INVESTIGATING THE ROLE OF EXTRACELLULAR VESICLES AND GLYCOSYLATION IN CANCER METASTASIS

Jamie Cooper

Supervisors: Professor Susan Brooks & Dr Ryan Pink

This thesis is submitted in partial fulfilment of the requirements of the award of

Doctor of Philosophy at

Oxford Brookes University

April 2024

AUTHOR'S DECLARATIONS

I, Jamie Cooper, declare that this thesis titled 'Investigating the role of extracellular vesicles and glycosylation in cancer metastasis' submitted for assessment is an original work of my own. Any works of other authors used within this thesis, or the work presented in it are properly acknowledged. A list of the references employed is included.

ACKNOWLEDGMENTS

Throughout my PhD journey I've been truly thankful for every moment. Through the ups and downs, I've gained invaluable experiences, met amazing people, and most importantly, I've embarked on a profound journey of self-discovery that has shaped me into the man I am today.

Humorously, it all began with a chance encounter with Professor Dave Carter on my way to the bathroom at UKEV. Little did I know that our seemingly casual conversation about my interests in the research field would set the stage for this incredible journey, culminating in my acceptance of an offer to undertake my PhD at Oxford Brookes University. Although you decided to transition away from the academic world, I will never forget what you said to me, 'You will be left the safest pair of hands in regards to your supervision and you will learn so much', and you couldn't have been righter. For that, I am profoundly grateful.

My supervisors Professor Susan Brooks and Dr Ryan Pink, how do I begin to express my sincere gratitude for all your guidance throughout my journey. Susan, your ability to systematically breakdown experimental goals, your excellent organisation and mastery of written work and alongside your profound understanding of the theory behind my work has been truly invaluable. Ryan, your inspiring innovative ideas for my experimental approach, your technological insights and boundless enthusiasm and expertise have been a driving force behind my academic and scientific growth. I cannot thank you both enough.

A special thanks to the various members of the lab including past and present. In particular, my heartfelt appreciation to Joanna, Elise, Manu, Lizzie, Ellie, Soozana, Amber H, Amber R, and Grace, for your unwavering support and encouragement over the past years. In addition, the fantastic bioimaging team, Dr. Flavia Moreira-Leite, Edward Rea, and Dr. Alan Marron. A special mention to Billy, Cam, Harry, Liam and Paddy. Your presence has not only been a source of friendship but also the positivity you bring to my life, which has been instrumental in shaping both my academic endeavours and personal growth.

But, above all, my deepest thanks go to my incredible family for the unwavering support they've shown me at every step of the way. My brother Joe, your profound emotional understanding has always been there to lift me up and provide motivational support when needed. My mum and dad, there's no words to express my gratitude for everything you've done for me throughout this journey. The support you've shown me is the bedrock of any of my success and I will be forever grateful.

Finally, I would like to dedicate this doctoral work to my two nieces, Adelynn Rose Cooper and Paislee Grace Cooper. This work is a testament to that if you can achieve anything you want to in life. Everything is possible your dreams, your ideas and your vision. Never let anyone tell you, you can't, because you can.

ABSTRACT

Breast and colorectal cancers present significant challenges for the UK healthcare system, placing a substantial burden on the National Health Service (NHS). Early detection is crucial for improving survival rates and reducing healthcare strain, especially as metastasis, the primary cause of mortality in these cancers, remains poorly understood. Protein glycosylation, a post-translational modification of carbohydrate structures, is a significant contributor to cancer progression and metastasis. Small extracellular vesicles (sEVs) have emerged as key players in cancer development and metastasis, facilitating intercellular communication through bioactive cargo transfer. Despite mounting evidence, the importance of sEV glycosylation in cancer and metastasis is often overlooked. Liquid biopsies, including sEV biomarker analysis, offer promising non-invasive diagnostic approaches. Given the prevalence of glycoproteins among blood cancer biomarkers, exploring the glycosylation of sEVs as biomarker targets could revolutionise early detection strategies for breast and colorectal cancer.

Efforts initially focused on optimising single-vesicle flow cytometry for sEV characterisation to ensure maximum signal and sample integrity. Subsequent analysis of breast and colorectal epithelial cells and their derived sEVs revealed pronounced HPA lectin binding in the metastatic phenotypes. A lectin microarray identified increased lectin binding of LCA and TL in 'CD81-positive' sEVs derived from breast cancer-associated cell lines in comparison to normal cells. Validation through single-vesicle flow cytometry confirmed these observations, prompting investigations into the diagnostic application of these lectins in distinguishing plasma-enriched sEVs from breast cancer patients and 'healthy' individuals. However, lectins alone did not show diagnostic significance, in contrast to the diagnostic capability of the well-established breast cancer marker EpCAM.

Overall, these results underscore the importance of optimising single-vesicle flow cytometry for comprehensive characterisation of sEVs. They also highlight the HPA lectin binding patterns of breast and colorectal metastatic cell phenotypes, which mirror the observations of their derived sEVs but with nuanced specificity attributed to tetraspanin composition. Additionally, other glycosylated targets, as recognised by TL and LCA, show increased abundance in breast cancer cell-derived sEVs. Moreover, the diagnostic capability of these lectins and EpCAM in distinguishing plasma-enriched sEVs derived from breast cancer patients from those derived from 'healthy' individuals is demonstrated.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 5

CHAPTER 6

SUPPLEMENTARY FIGURES

LIST OF TABLES

ABBREVIATIONS

ABA – *Agaricus bisporus* agglutinin

ACA – *Amaranthus caudatus* agglutinin

AFP-L3 – alpha-fetoprotein-L3

ALIX – ALG-2 interacting protein X

ASA – *Allium sativum* agglutinin

Asn – Asparagine

AUC – Area under the curve

B3GNT6 - β1,3-N-acetylglucosaminyltransferase 6

BCA – Bicinchoninic acid assay

BM – Basement membrane

BPA - *Bauhinia purpurea* agglutinin

BSA – Bovine serum albumin

CA15-3 - Cancer antigen 15-3

CEA - Carcinoembryonic antigen

CFDA-SE - Carboxyfluorescein diacetate succinimidyl ester

CfDNA – Cell-free DNA

CFSE- Carboxyfluorescein succinimidyl ester

C1GalT1 - β1,3-galactosyltransferase

CIN – Chromosomal instability

CISH - Chromogenic *in situ* hybridization

COSMC - Core 1 β3-galactosyltransferase Specific Molecular Chaperone

CTCs – Circulating tumour cells

CtDNA – Circulating tumour DNA

DCIS - Ductal carcinoma *in situ*

dH$_2$O – Distilled water

DMEM - Dulbecco's modified Eagle medium

DMSO - Dimethyl sulfoxide

DTT – Dithiothreitol

ECACC - European Collection of Authenticated Cell Cultures

ECL - Enhanced chemiluminescence

ECM – Extracellular matrix

EDTA - Ethylenediaminetetraacetic acid

EGFR – Epidermal growth factor receptor

ELISA – Enzyme linked immunosorbent assay

EMT – Epithelial-mesenchymal transition

EpCAM – Epithelial cell adhesion molecule

ER – Endoplasmic reticulum

ESCRT - Endosomal sorting complexes required for transport

EVs – Extracellular vesicles

EWI-2 – Ewing's Sarcoma-Associated Transcript 2

FDA - Food and Drug Administration

FBS – Fetal bovine serum

FPS – Frames per second

FSC-A - Forward scatter area

FSC-H – Forward scatter height

Fuc – Fucose

FUT8 - α1,6-fucosyltransferase

GALNTs - N-acetylgalactosaminyltransferases

Gal – Galactose

GalNAc – N-acetylgalactosamine

gFOBt - Guaiac faecal occult blood test

GlcNAc – N-acetylglucosamine

Glu – Glucose

GNA - *Galanthus nivalis* agglutinin

GnT-III - N-acetylglucosaminyltransferase III

GnT-V - N-acetylglucosaminyltransferase V

HAA – *Helix aspersa* agglutinin

HEWL – Hen egg white lysozyme

HD-PTP - His domain–containing tyrosine phosphatase

HPA - *Helix pomatia* agglutinin

ILVs – Intraluminal vesicles

IMS – Industrial methylated spirits

JAK/STAT - Janus kinase/signal transducer and activator of transcription

LBPA - Lysobisphosphatidic acid

LCA - *Lens culinaris* agglutinin

LDL – Ultralow density lipoprotein marker

Man – Mannose

MCED – Multi-cancer early detection

MESF - Molecules of equivalent soluble fluorochrome

MFI – Median fluorescence intensity

MHRA - Medicines and Healthcare products Regulatory Agency

MVs – Microvesicles

MVBs – Multivesicular bodies

NANA - N-acetylneuraminic acid

NGS – Next-generation sequencing

NHS – National Health Service

nSMase 2 - Neutral sphingomyelinase 2

NTA – Nanoparticle tracking analysis

PAR1 - Protease-activated receptor 1

PBS – Phosphate buffered saline

PCR – Polymerase chain reaction

PE-A – Phycoerythrin area

PFA – Paraformaldehyde

PHA-L – *Phaseolus vulgaris* leucoagglutinin

PI – Propidium iodide

PMLE - Amyloidogenic pigment cell-specific type I integral membrane protein

PI3K-Akt - Phosphatidylinositol 3-kinase/protein kinase B

PtdIns3P - Phosphatidylinositol-3-phosphate

PVDF - Polyvinylidene difluoride

RAB – Ras-associated binding

ROC – Receiver operating characteristic

RIPA - Radio-immunoprecipitation assay

RPA - R*obinia pseudoacacia* agglutinin

RT – Room temperature

RTKs – Receptor tyrosine kinases

SEC – Size exclusion chromatography

Ser – Serine

SDS - Sodium dodecyl sulphate

SHA - *Salivia horminum* agglutinin

Sia – Sialic acid

SMA – Styrene-maleic acid

SNA - *Sambucus nigra* agglutinin

SNAREs - N-ethylmaleimide-sensitive factor attachment protein receptors

SFM - Serum-free medium

sEVs – Small extracellular vesicles

SSA - *Sambucus sieboldiana* agglutinin

SSC-A – Side scatter area

Stdev – Standard deviation

TBS – Tris-buffered saline

TBST - Tris-buffered saline with Tween

TDI – Time-delay integration

TEM – Transmission electron microscopy

TEMs – Tetraspanin-enriched microdomains

TGF-β - Transforming growth factor-β

Thr – Threonine

TL – *Tulipa* lectin

UDA - *Urtica dioica* agglutinin

UDP - Uridine diphosphate

UEA – *Ulex europaeus* agglutinin

WGA – Wheat germ agglutinin

CHAPTER 1

INTRODUCTION

1.1. CANCER

1.1.1. HALLMARKS

Cancer is characterised by the accumulation of genetic alterations that trigger abnormal and uncontrolled cellular growth, leading to tumour formation (reviewed by Sarkar et al., 2013). The development of cancer is a complex and multifaceted process that can arise from the activation of proto-oncogenes and/or inactivation of tumour suppressor genes (Pishas et al., 2015). These functional alterations are attributed to the intricacy of cancer which is often described by its hallmarks; a set of characteristics proposed by Hanahan & Weinberg (2000) that is established to dictate malignant growth. These ten hallmarks comprise of eight acquired and two enabling capabilities (Figure 1.1). The acquired capabilities include sustaining proliferative signalling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, evading the immune system, deregulating metabolism, and activating invasion and metastasis (Hanahan & Weinberg, 2000). The two enabling capabilities are tumour-promoting inflammation and genome instability.
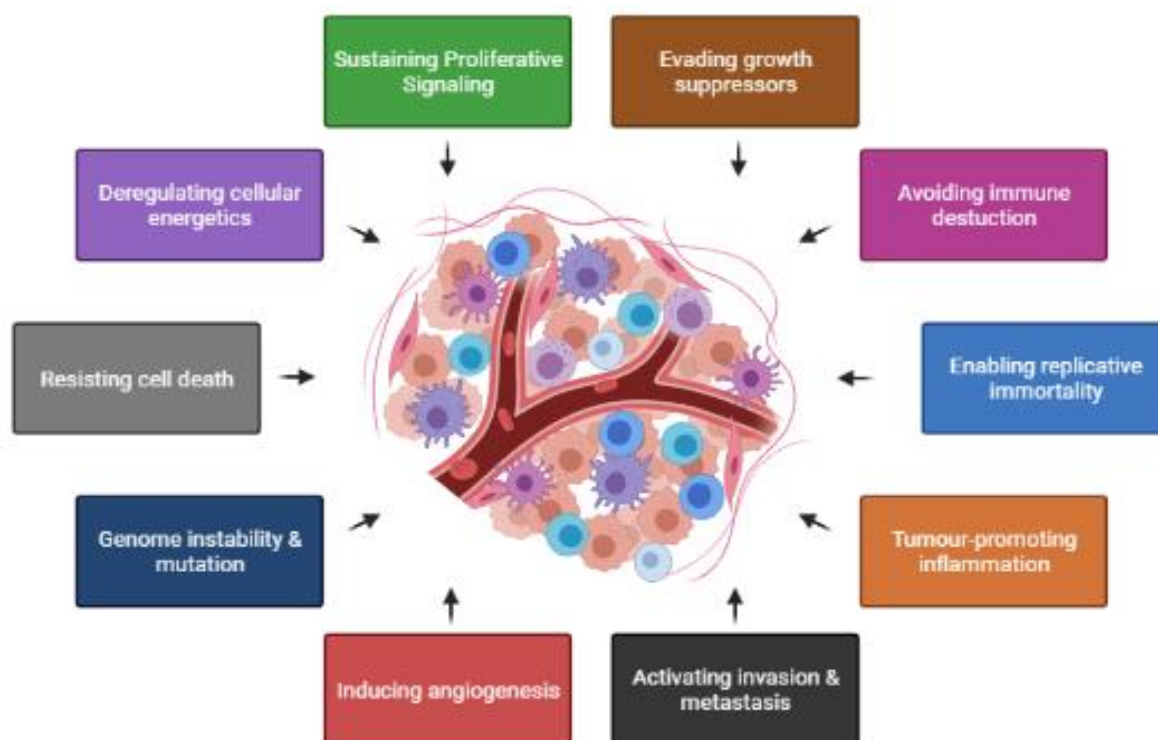
Figure 1.1. *The hallmarks of cancer: The ten hallmarks of cancer comprising of eight acquired capabilities: sustaining proliferative signalling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, evading the immune system, deregulating metabolism, and activating invasion and metastasis, as well as the two enabling capabilities, tumour-promoting inflammation and genome instability. Figure adapted from* (Bergstrom et al., 2015).

1.1.2. BREAST CANCER

1.1.2.1. BREAST CANCER INCIDENCE AND MORTALITY RATES

Breast cancer is the most frequently diagnosed cancer and the second leading cause of malignancy-related fatalities among women worldwide (DeSantis et al., 2019). In 2020, approximately 2.3 million women were diagnosed with breast cancer, resulting in 685,000 deaths globally (Sung et al., 2021). In the UK, breast cancer ranks as the second most common cause of cancer-related deaths among women, accounting for around 11,400 deaths per year (Cancer Research U.K, 2015). Despite advancements in therapeutic interventions, a significant proportion of patients still face an incurable prognosis, with a predicted ten-year cancer survival rate of 75.9% for females in the UK (Cancer Research U.K, 2015). Metastasis is the primary factor contributing to mortality in breast cancer, as discussed further in section 1.2.

1.1.2.2. BREAST ANATOMY AND CARCINOGENESIS

The breast is a complex organ comprising lobules, milk ducts, and supportive tissues (Figure 1.2) (reviewed by Biswas et al., 2022). Lobules are small, rounded glandular structures containing alveoli responsible for secreting milk during lactation into branching milk ducts and eventually into the nipple (Shah et al., 2023). The ductal network consists of two epithelial cell types: inner luminal epithelial cells and outer myoepithelial cells, surrounded by collagenous stroma and basement membrane (BM) (Figure 1.2-B) (Gudjonsson et al., 2005). Most breast cancers are carcinomas that originate from epithelial cells, such as ductal carcinoma *in situ* (DCIS) and lobular carcinoma *in situ* (LCIS) of the breast (Makki, 2015). The progression of both DCIS and lobular LCIS to invasive carcinoma, as proposed by Cowell et al. (2013), involves acquiring genetic mutations, changes in cellular morphology, and breakdown of the BM separating carcinoma cells from surrounding tissue. These changes can lead to invasion of carcinoma cells into surrounding tissue, where they proliferate and metastasise.
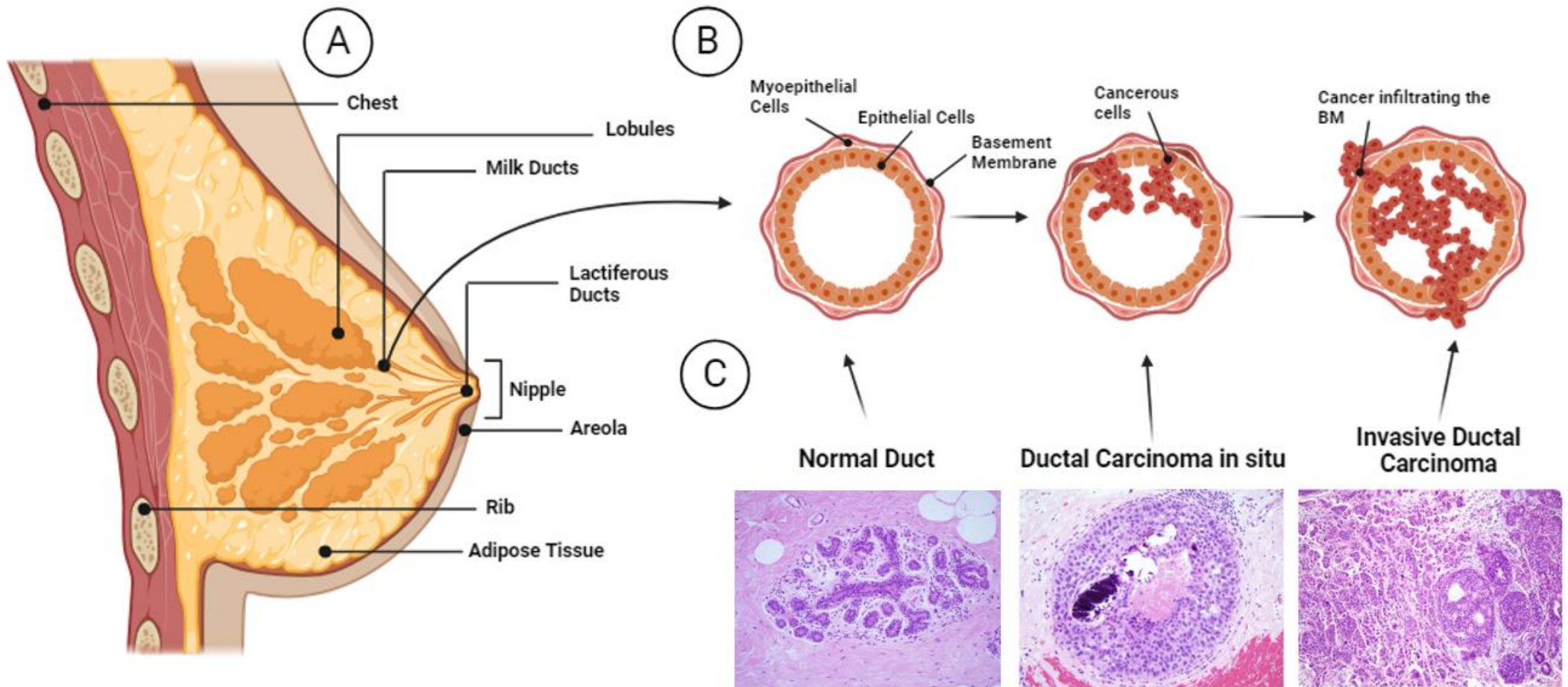
Figure 1.2. Anatomy and cellular progression of DCIS to invasive ductal carcinoma of the breast, with histological images. (A) *Overview of the anatomy of the breast, highlighting lobules, milk ducts, and other supportive tissues.* **(B)** *Detailed cellular components of the milk ducts, including the inner luminal epithelial cells, outer myoepithelial cells, and BM, with transformation of epithelial cells to DCIS and progression to invasive ductal carcinoma.* **(C)** *Histological images of normal duct, DCIS, and invasive ductal carcinoma. Histological images taken from* Tomlinson-Hansen et al. (2024).

1.1.3. COLORECTAL CANCER

1.1.3.1. COLORECTAL CANCER INCIDENCE AND MORTALITY RATES

Colorectal cancer, involving the colon and/or rectum, is the fourth most prevalent cancer in the UK, with over 42,900 diagnosed cases annually (Cancer Research U.K, 2015). Additionally, it accounts for approximately 16,800 cancer-related deaths annually in the UK, constituting 46 deaths per day. Globally, in 2020, colorectal cancer was ranked as the third most common cancer, with over 1.93 million new cases and 935,173 reported deaths (Xi & Xu, 2021). Despite progress in surgical and chemotherapeutic interventions for colorectal cancer, a significant proportion of patients still have an incurable prognosis, with a 52.9% predicted ten-year survival in the UK (Cancer Research U.K, 2015). The primary factor leading to mortality in colorectal cancer is metastasis, further discussed in section 1.2.

1.1.3.2. COLORECTAL ANATOMY AND CARCINOGENESIS

The large intestine, also known as the colon, is comprised of several major components, including the caecum, ascending colon, transverse colon, descending colon, and sigmoid colon. Additionally, the rectum and anus, which are separated from the colon, constitute the lower part of the gastrointestinal tract (Figure 1.3). The colon plays an important role in the digestive system by absorbing water, vitamins, and electrolytes, and compacting faeces for elimination (reviewed by Azzouz & Sharma, 2023). The colonic wall is composed of different layers including the mucosa, submucosa, muscular layer, and serosa, all of which support colonic motility. Colorectal cancer arises from the transformation of normal epithelial cells, specifically the inner epithelial lining, into cancerous neoplastic cells. The inner epithelial lining is composed of various cell types, including enterocytes, goblet cells, Paneth cells, and neuroendocrine cells, each serving distinct functions in the intestinal physiology, as described by Kong et al. (2018). The sequence of colorectal adenocarcinoma neoplasia, as proposed by O'Brien & Gibbons, (1996), describes colorectal cancer typically developing through a series of histopathological changes, progressing from benign polyps to invasive and metastatic cancer.
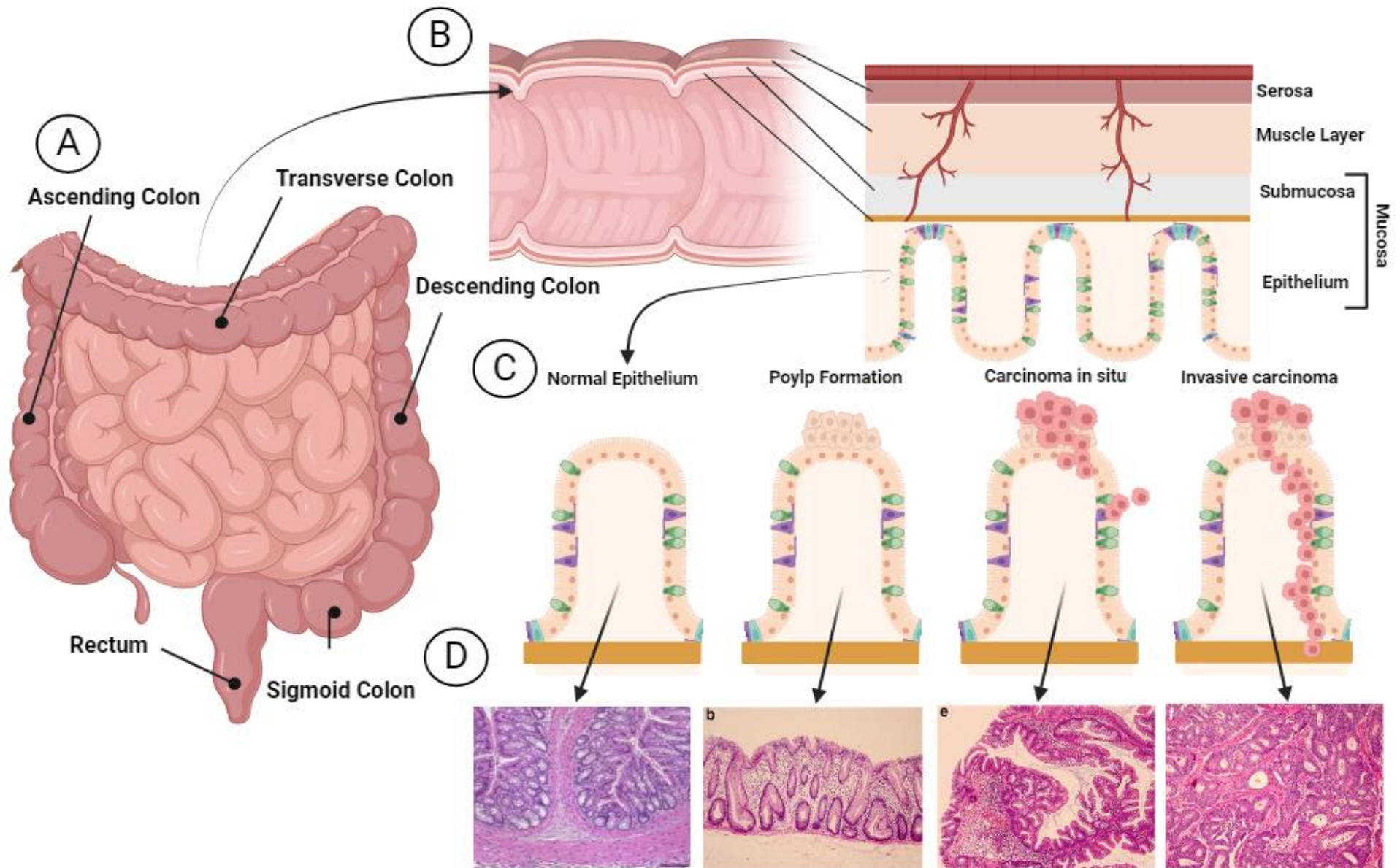
Figure 1.3. Anatomy and cellular progression of normal colorectal epithelium, to invasive carcinoma via poly formation, with histological images (A) *The anatomy of the colon* **(B)** *The layers of the colon wall including the mucosa, submucosa, muscular layer, and serosa, which supports colon motility through digestion* **(C)** *The stages of colorectal cancer transitioning from normal epithelial cells forming a polyp on the inner lining leading to the formation of carcinoma in situ and ultimately acquiring invasive capabilities.***(D)** *Histological images taken from* Erben et al. (2014) and Yamagishi et al. (2016).

1.2. CANCER METASTASIS

1.2.1. OVERVIEW

Cancer metastasis is a hallmark of tumour malignancy and accounts for the greatest number of cancer-related deaths; however, it remains poorly understood (Fares et al., 2020). Cancer invasion is a process whereby cells breach tissue barriers and subsequently infiltrate networks such as the blood, leading to haematogenous tumour dissemination (reviewed by McSherry et al., 2007). The metastasis cascade is characterised by several phases, including tumour angiogenesis, primary disaggregation of tumour cells from the primary tumour, invasion and migration through the BM, intravasation of tumour cells into the blood vasculature or lymphatic system, adhesion of circulatory tumour cells to endothelial cells, and extravasation of invasive tumour cells, as illustrated in Figure 1.4.
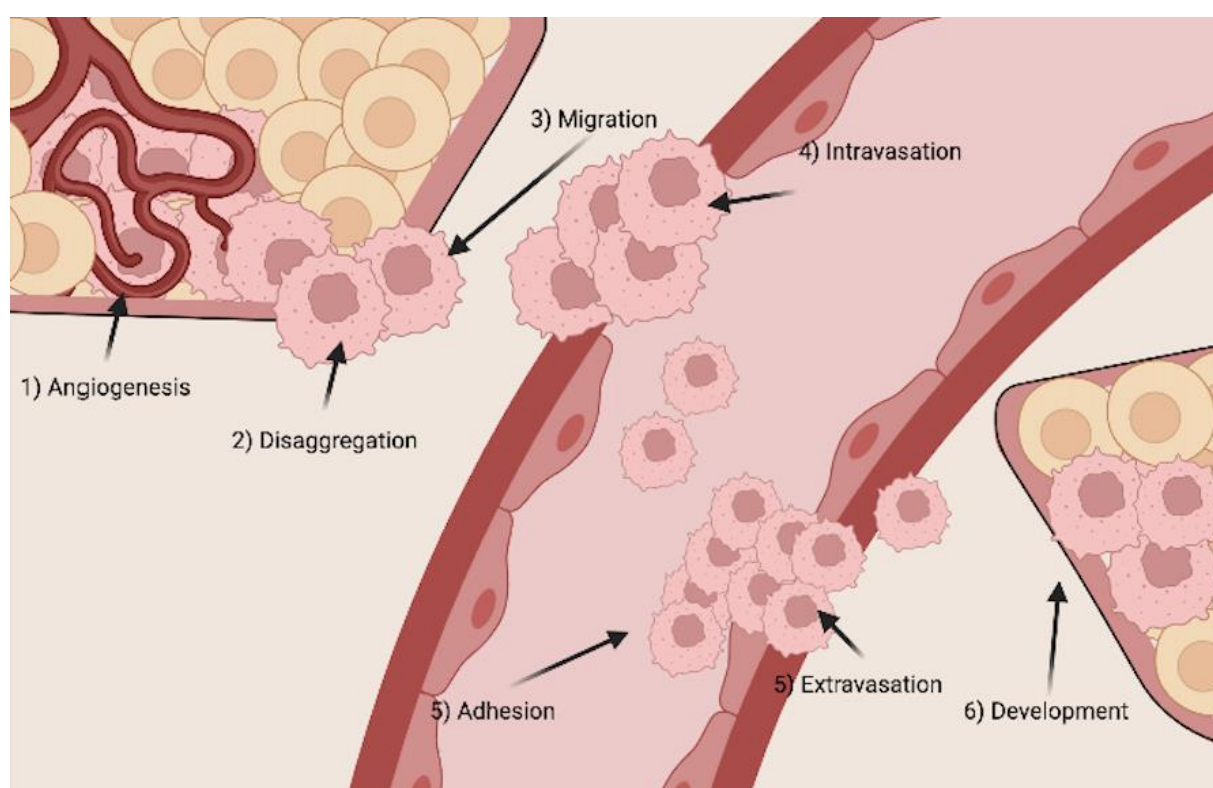
Figure 1.4. **The steps of metastasis.** *(1) Tumour angiogenesis (2) Disaggregation of tumour cells from the primary tumour (3) Migration through the BM (4) Intravasation of the tumour cells into the blood vessels (5) Adhesion of the circulating tumour cells to the endothelial cell lining (6) Extravasation of tumour cells at a target organ (7) Development of secondary tumour foci at the target organ site.*

1.3. GLYCOSYLATION OF PROTEINS

1.3.1. OVERVIEW

Glycosylation is characterised by the addition of a carbohydrate structure to a protein. This complex post-translational modification plays a crucial role in many important physiological processes such as cell differentiation and intracellular and intercellular signalling (reviewed by Ho et al., 2016). Glycosylation, a

fundamental process crucial for the proper functioning of numerous proteins, can modify protein characteristics, affecting their stability and solubility, both critical for functionality (reviewed by Jayaprakash & Surolia, 2017). The human glycome is composed of seven monosaccharides: glucose (Glu), mannose (Man), galactose (Gal), fucose (Fuc), sialic acids (Sia), N-acetylglucosamine (GlcNAc), and N-acetylgalactosamine (GalNAc). The most common glycan residues are hexoses, which are monosaccharides possessing six carbon atoms. The bond between a monosaccharide and another residue is called a glycosidic linkage. A glycosidic linkage is formed via activation of the glycosyl donor, which generates a reactive electrophilic form that couples with the hydroxyl group glycosylase acceptor, resulting in a condensation reaction (Seeberger et al., 2009). The formation of glycoproteins may be either O- or N-linked, depending on whether they are linked to the protein core via an oxygen atom of a serine (Ser) or threonine (Thr) (O-linked) or a nitrogen atom of asparagine (Asn) (N-linked) via a glycosidic bond, as illustrated in Figure 1.5 (Rachmilewitz, 2010). The synthesis of specific glycans involves a complex interplay of factors, including the availability, presence, and activity of glycosyltransferases, which are regulated by various mechanisms, such as transcription factors, miRNA, and DNA methylation (reviewed by Dall'Olio & Trinchera, 2017; Kasper et al., 2014; Neelamegham & Mahal, 2016; Rini et al., 2022). Additionally, the localisation of these enzymes within the endoplasmic reticulum (ER) or Golgi apparatus, the availability of nucleotide donors, and local environmental factors including pH further influence the glycan synthesis process (Colley et al., 2022; Rudd et al., 2022).
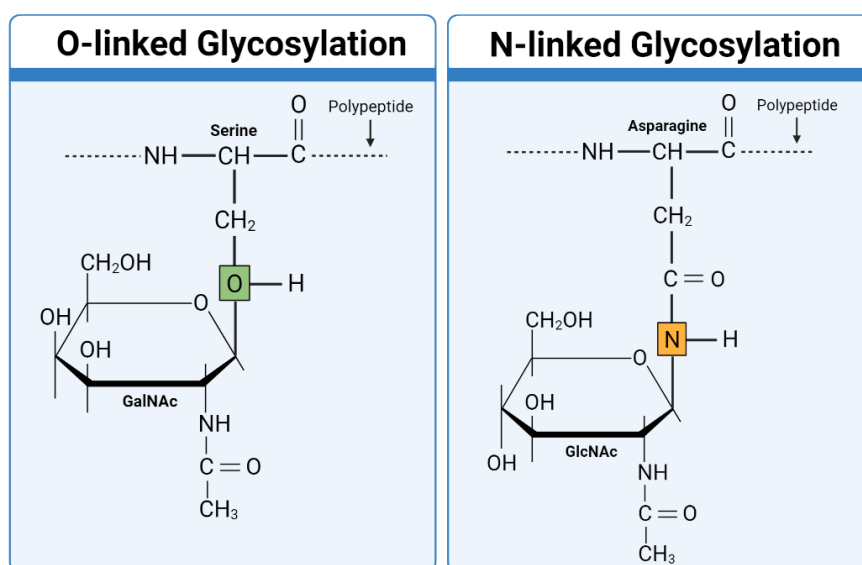
Figure 1.5. **Examples of O-linked and N-linked glycosylation.** *O-linked glycosylation showing the oligosaccharide linked via GalNAc to the oxygen of a Ser amino acid on the polypeptide chain. N-linked glycosylation occurs by the linkage between the oligosaccharide and the GlcNAc molecule of the amide of an Asn amino acid on the polypeptide chain.*

1.3.2. N-LINKED GLYCOSYLATION

N-linked glycosylation is a ubiquitous protein modification that is found in all domains of life. Highly diverse N-linked glycosylation structures are found in a large number of glycosylated proteins (reviewed by Schwarz & Aebi, 2011). The post-translational modification is key for a variety of protein functions, including membrane stability and folding (reviewed by Esmail & Manolson, 2021). The different structures of N-linked glycans often

contain a common core but differ in their extension on the terminal side. Taylor & Drickamer (2011) provided an in-depth review of the biosynthesis process of N-linked glycosylation, offering a comprehensive description of its various stages, including the formation of a dolichol-linked precursor oligosaccharide, *en bloc* transfer of the oligosaccharide to the polypeptide, and processing of the oligosaccharide, as illustrated in Figure 1.6.

