Surface Proteome of Extracellular Vesicles and Correlation Analysis for Identification of Breast Cancer Biomarkers

Nico Hüttmann

A thesis submitted 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

Breast cancer (BC) is the second leading cause of death in Canadian women. Detection of the disease at an early stage greatly increases the average 5-year survival rate, however non-invasive early detection methods are not available to-date. Cells release various types of extracellular vesicles (EVs) to mediate intercellular communication by transferring signals in the form of bioactive molecules such as proteins, metabolites, and nucleic acids. Understanding the composition of these biomolecules may shed light on the physiological state of the cell of origin. Therefore, EVs are a promising source of biomarkers for non-invasive detection of BC. However, the surface proteome of EVs is not yet understood well enough to propose BC biomarkers that could be detected directly from biofluids. In this study, small EVs (sEVs) and medium EVs (mEVs) were isolated by differential ultracentrifugation from breast cancer MDA-MB-231 and MCF7, and non-cancerous breast epithelial MCF10A cell lines and analyzed by nano-liquid chromatography coupled to tandem mass spectrometry. EV proteins were analyzed by two approaches: (1) global proteomic analysis and (2) enrichment of EV surface proteins by labelling surface-accessible proteins with a Sulfo-NHS-SS-Biotin reagent. Potential BC biomarkers were obtained from

the first approach (1) by identifying the presence of cell line specific sEV proteins, filtering for membrane/surface proteins using UniProt annotations, and predicting the co-localization of proteins on sEVs with known EV marker proteins (CD63, CD9, CD81) by correlation analysis. This resulted in 11 potential BC sEV biomarkers (C8A, AXL, ST14, FAM20B, PROM2, CLDN3, ITGA7, MEGF10, SHISA2, GJC1, IFNGR1); the presence of ST14, CLDN3 and ITGA7 was validated by Western blot analysis. The surface labelling approach (2) enriched proteins previously not identified using the first approach (1). Potential general BC biomarkers were selected from surface proteins commonly identified from MDA-MB-231 and MCF7, but not identified in MCF10A EVs. Annotation with known BC disease associations from DisGeNET yielded 9 and 2 potential surface proteins on sEVs and mEVs, respectively. This study demonstrates the emerging role of EVs as a rich source of known and novel biomarkers which may be used for non-invasive detection of BC.

Preface

This thesis is an original intellectual product of Nico Huttmann. No part of this thesis has been previously published.

Nico Huttmann designed the study, conducted the experiments, wrote the manuscript and analyzed the data.

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. John Pezacki provided access to the chemiluminescence imaging system.

Dr. Derrick Gibbings provided access to the NTA instrument.

Acknowledgements

I sincerely thank my supervisor, Dr. Maxim Berezovski, for his guidance, insight, and support. I am very grateful for all opportunities, challenges, and the scientific spirit for discoveries that was presented to me over the course of my Master program.

I would like to thank Dr. Zoran Minic for his support, trust, and the challenging attitude that formed me over the time.

I wish to thank Prof. Michael Przybylski for his support over the time of my Master program.

I would like to thank the examiners Dr. Derrick Gibbings, Dr. Mathieu Lavallee-Adam and Dr. John Pezacki for critical correction and fruitful additions to this thesis.

The staff of the University of Ottawa create supportive and flourishing environment for students. I would like to express my gratitude to the Science store and the administrative staff for all help and a good mood at all times.

To my *colleagues and friends*, each of you has played a special role in the past years, inside and outside the lab. Emil Zaripov, Vanessa Susevski, Zoran Minic, Yuchu Dou, Yaroslav Grechkin, Suttinee Poolsup, Lixuan Ren and Cici Li. I am looking forward to seeing you and your successes in the future. Also, I would like to thank Tiah Lee, Zoran Bert-Caszo and Tony Ohu for the time we spent together.

I finally thank my parents, Bettina and Michael Huttmann, as well as my brother Malte for their continuous support throughout the years of my Master program and beyond.

Table of Contents

List of Abbreviations

BC	Breast cancer
DDA	Data dependent acquisition
DDM	n-Dodecyl-B-D-Maltoside
DNA	Deoxyribonucleic acid
dUC	Differential ultracentrifugation
ELISA	Enzyme-linked immunoassay
ESCRT	Endosomal sorting complex required for transport
ESI	Electrospray ionization
HPLC	High-performance liquid chromatography
ILVs	Intraluminal vesicles
LC-MS/MS	Liquid chromatography coupled to tandem mass spectrometry
lEVs	large extracellular vesicles
mEVs	medium extracellular vesicles
MS	Mass spectrometry
MVBs	Multivesiclular bodies
MVs	Microvesicles

NHS	N-Hydroxysuccinimide
PAGE	Polyacrylamide gel electrophoresis
PPI	Protein-protein interaction
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulfate
sEVs	Small extracellular vesicles
TEM	Transmission electron microscopy
TFF	Tangential flow filtration
UF	Ultrafiltration
WGCNA	Weighted gene co-expression network analysis

List of Figures

List of Tables

Chapter 1

Introduction

Early detection of most diseases significantly improves the outcome for patients and may prevent worsening of symptoms (DeSantis et al. 2014). For this reason, screening and diagnostic methods are continuously developed to achieve better sensitivity and specificity, and minimize the degree of invasiveness lessening the burden to obtain clinical samples from patients (Vlastos and Verkooijen 2007). Breast cancer (BC) is an example of an illness that would greatly benefit from the development of novel and non-invasive early detection methods (DeSantis et al. 2014). Although many chemo-, radio-, hormone-, and immunotherapeutic options are available, (Mustacchi et al. 2015; Sonnenblick and Piccart 2015; Tangutoori, Baldwin, and Sridhar 2015), one third of women will die from the disease (Siegel, Miller, and Jemal 2015 Jan-Feb). Extracellular vesicles (EVs) have been hypothesized to play significant roles in BC and have been explored as a novel source of biomarkers for early detection of BC (Green et al. 2015).

This chapter provides an overview of EVs and their involvement in BC. Furthermore, it discusses the role of mass spectrometry (MS)-based proteomics as the primary analysis method to investigate the proteome and surface proteome of cell line-derived EVs. This chapter also serves to highlight the tremendous growth over the past two decades of our understanding of EVs and the field of MS-based proteomics, which roughly defines the scope of the following introduction.

1.1 Extracellular Vesicles

EVs is a collective term for phospholipid bilayer-enclosed structures that are secreted into the extracellular environment by essentially all cell types, in all domains of life (Doyle and Wang 2019; Gill, Catchpole, and Forterre 2019). Currently, EVs are understood to play a role in communication, between cells that are in close proximity or between distant cells (Maia et al. 2018). EVs were first described in 1967 as phospholipid-rich "platelet dust," separated from fresh platelet-free plasma using ultracentrifugation (Wolf 1967). The procoagulant activity attributed to EVs was notably distinct from the platelets or the plasma themselves. Today, EVs are classified into three main subtypes based on their intracellular origin and secretion mechanism: exosomes, microvesicles (MVs), and apoptotic bodies (Margolis and Sadovsky 2019). The term microvesicles was first coined in 1971 and it was used to describe membrane blebs that were released from maturing reticulocytes and peripheral blood platelets, initially thought of as a mechanism to eliminate excess cell membrane and surface receptors (Crawford 1971; Schrier et al. 1971; C. V. Harding, Heuser, and Stahl 2013). Similarly, ~50 nM sized vesicles associated with the transferrin receptor from maturing blood reticulocytes were referred to as exosomes in 1983 (C. V. Harding, Heuser, and Stahl 2013; C. Harding and Stahl 1983; Pan and Johnstone 1983). Although apoptosis and the formation of apoptotic bodies was partially observed and described many years prior, apoptotic bodies were formally discovered in 1972 as part of a regulated mechanism of cell death in multicellular organisms (Elmore 2007; Kerr, Wyllie, and Currie 1972).

Over the last two decades, most research within the field of EVs has focused on non-apoptotic EV subtypes exosomes and MVs, due to their broad involvement in physiological and pathophysiological processes (Ohno, Ishikawa, and Kuroda 2013). Moreover, EVs have been found in many biofluids such as urine (I. O. Sun and Lerman 2020), saliva (Chiabotto et al. 2019), blood (Arraud et al. 2014), and cerebral fluid (Guha et al. 2019), making them an attractive source of biomarkers to detect or monitor the progression of a disease in a non-invasive manner (Tian et al. 2021).

1.1.1 Characteristics of Extracellular Vesicles

By definition, all EVs share the presence of a phospholipid-bilayer membrane, are secreted into the extracellular environment from cells, do not have the ability to self-replicate, and contain different biomolecules such as proteins, various types of nucleic acids, and metabolites including lipids (Doyle and Wang 2019). The primary difference between exosomes and MVs lays in their biogenesis, which suggests different biological roles (Zaborowski et al. 2015). Advancements in isolation and characterization techniques have allowed for the dissemination of distinct properties and functions to each respective EV type (Kowal et al. 2016). Nevertheless, isolation techniques commonly enrich more than one EV subtype due to a natural overlap in biophysical properties. As a means to provide guidance to the scientific community, the International Society of Extracellular Vesicles (ISEV) published guidelines discouraging the interchangeable use of the terms exosomes and MVs. Instead, ISEV suggested the use of more general nomenclature that does not rely on knowing the

precise biogenesis pathway of the vesicles, such as small EVs (sEVs), medium EVs (mEVs), and large EVs (lEVs) or low- and high-density EVs (Théry et al. 2018). The following discussion on the biology of EVs will use the terms exosomes and MVs only where applicable, while the default will be to use the most up-to-date terminology, respecting the ISEV guidelines.

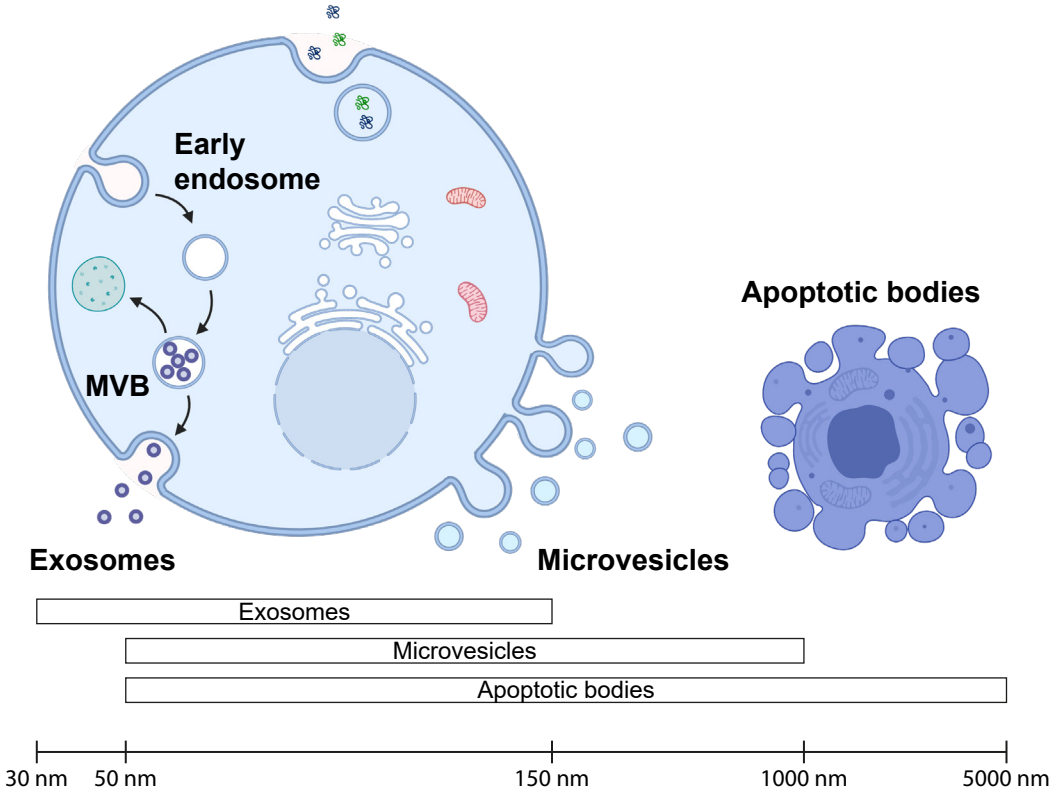

Fig. 1.1 Overview of biogenesis and size of EV subtypes based on their biogenesis and size. The figure was created with biorender.com.

Exosomes are a subtype of EVs that originate from the endolysosomal pathway. They are the smallest known secreted vesicles, with a diameter range between 30 to 150 nm, comparable to small viruses (Yáñez-Mó et al. 2015). MVs, also referred to as ectosomes, microparticles, or shedding vesicles, form directly by outward budding of

the plasma membrane and span a size range from typically 50 to 1000 nm (van Niel, D'Angelo, and Raposo 2018). Both subtypes have a similar density of 1.08 – 1.19 g/ml (Brennan et al. 2020). Morphologically, both exosomes and MVs are considered to have a round shape, whereas double-membrane or multilayer vesicles have been observed as well in EVs isolated from biofluids (Arraud et al. 2014; Emelyanov et al. 2020).

The biogenesis of EVs underlies a controlled sequence of steps including vesicle formation, cargo loading, and the release which occurs at distinct sites in the cell (Raposo and Stoorvogel 2013).

Exosomes form in the endosomal system as intraluminal vesicles (ILVs) by inward budding of the early endosome membrane. During this process early endosomes mature into multivesicular bodies (MVBs). Early endosomes form by inward budding of the plasma membrane, and along with MVBs, are involved in endocytosis and trafficking of cell material (Borges, Reis, and Schor 2013). MVBs are then either degraded in the lysosome or release their content including exosomes into the extracellular space by fusing with the cell membrane. The mechanism that determines the fate of individual MVB are not well understood (Doyle and Wang 2019). The formation of exosomes and cargo loading depend in part on the endosomal sorting complex required for transport (ESCRT) machinery, therefore some of these proteins (Alix, TSG101, HSC70, HSP90β) are expected to be present in exosomes (Tauro et al. 2012). Some studies report an alternative, ESCRT-independent mechanism for exosomes release, which is involves the sphingomyelinase enzyme (Trajkovic et al.

2008). Cells with impaired ESCRT machinery were able to produced CD63-positive exosomes, suggesting at least one more mode of exosome formation (Stuffers et al. 2009). Moreover, proteomic analysis identified CD63 along with other tetraspanins CD9 and CD81 to be highly enriched in exosomes compared to cell lysate [Witwer et al. (2013); sinhaIndepthProteomicAnalyses2014]. Nevertheless, these marker proteins have been found on the cell surface and in other EV types as well, shedding doubt on their ability to serve as exosome markers (Raposo and Stoorvogel 2013; Crescitelli et al. 2013).

MV formation occurs directly at the plasma membrane by pinching or outward budding. The budding process is suggested to require cytoskeleton proteins (actin, microtubules), molecular motors (kinesins, myosin), and fusion machinery (SNAREs and tethering factors) (Cai, Reinisch, and Ferro-Novick 2007). Overall, the process is not yet well understood. Other proteins found in MVs are mainly cytosolic and plasma membrane associated proteins, but may also include heatshock proteins, integrins and proteins containing post-translational modifications (PTMs) (Di Vizio et al. 2012; Morello et al. 2013). Still, there are no specific markers that allow MVs to be distinguished from exosomes (Jeppesen et al. 2019).

1.1.2 Methods for Isolation of EVs

Current EV isolation methods are able to enrich EVs from biological matrices such as blood, saliva, urine or cell culture supernatant (Brennan et al. 2020; Patel et al. 2019). These isolation methods are commonly based on on their size, density, or the

presence of biological molecules, typically surface proteins . Current EV isolation protocols include differential ultracentrifugation (dUC), density gradient separation, ultrafiltration (UF), polymer-based precipitation, affinity-based enrichment and microfluidic methods (P. Li et al. 2017).

Differential ultracentrifugation is the most commonly used technique for EV isolation (Willms et al. 2018). EVs are pelleted by sequential ultracentrifugation steps by subjecting the biological matrix to increasing centrifugal forces, which separates EVs based on their sedimentation coefficient, dependent on their size, density and shape (Livshits et al. 2015). Due to the overlap of these characteristics among EV subtypes, dUC yields a diverse population of EVs (P. Li et al. 2017). EVs isolated by dUC can be further separated by density gradients in form or centrifugation or flotation approaches (X. Zhang et al. 2020). The addition of a density gradient enabled the isolation of EV subpopulations with distinct density and protein profiles (Onódi et al. 2018).

Filtration-based isolation methods are typically combined with dUC protocols to achieve a more defined size distribution (J. Li et al. 2019). Used independently, UF protocols require much less time and have been demonstrated to enrich exosomes and microvesicles from cell line derived EVs (Zeringer et al. 2015). However, the low recovery of EVs from filter membranes (Martin-Jaular et al. 2020) and the potential to rupture or aggregate EVs makes it less suitable for some applications that require intact EVs (Xin et al. 2012).

Polymer-based EV precipitation methods have been used for fast EV isolation that

require only low speed centrifugation. The solubility of EVs is reduced by volume-excluding polymers such as polyethylene glycol, and both polymer and EVs are then pelleted by centrifugation (Brennan et al. 2020). Polymer-based methods were found to preserve the biological activity of EVs (P. N. Brown and Yin 2017), however the amount of coprecipitated proteins and insufficient removal of isolation reagent interfere with MS analysis and yield low numbers of detected proteins (Risha et al. 2020).

Affinity-based EV isolation techniques offer great potential to yield homogeneous subpopulations by capturing EVs that carry a specific protein of interest on their surface (Bellotti et al. 2021; Nakai et al. 2016). These methods have gained prominence despite being labor intensive and cost expensive (Greening et al. 2015). Furthermore, the choice of an appropriate surface target may be difficult if the source of EVs is not well understood.

Other methods such as flow field-flow fractionation (AF4) are being explored to isolated and characterize EVs (H. Zhang and Lyden 2019; P. Li et al. 2017). AF4 separates EVs based on hydrodynamic size and its capability to separate nanoparticles in size ranges from a few nanometers discovered non-membranous nanoparticles termed exomeres (H. Zhang et al. 2018). In another study, EVs below and above 200 nm were efficiently separated based on size and density using ultrasound standing waves (Lee et al. 2015). Microfluidic devices promise quick and specific isolation of EVs populations, but there is currently a lack of standardization across protocols and the inability to isolate EVs from large volumes make them insufficient for MS based

analysis of EVs (P. Li et al. 2017).

1.1.3 Extracellular Vesicles in Breast Cancer

The highly complex nature of healthy and diseased states requires sophisticated intercellular communication methods, from which BC is no exception (Giordano et al. 2020). In recent decades, EVs have been established as mediators of intercellular communication by playing an essential role in the initiation, progression, metastasis and drug resistance of BC (Yu et al. 2015; Adem, Vieira, and Melo 2020; Jabbari et al. 2020; Dong et al. 2020).

Cancer progression is a multistep process that depends on the cell-cell communication within the tumor environment (Tai et al. 2018). It has been suggested that exosomes in the tumor environment actively regulate cancer progression by inducing autocrine and paracrine signalling, reprogramming stromal cells and by promoting angiogenesis (Maia et al. 2018). The development of a diagnostic platform with the ability to detect exosomes, or EVs, with a molecular signature of cancer progression may serve as biomarker for early detection of the disease (Lu and Risch 2016).

Cancer-derived EVs have the ability to transform surrounding or distant cells to a more malignant phenotype by transferring their bioactive contents (Arena et al. 2017) with specific molecules, such as microRNAs (miRNAs) (Arena et al. 2017; Abdouh et al. 2017; Wei et al. 2017). The exosomal miR-222-3p has been demonstrated to functions as a regulator of gemcitabine resistance and malignant characteristics by targeting the promotor of SOCS3, a negative feedback regulator of the JAK/STAT sig-

naling pathway (Yoshimura, Naka, and Kubo 2007). EVs have the ability to remodel the extracellular matrix via matrix metalloproteases, which may promote metastatic niche formation (Nawaz et al. 2018). Moreover, the uptake of tumor-secreted EVs has been shown to promote metastatic niches at distant sites (Hoshino et al. 2015) and the profile of surface integrin proteins of exosomes influenced their organotrophic uptake from liver and lung (Hoshino et al. 2015). EVs may have the potential to provide a snapshot that gives insight on the metastatic state of BC and thus could be useful as a source of biomarkers that can predict the overall outcome of the disease. Furthermore, a better understanding of the organtrophic nature of EVs isolated from biofluids like blood may help prevent metastasis of BC to different organs.

Chemotherapy, among others, is a standard treatment for invasive BC, however there is a variety of mechanisms by which tumor cells evade drug-induced cell death (Dallavalle et al. 2020). Some of these mechanisms include efflux or inactivation of drugs (Dallavalle et al. 2020), enhancement of deoxyribonucleic acid (DNA) damage repair (Battista et al. 2020), and inducing epithelial–mesenchymal transition (Navas et al. 2020). HER2-overexpressing tumor cell lines SKBR3 and BT474 secrete active HER2 protein on exosomes, which was demonstrated to bind to and inhibit the monoclonal antibody treatment, Trastuzumab, which is used to treat HER2-positive BC patients (Ciravolo et al. 2012). This suggests EVs may be useful as predictive biomarkers to inform clinicians about drug sensitivity or resistance to possible treatments.

One determining factor of BC prognosis is the early detection of cancerous cells,

which can reduce the death rate significantly (L. Wang 2017). However, current approaches for detection and diagnosis of BC such as mammography, ultrasound, magnetic resonance imaging, computerized tomography or needle biopsy are costly, time consuming and not adapted to young women creating a burden to the healthcare system as well as individuals (Milosevic et al. 2018). These complications combined with a lack of sensitivity of early detection of new cancer cells make most methods unsuitable for broad screening for BC. While genetic mutations such as in the BRCA1 and BRCA2 gene are used for risk assessment of BC, there are currently no biomarkers that can be detected in liquid biopsy to diagnose BC in early stages (Duffy et al. 2015).

EVs are involved in a broad spectrum of processes in BC which suggests they may be a non-invasive source of diagnostic, prognostic and predictive biomarkers (Chitti and Nedeva 2021; Shaw et al. 2015). One of the most common techniques employed for protein biomarker discovery is MS (Crutchfield et al. 2016). Mass spectrometers are unique in their ability to analyze virtually any biological molecule and specifically quantify more than 1000 proteins in parallel.

1.2 Mass Spectrometry-Based Proteomics

Cells are composed of a quasi-innumerable number of biomolecules of distinct nature that act concordantly together allowing all forms of life to exist, reproduce, and react to external stimuli (Marth 2008). The understanding of the interplay between all these molecules is the fundamental question of molecular biology (Levin 2006). The discovery of DNA as the carrier of genetic material was revolutionary to our understanding about how cells proliferate to allow for organisms to develop. This understanding led to the postulation of the central dogma of molecular biology, which describes the unidirectional flow of information from DNA to RNA to proteins (Crick 1970). The exploration of the human genome has been motivated by recent advancements in sequencing technologies that have allowed for sequencing of the whole human genome, as well as its transcribed genetic material (Venter et al. 2001). The fields of genomics and transcriptomics study the totality of DNA or transcribed RNA of an organism, respectively. While the genome provides the information for all cellular processes, proteins are the main player executing cellular functions. Therefore, proteomics aims to understand the proteome, the complete set of proteins of a cell (Aebersold and Mann 2003).

Nowadays, the most popular technique to identify and quantify almost entire proteomes is based on the combination of liquid chromatography and tandem mass spectrometry (LC-MS/MS). LC-MS/MS follows in most cases the bottom-up strategy, in which proteins are first subject to proteolytic digestion to generate peptides that can be separated by LC and analyzed by MS. In contrast, the top-down approach

analyzes intact proteins, which remains a technical challenge applied to large scale analysis (K. A. Brown et al. 2020). Besides the growing number of subfields such as phosphoproteomics, spatial proteomics, single-cell proteomics, et cetera, the primary question addressed by MS-based proteomics is the identification and relative or absolute quantification of proteins. (Aebersold and Mann 2003)

1.2.1 Bottom-Up Proteomics

Bottom-up proteomics experiments can be applied to a wide number of sample sources, ranging from cell culture material to tissue or body fluids from patients. The workflow can be separated into three main steps: (1) Sample preparation of proteins from the crude biological matrix to generate defined peptides, that can be (2) analyzed by LC-MS/MS generating a set of MS1 and MS2 spectra, which are then used as input for (3) search engine matching of experimental spectra against in-silico predicted spectra to identify and quantify peptides and proteins.

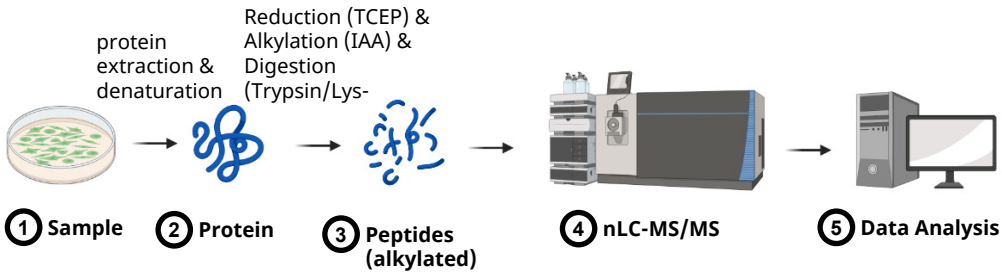

Fig. 1.2 General schematic of a proteomics experiment. (1) A sample from any biological matrix is subjected to a protein extraction method to obtain (2) proteins in solution. Proteins are then digested into (3) peptides to be analyzed by (4) LC-MS/MS. Generated mass spectra are then analyzed for (5) peptide and protein identification.

Sample Preparation for MS Analysis

The primary challenge of sample preparation for proteomics is performing efficient proteolytic digestion of proteins to generate peptides for MS analysis. To address this challenge, proteins must be solubilized and denatured, ideally to a linearized form, to allow the protease to access peptide bonds for cleavage. Here, only the in-solution digestion approach will be discussed, as advancements in MS technologies have made protein separation by SDS-polyacrylamide electrophoresis (SDS-PAGE) and accompanied in-gel digestion obsolete (Shevchenko et al. 2006).

The first step is to lyse the biological matrix, followed by protein extraction and separation from other cellular structures, such as the lipid membranes or protein-DNA complexes. The separation of proteins from other biomolecules may be facilitated by physical pressure, heat, solvents, or detergents such as sodium dodecyl sulfate (SDS) or n-dodecyl-β-D-maltoside (DDM) which act by disturbing the intermolecular interactions and facilitate bring proteins into solution (Niu et al. 2018). There are a wide range of detergents: ionic, non-ionic, and detergents that retain the native confirmation of proteins, the latter of which is colloquially known as MS-compatible detergents and are useful for affinity enrichment and MS analysis. The type of detergent should carefully consider the origin and composition of the sample, as well as the downstream application. Subsequently, proteins are denatured, if not already done by the detergent. Protein denaturation can be accomplished by chaotropes such as urea or guanidinium chloride. These agents disturb the secondary structure by interfering with intramolecular hydrogen bonding. To support denaturation and

prevent spontaneous refolding, disulfide bonds between cysteine residues are reduced with dithiothreitol (DTT) or tris(2-carboxyethyl)phosphine (TCEP) and are capped by a thiol-reactive halogen compound, iodoacetamide (IAA) for example. Following denaturation and capping of cysteine residues, one or multiple proteases are used to digest proteins into peptides that are identifiable by MS. The majority of studies uses trypsin, a pancreatic serine protease, due to its highly specific cleavage C-terminal to arginine and lysine residues. Both amino acids have a suitable frequency and distribution in the proteome and increase the positive charge of peptides due to their basic nature, which is necessary for ionization during MS analysis. Other proteases find application by either supporting tryptic digestion, in the case of endoproteinase Lys-C (Lys-C) which cleaves C-terminally to lysine with higher efficiency than trypsin (Saveliev et al. 2013), or to enable the analysis of proteins where lysine and arginine do not provide sufficient coverage, such as Chymotrypsin (C-terminal of tyrosine, phenylalanine, tryptophan, and leucine) for membrane proteins. After digestion of proteins, peptides are separated from interfering molecules, such as salts and buffer components by C18 cartridges prior to MS analysis.

Throughout the sample preparation process, the pH must be kept stable by compatible buffers to enable efficient and site-specific reduction, alkylation, and proteolytic cleavage. These controlled conditions aim to produce sequence and modification-defined peptides, which can be predicted and identified by the mass spectra processing software. The field has heavily invested in protocol development in an attempt to overcome various challenges during the sample preparation process, such as solvent-

induced precipitation of proteins or removal of endogenous and exogeneous interfering molecules, such as lipids, metabolites or MS-incompatible reagents like SDS (Chertov et al. 2004). A commonly adopted protocol known as filter-aided sample preparation (FASP) makes use of a molecular weight cut-off (MWCO) filter to rapidly exchange buffers and eliminate interfering substances during sample preparation (Wiśniewski et al. 2009).

LC-MS/MS Analysis

The development of MS, initially mass spectroscopy due to the qualitative character of the analysis, terms back to the manipulation of trajectories of neon ions by electric and magnetic fields, visualized on a photographic plate by J. J. Thomson. Modern mass spectrometers still rely on the same principles of (i) a source that ionizes analyte molecules and brings them into the gas phase, in which (ii) ions are separated by the mass analyzer based on their mass-to-charge (m/z) ratio through electric and magnetic fields, and (iii) a detector that measures the intensity of separated ion species (Glish and Vachet 2003). Modern mass spectrometers address a wide range of problems in science by application-specific combinations of different types of ion sources, mass analyzers, and detectors (Finehout and Lee 2004; Nadler et al. 2017). Additional parts such as ion traps, quadrupoles, and ion routing modules enable separation, isolation or even fragmentation of peptide ions for precise m/z determination within a few parts per million, quantification, and structural and sequence determination. Furthermore, the combination with other instruments such as high-performance liquid chromatography (HPLC), capillary electrophoresis, or biosensors helps separate

peptides and decreases the complexity of ions analyzed at a time. The combination of reversed-phase LC with MS/MS is the most commonly method for bottom-up proteomics experiments in which peptides are separated by their hydrophobicity on a C18 column, ionized by electrospray ionization (ESI) to form positive gas phase ions, and are subjected to MS/MS to determine the peptide precursor m/z and its sequence composition (Aebersold and Mann 2003). In the following, the combination of (nano) LC-MS/MS using an Orbitrap-Linear ion trap mass spectrometer operated in a data-dependent acquisition (DDA) mode is introduced.

