis C.M., Vishwanatha J.K., Chaudhary P. Clinical Significance of Annexin A2 Expression in Breast Cancer Patients. *Cancers.* 2020;13:2. doi: 10.3390/cancers13010002.
78. Long Y., Chong T., Lyu X., Chen L., Luo X., Faleti O.D., Deng S., Wang F., He M., Qian Z., et al. FOXD1-dependent RalA-ANXA2-Src complex promotes CTC formation in breast cancer. *J. Exp. Clin. Cancer Res.* 2022;41:301. doi: 10.1186/s13046-022-02504-0.
79. Peng B., Guo C., Guan H., Liu S., Sun M.-Z. Annexin A5 as a potential marker in tumors. *Clin. Chim. Acta.* 2014;427:42–48. doi: 10.1016/j.cca.2013.09.048.
80. Sato H., Ogata H., De Luca L.M. Annexin V inhibits the 12-O-tetradecanoylphorbol-13-acetate-induced activation of Ras/extracellular signal-regulated kinase (ERK) signaling pathway upstream of Shc in MCF-7 cells. *Oncogene.* 2000;19:2904–2912. doi: 10.1038/sj.onc.1203615.
81. Korolkova O.Y., Widatalla S.E., Whalen D.S., Nangami G.N., Abimbola A., Williams S.D., Beasley H.K., Reisenbichler E., Washington M.K., Ochieng J., et al. Reciprocal expression of Annexin A6 and RasGRF2 discriminates rapidly growing from invasive triple negative breast cancer subsets. *PLoS ONE.* 2020;15:e0231711. doi: 10.1371/journal.pone.0231711.
82. Cluntun A.A., Huang H., Dai L., Liu X., Zhao Y., Locasale J.W. The rate of glycolysis quantitatively mediates specific histone acetylation sites. *Cancer Metab.* 2015;3:10. doi: 10.1186/s40170-015-0135-3.
83. Elsheikh S.E., Green A.R., Rakha E.A., Powe D.G., Ahmed R.A., Collins H.M., Soria D., Garibaldi J.M., Paish C.E., Ammar A.A., et al. Global histone modifications in breast cancer correlate with tumor phenotypes, prognostic factors, and patient outcome. *Cancer Res.* 2009;69:3802–3809. doi: 10.1158/0008-5472.CAN-08-3907.
84. Abell A.N., Jordan N.V., Huang W., Prat A., Midland A.A., Johnson N.L., Granger D.A., Mieczkowski P.A., Perou C.M., Gomez S.M., et al. MAP3K4/CBP-regulated H2B acetylation controls epithelial-mesenchymal transition in trophoblast stem cells. *Cell Stem Cell.* 2011;8:525–537. doi: 10.1016/j.stem.2011.03.008.

3.4 Proteomics Approaches for the Discovery of Potential Enzymatic Biomarkers for Early Diagnosis of Breast Cancer

3.4.1 Abstract

Breast cancer (BC) is one of the leading causes of death in Canadian women, with an average survival rate of 5 years after diagnosis. Early detection of BC can greatly improve patient outcomes and survival. However, a non-invasive BC detection method is not currently available in clinics. Recent studies suggest that proteins from small extracellular vesicles (sEVs) could be promising biomarkers for non-invasive BC early-stage diagnosis. sEVs are membrane-enclosed vesicles secreted by cells that drive different stages of carcinogenesis in BC. The purpose of this work was to analyze different published proteomics datasets to identify enzymes that could be potentially used as diagnostic biomarkers. Three cell line studies were compared, and overlapping BC proteins were highlighted with proteins found in sEVs from blood and plasma. In total, 106 proteins were selected based on the cell line studies, of which 40 have been identified in blood/plasma sEVs. These 106 proteins were mostly enriched with cell–cell signaling and DNA repair terms based on GO analysis. Furthermore, these 40 proteins contained 11 enzymes that can be explored as potential BC biomarkers. Future validation of enzymes using cancer cell lines and blood from BC patients remains to be determined.