1.3.2.1. FORMATION OF A DOLICHOL-LINKED PRECURSOR

The process of biosynthesising N-linked oligosaccharides involves the use of a lipid carrier, dolichol phosphate, to transport precursor oligosaccharides. The monosaccharide donors uridine diphosphate (UDP)-GlcNAc and GDP-Man are first synthesised in the cytoplasm and then transferred step-by-step to the lipid carrier, forming a $Man_5GlcNAc_2$ branched oligosaccharide through the actions of glycosyltransferases. Translocation of the newly formed oligosaccharide across the ER is proposed to be facilitated by a specialised enzyme known as flippase (Rush, 2016). Flippase helps flip the lipid-linked oligosaccharide from the cytosol to the luminal side of the ER membrane, allowing it to be transferred to specific Asn residues on nascent proteins that are synthesised by ribosomes bound to the ER. Although the fundamental role of flippase in N-linked glycosylation is well established, its precise function and mechanism are still not fully understood.

1.3.2.2. *EN BLOC* TRANSFER OF THE OLIGOSACCHARIDE

The precursor oligosaccharide attached to the dolichol lipid carrier anchored in the ER membrane is transferred *en bloc* to nascent polypeptides synthesised by ribosomes on the ER membrane. With the help of an enzyme called oligosaccharyltransferase, the precursor oligosaccharide is transferred from the dolichol lipid to an Asn residue in the growing polypeptide chain. However, for attachment, three key conditions must be met: 1) the Asn must be in a specific sequence context (Asn-X-Ser/Thr sequence), where 'X' represents any amino acid except proline, within the primary structure of the protein; 2) the position of the Asn must be appropriately located in a three-dimensional structure; and 3) Asn must be found on the luminal side of the ER to initiate N-linked glycosylation.

1.3.2.3. PROCESSING OF THE OLIGOSACCHARIDE

Next, the newly formed polypeptide and N-linked sugar residues are trimmed by a series of exoglycosidases at the non-reducing termini of the glycan. First, glucosidase I removes the terminal α1-3-linked Glu, followed by glucosidase II, which removes the two inner α1-3-linked Glu molecules, indicating that the glycoprotein is available for transit from the ER for further processing. The common core structure is defined by a trimannosyl core and can be extended to produce heterogeneous branched glycans in the cis-Golgi, catalysed by a multitude of glycosyltransferases, and further elongated in the medial Golgi (Stanley, 2011). Furthermore, complex glycans with highly branched structures are typically extended with various monosaccharides, in the trans-Golgi secretory pathway, with Sia addition often terminating chain extension.
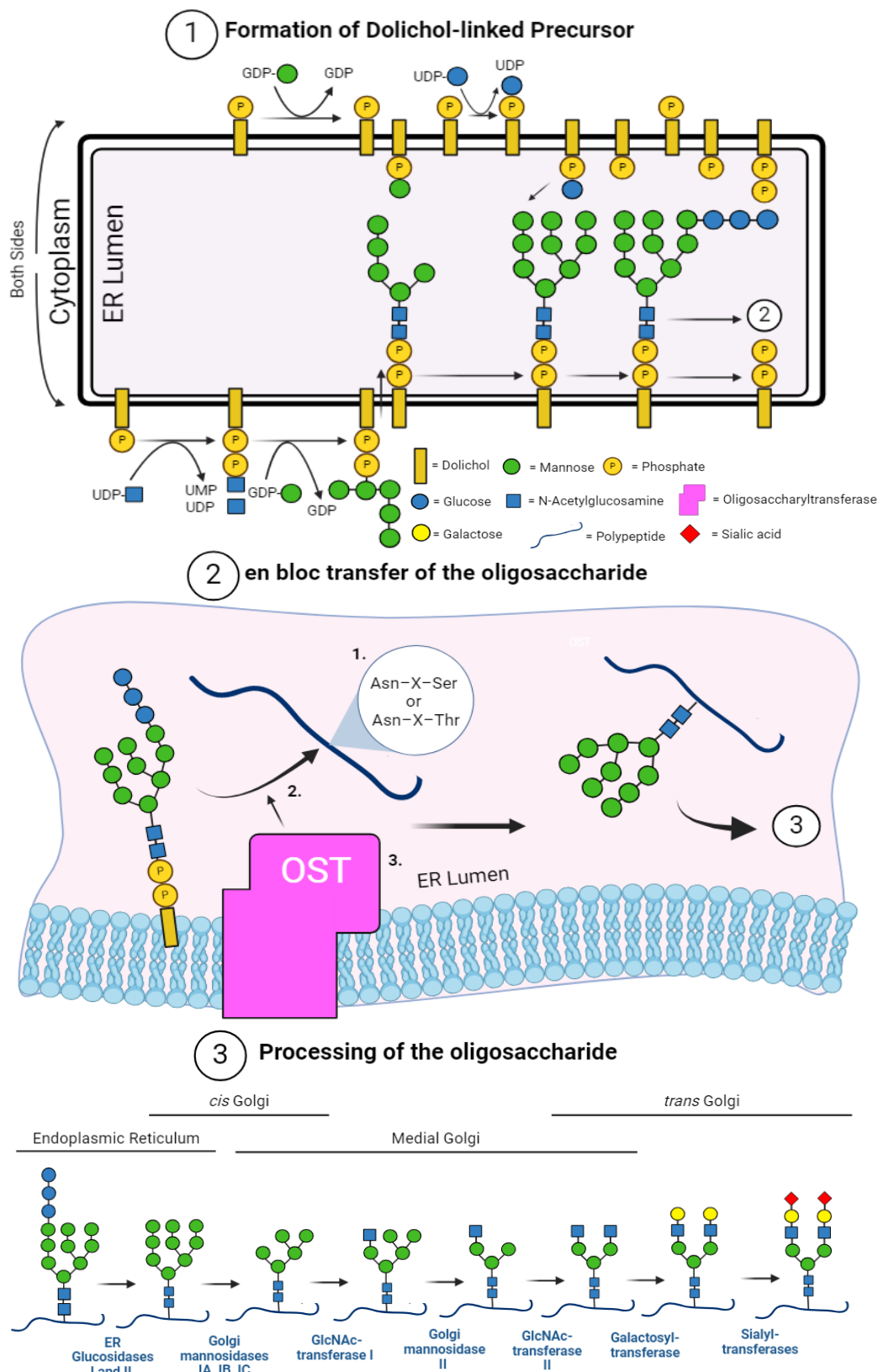
Figure 1.6. **Biosynthetic pathways of N-linked glycosylation:** *1) The biosynthesis of N-linked glycans begins with the stepwise transfer of monosaccharides from UDP-GlcNAc and GDP-Man, which are synthesised in the cytoplasm, to the lipid carrier dolichyl phosphate. The dolichol-anchored heptasaccharide is then flipped into the lumen of the endoplasmic reticulum, where additional monosaccharide transferases extend the oligosaccharide. 2) Key conditions which must be met for oligosaccharide to be transferred to Asn residues. 3) Processing of the oligosaccharide which involves sequential processing steps including glycosidase and glycosyltransferase action. Adapted from* (Taylor & Drickamer, 2011).

1.3.2.4. THE SIGNIFICANCE OF N-GLYCOSYLATION IN CANCER PROGRESSION

Abnormal N-linked glycosylation is recognised as a significant contributor to cancer progression; however, the precise molecular mechanisms involved remains elusive. This aberrant N-linked glycosylation can trigger disruptions in cellular signalling pathways, ultimately supporting carcinogenesis (reviewed by Scott & Drake, 2019). It is widely accepted that alterations in N-linked glycosylation patterns, coupled with gene mutations and genomic instability, collectively initiate oncogenic signalling pathways. These pathways include Wnt/β-catenin, Hippo, phosphatidylinositol 3-kinase/protein kinase B (PI3K/Akt), Janus kinase/signal transducer and activator of transcription (JAK/STAT), transforming growth factor-β (TGF-β) or SMAD, and Notch signalling (reviewed by Lin & Lubman, 2023). Moreover, aberrant glycosylation of cell surface proteins, such as transmembrane proteins and growth factor receptors, contribute to tumour cell growth, invasion, and metastasis through the activation of these signalling cascades (reviewed by Čaval et al., 2023).

Epidermal growth factor receptor (EGFR), a member of the receptor tyrosine kinases (RTKs), plays a pivotal role in cancer cell processes, including growth, proliferation, migration, and metabolism (reviewed by Hsu & Hung, 2016). N-glycosylation abnormalities in EGFR, resulting from changes in glycosylation patterns, affect the conformation of the extracellular domain of receptors, leading to aberrant activation and signalling transmission (Kaszuba et al., 2015). Similarly, N-glycosylation influences TGF-β receptor interactions and controls cell surface transport, affecting the downstream signalling cascade, specifically by promoting epithelial-mesenchymal transition (EMT) (Kim et al., 2012). Cadherins, which are essential calcium-dependent cell adhesion proteins, are pivotal in carcinogenesis, notably E-cadherin (reviewed by Rajwar et al., 2015). Decreased E-cadherin expression is linked to enhanced invasiveness in various cancer types, and is often associated with EMT (Na et al., 2020). The biological functions of E-cadherin are modulated by N-acetylglucosaminyltransferase III (GnT-III) and V (GnT-V), which affect N-linked glycosylation patterns (Pinho et al., 2009). Bisecting N-glycosylation, predominantly attributed to GnT-III, stabilises E-cadherin on the cell membrane, promotes functional adherens junctions, and increases cell aggregation (reviewed by Lin & Lubman, 2023). Integrins, pivotal adhesion receptors present across cell types, mediate cell attachment to the extracellular matrix (ECM) and influence migration and proliferation (reviewed by Huttenlocher & Horwitz, 2011). These receptors, composed of α- and β-subunits, form diverse combinations that affect cellular responses (reviewed by Mezu-Ndubuisi & Maheshwari, 2021). N-glycosylation of integrins affects heterodimeric formation and biological functions, as exemplified by the ability of a5b1 integrin to modulate cellular responses and contribute to tumour phenotypes (reviewed by Lin & Lubman, 2023).

1.3.3. O-LINKED GLYCOSYLATION

In eukaryotes, O-linked glycosylation occurs within the Golgi apparatus. O-linked glycosylation is characterised by the attachment of a sugar molecule to the oxygen atom of a Ser or Thr residue of a protein. O-linked glycosylation has a wide range of biological functions including in tissue development, homeostasis, signalling, cell identity, and immune regulation (reviewed by Wandall et al., 2021). The most common type of O-linked glycosylation begins with the attachment of a GalNAc monosaccharide to the protein, and is also known as O-

GalNAc or 'mucin-type'. Mucin type O-glycosylation is one of the most diverse types of glycosylation, playing essential roles in normal development, growth, and differentiation of cells and tissues (reviewed by Cummings, 2009; Schjoldager & Clausen, 2012).

1.3.3.1. O-GalNAc (MUCIN-TYPE) GLYCOSYLATION

O-GalNAc glycosylation involves a stepwise process that is initiated by the attachment of an α-GalNAc residue to the hydroxyl group of Ser or Thr residues by a family of 20 different N-acetylgalactosaminyltransferases (GALNTs). This produces GalNAcα-Ser/Thr, known as the Tn or Thompsen nouvelle antigen (Figure 1.7) (Raman et al., 2012). In normal adult cells, this structure is then extended in a stepwise manner to produce a variety of higher-order glycan structures. The specific glycosyltransferase responsible for catalysing the extension of the structure determines the ultimate glycan structure that is produced. The most common elaboration of the Tn antigen is the addition of a β1,3 linked Gal monosaccharide, catalysed by β1,3-galactosyltransferase (C1GalT1), to form of the core 1 or T (Thomsen-Friedenreich) antigen, Galβ1-3GalNAcα-Ser/Thr (Sun et al., 2021). The activity of C1GalT1 is highly dependent on an ER chaperone known as Core 1 β3-galactosyltransferase Specific Molecular Chaperone (COSMC). The addition of GlcNAc in a β1,3-linkage to Tn antigen is catalysed by β1,3-N-acetylglucosaminyltransferase 6 (B3GNT6) to form the core 3 structure (reviewed by Tran & Ten Hagen, 2013). The core 1 and core 3 structures can undergo further modifications, leading to their transformation into core 2 and core 4 structures, respectively, with the addition of GlcNAc via different isoforms of the β1,6-N-acetylglucosaminyltransferase family. There are other less common cores, 5 to 8. The core structures can also be further modified by sialyltransferases and fucosyltransferases which catalyse the addition of Sia and Fuc, respectively. The sialyl Tn antigen comprises of Sia α-2,6 linked to GalNAc (Neu5Acα2–6GalNAcα-Ser/Thr). Its synthesis involves the enzyme sialyltransferase ST6GalNAc1 (Munkley, 2016). Moreover, core 2 O-GalNAc glycans can undergo further elongation and sialylation to produce structures, such as Lewis and sialyl Lewis antigens. The activity of mucin-type associated transferases is highly regulated by a variety of factors, including the availability of donor substrates, localisation of enzymes within the Golgi apparatus, and interaction with other proteins.
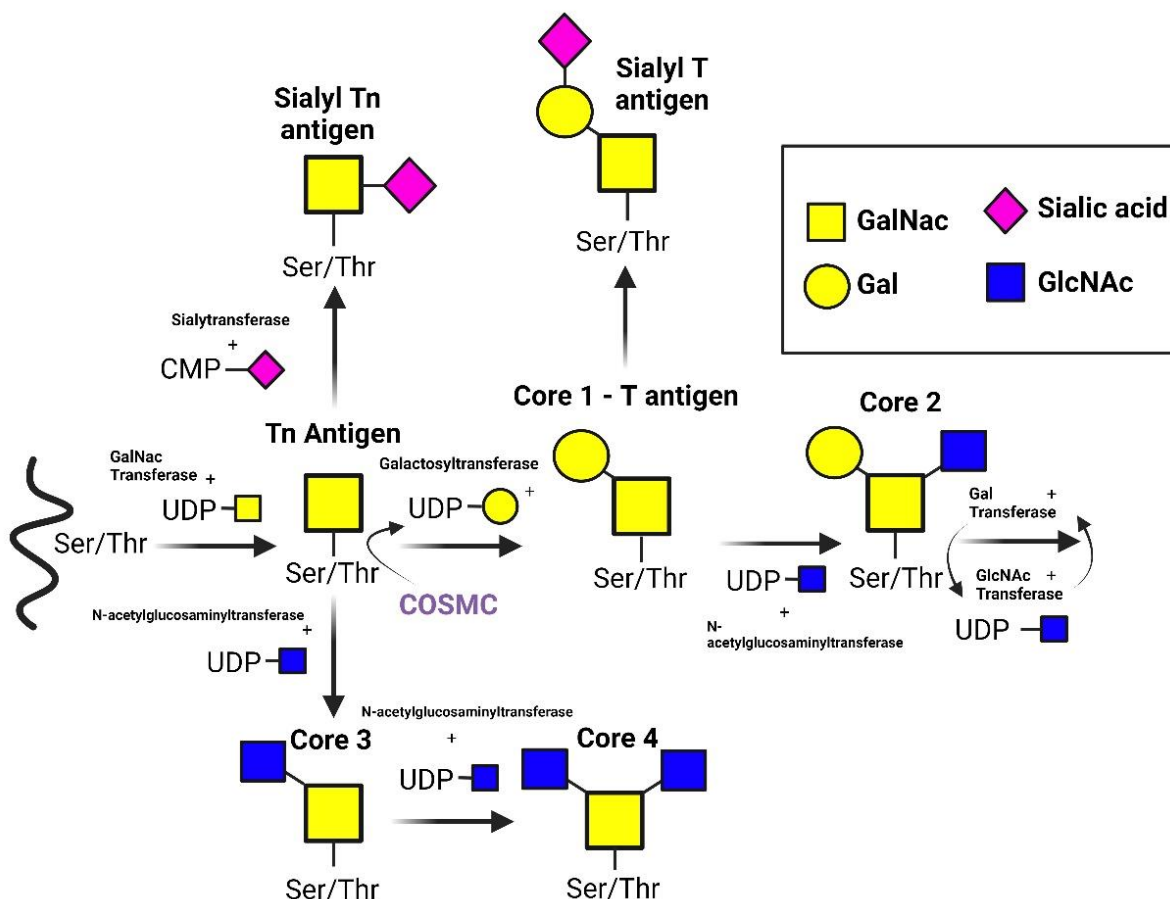
Figure 1.7. **Biosynthetic pathways of mucin-type O-Glycosylation:** *The synthesis initiates with GalNAc addition to Ser/Thr residues, forming the Tn antigen. Subsequently, core structures such as core 1, core 2, core 3, and core 4 are synthesised dependent on glycosyltransferase activity. Sialyltransferases extend the Tn antigen before core 1 synthesis, leading to the formation of the sialyl Tn antigen. Additionally, core 1 T antigen can also be extended to form the sialyl T antigen. This figure is adapted from* (Bergstrom et al., 2015).

1.3.3.2. THE SIGNIFICANCE OF O-GalNAc (MUCIN-TYPE) GLYCOSYLATION IN CANCER PROGRESSION

Abnormal mucin-type O-glycosylation is linked to numerous human diseases, including cancer development (Brockhausen, 2006). For instance, alterations in mucin-type O-glycosylation can impact the aggressiveness of tumour cells, influencing their capability to disseminate via circulation and undergo distant metastasis (reviewed by Häuselmann & Borsig, 2014). In normal cells, mucin-type O-glycan maturation involves elongation and branching of the chains, which are eventually capped with Sia or Fuc (reviewed by Zhang et al., 2022). Whereas, cancer cells frequently display truncated O-glycans, which represent an intermediate stage of biosynthesis and are commonly observed in various epithelial cell carcinomas (reviewed by Pinto & Parameswaran, 2023). In cancer, truncated glycans such as the Tn antigen, STn antigen and T antigen are most commonly observed (reviewed by Stowell et al., 2015; Varki et al., 2022).

Multiple investigations have reported elevated levels of the Tn antigen in approximately 90% of various human cancer types, such as breast, bladder, cervical, ovarian, colon, lung, gastric, and prostate cancers (Desai, 2000; Springer, 1984, 1997). Interestingly, Tn antigen has also been associated with metastasis associated with poor prognosis in cancers such as lung, cervical, colorectal, gastric, and breast cancers (reviewed by Fu et al., 2016). Interestingly, a study by Gill et al. (2013) demonstrated that the elevated levels of Tn antigen in breast cancer was strongly influenced by the subcellular localisation of GALNTs. As relocation of GALNTs from the Golgi apparatus to the ER induced synthesises of Tn antigen and promoted cell migration and invasion. The sTn antigen, a prevalent tumour-associated carbohydrate antigen has also been detected in 80% of various human cancer types and is rarely found in normal tissues (Cao et al., 1996). It is suggested that in certain cancers, the expression of ST6GalNAc1 may dominate over the expression of T synthase, leading to the formation of the STn antigen (Ogawa et al., 2017). The sTn antigen has been established to inhibit cytotoxicity of natural killer cells and has been identified to protect metastatic cells in the bloodstream from apoptosis through cell surface interactions (Ogata et al., 1992). T antigen is detected in in 90% of human cancer types, including precancerous lesions and is selectively presented on the surface of breast, colon, prostate, and bladder cancer cells, whilst also concealed in normal counterparts (Ferguson et al., 2014; Khaldoyanidi et al., 2003; Schindlbeck et al., 2005). The synthesis of T antigen is influenced not only by COSMC and C1GALT1 but also by Golgi pH gradient dissipation (Thorens & Vassalli, 1986).

1.4. LECTINS

1.4.1. OVERVIEW

Lectin are defined as a sugar-binding protein of non-immune origin that agglutinates cells or precipitates glycoconjugates (reviewed by Goldstein et al., 1980). These proteins possess at least one non-catalytic domain capable of reversibly binding to specific monosaccharides or oligosaccharides, enabling them to bind to carbohydrate components on erythrocyte surfaces and cause agglutination (reviewed by Lam & Ng, 2011). Lectins, which exhibit distinct carbohydrate specificities, have been isolated from diverse plant tissues and organisms. They are generally classified based on their carbohydrate specificity, or grouped according to their overall structure into merolectins, hololectins, chimerolectins, and superlectins, as illustrated in Figure 1.8 (reviewed by Santos et al., 2014). Lectins are ubiquitous in living organisms and have been extensively used as tools to investigate cellular glycosylation as reviewed by Brooks, (2024).
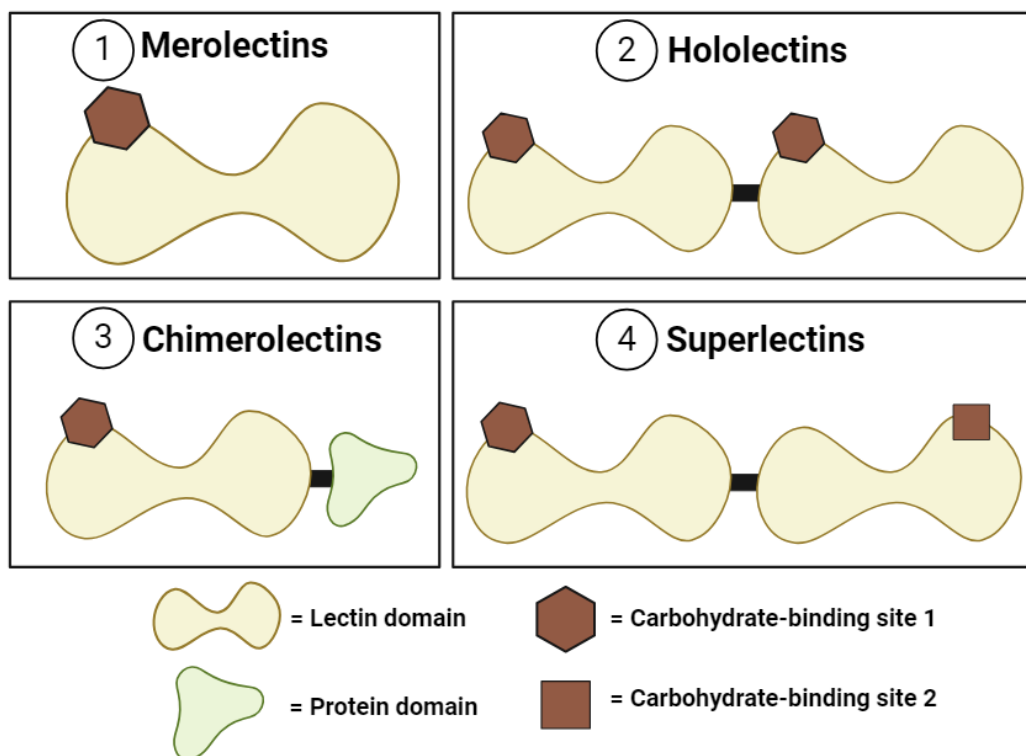
Figure 1.8. **Classification of lectins based on their structural diversity. 1)** *Merolectins are lectins composed of small monomeric lectins consisting exclusively of a single carbohydrate-binding domains.* **2)** *Hololectins comprise of at least two identical carbohydrate-binding domains.* **3)** *Chimeroletins have a carbohydrate-binding domain fused with another catalytical protein domain.* **4)** *Superlectins possess two non-identical carbohydrate-binding domains that recognise structurally different sugars.*

1.4.2. LECTINS CARBOHYDRATE-BINDING SPECIFICITY

The carbohydrate-binding specificity of lectins is typically described as a carbohydrate structure often consisting of monosaccharides, or occasionally disaccharides or trisaccharides, which best inhibits the agglutination process (reviewed by Brooks, 2024). When using lectins as biological tools, it is crucial to understand that their recognition in cells and tissues is typically more intricate than that of simple monosaccharides or disaccharides. The specificity of lectins for particular carbohydrate structures is also contingent upon various other factors such as the proximity of neighbouring glycans, their charges, and additional molecular attributes, highlighting the heightened complexity of these interactions (Roth, 2011). Moreover, the natural binding partners of lectins within biological systems remain largely unidentified and are likely to be more intricate than their defined 'inhibitory carbohydrate' or binding specificity. Example of the complexity of lectin binding is *Maackia amurensis* lectin, *Sambucus nigra* lectin, and *Polyporus squamosus* lectin which all selectively recognise Sia glycans (reviewed by Brooks, 2024). *Maackia amurensis* lectin shows a preference for α-2,3 linked Sia residues, whereas *Sambucus nigra* lectin recognises α-2,6 linked Sia residues on O-linked glycan chains, and *Polyporus squamosus* lectin preferentially binds to α-2,6 linked Sia residues on N-linked glycans. This underscores the intricate selectivity exhibited towards specific glycan structures, emphasising the influence of various binding factors at play.

1.4.3. HPA AS A PROGNOSTIC MARKER OF CANCER METASTASIS

Alterations in cellular O-linked GalNAc glycosylation are recognised as common occurrences in cancer development. These alterations in cellular glycosylation can be readily detected using sugar-binding proteins known as lectins. *Helix pomatia* agglutinin (HPA) is a lectin that has received considerable attention in the search for changes in glycosylation patterns that are associated with cancer metastasis. This lectin has a molecular weight of 79,000 kDa and comprises six identical polypeptide chains, each with a carbohydrate binding site (Sanchez et al., 2006). HPA is a hexameric glycoprotein which has previously demonstrated the ability to detect O-GalNAc (mucin-type) glycosylation, including the identification of Tn antigen (section 1.3.3.2) (reviewed by Springer, 1989). More specifically it's binding preferences include the Forssman antigen (GalNAc-a-1,3 GalNAc), blood group A substance (α-GalNAc1-3[α-Fuc1-2]Gal), Tn antigen (α-GalNAc-Ser/Thr), α-GalNAc, and α-GlcNAc (Wu & Sugii, 1991). Therefore, highlighting its recognition of a heterogeneous array of glycoproteins aberrantly GalNAc glycosylated suggests a fundamental disruption in O-linked glycosylation mechanisms (section 1.3.3.2). Brooks & Leathem, (1991) conducted a retrospective study spanning twenty-four years, involving 373 breast cancer patients, of which 293 (79%) had tumours positive for HPA and 80 (21%) had tumours negative for HPA, and significant differences in survival rates were observed between the two groups. Interestingly, their findings revealed a significant correlation between HPA-positivity and metastatic cancer. Moreover, Schumacher et al. (1994) detected the expression of HPA-binding sites using immunohistochemistry in patient tissues with colorectal cancer and revealed that binding to HPA correlated with patient survival and tumour recurrence. The authors suggested that the ability of HPA to bind GalNAc indicates that the sugar residue is at least partly involved in the process of metastasis of human colorectal carcinoma cells to regional lymph nodes and possibly to distant sites. This has also been demonstrated in a variety of other types of cancers, such as gastric (Kakeji et al., 1991), prostate (Shiraishi et al., 1992), oesophageal (Yoshida et al., 1993), and lung (Laack et al., 2002).

1.5. EXTRACELLULAR VESICLES

1.5.1. OVERVIEW

Extracellular vesicles (EVs) are membrane-surrounded structures released by all types of cells into the surrounding environment. These vesicles perform a wide array of biological functions by mediating communication between cells, delivering bioactive cargo, and inducing phenotypic changes (reviewed by Colombo et al., 2014). Bioactive cargo includes proteins, lipids, metabolites, DNA, RNA, and non-coding RNA's (miRNAs, tRNA, and rRNA) (reviewed by Chuo et al., 2018). The uptake of EVs has been shown to play a role in normal physiological processes including immune system function, angiogenesis, and stress responses (reviewed by Yates et al., 2022). However, there is growing evidence for its role in the pathological progression of several diseases, including cancer (reviewed by Osaki & Okada, 2019).

1.5.2. BIOGENESIS AND CLASSIFICATION

EVs are broadly classified into three main subpopulations (Figure 1.9), depending on their mechanism of biogenesis and size. Apoptotic bodies, the largest, range from 800 nm to 5000 nm and are formed during

programmed cell death (Poon et al., 2019). Microvesicles (MVs), ranging from 100 nm to 1000 nm, are generated by direct budding from the plasma membrane (reviewed by Tricarico et al., 2016). Exosomes, also known as sEVs, are synthesised in the endosomal compartment at 30–150 nm (reviewed by Doyle & Wang, 2019). Recently, other EV-like particle subtypes have also been described, including migrasomes, large oncosomes, exophers, supermeres, and exomeres (Di Vizio et al., 2012; Ma et al., 2015; Nicolás-Ávila et al., 2020; Zhang et al., 2018; Zhang et al., 2021).
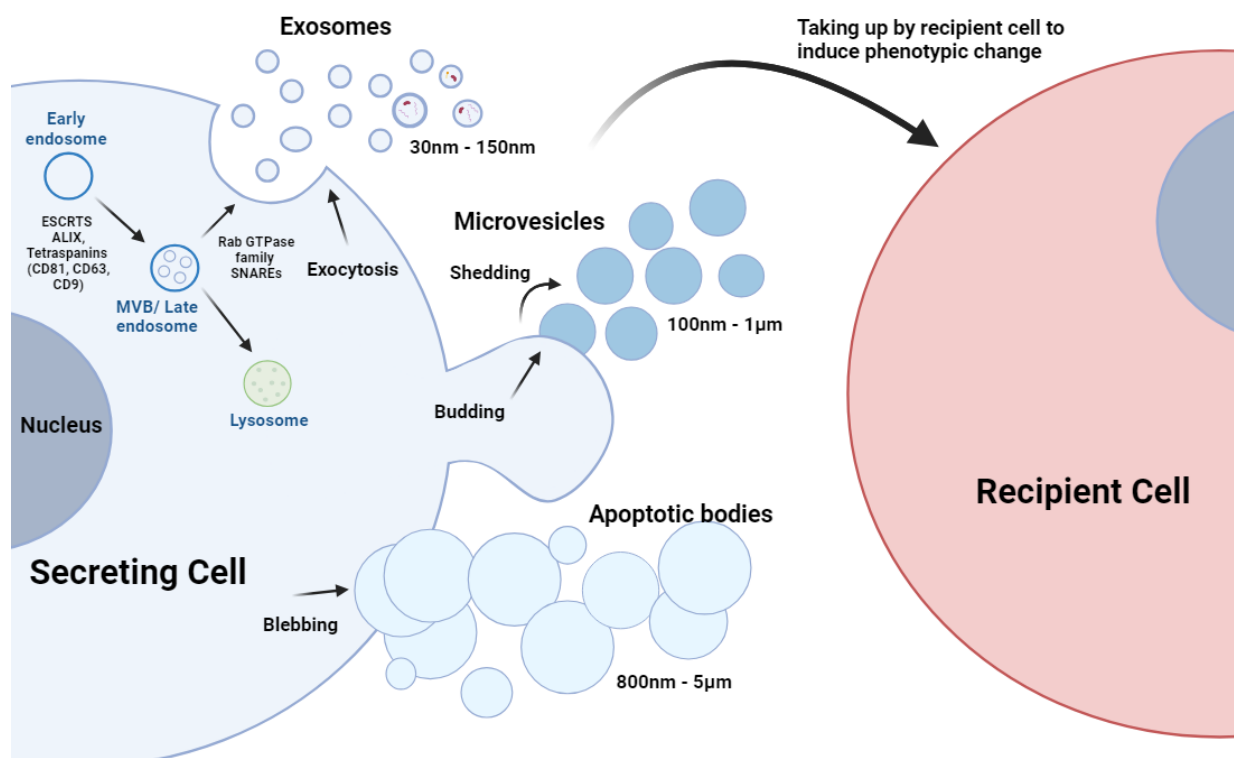
Figure 1.9. Schematic representation of EVs secretion: *Exosomes are formed by intraluminal vesicles budding into early endosomes that are released by exocytosis. Microvesicles are produced by the outwards budding and fission of the plasma membrane. Apoptotic bodies, which are the largest EVs, are formed by programmed cell death.*

1.5.2.1. EXOSOMES

The biogenesis of exosomes, often referred to as sEVs, with a size ranging from 30 to 150 nm, is a highly orchestrated process governed by the endosomal sorting complexes required for transport (ESCRT) machinery (reviewed by Doyle & Wang, 2019). Derived from the endocytic pathway, exosomes emerge as intraluminal vesicles (ILVs) within multivesicular bodies (MVBs), which are eventually released upon MVB fusion with the plasma membrane (reviewed by Zhang et al., 2019).