Desalted and dried peptides are reconstituted in MS-grade water acidified with e.g. formic acid to ensure peptide solubility and a positive net charge of peptides. Peptides are introduced on the C18 LC column and then eluted by a water-acetonitrile gradient and sprayed into the MS source. This leads to hydrophilic peptides being analyzed at the beginning and more hydrophobic, and generally longer peptides to elute later during the experiment. Once peptides elute from the C18 column, they are ionized and evaporated from the spray needle by ESI. This requires a potential of ~2 kV between the needle and the the MS source. An aerosol is formed from which the droplets burst due to the high internal density of positive charges, while water and acetonitrile evaporate as uncharged molecules (Ho et al. 2003). All molecules that enter the mass spectrometer are initially filtered for charged ions and uncharged molecules are being eliminated. Remaining ions are further filtered by a quadrupole to fit a specified m/z range excluding very small peptides and metabolites by a lower limit of ~300 m/z or very large peptides above ~2000 m/z. First, MS1 spectra are acquired for all

ions entering the mass spectrometer at a given time using an ultra-high resolution mass analyzer such as an Orbitrap mass analyzer, providing precise measurement of precursor m/z of peptides and their charge states from isotope patterns (Zubarev and Makarov 2013).

As precursor masses, obtained from m/z and charge state, do not suffice to identify tens of thousands of peptides from complex samples, insight into their respective sequence composition is required (Coon et al. 2005). This is obtained by isolation and fragmentation of peptide ions of a specific m/z, monitored as MS2 spectra. Fragmentation patterns can be predicted in-silico depending on the employed fragmentation method: collision-induced dissociation (CID), higher-energy collisional dissociation (HCD), or electron transfer dissociation (ETD). CID, the most frequently utilized fragmentation method in MS-based proteomics, relies on the collision of peptide ions with neutral gas molecules such as Helium, Nitrogen or Xenon (McLuckey 1992). Part of the kinetic energy from the impact of peptide ions with the collision gas is converted into vibrational energy, which is instantly distributed over all bonds of the peptide and leads to breakage of the weakest bonds, namely the C-N peptide bond resulting in so called b- and y-ions (Roepstorff and Fohlman 1984). Depending on specific applications such as isobaric labelling for quantification or to preserve PTMs, HCD or ETD are being used, respectively. However, primary reason for the use of CID in regular proteomics experiments is the short fragmentation time, allowing for a higher number of acquired MS2 spectra and consequently more identified peptides. Compared to the high-resolution requirements of the MS1 spectra, MS2 spectra can

be acquired quickly with a linear ion trap (LIT) and fragment mass deviation of up to 0.5 Da are accepted. The described sequence of acquiring a high resolution MS1 spectrum and subsequently isolating and fragmenting single ions is referred to as DDA and proved as an efficient mode for MS-based proteomics (Doerr 2015).

MS Data Analysis

Analysis of MS data can be split into spectra processing and downstream statistical and bioinformatics analysis, whereas the first provides peptide and protein identification and quantification information from raw mass spectra (Yates 2004), which is then further investigated to infer the state or the composition of a biological system (Schmidt, Forne, and Imhof 2014).

MS spectra processing is performed by comparing experimental peptide masses from MS1 spectra and their respective fragmentation patterns from MS2 spectra with an in-silico generated database of possible peptide sequences (Cox et al. 2011). Based on the biological system and the experimental conditions, a species-specific database, in format of a fasta file, of all protein sequences is required. All possible peptide sequences are then predicted based on in-silico digestion of given protein sequences with protease-specific cleavage sites, in combination with given modifications of residues. Peptide modifications are categorized as variable or fixed modifications, depending if they are optionally present on a residue such as oxidation of methionine or if they are expected on each residue such as carbamidomethylation of cysteine residues, the chemical modification introduced by IAA during sample processing, respectively. In

the case of variable modifications, unmodified residues are accepted as well. While analysis software such as MaxQuant (Cox and Mann 2008) and ProteomeDiscoverer (Orsburn 2021) provide an interface to set all experimental parameters and generate a list of monoisotipic masses from MS1 and MS2 spectra, each uses a separate search engine like Andromeda (Cox et al. 2011), Mascot (Brosch et al. 2009), or Sequest (Eng, McCormack, and Yates 1994) to match experimental data with predicted peptide sequences. Once all precursor masses and MS2 spectra have been compared, each experimental MS1-MS2 spectrum pair is scored based on precursor mass deviation and the number of matched fragment ions compared with one or more predicted peptide sequences. This results in the assignment of peptide-spectrum matches (PSMs). Identified peptides are then aggregated to proteins, however some proteins cannot be distinguished by the set of identified peptides due to sequence similarities. This challenge is overcome by assigning protein groups, a set of proteins all identified by the same set of peptides. In case of distinct protein groups, overlapping peptides, or so-called razor peptides are being assigned to the protein group with the higher number of total peptides (Beer et al. 2017). The estimation of a false discovery rate (FDR) is performed with a decoy database strategy both on the peptide and the protein level (Elias and Gygi 2007). A database of sequences not expected in the sample, commonly the reversed main search database, is analyzed in parallel with the true sequences. The distributions of scores of peptide identifications from both searches are then compared, while the decoy database generally exhibits a shift towards lower confidence scores. A confidence threshold for peptide identification for the desired FDR is then estimated based on the fraction of true and decoy sequences at a given

confidence score. Besides an FDR, contaminant databases are being used to filter out peptides and proteins that do not stem from the sample itself and may have been introduced during the sample processing (Mellacheruvu et al. 2013). The results from an analysis are provided as protein- and peptide-centric tables which contain sample-wise identification and quantitative information. Identification is the qualitative information if a peptide or a protein is present in any given sample and if so, the intensity measured by the mass spectrometer of peptides can be used to infer relative peptide and protein abundances. While peptide intensities can be compared between the same species of peptides, protein abundance must be inferred based on the intensity of the respective peptides identified for a given protein. A generic quantification technology implemented in MaxQuant is the MaxLFQ algorithm, which yields the so called label-free (protein) quantification (LFQ) intensity [Cox et al. (2014). Peptide intensities are normalized across all samples in order to yield the lowest overall variation of the proteomes and LFQ intensities for proteins are then obtained by the pair-wise median ratio of common peptides for each protein between samples. The resulting LFQ intensity can then be used in further statistical tests to compute significance and extend of differential protein abundance.

1.2.2 Proteomics Approaches Beyond Protein Identification

Advancements in the field of proteomics resulted in a wide range of biochemical and computational approaches, which enable MS-based proteomics to answer a broad spectrum of questions beyond protein identification and quantification (Altelaar, Munoz, and Heck 2013). The investigation of PTMs such as phosphorylation or acetylation requires enrichment of modified proteins, or peptides after proteolytic digestion. Besides PTMs, the function or role of a protein depends on its subcellular localization (Millar and Taylor 2014). Therefore, understanding of the subcellular distribution of proteins offers new insights into biological processes in health and disease states (Lundberg and Borner 2019). The primary mode of compartmentalization in the cell is driven by phospholipid membranes to which proteins can be associated integrally through transmembrane domains, or from the interior or exterior side through PTMs or by interaction with other proteins (Tan, Tan, and Chung 2008). As the relation of proteins to membranes is not ambiguous based on known sequence composition and PTMs, it must be experimentally determined in respective cellular compartments or under specific disease states (Pankow et al. 2019).

The discovery of EVs and their role in many biological processes has led to the in-depth profiling of their proteome (Hoshino et al. 2020), phosphoproteome (Andaluz Aguilar et al. 2020), transcriptome (Turchinovich, Drapkina, and Tonevitsky 2019), lipidome (Brzozowski et al. 2018), and metabolome (Puhka et al. 2017). Many individual proteins from EVs have been assigned to impose changes in recipient cells or in the extracellular environment, but still the proteome of EVs is considered to be

not fully understood (Willms et al. 2018; Martin-Jaular et al. 2020). In the following, different proteomics approaches are discussed, from which additional insight into the proteome of EVs can be gained. Methods are separated into biochemical and data analysis-based approaches.

Biochemical Approaches for Proteomics Applications

Biochemical approaches to problems addressed by MS include techniques to separate, enrich, fractionate, or modify proteins in their native or quasi native environment to gain information on their cellular or subcellular localization (Lundberg and Borner 2019). For example, introduced chemical modifications can be identified by MS or serve as affinity tags to enrich subsets of proteins. For this purpose, the chemical agent must selectively label proteins of interest such as surface exposed proteins by using an non-membrane permeable agent that only reacts with residues exposed on the surface of the membrane (Huang 2012). Many variantions of agents are available for such purposes, by providing differences in residue specificity, charged groups that impede permeation of membranes, affinity structures for enrichment, and chemical cleavage sites for convenient elution of affinity captured proteins. A wide range of approaches have been extensively used to identify cell surface proteins (Henry et al. 2018; M. Li et al. 2021), as well as EV surface proteins (Cvjetkovic et al. 2016; Rai et al. 2021). Since targeted groups must be accessible for successful labelling, it may in addition to the localization of proteins, also inform about the availability of residues for interaction with other proteins like antibodies.

Data Analysis Approaches for Proteomics Applications

Data analysis-based approaches for spatial or functional protein classification may be informative for the state or localization of proteins, without the need to for chemical modification. Mainly, dimension reduction methods like principal component analysis (PCA), clustering algorithms, or correlation analysis have been used to identify groups of proteins which are affected by cellular perturbations and mapped to pathways, cellular processes, or to organelles or cell types. Similarities in abundance profiles of proteins in biological replicates over several fractions are used to infer gene co-expression (Langfelder and Horvath 2008) or protein co-localization (Itzhak et al. 2016).

Weighted gene co-expression network analysis (WGCNA) is a popular framework introducing correlation network analysis initially developed for gene expression data (Langfelder and Horvath 2008). Gene or protein networks are created based on expression or abundance data, respectively, by measuring pairwise protein correlation and further transforming correlation coefficients to identify robust clusters. Such clusters can then be related to biological concepts through annotation data.

Dynamic Organellar Maps, a tool for dynamic and quantitative protein subcellular localization from protein fractionation was initially developed for organelles and intracellular localization prediction (Itzhak et al. 2016). In short, PCA is applied to protein abundance profiles over several fractions obtained by dUC. The resulting PCA plot was then annotated by known organelle markers to identify protein cluster for specific organelles. More recently, the described tool has been used to track changes

in protein composition of EVs upon HIV infection (Martin-Jaular et al. 2021).

1.3 Thesis rationale

EVs have emerged as a new mechanism to facilitate intercellular communication. Deducing their role in physiological and pathophysiological processes is of particular interest to researchers world-wide (Mallia et al. 2020; Ohno, Ishikawa, and Kuroda 2013). Proteins, along with other biomolecules, compose the cargo of EVs, making them a promising source of biomarkers for non-invasive detection of various diseases (Hoshino et al. 2020) given they can be isolated from all biofluids (Yuan, Li, and Wang 2021). Early on in the evolution of our understanding of EVs, MS-based proteomics discovered the protein markers that are widely used for confirming the presence of EVs: CD63, CD9, and CD81 (Kowal et al. 2016). Moreover, proteomics profiling of sEVs allowed classification of BC subtypes from cell lines (Rontogianni et al. 2019). This motivated further exploration of BC cell line-derived EVs to identify surface biomarkers that can be directly detected from biofluids for BC diagnosis by antibody- or aptamer-based diagnosis techniques (Tian et al. 2021). In this study, two proteomics-based approaches for the discovery of potential surface EV protein biomarker are demonstrated. EVs were isolated by dUC from three epithelial breast cell lines with varying degrees of malignancy: MDA-MB-231, MCF7 and non-malignant MCF10A cells.

The first approach is based on cell line specific identification of sEV proteins. Surface proteins are predicted by database annotations of membrane proteins and cluster

analysis predicted co-localization of potential biomarkers with EV marker proteins. The presence of selected markers are then validated by western blot analysis for their potential use as BC markers.

The second approach uses a biotinylation reagent to experimentally identify proteins that are accessible from the surface of EVs. Proteins common to EVs from MDA-MB-231 and MCF7, but not identified in MCF10A EVs, are considered general cancer EV markers. These proteins are then annotated with disease terms to identify known BC-associated proteins that can be directly detected from biofluids on the surface of EVs.

Chapter 2

Materials and Methods

2.1 Cell Lines and Cell Culture

MDA-MB-231 (ATCC® HTB-26™), MCF7 (ATCC® HTB-22™) and MCF10A (ATCC® CRL-10317™) cell lines were obtained from ATCC, Canada. MDA-MB-231 and MCF7 were cultured in DMEM/F12 growth medium supplemented with 10% (v/v) fetal bovine serum (FBS), 100 U/mL penicillin, 100 μg/mL streptomycin. MCF10A was cultured in DMEM/F12 growth medium supplemented with 5% (v/v) horse serum (HS), 100 U/mL penicillin, 100 μg/mL streptomycin, 20 ng/mL epidermal growth factor (EGF), 0.5 mg/mL hydrocortisone, 10 μg/mL insulin, and 100 ng/mL cholera toxin (Debnath, Muthuswamy, and Brugge 2003). FBS and HS were diluted 1:1 with growth medium and centrifuged for 20 h at 100,000 g to prepare EV-depleted growth medium. All cell lines were grown in EV-depleted cell medium in a humidified incubator at 37 °C with 5% CO_2 for 48 to 72 h until they reached ~90% confluence prior

to collection of cell culture supernatant.

2.2 EV Isolation by Differential Ultracentrifugation

The EV isolation protocol for small to large EVs was adapted from previous protocols by calculating the k-factor, a measure for the pelleting efficiency (Livshits et al. 2015). Approximately 37 mL EV-containing cell supernatant were collected per cell culture dish and transferred to Falcon tubes separately. Cells were depleted by centrifugation at 300 g for 10 min at 4 °C. The supernatant was decanted into another falcon tube and cell debris and larger vesicles such as apoptotic bodies were depleted at 2000 g for 20 min at 4 °C and the supernatant was frozen prior to further processing. Frozen samples were thawed overnight at 4 °C and centrifuged again at 2000 g for 20 min at 4 °C before the supernatant was transferred to ultracentrifuge tubes (Polypropylene, Open-Top Thinwall centrifuge tubes, 38.5 mL, Beckman Coulter). Following centrifugation steps were done using a Beckman Coulter XL-A Analytical Ultracentrifuge. Medium EVs were isolated for 1 h at 4 °C at 16,500 g_{avg} (SW 28 swinging bucket-rotor; k-factor: 1,541, 11,200 rpm, Beckman Coulter). The supernatant was collected for sEV isolation for 3 h at 4 °C at 100,000 g_{avg} (SW 28 swinging bucket-rotor; k-factor: 256, 11,200 rpm, Beckman Coulter). Pellets from both centrifugation steps were resuspended in PBS and subjected to a repeated centrifugation step at previous conditions. Pellets were resuspended in 100 μL PBS and frozen at

-20 °C until further processing.

2.3 Determination of Protein Concentration

The protein content of individual samples was measured with the Bradford assay. 4 μL of solubilized EV samples was incubated with 10 μL Coomassie reagent (Pierce™ Coomassie Plus (Bradford) Assay Reagent) and incubated for 10 min at RT. The absorption was read at a wavelength of 595 nm using a NanoDrop™ One Microvolume UV-Vis Spectrophotometer. A calibration curve was prepared with bovine serum albumin (BSA).

2.4 Nanoparticle Tracking Analysis

A ZetaView nanoparticle tracking microscope PMX-110 (Particle Metrix) was used for determining the size distribution of EVs. Camera shutter speeds of 85 and 40 were used. The Instrument was calibrated and the camera was focused with 102 nm polystyrene beads (Microtrac 900383).

2.5 Transmission Electron Microscopy

EVs were fixed in 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer (pH 7.4). Fixed suspensions were spotted on Formvar coated copper grids (200 mesh; Canemco, Lakefield, ON, Canada) for 30 s. Samples were negatively stained with 2% uranyl acetate in water for 6 min and dried with filter paper. Vesicles were examined on

a transmission electron microscope (JEOL JEM 1230, Japan) that was operated at 50 kV.

2.6 Western Blot Analysis

SDS electrophoresis and Western blot (WB) experiments were carried out as described previously (Risha et al. 2020). The dilution for primary antibodies was 1:1000 and the dilution for the secondary antibody was 1:5000. All antibodies were purchased from Abcam.

2.7 Sample Preparation for Proteomics Analysis

The sample preparation protocol was adapted from the filter-aided sample processing (FASP) protocol (Wiśniewski et al. 2009). Frozen EV samples in PBS, approximately 100 μL, were thawed at room temperature and adjusted to 0.1% DDM, 50 mM TRIS-HCl, pH 7.5. Samples were heated to 95 °C for 3 min and cooled to RT before transfer to a 10 kDa MWCO filter (Microcon® 10K device, Cat No. MRCPRT010, Millipore). Buffer was exchanged by centrifuging the filter devices for 15 min at 14,000 g and adding 100 μL denaturation buffer (8 M urea, 50 mM TRIS HCl, pH 8). Proteins were reduced by adding 4 μL 100 mM TCEP in ddH_2O and incubation for 30 min at RT followed by centrifugation for 15 min at 14,000 g. 100 μL denaturation buffer and 4 μL 500 mM IAA in H_2O were added to alkylate proteins in the dark for 45 min at RT. The reaction was completed during the 15 min centrifugation at 14,000 g.

Digestion buffer (50 mM TRIS HCl, 0.6% glycerol, pH 8) was added and remaining urea was washed by a repeated centrifugation step before digestion buffer and 300 ng trypsin/Lys-C were added. The digestion was maintained at 37 °C for 12 h. Peptides were collected by centrifugation at 14,000 g for 10 min, adding of 40 μL digestion buffer, and an additonal 15 min of centrifugation. Peptides were acidified with 2 μL of 100% formic acid and desalted for MS analysis using C18 TopTip microcolumns (Cat No. TT2C18, Glygen) according to the manufacturer's protocol. Purified peptides were vacuum dried and frozen at -20 °C.

2.8 Surface Protein Labelling of Extracellular Vesicles

Cell culture supernatant was collected and prepared as previously described without freezing the media. Following the first ultracentrifugation steps for mEVs and sEVs at 16,500 g and 100,000 g, respectively, pellets were resuspended in 10 mM EZ-Link™ Sulfo-NHS-SS-Biotin (ThermoFisher Scinetific, 21331) in PBS pH 7.4. Primary amines exposed on the surface of EVs were biotinylated for 2 h at RT. The reaction was quenched by adding an equal volume of 200 mM TRIS pH 8 and incubation for 15 min. Labelled EVs were separated from free Sulfo-NHS-SS-Biotin-TRIS conjugates by diluting EVs and free linker with PBS and and an additional cleanup ultracentrifugation step as previously described. Pellets were resuspended in PBS and frozen until further processing.

2.9 Surface Protein Enrichment

While labelled EV samples were thawed, 120 μL (binding capacity: 60 μg of biotinylated BSA) Pierce™ Streptavidin Agarose resin (ThermoFisher Scientific, 20349) was added on top of Pierce™ Micro-Spin Columns (ThermoFisher Scientific, #89879) and washed three times with 1X binding buffer (0.1% DDM, 50 mM TRIS, 150 mM NaCl, 1 mM $MgCl_2$, 5% glycerol, pH 7.5). Thawed EV samples were adjusted with 10X binding buffer and heated to 95 °C for 3 min. Samples were cooled down, added to the resin columns and labelled proteins were captured on streptavidin resin for 30 min at 4 °C while shaking. Columns were placed in new Eppendorf tube to collect unlabelled proteins. Unbound proteins were washed five times with 1X binding buffer with 2 M Urea and collected by gentle centrifugation. Following the last wash, columns were transferred to a new tube and 1X binding with 50 mM TCEP was added to release bound proteins. The cysteine bond of the biotin linker was reduced for 2 h at RT. In parallel, collected unbound proteins were added to a 10 kDa MWCO filter, on which the buffer was exchanged to 1X binding buffer with 50 mM TCEP and incubated for 2 h at RT. Biotinylated proteins were eluted from the streptavidin resin column by gentl centrifugation and an additional wash with binding buffer and 2 M Urea, and added on top of a 10 kDa MWCO filter for MS sample preparation.

From here, both washed non-biotinylated and eluted biotinylated samples were treated the same. The binding buffer with TCEP was exchanged by centrifugation for 15 min at 14,000 g with denaturation buffer (8 M urea, 50 mM TRIS HCl, pH 8) and an additional centrifugation for 15 min at 14,000 g. 100 μL denaturation buffer and

4 μL 500 mM IAA in H_2O were added to alkylate proteins in the dark for 45 min at RT. The reaction was completed during the 15 min centrifugation at 14,000 g. Digestion buffer (50 mM TRIS HCl, 0.6% glycerol, pH 8) was added and remaining urea was washed by a repeated centrifugation step before digestion buffer and 300 ng trypsin/Lys-C were added. The digestion was maintained at 37 °C for 12 h. Peptides were collected by two centrifugation steps at 14,000 g for 10 min, after which 40 μL digestion buffer was added, and 15 min. Remaining peptides on top of the filter, were digested by adding 100 μL digestion buffer with 10 mM $CaCl_2$ and Chymotrypsin for 4 h at 25 °C. Peptides were eluted as described above, acidified with 4 μL of 100% formic acid and desalted for MS analysis using C18 TopTip microcolumns (Cat No. TT2C18, Glygen) according to the manufacturer's protocol. Purified peptides were vacuum dried and frozen in -20 °C.

2.10 LC-MS/MS Analysis

The LC-MS/MS analysis procedure has been described previously (Risha et al. 2021). Protein samples with up to 20 μg of protein content as determined by Bradford assay were analyzed on an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific) coupled to an UltiMate 3000 nanoRSLC (Dionex, Thermo Fisher Scientific). Peptides were separated on an in-house packed column (Polymicro Technology), 15 cm $\times$ 70 μm ID, Luna C18(2), 3 μm, 100 Å (Phenomenex) employing a water/acetonitrile/0.1% formic acid gradient. Samples were loaded onto the column for 105 min at a flow rate of 0.30 μL/min. Peptides were separated using 2% acetonitrile in the first 7 min

and then using a linear gradient from 2 to 38% of acetonitrile for 70 min, followed by a gradient from 38 to 98% of acetonitrile for 9 min, then at 98% of acetonitrile for 10 min, followed by a gradient from 98 to 2% of acetonitrile for 3 min and wash 10 min at 2% of acetonitrile. 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 the 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 1×104 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%.

2.11 MS Data Processing

MS raw files were analyzed with MaxQuant (version 2.0.1.0) and the Andromeda search engine (Cox and Mann 2008; Cox et al. 2011). Peptides were searched against a human UniProt FASTA file containing 20,412 entries (03.01.2019) and a default contaminants database. Default parameters were used if not mentioned otherwise. N-terminal acetylation and methionine oxidation were set as variable modifications, and cysteine carbamidomethylation was set as a fixed modification. A minimum peptide length of 6 amino acids was required and false discovery rate (FDR) was set to 0.01 for

both the protein and peptide level, determined by searching against a reverse sequence decoy database. Enzyme specificity was set as C-terminal to arginine and lysine, for regular proteomics samples and including Chymotrypsin for labelled proteins, with a maximum of two missed cleavages. Peptides were identified with an initial precursor mass deviation of up to 4.5 ppm and a fragment mass deviation of 0.5 Da. The 'Match between runs' algorithm in MaxQuant was performed between all samples to increase peptide identification rate. (Nagaraj et al. 2012) Proteins and peptides matching to the reverse database were discarded. For label-free quantification, a minimum ratio count of 2 was required.

2.12 Data Analysis

All data analysis and visualization was done in R and the in-house built pOmics R package (github.com/nicohuttmann/pOmics). MaxQuant output tables proteinGroups.txt were loaded in R. Potential contaminants and reverse protein identifications, as annotated by MaxQuant, were excluded from the analysis. Proteins needed to be identified by at least one unique peptide in two or more samples to be considered identified.

PCA was performed with the prcomp R funtion based on scaled LFQ intensity values. Proteins with a LFQ value in at least 50% of samples were used and missing values were imputed by drawing values from a down shifted Gaussian distribution of log-transformed protein LFQ intensities (shift = 1.8 standard deviations (sd), width = 0.3 sd) to simulate low abundance profiles (Lazar et al. 2016).

Correlation matrices to predict protein co-localization were constructed for sEVs from each cell line separately by computing the pairwise Pearson correlation coefficient from protein LFQ intensity. A hierarchical clustering was computed from the resulting matrix using euclidean distance and complete linkage clustering. The correlation matrices were plotted and the position of potential biomarkers and EV proteins CD63, CD9 and CD81 was indicated to visualize protein clusters of similar abundance profiles.

Protein subcellular localizations were retrieved with the UniProt.ws R package (Carlson, Ortutay, and Maintainer 2021), Gene ontology annotations from the org.Hs.eg.db R package (Carlson 2019), and gene-disease relations with the disgenet2r package (Piñero et al. 2020).

Identification of surface-accessible proteins by the labelling approach was performed in biological triplicates for each cell line and EV fraction. The number of replicates in which a protein was identified was compared between wash and the eluted fractions. Proteins identified in at least two more elution fractions than wash fractions were considered surface accessible.

Chapter 3

Results

3.1 Characterization of Cell Line-Derived EVs

In previous studies, sEVs or mEVs from BC and breast epithelial cell lines were compared to identify potential new biomarkers for non-invasive liquid biopsies (Risha et al. 2021, 2020; Rontogianni et al. 2019). While the focus of this study lays on sEVs, mEVs were analyzed as well to provide a comprehensive view of the vesicles secreted from MDA-MB-231, MCF7 and MCF10A. Cells were grown in EV-depleted growth medium for 48 - 72 h to reach a confluence of 80 - 90%. The supernatant was collected and remaining cells, debris, and apoptotic bodies were depleted by centrifugations at 300 g and 2000 g (Fig. 3.1A). From sequential centrifugation steps, mEVs were collected from the 16,500 g pellet, whereas sEVs were collected from the 100,000 g pellet. The pellets of mEVs and sEVs were washed once by resuspension and ultracentrifugation at 16,500 g and 100,000 g, respectively. The centrifugal forces

selected for the EV isolation protocol described here considered values published in the literature but were modified to account for the available rotor type and dimensions (Théry et al. 2006). Following isolation, EV pellets were resuspended in PBS and stored at -20°C prior to further experiments.

The isolated mEVs and sEVs were each characterized by NTA, TEM, and WB analysis. The size range of the particles was measured by NTA. The median diameter of mEVs isolated from MDA-MB-231, MCF7 and MCF10A cells was 159 nm, 146 nm and 150 nm, while the median diameter of sEVs was 107 nm, 110 nm and 129 nm, respectively (Fig. 3.1B). TEM analysis of the isolated vesicles demonstrates the presence of an intact lipid membrane (Fig. 3.1D). WB analysis was used to confirm the presence of protein markers known to be present on EVs: CD63, CD9, and CD81 (Fig. 3.1C). All EV marker proteins were present in sEV pellets from each cell line, however CD63 was only found less prominent in mEVs from MCF7 cells, suggesting an enrichment of EVs in the sEV pellet.

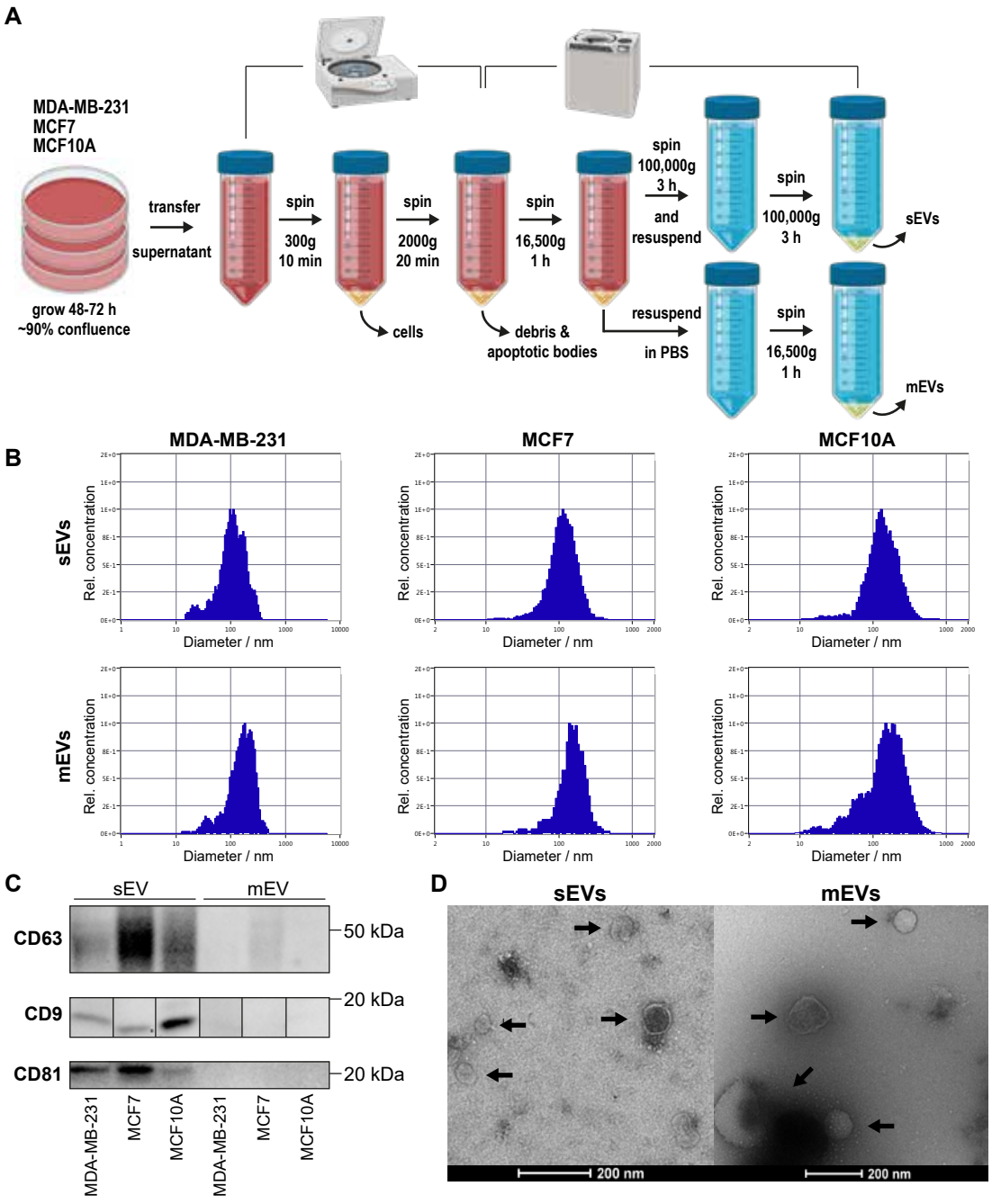

Fig. 3.1 Characterization of EVs isolated from cell culture supernatant by dUC. A) Workflow for EV isolation from cell culture supernatant based on dUC. B) Size distribution of isolated sEVs and mEVs obtained by NTA. C) WB analysis of standard EV markers CD63, CD9 and CD81 (full images in Fig. S2). D) TEM images of isolated sEVs and mEVs from MDA-MB-231 cells.