Keywords: breast cancer; biomarkers; enzymes; proteomics; extracellular vesicles

3.4.2 Introduction

Extracellular vesicles (EVs) are small, membrane-bound particles that are released by most cells in the body. They contain a variety of molecules, such as proteins, metabolites, and nucleic acids, and are thought to play important roles in cell–cell communication [1]. EVs can be classified into several subtypes based on their size, biogenesis, and cargo, including small extracellular vesicles (sEVs), microvesicles, and apoptotic bodies [1]. Research has shown that EVs can be involved in a range of physiological and pathological processes, such as immune regulation, tumor progression, and neurodegenerative diseases. EVs also have potential as diagnostic and therapeutic tools, as they can be isolated from biological fluids such as blood and urine, and their cargo can be manipulated for targeted delivery of therapeutic agents [2].

sEVs have emerged as promising biomarkers for cancer due to their ability to reflect the molecular signature of their cell of origin [3-5]. Cancer cells release sEVs that contain a variety of molecules, such as proteins, which can be detected in blood, urine, or other types of bodily fluids. By analyzing the proteins of these sEVs, specific biomarkers associated with cancer development, progression, and treatment can be identified. Breast cancer (BC) is one of the leading causes of death in women [3]. Early detection of BC can greatly improve patient outcomes and survival. However, a non-invasive BC detection method is not available in clinics yet [4]. Recent studies suggest that proteins in EVs could be promising biomarkers for non-invasive BC in early-stage diagnosis [5].

Various studies have been performed with the aim of discovering potential BC biomarkers in EVs, and MS-based proteomic analysis is one of the main approaches. For example, Risha et al. isolated EVs from a highly metastatic BC cell line and MCF-10A, a non-cancerous epithelial breast cell line [6]. The proteomic analysis of the isolated sEVs revealed 726 proteins unique to the BC cell line [6]. Minic et al. isolated sEVs and enriched phosphopetides and performed phosphoproteomic analysis using LC-MS/MS to expand the known proteome of such sEVs. The profiling of the phosphoproteome resulted in the identification of 2003 phosphopeptides that were mapped onto 855 different proteins, encompassing a broad range of functionalities [5]. Rontogianni et al. isolated EVs from nine subtypes of BC cell lines and MCF-10A, which were used to perform both quantitative proteomic analysis and phosphoproteomic analysis. Their results reveal that EVs are subtype-specific in BC cell lines, which indicates EVs could potentially play an important role in BC subtyping in clinical diagnostics [1]. In addition to BC cell lines, researchers also isolated EVs from human plasma and serum to perform proteomic analysis. For example, Muraoka et al. isolated EVs via the affinity capture isolation method from EDTA plasma and serum and successfully identified a total of 4079 proteins by quantitative proteomics [7]. This presents the deepest proteomic study of plasma and blood EVs besides the medium EV (mEV) study by Kverneland et al. [8].

Enzyme biomarkers can be very efficient for early detection, diagnosis, therapeutic treatment, and monitoring disease recurrence of cancer patients [9]. Although enzyme biomarkers can be efficiently used for precise measurement of cancer progression, a limited number of clinically approved cancer biomarkers are available for early diagnosis [9]. Advancement of proteomic technologies enables the identification of potential biomarkers using different human cancer cell lines and blood.

The purpose of this work was to find potential BC-associated proteins identified from cell line sEV studies that can be isolated from blood or plasma sEVs and may serve as diagnostic biomarkers.

3.4.3 Methods

3.4.3.1. Data Source

The data were obtained from the supplemental material from the studies presented in Table 3.4.1.

Table 3.4.1. Published BC proteomics studies used for data analysis.