The canonical ESCRT-dependent pathway of ILVs begins with the stepwise recruitment of ESCRT complexes to the endosomal membrane (reviewed by Teng & Fussenegger, 2021). Phosphatidylinositol-3-phosphate (PtdIns3P) first recruits the ESCRT-0 complex to early endosomes, followed by the subsequent recruitment of ESCRT-I, ESCRT-II, and ESCRT-IIII (reviewed by Babst et al., 2002; Bache et al., 2003; Gill et al., 2007; Katzmann et al., 2003; Lu et al., 2003; Wollert & Hurley, 2010). ESCRT-III assembly is initiated by ESCRT-II recruitment,

mediated by the direct binding of CHMP6 to EAP20 (Teo et al., 2004). ESCRT-III plays a crucial role in the scission of ILV into the MVB lumen via a tightly regulated process. ESCRT-III disassociates from the membrane after scission, a process facilitated by the AAA-ATPase (Babst, et al., 2002; Wollert et al., 2009). In addition to the canonical pathway, another non-canonical ESCRT-dependent pathway can also contribute to ILV biogenesis. ALG-2 interacting protein X (ALIX) functions as a ubiquitin receptor and operates in parallel with ESCRT-0 (Baietti et al., 2012). ALIX also participates in syndecan-syntenin-ALIX ILV biogenesis and cargo sorting, involving interactions with various cellular components including lysobisphosphatidic acid (LBPA) and protease-activated receptor 1 (PAR1) (Dores et al., 2012; Kobayashi et al., 1998). Syntenin-1, which is involved in diverse cellular processes such as protein trafficking, synergistically interacts with syndecan and ALIX to facilitate ILV formation and exosome release (reviewed by Lee et al., 2023).

Another ALIX domain-containing protein, His domain–containing tyrosine phosphatase (HD-PTP), provides an alternative pathway that bypasses ESCRT-II, recruiting ESCRT-I, ESCRT-II, and ESCRT-III (Ali et al., 2013; Doyotte et al., 2008; Gahloth et al., 2017; Stefani et al., 2011; Wenzel et al., 2018). It was also suggested that the ESCRT-independent mechanism is another mechanism of ILV formation, as the simultaneous inhibition of ESCRT complexes does not eliminate MVBs (Stuffers et al., 2009). Ceramide, generated by neutral sphingomyelinase 2 (nSMase 2), can also mediate exosome biogenesis (Trajkovic et al., 2008; van Niel et al., 2011). The ceramide can induces rigid lipid raft domains, causing spontaneous negative curvature of the MVB membrane and facilitating ILV formation (reviewed by Catalano & O'Driscoll, 2019).

Tetraspanins, including CD63, CD81, and CD9, are essential for ESCRT-independent endosomal sorting by clustering on the membrane surface to form tetraspanin-enriched microdomains (TEMs) that bud into MVB (reviewed by Andreu & Yáñez-Mó, 2014). Moreover, Ras-associated binding (RAB) GTPase, specifically RAB31, controls the ESCRT- and tetraspanin-independent ceramide-dependent ILV formation pathway (Kenific et al., 2021; Wei et al., 2021). Ultimately, MVBS can fuse with lysosomes for degradation or with the plasma membrane for EV release (reviewed by van Niel et al., 2018). MVBs undergo transportation to the plasma membrane for the release of their contents as exosomes through interactions with the actin and microtubule cytoskeleton, mediated by a variety of proteins, including cortactin and several members of the Rab GTPase family (Gurung et al., 2021; Wang et al., 2023). Upon reaching the PM, Ras GTPases interact with soluble N-ethylmaleimide-sensitive factor attachment protein receptors (SNAREs) to facilitate MVB plasma membrane fusion (reviewed by Xu et al., 2022). This intricate network of pathways highlights the heterogeneity of exosome biogenesis and secretion, posing challenges in defining specific EVs markers and isolating homogeneous populations for analysis. Significant size overlap among EVs subpopulations presents a challenge for current isolation methods that rely on factors such as size, density, and molecular composition.

1.5.3. THE GLYCOSYLATION OF sEVs

1.5.3.1. OVERVIEW

The glycosylation of sEVs is often overlooked despite growing evidence supporting its significance, mainly due to the challenging nature of glycan biochemistry analysis, particularly at the nano level, as required with sEVs.

Analysis of this complex micro-heterogeneity requires either advanced mass spectrometry methods or the utilisation of lectins (section 1.4) as probes (Zaia, 2010). These methods were first applied in 2009, with initial research consistently demonstrating that sEVs possessed an abundance of glycans (Krishnamoorthy et al., 2009). The glycosylated state is hypothesised to mirror cellular events and states, conveying information that influences neighbouring cells and the surrounding environment (reviewed by Wu & Gao, 2023). A variety of glycoconjugates, including oligosaccharides, polysaccharides, and glycoproteins, which are vital components on the surface of vesicles (Figure 1.10), have also been shown to play roles in the uptake of recipient cells. (Becker et al., 2016, 2016; Lin et al., 2020). Interestingly studies have also identified the presence of tumour-associated glycans on cancer-associated sEVs, such as the Tn antigen (Feng et al., 2018), sTn antigen (Freitas et al., 2019; Nagao et al., 2022), T antigen (Gomes et al., 2015), and increased N-glycan branching (Harada et al., 2019) (Figure 1.10).
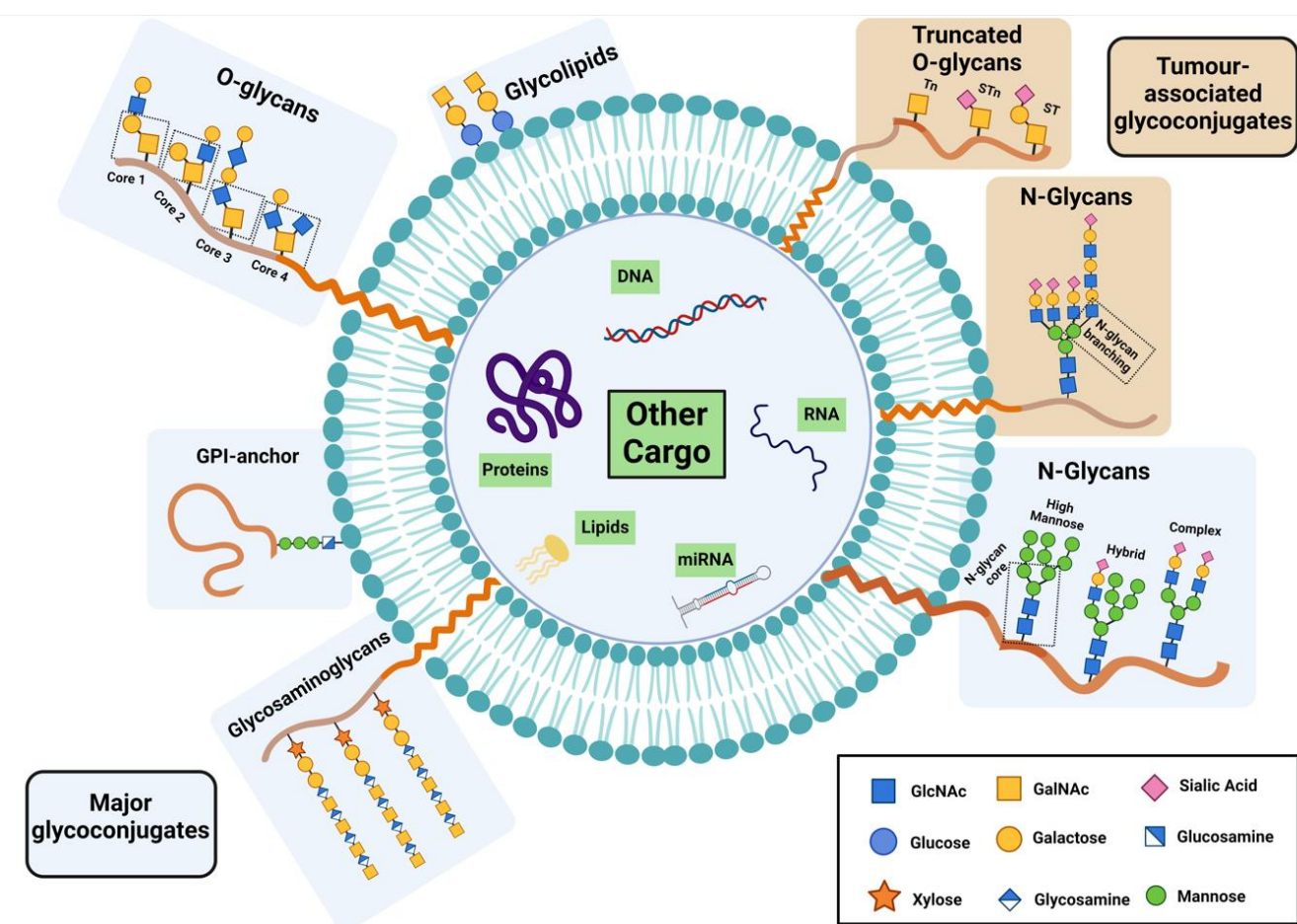
Figure 1.10. Schematic representation illustrating the glycosylation of sEVs and tumour related glycoconjugates. *sEVs carry a variety of bioactive cargo in the form of proteins, lipids, metabolites, DNA, and various RNA species. Moreover, sEVs are widely glycosylated with a variety of major common classes of glycoconjugates found in human cells such as O-glycans, N-glycans, glycosaminoglycans, glycolipids, and GPI-anchors. Tumour associated glycoconjugates have also been established on cancer associated EVs, including branching N-glycans and truncated O-glycans. Adapted from (Martins et al., 2021).*

1.6. COLORECTAL AND BREAST CANCER BLOOD BIOMARKERS

1.6.1. OVERVIEW

Breast and colorectal cancers often develop asymptomatically, with detection commonly occurring through routine screenings (reviewed by Alkabban & Ferguson, 2024; Duan et al., 2022). The National Health Service (NHS) breast screening initiative commenced in 1988, employing single-view mammography and extending invitations to women aged 50–64 years for screening once every three years (reviewed by Godley et al., 2017). By 2005, the program transitioned to two-view mammography, screening approximately 1.3 million women aged 50–70 years annually, equating to about 75% of those invited. Presently, the program detects around 10,000 breast cancers annually. (reviewed by Kekelidze et al., 2013). The NHS bowel cancer screening programme was introduced in 2006 to improve colorectal cancer mortality by earlier detection (reviewed by Koo et al., 2017). It is now offered to patients aged 60-74 years and involves a home-based guaiac faecal occult blood test (gFOBt), and if positive, patients are offered a colonoscopy, which has been associated with a 15% reduction in mortality. Moreover, in 2013 Bowel scope screening was offered to individuals aged 55 years for a one-off flexible sigmoidoscopy, and if several adenomas are found, the patients are offered a completion colonoscopy.

Despite the effectiveness of these screening programs in improving mortality rates, there remains a pressing need for new and improved diagnostic methods such as enhanced biomarkers. While the implementation of biomarkers in the UK is governed by the Medicines and Healthcare Products Regulatory Agency (MHRA), it is essential to acknowledge the regulatory oversight of biomarkers approved by agencies such as the Food and Drug Administration (FDA) in the United States. These diagnostic tests for breast and colorectal cancers encompass a wide range of biomarkers and assays for clinical application, as detailed in Supplementary Table 1 and 2. Notably, the sample types used to detect these biomarkers include blood samples. The emergence of liquid biopsies for the detection of blood-based biomarkers holds great promise in revolutionising cancer diagnosis and monitoring. By analysing circulating tumour cells (CTCs), circulating tumour DNA (ctDNA), circulatory carcinoma proteins, and other biomarkers present in blood samples, liquid biopsies offer a less invasive alternative to traditional tissue biopsies. Moreover, they provide valuable insights into tumour heterogeneity, treatment response, and disease progression, paving the way for personalised and targeted therapies.

1.6.2. ROUTINE BLOOD-BASED BIOMARKERS

Circulating carcinoma proteins play crucial roles as biomarkers in screening, diagnosis, prognosis, and therapeutic interventions for breast and colorectal cancers. For example, cancer antigen 15-3 (CA15-3) is sometimes used as a prognostic marker for breast cancer (Cancer Research U.K, 2015). CA15-3 is a large transmembrane glycoprotein frequently overexpressed and aberrantly glycosylated in breast cancer (Duffy et al., 2000). Increased serum CA15-3 levels at diagnosis correlate with advanced breast cancer stage, larger tumour size, positive axillary lymph nodes, and poorer overall and disease-free survival rates (Di Gioia et al., 2015; Fu & Li, 2016; Molina et al., 1998; Shao et al., 2015; Uehara et al., 2008). Despite its prognostic significance in early breast cancer, CA15-3 often exhibits poor sensitivity (less than 7%), highlighting the limitations of its use as a standalone diagnostic tool (Seale & Tkaczuk, 2022). Another example of a carcinoma protein biomarker is carcinoembryonic antigen

(CEA), a glycoprotein involved in cell adhesion that is often elevated in patients with colorectal cancer (Anoop et al., 2022; Cancer Research U.K, 2015). Serial measurements of CEA in the blood have demonstrated a sensitivity of approximately 80% and specificity of approximately 70% in detecting recurrent colorectal cancer, offering an average lead time of approximately 5 months (Duffy, 2001). Despite its importance in monitoring disease progression, CEA lacks the sensitivity and specificity required for screening asymptomatic individuals or independently diagnosing colorectal cancer (Kabel, 2017).

1.6.3.   THE MULTI-CANCER EARLY DETECTION TEST

The multi-cancer early detection (MCED) test, also referred to as the Galleri test, has demonstrated potential as a screening tool for early cancer detection as reviewed by Brito-Rocha et al. (2023). The 'SYMPLIFY' study marks the initial comprehensive assessment of this test among individuals referred for diagnostic follow-up following suspected cancer. The test works by examining cell-free DNA (cfDNA) released by cells from across the body into the bloodstream and other body fluids. Among individuals with cancer, an estimated 0.5% of cfDNA may originate from the tumour (Duffy et al., 2021) . The amount of DNA shed can vary based on the size of the tumour and different pathological features of the cancer (ctDNA), and is relatively small compared with DNA from other tissues and cells in the body (Pons-Belda et al., 2021). However, ctDNA can have a distinct molecular signature compared to DNA shed from healthy cells; such differences may include mutations, copy number variations, and epigenetic changes which serve as biomarkers (Lianidou, 2021). Specifically, the Galleri test uses DNA sequencing to analyse over 100,000 genomic regions that are susceptible to methylation variations (reviewed by Basharat & Horton, 2022).

After the initial discovery phase, a training and validation study evaluated the performance of the test by using blood samples from individuals with and without cancer (Liu et al., 2020). Conducted in a double-blinded manner, Liu et al. (2020) encompassed samples from over 50 different cancer types, including breast and colorectal cancer . Specifically, 12 cancer types, including colorectal cancer, were selected as high signals. For these 12 high-signal cancer types, the test demonstrated a sensitivity of 76.4% across all stages (I–IV). However, the sensitivity varied according to cancer stage, ranging from 39% for stage I samples to 92% for stage IV samples. The overall sensitivity for all 50 cancer types was lower (54.9 %) across stages I to IV. Specificity was reported to be 99.3%, with a false-positive rate of less than 1%. Further studies on the Galleri Test are currently underway to validate test performance across various populations, evaluate safety, explore its potential role in routine screening programs within healthcare systems, and gather more robust evidence through randomised trials. These ongoing investigations hold promise in enabling earlier detection and improving patient outcomes.

1.7. **PROJECT AIMS AND OBJECTIVES**

Extensive research has established the implication of significant alterations of N-linked and O-linked glycosylation throughout cancer progression. Moreover, numerous studies have revealed the impact of sEVs on the well-established hallmarks of cancer. These sEVs facilitate intercellular communication by transferring bioactive cargo such as glycoconjugates to recipient cells and inducing phenotypic changes that contribute to cancer progression and metastasis. Despite this understanding, the exploration of glycoconjugates as a predominant component of sEVs has received limited attention in comparison to other cargo. Therefore, this represents a promising frontier for further investigation, as elucidating the potential mechanisms underlying the glycosylation of sEVs could pave the way for novel diagnostic strategies in cancer management.

Therefore, the aims of this thesis were as follows: a) To develop and optimise characterisation of sEVs derived from breast and colorectal cell lines representing different phenotypes b) To investigate the influence of glycosylated targets on sEVs on *in vitro* models of cell behaviour modelling aspects of metastasis c) To identify novel glycosylated targets related to cancer and metastasis on sEVs D) To assess the diagnostic capability of sEVs glycosylated targets in patient clinical samples. The specific objectives were to:

1) To develop and optimise single-vesicle analysis for the accurate characterisation of tetraspanins of sEVs derived from breast and colorectal model cell lines representing normal, primary cancer and the metastatic tumour.

2) Analyse the HPA glycosylation patterns of cell surface in various breast and colorectal model cell lines by flow cytometry and confocal microscopy.

3) Determine the presence and abundance of HPA on sEVs derived from the breast and colorectal model cell lines by single-vesicle flow cytometry.

4) Assess the functional significance of sEVs and their HPA recognising glycans in promoting pro-metastatic characteristics.

5) Identify other unique glycans associated with cancer and metastatic breast phenotypes derived sEVs by lectin microarray and validate findings by single-vesicle analysis.

6) Assess whether the glycosylation patterns of sEVs determined from lectin microarray reflect those of the parental cell surface of the breast and colorectal model cell lines by flow cytometry.

7) Analyse glycosylated targets of plasma-enriched sEVs isolated from healthy individuals and breast cancer patients representing different stages of breast cancer to determine diagnostic significance.

CHAPTER 2

MATERIALS AND METHODS

2. MATERIALS AND METHODS

2.1. CELL CULTURE

2.1.1. BREAST AND COLORECTAL EPITHELIAL CELL LINES

Our research group has extensively characterised the breast epithelial cells used in this study for HPA lectin binding, confirming their ability to synthesise a heterogeneous profile of glycoproteins, similar to that found in clinical tumour samples (Brooks et al., 2001). Although colorectal epithelial cells have not previously been investigated in our laboratory, the existing literature supports their use as experimental models in this context, given their reported ability to display a heterogeneous profile of HPA-binding glycoproteins mirroring those observed in clinical tumour samples, therefore confirming their suitability as a model in these studies (Saint-Guirons et al., 2007; Schumacher et al., 1994; Schumacher & Adam, 1997).

The breast epithelial cell lines used in this study were chosen based on their established characteristics and origins (Table 2.1). Specifically, the hTERT-HME-1 cell line represents normal breast epithelial cells, providing a baseline comparison of cancer cell behaviour. The BT-474 cell line was selected as it represents primary breast cancer cells. Finally, the MCF7 cell line was originally derived from metastasis to the lung, representing a metastatic cancer cell line. In contrast to breast epithelial cell lines, a different approach was used to select colorectal epithelial cell lines based on their distinct characteristics (Table 2.1). The NCM460 cell line was selected to represent normal colorectal epithelial cells, providing a baseline for comparison. To represent colorectal cancer, both SW480 and HT-29 cell lines were selected. Notably, while both SW480 and HT-29 cells originate from primary tumours, HT-29 cells have demonstrated higher metastatic potential in animal models. The NCM460 cell line was kindly gifted from Dr. Heiner Schäfer, University Hospital Schleswig-Holstein, Kiel, Germany. All other epithelial cell lines were obtained from the European Collection of Authenticated Cell Cultures (ECACC).

For clarity and convenience throughout this thesis, the breast cell lines including hTERT-HME-1 will be referred to as 'normal', the BT-474 cell line as 'primary', and the MCF7 cell line as 'metastatic'. For the colorectal cell lines, the NCM460 cell line will be referred to as 'normal', the SW480 cell line as 'primary', and the HT-29 cell line as 'metastatic'. This shorthand terminology will be used for ease of reference and to maintain consistency in the discussion of the characteristics and behaviour of the cell lines in subsequent sections.

Table 2.1. Characteristics of the model cell lines used in this work

Organ of Origin	Cell line	Type	Derived from	Phenotypic characteristic references	Complete culture medium
Breast	hTERT-HME-1	Human mammary epithelial breast non-cancer cell line – immortalised using human telomerase reverse transcriptase – represents normal epithelial cells	Normal mammary epithelium of patient undergoing reduction mammoplasty surgery (no cancer history) (Band et al., 1990)	Observed to maintain normal karyotype *in vitro* (Detilleux et al., 2022)	DMEM/F12 (Sigma-Aldrich®) + 20 ng/mL (EGF) (PeproTech®) + 10 g/ml insulin (Sigma-Aldrich®) + 100 g/ml hydrocortisone (Sigma-Aldrich®) + 10% (FBS) (Gibco™)
	BT-474	Human epithelial breast cancer cell (ductal carcinoma) – represents lower metastatic potential	Derived from primary invasive ductal carcinoma (Lasfargues et al., 1978)	Do not readily produce tumours in nude mice (Lasfargues et al., 1978) In animal models, shows low or no lung metastatic potential via IV injection (Han et al., 2022)	Roswell Park Memorial Institute (RPMI)-1640 (Sigma-Aldrich®) + 10% FBS (Gibco™) + L-Glutamine 2mM (Gibco™) + Insulin 2% v/v (Sigma-Aldrich®)
	MCF-7	Human epithelial breast cancer cell line (adenocarcinoma) – represents higher metastatic potential	Derived from pleural effusion of metastatic breast cancer (Soule et al., 1973)	In animal models, showed high lung metastatic potential (Schumacher & Adam, 1997)	Dulbecco's Modified Eagle Medium (DMEM)/F12 (Sigma-Aldrich®) + 10% FBS (Gibco™) + L-Glutamine 2mM (Gibco™)
Colorectal	NCM460	Human epithelial colon non-cancer cell line – represents normal epithelial cells	Normal epithelium of patient undergoing partial gastrectomy (no cancer history) (Moyer et al., 1996)	Normal karyotype and stable genome (Huang & Wen, 2016)	McCoy's 5a (Sigma-Aldrich®) + 10% FBS (Gibco™) + L-Glutamine 2mM (Gibco™)
	SW480	Human epithelial colorectal cancer cell line (adenocarcinoma) – represents lower metastatic potential	Primary epithelial adenocarcinoma (Leibovltz et al., 1976)	In animal models, showed no lung metastatic potential via trochar implantation (Schumacher & Adam, 1997)	McCoy's 5a (Sigma-Aldrich®) + 10% FBS (Gibco™) + L-Glutamine 2mM (Gibco™)
	HT-29	Human epithelial colorectal cancer cell line (adenocarcinoma) – represents higher metastatic potential	Primary epithelial adenocarcinoma (Fogh & Trempe, 1975)	In animal models, showed high lung metastatic potential via trochar implantation (Schumacher & Adam, 1997)	RPMI-1640 (Sigma-Aldrich®) + 10% FBS (Gibco™) + L-Glutamine 2mM (Gibco™)

2.1.2. CELL LINE MAINTENANCE

The cells were cultured in T75 cm^2 cell culture flasks (Thermo Fisher Scientific) for routine passaging. The cell confluence was assessed daily visually using a primary inverted microscope (Carl Zeiss). The cell culture medium was refreshed every 48-72 hours by aspirating the spent medium and replenishing it with 10 mL of pre-warmed (37°C) culture medium (Table 2.1). This process was repeated until the cells reached 70% confluence, after which the cells were passaged to sustain cell growth. The culture medium was removed and the cells were thoroughly washed with phosphate buffered saline (PBS) (pH 7.4) (Gibco™). Subsequently, 2 mL of 0.5% w/v trypsin/ethylenediaminetetraacetic acid (EDTA) (Gibco™) in PBS (pH 7.4) solution was added to cover the entire surface of the T75cm^{2}flask. The flasks were incubated for 5 min at 37°C until the cells were fully detached. Once detached, pre-warmed culture medium was added to the flask to neutralise the trypsin, with the minimum culture media volume being at least twice the volume of trypsin/EDTA used. The cell suspension was carefully transferred into a 15 mL Falcon tube and centrifuged at 300 × g for 5 min (Megafuge 16 benchtop centrifuge, Thermo/Hereaus) to pellet the cells, and the supernatant was discarded. The cell pellet was resuspended in 10 mL of culture medium, and 1 mL of the resuspended cells was transferred to a new T75 cm^2 flask containing 10 mL of the preferred culture medium. For general subculturing, a seeding density of 2.1 x 10^6 cells was used per T75

cm^2 flask which was determined by mixing 10 μL of the cell suspension and 10 μL of trypan blue (Bio-Rad). This mix was then loaded into cell counting slides (Bio-Rad) in duplicate, and the average cell count was determined using a TC10™/TC20™ cell counter (Bio-Rad). Culturing was also conducted using T175cm^2 flasks adhering to the same protocol, but the seeding density was adjusted to 4.9×10^6 cells with 25 mL of preferred culture medium. Following seeding, flasks were placed in a humidified incubator maintained at 37°C in an atmosphere containing 5% CO_2.

2.1.3. CELL LINE STOCKS PREPARATION

In preparation for experiments, multiple cell stocks were generated to ensure a readily available supply of cells at early passages when needed. All experiments were carried out using cells at passages 1–15 to maintain consistency and reliability of the experimental outcomes. Cells were grown in flasks until confluency, trypsinised, and pelleted, as described in section 2.1.2. The cell pellet was resuspended in 1 mL freezing medium composed of 50% complete medium, 40% fetal bovine serum (FBS), and 10% dimethyl sulfoxide (DMSO) (Corning). The suspension was then pipetted into 1.2 mL cryovials (Thermo Fisher Scientific) and placed into a Mr Frosty™ freezing container (Nalgene™) which facilitated a gradual temperature decrease at -80°C. After 24 h, cryovials were placed in liquid nitrogen for long-term storage.

2.1.4. RETRIEVAL OF CELLS FROM FROZEN STOCK

Once a 1.2 mL cryovial containing a frozen cell suspension was retrieved from long-term storage in liquid nitrogen, it was thawed rapidly to ensure cell viability. This was performed by gently agitating the vial in a 37°C water bath. Subsequently, the thawed cell suspension was promptly aspirated from the cryovial and transferred into a 15 mL Falcon tube containing the preferred culture medium. This was centrifuged at $300 \times g$ for 5 min to pellet the cells and the supernatant was discarded. The resulting cell pellet was resuspended in pre-warmed (37°C) growth medium and dispensed into a T25 cm^2 cell culture flask (Thermo Fisher Scientific). The flask was then placed in an incubator at 37°C in a 5% CO_2 atmosphere for 24 h. The culture medium was aspirated and discarded, and 5 mL of pre-warmed (37°C) growth medium was added to replenish the flask. The cells were allowed to grow until they reached confluence, at which point they were passaged, as described in section 2.1.2.

2.1.5. MYCOPLASMA TESTING

Regular mycoplasma contamination testing of the cells was conducted using an e-Myco™ Mycoplasma PCR Detection Kit (iNtRON Biotechnology) following the manufacturer's instructions.

2.2. FLOW CYTOMETRY ANALYSIS OF LECTIN LABELLING ON EPITHELIAL CELL SURFACE

Flow cytometry was used to assess lectin binding to the surface of the breast and colorectal cell lines. The lectins used in the flow cytometry analysis are listed in Table 2.2. This table includes essential details such as the fluorescent labels, concentrations, sugar-binding preferences, and manufacturer information for each lectin. Cells

were brought into suspension as outlined in section 2.1.2, but instead of trypsinisation, cell detachment was achieved using 5 mL of Accutase^{TM} (StemCell Technologies) for 10 min at 37°C to ensure a gentle dissociation of cells. Following this, cells were counted as described in section 2.1.2, to obtain a cell density of 500,000 cells/ml and pelleted by centrifugation at 300 × g for 5 min. The remaining supernatant was removed, leaving a cell pellet which was resuspended in 500 μL of 3% w/v bovine serum albumin (BSA) and Tris-buffered saline (TBS) containing the lectin of interest, to be quantified in detail in Table 2.2. Controls were included for each positive lectin-labelled cell sample to validate lectin binding. These controls consisted of a negative control comprising the lectin of interest and 500 μL of 3% w/v BSA/TBS without any cells. This was done to evaluate non-specific binding of the lectin in the absence of cellular interactions. A competitive inhibitor control, varied based on the lectin; for HPA, it involved a combination of lectin and 500 μL of 0.1M GalNAc with the cell sample, to confirm lectin binding specificity. For *Lens culinaris* agglutinin (LCA) and *Tulipa* lectin (TL) analyses, since these lectins are reported to have complex glycan binding preferences, the control included 1 μL of lectin mixed with 1 mg/ml fetuin (Thermo Fisher Scientific) along with the cells, as fetuin highly glycosylated by both O-linked and N-linked glycans. Finally, the buffer-only control contained cells in 500 μL of 3% w/v BSA/TBS without any added lectin to determine the baseline fluorescence or background signal within the cell population. The conditions for lectin labelling were carefully optimised using MCF-7 cells exposed to lectins at a range of concentrations and incubation times, as illustrated in Supplementary Figure 1. MCF7 cell line was chosen because of its well-documented characteristics in previous HPA lectin binding studies by our research group as well as others. The optimisation aimed to identify the lectin concentration that provided the clearest distinction between positive (lectin-bound) and negative (non-lectin-bound) samples while also minimising nonspecific background signals. The optimal concentration was determined using a calculation known as the stain index. The stain index is a quantitative measure that helps evaluate the effectiveness of the staining process. It is calculated by subtracting the mean fluorescence intensity (MFI) of the background (unlabelled cells) from the MFI of the beads (labelled cells) and then dividing this value by the standard deviation (Stdev) of the background MFI. Using this formula, the most effective lectin concentration for the experiments was quantitatively determined. The optimised concentrations are listed in Table 2.2. Prior to analysis, each sample was prepared by adding 10 μL of a 1 mg/ml solution of propidium iodide (PI), a fluorescent dye that permeates dead cells with compromised membranes, and incubating for 10 min at room temperature (RT). This step is crucial for assessing the cell viability during flow cytometry. Subsequently, the samples were processed using a CytoFLEX flow cytometer (Beckman Coulter) equipped with 405, 488, and 638 nm lasers. For flow cytometry gating, single cells or singlets were first identified using a combination of forward scatter area (FSC-A) and forward scatter height (FSC-H) gates. These parameters help differentiate single cells from aggregates based on the size and shape of the particles. After isolating the singlets, live cell populations were identified by employing side scatter area (SSC-A) and phycoerythrin area (PE-A) gates. The PE-A gate was specifically used to detect cells stained with PI. Cells labelled with PI were considered dead and consequently excluded from the final analysis. The results were analysed using FlowJo_v10.8.1 software quantifying the median values. Each experiment consisted of biological triplicates, with three technical triplicates.

Table 2.2. Specifications of lectins used for flow cytometry analysis

Lectin	Conjugate	Preferential Sugar binding	Optimised Concentration	Manufacturer
Helix pomatia agglutinin (HPA)	Alexa Fluor 647	GalNAc	7.5ug/ml	ThermoFisher, L32454
Lens culinaris agglutinin (LCA)	DyLight 649	Glu and Man	4.0ug/ml	Vector Labs, DL1048
Tulipa lectin (TL)	Cy5 650	GalNAc, lactose, Fuc, and Gal	4.0ug/ml	Bio world, L22052

2.3. HPA LABELLING AND ASSESSMENT USING CONFOCAL MICROSCOPY

To evaluate HPA lectin binding in breast and colorectal epithelial cell lines, confocal microscopy was used for both qualitative and quantitative analyses. First, 3 mm diameter coverslips (Thermo Fisher Scientific) were sterilised with 100% industrial methylated spirits (IMS). After sterilisation, coverslips were carefully placed in 24-well plates (Corning). The coverslips were then left in a culture hood to air dry, ensuring that any residual IMS had evaporated completely. Once dried, the cells were seeded at a density of 50,000 cells/ml (section 2.1.2) in a 24-well plate and placed in a humidified tissue culture incubator set at 37°C in a 5% CO_2 atmosphere. Subsequently, the cells were washed three times with PBS (pH 7.4) to remove any unbound cells. To visualise the cellular membranes, the cells were stained with 1 μL of Cell Mask Deep Red (Thermo Fisher Scientific) diluted in 1 mL PBS (pH 7.4) to achieve a final concentration of 1:1000. Cell Mask Deep Red is a fluorescent lipid dye specifically designed to stain the plasma membranes of cells. The staining procedure was conducted for 5 min at 37°C in the dark. Following this, Alexa 488 conjugated to HPA lectin (Thermo Fisher Scientific, L11271) was added to the stained cells at a concentration of 7.5 μg/ml in 3% w/v BSA/TBS for 10 min, previously optimised by flow cytometry, as detailed in section 2.2. In addition to the HPA labelled cells, controls were also included to validate lectin binding. A negative control, where cells were processed without the addition of Alexa 488-HPA lectin to establish a baseline for staining, and sugar-specific control. In the sugar-specific control, cells were incubated with 7.5 μg/ml Alexa 488-HPA lectin in 3% w/v BSA/TBS with 0.1M GalNAc, to confirm lectin binding specificity through competitive inhibition. The cells were then washed three times with PBS (pH 7.4) to remove unbound HPA and fixed with 4% v/v paraformaldehyde (PFA) (pH 6.9, Sigma-Aldrich®) in PBS (pH 7.4) at RT for 10 min. The cells were then subjected to three washes with PBS (pH 7.4), the coverslips were gently removed from each well, mounted onto glass slides with 200 μL of Prolong Gold Antifade DAPI mounting media (Thermo Fisher Scientific), covered in tin foil to exclude light, and allowed to dry overnight at RT before long-term storage at 4°C. All cell images were acquired using an LSM 800 confocal microscope (Zeiss) with identical laser power, detector gain, and offset settings. Cell samples were imaged in x-, y-, z-, and Z-stacks of 30-40 (0.2 μm) planes. Images of the Z-stacks were processed using the FIJI software (v2.3.0), where the fluorescence signal in the channel corresponding to 488 nm excitation, indicative of HPA-lectin binding, was quantified. To analyse these images, each Z-stack was processed to select the membrane. This was achieved by creating a cell mask through membrane signal thresholding using the Huang method (Huang & Wang, 1995). The MFI from the 488

nm excitation channel was then measured and normalised to the area occupied by the cell mask stain for each cell line.