3.2 Proteomic Profiling of Cell Line-Derived EVs

Next, the proteomes of sEVs and mEVs from all three cell lines were compared. EVs of each cell line were prepared from twelve biological replicates for protein identification and quantitative analysis. Total protein content of each isolation, about 5 μg, was analyzed by nLC-MS/MS (Fig. S1).

A total of 2,459 proteins were identified with at least one unique peptide in two or more replicates in mEVs and sEVs from any of the three cell lines after removing contaminant proteins. Proteins were mapped to gene symbols and compared with existing proteomics databases for EVs. Vesiclepedia (last updated 2018) and ExoCarta (last updated 2015) contain proteins from 13,267 and 5,405 genes which were identified in EVs from human tissue or cell lines (Fig. 3.2A), respectively (Kalra et al. 2012; Keerthikumar et al. 2016). In addition, the Top 100 EV protein list was included, which consists of the most frequently identified EV proteins. All proteins have been identified, whereas 125 proteins have not been reported in any EV database.

On average, 1,483 proteins were identified per sample with higher numbers in sEVs (Fig. 3.2B). In sEVs and mEVs, 1315 and 1128 proteins were common to all cell lines (3.2C & D) and EV markers CD63, CD9 and CD81 probed by WB analysis were identified in sEVs and mEVs from each cell line. Combined, sEVs and mEVs shared 1928 proteins (Fig. 3.2E). To compare all cell line fractions, PCA was conducted based on scaled LFQ intensity (Fig. 3.2F). Principal components 1 and 2 distinctly separate sEVs from all three cell lines, while mEVs from MCF7 and MCF10A were not separated.

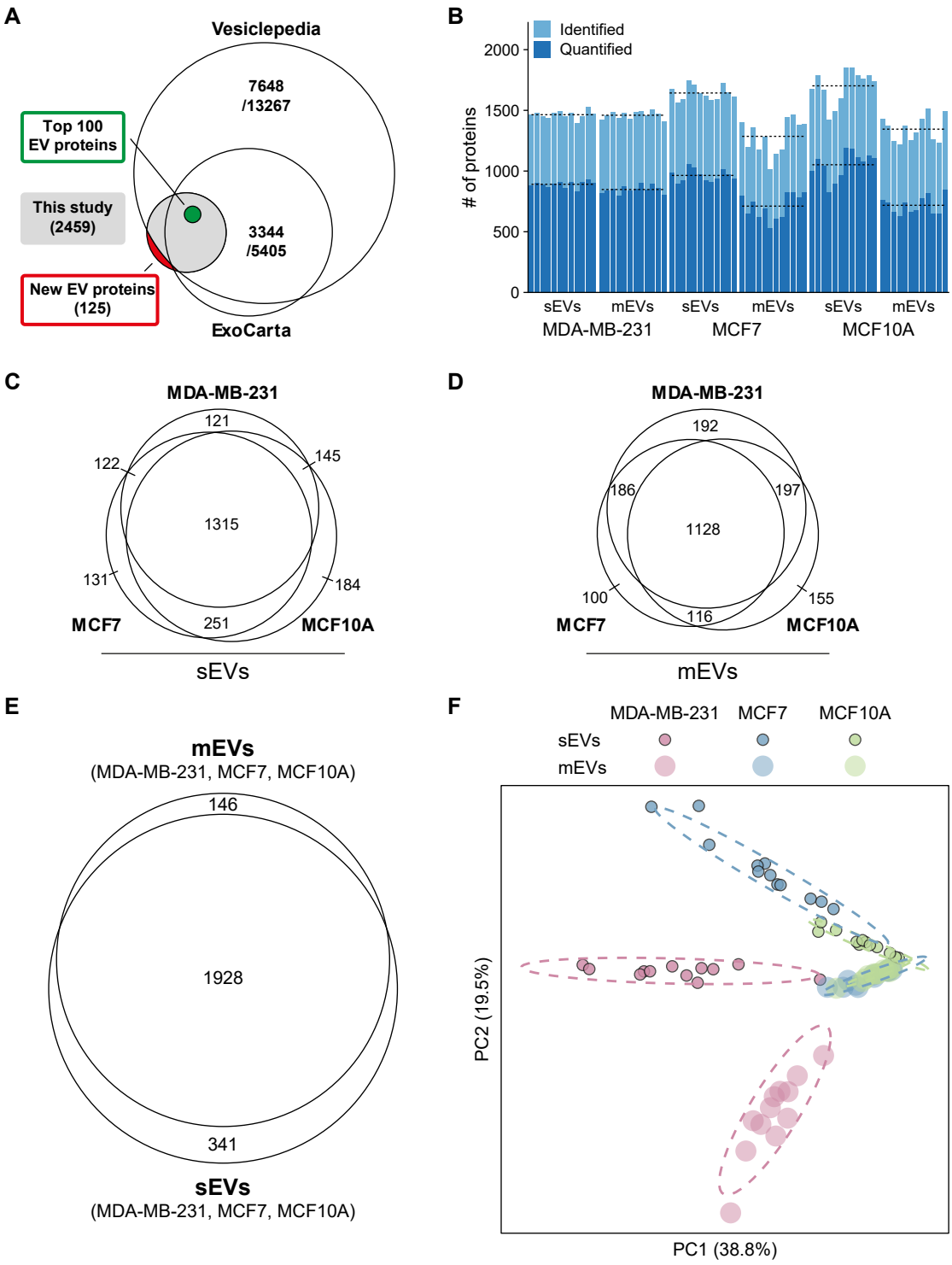

Fig. 3.2 Protein profiling overview. A) Venn diagram representing numbers of proteins in EV databases. B) Identified and quantified proteins per replicate. Dashed lines indicate average numbers per group. C) Overlap between identified proteins in sEVs and D) in mEVs. E) Number of identified proteins in sEVs and mEVs. F) PCA plot of all samples.

3.3 Selection Criteria for EV Breast Cancer Markers

The proteomic patterns of sEVS and mEVs from three breast cell lines were analyzed to derive a list of proteins that are present on the surface of EVs which may have diagnostic value for BC detection. To satisfy these criteria, a filtering strategy was formulated based on differences between cell lines, protein annotations, and a prediction of which proteins were likely to co-localize with known EV markers (Fig. 3.3).

The observed enrichment of EV marker proteins in sEVs shifted the focus of the following analysis to proteins identified in sEVs. Initially, sEV proteins were classified into positive (only identified in MDA-MB-231, MCF7, or both sEVs) and negative markers (only identified in MCF10A sEVs) based on Venn diagrams (Fig. 3.2C). Since the detected abundance should be consistent across samples containing sEVs, proteins needed to be identified by two or more peptides in all respective biological replicates, while identification of no peptide was required in sEVs and mEVs of the control cell lines. This resulted in a list of 44 proteins (Fig. S4).

Next, UniProt annotations for membrane proteins were used to select proteins accessible on the surface of EVs (Tab. S1). The final list of potential EV BC surface biomarkers contains 25 proteins. To predict co-localization of remaining 25 membrane proteins with EV marker proteins, correlation matrices were constructed for sEVs of each cell (Fig. S5). Pairwise Pearson's correlation coefficients were computed from protein abundances from 12 biological replicates and hierarchical clustering was per-

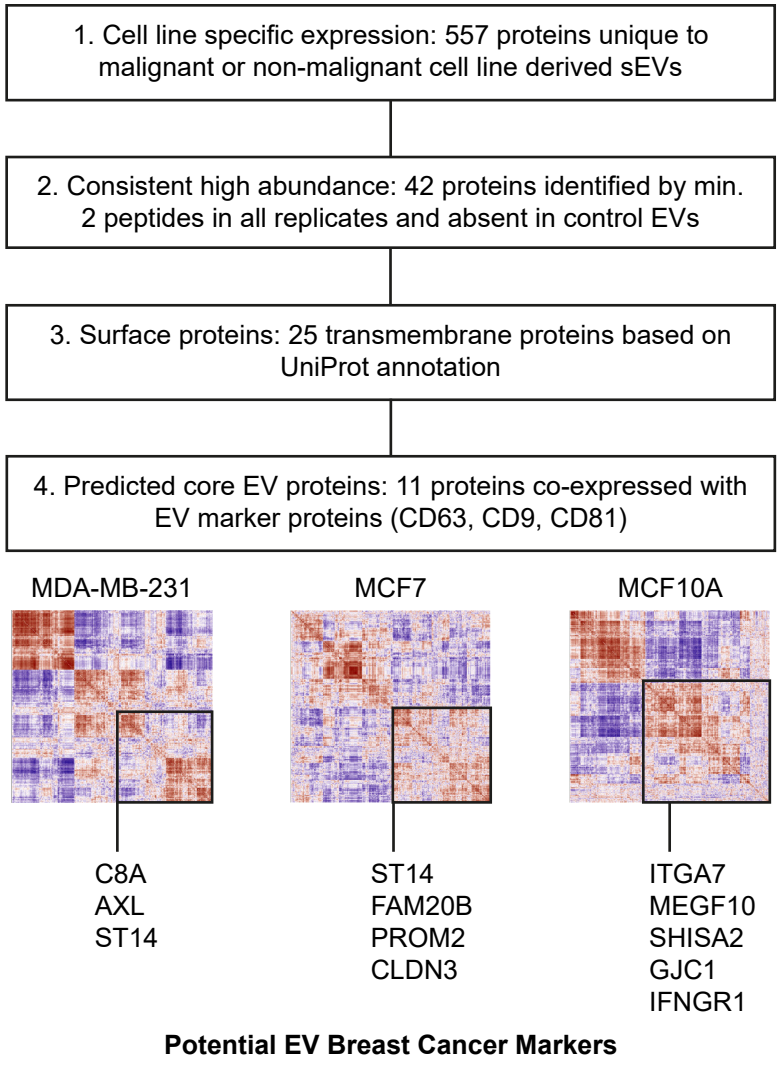

Fig. 3.3 Biomarker selection strategy. 1. Proteins uniquely identified in sEVs for each cell and common to MDA-MB-231 and MCF7 were selected (Fig. 3.3c). 2. Abundance by at least 2 peptides in all replicates and absence of any peptides in control cell lines was required (Fig. S3). 3. Proteins were classified by UniProt subcellular locations to identify membrane proteins (Fig. S3). 4. Only proteins with predicted co-localization with EV marker based on correlation matrices were used as potential biomarkers (Fig. S4).

formed on the resulting distance matrix to sort proteins into clusters of co-expression. Similar approaches have been used in the context of cell types in tissue samples, where

measured protein abundances or gene counts mainly depend on the overarching composition of subgroups like cell types in tissue samples, organelles within cells, and EV subtypes within isolated pellets (Itzhak et al. 2016; Tsoucas et al. 2019). clusters from the correlation matrices were then used to predict protein co-localization on vesicle subtypes (Fig. S5 similar to a recently published approach (Martin-Jaular et al. 2020, 2021). Finally, three, four, and five proteins were found within the same major cluster with EV marker proteins (CD63, CD9, CD81) in MDA-MB-231, MCF7 and MCF10A-derived sEVs, respectively (Tab. 3.1). An investigative literature search of

Table 3.1 List of cell line-specific sEV surface proteins proposed as biomarkers for BC.

Protein accession	Gene name	Protein name
MDA-MB-231 sEVs		
P07357	C8A	Complement C8 alpha chain
P30530	AXL	AXL receptor tyrosine kinase
MDA-MB-231 & MCF7 sEVs		
Q9Y5Y6	ST14	ST14 transmembrane serine protease matriptase
MCF7 sEVs		
O75063	FAM20B	FAM20B glycosaminoglycan xylosylkinase
Q8N271	PROM2	Prominin 2
O15551	CLDN3	Claudin 3
Q13683	ITGA7	Integrin subunit alpha 7
MCF10A sEVs		
Q96KG7	MEGF10	Multiple EGF like domains 10
Q6UWI4	SHISA2	Shisa family member 2
P36383	GJC1	Gap junction protein gamma 1
P15260	IFNGR1	Interferon gamma receptor 1

the working list of identified potential biomarkers revealed a subset of proteins with known relevance in cancer and BC, that have not been studied in the context of EVs: Suppressor of tumorigenicity 14 protein (ST14; MDA-MB-231 and MCF7 sEV-specific), Claudin-3 (CLDN3; MCF7 sEV-specific), Tyrosine-protein kinase receptor

UFO (AXL - MDA-MB-231 sEV-specific), and Integrin alpha-7 (ITGA7; MCF10A sEV-specific). Presence and and possibility of immunodetection of these proteins was probed by WB analysis in sEVs and mEVs from each cell line (Fig. 3.4 and full images in Fig. S6 and S7). AXL was not detected by WB in MDA-MB-231 EVs as found

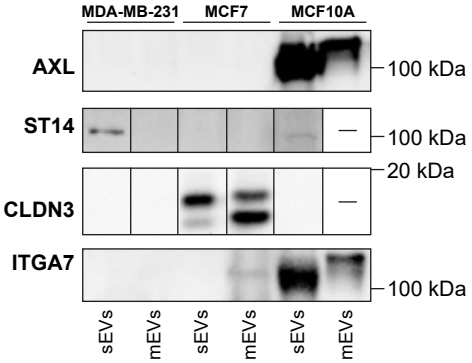

Fig. 3.4 WB analysis of potential BC EV markers: AXL, ST14, CLDN3 and ITGA7.

by MS, but was observed in MCF10A sEVs and mEVs. The presence of ST14 was detected by WB in sEVs from MDA-MB-231 and to lesser extent in MCF10A, but not in MCF7. It was fully absent in mEVs. This is in contrast with the proteomics results which identified ST14 in MDA-MB-231 and MCF7 sEVs. CLDN3 was only identified in sEVs and mEVs from MCF7, which is in agreement with the proteomics data. This result suggest the presence of two isoforms in sEVs and mEVs with approximate masses of 15.8 and 17.7 kDa. ITGA7 was identified only in MCF10A sEVs by proteomic analysis; notably, the full protein was found in MCF10A mEVs and the heavy chain with a lower mass was observed in MCF10A sEVs and a faint band in MCF7 mEVs. The full WB image shows that the 70 kDa form and the light chain of ITGA7 were identified in MCF7 mEVs as well. The different proteoforms of CLDN3 and

ITGA7 in different EV types exemplifies the heterogeneity of EVs and furthermore suggest the potential utility of specific proteoforms for diagnostic purposes.

3.4 Identification of Surface Proteins on EVs

Following the selection of potential EV biomarker proteins based on annotated transmembrane proteins, a second approach aims to experimentally identify proteins accessible from the surface of EVs using biotin labelling. A commonly used reagent, Sulfo-NHS-SS-Biotin, was used to react with free lysine residues and N-termini on the surface of EVs. Chemical modification of surface proteins with a biotin tag allows for affinity separation of labelled and unlabelled proteins, and subsequent MS analysis (Fig. 3.5). The labelling procedure of surface proteins on EVs has been incorporated into the ultracentrifugation protocol to remove free biotin, which would otherwise interfere with affinity enrichment. Isolated EVs were then lysed and incubated on a streptavidin-affinity column. Unlabelled proteins were washed, labelled proteins were retained, then finally eluted from the column by cleavage of the disulfide bond of the biotin label with TCEP. The trypsin digestion of both protein fractions was supplemented with chymotrypsin, to account for missing cleavage at labelled lysine residues. Both EV pellets, sEVs and mEVs, were prepared in biological triplicates from MDA-MB-231, MCF7, and MCF10A cells.

In total, 846 proteins were found on the surface of EVs. A comparison with EV databases revealed 154 proteins not yet reported in EV studies (Fig. 3.6A). Identified proteins in this experiment common to the Top 100 EV proteins include tetraspanins (CD63, CD9, CD81), 14-3-3 adapter proteins and heat shock proteins. On the contrary, proteins such as histones and proteins of the Rab GTPase family are not found to be located on the surface of EVs and therefore not accessible.

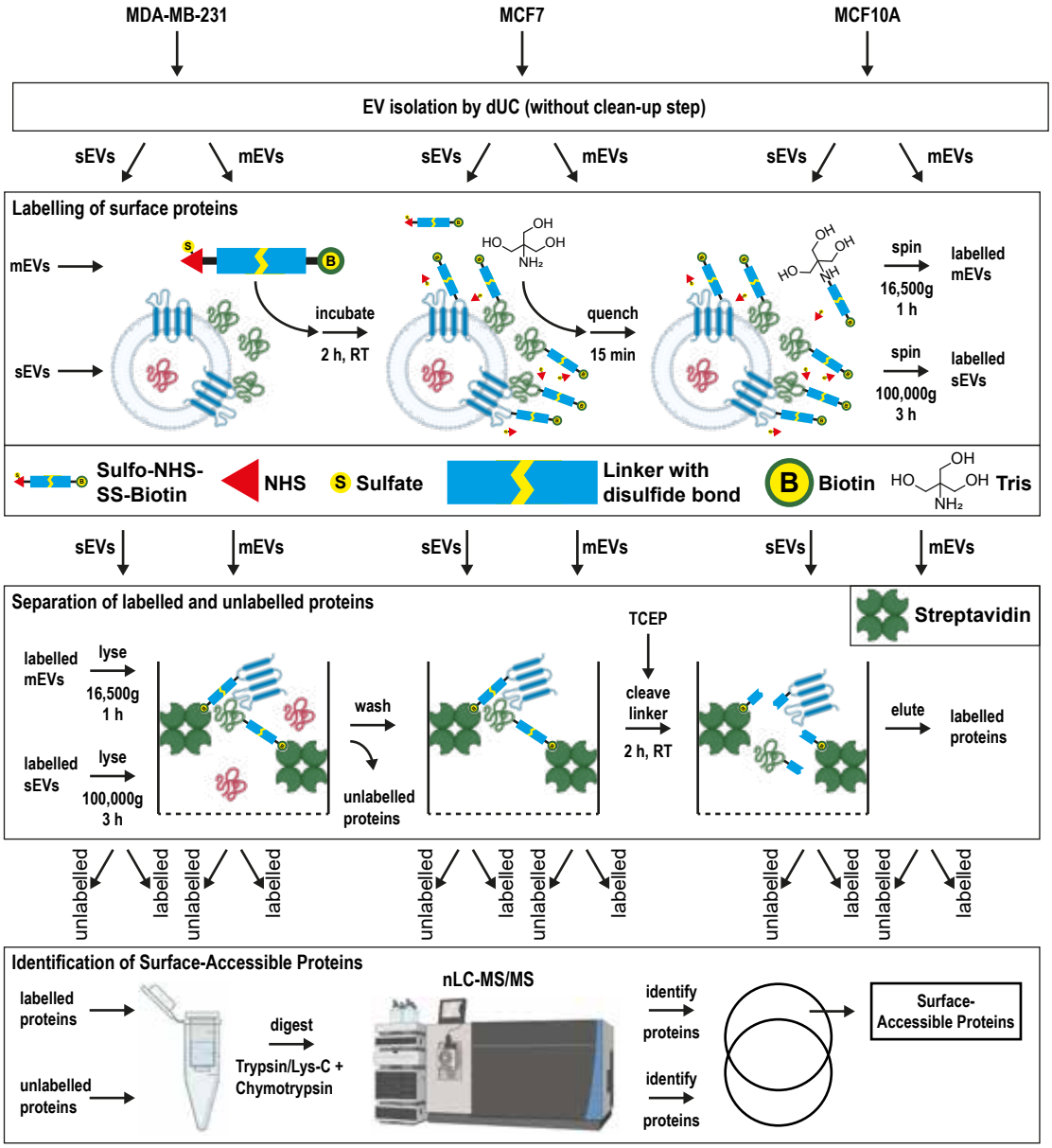

Fig. 3.5 Schematic workflow for identification of surface-accessible proteins on EVs. EVs were isolated by dUC as previously described, without resuspension and clean-up of the pellets. Surface proteins on EVs were then labelled using an amine-reactive Sulfo-NHS-SS-Biotin linker which was quenched by adding Tris. Free quenched biotin linker groups were removed by ultracentrifugation. Labelled surface proteins were enriched on a streptavidin-affinity column and eluted by cleaving the disulfide bond of the biotin linker. Both washed and eluted fractions were analyzed by MS. Proteins only identified in the eluted fraction were considered surface accessible.

Surface proteins identified by the labelling approach display a considerable number of proteins not identified by the regular proteomics approach (Fig. 3.6B). Many of these proteins may be underrepresented in global proteomic analysis, hence are enriched as the subset of surface proteins or identified due to different preparation conditions. Furthermore, the small overlap of proteins with the previous marker selection that was based on total proteomic analysis prompted independent analysis of these results.

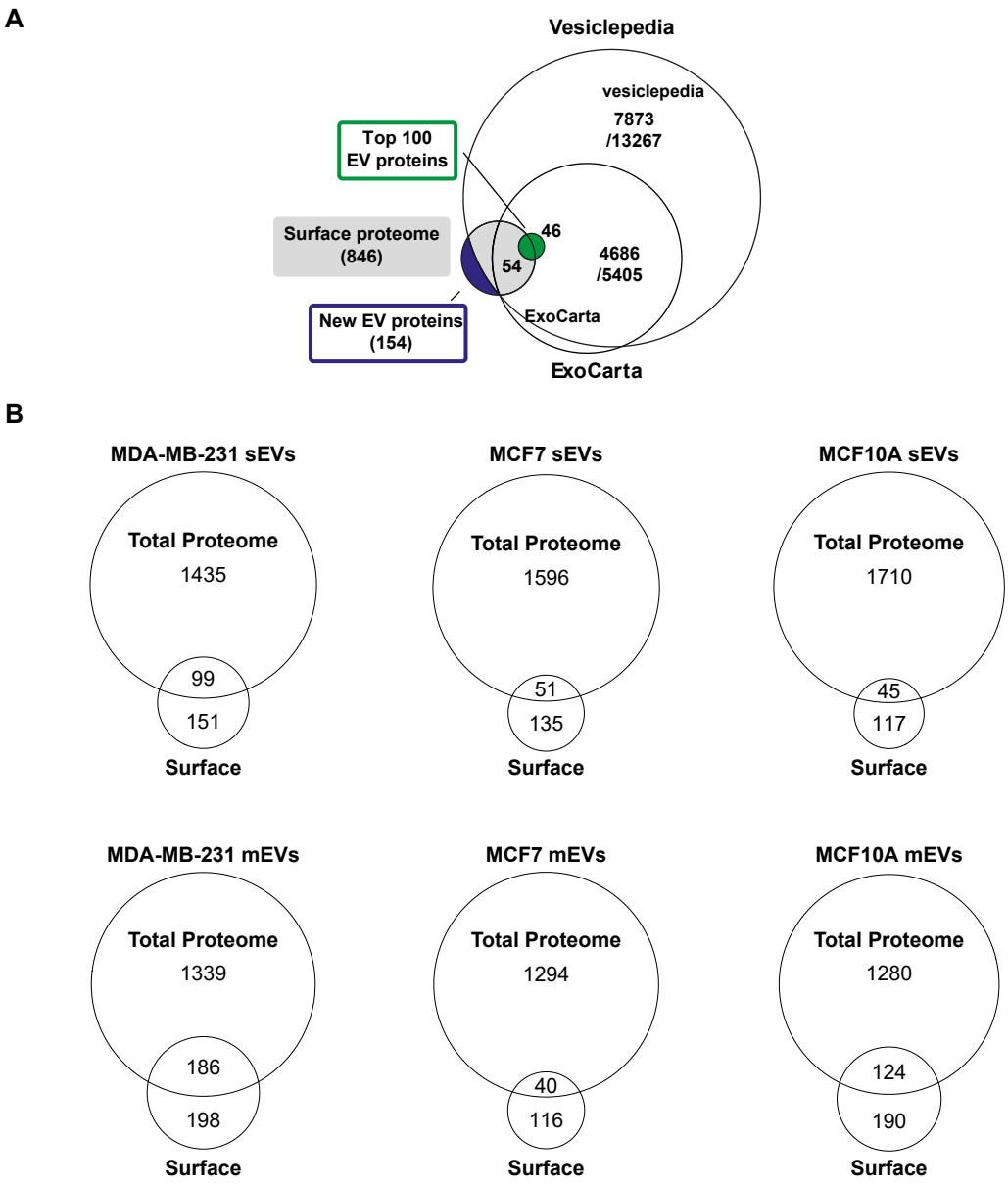

Fig. 3.6 An overview and comparison of the number of identified surface proteins on EVs from cell lines. A) Comparison of all identified surface proteins from three cell lines and two EV fractions with EV databases. B) Comparison of identified proteins from sEVs and mEVs from each cell line from previous total proteomics analysis with surface proteins identified by surface labelling and MS analysis.

Based on the experimentally identified surface proteins, proteins were compared be-

tween cell lines and compared to known disease-related proteins. With the goal of selecting general BC protein markers, the focus was first on finding which proteins were common to malignant cell lines. Surface proteins of sEVs and mEVs were separately compared between cell lines, showing a large overlap between MDA-MB-231 and MCF7 sEVs (Fig. 3.7A) and MDA-MB-231 and MCF10A mEVs (Fig. 3.7B).

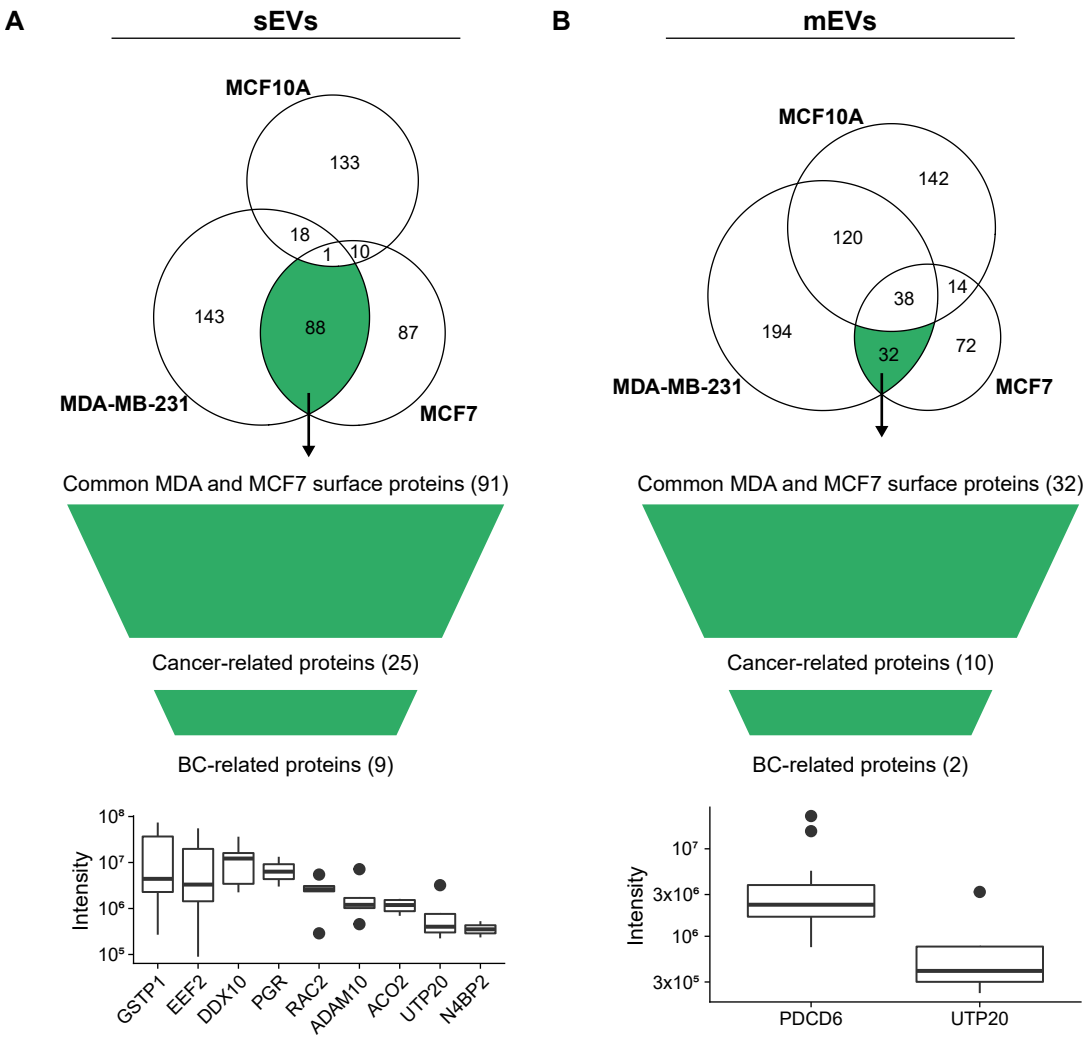

Fig. 3.7 Biomarker selection based on biotin-accessible surface proteins on A) sEVs and B) mEVs. Common surface proteins from MDA and MCF7, not identified in MCF10A sEVs, were annotated with known disease associations from the DisGeNET database and subsequently, proteins associated with cancer and breast cancer were selected. The final list of potential BC EV surface markers were ranked based on LFQ intensity to prioritize more abundant proteins.

Next, 88 and 32 proteins common to sEVs and mEVs from MDA-MB-231 and MCF7 and absent in MCF10A (Tab. S2 and S3), respectively, were compared with the Dis-

GeNET database for disease associations (Piñero et al. 2020). Among the 88 and 32 identified surface proteins, 25 and 10 were annotated to several cancers. Furthermore, 9 and 2 proteins in sEVs and mEVS had known implications in BC, respectively. These proteins were ranked based on their intensity to approximate the abundance of proteins on the surface of EVs, important for diagnostic approaches. The 4 most

Table 3.2 List of surface proteins common to MDA-MB-231 and MCF7 sEVS and mEVs with known BC-association.