Publication	Sample Types	Breast Cancer Subtypes	Control
Risha et al. [6]	BC cell lines	MDA-MB-231	MCF-10A
Minic et al. [5]	BC cell lines	MDA-MB-231, MCF-7	MCF-10A
Rontogianni et al. [1]	BC cell lines	MCF-7, Hs578T, BT549, MDA-MB-231, LM2, HCC1954, HCC1419, JIMT1, SKBR3	MCF-10A
Muraoka et al. [7]	Plasma and serum	Healthy human subjects	

3.4.3.2. Bioinformatics Analysis

The protein groups identified from all datasets were compared based on the first protein accession ID. Contaminant and reverse proteins were removed if still present [5]. In the cell line studies, proteins identified in the control MCF10A cell line, which was used in all datasets, were pooled and subtracted from all proteins identified in more than 50% of replicates for any given BC cell line. When only quantitative data were available [1], proteins were considered to not be identified in MCF10A if their log2 LFQ intensity was one SD below the average with a significant ANOVA test [1]. Then, only proteins identified in BC cell lines were compared between the datasets and only selected if present in two out of three studies. The resulting proteins were compared with proteins identified in blood or plasma.

Gene Ontology enrichment analysis was performed with the clusterProfiler R package using the 106 overlapping proteins against all identified proteins in cell lines. Enzyme commissions were retrieved from the org.Hs.eg.db R annotation package.

3.4.4 Results and Discussion

Usually, proteins are first identified in both non-cancerous and cancerous cell lines, and cancer-specific proteins are selected as potential biomarkers. Then, those biomarker candidates are matched in serum/plasma samples, which are among the most accessible biological samples from patients. EVs can be detected in different types of biofluids, such as blood and urine, making them an ideal subject for our study on the use of EVs for the early diagnosis of BC. In this work, we compared the proteins obtained from different studies on EVs derived from human BC cell lines. Three of them, Risha et al. [6], Minic et al. [5], and Rontogianni et al. [1], used EVs isolated from breast cancer cell lines, and Muraoka et al. [7] used EVs isolated from blood and plasma.

The three cell lines studies identified a total of 5309 proteins from both the non-cancerous epithelial breast cell line and various breast cancer cell lines reflecting different BC subtypes. Proteins that were identified in MCF-10A were removed from the total identified proteins, resulting in 831 proteins unique to the BC cell lines. This led to the identification of 106 proteins present in at least two out of the three cell line studies (Figure 3.4.1 and Table S1).

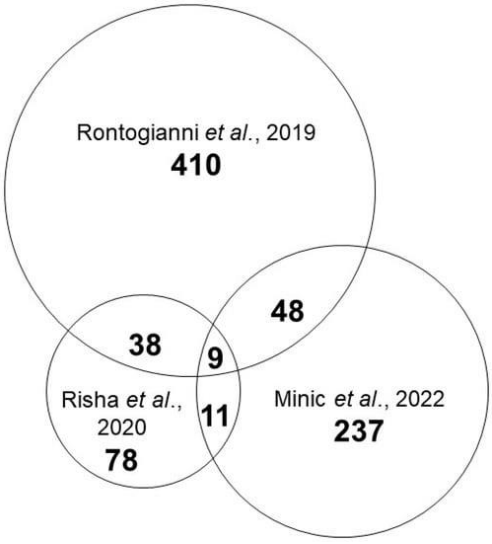

Figure 3.4.1. The overlap between identified sEV proteomes from three breast cancer cell line studies [1,5,6]. In total, 106 proteins were identified in at least two out of the three studies.