2.4. sEVs EXTRACTION

2.4.1. EXTRACTION OF sEVs FROM EPITHELIAL CELL LINE MODELS

Cells were cultured under conditions described in section 2.1.2 to achieve 50% confluence in T175 cm^2 culture flasks. Subsequently, the culture medium was aspirated, replaced with serum-free medium (SFM), and further cultured in a humidified tissue culture incubator set at 37°C in a 5% CO_2 atmosphere. After 48 h, the SFM was removed from each flask, pipetted into 50 mL Falcon tubes, and centrifuged at 300 × g for 5 min (Megafuge 16 benchtop centrifuge, Thermo/Hereaus) to remove dead cells. The supernatant was removed, and equal volumes were loaded into 50 mL high-speed 20,000 × g centrifuge tubes (Alpha Laboratories) and centrifuged at 16,500 × g for 20 min at 4°C using an Avanti JXN-26 centrifuge (Beckman Coulter) to pellet cellular debris and larger EVs of non-interest. Next, a 0.22-micron filter (Thermo Fisher Scientific) attached to a 20 mL syringe (Thermo Fisher Scientific) was pre-coated with a 0.1% w/v BSA solution in PBS (pH 7.4) to minimise non-specific binding of sEVs during the filtration of the supernatant from the previous centrifugation step into a fresh 50 mL Falcon tube. Subsequently, the filtered media were concentrated using Vivaspin 20 concentrator columns (100 kDa, GE Healthcare). After centrifugation at 3,000 × g for 15 min (Megafuge 16 benchtop centrifuge, Thermo/Hereaus) at RT, the sEVs were retained in the column, while molecules of non-interest passed through, reducing the total volume and yielding concentrated aliquots of 500 μL. The extraction of sEVs from concentrated media samples involved size exclusion chromatography (SEC). This technique separates molecules by size, as they travel through columns with different pore sizes. Larger molecules exit the column first due to exclusion from the pores, whereas smaller molecules, including sEVs, enter the pores, slowing their elution. By collecting fractions at specific intervals, sEVs were isolated based on size and separated from other media components for contaminant-free analysis. Prior to sEVs extraction, the SEC columns were prepared as follows: Econo-Pac® chromatography columns (14 cm, Bio-Rad) were loaded with 14 mL of Sepharose gel filtration media (GE Healthcare, particle size 45 μm-165 μm), topped with PBS (pH 7.4), and allowed to settle overnight. Subsequently, a column bed support was placed atop the Sepharose to prevent disruption during washing. The columns were thoroughly washed with 10 mL of PBS (pH 7.4) for three cycles before pipetting 500 μL of the concentrated media sample onto the column bed support. Once the sample had settled, the columns were topped with 10 mL of PBS (pH 7.4) and the flow-through was promptly collected in 500 μL aliquots, designated as fractions 1 through 15. Fractions 5 to 7 were recognised as sEV-containing fractions through extensive characterisation, as outlined in section 2.5. These fractions were pooled and further concentrated to the desired volume using Vivaspin 2 5kDa concentrators (GE Healthcare) at 3000 × g for 15 min (Megafuge 16 benchtop centrifuge, Thermo/Hereaus) at RT. For short-term storage (up to two days), the samples were stored at 4°C and discarded afterwards. Serum-free media processing for sEVs extraction and characterisation is shown in Figure 2.1.

2.4.2. HUMAN PLASMA SAMPLES

2.4.2.1. ETHICAL APPROVAL

Frozen human EDTA-K2 plasma clinical samples were purchased from Precision for Medicine. Clinicopathological characteristics of the plasma samples are outlined in Table 2.3. Ethical approval was obtained from the Oxford Brookes University Research Ethics Committee (Appendix 1).

2.4.2.2. HUMAN PLASMA sEVs EXTRACTION

The human plasma samples were defrosted on ice, and 500 μL of plasma was transferred into a 1.5 mL microcentrifuge tube. The tube was then centrifuged for 15 min at 1,500 × g at RT (Megafuge 16 benchtop centrifuge, Thermo/Hereaus) to pellet the larger cell debris and remove the remaining platelets. The supernatant was then subjected to a second centrifugation using an Avanti JXN-26 centrifuge (Beckman Coulter) centrifuge at 14,000 × g for 35 min at RT, and the resulting supernatant was transferred into another 1.5 mL microcentrifuge tube. sEVs were separated through SEC and concentrated as outlined in section 2.4.1. Clinical plasma sample processing for sEVs extraction and characterisation is outlined in Figure 2.1.

For brevity, clarity and consistency throughout this thesis, plasma-enriched sEVs derived from healthy individuals will be denoted as 'healthy' sEVs, while those obtained from patients diagnosed with stage IIA cancer will be designated as 'non-metastatic' sEVs, indicating the cancer that has not yet metastasised in the patients. Similarly, plasma-enriched sEVs derived from patients diagnosed with stage IV cancer will be termed 'metastatic' sEVs, indicating cancer that has metastasised in these patients. This nomenclature aims to provide a clear and standardised framework for referencing the different states of cancer progression represented by the sEV samples.

Table 2.3. Clinicopathological characteristics of human plasma samples (n=15)

Samples	Age, years	Sex	Stage at Diagnosis
Healthy Individuals (n=5)	44	F	-
	34	M	-
	47	M	-
	60	F	-
	64	M	-
Breast Cancer Patients (n=10)	64	F	IIA
	41	F	IIA
	31	F	IIA
	68	F	IIA
	67	F	IV
	77	F	IV
	75	F	IV
	60	F	IV
	62	F	IV
	54	F	IV
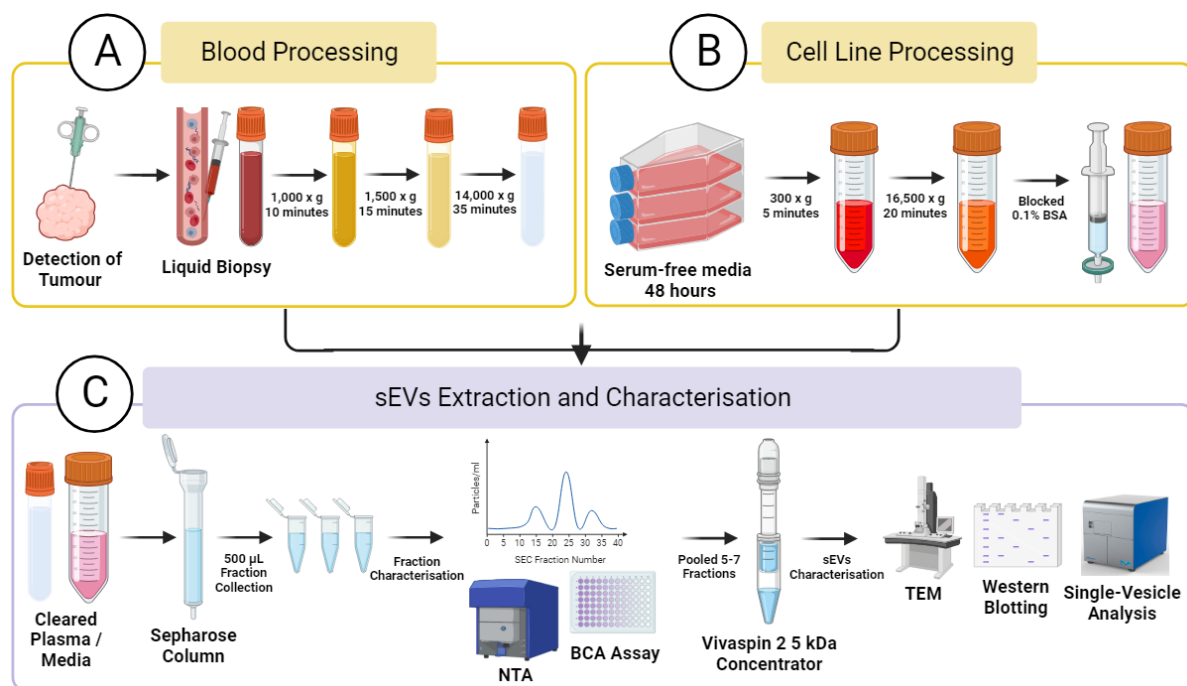
Figure 2.1. Workflow for the extraction and characterisation of sEVs (A) *Blood samples from patients with known detected tumours was centrifuged to isolate plasma and further centrifuged to obtain cleared plasma samples.* **(B)** *Serum-free media from the cell line models were processed by centrifugation and blocking with BSA.* **C)** *The processed samples were subjected to SEC. Fractions eluted during this process were collected in 500 µL aliquots and evaluated by NTA and BCA assays to determine the specific fractions containing sEVs. Fractions containing high particle counts and low protein concentrations were pooled together and subjected to further characterisation by TEM, western blotting, and single-vesicle analysis.*

2.5. sEVs CHARACTERISATION

2.5.1. NANOPARTICLE TRACKING ANALYSIS (NTA)

Nanoparticle tracking analysis (NTA) was used to assess the size distributions and concentrations of the pooled fractions containing sEVs. NTA enables the visualisation and analysis of particles in suspensions within the size range of 10 nm to 1 µm. This technique operates based on the principle of Brownian motion, which governs the random movement of particles in a fluid. It involves illuminating particles using a laser light-scattering microscope, and the light scattered by each particle is captured by a camera. The resulting motion data were then processed using specialised software to establish a connection between particle size and concentration. A ZetaView PMX 110 instrument (Particle Metrix GmbH) was used for this study. Calibration was performed using 100 nm polystyrene beads (Applied Microspheres) at a concentration ratio of 1:25,000. The pooled fractions were then diluted to a 1:1,000 ratio in a final volume of 1 mL of PBS (pH 7.4). The analysis involved running each sample through the instrument using 1 mL disposable sterile syringes (Terumo), and PBS (pH 7.4) was used to wash the system between measurements. Data were acquired at RT using the following settings: sensitivity 80, frame rate, 30 frames per second; shutter speed set to 100, minimum brightness 25, maximum pixel size 1,000, and minimum pixel size, 5.

2.5.2. WESTERN BLOTTING FOR sEVs MARKERS

To confirm the presence of sEVs, western blotting was performed for the common sEVs markers (Table 2.4). Samples containing sEVs were prepared and concentrated to a volume of 50-100 µL following the procedure described in section 2.4.1. In addition to the sEVs samples, the originating model cell lines were also processed for comparison. Following the procedure described in section 2.1.2, the cells were suspended, centrifuged at 300 × g for 5 min (Megafuge 16 benchtop centrifuge, Thermo/Hereaus) and snap-frozen at -80°C for protein extraction. Analysis of these cell lines alongside the sEVs provided a crucial positive control. This was particularly important for identifying potential cellular contaminants in the sEVs samples, such as cytochrome C, a mitochondrial marker expected in cells, but not on sEVs. sEVs were lysed by adding 50 µL of 2X radio-immunoprecipitation assay (RIPA) buffer (Sigma-Aldrich) and subjected to swirling for 20 min at 1000 rpm at 37°C using a thermomixer (Eppendorf). The snap frozen cell pellets were lysed by the addition of 500 µL of 1X RIPA buffer supplemented with Halt™ protease inhibitor cocktail (Thermo Fisher Scientific), followed by a brief vortex and incubation for 30 min on ice. Cell lysates were pelleted by centrifugation at 3000 × g for 10 min (Megafuge 16 benchtop centrifuge, Thermo/Hereaus). The protein concentration in the lysates and sEVs samples was assessed using the bicinchoninic acid assay (BCA) protein assay kit (Thermo Fisher Scientific), following the manufacturer's guidelines. Absorbance at 562 nm was measured using a SpectraMax i3x plate reader (Molecular Devices). For both cell and sEV samples, an equivalent concentration of 5µg of total protein was mixed in 4x Laemmli sample buffer (Bio-Rad) with or without 0.4mM dithiothreitol (DTT), depending on the need for reducing conditions for antibody binding (Table 2.4) at a final volume of 30 µL. Because of the typically low protein concentration observed in lysed sEVs samples, it was often necessary to use the maximum possible volume when loading the samples onto the gels. Following sample preparation, all the samples were heated at 95°C for 5 min using a dry block heater (Techne). Subsequently, they were loaded onto 8-16% gradient Mini-PROTEAN TGX stain-free pre-cast polyacrylamide gels (Bio-Rad). The 8-16% gradient refers to the concentration range of polyacrylamide in the gel designed to separate proteins of various sizes during the electrophoresis step. Alongside the samples, 5 µL of precision plus protein standards (Bio-Rad) was separated to serve as molecular weight markers for protein identification. Additionally, a H_2O control was included in each electrophoresis run to confirm the specificity of the observed bands. Electrophoresis was carried out at 125V for 90 min in running buffer consisting of Tris-glycine sodium dodecyl sulphate (SDS) (25 mM Tris, 190 mM glycine, and 0.1% SDS). The separated proteins were subsequently transferred from the gels using a Trans-Blot Turbo RTA midi polyvinylidene difluoride (PVDF) transfer kit (Bio-Rad), following the manufacturer's instructions. Proteins separated on the gel were transferred onto a membrane using the Trans-Blot Turbo transfer system (Bio-Rad) by operating the 'mixed molecular weight protocol' settings at 2.5 A and 25 V for 7 minutes. Membranes were removed from the transfer system and placed in weighing boats for a 1-hour blocking step at RT using a 5% w/v skimmed milk powder (Marvel) solution in TBS with 0.1% Tween (TBST). Following blocking, the membranes were incubated overnight at 4°C with primary antibodies diluted in the same blocking buffer (details in Table 2.4). This was followed by three 5-minute washes in TBST and a 1-hour RT exposure to secondary antibodies diluted in blocking buffer (details in Table 2.4). The washing steps were repeated, and the membranes were immersed in Clarity Max Western enhanced chemiluminescence (ECL) substrate (Bio-Rad) for 5 min before imaging using the ChemiDoc MP Bio-Rad imaging system (Bio-Rad).

Table 2.4. Specifications of antibodies for western blotting for sEVs markers

Target	Host	Dilution	Reducing Conditions	Manufacturer
CD63	Mouse, monoclonal	1:1000	NO	ThermoFisher, 10628D
CD81	Mouse, monoclonal	1:1000	NO	Abcam, ab79559
Syntenin-1	Rabbit, monoclonal	1:1000	YES	Abcam, ab133267
Cytochrome C	Rabbit, monoclonal	1:1000	YES	Abcam, ab150422
ApoB	Rabbit, monoclonal	1:1000	NO	Abcam, ab139401
Rabbit IgG	Goat, Polyclonal	1:5000	N/A	Promega, W4011
Mouse IgG	Goat, Polyclonal	1:20000	N/A	Promega, W4021

2.5.3. TRANSMISSION ELECTRON MICROSCOPY

Transmission electron microscopy (TEM) was used to assess the characteristics of sEVs in the pooled SEC fractions. Carbon 300 mesh copper grids (TAAB Laboratories Equipment Ltd) were glow-discharged for 20 seconds at 15mA and 8μl of sEVs (prepared as described in section 2.4.) were pipetted onto the grids and incubated for 2 min at RT. After blotting against filter paper to remove excess liquid, the grids were subjected to negative staining with 10 μL 2% v/v uranyl acetate in distilled water (dH$_2$O) for 10 s. The grids were gently dabbed with tissue paper, air-dried, and stored at RT in their original boxes to be disposed of after one month. The negatively stained sEVs samples were imaged using Jeol JEM-1400 Flash TEM, utilising a Gatan OneView 16-megapixel camera at 100 kV. Specifically, sEVs were measured using the line selection tool of FIJI software (v2.3.0) with a scale bar measurement set.

2.6. SINGLE-VESICLE FLOW CYTOMETRY

2.6.1. OVERVIEW

The analysis of sEVs at the single-vesicle level was performed using the Amnis® CellStream® flow cytometer (Luminex), equipped with 405, 488, 561, and 642 nm lasers. This instrument utilises a charge-coupled device employing time-delay integration (TDI), combining the benefits of high-throughput flow cytometry with high-sensitivity detection, making it ideal for the analysis of submicron particles such as sEVs. Details regarding the specifications of the lectins, antibodies, and dyes are provided in Table 2.5. The workflow for single vesicle flow cytometry is also outlined in Figure 2.2.
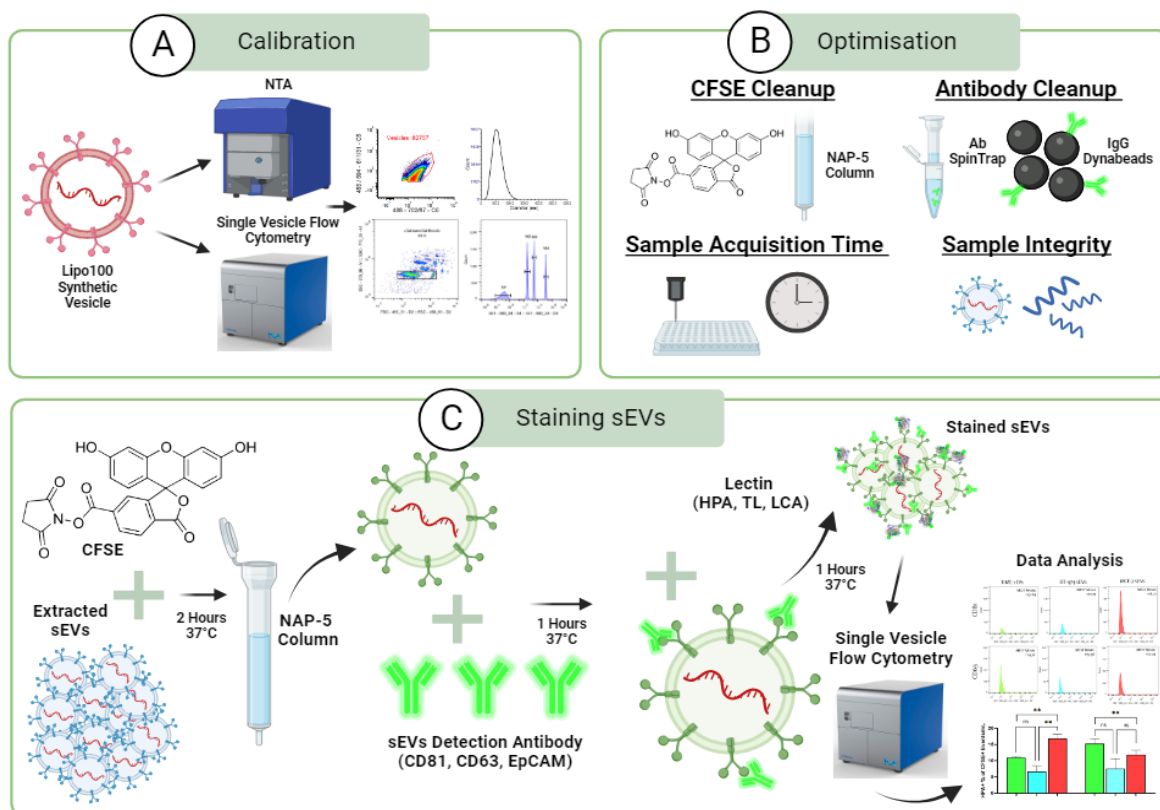
Figure 2.2. Workflow of single-vesicle analysis (A) *The calibration process involved measuring a synthetic vesicle (Lipo100) using the CellStream and NTA.* **(B)** *Optimisation was conducted to achieve the maximum signal quality during the analysis of sEVs samples. This included cleaning unbound label, defining the sample acquisition time, and ensuring sEVs sample integrity over extended periods of analysis.* **(C)** *Staining and labelling of sEVs samples involved several key steps, such as CFSE staining followed by SEC to remove unbound dye, treatment with the first detection antibody, and addition of the lectin of interest. The finalised labelled sEVs samples were then measured.*

Table 2.5. Specifications of reagents used for single-vesicle flow cytometry

Reagent	Conjugate	Isotype	Optimised Concentration	Manufacturer
Helix pomatia agglutinin (HPA)	Alexa Fluor 647	-	4.0 µg/ml	ThermoFisher, L32454
Lens culinaris agglutinin (LCA)	DyLight 649	-	7.0 µg/ml	Vector Labs, DL1048
Tulipa lectin (TL)	Cy5 650	-	4.0 µg/ml	Bio world, L22052
Carboxyfluorescein diacetate succinimidyl ester (CFDA-SE)	-	-	40 µM	Abcam, ab145291
Epithelial cellular adhesion molecule (EpCAM)	PE 561	Mouse IgG2b, κ	2.0 µg/ml	Biolegend, 324205
CD81	PE 561	Mouse IgG1, κ	0.75 µg/ml	Biolegend, 349505
CD63	PE 561	Mouse IgG1, κ	3.00 µg/ml	Biolegend, 353003
Mouse IgG2b κ isotype Ctrl	PE 561	-	-	Biolegend, 401207
Mouse IgG1 κ isotype Ctrl	PE 561	-	-	Biolegend, 981804

2.6.2. SIZING, FLUORESCENCE CALIBRATION AND GATING FOR SINGLE VESICLE ANALYSIS

The CellStream instrument was size-calibrated to confirm that it had the capability of resolving sEVs ranging from 30 to 150 nm. The instrument settings used for calibration and single-vesicle analysis were FSC turned off, SSC laser set to 5%, and all other lasers set to 100% operating in the small particle detection mode at a flow rate of 3.66 µL/min. Vesicle size was assessed using a vFC™ assay kit (Cellarcus Biosciences) containing a lipophilic probe vFRed™ used to stain lipo100™ (Cellarcus Biosciences), a vesicle size standard, as per the manufacturer's instructions. The stained lipo100™ size distribution was first measured by NTA (section 2.5.1), and the relationship between the population surface area and membrane fluorescence was determined by flow cytometry to calculate the surface area per fluorescence intensity unit (Supplementary Figure 2). This was then used to calibrate the membrane florescence axis in units of the surface area and diameter, assuming spherical particles, as illustrated in Figure 2.2. Arbitrary units of median fluorescence intensity (MFI) were converted to standardised molecules of equivalent soluble fluorochrome (MESF) units using vCal™ nanorainbow beads (Cellarcus Biosciences) (Supplementary Figure 3). To adhere to MIFlowCyt-EV guidelines for standardised sEVs flow cytometry reporting, including the latest minimal information for studies of EVs (MISEV) guidelines, the checklist is outlined in Supplementary Table 5 (Théry et al., 2018; Welsh et al., 2020).

2.6.3. CHARACTERISATION OF SURFACE MARKERS ON sEVs USING ANTIBODY AND LECTIN LABELLING

The characterisation of surface markers on sEVs using antibody and lectin labelling is illustrated in Figure 2.2. First, sEVs extracted from model cell lines or plasma samples, as detailed in section 2.4, were stained with carboxyfluorescein diacetate succinimidyl ester (CFDA-SE). This staining involves the conversion of CFDA-SE to its fluorescent form, carboxyfluorescein succinimidyl ester (CFSE), within sEVs through esterase-mediated cleavage of acetate groups. As previously optimised by Morales-Kastresana et al. (2017) sEVs were treated with 40 µM CFDA-SE in 100 µL of PBS (pH 7.4) for 2 h at 37°C in the dark. Subsequently, these samples were applied to a pre-washed NAP-5 SEC column (GE Healthcare) and 500 µL fractions containing CFSE-stained sEVs were collected. Once the entire 100 µL sample entered the column bed, 500 µL of PBS (pH 7.4) was added, and fractions were collected, with fraction 2 containing 80% of the eluted sEVs. The eluted sEVs were then labelled for 1 h at 37°C in the dark with a fluorescent detection antibody of either PE labelled anti-CD81 (BioLegend, 349505) or anti-CD63 (BioLegend, 353003) for probing cell line-derived sEVs, as detailed in Table 2.5. For plasma-enriched sEVs, a detection antibody of EpCAM (BioLegend, 324205) was used. Subsequently, the sEVs were labelled with fluorescently labelled lectin for 1 h at 37°C in the dark, as detailed in Table 2.5. All antibodies and lectins used were previously optimised by titrating on MCF-7 sEVs until the maximum signal was achieved against the background (Table 2.5 and Supplementary Figure 4). All sEVs sample labelling was performed in a final volume of 25 µL comprising 10 µL of sEVs samples, 3 µL of antibodies and lectins, followed by PBS (pH 7.4) to obtain the final volume if required. For each single-vesicle experiment run, controls were run alongside to comply with the MIFlowCyt-EV guidelines, which included sEVs negative controls, isotype, sugar specific, detergent lysis, buffer-only, buffer with reagent, and unstained samples (Supplementary Table 5) (Welsh et al.,

2020). All samples were then diluted 1:20 to a final volume of 400 μL, briefly vortexed, and 100 μL of each sample was loaded into individual wells of an ABgene Super Plate 96 well PCR plate (Thermo Fisher Scientific) and covered with X-Pierce sealing films (EXCEL Scientific Inc.) to prevent evaporation. All samples were run for 5 min with the previously calibrated instrument settings, as detailed in section 2.6.1. For single-vesicle gating (Supplementary Figure 5), the size range occupied by sEVs (30–150 nm) was initially identified using the SSC-A and FSC-A gates of Lipo100 (Cellarcus Biosciences). This step was followed by elimination of the CFSE background from the stained sEVs sample, utilising the negative control (which shows fluorescence in the absence of sEVs). This was achieved by gating on an SSC-A and 488 nm fluorescence channel, specifically used to detect the fluorescence emitted by CFSE at a wavelength of 488 nm. To further ensure specificity, non-specific binding of the PE-detection antibody was addressed. This involved a gating strategy for the isotype control samples to isolate the background signal. The gate was applied around the background fluorescence detected on the SSC-A and 561 nm fluorescence channel, which measured the background fluorescence emitted by the isotype control. This gate facilitated the quantification of events per millilitre and provided MFI values in the 647 nm channel, corresponding to the lectin used for labelling the sEVs sample. Two distinct analyses were performed throughout, comprising of MESF analysis to quantify the number of molecules on the surface of an sEV and the assessment of the proportion of sEVs in a population that is positive for each tetraspanin. To facilitate the comparison and normalisation of the varying amounts of loaded sEVs in each sample, the events/ml for the final gated lectin+ signal was calculated as a percentage of the CFSE+ events/ml. All results were analysed using FlowJo_v10.8.1, and median values were converted into standardised MESF units, as discussed in section 2.6.2.

For consistency and clarity throughout this thesis, sEVs as detected by CD81 tetraspanin will be referred to as 'CD81-positive' sEVs, sEVs as detected by CD63 tetraspanin will be denoted as 'CD63-positive' sEVs and sEVs as detected by EpCAM will be referred to as 'EpCAM-positive' sEVs.

2.7. CELL FUNCTIONAL ASSAYS

2.7.1. sEVs TREATMENT OF CELLS

For sEVs functional analysis experiments, including the wound healing assay (section 2.7.2) and static adhesion assay (section 2.7.3), sEVs were extracted from cells, as described in section 2.4. Subsequently, each sEVs sample underwent NTA to identify the particle concentration per millilitre (section 2.5.1). Following this, the samples were appropriately diluted to achieve a predetermined concentration for subsequent treatment in the cell functional assays, and both 'normal' and 'metastatic' cell lines were treated with each other's extracted sEVs, followed by their respective sEVs.

2.7.2. WOUND HEALING ASSAY

A wound-healing assay was performed to assess the motility of breast and colorectal cells following treatment with sEVs and sEVs incubated with HPA over a 24-hour period (Figure 2.3). Before cell seeding, the bottom of each well of a 12-well plate was marked with a circle so that the wound area was centred and consistent for each image. Cells were counted and then seeded into 12-well plates (Thermo Fisher Scientific) at a density of 1×10^5

cells per well (see section 2.1.2). The cells were then incubated in a humidified incubator maintained at 37°C in an atmosphere containing 5% CO_2 until they reached 90% confluence, and were then serum-starved for 24 h by treating the cells with their respective culture medium without FBS. A p1000 pipette tip was used to create the 'wound'. The tip was angled at approximately 30 °to create a controlled scratch and to limit the width of the 'wound'. This was followed by washing with PBS (pH 7.4) to remove the excess floating cells. To optimise the concentration of sEVs (see section 2.7.1) A range of concentrations of 1×10^9, 1×10^8, and 1×10^7 particles/ml was prepared in 1 mL of the respective cell line culture media (Table 2.1) alongside controls containing media (no treatment control). To investigate the potential functional effects of exposed GalNAc glycans on sEVs in modulating cell motility, sEVs were first labelled with HPA as outlined in section 2.6.3. This was followed by purification using NAP-5 SEC columns (GE Healthcare) to eliminate any unbound HPA. The HPA labelled sEVs were then used for cell treatment. A background control, with no sEVs or HPA treatment, was run concurrently to account for any functional effect observed via non-bound HPA. Images of the cells post-scratch were taken at 0, 4, 8- and 24-hours intervals using a Primovert inverted microscope. To quantify wound closure over time, FIJI software (v2.3.0) was used to draw the area around the open wound. To prevent bias in the gap quantification resulting from uneven scratch widths, the migration area was determined by subtracting the scratch area at 24 h from the initial area at 0 h.