Protein accession	Gene name	Protein name
sEVs		
P09211	GSTP1	Glutathione S-transferase pi 1
P13639	EEF2	Eukaryotic translation elongation factor 2
Q13206	DDX10	DEAD-box helicase 10
P06401	PGR	Progesterone receptor
P15153	RAC2	Rac family small GTPase 2
O14672	ADAM10	ADAM metallopeptidase domain 10
Q99798	ACO2	Aconitase 2
O75691	UTP20	UTP20 small subunit processome component
Q86UW6	N4BP2	NEDD4 binding protein 2
mEVs		
O75340	PDCD6	Programmed cell death 6
O75691	UTP20	UTP20 small subunit processome component

abundant biotin-accessible surface proteins common to sEVs from MDA-MB-231 and MCF7 (GSTP1, EEF2, DOX10, PGR) and two surface proteins common to mEVS from MDA-MB-231 and MCF7 (PDCD6, UTP20) can be explored for their potential as BC biomarkers in future studies. The experimental evidence of their surface accessibility increases the possibility that EVs carrying these proteins can be isolated or directly detected from body fluid samples of patients.

Chapter 4

Discussion

4.1 EV Isolation

Currently, EV isolation methods enrich EVs from biological matrices such as cell media or blood plasma based on size, density, or presence of biological molecules, commonly surface proteins (Brennan et al. 2020; Patel et al. 2019). Ideally, isolation methods are designed to separate EVs from whole cells, cellular debris and soluble proteins. However, available isolation protocols fail to fully discriminate between EVs of different cellular origins, such as exosomes and microvesicles, due to overlap and similarities in size, density, and biochemical composition (Doyle and Wang 2019; Willms et al. 2018).

EV isolation techniques that are based on affinity enrichment offer great potential because they allow for discrimination between EV subtypes, mainly exosomes and microvesicles, by probing for proteins on the surface specific to each subgroup (Bel-

lotti et al. 2021; Nakai et al. 2016). However, the surface proteome of EVs is not yet well understood, both within and between EV subgroups; this may introduce biases to analyses. Moreover, future studies may benefit from more sensitive diagnostic tests as a result of being able to specifically isolate known EV subtypes, such as exosomes or organ-derived EVs from complex biofluids (Sayyadi et al. 2021). Filtration-based isolation methods are typically combined with dUC protocols to achieve EV isolates with defined size ranges (J. Li et al. 2019), however there is generally a low recovery of EVs from filter membranes (Martin-Jaular et al. 2020). Polymer-based EV precipitation methods have been found to preserve the biological activity of EVs (P. N. Brown and Yin 2017), but insufficient removal of isolation reagent might interfere with MS analysis and only yields low numbers of detected proteins (Risha et al. 2020). Other methods such as density gradient isolation, flow cytometry and flow field-flow fractionation (AF4) hold great potential in defining or discovering EV subtypes and subpopulations, but rely on initial EV enrichment by dUC (Deville et al. 2021; Onódi et al. 2018; H. Zhang and Lyden 2019). Considering all factors, dUC is the most suitable method for isolation of EVs that will have downstream applications in proteomic studies (Risha et al. 2020).

The isolation of EVs by dUC achieves a sufficient yield and purity of EVs for proteomic analysis (J. Li et al. 2019). In addition to being the most commonly used technique for EV isolation (Willms et al. 2018), it proved useful in the past to define the currently used EV markers CD63, CD9, and CD81 (Kowal et al. 2016; Livshits et al. 2015). The originally published dUC protocols have been modified over the years

by many groups with the intentions of improving EV yield, sample processing time and specificity for EV subtypes (Théry et al. 2006). Likewise, the procedure used in previous studies of BC cell line-derived EVs (Risha et al. 2021, 2020) has been modified.

4.2 EV Characterization and Sample Preparation

Isolated vesicles of the 16,500 g and 100,000 g pellets were considered to be mEVs and sEVs, respectively, as suggested by MISEV guidelines to highlight the isolation of vesicles by dUC (Théry et al. 2018). Notably, centrifugation separates particles based on their mass, density and shape. In contrast, particles separated by density gradients are commonly characterized by their density. Both sEVs and mEVs have been previously found to be enriched in proteins related to exosomes and microvesicles, respectively (Doyle and Wang 2019). Characterization of isolated EV pellets by NTA confirmed an overlap in the size distributions of vesicles between sEVs and mEVs in all cell lines that were studied. Western blot analysis identified EV markers, sometimes referred to as exosome markers, CD63, CD9, and CD81 exclusively in sEVS of all cell lines, except the presence of CD63 in mEVs from MCF7 cells. Due to the higher sensitivity of proteomic analysis, these proteins were subsequently identified in all fractions due to higher sensitivity.

Following EV isolation by dUC, the protein composition was profiled by MS-based proteomics. In contrast to previous proteomics studies of BC cell line-derived EVs, the in-solution sample preparation protocol (Risha et al. 2021, 2020) has been modified to

a FASP protocol to remove interfering molecules (e.g. metabolites) during proteolytic digestion and MS analysis (Wiśniewski et al. 2009). The observed increase in the number of proteins identified per sample may be in part due to improved sample preparation methods, however it is known that the cell culturing conditions and chosen isolation protocol contribute to observed difference in the types and quantity of yielded protein (Zaborowski et al. 2015). Ultimately, the increased number of identified proteins per sample led to a higher number of commonly identified proteins between cell lines and EV fractions. Despite controlling for the quantity of protein for all cell lines and fractions, sEVs from MDA-MB-231 cells gave rise to a lower number of distinct proteins per sample compared to MCF7 and MCF10A sEV. This has been previously observed in another study for TNBC cell lines including MDA-MB-231 (Rontogianni et al. 2019). Furthermore, the number of identified proteins between EV fraction of each cell line only differed between MCF7 and MCF10A.

4.3 Protein Classification Methods for Selection of Potential EV biomarkers

Previous BC biomarker studies of cell line-derived EVs simply compared the proteomic composition of EVs from malignant MDA-MB-231 and control MCF10A cells (Risha et al. 2021, 2020). Potential marker proteins were selected based on cell line-specific identification of proteins, UniProt annotation of membrane proteins and known disease implications. Another study evaluated ten cell lines corresponding to

four groups of cell lines: TNBC, HER2-positive, ER/PR-positive, and control epithelial cells (Rontogianni et al. 2019) The proteomic analysis revealed the utility of BC cell line-derived sEVs to discriminate between BC subtypes. In this study, control breast epithelial cells MCF10A have been compared with two malignant cell lines MDA-MB-231 (TNBC) and MCF7 (ER/PR-positive) to account for subtype specific differences and to eventually identify general BC biomarkers.

The selection strategy included four steps leading to 12 potential marker proteins from sEVs of which four have been further investigated. Initially, cell line specific expression of proteins was used to find BC specific proteins. A common approach in proteomic studies is to primarily compare groups like control and case based on the relative abundance of proteins and impute missing quantitative values under the assumption that unidentified proteins fall under the limit of detection while still being present (L. Jin et al. 2021). Here, the absence of a protein was assumed as missing not at random (MNAR) and therefore considered informative for the analysis. Four subgroups of proteins were defined based on Venn diagram representation of identified proteins: Only identified in MDA-MB-231 sEVs, only identified in MCF7 sEVs, only identified in in MDA-MB-231 and MCF7, and only identified in MCF10A. Proteins identified in malignant cell line-derived EVs may serve as BC biomarkers and proteins only identified in control cell line EVs, potential tumor suppressor genes, were included as a control for following experiments.

The 557 EV specific proteins were identified by at least one unique peptide in two or more replicates per cell line EV fraction, including identification of low abundant

proteins. Therefore, the list was further restricted to only include proteins identified in all 12 replicates, by at least two unique peptides, in respective EV fractions. In addition, the condition for a protein not to be identified in a cell lines was extended to no peptide in any fraction of sEVs and mEVs. The purpose of these restrictions was to ensure a consistent presence of proteins detected from EVs and that low abundant protein identifications were accounted for; this resulted in 42 cell line specific sEV proteins (Fig. S4 and Tab. S1).

Surface proteins on EVs were predicted using a previously described approach by using UniProt subcellular location annotations (Risha et al. 2020), which resulted in 25 membrane proteins. Transmembrane proteins were assumed to expose an extracellular domain on the surface of EVs. It has been shown that some proteins exhibit an inverse orientation of the N- and C-terminus in EVs, compared to cells (Cvjetkovic et al. 2016). Furthermore, many proteins on the surface of EVs bind non-covalently to exposed proteins on EVs which creates a protein corona that further modifies the surface (Tóth et al. 2021). This challenge is addressed in the second selection approach: the biotinylation of surface proteins.

Until now, all proteins identified from the dUC isolated pellets were considered EV proteins. Still, the origin of each protein cannot be confirmed unless experimentally verified by a single-vesicle method, such as flow cytometry (Nolan and Duggan 2018). One method of predicting the cellular origin of a protein based on fractionation is referred to as protein correlation profiling, or more generally, correlation analysis. There are many applications of correlation analysis, such as the identification of co-

regulated gene modules (Langfelder and Horvath 2008), co-expressed plasma proteins (Wewer Albrechtsen et al. 2018) or organellar maps of proteins (Itzhak et al. 2016) and recently for the composition of EVs (Martin-Jaular et al. 2020, 2021). The analysis of protein localization by organellar maps and representation as PCA plots, that has recently been applied to the study of EVs, relies on fractionation of the biological sample (Martin-Jaular et al. 2021). EVs have been collected from 10,000g, 30,000g, and the 100,000g pellets and were analyzed by MS-based proteomics. Each fraction is then normalized to the total amount of proteins of each fraction. PCA was then applied to the protein levels and known protein markers of cellular organelles in the PCA plot are then used as a reference for a support-vector machine (SVM) classification. This results in a two-dimensional plot of all proteins, in which groups of proteins are annotated to specific cellular organelles or origins by the SVM. This approach has been applied to detect changes in the location of single proteins (Itzhak et al. 2016) and alterations in the composition of EVs upon HIV-1 infection of cells (Martin-Jaular et al. 2021). Here, correlation matrices were constructed from the LFQ intensity of each protein from 12 replicates for both EV fractions, sEVs and mEVs, per cell line separately. Pairwise correlation coefficients between all proteins were calculated and the resulting correlation matrix was subjected to hierarchical clustering. Within the ordered correlation matrix, proteins in the same clusters were considered to be co-expressed and assumed to be co-localized on the same EV types. Known tetraspanin marker proteins CD63, CD9 and CD81 were used to indicate clusters of "core EV proteins," which was compared with the list of potential BC markers. 14 proteins were predicted to be co-localized with EV markers and considered useful

for further investigation.

The PCA and SVM-based generation of organellar maps may provide a more detailed classification by taking different fractions into account. The collection and analysis of at least three EV fractions is often not a reasonable computational problem, however, the method presented here allows for the identification of co-localized proteins based on correlation analysis from only one type of fraction.

The interaction of EV membrane proteins with other proteins that ultimately forms a protein corona (Tóth et al. 2021), poses a challenge for the direct detection of EV proteins or enrichment of EVs positive markers through affinity-based approaches (Palviainen et al. 2020). To gain experimental evidence on EV proteins accessible from the surface, a biotinylation protocol was adopted to identify and enrich such proteins. A biotinylation reagent, Sulfo-NHS-SS-Biotin, has been used in many studies to identify surface proteins on cell and EVs (Henry et al. 2018; M. Li et al. 2021; Cvjetkovic et al. 2016; Rai et al. 2021).

The charged sulfonate group of the crosslinking reagent allows for the solubilization in water, while being membrane impermeable. The N-succinimidyl group reacts with primary amines from proteins and other molecules, labelling accessible lysine residues and N-termini. The biotin group enables a separation of labelled and unlabelled proteins on a streptavidin affinity column, since labelled proteins can be eluted by a reaction that induces cleavage of the disulfide bond within the linker. Prior to proteomics analysis, both washed and eluted proteins were digested with trypsin/Lys-C and chymotrypsin. Chymotrypsin cleaves proteins at the carboxyl side of tyrosine,

phenylalanine, tryptophan, and leucine residue and was used to overcome incomplete digestion of labelled proteins, as modified lysine residues did not serve as cleavage sites to trypsin or Lys-C anymore. Proteins were then identified by MS analysis and classified into surface accessible or non-accessible, depending if they were identified in the unbound or biotinylated fraction. The rigorous washing of the column resulted in a strong, nearly complete separation between labelled and unlabelled proteins, therefore relative quantification in washed and eluted fraction as proposed in other studies was not applicable (Keilhauer, Hein, and Mann 2015).

Following comparison of surface proteins among cell lines and EV types, common surface proteins between sEVs and mEVs from MDA-MB-231 and MCF7 were selected as potential markers of BC that could be directly detected on the surface of cancerous EVs. Known protein-disease annotations from the DisGeNET database were then retrieved to propose BC surface EV markers (Piñero et al. 2020). Finally, 9 proteins GSTP1, EEF2, DDX10, PGR, RAC2, ADAM10, ACO2, UTP20, N4BP2 and 2 proteins PDCD6, UTP20 were obtained by the selection approach described above in sEVS and mEVs, respectively. The 9 proteins in sEVS were then ranked by their detected intensity with the intention of selecting more abundant proteins. This approach, as it selects already known disease-associated proteins, does not discover new proteins involved in BC, however here the translation of known proteins into EV biomarkers was anticipated.

The proteins identified from both strategies resulted in relatively small overlap, therefore two separate biomarker selection strategies were considered. The first approach

predicted transmembrane proteins which were then assumed to be surface proteins, whereas the final list of the second approach only contained two membrane proteins, PGR and ADAM10. This suggests that the majority of directly accessible proteins on EVs may not be directly associated to the lipid bilayer (Rai et al. 2021). Other studies have used different approaches as well, such as digestion of surface proteins by Proteinase K (Cvjetkovic et al. 2016) and subsequent MS analysis of remaining proteins or EDTA treatment of isolated EVs prior to labelling with a biotin reagent (Tóth et al. 2021). Both approaches may yield a different surface profile for EVs, however it is not evident that all proteins may then be detectable under native conditions.

4.4 Initial Assessment of Potential EV Surface Biomarkers for Non-Invasive Breast Cancer Diagnosis

The final panel of potential biomarkers obtained through the first selection approach include the following proteins: ST14, CLDN3, AXL, ITGA7.

AXL has been identified by proteomics as an MDA-specific protein. The results could not be validated by WB analysis, thus the experiment will be repeated with a different antibody clone. Overexpression of AXL has been reported in many different cancers, including breast, ovarian, prostate, non-small cell lung cancer and head and neck squamous cell carcinoma (Gay, Balaji, and Byers 2017). Activation of the AXL signalling pathway promotes cell proliferation, migration, invasion, epithelial-

mesenchymal transition (EMT) and tumor angiogenesis (Gjerdrum et al. 2010). It has been shown that AXL is associated with multiple receptor tyrosine kinases, such as EGFR (Debruyne et al. 2016), HER2 (Guo et al. 2017), ALK (Liu et al. 2009), and VEGFR (X. Wu et al. 2015) resulting in activation of these receptor tyrosine kinases in a ligand-independent manner, which induces resistance to targeted therapies. Activation of AXL signalling has an immunosuppressive effect, which in turn reduces the antitumor immune response (Z. Zhang et al. 2012; Zhou et al. 2016). Quantitative phosphoproteomics of a panel of TNBC cell lines discovered that AXL was overexpressed and activated in ~40% of TNBC cell lines associated with high invasion and strong anti-aniokis oncogenic phenotypes (Aguilera et al. 2016). It has been shown that higher expression of AXL was significantly associated with poor prognosis of patients with TNBC and humanized anti-AXL monoclonal antibody was able to inhibit oncogenic phenotypes in vitro and reduces tumor formation in xenograft mouse models (Aguilera et al. 2016). Although, AXL is considered a predictive prognosis biomarker for breast cancer, it may not be suitable as a EV biomarker, as proteomics and WB analysis gave contradictory results.

The suppressor of tumorigenicity 14 protein (ST14), also called Matriptase, is a member of the transmembrane protease serine 3 (tmprss3) family (Alef et al. 2009). ST14 has been identified by proteomics in sEVs from MDA and MCF7 cells, and WB found expression in MDA sEVs and to less extent in MCF10A sEVs. The single-pass type II transmembrane protein plays an essential role in extracellular matrix degradation through its trypsin-like activity and is involved in formation and regulation of

oral epithelium, epidermis, hair follicles, and thymic epithelium (Arena et al. 2017; Bergum et al. 2012; Bhatt et al. 2003; Dai et al. 2021; Ge et al. 2006). It is speculated to interact with other proteins to regulate its activity and cellular localization, and has been reported to be anchored to the actin cytoskeleton in cells (J. S. Jin et al. 2007). ST14 was initially discovered in breast cancer cell lines, high protein expression has been shown in EMT from malignant cells isolated from human breast carcinomas (Petri and Klinge 2020; C. Kim et al. 2005; K. Y. Kim et al. 2019; List et al. 2006). Reduced expression of ST14 slowed tumor growth in mice by impairment of pro-HGF/c-Met signalling pathway and cell proliferation (List et al. 2003). Similar results were obtained by targeting ST14 with peptide epitopes derived from its autocatalytic loop and monoclonal antibodies, suggesting it may be suitable as a pharmaceutical target (List et al. 2002). Finally, ST14 gene methylation is predictive for survival rates in breast cancer patients (Petri and Klinge 2020). The presence of ST14 in sEVs from MDA sEVs may affect its activity due to different interaction partner and transfer the protein to distant sites, where it may be involved in metastatic niche formation (Takeuchi et al. 2000). This suggests that ST14 contributors to the transfer of malignant traits in metastatic BC (Zoratti et al. 2015).

Claudins are tight junction proteins of 20-24 kDa in mass, that are present on the apicolaterla membranes of epithelial, endothelial and mesothelial cells (Ylermi Soini 2011; Osanai et al. 2017). In malignancies, the expression of claudins varies depending on the type of the tumor (Y. Soini 2005). For example, overexpression of claudins 3 and 4 has been found in several carcinomas, including breast cancer (Y. Soini 2005).

In MCF7 cells, inhibiting CLDN3 overexpression decreased cell migration (Todd et al. 2015). Recent findings indicate an association between the outcome of BC patients and expressions of claudins 3, 4 and 7 in membranes and cytoplasm (Jääskeläinen et al. 2018). Moreover, CLDN3 expression is positively correlated with BRCA mutations in women (Danzinger et al. 2019). Proteomic analysis of human bile-derived EVs has been performed to identify novel biomarkers for cholangiocarcinoma (CCA) (Ikeda et al. 2021). From 166 CCA-specific proteins, ELISA for CLDN3 exhibited statistically significant differences and high diagnostic accuracy. In addition, a proteomics study of sEVs from malignant PC3 and benign PNT1A cells, combined with subsequent ranking of proteins specific to malignant cells based on public databases identified CLDN3 as a putative prostate cancer biomarker (Worst et al. 2017). Here, it was demonstrated that CLDN3 could be used as an exemplary EV-based biomarker for BC. Still, further analysis is needed to determine if CLDN3 is detectable in biofluid from BC patients, and if so, the degree of diagnostic value.

Beyond cell line-specific expression of CLDN3 in the context of BC, WB analysis revealed the presence of two isoforms of CLDN3 in MCF7 EVs. Both isoforms were identified in sEVs and mEVs, suggesting specificity of isoforms to exosomes and microvesicles, which are enriched in sEVs and mEVs, respectively. Most claudins are known to express multiple isoforms (Lal-Nag and Morin 2009; Tipsmark and Madsen 2012). However, an isoform-specific protein loading has not been described for EV subtypes yet. Further investigation could shed light on the mechanisms of EV loading through sequence motifs, PTMs or PPIs.

Integrins are a major family of cell adhesion receptors which consist of 18 alpha and 8 beta subunit genes in mammalians, which form at least 24 non-covalently linked, heterodimeric complexes from an alpha and a beta subunit (Humphries, Byron, and Humphries 2006). These heterodimers are transported from the endoplasmic reticulum to the Golgi apparatus and transferred to the cell surface in an inactive state (De Franceschi et al. 2015). As a receptor on the cell membrane, integrins primarily interact with ECM components to mediate cell adhesion (Harjunpää et al. 2019). Integrins are well known constituents of EVs and have been studied in different cancers (Wanessa Fernanda Altei et al. 2021; Ferrarelli 2020). In particular, the integrin content of EVs is able to distinguished different cancers (Hurwitz and Meckes 2019) and impaired adhesion of tumor-derived EVs (Wanessa F. Altei et al. 2020). Cellular ITGA7 has been found to influence migration, invasion, and EMT to function as a tumor suppressor gene in BC (Bhandari et al. 2018) and papillary thyroid carcinoma cells (Guan et al. 2020). Furthermore, transcriptomic analysis of stem-like cells from primary BC identified ITGA7 as a predictive marker for chemotheraphy response (Gwili et al. 2021). This study indicates that ITGA7 is highly expressed in MCF10A compared to MCF7 and MDA-MB-231 derived sEVs, which proposes that ITGA7 might be used as a control biomarker for BC assessments. Furthermore, different forms of ITGA7 have been detected by WB analysis, suggesting presence of the unprocessed form in MCF10A mEVs, the heavy chain in MCF10A sEVs and to less extent in MCF7 mEVs, and the 70 kDa form and the lighjt chain in MCF7 mEVs. The presence or presence of different isoforms of ITGA7 may both be used for diagnostic purposes if isoform specific antibodies are available.

The second approach, based on surface proteomics of EVs, highlighted nine proteins with known implications in BC. These proteins were only found on the surface of EVs MDA-MB-231 and MCF7, but not on EVs from control MCF10A cells. These include Glutathione S-transferase P1 (GSTP1), Elongation factor 2 (EEF2), DEAD/H box RNA helicase (DDX10), progesterone receptor (PGR), Ras-related C3 botulinum toxin substrate 2 (RAC2) and Disintegrin and metalloproteinase domain-containing protein 10 (ADAM10). The four proteins with the highest abundance were selected as promising biomarkers for BC detection. Notably, ADAM10 has been previously identified as a potential BC biomarker (Risha et al. 2020).

GSTP1 belongs to the family of phase II metabolic enzymes and has been described to inhibit several anti-cancer drugs by conjugating them to glutathione (Coles and Kadlubar 2003). Previous studies have demonstrated the role of GSTP1 in treatment-resistance of BC (Arai et al. 2008; G. Yang et al. 2005). High GSTP1 expression has been linked to resistance to docetaxel, paclitaxel, cyclophosphamide and doxorubicin in BC patients (Arai et al. 2008; Eralp et al. 2013; F. Su et al. 2003). After treating 30 BC patients with anthracycline/taxane-based neoadjuvant chemotherapy, the levels of GSTP1 in exosomes from patients in the progressive disease/stable disease group were significantly increased compared to those in the partial response/complete response group (S. Yang et al. 2017). This suggested GSTP1 may transferred by an exosome-based mechanism and be involved in of drug resistance, making it potentially useful as a predictive marker for chemoresistance. The identification of GSTP1 on the surface of EVs, to potentially act outside the cell, could further explain the

chemoresistant activity of GSTP1-carrying exosomes.

The elongation step of mRNA translation requires substantial amounts of energy as well as eukaryotic elongation factors (eFFs) such as EEF2 (Kaul, Pattan, and Rafeequi 2011). The activity of EEF2 is regulated by its phosphorylation state by a specific kinase known as eEF2 kinase (EEF2K). It has been reported that decreased eEF2 phosphorylation, mediated by increased PP2A activity, contributes to resistance against HER2 inhibition and may pose a novel target for therapeutic intervention in HER2-positive BC (McDermott et al. 2014). In addition, a recent study demonstrated that EEF2K plays a key role in the maintenance of aggressive tumor behaviour and chemoresistance in treatment-resistant TNBC (R.-X. Wang et al. 2019). Based on this research, EEF2K silencing was proposed as a novel strategy for the treatment of TNBC. In this study, EEF2 is identified in both cancerous cell lines, MDA-MB-231 and MCF7, but not in the non-cancerous control cell line MCF10A and may serve as a diagnostic biomarker.

DEAD/H box RNA helicases are a group of regulating mRNA translation with growing knowledge about their involvement in cancer (Heerma van Voss, van Diest, and Raman 2017). Gene mutation and changes in DDX10 expression have been previously investigated in osteosarcoma, hepatocellular carcinoma, leukemia, and ovarian cancer (Gorello et al. 2013; Shi and Hao 2019; Z. Wang et al. 2015; Y. Wang et al. 2017). Reported evidence suggests that proteins of the DDX family play a crucial role in tumorigenesis of cancer cells and stem cell differentiation (Yassin et al. 2010). However, little is known about the detailed molecular mechanisms of DDX protein

regulation. Suppression of five candidate genes (DDX10, SKA3, EPHA5, CLTC and TNIK) impacted the growth and proliferation of breast cancer cells (Jiao et al. 2013). Moreover, downregulation of DDX10 in breast cancer cells led to an increased frequency in apoptosis. Therefore, the mechanism and utility of DDX10 in BC sEVs is not evident.

Estrogen receptorα (ERα or ESR1) and progesterone receptor (PGR) are crucial prognostic and predictive biomarkers in breast cancer (BC) (Kunc et al. 2021). It has been shown that ESR1 and PGR were highly expressed in MCF7, T47D and MDA-MB-361 metastatic cell lines, suggesting that these cancer cells represent models of estrogen- and progesterone-dependent cancers (Hevir et al. 2011). Moreover, high ESR1 and PGR expression levels were observed in MCF7, BT-474 and T-47D, and low ESR1 and PGR expression was found in MDA-MB-231 (J.-R. Wu et al. 2020). These results suggest that mRNA expression levels of ESR1 and PGR can be considered distinct biomarkers and essential prognostic factors for ER-positive BC. This work identified PGR on the surface of sEVs from MCF7 and MDA-MB-231 cells and suggests potential as EV biomarker for BC.

Programmed cell death protein 6 (PDCD6) is a pro-apoptotic protein contributing to T-cell receptor-, Fas-, and glucocorticoid-induced programmed cell death (K. Zhang et al. 2013; Shibata et al. 2004), as well as endoplasmic reticulum stress induced apoptosis during organ formation (Rao et al. 2004; Mahul-Mellier et al. 2006). PDCD6 expression was found to be upregulated in tumor tissue samples from lung, breast, colon cancer, and ovarian cancer, which suggested that PDCD6 may be involved in

maintenance of cellular viability (la Cour et al. 2008, 2003; Høj et al. 2009; D. Su et al. 2012). Recently, it has been demonstrated that miR-124-3p attenuated tumor metastasis by inhibiting PDCD6 expression in BC, suggesting potential as a therapeutic target in patients with advanced disease (L. Zhang et al. 2018). Another study showed that the over-expression of the same microRNA inhibited cell proliferation, migration and invasion, and induced apoptosis by suppressing PDCD6 expression in SKOV3 and OCVAR3 cells *in vitro*.

Small subunit processome component 20 (UTP20) is present in both 90S and 40S preribosome particles (Dez, Dlakić, and Tollervey 2007). It has been hypothesized to be important for the building of early 90S particle intermediates, which is suggested by proteomic and bioinformatics analyses that showed UTP20 to be heavily interconnected and in close physical proximity to components of the tUTP, UTP-B and UTP-C subunits (Pérez-Fernández et al. 2007). The UTP20 gene has been identified to be associated with colorectal cancer (L.-C. Sun and Qian 2018).

Some potential biomarkers that are proposed here have already been investigated extensively in BC studies (ST14, CLDN3, AXL), making their translation to EV based diagnostics more likely than others. Nevertheless, our understanding of EVs is far from conclusive and mainly consists of studies highlighting the importance of individual proteins in EVs or diseases (Y. Zhang et al. 2019). Therefore, functional aspects for the surface localization of some proteins can only be speculated. While the surface expression of several proteins may be the specific mode of transfer due to protein interaction partners (DDX10, EEF2, PDCD6, UTP20), some studies demon-

strated functional aspects of surface proteins on EVs such as presenting antigens to evade chemotherapy against therapeutic or endogenous antibodies (Ciravolo et al. 2012; Marleau et al. 2012) or degradation of the extracellular environment (Nawaz et al. 2018; Sanderson, Bandari, and Vlodavsky 2019), emphasizing the importance of understanding the role of EV surface proteins for diagnostic and therapeutic purposes.

Chapter 5

Conclusions

5.1 Summary

The aim of this study was to identify EV surface proteins with potential as BC biomarkers. EVs were isolated by dUC from three cell lines, malignant MDA-MB-231 and MCF7 cells and control non-malignant MCF10A cells. Two selection strategies for the selection of biomarkers were applied to derive cell line specific EV proteins with known implications in BC. The first approach was based on cell line specific identification of sEV proteins, UniProt annotation of transmembrane proteins and cluster analysis to predict co-localization with EV markers (CD63, CD9, CD81), and resulted in 11 proteins, four of which were further investigated by WB analysis. ST14, CLDN3, and ITGA7 were validated as potential biomarkers, while AXL was detected in MCF10A EVs instead of MDA-MB-231 and requires reevaluation. The second approach used biotinylation labelling to experimentally identify proteins that are

accessible from the surface of EVs. Proteins common to MDA-MB-231 and MCF7, but not identified in MCF10A were subjected to database annotation for BC resulting in 9 proteins (GSTP1, EEF2, DDX10, PGR, RAC2, ADAM10, ACO2, UTP20, N4BP2) and 2 proteins (PDCD6, UTP20) as potential surface markers of BC in sEVs and mEVs, respectively.

5.2 Limitations

This study is based on a limited number of cell lines to identify potential markers that could be further validated in EVs derived from plasma from BC patients. Increasing the number of cell lines studied would allow for more powerful biomarker selection as it could enable the identification of BC subtype specific markers (Rontogianni et al. 2019). Furthermore, a direct comparison of primary cell lines derived from tumor and adjacent tissue could uncover the differences in EV in patients (Hoshino et al. 2020). Lastly, cell line-based studies are generally considered initial exploration studies for biomarker discovery, however they require further validation with clinical samples. This poses a fundamental challenge to the analysis of EVs from biological samples like plasma or urine since they contain EVs from many different tissues and cell types of the body. Furthermore, body fluids are accessible to all cell types, hence EVs from some cell types may not be detectable at all.

The current definition of EVs depends on accepted EV markers like CD63, CD9 and CD81. EVs are known to be heterogeneous, containing subtypes like exosomes and microvesicles, each containing very different cargoes despite being within the same

"subgroup," but cannot be separated by current isolation methods (Willms et al. 2018). To respect the most recently published MISEV guidelines, the terms sEVs and mEVS were used to specify the studied biomaterial.