These 106 proteins were then compared to the proteomic data from the blood and plasma study. In total, 40 proteins overlapped, from which 11 enzymes were identified as enzymes by their enzyme commission number (Table 3.4.2). Four of these enzymes belong to the family of protein kinases: unc-51-like kinase 3, protein kinase C beta, G-protein-coupled receptor kinase 2, and glycogen synthase kinase 3 alpha. Several reports found that targeted inhibition of unc-51-like kinase 3 contributed to the inhibition of metastasis and tumor growth in a diverse range of cancers [10-12]. It has been found that the expression of protein kinase C beta (PKCβ) and G-protein-coupled receptor kinase 2 promotes tumorigenesis in BC [13,14]. Using MCF7 cell lines, it has been demonstrated that glycogen synthase kinase-3 protects estrogen receptor alpha from proteasomal degradation and is required for full transcriptional activity of the receptor [15]. This estrogen receptor is known to play a significant role in the formation of BC [16,17]. Two enzymes are associated with ligase in the ubiquitin pathway: thyroid hormone receptor interactor 12 and WW domain containing E3 ubiquitin protein ligase 2. Ubiquitination-related proteins (URGs) have been proposed as important biomarkers and therapeutic targets in cancer, including BC [18]. The WW domain containing E3 ubiquitin protein ligase 2 has been proposed to play a central role in tumourigenesis and has potential as a prognostic marker and molecular therapeutic target [19]. Other identified enzymes include galactosidase alpha, dipeptidyl peptidase 4, peptidylprolyl cis/trans isomerase—NIMA-interacting 1, acyl-CoA synthetase short-chain family member 2, and CTP synthase 2. All these enzymes have been reported to have different roles in cancer biology [20-24].

Table 3.4.2. List of enzymes identified in EVs from blood and plasma suggested as potential BC biomarkers. Eleven enzymes were identified in sEVs from both BC cell lines and blood and plasma. T indicates that the enzymes were identified in the proteomic data from the published BC cell lines studies.

UniProt ID	Enzyme	Gene	Name	Ref. 6	Ref. 5	Ref. 1
Q14669	6.3.2.19	TRIP12	Thyroid hormone receptor interactor 12	T	T	T
Q6PHR2	2.7.11.1	ULK3	Unc-51-like kinase 3	T		T

UniProt ID	Enzyme	Gene	Name	Ref. 6	Ref. 5	Ref. 1
P05771	2.7.11.13	PRKCB	Protein kinase C beta	T		T
P25098	2.7.11.15	GRK2	G-protein-coupled receptor kinase 2		T	T
P49840	2.7.11.26	GSK3A	Glycogen synthase kinase 3 alpha		T	T
P06280	3.2.1.22	GLA	Galactosidase alpha	T	T	
P27487	3.4.14.5	DPP4	Dipeptidyl peptidase 4		T	T
Q13526	5.2.1.8	PIN1	Peptidylprolyl cis/trans isomerase, NIMA-interacting 1	T		T
Q9NR19	6.2.1.1	ACSS2	Acyl-CoA synthetase short-chain family member 2		T	T
O00308	6.3.2.19	WWP2	WW domain containing E3 ubiquitin protein ligase 2		T	T
Q9NRF8	6.3.4.2	CTPS2	CTP synthase 2		T	T

The functional enrichment analysis of the selected 106 proteins (Table S2) revealed many enriched terms, including cell–cell signaling (Table S3) and DNA repair. EVs are known to play an important role in mediating cell–cell signaling, immune response, and tumor metastasis [25]. Additionally, altered DNA repair systems are often related to an increase in the cancer rate [26]. Our findings suggest that EVs play a role in DNA repair during cancer progression.

3.4.5 Conclusions

In conclusion, the results of this study and previously reported data strongly suggest that 11 identified enzymes may be potential candidates as biomarkers for the early diagnosis of BC. In our previous report [5], we also identified and validated some enzymatic biomarkers from sEVs derived

from BC cell lines. All these candidates, including those reported here, provide a feasible panel of potential biomarkers that can be tested using plasma/blood from BC patients.

3.4.6 Supplementary Materials

The following supporting information can be downloaded at: **https://www.mdpi.com/article/10.3390/ECB2023-14099/s1**. Table S1: the overlap between identified sEV proteomes from three breast cancer cell lines studies identified in at least two out of the three studies; Table S2: gene ontology enrichment result using Biological Process terms for 106 overlapping BC proteins vs. all identified cell line proteins. Table S3: proteins that are related to cell–cell signaling function according to Gene Ontology analysis.