2.7.3. STATIC ADHESION ASSAY

A static adhesion assay was performed to evaluate the ability of breast and colorectal epithelial cells to adhere to endothelial cells treated with sEVs of different phenotypes. Briefly, glass coverslips (13 mm diameter; Thermo Fisher Scientific) were sterilised in 70% v/v isopropanol in dH_2O, placed in the wells of a 24-well plate (Thermo Fisher Scientific), and air-dried in a cell culture hood. Subsequently, each coverslip was coated with 500 μL of 0.2% w/v bovine gelatin in PBS (pH 7.4) and allowed to adhere for 30 min. After removing excess gelatin, 100,000 HUVEC (refer to section 2.1.2) were seeded per well and cultured in a humidified incubator maintained at 37°C in an atmosphere containing 5% CO_2 until they reached 90% confluence. Upon reaching 90% confluence, the endothelial cells were pre-stimulated with 10 ng/ml of tumour necrosis factor alpha (TNF-α) (Sigma) for 2 h at 37°C using the culture medium specified in Table 2.1. Concurrently, epithelial cells were treated with a 10 mg/ml solution of the fluorescent dye 8-hydroxypyrenetrisulphonic acid (PTS) (Sigma), also in the respective culture medium as detailed in Table 2.1. Following incubation, the PTS solution was removed and the epithelial cells were washed five times with PBS (pH 7.4) to eliminate excess PTS. The cells were then detached using 5 mL Accutase™ (StemCell Technologies) and aspirated into 15 mL Falcon tubes. Following centrifugation at $300 \times g$ for 5 min (Megafuge 16 benchtop centrifuge, Thermo/Hereaus), the pellet was resuspended in 10 mL PBS (pH 7.4). PTS-stained epithelial cells were counted, as described in section 2.1.2, to determine their concentration. Following cell counting, the cell suspension was adjusted to obtain 100,000 cancer cells per mL in 10 mL of the corresponding culture media. Subsequently, the epithelial cells were treated with 1×10^9 sEVs (refer to section 2.7) or a control with no treatment for 1 h at 37°C incubation, TNF-α was removed from the HUVEC wells and the treated epithelial cells were co-incubated with HUVEC for 10 min at 37°C in a humidified incubator with 5% CO_2. The non-bound epithelial cells were aspirated from the wells and washed three times with PBS (pH 7.4). Epithelial cells adherent to the endothelial cells were then fixed with 1 mL of 4% v/v PFA in PBS (pH 7.4) for 10 min at RT and washed with PBS (pH 7.4) three times. The coverslips were then carefully removed from each

well, mounted onto microscope glass slides using 100 µL fluoromount anti-fade mountant (Sigma), covered in tin foil and allowed to dry overnight at RT before long-term storage at 4°C. Within 48 h, the coverslips were imaged using an upright wide-field fluorescence microscope (Zeiss). Tiling images of the entire coverslip were captured using a 10X objective lens, resulting in a total magnification of 100X. Images were analysed using the FIJI software (v2.3.0) following a custom protocol. First, composite images stitched from individual tiles were imported into FIJI and converted to 8-bit grayscale images. A specific 4 × 4 grid region displaying stained cells (16 images) was selected for a detailed analysis. Otsu's auto-threshold clustering algorithm, through the 'auto threshold' plugin, was applied to these images for segmentation (Otsu, 1979). Adherent epithelial cells were counted using the 'analysis of particle' function. For data interpretation, the number of adherent cells after sEVs treatment was normalised against the control (without sEVs) and the average fold changes were calculated.
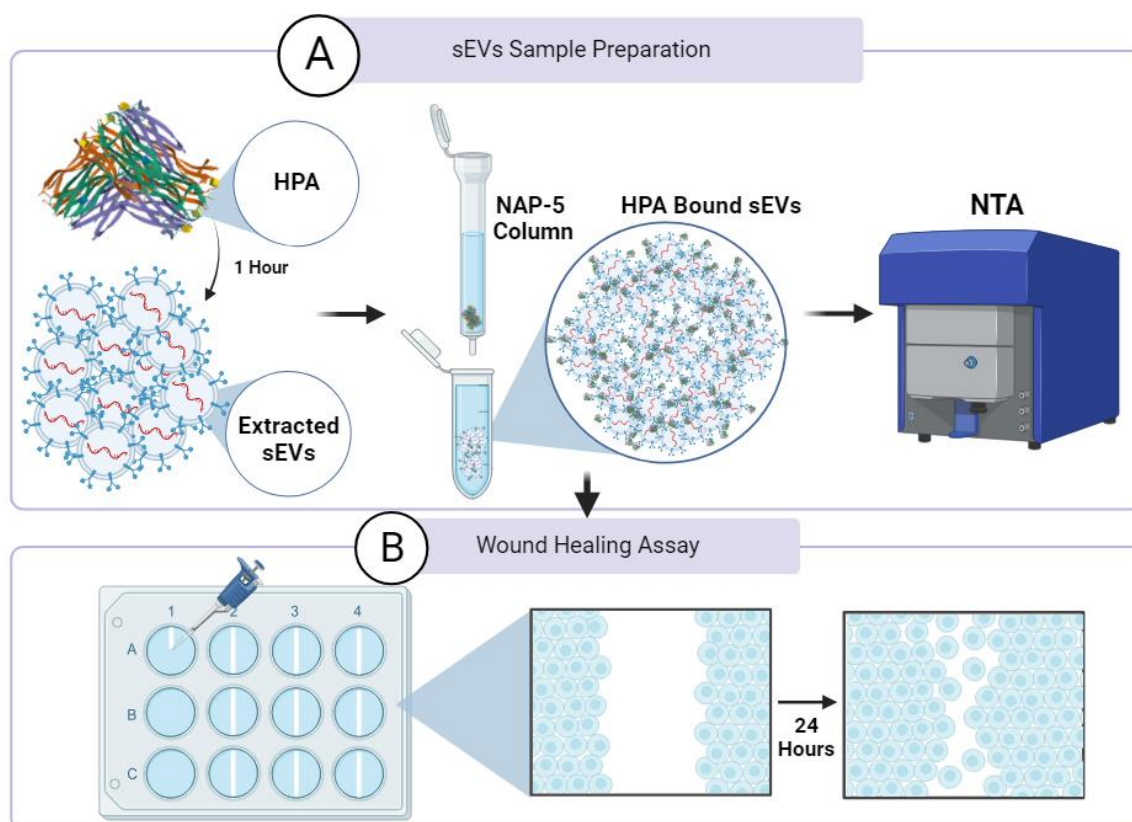
Figure 2.3. Workflow of the wound healing assay with sEVs treatment (A) *Sample preparation involved treating sEVs with HPA to mask truncated GalNAc glycans, followed by SEC to remove excess HPA. The collected HPA-bound sEVs were subjected to NTA to determine seeding densities.* **(B)** *A pipette was used to create a wound-like scratch in each seeded cell well, and measurements were taken after 24 h with various sEVs treatments.*

2.8. LECTIN MICROARRAY

Use of a lectin microarray 95 kit (RayBiotech Life) enabled rapid and sensitive analysis of carbohydrate-binding proteins, allowing for the simultaneous detection of sEVs bound to a panel of 95 different lectins. This kit was used according to manufacturer's instructions to identify surface glycans on sEVs derived from the breast cell lines under study. Supplementary Table 3 presents a comprehensive summary detailing both the identity of each lectin included in the microarray and their respective glycan-binding preferences. First, breast sEVs were extracted

from model cell lines, as outlined in section 2.4. Subsequently, the extracted breast sEVs were assessed for protein concentration using a BCA protein assay kit (Thermo Fisher Scientific) according to the manufacturer's instructions. The absorbance at 562 nm was measured using a SpectraMax i3x plate reader (Molecular Devices). The sEVs were appropriately diluted to obtain a 2 µg/ml total protein sEVs sample in 100 µL PBS (pH 7.4) and maintained on ice. The microarray slide coated with 95 lectins was left to equilibrate to RT inside a sealed plastic bag for 20-30 minutes. The microarray was then removed from the sealed plastic bag and left to air-dry at RT for an additional 2 h. Subsequently, 100 µL of sEV sample extracted from each breast cell line, in quadruplicate, alongside diluent controls, was added to each well of the microarray slide, covered with adhesive tape to prevent evaporation, and incubated overnight at 4°C on a rocking platform (Bio-Rad). The following day, the adhesive tape was removed, and each well underwent five gentle wash steps with 150 µL of 1x wash buffer I, provided as part of the lectin array 95 kit, at RT. To visualise bound sEVs, a labelling technique was employed using allophycocyanin-conjugated anti-CD81 (BioLegend, 349509). This method involves the application of fluorescently labelled antibodies specifically designed to bind to the CD81 protein on the surface of sEVs. Anti-CD81, previously optimised by single vesicle flow cytometry as detailed in section 2.6.3, was utilised at a concentration of 0.75 µg/mL. For the labelling process, the antibody was diluted in a 3% w/v BSA/TBS solution and added to each well of the microarray slide containing breast sEVs. This approach enabled the specific detection and analysis of breast sEVs through fluorescence emitted by the allophycocyanin-conjugated anti-CD81. The corresponding control well was designated for each tested sEV sample. This well was treated with the same concentration and volume of APC mouse IgG1, κ isotype control (BioLegend, 981806) as the sEV samples. Additionally, a diluent control was included. These controls were used to verify the specificity of the observed binding in the sEV samples and to account for potential non-specific interactions or background fluorescence. Each well was then incubated for 1 h at 37°C, the washing steps with 150 µl of 1x wash buffer I, provided as part of the lectin array 95 kit, were repeated at RT. The microarray slide was then placed in a microarray high-speed centrifuge (ArrayIt) and centrifuged at 1,000 rpm for 3 min to dry the microarray slide. An Innoscan 700 scanner (Carbonne) was used to scan the microarray slide and data were then analysed using the Mapix software (Innopsys). Data extraction was performed using the GAL file provided with the lectin microarray kit, which identified each spot for its respective lectin. Fluorescence intensity was obtained for each spot, and raw data were normalised against the fluorescence intensities obtained from the respective sEVs isotype control samples. A two-tailed t-test was applied across sEVs derived from different breast cell lines (BT-474 vs. MCF-7, hTERT-HME-1 vs. BT-474, and hTERT-HME-1 vs. MCF-7), and the $\log_2$ fold change was calculated for each group comparison. A Python environment was used to generate a volcano plot for each treatment with thresholds set at $\log_2$ fold-change higher than 1.5 and a p-value higher than 2 using Bioinfokit v2.0.8. Heatmaps were generated using GraphPad Prism v9 software.
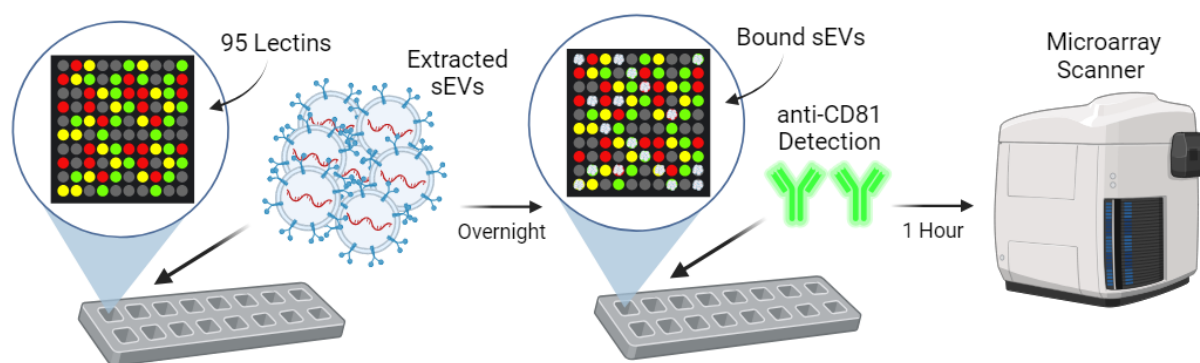
Figure 2.4. Analysis of sEVs surface glycans by using a lectin microarray. *Extracted sEVs were introduced into the wells of a glass slide containing 95 lectins and incubated overnight. To visualise the sEVs, fluorescent anti-CD81 was added to each well and the glass slides were analysed using a microarray scanner.*

2.9. MACSPLEX ANALYSIS

The MACSPlex exosome kit, human (Miltenyi Biotec), is a multiplex bead-based flow cytometry assay kit that allows for the semi-quantitative detection of 37 distinct potential surface markers on sEVs simultaneously (Supplementary Table 4). Plasma-enriched sEVs were isolated from 'healthy' individuals and patients with 'non-metastatic' and 'metastatic' cancer by SEC (section 2.4.2.2). A concentration of $1x10^9$ particles/ml, as determined by NTA (section 2.5.1), was used for the assay. The sEVs suspension (120 µL) was mixed with 15 µL MACSPlex exosome capture beads (Miltenyi Biotec). The mixture was agitated at RT in an orbital shaker at 450 rpm in a light-protected environment. Following incubation, the samples were washed with 500 µL MACSPlex buffer (Miltenyi Biotec) and centrifuged at 3000 × g for 5 min at RT (Megafuge 16 benchtop centrifuge, Thermo/Hereaus). The supernatant was removed, and the bead pellet was exposed to lectins, including HPA, LCA, and TL (previously optimised by single vesicle flow cytometry as outlined in Table 2.5) for 1 h at RT with continuous shaking. Buffer controls were also run alongside each sample to ascertain the background binding levels for data interpretation. After washing with MACSPlex buffer, the fluorescence intensity of each sample was assessed using a CytoFLEX flow cytometer (Beckman Coulter), recording at least 10,000 single-bead events per sample. Data analysis was performed using FlowJo_v10.8.1 software. Bead populations were identified based on their fluorescence characteristics in the PE and FITC channels. The background values of MACSPlex buffer and isotype controls on MACSPlex exosome capture beads were subtracted from each sample. The values were normalised to the mean MFI of the expressed markers to determine the relative expression of each marker.

2.10. STATISTICAL ANALYSIS

For statistical analysis, GraphPad Prism v9 software was used to analyse all sets of data which are presented as mean ± SD. Normality of the data sets was assessed using the Shapiro-Wilk test, and the appropriate parametric or non-parametric alternative with a multiple comparison test was then applied. To compare two sample means, a two-tailed paired t-test was applied for the analysis of paired parametric data. For paired non-parametric data, a Wilcoxon matched-pair signed-rank test was performed. An unpaired t-test with Welch's correction was used for unpaired parametric data without assumptions of Stdev. Receiver operating characteristic (ROC) curves were

constructed to evaluate the diagnostic performance of the detection antibodies with lectins to determine their prognostic value. The area under the curve (AUC) was calculated. All statistical tests are outlined in the figure legends, and p-values are indicated using asterisks denoting the following: $p < 0.05$ *, $p < 0.01$ **, $p < 0.001$ ***, and $p < 0.0001$ ****.

CHAPTER 3

THE OPTIMISATION OF SINGLE-VESICLE FLOW

CYTOMETRY FOR THE CHARACTERISATION OF

sEVs

3. THE OPTIMISATION OF SINGLE-VESICLE FLOW CYTOMETRY FOR THE CHARACTERISATION OF sEVs

3.1. BACKGROUND

The field of sEVs research has surged exponentially in recent decades owing to their key role in intercellular communication in various physiological processes and disease states, notably in the context of cancer. This heightened interest on sEVs has catalysed the development of various labelling methodologies coupled with advanced technologies. Notably, the integration of single-vesicle flow cytometry analysis for sEVs characterisation has emerged as a significant development in this field. The work described in this chapter includes the characterisation of sEVs derived from breast and colorectal epithelial cell lines. This was conducted by implementing by NTA, TEM, and western blotting. Subsequently, optimisation efforts focused on enhancing sEVs characterisation through single-vesicle flow cytometry by exploring cleanup methods of unbound labelling, optimising acquisition sample running time, and ensuring the maintenance of sEV sample integrity over time. Key tetraspanin markers of sEVs, including CD81 and CD63, were comprehensively characterised using the optimised single-vesicle flow cytometry methodology.

3.1.1. CLEANUP OF UNBOUND LABELLING

Operating at the nano level, as is necessary for sEVs, the presence of unbound dye and antibodies can present challenges in accurately visualising sEVs, potentially introducing biases into experimental outcomes. Consequently, standardised methodologies for removing excess dye and antibodies from labelled sEVs are crucial to ensure the reliability and validity of sEV characterisation outcomes. However, there is currently no standardised methodology for selecting a suitable protocol for sEV cleanup after staining and labelling (Théry et al., 2018). Some studies have employed sophisticated technologies to address these challenges, such as asymmetric flow field–flow fractionation coupled with a multi-angle light-scattering detector (Dominkuš et al., 2018). However, implementation of these techniques is often laborious and requires a high level of expertise. Therefore, there is a growing need for simpler purification methods for sEVs after staining and labelling. To address this, a method previously optimised by Morales-Kastresana et al. (2017) involving the cleanup of CFDA-SE-labelled sEVs was chosen for further optimisation along with the use of established commercial kits to improve the characterisation of sEVs by single-vesicle flow cytometry analysis.

3.1.2. OPTIMISATION OF SINGLE-VESICLE FLOW CYTOMETRY ACQUISITION

While conventional flow cytometers are primarily designed to detect cells that are significantly larger than sEVs, signals originating from sEVs often overlap with background noise, making their detection more challenging. This disparity in size and detection poses limitations and obstacles to conducting sEV experiments using conventional flow cytometers. To overcome these challenges, researchers have developed more sophisticated imaging flow cytometers. These instruments collect signals through microscope objectives and quantify them based on the images captured by charge-coupled device cameras. This approach allows for the visualisation and

individual quantification of pixels or areas within the images, making these flow cytometers ideal for sEV analysis (Görgens et al., 2019). However, the exact optimised acquisition settings for single-vesicle analysis using this type of instrument platform are limited. Given the sensitivity required for this analysis, it is crucial to understand the precise parameters necessary to accurately identify sEVs. Moreover, a notable feature of the CellStream instrument is its robust high-throughput plate-mode capability that enables simultaneous analysis of multiple sEV samples. However, the impact of this high-throughput capability on the integrity of sEVs over time remains unclear. To address this, acquisition settings were optimised by fine-tuning the running time and amount of sEV samples processed on the flow cytometer to ensure sample integrity when implementing the plate mode for analysis.

3.1.3. TETRASPANINS AND sEVs

The tetraspanin protein family is ubiquitously distributed across cell types, and are found either on the plasma membrane or within the endosomal or lysosomal compartments (reviewed by Stipp et al., 2003). Tetraspanins are recognised as crucial organisers, forming multimolecular membrane networks known as the tetraspanin web, which involves other transmembrane and cytosolic proteins, such as integrins, members of the immunoglobulin superfamily, and proteases (reviewed by Boucheix & Rubinstein, 2001; Charrin et al., 2009; Hemler, 1998, 2001, 2005). Tetraspanins play diverse roles in cellular processes like adhesion, motility regulation, membrane fusion, and signalling (reviewed by Berditchevski, 2001; Boucheix et al., 2001; Hemler, 2005, 2008; Yáñez-Mó et al., 2001). Evidence suggests that tetraspanins promote multiple cancer stages, as reviewed by Hemler, (2014). Tetraspanins such as CD81 and CD63 are specifically enriched in the membrane of sEVs and commonly used as markers (reviewed by Andreu & Yáñez-Mó, 2014). Tetraspanins significantly influence sEV functionality, participating in protein biogenesis, sEV generation, cargo sorting, binding, fusion, targeting, and selective uptake by recipient cells (Rana et al., 2012; van Niel et al., 2011). In the ESCRT-independent pathway of exosome biogenesis, tetraspanins have been established to play a pivotal role, as highlighted in section 1.5.2.1. Moreover, in mouse models of breast cancer-associated fibroblast-derived sEVs, CD81 was identified as responsible for Wnt 11 cargo sorting into sEVs (Luga et al., 2012). These vesicles were subsequently released into the tumour stroma, internalised by breast cancer cells, and contributed to cell migration and metastasis. Moreover, tetraspanins CD63 was established to sort the amyloidogenic pigment cell-specific type I integral membrane protein (PMLE) into sEVs and was essential during melanogenesis (van Niel et al., 2011). Tetraspanins were also implicated in the selection of vesicular cargo in neuroblastoma cells, with sEVs targeting specific cells based on the presence of CD63 tetraspanin (Laulagnier et al., 2004). CD81, alongside integrin αvβ3, was also identified as playing a role in directing and facilitating the uptake of sEVs by dendritic cells. The composition of tetraspanins varies among sEVs derived from different cell phenotypes, raising questions of their abundance in the context of cancer progression (Breitwieser et al., 2022). As CD81 and CD63 are specifically enriched in the membrane of sEVs and have demonstrated functional significance, this study employed a single-vesicle flow cytometry approach to quantify their presence on sEVs derived from breast and colorectal epithelial cells. The objective was to gain deeper insights into their expression across various cell phenotypes, encompassing 'normal', 'primary', and 'metastatic' cells.

3.2. AIMS AND OBJECTIVES

The experiments detailed in this chapter aimed to optimise single-vesicle flow cytometry to facilitate the characterisation of sEVs derived from breast and colorectal epithelial cell models in addition to well-established characterisation methods.

The objectives were:

1) To isolate sEVs from the model cell lines and perform characterisation of each SEC fraction by NTA and BCA assay to identify fractions that are most likely to contain sEVs.

2) To characterise the pooled fraction that is most likely to contain sEVs by NTA, western blotting for key sEVs markers, and TEM.

3) To optimise the cleanup of unbound labelling of sEVs with CFSE and antibodies, identify the appropriate sample acquisition parameters for single-vesicle flow cytometry analysis.

4) To quantify tetraspanin CD81 and CD63 profiles of the sEVs derived from the 'normal', 'primary' and 'metastatic' cancer cell phenotypes of breast and colorectal by single-vesicle analysis.

3.3. METHODS

3.3.1. EXTRACTION AND CHARACTERISATION OF EPITHELIAL CELL LINE DERIVED sEVs

3.3.1.1. EXTRACTION OF sEVs FROM EPITHELIAL CELL LINE MODELS

To extract sEVs from the epithelial cell line models of both breast and colorectal cell lines, SEC was adopted, as described in section 2.4.1.

3.3.1.2. NTA AND BCA ASSAY

NTA and BCA assay was performed on the SEC fractions to confirm effective isolation of the pooled fractions to assess their concentration and size, as described in section 2.5.1 and section 2.4.

3.3.1.3. WESTERN BLOTTING KEY sEVs MARKERS

To identify key markers, western blotting was performed on the pooled fractions to identify the presence of well-established markers of sEVs (Table 2.4), as described in section 2.5.2.

3.3.1.4. TEM

To effectively identify the morphological and size characteristics of the pooled fractions containing sEVs, TEM was used as described in section 2.5.3.

3.3.2. SINGLE-VESICLE FLOW CYTOMETRY

3.3.2.1. OPTIMISATION OF sEVs STAINING

To maximise the visualisation of sEVs by single-vesicle flow cytometry, optimisation experiments were conducted to identify the most effective methodology for labelling sEVs. The method described in section 2.6. including the instrument settings (section 2.6.1), calibration of sizing, and fluorescence in preparation for acquisition of sEVs (section 2.6.2) were applied. First, MCF-7 sEVs were extracted from their parental cells, as detailed in section 2.4.1, bulked stained with CFSE-SE, and processed using NAP-5 SEC columns, as outlined in section 2.6.3, alongside stained CFDA-SE-labelled sEVs that were not subjected to the column, to assess the necessity of the cleanup step. Next, both CFDA-SE-labelled sEVs either processed by the NAP-5 SEC column or not were then subjected to antibody labelling at 1 mg/ml for 1 h at 37°C in the dark with a fluorescent detection antibody of PE labelled anti-CD81 (BioLegend, 349505) in 25 µL comprising 10 µL of sEVs samples, 3 µL of CD81, followed by PBS (pH 7.4) to obtain the final volume. Controls were included in all experimental runs to adhere to the MIFlowCyt-EV guidelines, encompassing sEVs negative controls, isotype controls, detergent lysis controls, buffer-only controls, buffer with reagent controls, and unstained samples (Welsh et al., 2020). All samples were appropriately diluted and run on the CellStream instrument, followed by strategic gating and analysis, as outlined in section 2.6.3.

3.3.2.2. OPTIMISATION OF ANTIBODY CLEANUP

To assess whether the purification of commercially purchased antibodies would enhance the signal of interest, a series of optimisation experiments were conducted. The method described in section 2.6, including the instrument settings (section 2.6.2) and calibration of sizing and fluorescence for sEV acquisition was implemented. Subsequently, MCF-7 sEVs were extracted from their parental cells, as detailed in section 2.4.1, stained with CFDA-SE, and processed using NAP-5 SEC columns, as outlined in section 2.6.3. Prior to antibody labelling, fluorescently labelled PE anti-CD81 (BioLegend, 349505) was purified using Dynabeads^{TM} Protein G (Invitrogen) and subjected to a SpinTrap^{TM} column (Cytiva) according to the manufacturer's instructions. The purified anti-CD81 was assessed for protein concentration using a NanoDrop 2000c spectrophotometer (Thermo Scientific). Next, the stained sEVs samples labelled for anti-CD81 were subjected to DynabeadsTM Protein G (Invitrogen) or SpinTrapTM column (Cytiva). Additionally, anti-CD81 not subjected to any purification method was added to the stained sEVs samples at a concentration of 1 mg/ml for 1 h at 37°C in the dark. A fluorescent detection antibody of PE labelled anti-CD81 (Biolegend, 349505) in 25 μL, comprising 10 μL of sEVs samples, 3 μL of CD81, followed by PBS (pH 7.4) to obtain the final volume. All samples were then appropriately diluted, aliquoted into individual wells of an ABgene Super Plate 96 well PCR plate (Thermo Fisher Scientific), and run on the CellStream instrument, followed by strategic gating and analysis, as outlined in section 2.6.3.

3.3.2.3. OPTIMISATION OF sEVs SAMPLE INTEGRITY OVER TIME

To ensure the integrity of the sEVs samples throughout the acquisition period, an optimisation experiment was conducted. The method described in section 2.6 including the instrument settings (2.6.1), calibration of sizing and fluorescence in preparation of acquisition of sEVs (2.6.2) were applied. sEVs were isolated from MCF-7 cells, as detailed in section 2.4.1. Staining with CFDA-SE and processing through NAP-5 SEC columns was performed, followed by antibody labelling with PE-labelled anti-CD81 (Biolegend, 349505), as outlined in section 2.6.3. All samples were diluted 1:20 to a final volume of 500 μL, briefly vortexed, and 100 μL of each sample was loaded into individual wells of an ABgene Super Plate 96 well PCR plate (Thermo Fisher Scientific) in triplicate, with 9 H$_2$O samples interspersed between each triplicate. This design resulted in 45-minute intervals between each triplicate run, extending up to the 315-minute time point, at which the last triplicate measurement was taken. All samples were run on a CellStream instrument, followed by strategic gating and analysis, as outlined in section 2.6.3.

3.3.2.4. TETRASPANIN PROFILING OF EPITHELIAL DERIVED sEVs

To quantify CD81 and CD63 tetraspanin composition of the epithelial-derived sEVs of breast and colorectal cell lines, single-vesicle flow cytometry was used. The methodology for extracting sEVs from their parental cells is detailed in section 2.4.1. The methodology for single-vesicle flow cytometry is outlined in section 2.6 including the instrument settings (section 2.6.1), calibration of sizing and fluorescence in preparation of acquisition of sEVs (section 2.6.2), and sEVs lipid staining and antibody labelling methodology to identify CD81 and CD63 (section 2.6.3).

3.4. RESULTS

3.4.1. CHARACTERISATION OF BREAST AND COLORECTAL EPITHELIAL DERIVED sEVs

The effectiveness of the methodology adopted to isolate sEVs was evaluated by analysing the particle and protein concentrations of each fraction eluted from the SEC column using NTA and BCA. Analysis revealed that fractions 5, 6, and 7, each collected with a fraction size of 500 μL, had the highest particle counts, indicating that they represented points at which sEVs were most effectively separated from the column with the lowest protein contamination for both breast and colorectal epithelial sEVs (Figure 3.1 A-C and Figure 3.2 A-C). Subsequently, these fractions were pooled for further characterisation. The methods used to characterise size and morphological characteristics included TEM and NTA, whereas western blotting and single-vesicle flow cytometry were used to identify the presence of key sEV markers. TEM analysis revealed the detection of negatively stained, 'cup-shaped' particles within the size range of 30-150 nm for both breast and colorectal sEVs (Figure 3.1 D-E and Figure 3.2 D-E). Western blot analysis was conducted to confirm the presence of key markers on sEVs extracted from various epithelial cell lines, each representing distinct phenotypes. In the analysis of breast sEVs, BT-474 and hTERT-HME1 samples exhibited strong bands for both CD81 and CD63, indicating a significant presence of these markers (Figure 3.1-F). In contrast, the MCF-7 sEVs displayed weaker bands for the same markers. However, strong bands were observed for syntenin-1 in MCF-7 and BT-474 sEVs, whereas weaker bands were observed for hTERT-HME1 sEVs (Figure 3.1-G). When examining western blots of whole breast cell lysates, strong bands were evident for CD63, whereas the bands for CD81 and syntenin-1 were notably less pronounced. Western blot analysis of colorectal epithelial sEVs revealed varied marker expressions: HT-29 sEVs displayed strong CD63 and CD81 bands, SW480 sEVs had weak CD63 but strong CD81 bands, and NCM460 sEVs showed strong bands for both markers (Figure 3.2-F). Strong bands for Syntenin-1 were detected in all colorectal sEVs samples. Comparisons of whole colorectal cell lysates revealed that HT-29 cells had strong CD63 bands but lacked CD81 and syntenin-1 bands, SW480 cells exhibited strong CD63 but weak CD81 and syntenin-1 bands, and NCM460 cells showed strong CD63 with weak bands for both CD81 and syntenin-1 (Figure 3.2-G). The mitochondrial marker cytochrome C was undetectable in blots of all breast and colorectal sEVs, suggesting the absence of cellular contamination in the sEVs samples. (Figure 3.1 F-G and 3.2 F-G). In addition, an H_2O control was included in each western blot to confirm the specificity of the observed bands, revealing no non-specific binding. NTA was used to determine the particle sizes of the epithelial sEVs. Specifically, the average particle size of the breast-derived sEVs was determined as follows: hTERT-HME1-derived sEVs exhibited an average size of 75 nm, whereas both BT-474- and MCF-7-derived sEVs had an average size of 105 nm (Figure 3.1 H-J). In the case of colorectal-derived sEVs, NCM460 sEVs displayed an average size of 75 nm, while SW480 and HT-29 sEVs each showed an average size of 105 nm (Figure 3.2 H-J).
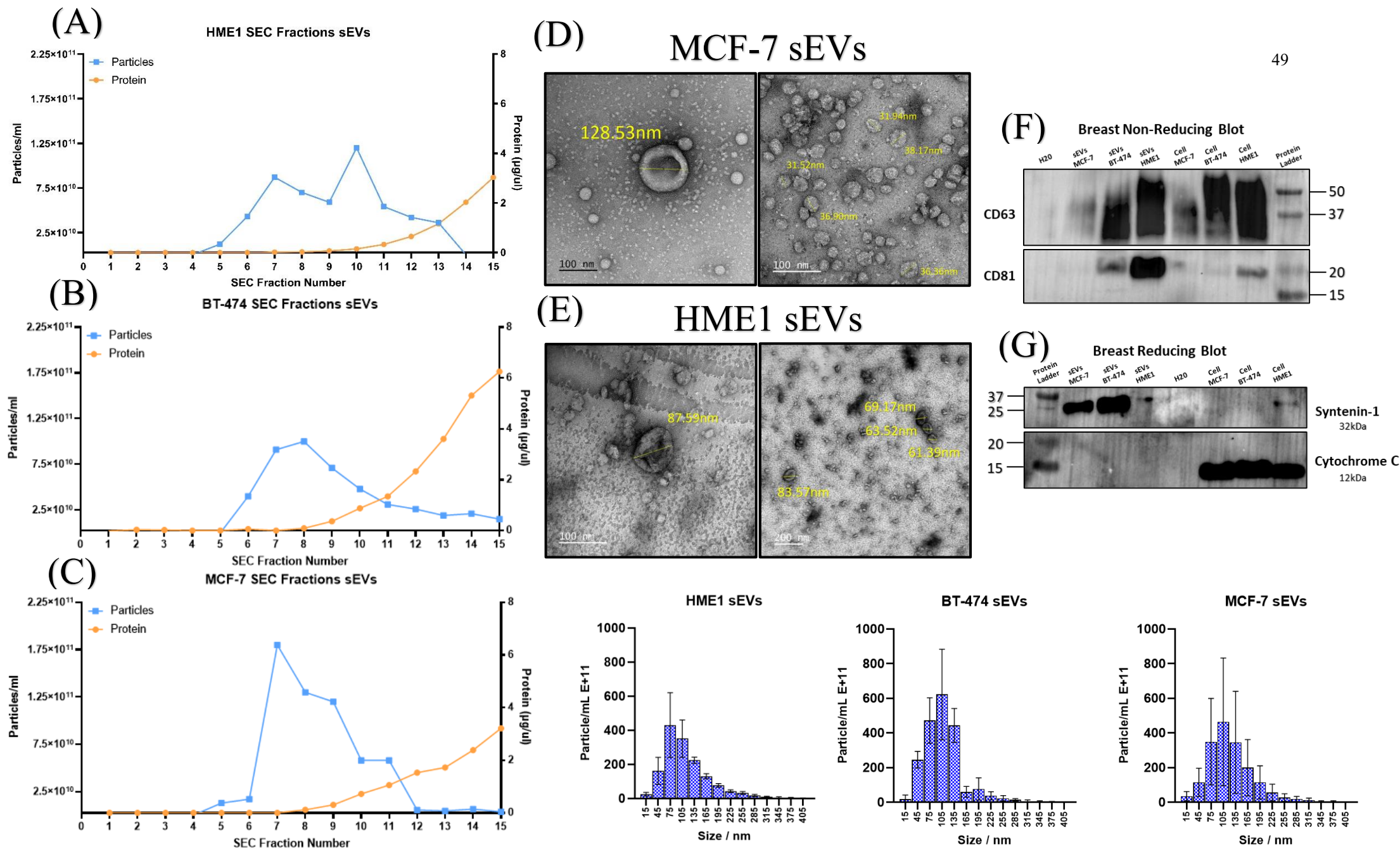
Figure 3.1. **Characterisation of SEC fractions for the isolation of breast cell line-derived sEVs by NTA, BCA, TEM and western blotting. (A-C)** *SEC fractions (500ul) containing hTERT-HME1, BT-474, and MCF-7 sEVs were subjected to NTA and BCA analysis to determine the particle and protein concentrations to identify fractions which most likely contained sEVs. **(D-E)** TEM analysis of MCF-7 and hTERT-HME1 sEVs from fractions 5, 6, and 7 pooled together. (Scale bars = 100 and 200 nm).* **(F)** *Western blot of sEVs protein markers (CD63 and CD81) under non-reducing conditions. (G) Western blotting of sEVs protein markers (syntenin-1) under reducing conditions and confirmation of sEVs lysates are clear cytochrome C marker indicating no mitochondria contaminates were present from dead cells.* **(H-J)** *NTA size characterisation of breast cell-line-derived sEVs from pooled fractions 5, 6, and 7.*
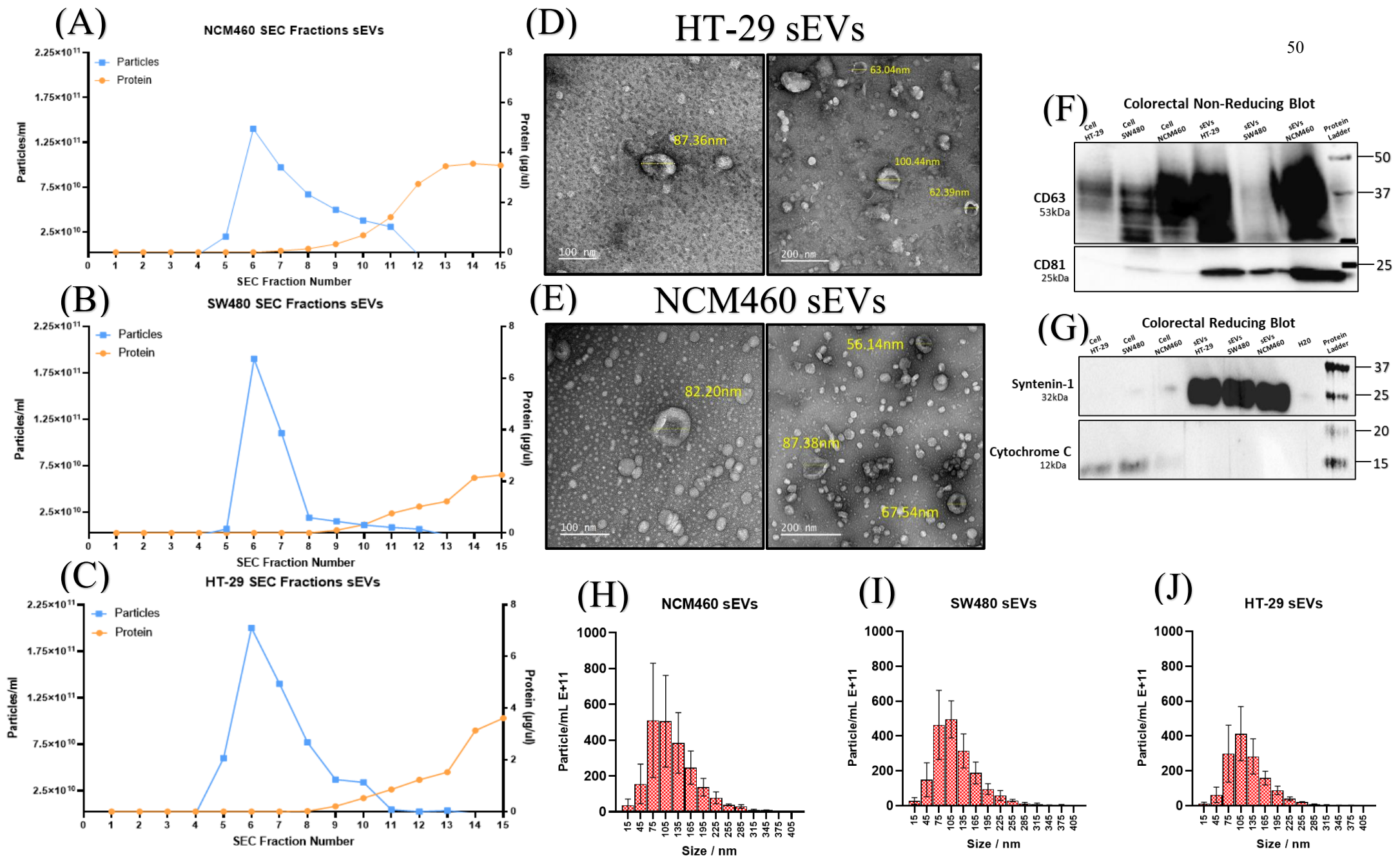
Figure 3.2. Characterisation of SEC fractions for the isolation of colorectal cell line derived sEVs by NTA, BCA, TEM and western blotting. *(A-C) SEC fractions (500ul) containing NCM460, SW480, and HT-29 sEVs were subjected to NTA and BCA analysis to determine the particle and protein concentrations to identify fractions which most likely contained sEVs. (D-E) TEM analysis of HT-29 and NCM460 sEVs from fractions 5, 6, and 7 pooled together. (Scale bars = 100 nm and 200 nm).* **(F)** *Non-reducing conditions western blotting of sEVs protein markers (CD63 and CD81)* **(G)** *Reducing conditions western blotting of sEVs protein markers (syntenin-1), and confirmation of sEVs lysates are clear cytochrome C marker indicating no mitochondria contaminates were present from dead cells. (H-J) NTA size characterisation of colorectal cell-line-derived sEVs from pooled fractions 5, 6, and 7.*