Addressing the heterogeneity of EVs, this study presented an approach to predict co-localization of such marker proteins with selected BC biomarkers similar to recent studies (Martin-Jaular et al. 2020, 2021). Still, predictions based on bulk analysis of EVs and correlation analysis must be further validated by single-vesicle techniques such as flow cytometry or analysis of affinity-enriched EVs fractions. For this reason, a binary comparison of identified and not identified proteins was chosen in both approaches, to overcome differential abundance due to subtype variation and to simplify possible future biomarker assays, which may not discriminate based on an abundance threshold but rather based on the mere presence of a protein. Likewise, flow cytometry or similar approaches are needed to confirm the surface expression and accessibility of proteins identified by the second approach.

5.3 Future Work

In this study, the presence of potential marker proteins was validated by WB analysis, however the result for AXL contradicted the proteomics results and needs to be reevaluated. A different clone of anti-AXL antibody can be used to help confirm AXL as MDA-MB-231 specific marker. Alternatively, the kinase activity of AXL can be investigated by enzymatic assays and getting functional insights into the difference between cell line EVs. Similarly for the second selection approach, proposed BC

biomarkers must be further investigated by WB analysis or flow cytometry to confirm the presence of respective proteins on the surface of EVs.

The two proteomics and biomarker selection approaches yielded promising lists of potential BC EV biomarkers. For further validation of their potential use for BC detection from plasma derived EVs in patients, the proteomics analysis could first be extended to include more BC cell lines such as BT549, Hs578T, LM2 (all triple-negative subtype), or HER2 positive subtype cell lines (HCC1954, HCC1419, JIMT1, SKBR3).

Future studies on plasma derived EVs may be the final goal of such exemplified biomarker discovery studies, however they pose major challenges for the analysis of EVs. Cell line studies, including other epithelial cell lines from different organs may be used to enrich a subclass of breast tissue specific EVs from plasma. Such enrichment of EV subsets would require understanding of their surface proteome which can be achieved with the presented labelling approach. Here, only native conditions were used during labelling of EVs, however other studies proposed EDTA treatment of EVs prior to labelling to remove the protein corona (Rai et al. 2021), which might be more similar to conditions used for plasma derived EVs and facilitate downstream applications.

In this study, the surface of EVs and their quantitative composition have been analyzed separately. However, an understanding of EV surface protein distribution could support biomarker studies as well as holds the potential to gain insight into the heterogeneity of EVs. It will therefore be important to adjust proteomics protocols to

obtain quantitative data of the surface proteins (Ravenhill et al. 2020). Finally, described large scale proteomics data will require extensive validation by single-vesicle methods like flow cytometry to confirm surface expression and to analyze the distribution of surface markers on EVs [(Maia et al. 2020); Nolan and Duggan (2018)]. The presented approaches combined may uncover the heterogeneity of EVs and facilitate biomarker studies of EVs on a large scale.

References

10 Abdouh, Mohamed, Dana Hamam, Zu-Hua Gao, Vincenzo Arena, Manuel Arena, and Goffredo Orazio Arena. 2017. "Exosomes Isolated from Cancer Patients' Sera Transfer Malignant Traits and Confer the Same Phenotype of Primary Tumors to Oncosuppressor-Mutated Cells." *Journal of Experimental & Clinical Cancer Research : CR* 36 (August): 113. https://doi.org/10.1186/s13046-017-0587-0.

Adem, Bárbara, Patricia F. Vieira, and Sonia A. Melo. 2020. "Decoding the Biology of Exosomes in Metastasis." *Trends in Cancer* 6 (1): 20–30. https://doi.org/10.1016/j.trecan.2019.11.007.

Aebersold, Ruedi, and Matthias Mann. 2003. "Mass Spectrometry-Based Proteomics." *Nature* 422 (6928, 6928): 198–207. https://doi.org/10.1038/nature01511.

Aguilera, Todd A., Marjan Rafat, Laura Castellini, Hussein Shehade, Mihalis S. Kariolis, Angela Bik-Yu Hui, Henning Stehr, et al. 2016. "Reprogramming the Immunological Microenvironment Through Radiation and Targeting Axl." *Nature Communications* 7 (1, 1): 13898. https://doi.org/10.1038/ncomms13898.

Alef, Thomas, Serena Torres, Ingrid Hausser, Dieter Metze, Umit Türsen, Gilles G Lestringant, and Hans Christian Hennies. 2009. "Ichthyosis, Follicular Atrophoderma, and Hypotrichosis Caused by Mutations in St14 Is Associated with Impaired Profilaggrin Processing." *The Journal of Investigative Dermatology* 129 (4): 862–69. https://doi.org/10.1038/jid.2008.311.

Altei, Wanessa F., Bianca C. Pachane, Patty K. dos Santos, Lígia N. M. Ribeiro, Bong Hwan Sung, Alissa M. Weaver, and Heloisa S. Selistre-de-Araújo. 2020. "Inhibition of Alpha-v Beta-3 Integrin Impairs Adhesion and Uptake of Tumor-Derived Small Extracellular Vesicles." *Cell Communication and Signaling* 18 (1): 158. https://doi.org/10.1186/s12964-020-00630-w.

Altei, Wanessa Fernanda, Bianca Cruz Pachane, Patty Karina dos Santos, and Heloisa Sobreiro Selistre de Araújo. 2021. "Extracellular Vesicles and Integrins: Partners in Cancer Progression." In *Role of Exosomes in Biological Communication Systems*, edited by Faisal A. Alzahrani and Islam M. Saadeldin, 293–310. Singapore: Springer. https://doi.org/10.1007/978-981-15-6599-1_14.

Altelaar, A. F. Maarten, Javier Munoz, and Albert J. R. Heck. 2013. "Next-Generation Proteomics: Towards an Integrative View of Proteome Dynamics." *Nature Reviews Genetics* 14 (1, 1): 35–48. https://doi.org/10.1038/nrg3356.

Andaluz Aguilar, Hillary, Anton B. Iliuk, I.-Hsuan Chen, and W. Andy Tao. 2020. "Sequential Phosphoproteomics and N-glycoproteomics of Plasma-Derived Extracellular Vesicles." *Nature Protocols* 15 (1, 1): 161–80. https://doi.org/10.1038/

s41596-019-0260-5.

Arai, T., Y. Miyoshi, S. J. Kim, K. Akazawa, N. Maruyama, T. Taguchi, Y. Tamaki, and S. Noguchi. 2008. "Association of Gstp1 Expression with Resistance to Docetaxel and Paclitaxel in Human Breast Cancers." *European Journal of Surgical Oncology: The Journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology* 34 (7): 734–38. https://doi.org/10.1016/j.ejso.2007.07.008.

Arena, Goffredo Orazio, Vincenzo Arena, Manuel Arena, and Mohamed Abdouh. 2017. "Transfer of Malignant Traits as Opposed to Migration of Cells: A Novel Concept to Explain Metastatic Disease." *Medical Hypotheses* 100 (March): 82–86. https://doi.org/10.1016/j.mehy.2017.01.019.

Arraud, N., R. Linares, S. Tan, C. Gounou, J.-M. Pasquet, S. Mornet, and A. R. Brisson. 2014. "Extracellular Vesicles from Blood Plasma: Determination of Their Morphology, Size, Phenotype and Concentration." *Journal of Thrombosis and Haemostasis: JTH* 12 (5): 614–27. https://doi.org/10.1111/jth.12554.

Battista, Theo, Annarita Fiorillo, Valerio Chiarini, Ilaria Genovese, Andrea Ilari, and Gianni Colotti. 2020. "Roles of Sorcin in Drug Resistance in Cancer: One Protein, Many Mechanisms, for a Novel Potential Anticancer Drug Target." *Cancers* 12 (4): 887. https://doi.org/10.3390/cancers12040887.

Beer, Lynn A., Pengyuan Liu, Bonnie Ky, Kurt T. Barnhart, and David W. Speicher. 2017. "Efficient Quantitative Comparisons of Plasma Proteomes Using Label-Free

Analysis with MaxQuant." *Methods in Molecular Biology (Clifton, N.J.)* 1619: 339–52. https://doi.org/10.1007/978-1-4939-7057-5_23.

Bellotti, Cristina, Kristina Lang, Nataliya Kuplennik, Alejandro Sosnik, and Robert Steinfeld. 2021. "High-Grade Extracellular Vesicles Preparation by Combined Size-Exclusion and Affinity Chromatography." *Scientific Reports* 11 (1, 1): 10550. https://doi.org/10.1038/s41598-021-90022-y.

Bergum, Christopher, Gina Zoratti, Julie Boerner, and Karin List. 2012. "Strong Expression Association Between Matriptase and Its Substrate Prostasin in Breast Cancer." *Journal of Cellular Physiology* 227 (4): 1604–9. https://doi.org/10.1002/jcp.22877.

Bhandari, Adheesh, Erjie Xia, Yuying Zhou, Yaoyao Guan, Jingjing Xiang, Lingguo Kong, Yinghao Wang, Fan Yang, Ouchen Wang, and Xiaohua Zhang. 2018. "Itga7 Functions as a Tumor Suppressor and Regulates Migration and Invasion in Breast Cancer." *Cancer Management and Research* 10 (May): 969–76. https://doi.org/10.2147/CMAR.S160379.

Bhatt, A. S., T. Takeuchi, B. Ylstra, D. Ginzinger, D. Albertson, M. A. Shuman, and C. S. Craik. 2003. "Quantitation of Membrane Type Serine Protease 1 (MT-Sp1) in Transformed and Normal Cells" 384 (2): 257–66. https://doi.org/10.1515/BC.2003.029.

Borges, F. T., L. A. Reis, and N. Schor. 2013. "Extracellular Vesicles: Structure, Function, and Potential Clinical Uses in Renal Diseases." *Brazilian Journal of*

Medical and Biological Research 46 (October): 824–30. https://doi.org/10.1590/1414-431X20132964.

Brennan, K., K. Martin, S. P. FitzGerald, J. O'Sullivan, Y. Wu, A. Blanco, C. Richardson, and M. M. Mc Gee. 2020. "A Comparison of Methods for the Isolation and Separation of Extracellular Vesicles from Protein and Lipid Particles in Human Serum." *Scientific Reports* 10 (1, 1): 1039. https://doi.org/10.1038/s41598-020-57497-7.

Brosch, Markus, Lu Yu, Tim Hubbard, and Jyoti Choudhary. 2009. "Accurate and Sensitive Peptide Identification with Mascot Percolator." *Journal of Proteome Research* 8 (6): 3176–81. https://doi.org/10.1021/pr800982s.

Brown, Kyle A., Jake A. Melby, David S. Roberts, and Ying Ge. 2020. "Top-down Proteomics: Challenges, Innovations, and Applications in Basic and Clinical Research." *Expert Review of Proteomics* 17 (10): 719–33. https://doi.org/10.1080/14789450.2020.1855982.

Brown, Peter N., and Hang Yin. 2017. "Polymer-Based Purification of Extracellular Vesicles." *Methods in Molecular Biology (Clifton, N.J.)* 1660: 91–103. https://doi.org/10.1007/978-1-4939-7253-1_8.

Brzozowski, Joshua S., Helen Jankowski, Danielle R. Bond, Siobhan B. McCague, Benjamin R. Munro, Melanie J. Predebon, Christopher J. Scarlett, Kathryn A. Skelding, and Judith Weidenhofer. 2018. "Lipidomic Profiling of Extracellular Vesicles Derived from Prostate and Prostate Cancer Cell Lines." *Lipids in Health*

and Disease 17 (1): 211. https://doi.org/10.1186/s12944-018-0854-x.

Cai, Huaqing, Karin Reinisch, and Susan Ferro-Novick. 2007. "Coats, Tethers, Rabs, and SNAREs Work Together to Mediate the Intracellular Destination of a Transport Vesicle." *Developmental Cell* 12 (5): 671–82. https://doi.org/10.1016/j.devcel.2007.04.005.

Carlson, Marc. 2019. *Org.Hs.eg.db: Genome Wide Annotation for Human. R Package Version 3.8.2.*

Carlson, Marc, Csaba Ortutay, and Bioconductor Package Maintainer. 2021. *UniProt.ws: R Interface to UniProt Web Services* (version 2.34.0). Bioconductor version: Release (3.14). https://doi.org/10.18129/B9.bioc.UniProt.ws.

Chertov, Oleg, Arya Biragyn, Larry W. Kwak, John T. Simpson, Tatiana Boronina, Van M. Hoang, DaRue A. Prieto, Thomas P. Conrads, Timothy D. Veenstra, and Robert J. Fisher. 2004. "Organic Solvent Extraction of Proteins and Peptides from Serum as an Effective Sample Preparation for Detection and Identification of Biomarkers by Mass Spectrometry." *Proteomics* 4 (4): 1195–1203. https://doi.org/10.1002/pmic.200300677.

Chiabotto, Giulia, Chiara Gai, Maria Chiara Deregibus, and Giovanni Camussi. 2019. "Salivary Extracellular Vesicle-Associated exRNA as Cancer Biomarker." *Cancers* 11 (7): 891. https://doi.org/10.3390/cancers11070891.

Chitti, Sai V., and Christina Nedeva. 2021. "Extracellular Vesicles Contain Putative Cancer Biomarkers." *Sub-Cellular Biochemistry* 97: 363–74. https://doi.org/

10.1007/978-3-030-67171-6_13.

Ciravolo, Valentina, Veronica Huber, Gaia C. Ghedini, Elisabetta Venturelli, Francesca Bianchi, Manuela Campiglio, Daniele Morelli, et al. 2012. "Potential Role of HER2-overexpressing Exosomes in Countering Trastuzumab-Based Therapy." *Journal of Cellular Physiology* 227 (2): 658–67. https://doi.org/10.1002/jcp.22773.

Coles, Brian F., and Fred F. Kadlubar. 2003. "Detoxification of Electrophilic Compounds by Glutathione S-transferase Catalysis: Determinants of Individual Response to Chemical Carcinogens and Chemotherapeutic Drugs?" *BioFactors (Oxford, England)* 17 (1-4): 115–30. https://doi.org/10.1002/biof.5520170112.

Coon, Joshua J., John E. P. Syka, Jeffrey Shabanowitz, and Donald F. Hunt. 2005. "Tandem Mass Spectrometry for Peptide and Protein Sequence Analysis." *BioTechniques* 38 (4): 519–23. https://doi.org/10.2144/05384TE01.

Cour, Jonas M. la, Berit R. Høj, Jens Mollerup, Ronald Simon, Guido Sauter, and Martin W. Berchtold. 2008. "The Apoptosis Linked Gene ALG-2 Is Dysregulated in Tumors of Various Origin and Contributes to Cancer Cell Viability." *Molecular Oncology* 1 (4): 431–39. https://doi.org/10.1016/j.molonc.2007.08.002.

Cour, Jonas M. la, Jens Mollerup, Pernille Winding, Svetlana Tarabykina, Maxwell Sehested, and Martin W. Berchtold. 2003. "Up-Regulation of ALG-2 in Hepatomas and Lung Cancer Tissue." *The American Journal of Pathology* 163 (1):

81–89. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1868153/.

Cox, Jürgen, Marco Y. Hein, Christian A. Luber, Igor Paron, Nagarjuna Nagaraj, and Matthias Mann. 2014. "Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ." *Molecular & Cellular Proteomics : MCP* 13 (9): 2513–26. https://doi.org/10.1074/mcp.M113.031591.

Cox, Jürgen, and Matthias Mann. 2008. "MaxQuant Enables High Peptide Identification Rates, Individualized p.p.b.-Range Mass Accuracies and Proteome-Wide Protein Quantification." *Nature Biotechnology* 26 (12, 12): 1367–72. https://doi.org/10.1038/nbt.1511.

Cox, Jürgen, Nadin Neuhauser, Annette Michalski, Richard A. Scheltema, Jesper V. Olsen, and Matthias Mann. 2011. "Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment." *Journal of Proteome Research* 10 (4): 1794–1805. https://doi.org/10.1021/pr101065j.

Crawford, N. 1971. "The Presence of Contractile Proteins in Platelet Microparticles Isolated from Human and Animal Platelet-Free Plasma." *British Journal of Haematology* 21 (1): 53–69. https://doi.org/10.1111/j.1365-2141.1971.tb03416.x.

Crescitelli, Rossella, Cecilia Lässer, Tamas G. Szabó, Agnes Kittel, Maria Eldh, Irma Dianzani, Edit I. Buzás, and Jan Lötvall. 2013. "Distinct RNA Profiles in Subpopulations of Extracellular Vesicles: Apoptotic Bodies, Microvesicles and Exosomes."

Journal of Extracellular Vesicles 2 (1): 20677. https://doi.org/10.3402/jev.v2i0.20677.

Crick, Francis. 1970. "Central Dogma of Molecular Biology." *Nature* 227 (5258, 5258): 561–63. https://doi.org/10.1038/227561a0.

Crutchfield, Christopher A., Stefani N. Thomas, Lori J. Sokoll, and Daniel W. Chan. 2016. "Advances in Mass Spectrometry-Based Clinical Biomarker Discovery." *Clinical Proteomics* 13 (1): 1. https://doi.org/10.1186/s12014-015-9102-9.

Cvjetkovic, Aleksander, Su Chul Jang, Barbora Konečná, Johanna L. Höög, Carina Sihlbom, Cecilia Lässer, and Jan Lötvall. 2016. "Detailed Analysis of Protein Topology of Extracellular Vesicles–Evidence of Unconventional Membrane Protein Orientation." *Scientific Reports* 6 (1, 1): 36338. https://doi.org/10.1038/srep36338.

Dai, Yang-Hong, Ying-Fu Wang, Po-Chien Shen, Cheng-Hsiang Lo, Jen-Fu Yang, Chun-Shu Lin, Hsing-Lung Chao, and Wen-Yen Huang. 2021. "Gene-Associated Methylation Status of St14 as a Predictor of Survival and Hormone Receptor Positivity in Breast Cancer." *BMC Cancer* 21 (1): 945. https://doi.org/10.1186/s12885-021-08645-3.

Dallavalle, Sabrina, Vladimir Dobričić, Loretta Lazzarato, Elena Gazzano, Miguel Machuqueiro, Ilza Pajeva, Ivanka Tsakovska, Nace Zidar, and Roberta Fruttero. 2020. "Improvement of Conventional Anti-Cancer Drugs as New Tools Against Multidrug Resistant Tumors." *Drug Resistance Updates: Reviews and Commen-*

taries in Antimicrobial and Anticancer Chemotherapy 50 (May): 100682. https://doi.org/10.1016/j.drup.2020.100682.

Danzinger, Sabine, Yen Yen Tan, Margaretha Rudas, Marie-Theres Kastner, Sigrid Weingartshofer, Daniela Muhr, and Christian F. Singer. 2019. "Association of Cytokeratin 5 and Claudin 3 Expression with Brca1 and Brca2 Germline Mutations in Women with Early Breast Cancer." *BMC Cancer* 19 (1): 695. https://doi.org/10.1186/s12885-019-5908-6.

De Franceschi, Nicola, Hellyeh Hamidi, Jonna Alanko, Pranshu Sahgal, and Johanna Ivaska. 2015. "Integrin Traffic – the Update." *Journal of Cell Science* 128 (5): 839–52. https://doi.org/10.1242/jcs.161653.

Debnath, Jayanta, Senthil K. Muthuswamy, and Joan S. Brugge. 2003. "Morphogenesis and Oncogenesis of MCF-10a Mammary Epithelial Acini Grown in Three-Dimensional Basement Membrane Cultures." *Methods (San Diego, Calif.)* 30 (3): 256–68. https://doi.org/10.1016/s1046-2023(03)00032-x.

Debruyne, D. N., N. Bhatnagar, B. Sharma, W. Luther, N. F. Moore, N.-K. Cheung, N. S. Gray, and R. E. George. 2016. "ALK Inhibitor Resistance in ALK(F1174L)-Driven Neuroblastoma Is Associated with AXL Activation and Induction of EMT." *Oncogene* 35 (28): 3681–91. https://doi.org/10.1038/onc.2015.434.

DeSantis, Carol E., Chun Chieh Lin, Angela B. Mariotto, Rebecca L. Siegel, Kevin D. Stein, Joan L. Kramer, Rick Alteri, Anthony S. Robbins, and Ahmedin Jemal. 2014. "Cancer Treatment and Survivorship Statistics, 2014." *CA: A Cancer Jour-*

nal for Clinicians 64 (4): 252–71. https://doi.org/10.3322/caac.21235.

Deville, Sarah, Pascale Berckmans, Rebekka Van Hoof, Ivo Lambrichts, Anna Salvati, and Inge Nelissen. 2021. "Comparison of Extracellular Vesicle Isolation and Storage Methods Using High-Sensitivity Flow Cytometry." *PLOS ONE* 16 (2): e0245835. https://doi.org/10.1371/journal.pone.0245835.

Dez, Christophe, Mensur Dlakić, and David Tollervey. 2007. "Roles of the HEAT Repeat Proteins Utp10 and Utp20 in 40s Ribosome Maturation." *RNA* 13 (9): 1516–27. https://doi.org/10.1261/rna.609807.

Di Vizio, Dolores, Matteo Morello, Andrew C. Dudley, Peter W. Schow, Rosalyn M. Adam, Samantha Morley, David Mulholland, et al. 2012. "Large Oncosomes in Human Prostate Cancer Tissues and in the Circulation of Mice with Metastatic Disease." *The American Journal of Pathology* 181 (5): 1573–84. https://doi.org/10.1016/j.ajpath.2012.07.030.

Doerr, Allison. 2015. "DIA Mass Spectrometry." *Nature Methods* 12 (1, 1): 35–35. https://doi.org/10.1038/nmeth.3234.

Dong, Xingli, Xupeng Bai, Jie Ni, Hao Zhang, Wei Duan, Peter Graham, and Yong Li. 2020. "Exosomes and Breast Cancer Drug Resistance." *Cell Death & Disease* 11 (11, 11): 1–14. https://doi.org/10.1038/s41419-020-03189-z.

Doyle, Laura M., and Michael Zhuo Wang. 2019. "Overview of Extracellular Vesicles, Their Origin, Composition, Purpose, and Methods for Exosome Isolation and Analysis." *Cells* 8 (7): 727. https://doi.org/10.3390/cells8070727.

Duffy, Michael J., Siun Walsh, Enda W. McDermott, and John Crown. 2015. "Biomarkers in Breast Cancer: Where Are We and Where Are We Going?" *Advances in Clinical Chemistry* 71: 1–23. https://doi.org/10.1016/bs.acc.2015.05.001.

Elias, Joshua E., and Steven P. Gygi. 2007. "Target-Decoy Search Strategy for Increased Confidence in Large-Scale Protein Identifications by Mass Spectrometry." *Nature Methods* 4 (3, 3): 207–14. https://doi.org/10.1038/nmeth1019.

Elmore, Susan. 2007. "Apoptosis: A Review of Programmed Cell Death." *Toxicologic Pathology* 35 (4): 495–516. https://doi.org/10.1080/01926230701320337.

Emelyanov, Anton, Tatiana Shtam, Roman Kamyshinsky, Luiza Garaeva, Nikolai Verlov, Irina Miliukhina, Anastasia Kudrevatykh, et al. 2020. "Cryo-Electron Microscopy of Extracellular Vesicles from Cerebrospinal Fluid." *PLoS ONE* 15 (1): e0227949. https://doi.org/10.1371/journal.pone.0227949.

Eng, J. K., A. L. McCormack, and J. R. Yates. 1994. "An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database." *Journal of the American Society for Mass Spectrometry* 5 (11): 976–89. https://doi.org/10.1016/1044-0305(94)80016-2.

Eralp, Yeşim, Serkan Keskin, Ebru Akışık, Elif Akışık, Abdullah İğci, Mahmut Müslümanoğlu, Sabri Yılmaz, et al. 2013. "Predictive Role of Midtreatment Changes in Survivin, Gstp1, and Topoisomerase 2alpha Expressions for Pathologic Complete Response to Neoadjuvant Chemotherapy in Patients with Locally

Advanced Breast Cancer." *American Journal of Clinical Oncology* 36 (3): 215–23. https://doi.org/10.1097/COC.0b013e318243913f.

Ferrarelli, Leslie K. 2020. "Linking Integrins and EVs in Metastasis." *Science Signaling* 13 (649): eabe3353. https://doi.org/10.1126/scisignal.abe3353.

Finehout, Erin J., and Kelvin H. Lee. 2004. "An Introduction to Mass Spectrometry Applications in Biological Research." *Biochemistry and Molecular Biology Education* 32 (2): 93–100. https://doi.org/10.1002/bmb.2004.494032020331.

Gay, Carl M., Kavitha Balaji, and Lauren Averett Byers. 2017. "Giving AXL the Axe: Targeting AXL in Human Malignancy." *British Journal of Cancer* 116 (4): 415–23. https://doi.org/10.1038/bjc.2016.428.

Ge, Weiting, Hanguang Hu, Kefeng Ding, Lifeng Sun, and Shu Zheng. 2006. "Protein Interaction Analysis of St14 Domains and Their Point and Deletion Mutants." *The Journal of Biological Chemistry* 281 (11): 7406–12. https://doi.org/10.1074/jbc.M510687200.

Gill, Sukhvinder, Ryan Catchpole, and Patrick Forterre. 2019. "Extracellular Membrane Vesicles in the Three Domains of Life and Beyond." *FEMS Microbiology Reviews* 43 (3): 273–303. https://doi.org/10.1093/femsre/fuy042.

Giordano, Cinzia, Giusi La Camera, Luca Gelsomino, Ines Barone, Daniela Bonofiglio, Sebastiano Andò, and Stefania Catalano. 2020. "The Biology of Exosomes in Breast Cancer Progression: Dissemination, Immune Evasion and Metastatic Colonization." *Cancers* 12 (8): 2179. https://doi.org/10.3390/cancers12082179.

Gjerdrum, Christine, Crina Tiron, Torill Høiby, Ingunn Stefansson, Hallvard Haugen, Tone Sandal, Karin Collett, et al. 2010. "Axl Is an Essential Epithelial-to-Mesenchymal Transition-Induced Regulator of Breast Cancer Metastasis and Patient Survival." *Proceedings of the National Academy of Sciences of the United States of America* 107 (3): 1124–29. https://doi.org/10.1073/pnas.0909333107.

Glish, Gary L., and Richard W. Vachet. 2003. "The Basics of Mass Spectrometry in the Twenty-First Century." *Nature Reviews Drug Discovery* 2 (2, 2): 140–50. https://doi.org/10.1038/nrd1011.

Gorello, Paolo, Valeria Nofrini, Lucia Brandimarte, Valentina Pierini, Barbara Crescenzi, Filomena Nozza, Giulia Daniele, et al. 2013. "Inv(11)(P15q22)/Nup98-Ddx10 Fusion and Isoforms in a New Case of de Novo Acute Myeloid Leukemia." *Cancer Genetics* 206 (3): 92–96. https://doi.org/10.1016/j.cancergen.2013.02.001.

Green, Toni M., Mary L. Alpaugh, Sanford H. Barsky, Germana Rappa, and Aurelio Lorico. 2015. "Breast Cancer-Derived Extracellular Vesicles: Characterization and Contribution to the Metastatic Phenotype." *BioMed Research International* 2015: 634865. https://doi.org/10.1155/2015/634865.

Greening, David W., Rong Xu, Hong Ji, Bow J. Tauro, and Richard J. Simpson. 2015. "A Protocol for Exosome Isolation and Characterization: Evaluation of Ultracentrifugation, Density-Gradient Separation, and Immunoaffinity Capture Methods." *Methods in Molecular Biology (Clifton, N.J.)* 1295: 179–209.

https://doi.org/10.1007/978-1-4939-2550-6_15.

Guan, Yaoyao, Adheesh Bhandari, Erjie Xia, Lingguo Kong, Xiaohua Zhang, and Ouchen Wang. 2020. "Downregulating Integrin Subunit Alpha 7 (Itga7) Promotes Proliferation, Invasion, and Migration of Papillary Thyroid Carcinoma Cells Through Regulating Epithelial-to-Mesenchymal Transition." *Acta Biochimica Et Biophysica Sinica* 52 (2): 116–24. https://doi.org/10.1093/abbs/gmz144.

Guha, Debjani, David R. Lorenz, Vikas Misra, Sukrutha Chettimada, Susan Morgello, and Dana Gabuzda. 2019. "Proteomic Analysis of Cerebrospinal Fluid Extracellular Vesicles Reveals Synaptic Injury, Inflammation, and Stress Response Markers in HIV Patients with Cognitive Impairment." *Journal of Neuroinflammation* 16 (1): 254. https://doi.org/10.1186/s12974-019-1617-y.

Guo, Zhiqiang, Yan Li, Dandan Zhang, and Jiaying Ma. 2017. "Axl Inhibition Induces the Antitumor Immune Response Which Can Be Further Potentiated by PD-1 Blockade in the Mouse Cancer Models." *Oncotarget* 8 (52): 89761–74. https://doi.org/10.18632/oncotarget.21125.

Gwili, Noha, Stacey J. Jones, Waleed Al Amri, Ian M. Carr, Sarah Harris, Brian V. Hogan, William E. Hughes, et al. 2021. "Transcriptome Profiles of Stem-Like Cells from Primary Breast Cancers Allow Identification of Itga7 as a Predictive Marker of Chemotherapy Response." *British Journal of Cancer* 125 (7, 7): 983–93. https://doi.org/10.1038/s41416-021-01484-w.

Harding, C., and P. Stahl. 1983. "Transferrin Recycling in Reticulocytes: pH and

Iron Are Important Determinants of Ligand Binding and Processing." *Biochemical and Biophysical Research Communications* 113 (2): 650–58. https://doi.org/10.1016/0006-291x(83)91776-x.

Harding, Clifford V., John E. Heuser, and Philip D. Stahl. 2013. "Exosomes: Looking Back Three Decades and into the Future." *The Journal of Cell Biology* 200 (4): 367–71. https://doi.org/10.1083/jcb.201212113.

Harjunpää, Heidi, Marc Llort Asens, Carla Guenther, and Susanna C. Fagerholm. 2019. "Cell Adhesion Molecules and Their Roles and Regulation in the Immune and Tumor Microenvironment." *Frontiers in Immunology* 10: 1078. https://doi.org/10.3389/fimmu.2019.01078.

Heerma van Voss, Marise R., Paul J. van Diest, and Venu Raman. 2017. "Targeting RNA Helicases in Cancer: The Translation Trap." *Biochimica Et Biophysica Acta. Reviews on Cancer* 1868 (2): 510–20. https://doi.org/10.1016/j.bbcan.2017.09.006.