3.4.7 Funding

The study was supported by the John L. Holmes Mass Spectrometry Facility for the sample preparation and MS analysis and granted by Natural Sciences and Engineering Research Council of Canada (M.V.B.: grant # RGPIN-2020-05775).

3.4.8 References

1. Rontogianni, S.; Synadaki, E.; Li, B.; Liefaard, M.C.; Lips, E.H.; Wesseling, J.; Wu, W.; Altelaar, M. Proteomic profiling of extracellular vesicles allows for human breast cancer subtyping. *Commun. Biol.* **2019**, *2*, 325.
2. Marar, C.; Starich, B.; Wirtz, D. Extracellular vesicles in immunomodulation and tumor progression. *Nat. Immunol.* **2021**, *22*, 560–570.
3. Jordan, K.R.; Hall, J.K.; Schedin, T.; Borakove, M.; Xian, J.J.; Dzieciatkowska, M.; Lyons, T.R.; Schedin, P.; Hansen, K.C.; Borges, V.F. Extracellular vesicles from young women's breast cancer patients drive increased invasion of non-malignant cells via the Focal Adhesion Kinase pathway: A proteomic approach. *Breast Cancer Res.* **2020**, *22*, 128.
4. Li, J.; Guan, X.; Fan, Z.; Ching, L.M.; Li, Y.; Wang, X.; Cao, W.M.; Liu, D.X. Non-Invasive Biomarkers for Early Detection of Breast Cancer. *Cancers* **2020**, *12*, 2767.
5. Minic, Z.; Hüttmann, N.; Poolsup, S.; Li, Y.; Susevski, V.; Zaripov, E.; Berezovski, M.V. Phosphoproteomic Analysis of Breast Cancer-Derived Small Extracellular Vesicles Reveals Disease-Specific Phosphorylated Enzymes. *Biomedicines* **2022**, *10*, 408.
6. Risha, Y.; Minic, Z.; Ghobadloo, S.M.; Berezovski, M.V. The proteomic analysis of breast cell line exosomes reveals disease patterns and potential biomarkers. *Sci. Rep.* **2020**, *10*, 13572.
7. Muraoka, S.; Hirano, M.; Isoyama, J.; Nagayama, S.; Tomonaga, T.; Adachi, J. Comprehensive proteomic profiling of plasma and serum phosphatidylserine-positive extracellular vesicles reveals tissue-specific proteins. *iScience* **2022**, *25*, 104012.