3.4.2. OPTIMISATION OF CFSE STAINING AND ACQUISITION TIME FOR SINGLE-VESICLE FLOW CYTOMETRY

To identify the optimal methodology for single-vesicle flow cytometry, a series of optimization experiments were conducted. The primary objective was to establish a robust staining and acquisition process for the accurate detection of sEVs. In this context, the staining process was first evaluated to determine the necessity of removing the excess dye using Cytiva NAP™-5 (NAP-5) columns for downstream quantification. Subsequently, the presence of CD81 on stained sEVs samples was assessed to determine the effectiveness of removing excess dye in achieving optimal signal. Notably, when no cleanup was applied to the CFSE-stained samples, a significant fold increase of 2.3 was observed compared with the implementation of NAP-5 ($p < 0.001$), suggesting that an additional step of cleanup yielded less CFSE+ events/ml (Figure 3.3-A). Conversely, analyses of 'CD81-positive' events/ml revealed a contrasting trend, as cleanup procedures yielded a substantial increase of 3.8 ($p < 0.05$), suggesting an increased ratio of CD81 positive particles (Figure 3.3-B). This observation emphasises the essential cleanup steps of excess CFSE when introducing an additional antibody, such as CD81, for the specific detection of the protein of interest. The significance of cleanup in this context is evident because enhancing the signal-to-noise ratio facilitates the identification of the targeted protein on sEVs. In addition to refining the staining procedure, the instrumental high-throughput capability for sEVs measurement was assessed. Initially, the measurement of 'CD81-positive' events/ml was conducted across varying time intervals to optimise the sample acquisition (Figure 3.3-C). Interestingly, it was determined that a minimum acquisition duration of 5 min is essential to attain the maximum signal from the sEVs sample. Furthermore, the integrity of the sEVs samples was assessed by measuring 'CD81-positive' events/ml in the same samples over time. The findings indicated that samples assessed beyond the 67-minute and 30-second mark exhibited a decline in integrity, resulting in a diminished signal of interest (Figure 3.3-D). The efficacy of the antibody cleanup method was also investigated, recognising the necessity for increased sensitivity at the single-vesicle level. However, the analysis revealed that no significant increase in the signal was attained compared with the conditions without antibody cleanup (Figure 3.3-E).

Cytiva NAP™ NAP-5 Columns CFSE Cleanup
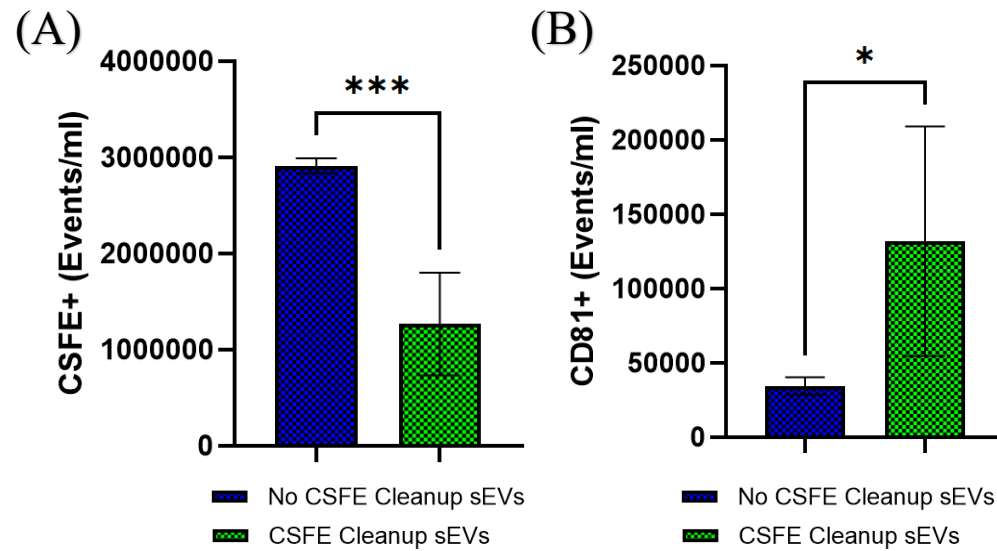
Figure 3.3. **Optimisation of single vesicle analysis** *(A-B) The optimisation of bulk CFSE staining with and without cleanup by Cytiva NAP™ NAP-5 Columns quantifying CFSE+ (events/ml) and CD81+ (events/ml) (C) The optimisation of acquisition time of each sample by quantification of CD81 (events/ml) (D) The evaluation of sEVs samples integrity over time by quantification of CD81 (events/ml) (E) The comparison of different antibody cleanup methods by quantifying CD81+ (events/ml). Each experiment consisted of biological triplicates with three technical triplicates, and the error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. Welch's correction was used for parametric t-tests, and the Mann-Whitney test was used for nonparametric t-tests. Significance levels were set at *p < 0.05, **p < 0.01 and ***p < 0.001.*

3.4.3. COMPARATIVE ANALYSIS OF CD81 AND CD63 BINDING OF BREAST AND COLORECTAL sEVs

Single-vesicle flow cytometry was performed to assess the CD81 and CD63 tetraspanin profiles of breast and colorectal epithelial cell lines. The utilisation of single-vesicle flow cytometry allowed for detailed analysis of individual vesicles, shedding light on the specific CD81 and CD63 profiles. The raw results, including the mean ± SD of CD81 and CD63 normalised to the CFSE events/ml and MESF values, are outlined in Table 3.1. Comparative analysis of tetraspanin levels revealed distinct differences between breast and colorectal epithelial sEVs (Figure 3.4 and Figure 3.5). For 'normal' breast hTERT-HME1 sEVs, the abundance of CD63 was notably higher, showing a 1.4-fold increase compared with that of CD81 ($p < 0.05$) (Figure 3.4-A). In contrast, cancer-associated MCF-7 sEVs exhibited a 1.5-fold higher abundance of CD81 than CD63 ($p < 0.001$) (Figure 3.4-A). However, no significant variations were observed in the BT-474 sEVs and colorectal sEVs for tetraspanin. Upon further extending the analysis to different sEV phenotypes, remarkable differences were noted in CD81 levels among breast epithelial sEVs. Specifically, MCF-7 sEVs displayed 1.8-fold and 3.1-fold higher CD81 abundance than hTERT-HME1 ($p < 0.01$) and BT-474 sEVs ($p < 0.0001$), respectively (Figure 3.4-A). Conversely, the comparison of CD81 levels between hTERT-HME1 and BT-474 sEVs did not show any significant difference. MESF analysis revealed that in both hTERT-HME1 and BT-474 sEVs, CD63 levels were 1.2-folder higher than those of CD81 (hTERT-HME1: $p < 0.001$; BT-474: $p < 0.0001$) (Figure 3.4-B). Interestingly, no significant differences in MESF values were observed in MCF-7 sEVs. Additionally, hTERT-HME1 sEVs exhibited 1.1-fold higher MESF values for CD81 and CD63 than BT-474 sEVs ($p < 0.01$. In the case of colorectal sEVs, the only significant differences observed were through MESF analysis which indicated a consistent increase in CD63 levels, with a 1.2 higher value in NCM460 ($p < 0.01$), SW480 ($p < 0.01$), and HT-29 ($p < 0.05$) sEVs compared to CD81 (Figure 3.5-B).

Table 3.1. Tetraspanin analysis of CD81 and CD63 on breast and colorectal epithelial sEVs by single-vesicle flow cytometry.

Organ of Origin	Cell Line sEVs	CD81 PE MESF (mean ± SD)	CD63 PE MESF (mean ± SD)	CD81 events/ml % of CFSE events/ml (mean ± SD)	CD63 events/ml % of CFSE events/ml (mean ± SD)	CD81 vs CD63 Expression	Vs HME / NCM460 sEVs	Vs MCF-7/HT-29 sEVs
Breast	hTERT-HME1 sEVs	108.16 ± 4.10	130.54 ± 8.96	16.033%±0.237	21.66%±4.96	**Events/ml** * **MESF** ***		**Events/ml** **CD81= **** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**
	BT-474 sEVs	101.17 ± 3.62	119.43 ± 1.66	9.78%±5.09	10.41%±8.15	**Events/ml** NS **MESF** ****	**Events/ml** **CD81= NS** **CD63= NS** **MESF** **CD81= *** **CD63= ***	**Events/ml** **CD81= ***** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**
	MCF-7 sEVs	117.02 ± 26.42	129.02 ± 8.95	28.94%±2.59	18.16%±3.18	**Events/ml** *** **MESF** NS	**Events/ml** **CD81= **** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**	
Colorectal	NCM460 sEVs	104.22 ± 3.39	125.00 ± 14.50	17.05%±6.61	21.42±8.96	**Events/ml** NS **MESF** **		**Events/ml** **CD81= NS** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**
	SW480 sEVs	104.16 ± 4.74	121.97 ± 13.29	18.49%±6.05	23.98±11.12	**Events/ml** NS **MESF** **	**Events/ml** **CD81= NS** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**	**Events/ml** **CD81= NS** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**
	HT-29 sEVs	106.58 ± 7.71	124.23 ± 14.97	20.66%±6.00	24.68±8.38	**Events/ml** NS **MESF** *	**Events/ml** **CD81= NS** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**	

The asterisks (*) in the table denote significance levels.

The NS in the table indicates no significance.
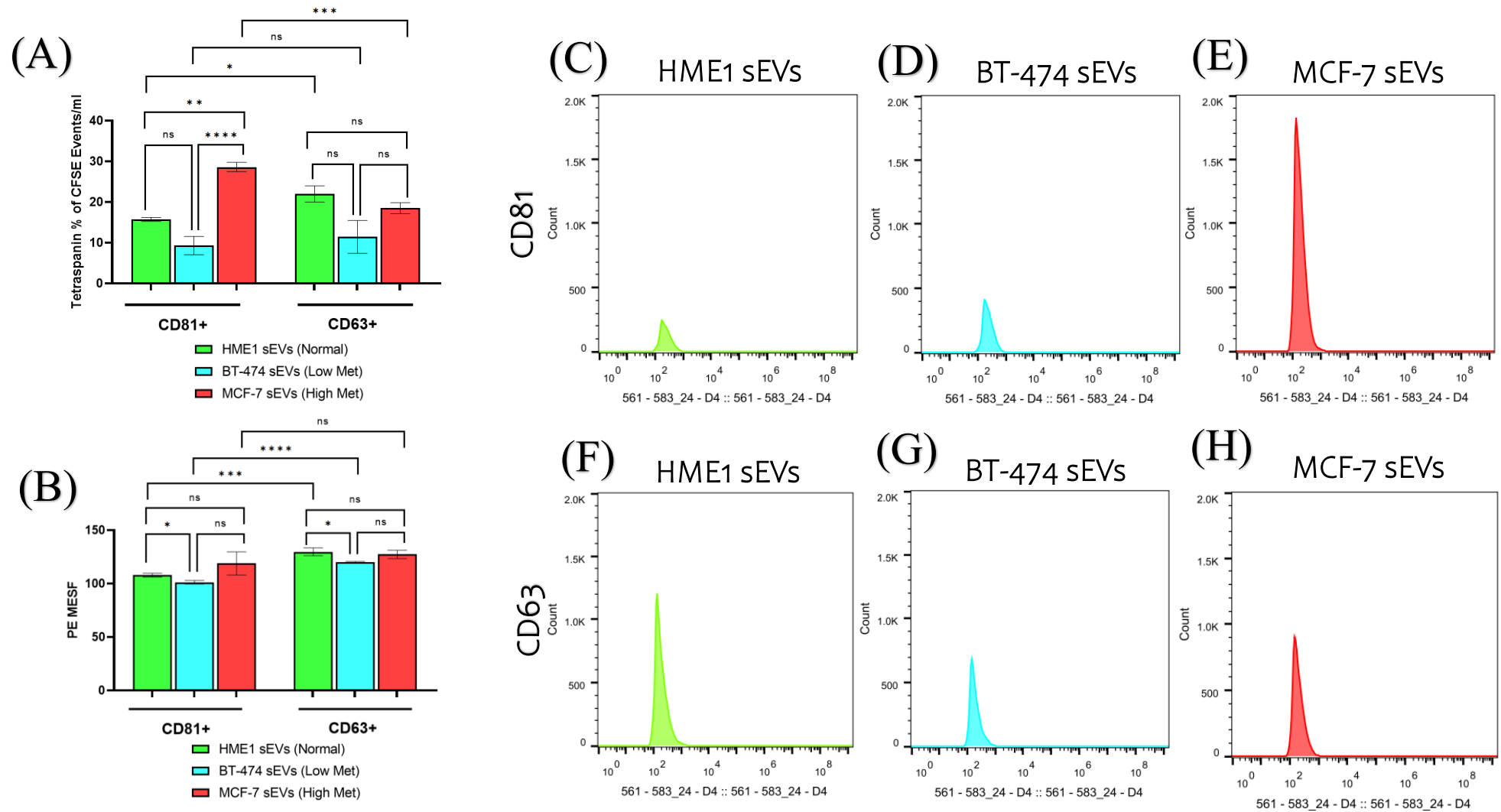
Figure 3.4. Single-vesicle flow cytometric analysis of CD81 and CD63 tetraspanin composition in breast sEVs. (A) *Quantification of CD81+ or CD63+ events/ml normalised to the percentage of CFSE events/ml in hTERT-HME1, BT-474, and MCF-7 sEVs.* **(B)** *PE MESF values of CD81+ or CD63+ for hTERT-HME1, BT-474, and MCF-7 sEVs with arbitrary units of MFI converted to standardised MESF units using vCal^{TM} nanorainbow beads.* **(C-E)** *Histogram profiles of breast sEVs showing CD81 fluorescence after gating.* **(F-H)** *Histogram profiles of breast sEVs showing CD63 fluorescence after gating. Each experiment consisted of biological triplicates with three technical triplicates, and error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. For parametric t-tests, Welch's correction was utilised, and for nonparametric t-tests, a Mann-Whitney test was used. Significance levels were set at *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.*

Figure 3.5. Single-vesicle flow cytometric analysis of CD81 and CD63 tetraspanin composition in colorectal sEVs. (A) *Quantification of CD81+ and CD63+ events/ml normalised to the percentage of CFSE events/ml for NCM460, SW480, and HT-29 sEVs.* **(B)** *PE MESF values of CD81+ or CD63+ for NCM460, SW480, and HT-29 sEVs with arbitrary units of MFI converted to standardised MESF units using vCal^{TM} nanorainbow beads.* **(C-E)** *Histogram profiles of colorectal sEVs showing CD81 fluorescence after gating.* **(F-H)** *Histogram profiles of colorectal sEVs showing CD63 fluorescence after gating. Each experiment consisted of biological triplicates with three technical triplicates, and error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. For parametric t-tests, Welch's correction was utilised, and for nonparametric t-tests, a Mann-Whitney test was used. Significance levels were set at *p < 0.05 and **p < 0.01.*

3.5. DISCUSSION

3.5.1. CHARACTERISATION OF sEVs DERIVED FROM BREAST AND COLORECTAL EPITHELIAL CELLS SHOWING TYPICAL MORPHOLOGICAL CHARACTERISTICS AND KEY MARKERS

To date, there is no standardised method for isolating and characterising sEVs (Théry et al., 2018). Ultracentrifugation, despite being the most commonly employed extraction technique, presents challenges such as contamination with non-vesicular macromolecules, potentially impacting '-omics' analyses and functional sEV studies (reviewed by Webber & Clayton, 2013). Therefore, this study opted to use SEC as the preferred method to reduce contamination by nonvesicular macromolecules. Analysis of the fractions using NTA and BCA suggested that fractions 5, 6, and 7 from the SEC separation were most likely to contain sEVs, as indicated by the highest particle counts and lowest protein contamination in both breast and colorectal epithelial-derived sEVs. These fractions were pooled for further characterisation.

NTA revealed that sEVs derived from epithelial cell lines fell within the established size range of 30-150 nm (reviewed by Doyle & Wang, 2019). Notably, a significant increase in size was observed for 'primary' and 'metastatic' sEVs from both breast and colorectal epithelial cell lines, averaging 105 nm, compared to 75 nm for 'normal' sEVs. This variation in size distribution could be partly attributed to the higher concentration of sEVs in cancer samples. It is well documented that cancer cells secrete a greater volume of EVs with altered composition than their non-malignant counterparts (reviewed by Bebelman et al., 2021; Riches et al., 2014). However, NTA cannot discriminate between vesicles and non-vesicular materials, and the presence of large crystals of salts and protein aggregates can potentially overestimate the true number of vesicles. In terms of morphological characteristics TEM analysis revealed 'cup-shaped' particles, within the size range of 30 nm-150 nm, presumed to be sEVs, from both breast and colorectal cell lines.

The isolated sEVs were further characterised by blotting for key marker proteins such as syntenin-1 protein which is highly abundant on sEVs and plays a fundamental role in sEVs biogenesis (section 1.5.2.1). Moreover, tetraspanins CD81 and CD63 were also characterised due to also their high abundance on sEVs (section 3.3.2.4). Notably, syntenin-1 was prominently detectable in both 'metastatic' MCF-7 and 'primary' BT-474 sEVs but less so in 'normal' hTERT-HME1 sEVs. This variation could be ascribed to the altered biogenesis and cargo selection mechanisms in cancer-derived sEVs (reviewed by Bebelman et al., 2018). Interestingly, a comprehensive analysis of 14 cell lines by Kugeratski et al. (2021) revealed that syntenin-1 was the most abundant protein on sEVs. In the current study, western blotting of 'normal' hTERT-HME1 and 'primary' BT-474 sEVs revealed pronounced bands for CD81 and smeared bands for CD63, confirming their presence. Notably, the CD63 band appeared as a broad smear, a characteristic attributed to its heavy glycosylation which leads to various glycoforms (Ageberg & Lindmark, 2003). In contrast, western blotting of MCF-7 sEVs revealed weaker bands for these markers, in contrast to the findings of González-King et al. (2022) who reported a distinct band for CD81 and a more prominent band for CD63. Li et al. (2018) identified highly visible bands of CD81 but less pronounced bands of CD63 in MCF-7 sEVs. The discrepancies between the band intensities of CD81 and CD63 in this study and others

may be attributed to variations in the isolation techniques used to obtain the sEVs (Brennan et al., 2020). In this study, SEC was used for sEVs isolation, whereas González-King et al. (2022) and Li et al. (2018) used ultracentrifugation. However, both studies used different speed, duration, and temperature parameters in their isolation protocols. Colorectal epithelial-derived sEVs derived from 'normal' NCM460, 'primary' SW480 and 'metastatic' HT-29 cells showed consistent intensities of all bands for CD81, CD63, and syntenin-1. Moreover, all breast and colorectal epithelial-derived sEVs samples were devoid of the mitochondrial marker cytochrome C, indicating an absence of cellular contamination. These findings corroborate the efficacy of the SEC method in isolating high-purity sEVs but emphasise the need to further investigate tetraspanin composition using more sophisticated methods at the single-vesicle level.

3.5.2. OPTIMISATION OF SINGLE-VESICLE FLOW CYTOMETRY IS ESSENTIAL FOR MAXIMUM EFFICIENCY AND PRECISION IN sEVs QUANTIFICATION

This study aimed to establish an effective staining and acquisition protocol for sEV detection by single-vesicle flow cytometry. For this purpose, CFDA-SE was selected as the staining agent for epithelial-derived sEVs because of its ability to stain sEVs without forming aggregates and to preserve their natural light-scattering properties, ensuring accurate detection (Morales-Kastresana et al., 2017). This choice is supported by previous research demonstrating the proficiency of CFDA-SE in identifying sEVs and serving as an initial pan-marker (Mastoridis et al., 2018). Some studies which have implemented CFDA-SE staining of sEVs have focused on removing excess dye during the isolation process using ultracentrifugation (Ender et al., 2019; Kormelink et al., 2016; Morales-Kastresana et al., 2017; Ragni et al., 2020). Whereas other studies have employed SEC after CFDA-SE staining (Maia et al., 2020; Morales-Kastresana et al., 2017; Welsh et al., 2021). In this study SEC was chosen to remove excess dye after CFDA-SE staining of sEVs to enable the use of sEVs in downstream applications beyond single-vesicle flow cytometry in which CFDA-SE staining wasn't required. SEC post-staining efficiently separates unbound CFDA-SE, which is prone to hydrolysis and artefact formation, from sEVs, leading to more accurate detection without interference from non-specific fluorescent signals (Morales-Kastresana et al., 2017). To evaluate the capability of SEC to remove excess dye after CFDA-SE staining of sEVs, tetraspanin CD81 was detected in conjunction with CFDA-SE. The results indicated a notable increase in the detection of 'CD81-positive' sEVs with the introduction of the SEC cleanup procedure compared to its absence. This confirmed the enhanced specificity of the combined staining and cleanup protocol for quantifying labelled sEVs.

Building on the foundational work by Görgens et al. (2019), who demonstrated the efficacy of enhanced green fluorescent protein-labelled sEVs and refined the optimum parameters for imaging flow cytometry, this study aimed to further optimise the acquisition settings for high-throughput capability. This involved fine-tuning the running time and ensuring sample integrity when implementing plate mode for analysis. The results indicated that a minimum duration of 5 min running time was essential to achieve the maximum signal from the sEVs sample, which was consistent with similar studies that used the same platform (Dar et al., 2021; Görgens et al., 2019; Li et al., 2023; van de Wakker et al., 2022). Whilst the CellStream instrument excels in high-throughput sEV analysis,

particularly with its plate mode enabling simultaneous measurement of multiple samples, there's a notable concern about potential sEV degradation whilst measuring over a period of time. Therefore, this study investigated stability under the specified conditions, focusing on the duration before sEV degradation affected experimental result outcome. Results showed that samples analysed beyond 67 minutes and 30 seconds displayed reduced integrity and weaker signals. This finding is crucial for adhering to MIFlowCyt-EV guidelines, which stresses comprehensive control for valid and reproducible EV-related experiments (Welsh et al., 2020). Understanding sEV stability in plate mode influences experimental design, particularly in determining the necessary number of control samples. Notably, several attempts have been made to enhance the stability of sEVs. For example, Görgens et al. (2022) noted that storing sEVs in PBS, as in this study, could result in decreased sample stability when assessed for particle concentration via NTA over a period of time. The research proposed alternative storage conditions, such as incorporating albumin and additional additives such as trehalose, to stabilise sEVs. While the study found no discernible differences in sEV shape, diameter, or integrity under varying storage buffer conditions, it reported a lower sEVs concentration, which hindered quantitative analysis. Given the intended downstream glycosylation assessment of sEVs in the current study and the absence of evidence suggesting any impact on glycan content under these conditions, the inclusion of these storage preservatives was avoided. Therefore, experiments utilising single-vesicle flow cytometry were meticulously designed to not only adhere to MIFlowCyt-EV guidelines, but also to account for potential sEV degradation over time when stored in PBS (Welsh et al., 2020). This study emphasises the importance of preserving sample integrity over time, irrespective of the analytical platform employed, to ensure the accuracy and reliability of sEV sample quality in high-throughput methodologies.

3.5.3. SINGLE-VESICLE ANALYSIS OF TETRASPANIN COMPOSITION REVEALS CD81-POSITIVE BREAST sEVs CORRELATES WITH METASTATIC PHENOTYPE

For further in-depth characterisation of the CD81 and CD63 tetraspanins in breast and colorectal epithelial-derived sEVs, single-vesicle flow cytometry was performed. Interestingly, MESF analysis revealed that a higher density of CD63 was present on the surface of colorectal epithelial sEVs of NCM460, SW480, and HT-29 in comparison to CD81. These observations were also mirrored for the breast epithelial sEVs hTERT-HME1 and BT-474. These findings suggest an increased abundance of CD63 relative to CD81 on the surface of these sEVs. Comparative analysis of the proportion of sEVs positive for each tetraspanin revealed a significantly higher abundance of 'CD81-positive' sEVs derived from MCF-7 cells than 'CD63-positive' sEVs. These findings align with the observations of Fan et al. (2023), who similarly noted increased CD81 enrichment through western blot analysis, suggesting that a potential membrane localisation of CD81 contributes to this phenomenon. However, contrasting results emerged in the tetraspanin characterisation between single-vesicle flow cytometry and western blotting in this study. This discrepancy likely stems from the differing treatment of sEVs, with western blotting involving lysis compared to intact analysis in single-vesicle flow cytometry. While single-vesicle analysis quantifies tetraspanins on the external membrane of intact sEVs, western blotting measures the total protein content, including intravesicular components. Moreover, the potential limitations of antibody binding efficiency and

specificity in western blot assays underscore the importance of cross validation across various analytical methods. One explanation for the varied abundance of these tetraspanin sEVs markers can be attributed to the diverse cellular phenotypes from which these sEVs originated, spanning from 'normal', 'primary' and 'metastatic' cells.

The abundance of breast 'metastatic' MCF-7 'CD81-positive' sEVs was notably higher than that of 'normal' hTERT-HME1 and 'primary' BT-474 'CD81-positive' sEVs, suggesting a correlation between the increased abundance of 'CD81-positive' sEVs and metastatic phenotype of breast sEVs. Interestingly, CD81 plays a significant role in addition to its presence on sEVs, as research has shown that CD81 is involved in regulating cell migration and invasion, suggesting a role in cancer progression (Vences-Catalán et al., 2017; Zhang et al., 2018). Recent studies have highlighted the key role of CD81 in promoting cancer stemness and metastasis by interacting with CD44, a cell-surface adhesion receptor which is highly expressed in many cancers (Ramos et al., 2022). It was also noted that 'CD63-positive' sEVs were elevated in 'normal' hTERT-HME1 and 'primary' BT-474 cells compared to 'CD81-positive' sEVs. This observation is intriguing, as CD63 is often associated with less aggressive cancer phenotypes and its expression tends to be inversely correlated with cancer metastasis (Berditchevski et al., 1995; Chirco et al., 2006; Jung et al., 2006; Pols & Klumperman, 2009; Radford et al., 1997). Collectively, these findings highlight the complexity of sEVs tetraspanin subpopulations across various cell types, emphasising the need for further investigation with a broader range of breast and colorectal epithelial cell-derived sEVs.

3.6. KEY FINDINGS

- Breast and colorectal 'primary' and 'metastatic' sEVs extracted by SEC showed increased sizes compared to 'normal' sEVs.
- Cleanup by SEC of excess CFDA-SE staining of sEVs is essential to obtain maximum signal of antibody of interest.
- A minimum of 5 minutes acquisition time is required for plate mode on the CellStream instrument for operating single-vesicle analysis of sEVs samples.
- Sample acquisition of sEVs by single-vesicle analysis past the 67-minute and 30-second mark exhibited a decline in integrity results in a diminished signal of interest.
- Increased 'CD81-positive' sEVs abundance in 'metastatic' MCF-7 cells in comparison to 'primary' BT-474 and 'normal' hTERT-HME1 sEVs.
- There was a higher abundance of CD63 than CD81 on the surface of colorectal epithelial sEVs from NCM460, SW480, and HT-29 cells.

CHAPTER 4

HPA BINDING sEVs AND THEIR FUNCTIONAL

ROLE IN CELL MOTILITY

4. HPA BINDING sEVs AND THEIR FUNCTIONAL ROLE IN CELL MOTILITY

4.1. BACKGROUND

In cancer, truncated O-linked glycans are frequently detected and serve as prominent markers of malignancy (reviewed by Stowell et al., 2015; Varki et al., 2022). HPA, a lectin of particular interest, has been extensively studied in the context of the altered glycosylation patterns associated with cancer metastasis (section 1.4.3). Although sEVs are known to play a significant role in cancer development, a comprehensive understanding of their overall glycosylation profile is often overlooked. Moreover, the exact biological functional role of the glycosylation profile of sEVs is limited, particularly in facilitating key pro-metastatic characteristics such as motility and adhesion changes. In this chapter, epithelial model cell lines were initially characterised for HPA-lectin binding using confocal microscopy and flow cytometry. Subsequently, efforts shifted to characterising HPA lectin binding of sEVs derived from these epithelial model cell lines was assessed, specifically assessing sEVs tetraspanin-positive CD81 and CD63 sEVs subpopulations. To further understand the functional involvement of sEVs taken up by recipient cells in acquiring pro-metastatic capabilities, wound healing and static adhesion assays were conducted. The potential involvement of HPA-recognising glycans on sEVs on cell motility was then explored in the wound healing assays by masking these glycans.

4.1.1. CANCER-ASSOCIATED sEVs AND TRUNCATED O-GalNAc (MUCIN-TYPE) GLYCANS

Despite the extensive literature documenting abnormal mucin-type O-glycosylation in breast and colorectal cancer cells through HPA lectin binding (section 1.4.3), the glycosylation patterns of their derived sEVs remain largely unexplored. Few studies have suggested the presence of mucin-type O-glycans on sEVs from other cancer types, indicating a potential role in cancer pathogenesis (Dusoswa et al., 2019; Feng et al., 2018). In this study, breast and colorectal epithelial cell lines representing 'normal', 'primary', and 'metastatic' cancer were qualitatively and quantitatively assessed for HPA lectin binding using confocal microscopy and flow cytometry. Subsequently, sEVs derived from these cell lines underwent HPA lectin labelling using single-vesicle flow cytometry, specifically targeting sEVs tetraspanin-positive CD81 and CD63 subpopulations.

4.1.2. EVALUATING THE FUNCTIONAL ROLE OF sEVS IN THE METASTATIC CASCADE *in vitro*

Metastasis is a hallmark of tumour malignancy and is responsible for most cancer-related fatalities; however, the underlying mechanisms remain poorly understood (reviewed by Fares et al., 2020). Over the last two decades, research has revealed the significant involvement of sEVs in facilitating metastasis by mediating intercellular communication (reviewed by Becker et al., 2016; Li et al., 2019). To gain insight into the role of sEVs in cancer metastasis, researchers have implemented *in vitro* experimental assays that mimic pathological events during the metastasis cascade offering a valuable approach. For example, O'Brien et al. (2013) conducted a wound healing

assay to evaluate the impact of sEV uptake from aggressive breast cancer cells on recipient, less aggressive breast cancer cells, revealing enhanced cell motility in recipient cells. The wound healing assay involves creating an artificial 'wound' by scratching a monolayer of cells and subsequently measuring the migration of cells over the 'wound' area over a period. Another example of an *in vitro* experimental assay is the static adhesion assay, which mimics a key stage in the metastatic cascade where metastasising cells reach the endothelial walls, adhere to them, and eventually intravasate and extravasate to form distant tumour foci (Kramer & Nicolson, 1979). The assay involves culturing endothelial cells in a monolayer, pre-activating them with pro-inflammatory cytokines, introducing fluorescently labelled cancer cells into the endothelial monolayer, and then removing non-bound cells before fixing and counting the remaining bound cells. In this study, both assays were used to test the effect of recipient cells treated with sEVs on cellular motility and the ability to adhere to the endothelium. To specifically investigate the impact of GalNAc glycans on sEVs to modulate cell motility characteristics on recipient cells, these glycans as recognised by HPA lectin were masked on sEVs, treated on recipient cells and evaluated by the wound healing assay.