Henry, Stephen, Eleanor Williams, Katie Barr, Elena Korchagina, Alexandr Tuzikov, Natalia Ilyushina, Sidahmed A. Abayzeed, Kevin F. Webb, and Nicolai Bovin. 2018. "Rapid One-Step Biotinylation of Biological and Non-Biological Surfaces." *Scientific Reports* 8 (1, 1): 2845. https://doi.org/10.1038/s41598-018-21186-3.

Hevir, N., N. Trošt, N. Debeljak, and T. Lanišnik Rižner. 2011. "Expression of Estrogen and Progesterone Receptors and Estrogen Metabolizing Enzymes in Different Breast Cancer Cell Lines." *Chemico-Biological Interactions* 191 (1-3): 206–16.

https://doi.org/10.1016/j.cbi.2010.12.013.

Ho, CS, CWK Lam, MHM Chan, RCK Cheung, LK Law, LCW Lit, KF Ng, MWM Suen, and HL Tai. 2003. "Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications." *The Clinical Biochemist Reviews* 24 (1): 3–12. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1853331/.

Hoshino, Ayuko, Bruno Costa-Silva, Tang-Long Shen, Goncalo Rodrigues, Ayako Hashimoto, Milica Tesic Mark, Henrik Molina, et al. 2015. "Tumour Exosome Integrins Determine Organotropic Metastasis." *Nature* 527 (7578): 329–35. https://doi.org/10.1038/nature15756.

Hoshino, Ayuko, Han Sang Kim, Linda Bojmar, Kofi Ennu Gyan, Michele Cioffi, Jonathan Hernandez, Constantinos P. Zambirinis, et al. 2020. "Extracellular Vesicle and Particle Biomarkers Define Multiple Human Cancers." *Cell* 182 (4): 1044–1061.e18. https://doi.org/10.1016/j.cell.2020.07.009.

Huang, Guo N. 2012. "Biotinylation of Cell Surface Proteins." *Bio-Protocol* 2 (9): e170. https://doi.org/10.21769/BioProtoc.170.

Humphries, Jonathan D., Adam Byron, and Martin J. Humphries. 2006. "Integrin Ligands at a Glance." *Journal of Cell Science* 119 (19): 3901–3. https://doi.org/10.1242/jcs.03098.

Hurwitz, Stephanie N., and David G. Meckes. 2019. "Extracellular Vesicle Integrins Distinguish Unique Cancers." *Proteomes* 7 (2): 14. https://doi.org/10.3390/proteomes7020014.

Høj, Berit R., Jonas M. la Cour, Jens Mollerup, and Martin W. Berchtold. 2009. "ALG-2 Knockdown in HeLa Cells Results in G2/M Cell Cycle Phase Accumulation and Cell Death." *Biochemical and Biophysical Research Communications* 378 (1): 145–48. https://doi.org/10.1016/j.bbrc.2008.11.021.

Ikeda, Chisaki, Hiroaki Haga, Naohiko Makino, Tatsutoshi Inuzuka, Ayako Kurimoto, Toshiki Ueda, Akiko Matsuda, et al. 2021. "Utility of Claudin-3 in Extracellular Vesicles from Human Bile as Biomarkers of Cholangiocarcinoma." *Scientific Reports* 11 (1, 1): 1195. https://doi.org/10.1038/s41598-021-81023-y.

Itzhak, Daniel N, Stefka Tyanova, Jürgen Cox, and Georg HH Borner. 2016. "Global, Quantitative and Dynamic Mapping of Protein Subcellular Localization." Edited by Ramanujan S Hegde. *eLife* 5 (June): e16950. https://doi.org/10.7554/eLife.16950.

Jabbari, Nasrollah, Elinaz Akbariazar, Maryam Feqhhi, Reza Rahbarghazi, and Jafar Rezaie. 2020. "Breast Cancer-Derived Exosomes: Tumor Progression and Therapeutic Agents." *Journal of Cellular Physiology* 235 (10): 6345–56. https://doi.org/10.1002/jcp.29668.

Jääskeläinen, Anniina, Ylermi Soini, Arja Jukkola-Vuorinen, Päivi Auvinen, Kirsi-Maria Haapasaari, and Peeter Karihtala. 2018. "High-Level Cytoplasmic Claudin 3 Expression Is an Independent Predictor of Poor Survival in Triple-Negative Breast Cancer." *BMC Cancer* 18 (1): 223. https://doi.org/10.1186/s12885-018-4141-z.

Jeppesen, Dennis K., Aidan M. Fenix, Jeffrey L. Franklin, James N. Higginbotham,

Qin Zhang, Lisa J. Zimmerman, Daniel C. Liebler, et al. 2019. "Reassessment of Exosome Composition." *Cell* 177 (2): 428–445.e18. https://doi.org/10.1016/j.cell.2019.02.029.

Jiao, Xiang, Sean D. Hooper, Tatjana Djureinovic, Chatarina Larsson, Fredrik Wärnberg, Christian Tellgren-Roth, Johan Botling, and Tobias Sjöblom. 2013. "Gene Rearrangements in Hormone Receptor Negative Breast Cancers Revealed by Mate Pair Sequencing." *BMC Genomics* 14 (March): 165. https://doi.org/10.1186/1471-2164-14-165.

Jin, J. S., T. F. Cheng, W. C. Tsai, L. F. Sheu, H. Chiang, and C. P. Yu. 2007. "Expression of the Serine Protease, Matriptase, in Breast Ductal Carcinoma of Chinese Women: Correlation with Clinicopathological Parameters." *Histology and Histopathology*. https://digitum.um.es/digitum/handle/10201/27559.

Jin, Liang, Yingtao Bi, Chenqi Hu, Jun Qu, Shichen Shen, Xue Wang, and Yu Tian. 2021. "A Comparative Study of Evaluating Missing Value Imputation Methods in Label-Free Proteomics." *Scientific Reports* 11 (1, 1): 1760. https://doi.org/10.1038/s41598-021-81279-4.

Kalra, Hina, Richard J. Simpson, Hong Ji, Elena Aikawa, Peter Altevogt, Philip Askenase, Vincent C. Bond, et al. 2012. "Vesiclepedia: A Compendium for Extracellular Vesicles with Continuous Community Annotation." *PLOS Biology* 10 (12): e1001450. https://doi.org/10.1371/journal.pbio.1001450.

Kaul, Gautam, Gurulingappa Pattan, and Towseef Rafeequi. 2011. "Eukaryotic

Elongation Factor-2 (eEF2): Its Regulation and Peptide Chain Elongation." *Cell Biochemistry and Function* 29 (3): 227–34. https://doi.org/10.1002/cbf.1740.

Keerthikumar, Shivakumar, David Chisanga, Dinuka Ariyaratne, Haidar Al Saffar, Sushma Anand, Kening Zhao, Monisha Samuel, et al. 2016. "ExoCarta: A Web-Based Compendium of Exosomal Cargo." *Journal of Molecular Biology*, Computation Resources for Molecular Biology, 428 (4): 688–92. https://doi.org/10.1016/j.jmb.2015.09.019.

Keilhauer, Eva C., Marco Y. Hein, and Matthias Mann. 2015. "Accurate Protein Complex Retrieval by Affinity Enrichment Mass Spectrometry (AE-MS) Rather Than Affinity Purification Mass Spectrometry (AP-MS)." *Molecular & Cellular Proteomics: MCP* 14 (1): 120–35. https://doi.org/10.1074/mcp.M114.041012.

Kerr, J. F. R., A. H. Wyllie, and A. R. Currie. 1972. "Apoptosis: A Basic Biological Phenomenon with Wideranging Implications in Tissue Kinetics." *British Journal of Cancer* 26 (4, 4): 239–57. https://doi.org/10.1038/bjc.1972.33.

Kim, Chungho, Yongcheol Cho, Chan-Hee Kang, Moon Gyo Kim, HyoSeon Lee, Eun-Gyung Cho, and Dongeun Park. 2005. "Filamin Is Essential for Shedding of the Transmembrane Serine Protease, Epithin." *EMBO Reports* 6 (11): 1045–51. https://doi.org/10.1038/sj.embor.7400534.

Kim, Ki Yeon, Minsang Yoon, Youngkyung Cho, Kwang-Hoon Lee, Sora Park, Sera Lee, So-Young Choi, et al. 2019. "Targeting Metastatic Breast Cancer with

Peptide Epitopes Derived from Autocatalytic Loop of Prss14/St14 Membrane Serine Protease and with Monoclonal Antibodies." *Journal of Experimental & Clinical Cancer Research* 38 (1): 363. https://doi.org/10.1186/s13046-019-1373-y.

Kowal, Joanna, Guillaume Arras, Marina Colombo, Mabel Jouve, Jakob Paul Morath, Bjarke Primdal-Bengtson, Florent Dingli, Damarys Loew, Mercedes Tkach, and Clotilde Théry. 2016. "Proteomic Comparison Defines Novel Markers to Characterize Heterogeneous Populations of Extracellular Vesicle Subtypes." *Proceedings of the National Academy of Sciences* 113 (8): E968–77. https://doi.org/10.1073/pnas.1521230113.

Kunc, Michał, Marta Popęda, Wojciech Biernat, and Elżbieta Senkus. 2021. "Lost but Not Least-Novel Insights into Progesterone Receptor Loss in Estrogen Receptor-Positive Breast Cancer." *Cancers* 13 (19): 4755. https://doi.org/10.3390/cancers13194755.

Lal-Nag, Madhu, and Patrice J. Morin. 2009. "The Claudins." *Genome Biology* 10 (8): 235. https://doi.org/10.1186/gb-2009-10-8-235.

Langfelder, Peter, and Steve Horvath. 2008. "WGCNA: An R Package for Weighted Correlation Network Analysis." *BMC Bioinformatics* 9 (1): 559. https://doi.org/10.1186/1471-2105-9-559.

Lazar, Cosmin, Laurent Gatto, Myriam Ferro, Christophe Bruley, and Thomas Burger. 2016. "Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies." *Journal of Pro-*

teome Research 15 (4): 1116–25. https://doi.org/10.1021/acs.jproteome.5b00981.

Lee, Kyungheon, Huilin Shao, Ralph Weissleder, and Hakho Lee. 2015. "Acoustic Purification of Extracellular Microvesicles." *ACS Nano* 9 (3): 2321–27. https://doi.org/10.1021/nn506538f.

Levin, Simon A. 2006. "Fundamental Questions in Biology." *PLOS Biology* 4 (9): e300. https://doi.org/10.1371/journal.pbio.0040300.

Li, Jing, Xianqing He, Yuanyuan Deng, and Chenxi Yang. 2019. "An Update on Isolation Methods for Proteomic Studies of Extracellular Vesicles in Biofluids." *Molecules* 24 (19): 3516. https://doi.org/10.3390/molecules24193516.

Li, Maoyu, Fang Peng, Guoqiang Wang, Xujun Liang, Meiying Shao, Zhuchu Chen, and Yongheng Chen. 2021. "Coupling of Cell Surface Biotinylation and SILAC-Based Quantitative Proteomics Identified Myoferlin as a Potential Therapeutic Target for Nasopharyngeal Carcinoma Metastasis." *Frontiers in Cell and Developmental Biology* 9: 1327. https://doi.org/10.3389/fcell.2021.621810.

Li, Pin, Melisa Kaslan, Sze Han Lee, Justin Yao, and Zhiqiang Gao. 2017. "Progress in Exosome Isolation Techniques." *Theranostics* 7 (3): 789–804. https://doi.org/10.7150/thno.18133.

List, Karin, Christian C. Haudenschild, Roman Szabo, WanJun Chen, Sharon M. Wahl, William Swaim, Lars H. Engelholm, Niels Behrendt, and Thomas H. Bugge. 2002. "Matriptase/MT-Sp1 Is Required for Postnatal Survival, Epidermal Barrier

Function, Hair Follicle Development, and Thymic Homeostasis." *Oncogene* 21 (23): 3765–79. https://doi.org/10.1038/sj.onc.1205502.

List, Karin, Roman Szabo, Alfredo Molinolo, Boye Schnack Nielsen, and Thomas H. Bugge. 2006. "Delineation of Matriptase Protein Expression by Enzymatic Gene Trapping Suggests Diverging Roles in Barrier Function, Hair Formation, and Squamous Cell Carcinogenesis." *The American Journal of Pathology* 168 (5): 1513–25. https://doi.org/10.2353/ajpath.2006.051071.

List, Karin, Roman Szabo, Philip W. Wertz, Julie Segre, Christian C. Haudenschild, Soo-Youl Kim, and Thomas H. Bugge. 2003. "Loss of Proteolytically Processed Filaggrin Caused by Epidermal Deletion of Matriptase/MT-Sp1." *The Journal of Cell Biology* 163 (4): 901–10. https://doi.org/10.1083/jcb.200304161.

Liu, Li, James Greger, Hong Shi, Yuan Liu, Joel Greshock, Roland Annan, Wendy Halsey, Ganesh M. Sathe, Anne-Marie Martin, and Tona M. Gilmer. 2009. "Novel Mechanism of Lapatinib Resistance in HER2-positive Breast Tumor Cells: Activation of AXL." *Cancer Research* 69 (17): 6871–78. https://doi.org/10.1158/0008-5472.CAN-08-4490.

Livshits, Mikhail A., Elena Khomyakova, Evgeniy G. Evtushenko, Vassili N. Lazarev, Nikolay A. Kulemin, Svetlana E. Semina, Edward V. Generozov, and Vadim M. Govorun. 2015. "Isolation of Exosomes by Differential Centrifugation: Theoretical Analysis of a Commonly Used Protocol." *Scientific Reports* 5 (1, 1): 17319. https://doi.org/10.1038/srep17319.

Lu, Lingeng, and Harvey A. Risch. 2016. "Exosomes: Potential for Early Detection in Pancreatic Cancer." *Future Oncology (London, England)* 12 (8): 1081–90. https://doi.org/10.2217/fon-2015-0005.

Lundberg, Emma, and Georg H. H. Borner. 2019. "Spatial Proteomics: A Powerful Discovery Tool for Cell Biology." *Nature Reviews Molecular Cell Biology* 20 (5, 5): 285–302. https://doi.org/10.1038/s41580-018-0094-y.

Mahul-Mellier, Anne-Laure, Fiona J. Hemming, Béatrice Blot, Sandrine Fraboulet, and Rémy Sadoul. 2006. "Alix, Making a Link Between Apoptosis-Linked Gene-2, the Endosomal Sorting Complexes Required for Transport, and Neuronal Death in Vivo." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 26 (2): 542–49. https://doi.org/10.1523/JNEUROSCI.3069-05.2006.

Maia, Joana, Silvia Batista, Nuno Couto, Ana C. Gregório, Cristian Bodo, Julia Elzanowska, Maria Carolina Strano Moraes, and Bruno Costa-Silva. 2020. "Employing Flow Cytometry to Extracellular Vesicles Sample Microvolume Analysis and Quality Control." *Frontiers in Cell and Developmental Biology* 8: 1165. https://doi.org/10.3389/fcell.2020.593750.

Maia, Joana, Sergio Caja, Maria Carolina Strano Moraes, Nuno Couto, and Bruno Costa-Silva. 2018. "Exosome-Based Cell-Cell Communication in the Tumor Microenvironment." *Frontiers in Cell and Developmental Biology* 6: 18. https://doi.org/10.3389/fcell.2018.00018.

Mallia, Alice, Erica Gianazza, Beatrice Zoanni, Maura Brioschi, Silvia Stella Barbieri, and Cristina Banfi. 2020. "Proteomics of Extracellular Vesicles: Update on Their Composition, Biological Roles and Potential Use as Diagnostic Tools in Atherosclerotic Cardiovascular Diseases." *Diagnostics* 10 (10): 843. https://doi.org/10.3390/diagnostics10100843.

Margolis, Leonid, and Yoel Sadovsky. 2019. "The Biology of Extracellular Vesicles: The Known Unknowns." *PLOS Biology* 17 (7): e3000363. https://doi.org/10.1371/journal.pbio.3000363.

Marleau, Annette M., Chien-Shing Chen, James A. Joyce, and Richard H. Tullis. 2012. "Exosome Removal as a Therapeutic Adjuvant in Cancer." *Journal of Translational Medicine* 10 (1): 134. https://doi.org/10.1186/1479-5876-10-134.

Marth, Jamey D. 2008. "A Unified Vision of the Building Blocks of Life." *Nature Cell Biology* 10 (9): 1015–16. https://doi.org/10.1038/ncb0908-1015.

Martin-Jaular, Lorena, Nathalie Nevo, Julia P. Schessner, Mercedes Tkach, Mabel Jouve, Florent Dingli, Damarys Loew, et al. 2021. "Unbiased Proteomic Profiling of Host Cell Extracellular Vesicle Composition and Dynamics Upon HIV-1 Infection." *The EMBO Journal* 40 (8): e105492. https://doi.org/10.15252/embj.2020105492.

Martin-Jaular, Lorena, Nathalie Nevo, Mercedes Tkach, Mabel Jouve, Florent Dingli, Damarys Loew, Kenneth W. Witwer, Matias Ostrowski, Georg H. H. Borner, and Clotilde Théry. 2020. "Proteomic Profiling Allows Unbiased Compositional

Analysis of Extracellular Vesicle Subtypes." SSRN Scholarly Paper ID 3516102. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3516102.

McDermott, Martina Sj, Brigid C. Browne, Neil T. Conlon, Neil A. O'Brien, Dennis J. Slamon, Michael Henry, Paula Meleady, et al. 2014. "Pp2a Inhibition Overcomes Acquired Resistance to Her2 Targeted Therapy." *Molecular Cancer* 13 (June): 157. https://doi.org/10.1186/1476-4598-13-157.

McLuckey, Scott A. 1992. "Principles of Collisional Activation in Analytical Mass Spectrometry." *Journal of the American Society for Mass Spectrometry* 3 (6): 599–614. https://doi.org/10.1016/1044-0305(92)85001-Z.

Mellacheruvu, Dattatreya, Zachary Wright, Amber L. Couzens, Jean-Philippe Lambert, Nicole St-Denis, Tuo Li, Yana V. Miteva, et al. 2013. "The CRAPome: A Contaminant Repository for Affinity Purification Mass Spectrometry Data." *Nature Methods* 10 (8): 730–36. https://doi.org/10.1038/nmeth.2557.

Millar, Harvey, and Nicolas Taylor. 2014. "Subcellular Proteomics—Where Cell Biology Meets Protein Chemistry." *Frontiers in Plant Science* 5: 55. https://doi.org/10.3389/fpls.2014.00055.

Milosevic, Marina, Dragan Jankovic, Aleksandar Milenkovic, and Dragan Stojanov. 2018. "Early Diagnosis and Detection of Breast Cancer." *Technology and Health Care: Official Journal of the European Society for Engineering and Medicine* 26 (4): 729–59. https://doi.org/10.3233/THC-181277.

Morello, Matteo, Valentina Minciacchi, Paola de Candia, Julie Yang, Edwin Posadas, Hyung Kim, Duncan Griffiths, et al. 2013. "Large Oncosomes Mediate Intercellular Transfer of Functional microRNA." *Cell Cycle* 12 (22): 3526–36. https://doi.org/10.4161/cc.26539.

Mustacchi, G., L. Biganzoli, P. Pronzato, F. Montemurro, M. Dambrosio, M. Minelli, L. Molteni, and L. Scaltriti. 2015. "HER2-positive Metastatic Breast Cancer: A Changing Scenario." *Critical Reviews in Oncology/Hematology* 95 (1): 78–87. https://doi.org/10.1016/j.critrevonc.2015.02.002.

Nadler, Wiebke Maria, Dietmar Waidelich, Alexander Kerner, Sabrina Hanke, Regina Berg, Andreas Trumpp, and Christoph Rösli. 2017. "MALDI Versus ESI: The Impact of the Ion Source on Peptide Identification." *Journal of Proteome Research* 16 (3): 1207–15. https://doi.org/10.1021/acs.jproteome.6b00805.

Nagaraj, Nagarjuna, Nils Alexander Kulak, Juergen Cox, Nadin Neuhauser, Korbinian Mayr, Ole Hoerning, Ole Vorm, and Matthias Mann. 2012. "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: MCP* 11 (3): M111.013722. https://doi.org/10.1074/mcp.M111.013722.

Nakai, Wataru, Takeshi Yoshida, Diego Diez, Yuji Miyatake, Takahiro Nishibu, Naoko Imawaka, Ken Naruse, Yoshifusa Sadamura, and Rikinari Hanayama. 2016. "A Novel Affinity-Based Method for the Isolation of Highly Purified Extracellu-

lar Vesicles." *Scientific Reports* 6 (1, 1): 33935. https://doi.org/10.1038/srep33935.

Navas, Tony, Robert J. Kinders, Scott M. Lawrence, Katherine V. Ferry-Galow, Suzanne Borgel, Melinda G. Hollingshead, Apurva K. Srivastava, et al. 2020. "Clinical Evolution of Epithelial–Mesenchymal Transition in Human Carcinomas." *Cancer Research* 80 (2): 304–18. https://doi.org/10.1158/0008-5472.CAN-18-3539.

Nawaz, Muhammad, Neelam Shah, Bruna Riedo Zanetti, Marco Maugeri, Renata Nacasaki Silvestre, Farah Fatima, Luciano Neder, and Hadi Valadi. 2018. "Extracellular Vesicles and Matrix Remodeling Enzymes: The Emerging Roles in Extracellular Matrix Remodeling, Progression of Diseases and Tissue Repair." *Cells* 7 (10): 167. https://doi.org/10.3390/cells7100167.

Niel, Guillaume van, Gisela D'Angelo, and Graça Raposo. 2018. "Shedding Light on the Cell Biology of Extracellular Vesicles." *Nature Reviews Molecular Cell Biology* 19 (4, 4): 213–28. https://doi.org/10.1038/nrm.2017.125.

Niu, Liangjie, Huayi Yuan, Fangping Gong, Xiaolin Wu, and Wei Wang. 2018. "Protein Extraction Methods Shape Much of the Extracted Proteomes." *Frontiers in Plant Science* 0. https://doi.org/10.3389/fpls.2018.00802.

Nolan, John P., and Erika Duggan. 2018. "Analysis of Individual Extracellular Vesicles by Flow Cytometry." *Methods in Molecular Biology (Clifton, N.J.)* 1678: 79–92. https://doi.org/10.1007/978-1-4939-7346-0_5.

Ohno, Shin-ichiro, Akio Ishikawa, and Masahiko Kuroda. 2013. "Roles of Exosomes

and Microvesicles in Disease Pathogenesis." *Advanced Drug Delivery Reviews* 65 (3): 398–401. https://doi.org/10.1016/j.addr.2012.07.019.

Onódi, Zsófia, Csilla Pelyhe, Csilla Terézia Nagy, Gábor B. Brenner, Laura Almási, Ágnes Kittel, Mateja Manček-Keber, Péter Ferdinandy, Edit I. Buzás, and Zoltán Giricz. 2018. "Isolation of High-Purity Extracellular Vesicles by the Combination of Iodixanol Density Gradient Ultracentrifugation and Bind-Elute Chromatography From Blood Plasma." *Frontiers in Physiology* 9: 1479. https://doi.org/10.3389/fphys.2018.01479.

Orsburn, Benjamin C. 2021. "Proteome Discoverer—A Community Enhanced Data Processing Suite for Protein Informatics." *Proteomes* 9 (1): 15. https://doi.org/10.3390/proteomes9010015.

Osanai, Makoto, Akira Takasawa, Masaki Murata, and Norimasa Sawada. 2017. "Claudins in Cancer: Bench to Bedside." *Pflugers Archiv: European Journal of Physiology* 469 (1): 55–67. https://doi.org/10.1007/s00424-016-1877-7.

Palviainen, Mari, Mayank Saraswat, Zoltán Varga, Diána Kitka, Maarit Neuvonen, Maija Puhka, Sakari Joenväärä, et al. 2020. "Extracellular Vesicles from Human Plasma and Serum Are Carriers of Extravesicular Cargo—Implications for Biomarker Discovery." Edited by John Matthew Koomen. *PLOS ONE* 15 (8): e0236439. https://doi.org/10.1371/journal.pone.0236439.

Pan, B. T., and R. M. Johnstone. 1983. "Fate of the Transferrin Receptor During Maturation of Sheep Reticulocytes in Vitro: Selective Externalization of the

Receptor." *Cell* 33 (3): 967–78. https://doi.org/10.1016/0092-8674(83)90040-5.

Pankow, Sandra, Salvador Martínez-Bartolomé, Casimir Bamberger, and John R Yates. 2019. "Understanding Molecular Mechanisms of Disease Through Spatial Proteomics." *Current Opinion in Chemical Biology* 48 (February): 19–25. https://doi.org/10.1016/j.cbpa.2018.09.016.

Patel, Girijesh Kumar, Mohammad Aslam Khan, Haseeb Zubair, Sanjeev Kumar Srivastava, Moh'd Khushman, Seema Singh, and Ajay Pratap Singh. 2019. "Comparative Analysis of Exosome Isolation Methods Using Culture Supernatant for Optimum Yield, Purity and Downstream Applications." *Scientific Reports* 9 (1, 1): 5335. https://doi.org/10.1038/s41598-019-41800-2.

Petri, Belinda J., and Carolyn M. Klinge. 2020. "Regulation of Breast Cancer Metastasis Signaling by miRNAs." *Cancer Metastasis Reviews* 39 (3): 837–86. https://doi.org/10.1007/s10555-020-09905-7.

Pérez-Fernández, Jorge, Ángel Román, Javier De Las Rivas, Xosé R. Bustelo, and Mercedes Dosil. 2007. "The 90s Preribosome Is a Multimodular Structure That Is Assembled Through a Hierarchical Mechanism." *Molecular and Cellular Biology* 27 (15): 5414–29. https://doi.org/10.1128/MCB.00380-07.

Piñero, Janet, Juan Manuel Ramírez-Anguita, Josep Saüch-Pitarch, Francesco Ronzano, Emilio Centeno, Ferran Sanz, and Laura I. Furlong. 2020. "The DisGeNET Knowledge Platform for Disease Genomics: 2019 Update." *Nucleic Acids Research*

48 (D1): D845–55. https://doi.org/10.1093/nar/gkz1021.

Puhka, Maija, Maarit Takatalo, Maria-Elisa Nordberg, Sami Valkonen, Jatin Nandania, Maria Aatonen, Marjo Yliperttula, et al. 2017. "Metabolomic Profiling of Extracellular Vesicles and Alternative Normalization Methods Reveal Enriched Metabolites and Strategies to Study Prostate Cancer-Related Changes." *Theranostics* 7 (16): 3824–41. https://doi.org/10.7150/thno.19890.

Rai, Alin, Haoyun Fang, Bethany Claridge, Richard J. Simpson, and David W Greening. 2021. "Proteomic Dissection of Large Extracellular Vesicle Surfaceome Unravels Interactive Surface Platform." *Journal of Extracellular Vesicles* 10 (13): e12164. https://doi.org/10.1002/jev2.12164.

Rao, Rammohan V., Karen S. Poksay, Susana Castro-Obregon, Birgit Schilling, Richard H. Row, Gabriel del Rio, Bradford W. Gibson, H. Michael Ellerby, and Dale E. Bredesen. 2004. "Molecular Components of a Cell Death Pathway Activated by Endoplasmic Reticulum Stress." *The Journal of Biological Chemistry* 279 (1): 177–87. https://doi.org/10.1074/jbc.M304490200.

Raposo, Graça, and Willem Stoorvogel. 2013. "Extracellular Vesicles: Exosomes, Microvesicles, and Friends." *Journal of Cell Biology* 200 (4): 373–83. https://doi.org/10.1083/jcb.201211138.

Ravenhill, Benjamin J., Lior Soday, Jack Houghton, Robin Antrobus, and Michael P. Weekes. 2020. "Comprehensive Cell Surface Proteomics Defines Markers of Classical, Intermediate and Non-Classical Monocytes." *Scientific Reports* 10 (1,

1): 4560. https://doi.org/10.1038/s41598-020-61356-w.

Risha, Yousef, Zoran Minic, Shahrokh M. Ghobadloo, and Maxim V. Berezovski. 2020. "The Proteomic Analysis of Breast Cell Line Exosomes Reveals Disease Patterns and Potential Biomarkers." *Scientific Reports* 10 (1, 1): 13572. https://doi.org/10.1038/s41598-020-70393-4.

Risha, Yousef, Vanessa Susevski, Nico Hüttmann, Suttinee Poolsup, Zoran Minic, and Maxim V. Berezovski. 2021. "Breast Cancer-Derived Microvesicles Are the Source of Functional Metabolic Enzymes as Potential Targets for Cancer Therapy." *Biomedicines* 9 (2, 2): 107. https://doi.org/10.3390/biomedicines9020107.

Roepstorff, P., and J. Fohlman. 1984. "Proposal for a Common Nomenclature for Sequence Ions in Mass Spectra of Peptides." *Biomedical Mass Spectrometry* 11 (11): 601. https://doi.org/10.1002/bms.1200111109.

Rontogianni, Stamatia, Eleni Synadaki, Bohui Li, Marte C. Liefaard, Esther H. Lips, Jelle Wesseling, Wei Wu, and Maarten Altelaar. 2019. "Proteomic Profiling of Extracellular Vesicles Allows for Human Breast Cancer Subtyping." *Communications Biology* 2 (1, 1): 1–13. https://doi.org/10.1038/s42003-019-0570-8.

Sanderson, Ralph D., Shyam K. Bandari, and Israel Vlodavsky. 2019. "Proteases and Glycosidases on the Surface of Exosomes: Newly Discovered Mechanisms for Extracellular Remodeling." *Matrix Biology : Journal of the International Society for Matrix Biology* 75–76 (January): 160–69. https://doi.org/10.1016/j.matbio.2017.10.007.

Saveliev, Sergei, Mark Bratz, Roman Zubarev, Matt Szapacs, Harshavardhan Budamgunta, and Marjeta Urh. 2013. "Trypsin/Lys-C Protease Mix for Enhanced Protein Mass Spectrometry Analysis." *Nature Methods* 10 (11, 11): i–ii. https://doi.org/10.1038/nmeth.f.371.

Sayyadi, Nima, Sareh Zhand, Sajad Razavi Bazaz, and Majid Ebrahimi Warkiani. 2021. "Affibody Functionalized Beads for the Highly Sensitive Detection of Cancer Cell-Derived Exosomes." *International Journal of Molecular Sciences* 22 (21): 12014. https://doi.org/10.3390/ijms222112014.