8. Kverneland, A.H.; Østergaard, O.; Emdal, K.B.; Svane, I.M.; Olsen, J.V. Differential ultracentrifugation enables deep plasma proteomics through enrichment of extracellular vesicles. *Proteomics* **2022**, *23*, e2200039.
9. Liang, S.-L.; Chan, D.W. Enzymes and related proteins as cancer biomarkers: A proteomic approach. *Clin. Chim. Acta* **2007**, *381*, 93–97.
10. Dower, C.M.; Bhat, N.; Gebru, M.T.; Chen, L.; Wills, C.A.; Miller, B.A.; Wang, H.G. Targeted Inhibition of ULK1 Promotes Apoptosis and Suppresses Tumor Growth and Metastasis in Neuroblastoma. *Mol. Cancer* **2018**, *17*, 2365–2376.
11. Hwang, D.Y.; Eom, J.I.; Jang, J.E.; Jeung, H.K.; Chung, H.; Kim, J.S.; Cheong, J.W.; Min, Y.H. ULK1 inhibition as a targeted therapeutic strategy for FLT3-ITD-mutated acute myeloid leukemia. *J. Exp. Clin. Cancer Res.* **2020**, *39*, 85.
12. Tang, F.; Hu, P.; Yang, Z.; Xue, C.; Gong, J.; Sun, S.; Shi, L.; Zhang, S.; Li, Z.; Yang, C.; et al. SBI0206965, a novel inhibitor of Ulk1, suppresses non-small cell lung cancer cell growth by modulating both autophagy and apoptosis pathways. *Oncol. Rep.* **2017**, *37*, 3449–3458.
13. Wallace, J.A.; Pitarresi, J.R.; Sharma, N.; Palettas, M.; Cuitiño, M.C.; Sizemore, S.T.; Yu, L.; Sanderlin, A.; Rosol, T.J.; Mehta, K.D.; et al. Protein Kinase C Beta in the Tumor Microenvironment Promotes Mammary Tumorigenesis. *Front. Oncol.* **2014**, *4*, 87.
14. Nogués, L.; Reglero, C.; Rivas, V.; Salcedo, A.; Lafarga, V.; Neves, M.; Ramos, P.; Mendiola, M.; Berjón, A.; Stamatakis, K.; et al. G Protein-coupled Receptor Kinase 2 (GRK2) Promotes Breast Tumorigenesis Through a HDAC6-Pin1 Axis. *eBioMedicine* **2016**, *13*, 132–145.
15. Grisouard, J.; Medunjanin, S.; Hermani, A.; Shukla, A.; Mayer, D. Glycogen Synthase Kinase-3 Protects Estrogen Receptor α from Proteasomal Degradation and Is Required for Full Transcriptional Activity of the Receptor. *Mol. Endocrinol.* **2007**, *21*, 2427–2439.
16. Nilsson, S.; Gustafsson, J.A. Biological role of estrogen and estrogen receptors. *Crit. Rev. Biochem. Mol. Biol.* **2002**, *37*, 1–28.
17. Couse, J.F.; Korach, K.S. Estrogen receptor null mice: What have we learned and where will they lead us? *Endocr. Rev.* **1999**, *20*, 358–417.
18. Zhao, K.; Zheng, Y.; Lu, W.; Chen, B. Identification of ubiquitination-related gene classification and a novel ubiquitination-related gene signature for patients with triple-negative breast cancer. *Front. Genet.* **2023**, *13*, 932027.
19. Soond, S.M.; Smith, P.G.; Wahl, L.; Swingler, T.E.; Clark, I.M.; Hemmings, A.M.; Chantry, A. Novel WWP2 ubiquitin ligase isoforms as potential prognostic markers and molecular targets in cancer. *Biochim. Biophys. Acta (BBA)—Mol. Basis Dis.* **2013**, *1832*, 2127–2135.
20. Tobler, J.E.; Jacobson, K.B. Multiple Forms of Alpha-Galactosidase of the Mouse and Their Use as a Cell Marker in Tumorigenesis2. *JNCI J. Natl. Cancer Inst.* **1978**, *61*, 1263–1268.
21. Mezawa, Y.; Daigo, Y.; Takano, A.; Miyagi, Y.; Yokose, T.; Yamashita, T.; Morimoto, C.; Hino, O.; Orimo, A. CD26 expression is attenuated by TGF-β and SDF-1 autocrine signaling on stromal myofibroblasts in human breast cancers. *Cancer Med.* **2019**, *8*, 3936–3948.
22. Liu, C.; Mu, C.; Li, Z.; Xu, L. Imazamethabenz inhibits human breast cancer cell proliferation, migration and invasion via combination with Pin1. *Mol. Med. Rep.* **2017**, *15*, 3210–3214.

23. Liu, M.; Liu, N.; Wang, J.; Fu, S.; Wang, X.; Chen, D. Acetyl-CoA Synthetase 2 as a Therapeutic Target in Tumor Metabolism. *Cancers* **2022**, *14*, 2896.
24. Hu, X.; Han, Y.; Liu, J.; Wang, H.; Tian, Z.; Zhang, X.; Zhang, Y.; Wang, X. CTP synthase 2 predicts inferior survival and mediates DNA damage response via interacting with BRCA1 in chronic lymphocytic leukemia. *Exp. Hematol. Oncol.* **2023**, *12*, 6.
25. Heydari, Z.; Peshkova, M.; Gonen, Z.B.; Coretchi, I.; Eken, A.; Yay, A.H.; Dogan, M.E.; Gokce, N.; Akalin, H.; Kosheleva, N.; et al. EVs vs. EVs: MSCs and Tregs as a source of invisible possibilities. *J. Mol. Med.* **2022**, *101*, 51–63.
26. Kitagishi, Y.; Kobayashi, M.; Matsuda, S. Defective DNA repair systems and the development of breast and prostate cancer (Review). *Int. J. Oncol.* **2013**, *42*, 29–34.