4.2. AIMS AND OBJECTIVES

The experiments detailed in this chapter aimed to characterise the O-linked GalNAc glycosylation of breast and colorectal cells by HPA lectin binding. In addition, the experiments aimed to determine the impact of sEVs and their GalNAc glycans in modulating motility cell changes on treated recipient cells. Furthermore, the study also evaluated the impact of sEVs treated on recipient cells on inducing changes for their capability to adhere to endothelial monolayers.

The objectives were:

1) To assess HPA lectin binding to GalNAc-glycans on the surface of breast and colorectal epithelial cell lines derived from 'normal', 'primary' and 'metastatic' cancer cell phenotypes using confocal microscopy and flow cytometry.

2) To quantify the abundance of HPA-binding GalNAc glycans on the surface of 'CD81-positive' and 'CD63-positive' sEVs derived from 'normal'. 'primary' and 'metastatic' cancer cell phenotypes by HPA lectin binding and single-vesicle flow cytometry.

3) The identify the effects of sEVs derived from 'normal' and 'metastatic' cancer-associated cell phenotypes on inducing motility and adhesion changes in treated recipient cells were evaluated by wound-healing assay and static adhesion assay.

4) To identify the effect of HPA lectin recognising glycans on sEVs derived from 'normal' and 'metastatic' cancer-associated cell phenotypes on motility in treated recipient cells in a wound-healing assay.

4.3. METHODS

4.3.1. HPA CHARACTERISATION OF EPITHELIAL CELL LINE MODELS

4.3.1.1. HPA LABELLING AND ASSESSMENT USING CONFOCAL MICROSCOPY

To quantify HPA lectin binding and qualitatively confirm cell surface binding in 'normal', 'primary', and 'metastatic' cancer cell phenotypes of breast and colorectal cancer, confocal microscopy was utilised, following the methodology outlined in section 2.3.

4.3.1.2. FLOW CYTOMETRY ANALYSIS OF HPA LECTIN LABELLING ON EPITHELIAL CELL SURFACES

To quantify HPA lectin binding on the cell surface of breast and colorectal epithelial cell models, flow cytometry was employed following the methodology described in section 2.2.

4.3.2. SINGLE-VESICLE FLOW CYTOMETRY

4.3.2.1. HPA LECTIN PROFILING OF EPITHELIAL DERIVED sEVs

To quantify the HPA-lectin binding on epithelial-derived sEVs of both breast and colorectal cells, single-vesicle flow cytometry was performed. The methodology for extracting sEVs from their parental cells is detailed in section 2.4.1. The single-vesicle flow cytometry methodology is outlined in section 2.6. including the instrument settings (section 2.6.1), calibration of sizing and fluorescence in preparation for acquisition of sEVs (section 2.6.2), and sEVs lipid staining and detection antibody labelling of CD81 and CD63, followed by HPA lectin labelling is outlined in section 2.6.3.

4.3.3. WOUND HEALING ASSAY

4.3.3.1. HPA LECTIN PROFILING OF EPITHELIAL DERIVED sEVs ASSESSMENT OF sEVs TREATMENT ON CELL MOTILITY

A wound-healing assay was conducted to investigate the functional involvement of sEVs derived from breast and colorectal epithelial cell lines on cell motility. First, sEVs were extracted from the parental cells of both breast hTERT-HME1 and MCF-7 cells and colorectal NCM460 and HT-29 cells, as outlined in section 2.4.1. The details of the experimental procedure of the wound healing assay, including the preparation of plates, seeding of cells, wound creation, and appropriate control of image acquisition and processing, are detailed in section 2.7.2. The details of the treatment of cells with sEVs are detailed in section 2.7.1.

4.3.3.2. ASSESSMENT OF sEVs HPA LECTIN RECOGNISING GLYCANS TREATMENT ON CELL MOTILITY

A wound healing assay was conducted to investigate the functional involvement of HPA lectin recognising glycans on breast epithelial derived sEVs on recipient cells motility. Firstly, sEVs were extracted from breast hTERT-HME1 and MCF-7 cells and colorectal NCM460 and HT-29 cells as outlined in section 2.4.1. Breast epithelial sEVs GalNAc-glycans were then blocked by incubating with HPA (section 2.6.3). This was followed by purification using NAP-5 SEC columns (GE Healthcare) to eliminate any unbound HPA. The HPA-bound sEVs and background controls, with no sEVs or HPA treatment, were run concurrently to account for any functional effects observed via non-bound HPA. The methodology of the wound healing assay, including preparation of plates, seeding of cells, wound creation, appropriate controls, and image acquisition and processing, are detailed in section 2.7.2. The treatment of cells with sEVs labelled with HPA is described in section 2.7.1

4.3.4. STATIC ADHESION ASSAY

4.3.4.1. ASSESSMENT OF sEVs TREATMENT ON CELL ADHESION

To investigate the effect of sEVs derived from 'normal' and 'metastatic' breast and colorectal phenotypes to induce adhesion changes in recipient cells to endothelial cell monolayers, a static adhesion assay was conducted. The experimental outline, including the preparation of plates, seeding of endothelial cells, appropriate controls, and image acquisition and processing, is detailed in section 2.7.3. The methodology for treating cells with sEVs is outlined in section 2.7.1

4.4. RESULTS

4.4.1. CELL SURFACE LOCALISATION OF HPA BINDING IN BREAST AND COLORECTAL MODEL CELL LINES

To identify the localisation of HPA lectin binding on the model cell lines, lectin cytochemistry and confocal microscopy Z-stack analysis were adopted, and results are illustrated in Figure 4.1. Notably, an initial examination of the Z-stack images revealed that the HPA lectin recognised glycans on the cell surface rather than in an intracellular localisation (Figure 4.1-D). Controls were included, as described in section 2.3. Analysis of all cell lines showed a significant increase in the mean fluorescence for HPA binding compared to the controls, confirming the presence of GalNAc-glycans. Analysis of the breast cell lines revealed that MCF-7 cells exhibited higher levels of GalNAc-glycans compared to BT-474 and hTERT-HME1 breast cells, with average fold increases of 1.7 and 2.7, respectively ($p < 0.0001$) (Figure 4.1-C). Similarly, among the colorectal cell lines, HT-29 cells exhibited the highest levels of HPA binding, with an average 2.9-fold and 2.1-fold increase in normalised fluorescence compared with SW480 and NCM460 cells, respectively ($p < 0.0001$) (Figure 4.2-C).
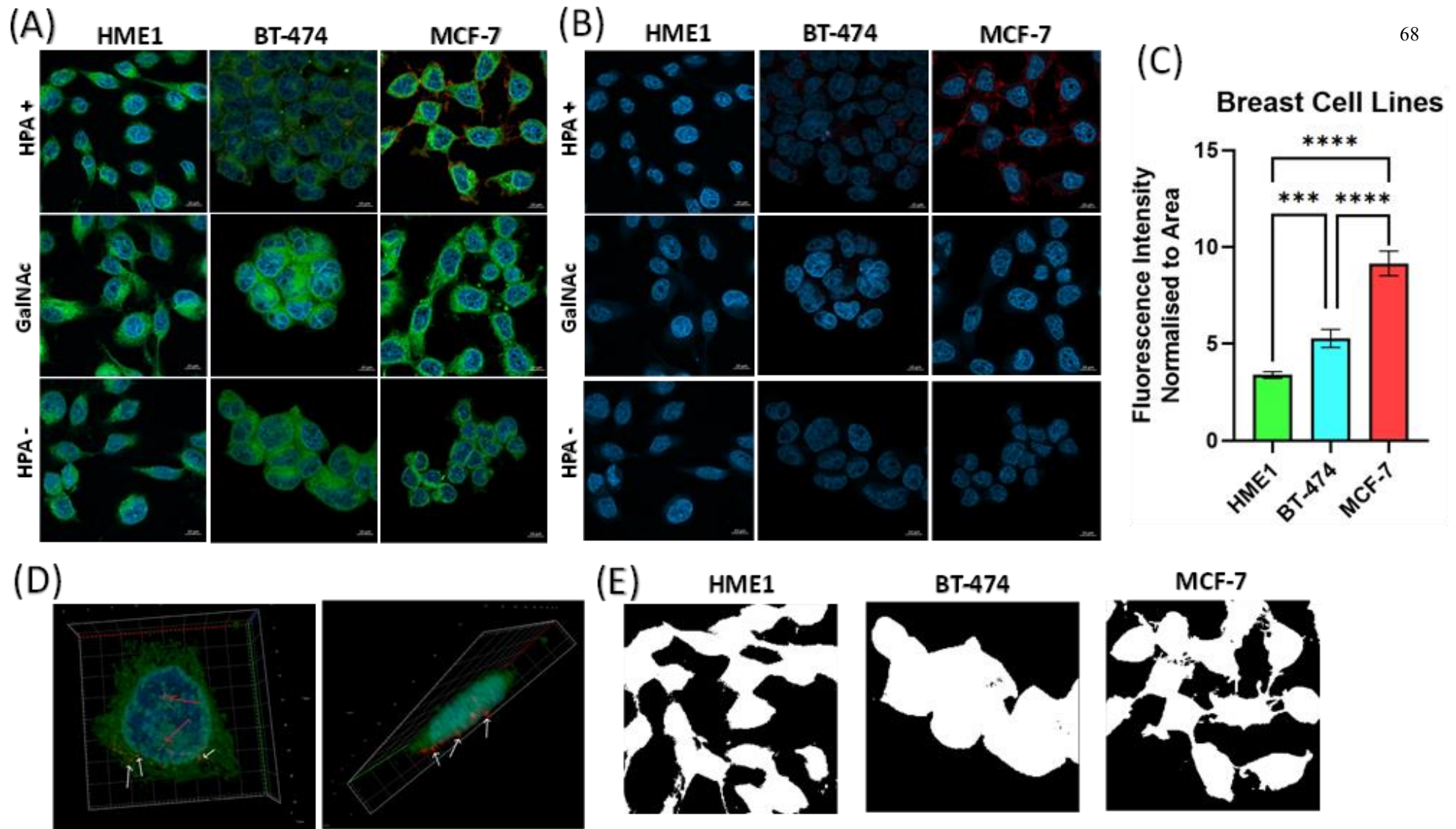
Figure 4.1. HPA glycosylation profiling of breast epithelial cell lines by confocal microscopy. (A) *Breast epithelial cell lines (hTERT-HME1, BT-474, and MCF-7) were probed with HPA lectin (red), CellMask*^{TM}
*to label the plasma membrane (green), and DAPI to label the nuclei (blue). The images show the HPA negative control (cells without HPA lectin), GalNAc inhibitor control (cells with HPA and sugar-specific control GalNAc), and HPA positive control (cells with HPA to identify the presence of GalNAc glycans). Scale bars = 10 µm. **(B)** The same images as in panel A, without plasma membrane overlay. **(C)** Quantification of fluorescence intensity normalised to the area occupied by the plasma membrane (green) for hTERT-HME1, BT-474, and MCF-7 cells. **(D)** An example of a Z-stacking image of a cell highlighting HPA lectin binding on the cell surface (red arrows indicate what appears to be internal labelling, but when the image is flipped, they appear to be external, represented by the white arrows). **(E)** Huang's thresholding method was used to highlight the plasma membrane region of breast cells. Each experiment was performed in triplicate, with error bars representing the Stdev of the data. A t-test followed by Welch's correction was used for statistical analysis, and significance was set at *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.*
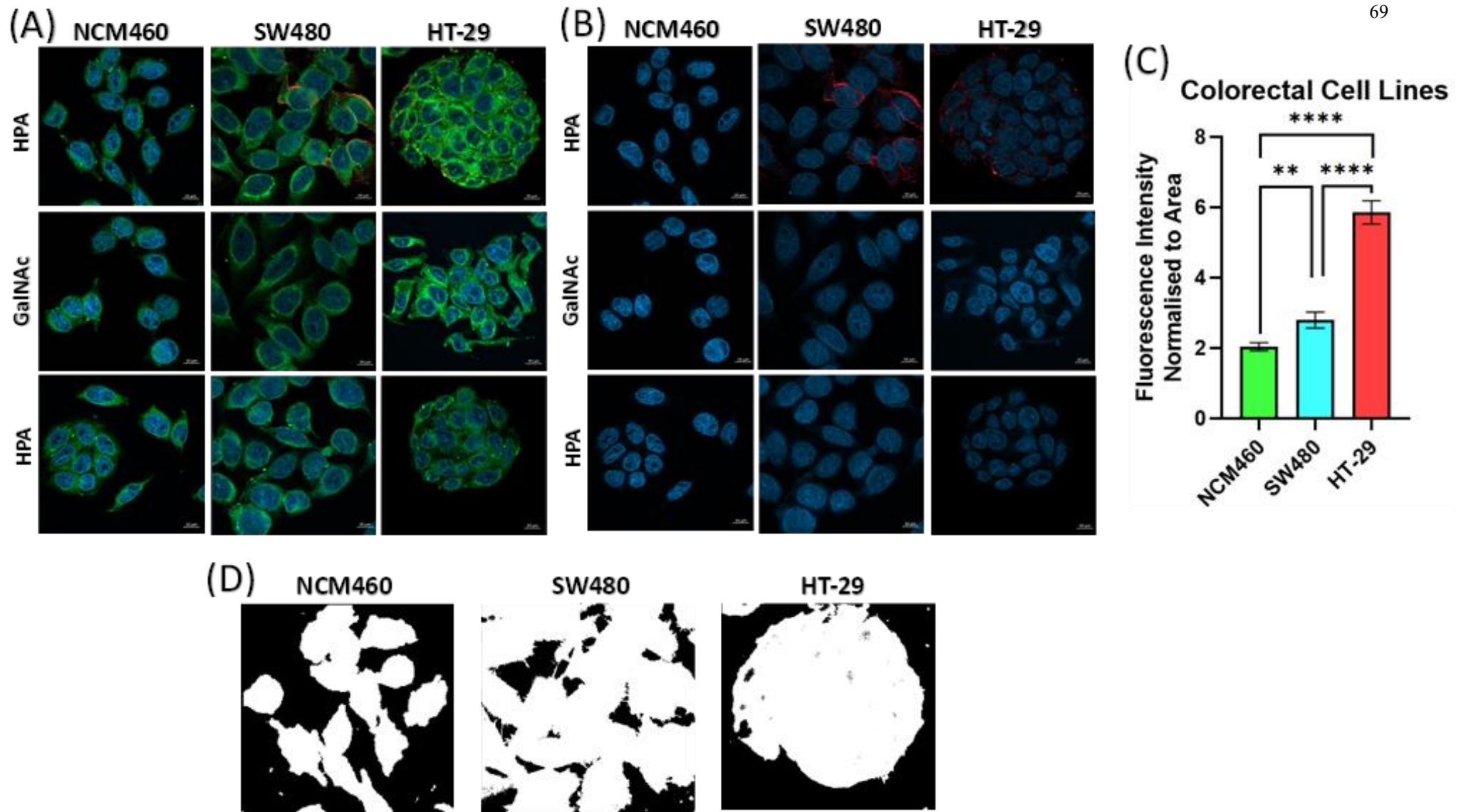
Figure 4.2. HPA glycosylation profiling of colorectal epithelial cell lines by confocal microscopy. (A) *Colorectal epithelial cell lines (NCM460, SW480, and HT-29) were probed with HPA lectin (red), CellMask*[TM] *to label the plasma membrane (green), and DAPI to label the nuclei (blue). The images show HPA negative control (cells without HPA lectin), GalNAc inhibitor control (cells with HPA and sugar-specific control GalNAc), and HPA positive control (cells with HPA to identify the presence of GalNAc glycans). Scale bars = 10 µm.* **(B)** *The same images as in panel A, without the plasma membrane overlay.* **(C)** *Quantification of fluorescence intensity normalised to the area occupied by the plasma membrane (green) for NCM460, SW480, and HT-29 cells.* **(D)** *Huang's thresholding method was used to highlight the plasma membrane region of breast cells. Each experiment was performed with biological triplicates with three technical triplicates with error bars representing the Stdev of the data. T-test followed by Welch's correction was used for statistical analysis and significance were set at *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.*

4.4.2. FLOW CYTOMETRY ANALYSIS OF HPA LECTIN BINDING OF BREAST AND COLORECTAL EPITHELIAL CELLS

To quantify the binding profile of HPA lectin to human breast (hTERT-HME1, BT-474, and MCF-7) and colorectal (NCM460, SW480 and HT-29) cell lines, flow cytometry was performed and results are presented in Figure 4.3. Competitive inhibition of HPA binding in the presence of GalNAc was used as a lectin specificity control, and an HPA negative control (omission of lectin incubation) was employed to accurately identify positive fluorescence signals. The resulting histograms, depicted in Figure 4.3 A-F, revealed a significant increase in HPA-binding fluorescence in all cell lines compared with both controls, indicating the presence of GalNAc-glycans. Interestingly, MCF-7 breast cancer cells exhibited an MFI average fold increase of 2 and 11 relative to BT-474 and hTERT-HME1 cells, respectively ($p < 0.0001$) (Figure 4.3-G). In the colorectal cell line HT-29, the average MFI-fold increases were 2 and 5 relative to SW480 ($p<0.001$) and NCM460 ($p < 0.0001$), respectively (Figure 4.3-H).
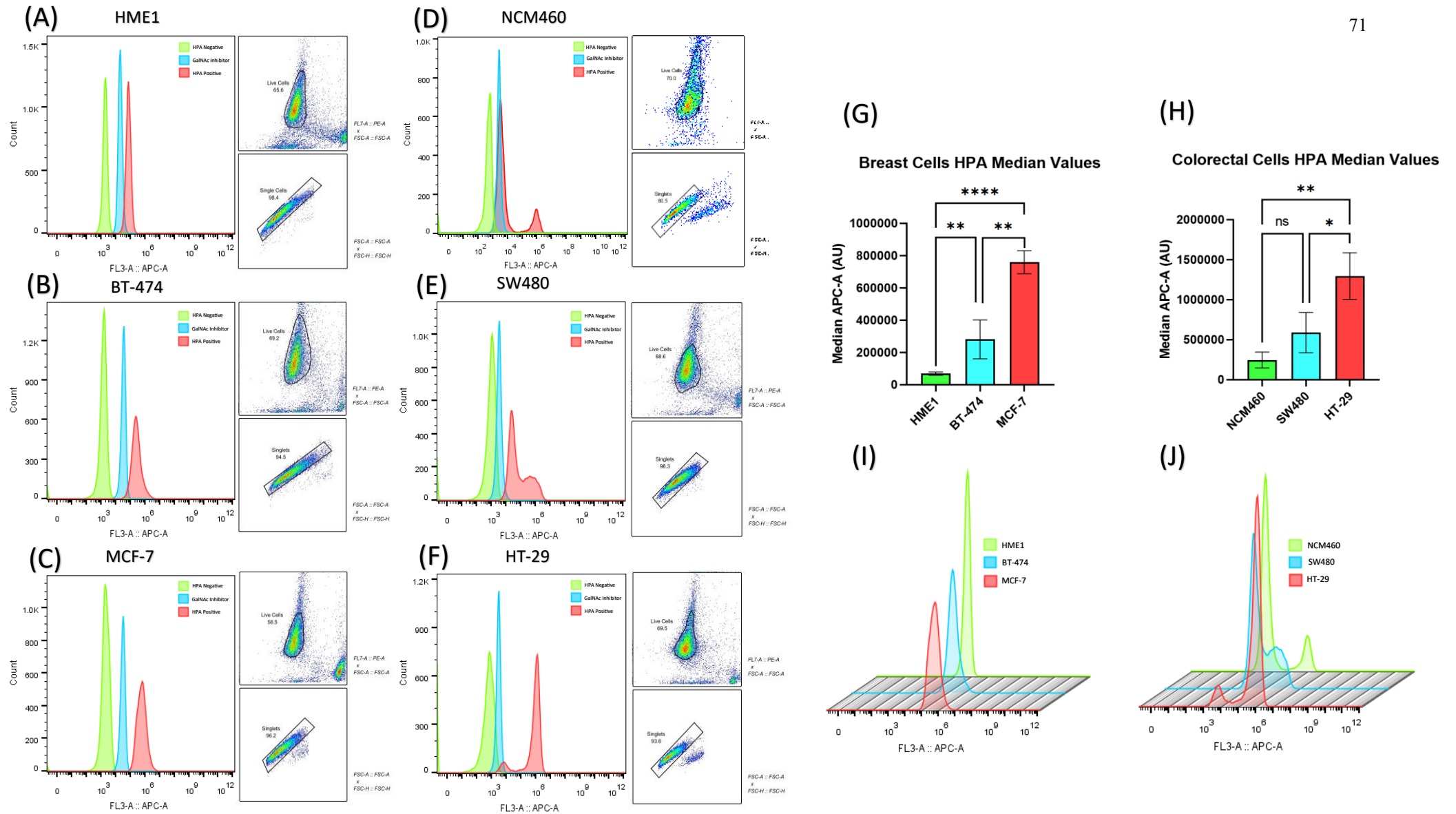
Figure 4.3. HPA glycosylation profiling on epithelial cell lines by flow cytometry. (A-C) *Histogram profiles of breast cell lines showing HPA binding alongside ancestral gating of live and single-cell exclusion.* **(D-F)** *Colorectal cell lines.* **(G)** *Median fluorescence (AU) of breast cancer cell lines with the HPA negative control (cells without HPA lectin), GalNAc inhibitor control (cells with HPA and sugar-specific control GalNAc), and HPA positive control (cells with HPA to identify the presence of GalNAc glycans)* **(H)** *Median fluorescence (AU) of colorectal cell lines with the HPA negative control, GalNAc inhibitor control, and HPA positive control.* **(I)** *Staggered histogram comparison of breast cancer cell lines for HPA positivity.* **(J)** *Staggered histogram comparison of HPA positivity in colorectal cell lines. Each experiment was performed in triplicate, with error bars representing the Stdev of the data. T-test followed by Welch's correction was used for statistical analysis and significance were set at $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, and $****p < 0.0001$.*

4.4.3. TETRASPANIN SUBPOPULATION ANALYSIS OF BREAST AND COLORECTAL EPITHELIAL sEVs WITH FLUORESCENTLY LABELLED HPA

This study focused on analysing HPA lectin recognising glycans glycans on sEVs by refining the single-vesicle flow cytometry analytics described in section 2.6. This analysis employed separate CD81 and CD63 antibody detections, each paired with HPA lectin. To reveal distinct glycosylation patterns associated with sEVs positive for each tetraspanin, providing deeper insights into GalNAc glycan enrichment of sEVs from breast and colorectal epithelial cell lines. The raw results, including the mean ± SD of HPA binding via CD81 or CD63 detection normalised to CFSE events/ml and MESF values, are shown in Table 4.1.

The binding of HPA lectin to breast and colorectal 'CD81-positive' and 'CD63-positive' sEVs exhibited a notable increase in fluorescent intensity compared to the lectin sugar specific control samples. In these control samples, GalNAc was utilised to inhibit the binding of HPA prior to labelling, confirming, through competitive inhibition, the recognition of GalNAc glycans on the sEVs by HPA (Figure 4.4 and 4.5 C-E, F-G). Analysis of breast epithelial-derived sEVs revealed varying HPA-binding profiles, which were observed in association with their tetraspanin composition (Figure 4.4-A). In MCF-7 sEVs, a pronounced increase in HPA binding was observed in CD81-postive sEVs showing a 1.4-fold higher level than that of 'CD63-positive' sEVs ($p < 0.05$). For hTERT-HME1 and BT-474 sEVs, there were no significant differences in HPA lectin binding between the 'CD81-positive' and CD63-postive sEVs. This indicates a uniform pattern of HPA-lectin interactions across these tetraspanins. Interestingly, comparisons between the different phenotypes of breast sEVs HPA lectin binding were observed for 'CD81-positive' sEVs in 'metastatic' cancer-associated MCF-7 sEVs, which showed a significant increase of 1.5-fold higher than that in 'normal' hTERT-HME1 sEVs, and 2.5-fold higher than that in primary cancer-associated BT-474 sEVs ($p < 0.01$). However, no differences in HPA lectin binding were observed between BT-474-and HME 'CD81-positive' sEVs, reflecting similar findings in 'CD63-positive' sEVs for BT-474 and MCF-7 sEVs. Similarly, no variations were observed in the MESF values across the different breast sEV phenotypes (Figure 4.4-B).

The same comparisons were made for colorectal epithelial-derived sEVs, although the only significant differences were observed in the MESF values (Figure 4.5 A-B). Comparisons of the HPA MESF values of 'CD81-positive' and 'CD63-positive' sEVs revealed an increase of 1.2- in HT-29 compared to NCM460 sEVs ($p < 0.01$, CD81; $p < 0.0001$ for CD63) and SW480 sEVs ($p < 0.001$ for CD81 and $p < 0.01$ for CD63) (Figure 4.5-B).

73

Table 4.1. HPA lectin binding quantity (MESF) or percentage of total CFSE positive EVs split into CD81-positive and CD63-positive sEVs from breast and colorectal epithelial cells by single-vesicle flow cytometry.

Organ of Origin	Cell Line sEVs	CD81-positive sEVs HPA MESF (mean ± SD)	CD63-positive sEVs HPA MESF (mean ± SD)	CD81-positive sEVs HPA events/ml % of CFSE events/ml (mean ± SD)	CD63-positive sEVs HPA events/ml % of CFSE events/ml (mean ± SD)	CD81-positive vs CD63-positive sEVs HPA	Vs HME/NCM460 sEVs	Vs MCF-7/HT-29 sEVs
Breast	hTERT-HME1 sEVs	56.62 ± 10.88	49.86 ± 7.19	10.84 ± 1.00	14.52 ± 3.91	**Events/ml** NS **MESF** NS		**Events/ml** CD81= ** CD63= ** **MESF** CD81= NS CD63= NS
	BT-474 sEVs	52.21 ± 13.56	43.14 ± 17.16	5.85 ± 3.41	6.79 ± 6.36	**Events/ml** NS **MESF** NS	**Events/ml** CD81= NS CD63= NS **MESF** CD81= NS CD63= NS	**Events/ml** CD81= ** CD63= NS **MESF** CD81= NS CD63= NS
	MCF-7 sEVs	49.44 ± 6.15	47.06 ± 1.64	16.82 ± 2.88	11.34 ± 3.28	**Events/ml** * **MESF** NS	**Events/ml** CD81= ** CD63= ** **MESF** CD81= NS CD63= NS	
Colorectal	NCM460 sEVs	52.09 ± 6.36	53.15 ± 1.05	10.85 ± 6.30	11.46 ± 6.23	**Events/ml** NS **MESF** NS		**Events/ml** CD81= NS CD63= NS **MESF** CD81= ** CD63= ****
	SW480 sEVs	50.88 ± 3.09	52.44 ± 6.16	13.05 ± 6.06	13.43 ± 8.47	**Events/ml** NS **MESF** NS	**Events/ml** CD81= NS CD63= NS **MESF** CD81= NS CD63= NS	**Events/ml** CD81= NS CD63= NS **MESF** CD81= *** CD63= **
	HT-29 sEVs	61.48 ± 3.48	65.06 ± 2.95	15.27 ± 3.20	17.44 ± 6.82	**Events/ml** NS **MESF** NS	**Events/ml** CD81= NS CD63= NS **MESF** CD81= ** CD63= ****	

The asterisks (*) in the table denote significance levels.

The NS in the table indicates no significance.
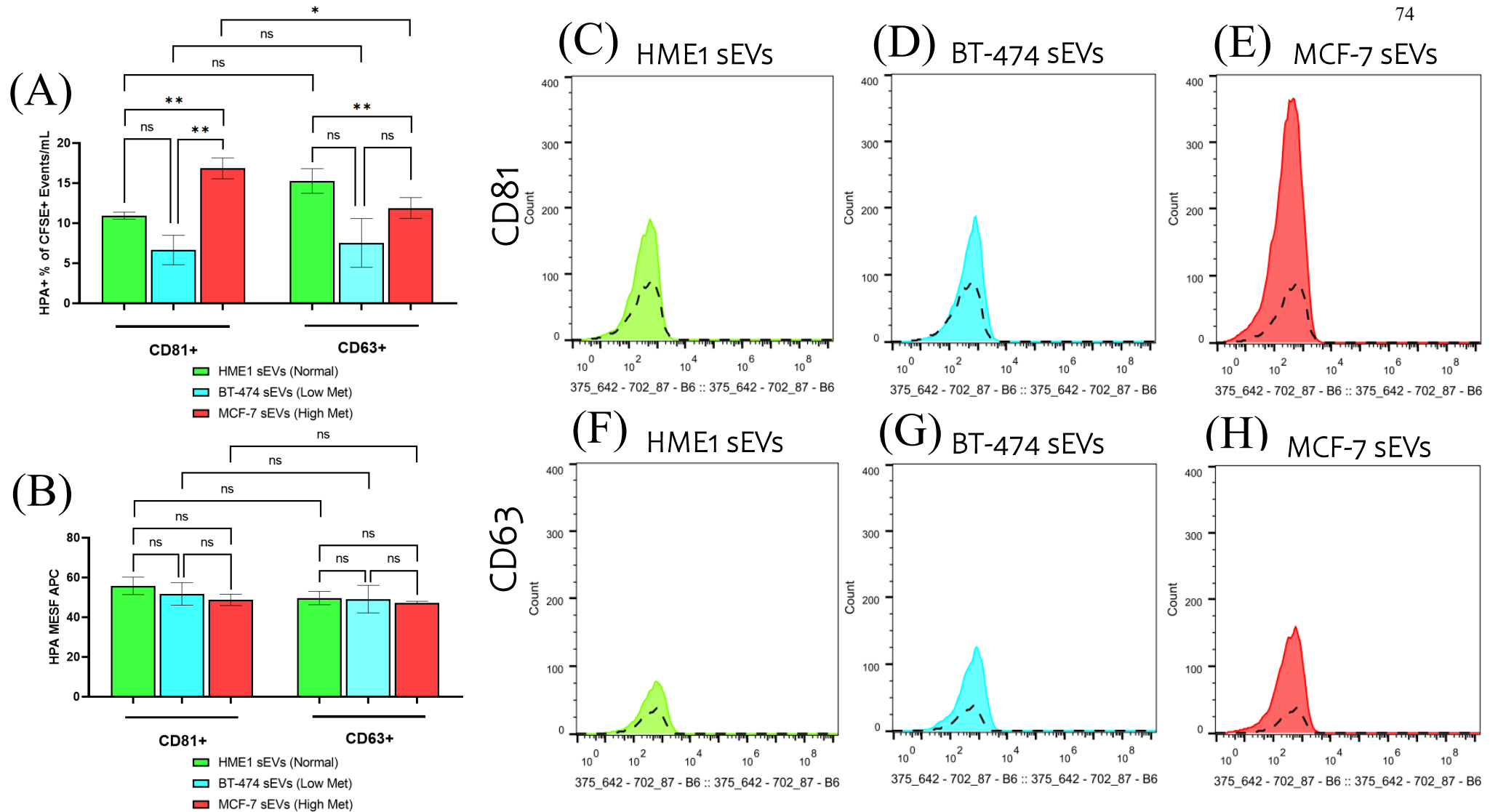
Figure 4.4. Single-vesicle flow cytometry analysis of HPA lectin binding to CD81-positive or CD63-positive breast sEVs. (A) *Quantification of HPA+ events/ml normalised to the percentage of CFSE events/ml for hTERT-HME1, BT-474, and MCF-7 sEVs which are either CD81-postive or CD63-positive.* **(B)** *APC MESF values of HPA+ for hTERT-HME1, BT-474, and MCF-7 sEVs with arbitrary units of MFI converted to standardised MESF units using vCal^{TM} nanorainbow beads.* **(C-E)** *Histogram profiles of the breast sEVs after gating showing HPA lectin binding of CD81-positive sEVs fluorescence alongside the sugar-specific control GalNAc outlined in black.* **(F-H)** *Histogram profiles of the breast sEVs after gating showing HPA lectin binding of CD63-positive sEVs fluorescence alongside the sugar-specific control GalNAc outlined in black. Each experiment consisted of biological triplicates with three technical triplicates, and error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. For parametric t-tests, Welch's correction was utilised, and for nonparametric t-tests, a Mann-Whitney test was used. Significance levels were set at *p < 0.05 and **p < 0.01.*

Figure 4.5. Single-vesicle flow cytometry analysis of HPA lectin binding to CD81-positive or CD63-positive colorectal sEVs. *(A) Quantification of HPA+ events/ml normalised to the percentage of CFSE events/ml for NCM460, SW480, and HT-29 sEVs which are either CD81-postive or CD63-positive. (B) APC MESF values of HPA+ for NCM460, SW480, and HT-29 sEVs with arbitrary units of MFI converted to standardised MESF units using vCal^{TM} nanorainbow beads. (C-E) Histogram profiles of the colorectal sEVs after gating showing HPA lectin binding of CD81-positive sEVs fluorescence alongside the sugar-specific control GalNAc outlined in black. (F-H) Histogram profiles of the colorectal sEVs after gating showing HPA lectin binding of CD63-positive sEVs fluorescence alongside the sugar-specific control GalNAc outlined in black. Each experiment consisted of biological triplicates with three technical triplicates, and error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. For parametric t-tests, Welch's correction was utilised, and for nonparametric t-tests, a Mann-Whitney test was used. Significance levels were set at *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.*

4.4.4. THE EFFECT OF sEVs TREATMENT ON RECIPIENT CELL MOTILITY

To investigate the potential functional involvement of sEVs derived from breast and colorectal epithelial cell line in cell motility, a wound healing assay was conducted. In this study, the assay was used to investigate the potential functional involvement of sEVs treated on recipient cell motility.