Schmidt, Andreas, Ignasi Forne, and Axel Imhof. 2014. "Bioinformatic Analysis of Proteomics Data." *BMC Systems Biology* 8 (2): S3. https://doi.org/10.1186/1752-0509-8-S2-S3.

Schrier, S. L., D. Godin, R. G. Gould, B. Swyryd, I. Junga, and M. Seeger. 1971. "Characterization of Microvesicles Produced by Shearing of Human Erythrocyte Membranes." *Biochimica Et Biophysica Acta* 233 (1): 26–36. https://doi.org/10.1016/0005-2736(71)90354-3.

Shaw, Abigail, Marcus D. Bradley, Sean Elyan, and Kathreena M. Kurian. 2015. "Tumour Biomarkers: Diagnostic, Prognostic, and Predictive." *BMJ* 351 (July): h3449. https://doi.org/10.1136/bmj.h3449.

Shevchenko, Andrej, Henrik Tomas, Jan Havli, Jesper V. Olsen, and Matthias Mann. 2006. "In-Gel Digestion for Mass Spectrometric Characterization of Proteins and Proteomes." *Nature Protocols* 1 (6, 6): 2856–60. https://doi.org/10.1038/

nprot.2006.468.

Shi, Jian-Hui, and Ying-Jie Hao. 2019. "Ddx10 Overexpression Predicts Worse Prognosis in Osteosarcoma and Its Deletion Prohibits Cell Activities Modulated by MAPK Pathway." *Biochemical and Biophysical Research Communications* 510 (4): 525–29. https://doi.org/10.1016/j.bbrc.2019.01.114.

Shibata, Hideki, Keiko Yamada, Takako Mizuno, Chiharu Yorikawa, Hiroshi Takahashi, Hirokazu Satoh, Yasuyuki Kitaura, and Masatoshi Maki. 2004. "The Penta-EF-hand Protein ALG-2 Interacts with a Region Containing PxY Repeats in Alix/Aip1, Which Is Required for the Subcellular Punctate Distribution of the Amino-Terminal Truncation Form of Alix/Aip1." *Journal of Biochemistry* 135 (1): 117–28. https://doi.org/10.1093/jb/mvh014.

Siegel, Rebecca L., Kimberly D. Miller, and Ahmedin Jemal. 2015 Jan-Feb. "Cancer Statistics, 2015." *CA: A Cancer Journal for Clinicians* 65 (1): 5–29. https://doi.org/10.3322/caac.21254.

Soini, Y. 2005. "Expression of Claudins 1, 2, 3, 4, 5 and 7 in Various Types of Tumours." *Histopathology* 46 (5): 551–60. https://doi.org/10.1111/j.1365-2559.2005.02127.x.

Soini, Ylermi. 2011. "Claudins in Lung Diseases." *Respiratory Research* 12 (1): 70. https://doi.org/10.1186/1465-9921-12-70.

Sonnenblick, A., and M. Piccart. 2015. "Adjuvant Systemic Therapy in Breast Cancer: Quo Vadis?" *Annals of Oncology: Official Journal of the European So-*

ciety for Medical Oncology 26 (8): 1629–34. https://doi.org/10.1093/annonc/mdv108.

Stuffers, Susanne, Catherine Sem Wegner, Harald Stenmark, and Andreas Brech. 2009. "Multivesicular Endosome Biogenesis in the Absence of ESCRTs." *Traffic* 10 (7): 925–37. https://doi.org/10.1111/j.1600-0854.2009.00920.x.

Su, Dan, Haiyan Xu, Jianguo Feng, Yun Gao, Linhui Gu, Lisha Ying, Dionyssios Katsaros, Herbert Yu, Shenhua Xu, and Ming Qi. 2012. "Pdcd6 Is an Independent Predictor of Progression Free Survival in Epithelial Ovarian Cancer." *Journal of Translational Medicine* 10 (February): 31. https://doi.org/10.1186/1479-5876-10-31.

Su, Fengxi, Xiaoqu Hu, Weijuan Jia, Chang Gong, Erwei Song, and Peter Hamar. 2003. "Glutathion S Transferase Pi Indicates Chemotherapy Resistance in Breast Cancer." *The Journal of Surgical Research* 113 (1): 102–8. https://doi.org/10.1016/s0022-4804(03)00200-2.

Sun, In O., and Lilach O. Lerman. 2020. "Urinary Extracellular Vesicles as Biomarkers of Kidney Disease: From Diagnostics to Therapeutics." *Diagnostics* 10 (5): 311. https://doi.org/10.3390/diagnostics10050311.

Sun, Long-Ci, and Hai-Xin Qian. 2018. "Screening for Implicated Genes in Colorectal Cancer Using Whole-Genome Gene Expression Profiling." *Molecular Medicine Reports* 17 (6): 8260–68. https://doi.org/10.3892/mmr.2018.8862.

Tai, Yu-Ling, Ko-Chien Chen, Jer-Tsong Hsieh, and Tang-Long Shen. 2018. "Exo-

somes in Cancer Development and Clinical Applications." *Cancer Science* 109 (8): 2364–74. https://doi.org/10.1111/cas.13697.

Takeuchi, T., J. L. Harris, W. Huang, K. W. Yan, S. R. Coughlin, and C. S. Craik. 2000. "Cellular Localization of Membrane-Type Serine Protease 1 and Identification of Protease-Activated Receptor-2 and Single-Chain Urokinase-Type Plasminogen Activator as Substrates." *The Journal of Biological Chemistry* 275 (34): 26333–42. https://doi.org/10.1074/jbc.M002941200.

Tan, Sandra, Hwee Tong Tan, and Maxey C. M. Chung. 2008. "Membrane Proteins and Membrane Proteomics." *PROTEOMICS* 8 (19): 3924–32. https://doi.org/10.1002/pmic.200800597.

Tangutoori, Shifalika, Paige Baldwin, and Srinivas Sridhar. 2015. "PARP Inhibitors: A New Era of Targeted Therapy." *Maturitas* 81 (1): 5–9. https://doi.org/10.1016/j.maturitas.2015.01.015.

Tauro, Bow J., David W. Greening, Rommel A. Mathias, Hong Ji, Suresh Mathivanan, Andrew M. Scott, and Richard J. Simpson. 2012. "Comparison of Ultracentrifugation, Density Gradient Separation, and Immunoaffinity Capture Methods for Isolating Human Colon Cancer Cell Line LIM1863-derived Exosomes." *Methods*, Immunoaffinity methods and related methods, 56 (2): 293–304. https://doi.org/10.1016/j.ymeth.2012.01.002.

Théry, Clotilde, Sebastian Amigorena, Graça Raposo, and Aled Clayton. 2006. "Isolation and Characterization of Exosomes from Cell Culture Supernatants and Bi-

ological Fluids." *Current Protocols in Cell Biology* 30 (1). https://doi.org/10.1002/0471143030.cb0322s30.

Théry, Clotilde, Kenneth W Witwer, Elena Aikawa, Maria Jose Alcaraz, Johnathon D Anderson, Ramaroson Andriantsitohaina, Anna Antoniou, et al. 2018. "Minimal Information for Studies of Extracellular Vesicles 2018 (Misev2018): A Position Statement of the International Society for Extracellular Vesicles and Update of the Misev2014 Guidelines." *Journal of Extracellular Vesicles* 7 (1): 1535750. https://doi.org/10.1080/20013078.2018.1535750.

Tian, Fei, Shaohua Zhang, Chao Liu, Ziwei Han, Yuan Liu, Jinqi Deng, Yike Li, et al. 2021. "Protein Analysis of Extracellular Vesicles to Monitor and Predict Therapeutic Response in Metastatic Breast Cancer." *Nature Communications* 12 (1, 1): 2536. https://doi.org/10.1038/s41467-021-22913-7.

Tipsmark, C. K., and S. S. Madsen. 2012. "Tricellulin, Occludin and Claudin-3 Expression in Salmon Intestine and Kidney During Salinity Adaptation." *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* 162 (4): 378–85. https://doi.org/10.1016/j.cbpa.2012.04.020.

Todd, Maria C., Heather M. Petty, Jonathan M. King, Brytanie N. Piana Marshall, Rebecca A. Sheller, and Maria E. Cuevas. 2015. "Overexpression and Delocalization of Claudin-3 Protein in MCF-7 and MDA-MB-415 Breast Cancer Cell Lines." *Oncology Letters* 10 (1): 156–62. https://doi.org/10.3892/ol.2015.3160.

Tóth, Eszter Á., Lilla Turiák, Tamás Visnovitz, Csaba Cserép, Anett Mázló, Barbara W. Sódar, András I. Försönits, et al. 2021. "Formation of a Protein Corona on the Surface of Extracellular Vesicles in Blood Plasma." *Journal of Extracellular Vesicles* 10 (11): e12140. https://doi.org/10.1002/jev2.12140.

Trajkovic, Katarina, Chieh Hsu, Salvatore Chiantia, Lawrence Rajendran, Dirk Wenzel, Felix Wieland, Petra Schwille, Britta Brügger, and Mikael Simons. 2008. "Ceramide Triggers Budding of Exosome Vesicles into Multivesicular Endosomes." *Science (New York, N.Y.)* 319 (5867): 1244–47. https://doi.org/10.1126/science.1153124.

Tsoucas, Daphne, Rui Dong, Haide Chen, Qian Zhu, Guoji Guo, and Guo-Cheng Yuan. 2019. "Accurate Estimation of Cell-Type Composition from Gene Expression Data." *Nature Communications* 10 (1, 1): 2975. https://doi.org/10.1038/s41467-019-10802-z.

Turchinovich, Andrey, Oxana Drapkina, and Alexander Tonevitsky. 2019. "Transcriptome of Extracellular Vesicles: State-of-the-Art." *Frontiers in Immunology* 10 (February): 202. https://doi.org/10.3389/fimmu.2019.00202.

Venter, J. Craig, Mark D. Adams, Eugene W. Myers, Peter W. Li, Richard J. Mural, Granger G. Sutton, Hamilton O. Smith, et al. 2001. "The Sequence of the Human Genome." *Science* 291 (5507): 1304–51. https://doi.org/10.1126/science.1058040.

Vlastos, Georges, and Helena M. Verkooijen. 2007. "Minimally Invasive Approaches

for Diagnosis and Treatment of Early-Stage Breast Cancer." *The Oncologist* 12 (1): 1–10. https://doi.org/10.1634/theoncologist.12-1-1.

Wang, Lulu. 2017. "Early Diagnosis of Breast Cancer." *Sensors (Basel, Switzerland)* 17 (7): E1572. https://doi.org/10.3390/s17071572.

Wang, Ruo-Xi, Xiao-En Xu, Liang Huang, Sheng Chen, and Zhi-Ming Shao. 2019. "eEF2 Kinase Mediated Autophagy as a Potential Therapeutic Target for Paclitaxel-Resistant Triple-Negative Breast Cancer." *Annals of Translational Medicine* 7 (23): 783. https://doi.org/10.21037/atm.2019.11.39.

Wang, Ying, Wen-Ming Xiao, Shu-Ming Wei, Xiang-Ming Chen, Lin Wei, Rui-Hua Tian, Li Meng, Bao-Rong Xiao, Ping-Xia Wu, and Yong-Hua Yu. 2017. "Highly Expressed Ddx10 Promotes Hepatocellular Carcinoma Cell Proliferation Through Wnt/Beta-Catenin Signaling," 7.

Wang, Zhendong, Zhonghua Luo, Lin Zhou, Xiaofei Li, Tao Jiang, and Enqing Fu. 2015. "Ddx5 Promotes Proliferation and Tumorigenesis of Non-Small-Cell Lung Cancer Cells by Activating Beta-Catenin Signaling Pathway." *Cancer Science* 106 (10): 1303–12. https://doi.org/10.1111/cas.12755.

Wei, Feng, Chengyuan Ma, Tong Zhou, Xuechao Dong, Qinghua Luo, Li Geng, Lijuan Ding, et al. 2017. "Exosomes Derived from Gemcitabine-Resistant Cells Transfer Malignant Phenotypic Traits via Delivery of miRNA-222-3p." *Molecular Cancer* 16 (1): 132. https://doi.org/10.1186/s12943-017-0694-8.

Wewer Albrechtsen, Nicolai J., Philipp E. Geyer, Sophia Doll, Peter V. Treit, Kirs-

tine N. Bojsen-Møller, Christoffer Martinussen, Nils B. Jørgensen, et al. 2018. "Plasma Proteome Profiling Reveals Dynamics of Inflammatory and Lipid Homeostasis Markers After Roux-En-Y Gastric Bypass Surgery." *Cell Systems* 7 (6): 601–612.e3. https://doi.org/10.1016/j.cels.2018.10.012.

Willms, Eduard, Carlos Cabañas, Imre Mäger, Matthew J. A. Wood, and Pieter Vader. 2018. "Extracellular Vesicle Heterogeneity: Subpopulations, Isolation Techniques, and Diverse Functions in Cancer Progression." *Frontiers in Immunology* 9: 738. https://doi.org/10.3389/fimmu.2018.00738.

Wiśniewski, Jacek R., Alexandre Zougman, Nagarjuna Nagaraj, and Matthias Mann. 2009. "Universal Sample Preparation Method for Proteome Analysis." *Nature Methods* 6 (5, 5): 359–62. https://doi.org/10.1038/nmeth.1322.

Witwer, Kenneth W., Edit I. Buzás, Lynne T. Bemis, Adriana Bora, Cecilia Lässer, Jan Lötvall, Esther N. Nolte-'t Hoen, et al. 2013. "Standardization of Sample Collection, Isolation and Analysis Methods in Extracellular Vesicle Research." *Journal of Extracellular Vesicles* 2 (1): 20360. https://doi.org/10.3402/jev.v2i0.20360.

Wolf, Peter. 1967. "The Nature and Significance of Platelet Products in Human Plasma." *British Journal of Haematology* 13 (3): 269–88. https://doi.org/10.1111/j.1365-2141.1967.tb08741.x.

Worst, Thomas Stefan, Jost von Hardenberg, Julia Christina Gross, Philipp Erben, Martina Schnölzer, Ingrid Hausser, Peter Bugert, Maurice Stephan Michel, and

Michael Boutros. 2017. "Database-Augmented Mass Spectrometry Analysis of Exosomes Identifies Claudin 3 as a Putative Prostate Cancer Biomarker." *Molecular & Cellular Proteomics: MCP* 16 (6): 998–1008. https://doi.org/10.1074/mcp.M117.068577.

Wu, Jun-Rong, Yang Zhao, Xiao-Ping Zhou, and Xue Qin. 2020. "Estrogen Receptor 1 and Progesterone Receptor Are Distinct Biomarkers and Prognostic Factors in Estrogen Receptor-Positive Breast Cancer: Evidence from a Bioinformatic Analysis." *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie* 121 (January): 109647. https://doi.org/10.1016/j.biopha.2019.109647.

Wu, Xinyan, Muhammad Saddiq Zahari, Binyun Ma, Ren Liu, Santosh Renuse, Nandini A. Sahasrabuddhe, Lily Chen, et al. 2015. "Global Phosphotyrosine Survey in Triple-Negative Breast Cancer Reveals Activation of Multiple Tyrosine Kinase Signaling Pathways." *Oncotarget* 6 (30): 29143–60. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4745717/.

Xin, Hongqi, Yi Li, Ben Buller, Mark Katakowski, Yi Zhang, Xinli Wang, Xia Shang, Zheng Gang Zhang, and Michael Chopp. 2012. "Exosome-Mediated Transfer of miR-133b from Multipotent Mesenchymal Stromal Cells to Neural Cells Contributes to Neurite Outgrowth." *Stem Cells (Dayton, Ohio)* 30 (7): 1556–64. https://doi.org/10.1002/stem.1129.

Yang, Gong, Xiao-Ou Shu, Zhi-Xian Ruan, Qiu-Yin Cai, Fan Jin, Yu-Tang Gao, and Wei Zheng. 2005. "Genetic Polymorphisms in Glutathione-S-transferase Genes

(Gstm1, Gstt1, Gstp1) and Survival After Chemotherapy for Invasive Breast Carcinoma." *Cancer* 103 (1): 52–58. https://doi.org/10.1002/cncr.20729.

Yang, Su-jin, Dan-dan Wang, Jian Li, Han-zi Xu, Hong-yu Shen, Xiu Chen, Si-ying Zhou, Shan-liang Zhong, Jian-hua Zhao, and Jin-hai Tang. 2017. "Predictive Role of GSTP1-containing Exosomes in Chemotherapy-Resistant Breast Cancer." *Gene* 623 (August): 5–14. https://doi.org/10.1016/j.gene.2017.04.031.

Yassin, E. R., A. M. Abdul-Nabi, A. Takeda, and N. R. Yaseen. 2010. "Effects of the Nup98-Ddx10 Oncogene on Primary Human Cd34+ Cells: Role of a Conserved Helicase Motif." *Leukemia* 24 (5): 1001–11. https://doi.org/10.1038/leu.2010.42.

Yates, John R. 2004. "Mass Spectral Analysis in Proteomics." *Annual Review of Biophysics and Biomolecular Structure* 33: 297–316. https://doi.org/10.1146/annurev.biophys.33.111502.082538.

Yáñez-Mó, María, Pia R.-M. Siljander, Zoraida Andreu, Apolonija Bedina Zavec, Francesc E. Borràs, Edit I. Buzas, Krisztina Buzas, et al. 2015. "Biological Properties of Extracellular Vesicles and Their Physiological Functions." *Journal of Extracellular Vesicles* 4 (1): 27066. https://doi.org/10.3402/jev.v4.27066.

Yoshimura, Akihiko, Tetsuji Naka, and Masato Kubo. 2007. "SOCS Proteins, Cytokine Signalling and Immune Regulation." *Nature Reviews. Immunology* 7 (6): 454–65. https://doi.org/10.1038/nri2093.

Yu, Dan-dan, Ying Wu, Hong-yu Shen, Meng-meng Lv, Wei-xian Chen, Xiao-hui

Zhang, Shan-liang Zhong, Jin-hai Tang, and Jian-hua Zhao. 2015. "Exosomes in Development, Metastasis and Drug Resistance of Breast Cancer." *Cancer Science* 106 (8): 959–64. https://doi.org/10.1111/cas.12715.

Yuan, Fumin, Ya-Min Li, and Zhuhui Wang. 2021. "Preserving Extracellular Vesicles for Biomedical Applications: Consideration of Storage Stability Before and After Isolation." *Drug Delivery* 28 (1): 1501–9. https://doi.org/10.1080/10717544.2021.1951896.

Zaborowski, MikoŁaj P., Leonora Balaj, Xandra O. Breakefield, and Charles P. Lai. 2015. "Extracellular Vesicles: Composition, Biological Relevance, and Methods of Study." *Bioscience* 65 (8): 783–97. https://doi.org/10.1093/biosci/biv084.

Zeringer, Emily, Timothy Barta, Mu Li, and Alexander V. Vlassov. 2015. "Strategies for Isolation of Exosomes." *Cold Spring Harbor Protocols* 2015 (4): 319–23. https://doi.org/10.1101/pdb.top074476.

Zhang, Haiying, Daniela Freitas, Han Sang Kim, Kristina Fabijanic, Zhong Li, Haiyan Chen, Milica Tesic Mark, et al. 2018. "Identification of Distinct Nanoparticles and Subsets of Extracellular Vesicles by Asymmetric-Flow Field-Flow Fractionation." *Nature Cell Biology* 20 (3): 332–43. https://doi.org/10.1038/s41556-018-0040-4.

Zhang, Haiying, and David Lyden. 2019. "Asymmetric Flow Field-Flow Fractionation Technology for Exomere and Small Extracellular Vesicle Separation and Char-

acterization." *Nature Protocols* 14 (4): 1027–53. https://doi.org/10.1038/s41596-019-0126-x.

Zhang, Kui, Bin Zhou, Shaoqing Shi, Yaping Song, and Lin Zhang. 2013. "Variations in the Pdcd6 Gene Are Associated with Increased Uterine Leiomyoma Risk in the Chinese." *Genetic Testing and Molecular Biomarkers* 17 (7): 524–28. https://doi.org/10.1089/gtmb.2012.0461.

Zhang, Ling, Xiangming Chen, Baoli Liu, and Junqing Han. 2018. "MicroRNA-124-3p Directly Targets Pdcd6 to Inhibit Metastasis in Breast Cancer." *Oncology Letters* 15 (1): 984–90. https://doi.org/10.3892/ol.2017.7358.

Zhang, Xiaogang, Ellen G. F. Borg, A. Manuel Liaci, Harmjan R. Vos, and Willem Stoorvogel. 2020. "A Novel Three Step Protocol to Isolate Extracellular Vesicles from Plasma or Cell Culture Medium with Both High Yield and Purity." *Journal of Extracellular Vesicles* 9 (1): 1791450. https://doi.org/10.1080/20013078.2020.1791450.

Zhang, Yuan, Yunfeng Liu, Haiying Liu, and Wai Ho Tang. 2019. "Exosomes: Biogenesis, Biologic Function and Clinical Potential." *Cell & Bioscience* 9 (1): 19. https://doi.org/10.1186/s13578-019-0282-2.

Zhang, Zhenfeng, Jae Cheol Lee, Luping Lin, Victor Olivas, Valerie Au, Thomas LaFramboise, Mohamed Abdel-Rahman, et al. 2012. "Activation of the AXL Kinase Causes Resistance to EGFR-targeted Therapy in Lung Cancer." *Nature Genetics* 44 (8): 852–60. https://doi.org/10.1038/ng.2330.

Zhou, L., X.-D. Liu, M. Sun, X. Zhang, P. German, S. Bai, Z. Ding, et al. 2016. "Targeting MET and AXL Overcomes Resistance to Sunitinib Therapy in Renal Cell Carcinoma." *Oncogene* 35 (21): 2687–97. https://doi.org/10.1038/onc.2015.343.

Zoratti, Gina L., Lauren M. Tanabe, Fausto A. Varela, Andrew S. Murray, Christopher Bergum, Éloïc Colombo, Julie E. Lang, et al. 2015. "Targeting Matriptase in Breast Cancer Abrogates Tumour Progression via Impairment of Stromal-Epithelial Growth Factor Signalling." *Nature Communications* 6 (1, 1): 6776. https://doi.org/10.1038/ncomms7776.

Zubarev, Roman A., and Alexander Makarov. 2013. "Orbitrap Mass Spectrometry." *Analytical Chemistry* 85 (11): 5288–96. https://doi.org/10.1021/ac4001223.

Appendix

Figures

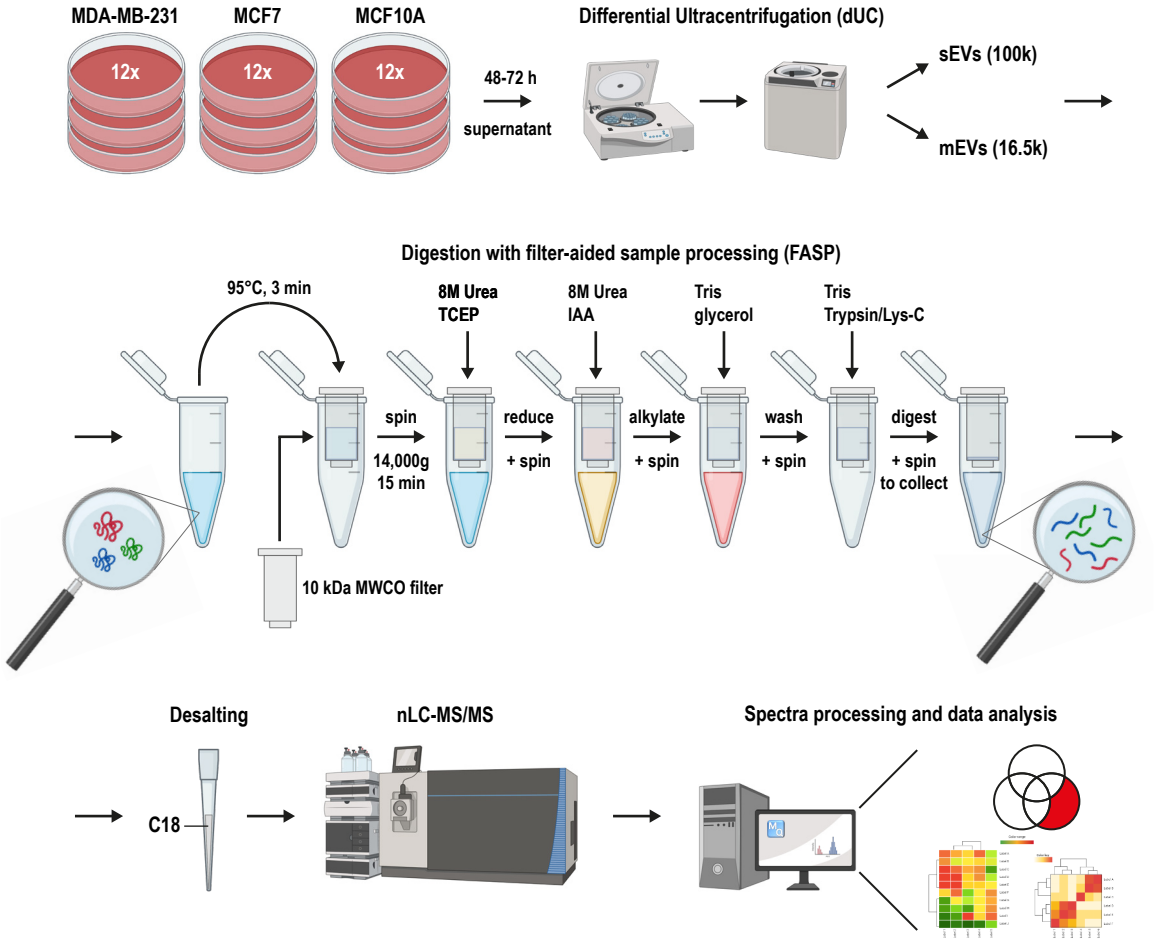

Fig. S1 Schematic workflow from cell culture to MS analysis of isolated EVs. Cell culture supernatant was collected and dUC was used to isolate sEVs and mEVs. EV proteins were digested with a FASP protocol, desalted with C18 TopTips and analyzed by nLC-MS/MS. Raw MS spectra were processed with MaxQuant and data analysis was conducted in R.

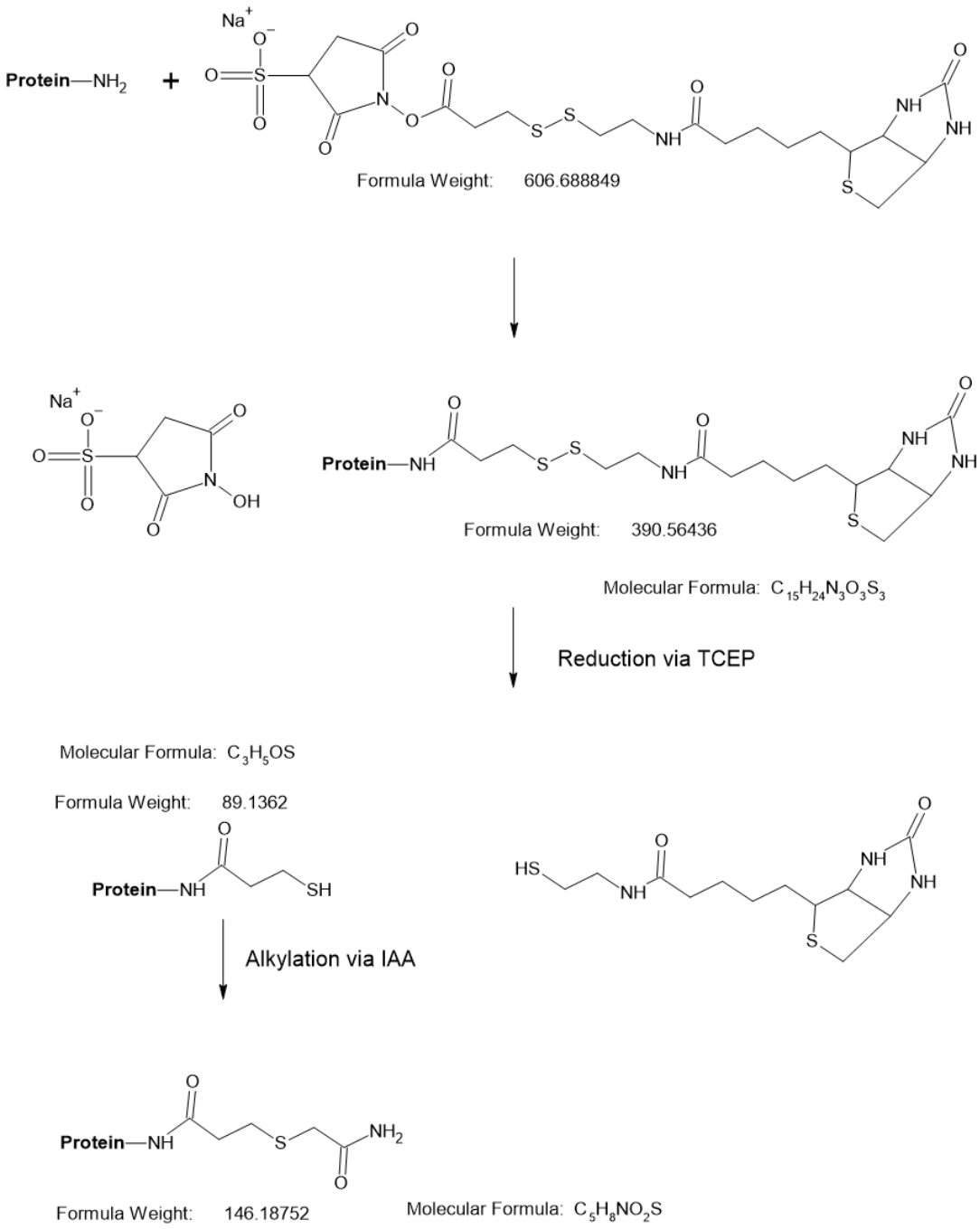

Fig. S2 Biotinylation of primary amines for surface protein identification. Sulfo-NHS-SS-Biotin reacts with lysine residues and protein N-termini via Michael addition.