3.5 Conclusions

In this chapter, different proteomic approaches were used to analyze proteins from breast cancer-derived extracellular vesicles. 855 distinct phosphoproteins and 60 distinct acetylated proteins were identified in BC cell lines. Nine of these proteins were selected for validation. Among them, three phosphorylated enzymes (ACLY, SIRT1, and SIRT6) and three acetylated enzymes (ALDOA, PGK1 and ENO) showed a significantly higher specific enzymatic activity in BC-derived sEVs in comparison to non-cancerous. Our quantitative proteomics analyses and previously reported data strongly suggest that 11 identified enzymes may be potential candidates as biomarkers for the diagnosis of BC. Future validation of these enzymes using both cancer cell lines and blood from BC patients remains to be determined.

Chapter 4: General Conclusion: Enzymatic Activity Assays for the Validation of Identified Proteomic Biomarker

An enzymatic assay is a laboratory technique used to measure the activity of enzymes in a sample, providing insight into their function, regulation, and kinetic properties [1]. These assays are widely utilized in biochemical research, drug development, and clinical diagnostics [2,3]. In this thesis, enzymatic assays were used as the validation method for studies in both Chapter 2 and Chapter 3.

Enzymatic assays provide several advantages over Western blot when studying non-model organisms like wood frogs, particularly because they measure enzyme activity directly without the need for species reagents. One of the primary challenges in working with non-model organisms is the lack of readily available antibodies tailored to their proteins, which is essential for Western blotting [4]. Enzymatic assays bypass this limitation by relying on the universal principles of enzyme-substrate interaction, which are often conserved across species. This flexibility allows researchers to study organisms where specific antibodies for protein detection are unavailable or costly to develop. Enzymatic assays can also be easily adapted for different species with minimal modifications, making them especially valuable in comparative studies of biodiversity or in fields like evolutionary biology and ecology.

Another key advantage of enzymatic assays is their ability to provide functional data based on enzyme activity [5], which is often more insightful than just detecting the presence of protein. Western blot can confirm the expression of an enzyme, but it doesn't reveal whether the enzyme is active or functional under given conditions. Enzymatic assays, however, directly measure the catalytic activity of the enzyme, allowing researchers to quantify how well an enzyme is functioning in real-time [5]. This is particularly important in the study of non-model organisms, where researchers may be interested in understanding how unique metabolic pathways or adaptive mechanisms operate in environments or under stresses that differ from those of well-studied species. Enzyme activity measurements can provide critical insights into the physiological and ecological adaptations of these organisms.

In addition to functional insights, enzymatic assays are often more straightforward and cost-effective to implement compared to Western blot, particularly for high-throughput applications. Western blot requires complex procedures such as protein extraction, gel electrophoresis, transfer to membranes, and antibody incubation, all of which can be time-consuming and require specialized reagents that may not be readily available for non-model species [6]. Enzymatic assays, on the other hand, typically involve simpler workflows, such as adding a substrate to a sample and measuring the product formation, often using a colorimetric or fluorometric readout [7]. This simplicity and scalability make enzymatic assays more practical. These factors together make enzymatic assays a powerful tool for studying non-model organisms, offering both practical and scientific advantages over traditional validation methods like Western blot.

Despite their many advantages, enzymatic assays come with several limitations that researchers must consider. One major limitation is that enzymatic assays are highly sensitive to experimental conditions such as pH, temperature, and ionic strength [7]. Since enzyme activity can fluctuate based on these factors, even small deviations from optimal conditions can significantly affect the accuracy and reproducibility of the results [7]. This makes it necessary to carefully control the assay environment, which may not always be feasible when working with complex or heterogeneous biological samples, such as tissue extracts from non-model organisms. In studies involving environmental stressors (e.g., freeze-thaw cycles in wood frogs), maintaining appropriate conditions for the assay can be particularly challenging.

Another limitation is that enzymatic assays often require purified or relatively simple samples, as the presence of other enzymes, proteins, or small molecules can interfere with the results. Complex biological samples may contain inhibitors, activators, or other compounds that can alter enzyme activity, leading to false positives or negatives [8]. For instance, in whole tissue homogenates, it might be difficult to isolate the specific enzyme of interest without interference from other cellular components. This can necessitate additional sample preparation steps, such as partial purification or dialysis [8].

In this thesis, we performed the proteomics analysis of protein phosphorylation and acetylation from EVs derived of control and cancerous BC cell lines. These PTMs of proteins are major regulatory mechanisms of cell signaling networks, thus essential for almost all cellular functions. Therefore, protein phosphorylation and acetylation play crucial roles in the regulation of protein functions. Enzymatic assays are predominately used to measure the functional

properties of proteins. Enzyme based biomarkers have been used for disease diagnosis for many decades. Including. In this thesis, we identified several potential enzyme-based biomarkers for BC.

Finally, enzymatic assays may struggle to detect enzymes that have low activity or are present at very low concentrations. While the assay measures functional activity, enzymes that are not highly expressed or only active under specific conditions might not generate detectable levels of product, leading to underestimation or failure to detect the enzyme's presence. Advanced detection techniques, such as fluorescence-based methods, can sometimes overcome this limitation [9], but these approaches may require expensive equipment or specialized reagents, making them less accessible to all laboratories. As a result, enzymatic assays can sometimes lack the sensitivity needed for detecting low-abundance enzymes. In conclusion, enzymatic assays are powerful tools to validate results from the proteomic fields, offering precise and quantitative insights into enzyme behavior. However, their effectiveness is often contingent upon carefully controlled experimental conditions, appropriate sample preparation, and consideration of potential interferences.

4.1 References – Chapter 4

1. Bisswanger, H. Enzyme assays. *Perspectives in Science* **1**, 41-55, doi:https://doi.org/10.1016/j.pisc.2014.02.005 (2014).
2. Rufer, A. C. Drug discovery for enzymes. *Drug Discovery Today* **26**, 875-886, doi:https://doi.org/10.1016/j.drudis.2021.01.006 (2021).
3. Danese, E. et al. Analytical evaluation of three enzymatic assays for measuring total bile acids in plasma using a fully-automated clinical chemistry platform. *PLoS One* **12**, e0179200, doi:10.1371/journal.pone.0179200 (2017).
4. Baker, M. Reproducibility crisis: Blame it on the antibodies. *Nature* **521**, 274-276, doi:10.1038/521274a (2015).
5. Harris, T. K. & Keshwani, M. M. Measurement of enzyme activity. *Methods Enzymol* **463**, 57-71, doi:10.1016/s0076-6879(09)63007-x (2009).
6. Mahmood, T. & Yang, P. C. Western blot: technique, theory, and trouble shooting. *N Am J Med Sci* **4**, 429-434, doi:10.4103/1947-2714.100998 (2012).
7. Onyeogaziri, F. C. & Papaneophytou, C. A General Guide for the Optimization of Enzyme Assay Conditions Using the Design of Experiments Approach. *SLAS Discov* **24**, 587-596, doi:10.1177/2472555219830084 (2019).
8. Kruger, N. J. Errors and artifacts in coupled spectrophotometric assays of enzyme activity. *Phytochemistry* **38**, 1065-1071, doi:10.1016/0031-9422(94)00787-t (1995).
9. Kovar, P. et al. Development of a sensitive high-throughput enzymatic assay capable of measuring sub-nanomolar inhibitors of SARS-CoV2 Mpro. *SLAS Discov* **29**, 100179, doi:10.1016/j.slasd.2024.100179 (2024).