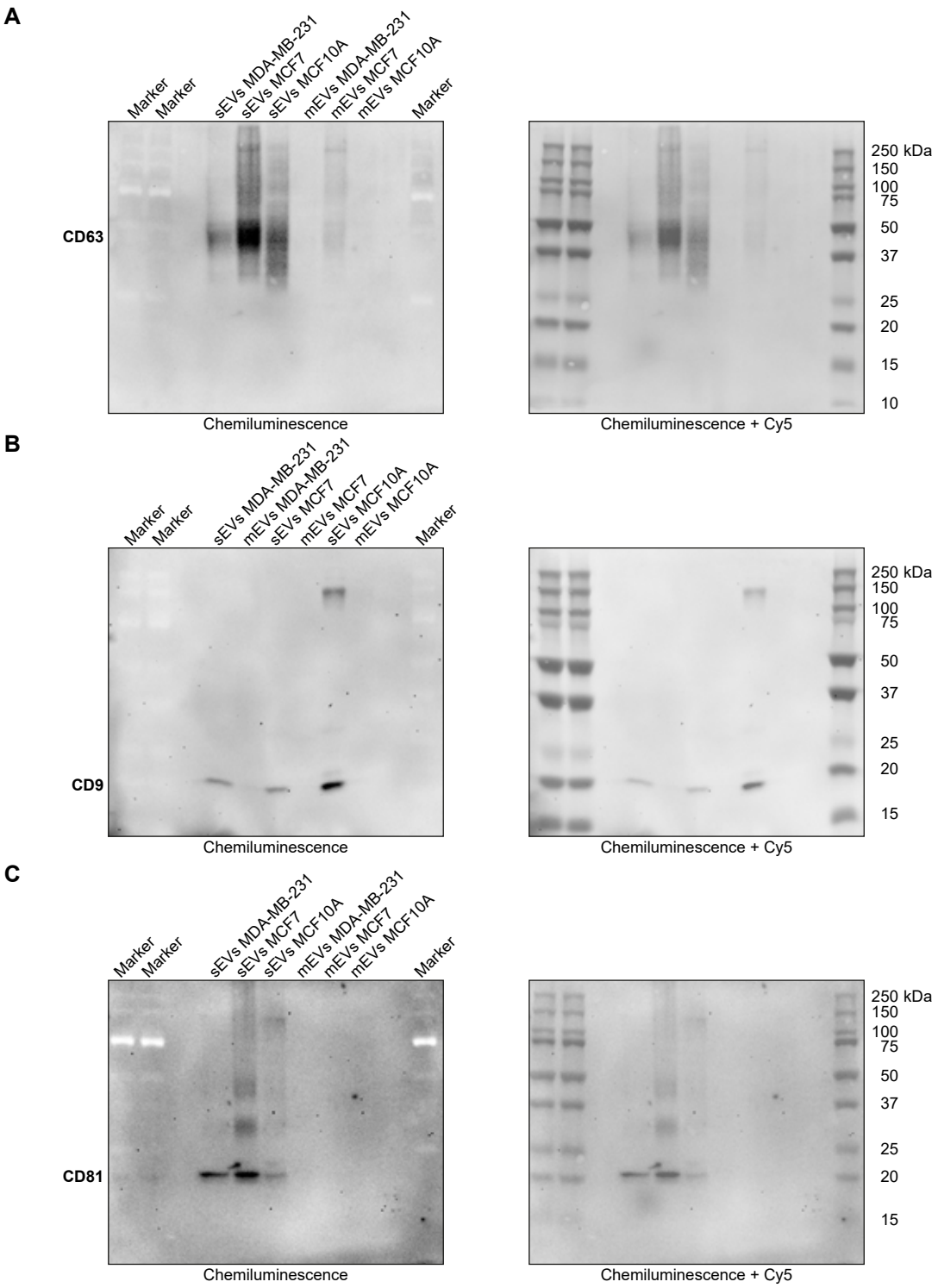

Fig. S3 Full WB images of CD63 (A), CD81 (B), and CD9 (C) presented in Fig. 3.1C.

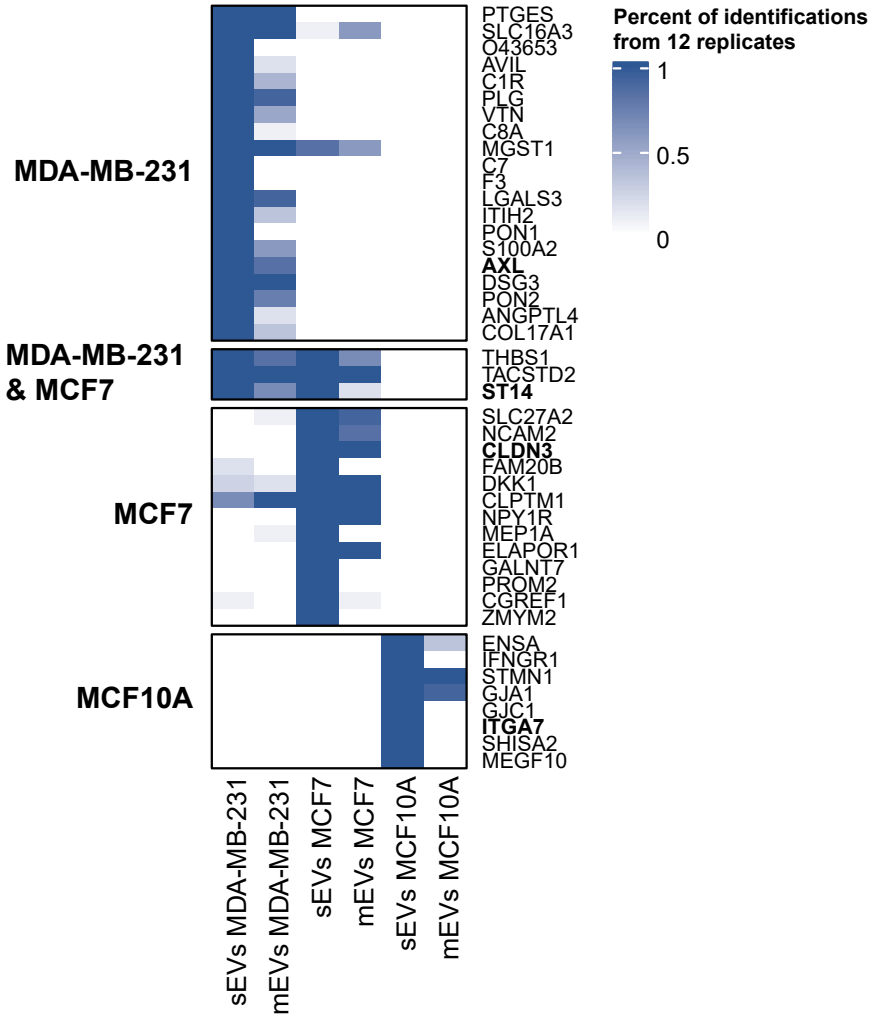

Fig. S4 Heatmap of cell line specific sEV proteins based on global proteomic profiling. Surface proteins are annotated as single- or multi-pass transmembrane proteins.

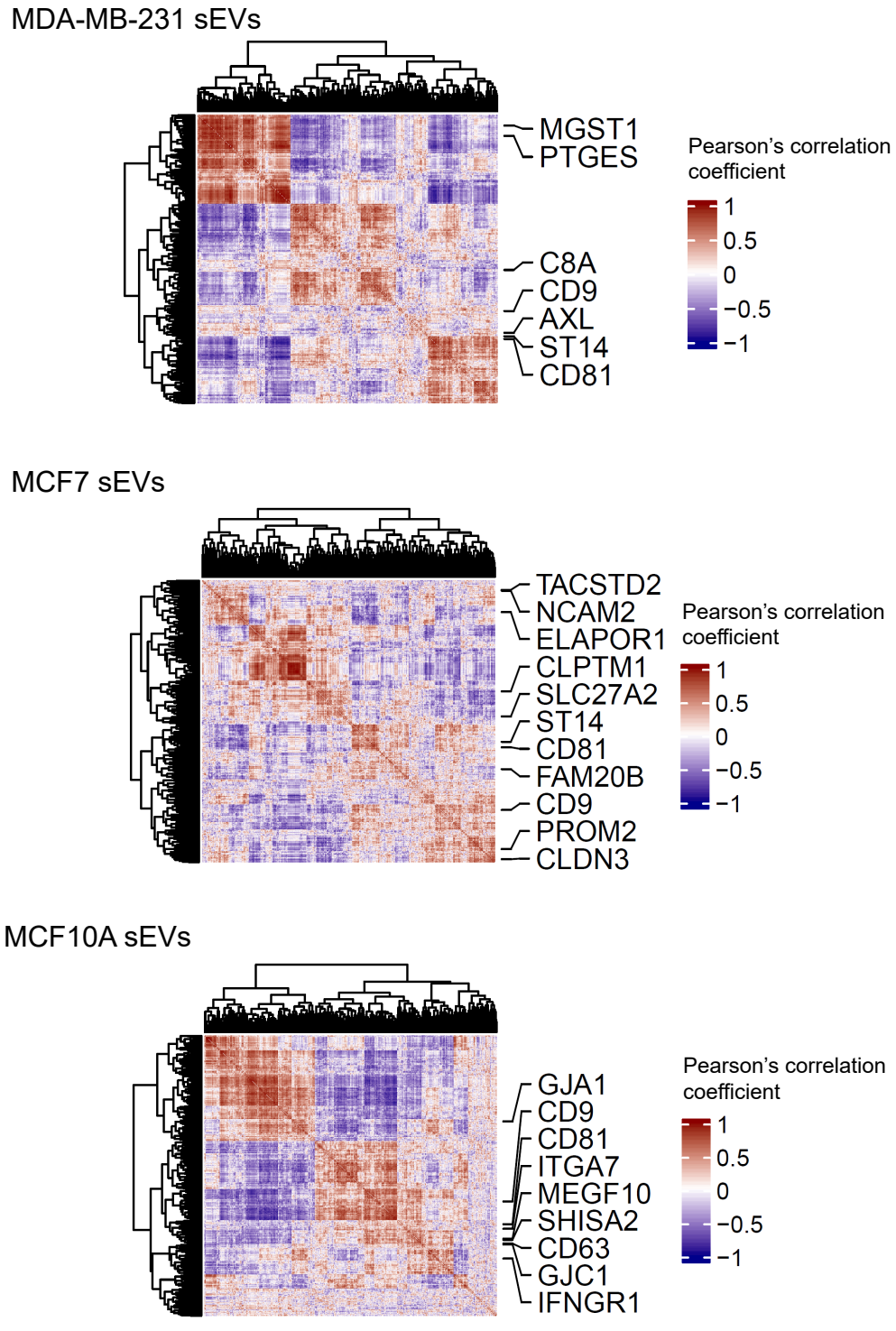

Fig. S5 Correlation matrices of sEV proteins predicting co-localization based on Pearson's correlation coefficient. The position of cell line-specific surface proteins and EV markers CD63, CD9 and CD81 has been indicated to highlight similar abundance profiles.

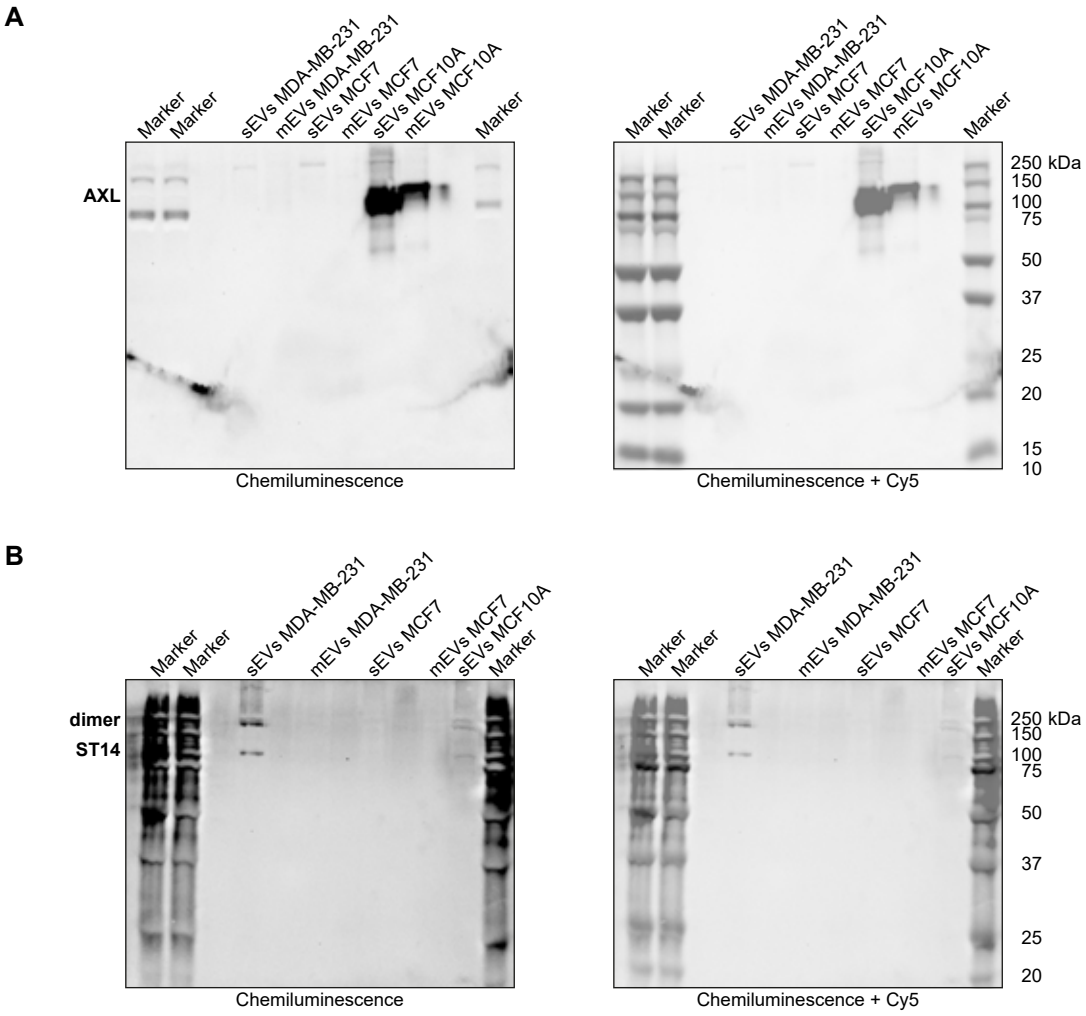

Fig. S6 Full WB images of AXL (A) and ST14 (B) presented in Fig. 3.5.

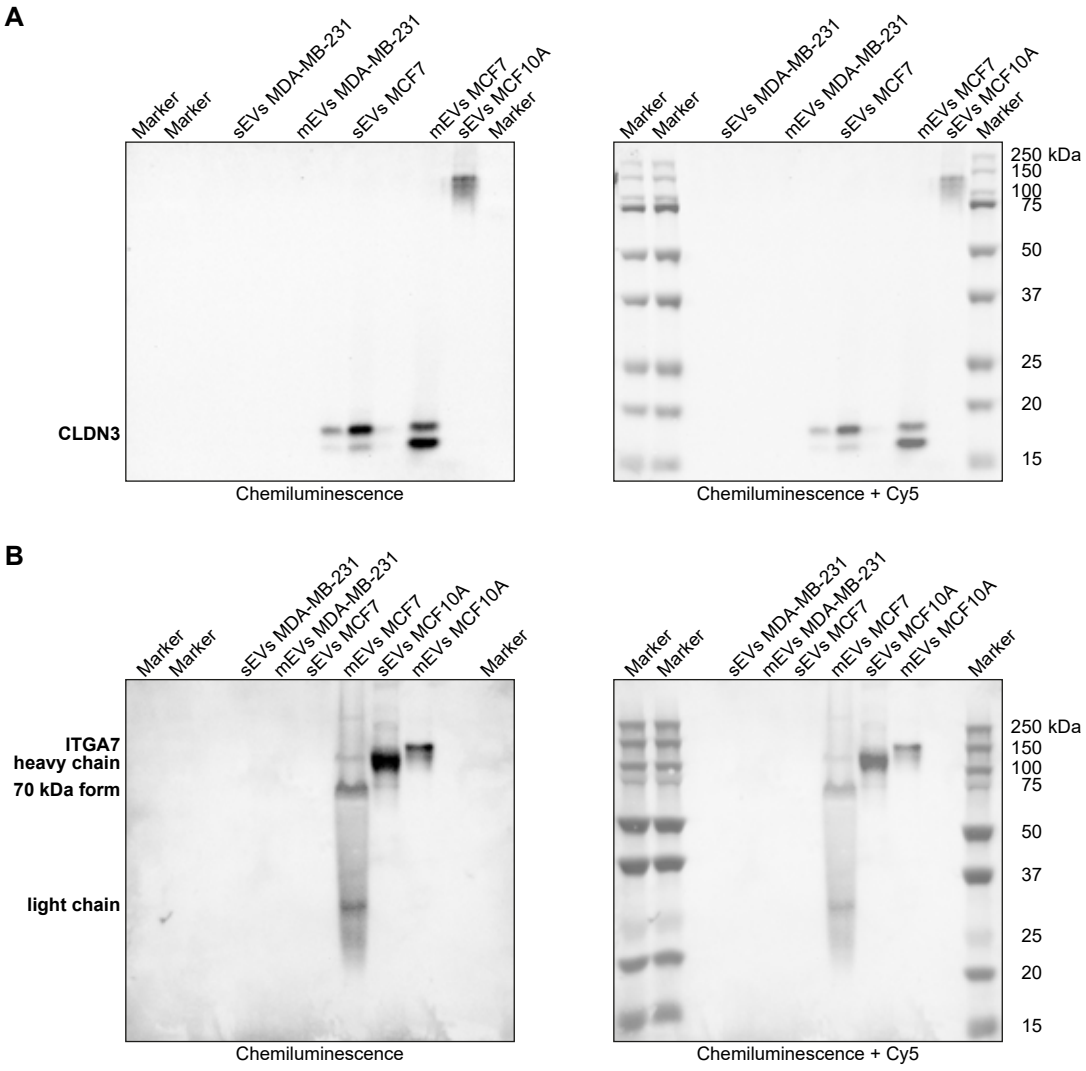

Fig. S7 Full WB images of CLDN3 (A) and ITGA7 (B) presented in Fig. 3.5.

Tables

Table S1 List of cell line-specific sEV proteins identified form global proteomics profiling. Single- and multi-pass transmembrane proteins are annotated based on UniProt subcellular-locations.

Protein accession	Gene name	Multi-pass membrane	Single-pass membrane
MDA-MB-231			
O14684	PTGES		
O14975	SLC27A2		
O15394	NCAM2	x	x
O15427	SLC16A3		
O15551	CLDN3		
O43653	PSCA		
O43768	ENSA		
O75063	FAM20B	x	x
O75366	AVIL		
O94907	DKK1		
O96005	CLPTM1		
P00736	C1R		
P00747	PLG		
P04004	VTN		
P07357	C8A		
P07996	THBS1		
P09758	TACSTD2	x	x
P10620	MGST1		
P10643	C7		
P13726	F3		
MDA-MB-231 & MCF7			
P15260	IFNGR1	x	x
P16949	STMN1		
P17302	GJA1		
MCF7			
P17931	LGALS3		
P19823	ITIH2		
P25929	NPY1R		
P27169	PON1		
P29034	S100A2		
P30530	AXL	x	x
P32926	DSG3	x	x
P36383	GJC1		
Q13683	ITGA7	x	x
Q15165	PON2		

Q16819	MEP1A	x	x
Q6UWI4	SHISA2	x	x
Q6UXG2	ELAPOR1	x	x
MCF10A			
Q86SF2	GALNT7	x	x
Q8N271	PROM2		
Q96KG7	MEGF10	x	x
Q99674	CGREF1		
Q9BY76	ANGPTL4		
Q9UBW7	ZMYM2		
Q9UMD9	COL17A1	x	x
Q9Y5Y6	ST14	x	x

Table S2 List of surface marker proteins of sEVs common to MDA-MB-231 and MCF7, and absent in MCF10A.

Protein accession	Gene name	Protein name
O60706	ABCC9	ATP binding cassette subfamily C member 9
Q96SE0	ABHD1	Abhydrolase domain containing 1
Q99798	ACO2	Aconitase 2
O14672	ADAM10	ADAM metallopeptidase domain 10
Q9UEY8	ADD3	Adducin 3
Q8N1W1	ARHGEF28	Rho guanine nucleotide exchange factor 28
Q6ICH7	ASPHD2	Aspartate beta-hydroxylase domain containing 2
P20020	ATP2B1	ATPase plasma membrane Ca2+ transporting 1
P56378	ATP5MJ	ATP synthase membrane subunit j
P35613	BSG	Basigin (Ok blood group)
P0C0L4	C4A	Complement C4A (Rodgers blood group)
P07358	C8B	Complement C8 beta chain
Q8IXQ3	C9orf40	Chromosome 9 open reading frame 40
P27352	CBLIF	Cobalamin binding intrinsic factor
P08962	CD63	CD63 molecule
Q14004	CDK13	Cyclin dependent kinase 13
Q8N8U2	CDYL2	Chromodomain Y like 2
Q7L2Z9	CENPQ	Centromere protein Q
O94986	CEP152	Centrosomal protein 152
Q9UL16	CFAP45	Cilia and flagella associated protein 45
P10909	CLU	Clusterin
Q9NXG0	CNTLN	Centlein
P07339	CTSD	Cathepsin D
Q13618	CUL3	Cullin 3
P98082	DAB2	DAB adaptor protein 2
Q13206	DDX10	DEAD-box helicase 10
Q09013	DMPK	DM1 protein kinase
Q9NX36	DNAJC28	DnaJ heat shock protein family (Hsp40) member C28
Q92784	DPF3	Double PHD fingers 3
Q008S8	ECT2L	Epithelial cell transforming 2 like
P13639	EEF2	Eukaryotic translation elongation factor 2
Q04637	EIF4G1	Eukaryotic translation initiation factor 4 gamma 1
Q9HB03	ELOVL3	ELOVL fatty acid elongase 3
P02794	FTH1	Ferritin heavy chain 1
O94808	GFPT2	Glutamine-fructose-6-phosphate transaminase 2
P09211	GSTP1	Glutathione S-transferase pi 1

Q99525	H4C7	H4 clustered histone 7
Q9Y4L1	HYOU1	Hypoxia up-regulated 1
Q9NRR6	INPP5E	Inositol polyphosphate-5-phosphatase E
P26006	ITGA3	Integrin subunit alpha 3
P19823	ITIH2	Inter-alpha-trypsin inhibitor heavy chain 2
Q6ZWJ8	KCP	Kielin cysteine rich BMP regulator
Q15058	KIF14	Kinesin family member 14
A6NM62	LRRC53	Leucine-rich repeat-containing protein 53
Q5JTD7	LRRC73	Leucine rich repeat containing 73
Q9UJA3	MCM8	Minichromosome maintenance 8 homologous recombination
O15146	MUSK	Muscle associated receptor tyrosine kinase
O43795	MYO1B	Myosin IB
Q8NAQ8	N/A	Putative uncharacterized protein FLJ34945
Q86UW6	N4BP2	NEDD4 binding protein 2
Q8WX92	NELFB	Negative elongation factor complex member B
Q6UWY5	OLFML1	Olfactomedin like 1
P06401	PGR	Progesterone receptor
A6NC86	PINLYP	Phospholipase A2 inhibitor and LY6/PLAUR domain
P48426	PIP4K2A	Phosphatidylinositol-5-phosphate 4-kinase type 2 alpha
P54277	PMS1	PMS1 homolog 1, mismatch repair system component
P36873	PPP1CC	Protein phosphatase 1 catalytic subunit gamma
O14818	PSMA7	Proteasome 20S subunit alpha 7
P35998	PSMC2	Proteasome 26S subunit, ATPase 2
Q96EY7	PTCD3	Pentatricopeptide repeat domain 3
Q9P2B2	PTGFRN	Prostaglandin F2 receptor inhibitor
Q2TAK8	PWWP3A	PWWP domain containing 3A, DNA repair factor
P15153	RAC2	Rac family small GTPase 2
Q96EQ8	RNF125	Ring finger protein 125
Q9ULK6	RNF150	Ring finger protein 150
P18124	RPL7	Ribosomal protein L7
Q59EK9	RUNDC3A	RUN domain containing 3A
Q5TEA6	SEL1L2	SEL1L2 adaptor subunit of ERAD E3 ligase
Q15758	SLC1A5	Solute carrier family 1 member 5
Q9UHI7	SLC23A1	Solute carrier family 23 member 1
Q9H2H9	SLC38A1	Solute carrier family 38 member 1
P13866	SLC5A1	Solute carrier family 5 member 1
Q9C093	SPEF2	Sperm flagellar 2
Q9BXG8	SPZ1	Spermatogenic leucine zipper 1
Q9H3Y6	SRMS	Src-related kinase lacking C-terminal regulatory tyrosine
Q9UQ35	SRRM2	Serine/arginine repetitive matrix 2
Q93045	STMN2	Stathmin 2

Q15431	SYCP1	Synaptonemal complex protein 1
Q8NBN3	TMEM87A	Transmembrane protein 87A
Q6ZUX3	TOGARAM2	TOG array regulator of axonemal microtubules 2
Q9BX84	TRPM6	Transient receptor potential cation channel subfamily
Q99816	TSG101	Tumor susceptibility 101
O43657	TSPAN6	Tetraspanin 6
Q96AE7	TTC17	Tetratricopeptide repeat domain 17
P02766	TTR	Transthyretin
Q70EL2	USP45	Ubiquitin specific peptidase 45
O75691	UTP20	UTP20 small subunit processome component
O60765	ZNF354A	Zinc finger protein 354A

Table S3 List of surface marker proteins of mEVs common to MDA-MB-231 and MCF7, and absent in MCF10A.

Protein accession	Gene name	Protein name
Q86UQ4	ABCA13	ATP binding cassette subfamily A member 13
P55008	AIF1	Allograft inflammatory factor 1
P54886	ALDH18A1	Aldehyde dehydrogenase 18 family member A1
Q8WW43	APH1B	Aph-1 homolog B, gamma-secretase subunit
Q9Y2Y0	ARL2BP	ADP ribosylation factor like GTPase 2 binding protein
Q6ICH7	ASPHD2	Aspartate beta-hydroxylase domain containing 2
P21926	CD9	CD9 molecule
Q9UL16	CFAP45	Cilia and flagella associated protein 45
Q96RY5	CRAMP1	Cramped chromatin regulator homolog 1
Q09013	DMPK	DM1 protein kinase
Q9Y295	DRG1	Developmentally regulated GTP binding protein 1
Q8TE68	EPS8L1	EPS8 like 1
O75426	FBXO24	F-box protein 24
Q15751	HERC1	HECT and RLD domain containing E3 ubiquitin protein
P51659	HSD17B4	Hydroxysteroid 17-beta dehydrogenase 4
Q8ND56	LSM14A	LSM14A mRNA processing body assembly factor
A8MW99	MEI4	Meiotic double-stranded break formation protein 4
Q9ULR5	PAIP2B	Poly(A) binding protein interacting protein 2B
Q15113	PCOLCE	Procollagen C-endopeptidase enhancer
O75340	PDCD6	Programmed cell death 6
P48426	PIP4K2A	Phosphatidylinositol-5-phosphate 4-kinase type 2 alpha
P14314	PRKCSH	Protein kinase C substrate 80K-H
Q92729	PTPRU	Protein tyrosine phosphatase receptor type U
Q96EQ8	RNF125	Ring finger protein 125
Q9BZW2	SLC13A1	Solute carrier family 13 member 1
Q15758	SLC1A5	Solute carrier family 1 member 5
O43173	ST8SIA3	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 3
Q9H2K2	TNKS2	Tankyrase 2
Q8NB66	UNC13C	Unc-13 homolog C
Q8N2C7	UNC80	Unc-80 homolog, NALCN channel complex subunit
O75691	UTP20	UTP20 small subunit processome component
P98182	ZNF200	Zinc finger protein 200

My Publications

1. P. Wiegand, L. Lupu, N. Hüttmann, J. Wack, S. Rawer, M. Przybylski, and K. Schmitz, "Epitope Identification and Affinity Determination of an Inhibiting Human Antibody to Interleukin IL8 (CXCL8) by SPR-Biosensor–Mass Spectrometry Combination" *Journal of the American Society for Mass Spectrometry* **31** (1), 109-116 (2020).
2. L. Lupu, P. Wiegand, N. Hüttmann, S. Rawer, W. Kleinekofort, I Shugureva, A. S. Kichkailo, F. N. Tomilin, A. Lazarev, M. V. Berezovski, and M. Przybylski, "Molecular Epitope Determination of Aptamer Complexes of the Multidomain Protein C-Met by Proteolytic Affinity-Mass Spectrometry," *ChemMedChem* **15** (4), 363-369 (2020).
3. Y. Risha, V. Susevski, N. Hüttmann, S. Poolsup, Z. Minic, and M. V. Berezovski, "Breast Cancer-Derived Microvesicles Are the Source of Functional Metabolic Enzymes as Potential Targets for Cancer Therapy," *Biomedicines* **9** (2), 107 (2021).
4. A. Allameh*, N. Hüttmann*, E. Charlebois*, A. Katsarou, W. Gu, K. Gkouvatsos, E. Pasini, Mamatha Bhat, Z. Minic, M. V. Berezovski, M. Guido, C. Fillebeen, and K. Pantopoulos, "Hemojuvelin deficiency promotes liver mito-

chondrial dysfunction and predisposes mice to hepatocellular carcinoma" *Communications Biology* (in review).

5. R. P. Parajulia, Y. Risha, N. Hüttmann, Z. Minic, S. M. Ghobadloo, L. Xin, M. V. Berezovski, and L. H. M. Chan, "Association between Urine Exosome Proteins and Body Burden of Arsenic among the participant from Yellowknife region in the Canadian Arctic" (manuscript in preparation).
6. A. Medić, N. Hüttmann, M. Lješević, Y. Risha, M. V. Berezovski, Z. Minic, and I. Karadžić, "A study of the flexibility of the carbon catabolic pathways of extremophilic P. aeruginosa san ai exposed to benzoate versus glucose as sole carbon sources by multi omics analytical platform" (submitted).
7. E. Di Stefano, N. Hüttmann, T. Oliviero, V. Fogliano, and C. C. Udenigwe, "Impact of gastrointestinal digestion and L. plantarum fermentation on peptides bioaccessibility and transport by human intestinal Caco-2 cells" (manuscript in preparation).
8. Z. Minic, N. Hüttmann, S. Poolsup, Y. Li, V. Susevski, E. Zaripov, and M. V. Berezovski, "Phosphoproteomic analysis of breast cancer-derived extracellular vesicles reveals disease-specific phosphorylated enzymes" (manuscript in preparation).