After the 24-hour time point, which corresponded to the peak of the observed gap closure in breast cells, colorectal exhibited minimal gap closure, indicating slower motility. hTERT-HME1 cells treatment with 1×10^8 particles/ml of MCF-7 sEVs induced motility by increasing gap closure by a fold difference of 4.8 times compared to hTERT-HME1 sEVs at the same density (p < 0.01) (Figure 4.6-A). In contrast, treatment of MCF-7 cells with their respective sEVs was observed to induce an inhibitory effect on motility in comparison to the no-treatment group for the different densities, with a fold difference of 2.8 at 1×10^9 particles/ml (p < 0.001), 2.5 at 1×10^8 particles/ml (p < 0.01), and 2.6 at 1×10^7 particles/ml (p < 0.01) (Figure 4.7-A). Conversely, MCF-7 cells treated with hTERT-HME1 sEVs at varying densities did not exhibit any significant differences, suggesting a consistent effect, irrespective of the concentration used. In addition, MCF-7 sEVs showed a greater inhibitory impact on cell motility, marked by fold differences of 2.9 at 1×10^9 particles/ml (p < 0.001), 3.1 at 1×10^8 particles/ml (p < 0.001), and 3.7 at 1×10^7 particles/ml (p < 0.001) compared to hTERT-HME1 sEVs treatment at the same densities. The same comparisons were also made for sEVs treatment on the colorectal cell lines; however, no statistically significant differences were revealed among the various sEVs densities and treatments for both NCM460 and HT-29 cells (Supplementary Figure 6 and 7).

Because of the observed motility changes triggered by breast sEVs, HPA lectin was utilised to target and obstruct HPA-binding glycans on the surface of sEVs before treatment of recipient cells. To account for any residual lectin not removed during cleanup, an additional sample, labelled as the 'background control', was processed with HPA and no sEVs, alongside the no-HPA and sEVs-treated samples. However, no significant differences were detected (Figure 4.6-B and 4.7-B).
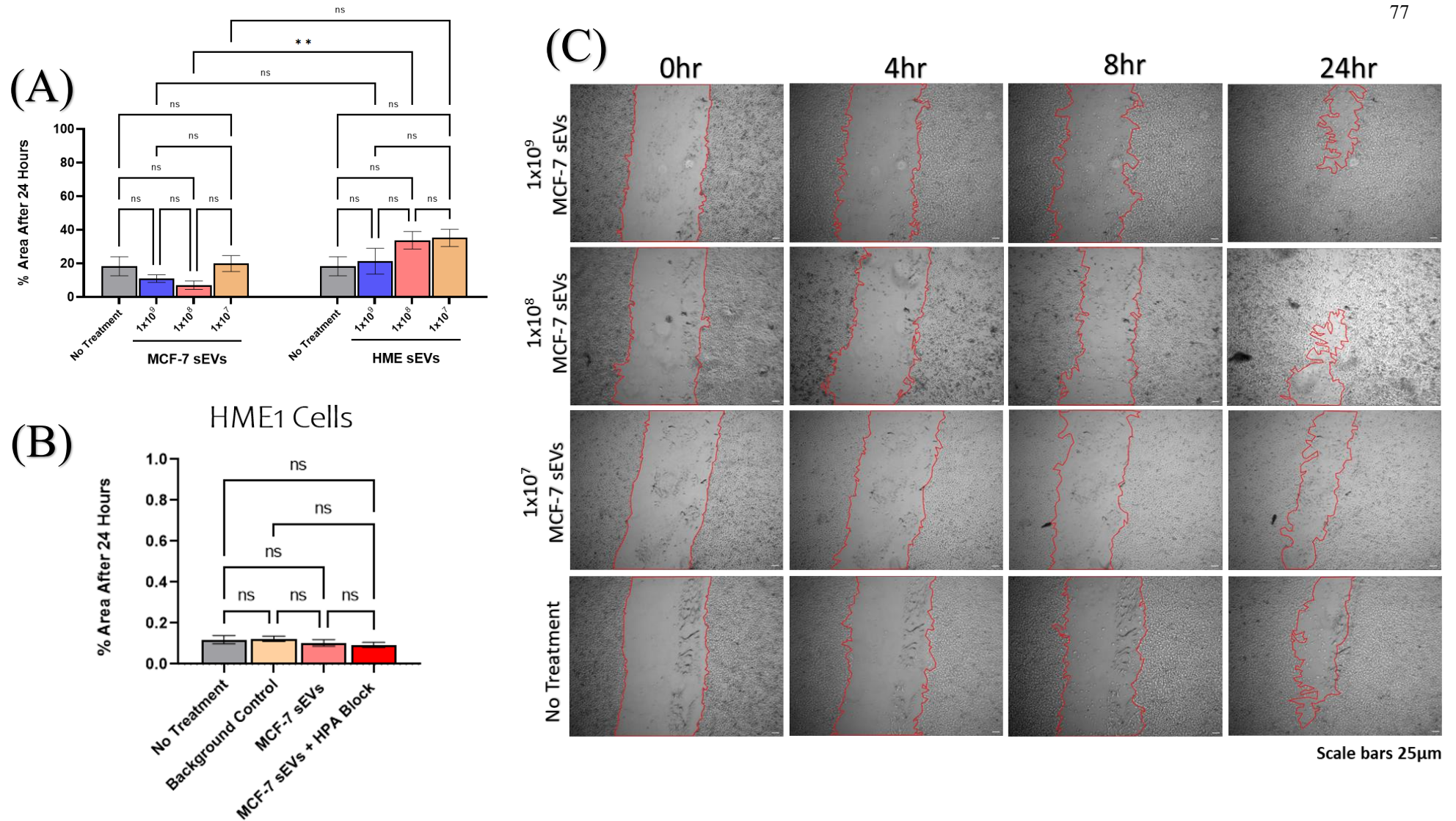
Figure 4.6. Wound healing assay using breast hTERT-HME1 cells. (A) *The migration area of hTERT-HME1 cells after treatment with different densities of MCF-7 sEVs and hTERT-HME1 sEVs (1×10^9, 1×10^8, and 1×10^7 particles/ml as determined by NTA) alongside a no-treatment control comprising growth media.* **(B)** *The migration area of hTERT-HME1 cells after treatment with MCF-7 sEVs labelled with HPA lectin alongside MCF-7 sEVs without HPA labelling, a background control with no sEVs subjected to the same procedure as the HPA-labelled sEVs, and a no-treatment control consisting of growth medium.* **(C)** *Representative images of hTERT-HME1 wound healing assay used to quantify cell migration at 0, 4, 8, and 24 h time points. The quantified areas are shown in red. Scale bars are 25µm. Each experiment consisted of biological triplicates with three technical triplicates, and error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. For parametric t-tests, Welch's correction was utilised, and for nonparametric t-tests, a Mann-Whitney test was used. Significance levels were set at *p < 0.05 and **p < 0.01.*
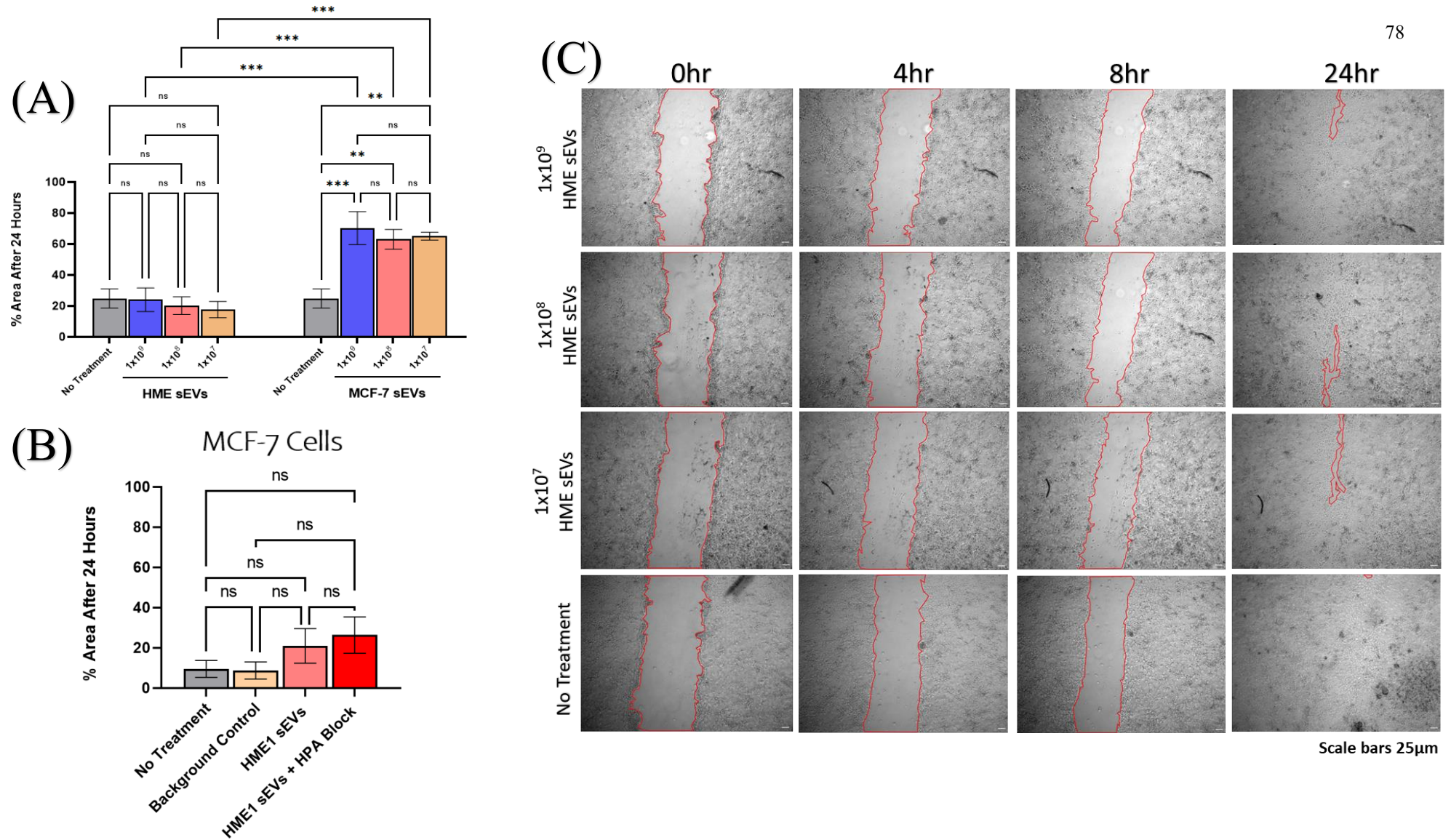
Figure 4.7. Wound healing assay assessment of breast MCF-7 cells. (A) *The migration area of MCF-7 cells after treatment with different densities of hTERT-HME1 sEVs and MCF-7 sEVs (1 × 10⁹, 1 × 10⁸, and 1 × 10⁷ particles/ml as determined by NTA) alongside a no-treatment control comprising growth media.* **(B)** *The migration area of MCF-7 cells after treatment with hTERT-HME1 sEVs labelled with HPA lectin alongside hTERT-HME1 sEVs without HPA labelling, a background control with no sEVs subjected to the same procedure as HPA-labelled sEVs, and a no-treatment control consisting of growth medium.* **(C)** *Representative images of the MCF-7 wound healing assay used to quantify cell migration at 0, 4, 8, and 24 h time points. Quantified area shown in red. Scale bars are 25µm. Each experiment consisted of biological triplicates with three technical triplicates, and error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. For parametric t-tests, Welch's correction was utilised, and for nonparametric t-tests, a Mann-Whitney test was used. The significance levels were set at *p < 0.05, **p < 0.01, and ***p < 0.001.*

4.5. DISCUSSION

4.5.1. HPA LECTIN BINDING CORRELATES WITH METASTATIC PHENOTYPE IN BREAST AND COLORECTAL EPITHELIAL CELL LINES

Breast and colorectal epithelial cell lines representing different phenotypes, including 'normal', 'primary', and 'metastatic' cancer, were assessed for HPA lectin binding using flow cytometry and confocal microscopy. The results revealed that both breast and colorectal cancer epithelial cell lines exhibited a higher HPA-binding capacity than 'normal' cell lines. HPA labelling was most pronounced in cell lines with a 'metastatic' phenotype, indicating the significant presence of GalNAc glycans. Confocal microscopy revealed that HPA binding specifically localised to the cell surface glycocalyx. These results align with prior investigations indicating a correlation between HPA binding and the metastatic cancer phenotype of breast and colorectal cells. Similarly, Brooks et al. (2001) conducted a study characterising various human breast cancer cell lines, such as BT-474 and MCF-7, along with normal breast epithelium, to assess HPA lectin binding using confocal microscopy and flow cytometry. Their findings revealed significant HPA lectin binding in 'metastatic' breast cells (MCF-7) compared to 'primary' BT-474 cells, whereas HPA binding was nearly absent in selected normal breast epithelial cells (HMT 3522). Xenograft models of colorectal cancer, specifically HT-29 cells, have also confirmed that HPA-positive tumour cells have a higher metastatic potential, whereas HPA-negative tumour cells typically do not exhibit metastatic behaviour (Schumacher & Adam, 1997). In addition, investigations of clinical tissue samples have demonstrated a strong association between HPA binding and metastasis in breast and colorectal cancers as discussed in section 1.4.3. However, it's important to note that these clinical studies have utilised a subjective scoring system and a positive control to evaluate HPA labelling intensity. In contrast, this study demonstrates a computational workflow for confocal microscopy quantification, offering advantages like reduced technical variability, minimised bias, and increased efficiency. Additionally, employing advanced technologies like flow cytometry, as shown in this study, could offer a more reliable and quantitative approach.

The specific recognition of HPA glycoproteins associated with cancer and metastasis, which is distinct from that found in normal cells, represents a crucial area for further exploration and understanding. Recently, by implementing lectin pulldown and proteomic tools, Khosrowabadi et al. (2022) uncovered ~85 glycoproteins with either known or potential binding for HPA in a variety of cancer cell line models. Among the glycoproteins, they identified several glycoproteins associated with metastasis, including MMP-14, EGFR, αV, β1, and β4 integrins, which displayed higher levels of the GalNAc glycotope in lectin pull-down samples of highly invasive cancer cell lines than in poorly invasive cells. However, the study highlighted the necessity for better identification of the glycoproteins that bind HPA, and a deeper understanding of how altered glycosylation regulates the activity of these proteins. Future research should focus on implementing advanced techniques to precisely identify glycoproteins associated with HPA binding and cancer metastasis.

4.5.2. CORRELATION OF HPA LECTIN BINDING WITH METASTATIC PHENOTYPE IN BREAST sEVs

Confocal microscopy and flow cytometry analysis of HPA lectin binding revealed a higher abundance of GalNAc glycans in 'metastatic' phenotypes of both breast and colorectal cells. Consequently, this study sought to investigate whether sEVs derived from breast and colorectal cell lines exhibited a similar correlative effect in terms of HPA lectin binding across different phenotypes, with a particular focus on CD81 and CD63 tetraspanin composition.

Similar to the tetraspanin analysis, two distinct analyses were conducted: quantification of the MESF and assessment of the proportion of sEVs in a population positive for HPA binding. However, concerns emerged in regards to the MESF analysis which revealed minimal significant differences. Such analyses typically involve antibodies directly, and no studies have used this approach with lectins. With extensive labelling involving CFDA-SE, tetraspanins, and HPA, questions were raised about the instruments capability to directly measure single vesicles at such sensitivity levels, raising concerns about measurement reliability. Consequently, while these analyses were conducted to meet MIFlowCyt-EV guidelines, analysis focused primarily on assessing the proportion of sEVs positive for lectin binding by the tetraspanin population (Welsh et al., 2020). This analysis was deemed more reliable because of the implemented normalisation process, which was absent in the MESF analysis.

Notably, 'metastatic' MCF-7 'CD81-positive' sEVs showed increased HPA lectin binding in comparison to 'CD63-positive' sEVs. Moreover, 'metastatic' MCF-7 'CD81-positive' sEVs showed increased HPA lectin binding compared with those from 'normal' hTERT-HME1 and 'primary' BT-474 cells. Interestingly, other studies have suggested the presence of mucin-type O-glycans on sEVs from other cancer types. For instance, Tn antigen has been detected on sEVs from brain, cervical, and pancreatic cancer cells (Dusoswa et al., 2019; Feng et al., 2018). Moreover, sTn has been described on sEVs derived from gastric and lung cancer cells (Freitas et al., 2019; Nagao et al., 2022). T antigen has also been identified on ovarian and cervical cancer cell-derived sEVs (Feng et al., 2018; Gomes et al., 2015). The study described here, however, marks the first specific identification of the increased abundance of GalNAc glycans recognised by HPA lectin on 'CD81-positive' sEVs derived from breast 'metastatic' cancer cells in comparison to those derived from 'normal' cells.

Moreover, analysis of the GalNac-glycans on sEVs align with what is seen at the cellular level, suggesting a potential parallel in glycan presentation between the cell surface and sEVs, but with nuanced specificity attributed to tetraspanin composition. However, recent evidence has suggested a more intricate relationship. For example, Nishida-Aoki et al. (2020) identified distinct glycosylation patterns on sEVs derived from BMD2a cells, a brain metastatic subline of the human triple-negative breast cancer cell line, MDA-MB-231. Their study revealed unique glycosylation profiles in BMD2a sEVs compared to those in their parental cells, suggesting that sEVs possess a distinctive glycosylation pattern. Moreover, Gomes et al. (2015) demonstrated that sEVs derived from ovarian carcinoma cells, OVMz, exhibit distinct glycosignatures, such as complex N-glycans and T-antigen O-glycans, differentiating them from whole-cell membranes. These findings are consistent with those of Escrevente et al.

(2011) who showed that sEVs from ovarian cancer were enriched in specific Man- and Sia-containing glycoproteins. Collectively, these contrasting findings underscore the likelihood of a selective mechanism dictating the glycan composition of sEVs, revealing the complex and highly specific processes of sEV glycosylation.

Further analysis of HPA lectin binding in 'CD63-positive' sEVs derived from 'normal' hTERT-HME1 cells revealed a higher level of binding compared to that in 'CD63-positive' sEVs derived from 'metastatic' MCF-7 cells. In the experiments described in Chapter 3, sEVs were subjected to characterisation of CD63 by western blot analysis and appeared as smears, suggesting the presence of glycovariants. Despite the evidence that CD63 is highly glycosylated, it's suggested that the glycosylation pattern might not favour HPA lectin binding. A study conducted by Terävä et al. (2022) determined that CD63 glycovariants in MCF-7 cells showed pronounced binding to *Ulex europaeus* agglutinin (UEA) lectin, which primarily recognises Fuc(α1-2)Gal, but exhibited minimal interaction with *Vicia villosa* (VVL) lectin, a lectin with carbohydrate specificity, for GalNAc, similar to HPA lectin. This observation may partly explain the reduced HPA detection in 'CD63-positive' sEVs from MCF-7 cells. This suggests that the glycosylation profile of CD63 on these sEVs is skewed towards glycan structures that are not preferred substrates for HPA lectin.

4.5.3. BREAST CELL LINE TREATMENT WITH sEVs IMPACTS MOTILITY CHARACTERISITICS

Initially, attempts were made to investigate the role of sEVs and their uptake by recipient cells in altering adhesive capabilities to endothelial monolayers, using a static adhesion assay (Supplementary Figure 8 and 9). However, no significant effects were observed, leading to a focus on investigating a potential effect of sEVs in modulation of cell motility. Therefore, to investigate the specific role of sEVs in facilitating changes in cell motility, a wound-healing assay was performed. Owing to the current lack of definitive understanding of sEVs secretion rates for each cell line, which can be influenced by a variety of physiological conditions, the experimental design of the wound healing assay incorporated a range of sEV concentrations. Interestingly, significant changes in cell motility associated with sEVs treatment were observed only in 'normal' breast hTERT-HME1 and 'metastatic' MCF-7 breast cancer cells. Specifically treating hTERT-HME1 cells with MCF-7 sEVs induced motility by increasing gap closure compared to hTERT-HME1 sEVs at 1×10^8 particles/ml. However, this effect was not evident at higher (1×10^9 particles/ml) or lower (1×10^7 particles/ml) concentrations of the MCF-7 sEVs. This pattern may indicate the presence of a specific concentration threshold or limit within which sEVs induce changes in cell motility. These findings contrast with those of a study conducted by Bertolini et al. (2020) who implemented a wound healing assay to assess normal MCF10A breast epithelial cells treated with MCF-7 sEVs and observed a minimal increase in cell migration compared to controls. This discrepancy might be attributed to factors such as the higher sEVs density used in their study, set at 3×10^8 particles/ml sEVs which may further support the notion of a concentration-dependent threshold for the role of sEVs in modulating recipient cell motility changes. In addition, the choice of a different representative 'normal' breast epithelial cell line (MCF10A vs. hTERT-HME1) was also implemented, as hTERT-HME1 cells are diploid, whereas MCF10A cells exhibit aneuploidy. Chromosomal instability (CIN) inherent in aneuploid cells could potentially influence their signalling pathways

and cellular behaviours, thereby affecting their responses to sEV-mediated motility changes. Another study by Galindo-Hernandez et al. (2015) adopted a wound healing assay with MCF10A cells treated with sEVs from breast cancer patients and showed a significant increase in the migration pattern compared to sEVs isolated from 'healthy' individuals. The study deduced that enhanced motility was attributed to sEVs harbouring molecules which facilitated the expression of mesenchymal markers such as N-cadherin and vimentin, which are indicative of EMT. These findings also suggest that the MCF-7 sEVs used in this study may carry bioactive cargo which facilitates EMT to induce increased motility. Dalla et al. (2020) provided further insights into the molecular composition of MCF-7 sEVs using peptide-centric LC-MS/MS proteomics. Amongst the identified proteins of MCF-7 sEVs was CD147 (EMMPRIN), which several studies have described as playing a significant role in cancer cell adhesion and migration (Landras et al., 2019). MCF-7 sEV cargo, such as CD147, may influence the motility of recipient cells as observed in this study.

4.5.4. MASKING HPA LECTIN RECOGNISING GLYCANS OF BREAST sEVs DO NOT ALTER RECIPIENT CELL MOTILITY PHENOTYPE

As a result of the motility changes observed following the administration of breast sEVs, the focus shifted towards masking GalNAc glycans on the sEVs by the application of HPA before initiating treatment. However, results revealed no significant differences, suggesting that HPA-binding glycans on sEVs had no effect on treated recipient cell motility. Notably, carbohydrate structures, such as oligosaccharides, polysaccharides, and glycoproteins, have been established to play pivotal roles on the vesicle surface, influencing the recipient cell uptake of sEVs (Becker et al., 2016; Lin et al., 2020; Martins et al., 2021). Specifically, Williams et al. (2019) investigated the role of surface glycans of sEVs in cellular uptake by measuring cellular sEVs uptake by flow cytometry after treatment with peptide-N-glycosidase F, which cleaves N-glycans and assesses glycan content by lectin microarray and established an increase uptake of sEVs. They attributed the increased uptake of sEVs to the removal of glycan structures which removes the steric hindrance of other vesicle surface ligands that are able to encounter their cell surface receptors and through differences in charge-based effects. However, it is important to note that the exact digestion efficacy of peptide-N-glycosidase F cannot be determined using lectin microarray alone. Nevertheless, considering this, it is plausible that the addition of HPA, designed to obstruct specific glycans, may have introduced increased steric hindrance on the vesicle surface. This impedes the interaction of other vesicle surface ligands with their respective cell surface receptors, thereby affecting their cellular uptake. Furthermore, alteration of surface glycans through fluorescently labelled HPA lectin binding may induce changes in interaction properties. For instance, Romanowska et al. (2015) compared hen egg white lysozyme (HEWL) and HEWL labelled with fluorescein isothiocyanate, and demonstrated that a small label can significantly change the interaction properties of a protein. Collectively, the increased steric hindrance and changes in interaction properties through fluorescently labelled HPA lectin may have contributed to the observed absence of significant differences in cellular motility.

4.6. KEY FINDINGS

- Increased HPA lectin binding in breast and colorectal cancer epithelial cells correlates with 'metastatic' phenotype, specifically MCF-7 (breast) and HT-29 (colorectal) in comparison to 'primary' BT-474 (breast) and SW480 (colorectal) and 'normal' hTERT-HME1 (breast) and NCM460 (colorectal) cells.

- Increased HPA lectin binding to breast 'metastatic' MCF-7 'CD81-positive' sEVs in comparison to 'primary' BT-474 and 'normal' hTERT-HME1 'CD81-positive' sEVs.

- Increased HPA lectin binding to breast 'normal' hTERT-HME1 'CD63-positive' sEVs in comparison to 'metastatic' MCF-7 'CD63-positive' sEVs.

- Breast model epithelial cells have increase migratory capabilities in comparison to colorectal model epithelial cells.

- Treatment of breast cells with their respective sEVs resulted and facilitated an inhibitory effect on cellular motility.

- Masking HPA-glycans on breast sEVs does not impact recipient cell motility.

CHAPTER 5

INVESTIGATING THE GLYCOME OF CANCER-ASSOCIATED sEVs

5. INVESTIGATING THE GLYCOME OF CANCER-ASSOCIATED sEVs

5.1. BACKGROUND

The experimental findings in Chapter 3 and Chapter 4 revealed higher levels of 'metastatic' MCF-7 'CD81-positive' sEVs and HPA lectin binding, compared to 'primary' BT-474 and 'normal' hTERT-HME1 'CD81-positive' sEVs. This pattern was also observed in parental cells with respect to HPA-lectin binding. Hence, it was hypothesised that other glycosylated targets on 'CD81-positive' breast sEVs might similarly reflect the glycan composition of their parental cell surfaces. In this chapter, 'CD81-positive' breast sEVs derived from the model cell lines were profiled for their glycosylated targets, by implementing a lectin microarray, aiming to identify lectins specifically bound to cancer-derived sEVs and specifically 'metastatic' cancer-derived sEVs. Subsequently, efforts shifted to validate the lectin microarray findings by adopting single-vesicle flow cytometry analysis. Additionally, flow cytometry was performed to determine whether the observed glycosylation patterns at the sEVs level reflected those of their parental cell surface glycosylation, in the same way as previously observed with HPA lectin binding of 'CD81-positive' breast sEVs.

5.1.1. LECTIN MICROARRAY ANALYSIS

Lectin microarray analysis has been adopted for a variety of sEVs studies as it allows targeting of all surface glycoconjugates of intact vesicles in a single analysis, making lectin microarrays advantageous for unbiased analysis (reviewed by Williams et al., 2018). For example, Shimoda et al. (2022) employed a lectin microarray to analyse the glycan content of sEVs from 20 distinct cell lines and compared it with the glycan profiles of parental cell membranes. Their study revealed notable differences in glycan profiles between the sEV groups and their respective original cells. sEVs and their parental cell membranes, in contrast to the observations of HPA lectin binding in the present study. This is intriguing not only for the clear application of lectin microarrays for sEVs analysis but also for the striking differences observed in glycan content between sEVs and their parental cell membranes.

5.2. AIMS AND OBJECTIVES

The aim of this chapter is to explore alternative glycosylation targets of 'CD81-positive' breast sEVs that correlate specifically to cancer-associated sEVs or 'metastatic' cancer associated sEVs and assess whether these distinct glycans also reflect the surface glycosylation of their parental cells.

The objectives were:

1) To identify lectins binding 'CD81-positive' breast sEVs derived from 'normal', 'primary' and 'metastatic' cancer cell phenotypes of breast epithelial cell lines using lectin microarray analysis.

2) Validate the lectins binding to 'CD81-positive' breast and colorectal sEVs derived from 'normal', 'primary' and 'metastatic' cancer cell phenotypes using single-vesicle flow cytometry.

3) To quantify lectins binding to breast and colorectal epithelial cell lines derived from normal', 'primary' and 'metastatic' cancer cell phenotypes using single-vesicle flow cytometry.

5.3. METHODS

5.3.1. LECTIN MICROARRAY ANALYSIS OF CD81-POSITIVE BREAST sEVs

To examine the surface glycan profiles of sEVs derived from breast cell lines representing different phenotypes, a lectin microarray comprising 95 different lectins was adopted following the methodology outlined in section 2.8. (Figure 2.4 and Supplementary Table 3).

5.3.2. SINGLE-VESICLE FLOW CYTOMETRY

5.3.2.1. LCA LECTIN AND TL PROFILING OF EPITHELIAL DERIVED sEVs

To quantify LCA and TL binding of epithelial-derived sEVs from breast and colorectal cells, single-vesicle flow cytometry was employed. The methodology for extracting sEVs from their parent cells is detailed in section 2.4.1, while section 2.6 outlines the procedure, including instrument settings (section 2.6.1), calibration for sizing and fluorescence before sEV acquisition (section 2.6.2), and staining and labelling steps for sEVs lipids, CD81, CD63, LCA and TL (section 2.6.3).

5.3.3. FLOW CYTOMETRY ANALYSIS OF LCA LECTIN AND TL LABELLING ON EPITHELIAL CELL SURFACES

To quantify LCA and TL lectin binding on the cell surface of epithelial cell models of breast and colorectal, flow cytometry was performed following the methodology outlined in section 2.2.

5.4. RESULTS

5.4.1. LECTIN MICROARRAY ANALYSIS OF BREAST sEVs

To analyse the surface glycan profiles of breast cell-line-derived sEVs, a lectin microarray comprising of 95 different lectins was used (Figure 2.4 and Supplementary Table 3). MCF-7-derived sEVs displayed a higher overall fluorescence intensity for lectin binding than BT-474 and hTERT-HME1 sEVs (Figure 5.1 A-C). Volcano plots were generated to visualise the differences in the surface glycan patterns of sEVs between breast cell lines. This comparison was conducted using a two-tailed t-test across sEVs (BT-474 vs. MCF-7, hTERT-HME1 vs. BT-474, and hTERT-HME1 vs. MCF-7), and the $\log_2$ fold change was calculated. A p-value < 0.05, and an absolute value of the average $\log_2$ fold-change > 1.5 were used to identify the lectins bound to each breast sEVs (Table 5.1, Figure 5.2 A-C). Sixteen lectins showed a significant fold-change in recognising MCF-7 sEVs compared with hTERT-HME1 sEVs, including RPA, PTL-1, SNA-II, WGA, BPA, lentil, BANLEC, TL, UEA-II, PHA-L, GNA, UDA, GAL3 MNAM, ORYSATA, TL, and HAA. Furthermore, five lectins showed a significant fold-change in binding MCF-7 sEVs compared with BT-474 sEVs, including RPA, BPA, lentil, SHA, ABA, and ASA. Additionally, 12 lectins demonstrated a significant fold-change in recognising BT-474 sEVs compared with hTERT-HME1 sEVs, including BANLEC TL, WGA, PHA-L, GAL2, CSA, GNA, PTL-1, UDA, ASA, GAL3C-S, and HAA. A heatmap was also generated to depict the binding profiles of each breast cell line-derived sEV to the lectin microarray, as visualised through CD81 detection and is presented in Figure 5.3.

Following the observation of higher overall fluorescence intensity for lectin binding in MCF-7-derived sEVs compared to BT-474 and hTERT-HME1 sEVs (Figure 5.1 A-C), further analysis focused on lectins, lentil and TL. Notably, Lentil lectin, also denoted LCA lectin, demonstrated significant binding to 'metastatic' MCF-7 'CD81-positive' sEVs in comparison to 'primary' BT-474 sEVs, and is the only lectin among those mentioned with reported applications in cancer diagnostics. Whereas for TL, which demonstrated significant binding to 'primary' BT-474 and 'metastatic' MCF-7 'CD81-positive' sEVs in comparison to 'normal' hTERT-HME1 sEVs, no studies have investigated these findings in the context of its preferential binding to cancer-associated sEVs compared to normal sEVs. Therefore, given the established literature supporting the specific binding of LCA in a tumour context and the absence of existing studies on TL, further investigation of these lectins is warranted.

Table 5.1. Significant differences observed in lectin microarray analysis of CD81-positive breast sEVs

COMPARISON OF BREAST sEVs	SIGNIFICANT LECTINS BOUND
MCF-7 sEVs Vs hTERT-HME1 sEVs	RPA, PTL-1, SNA-II, WGA, BPA, LENTIL, BANLEC TL, UEA-II, PHA-L, GNA, UDA, GAL3 MNAM, ORYSATA, TL, HAA
MCF-7 sEVs Vs BT-474 sEVs	RPA, BPA, LENTIL, SHA, ABA, ASA
BT-474 sEVs VS hTERT-HME1 sEVs	BANLEC TL, WGA, PHA-L, GAL2, CSA, GNA, PTL-1, UDA, ASA, GAL3C-S, and HAA
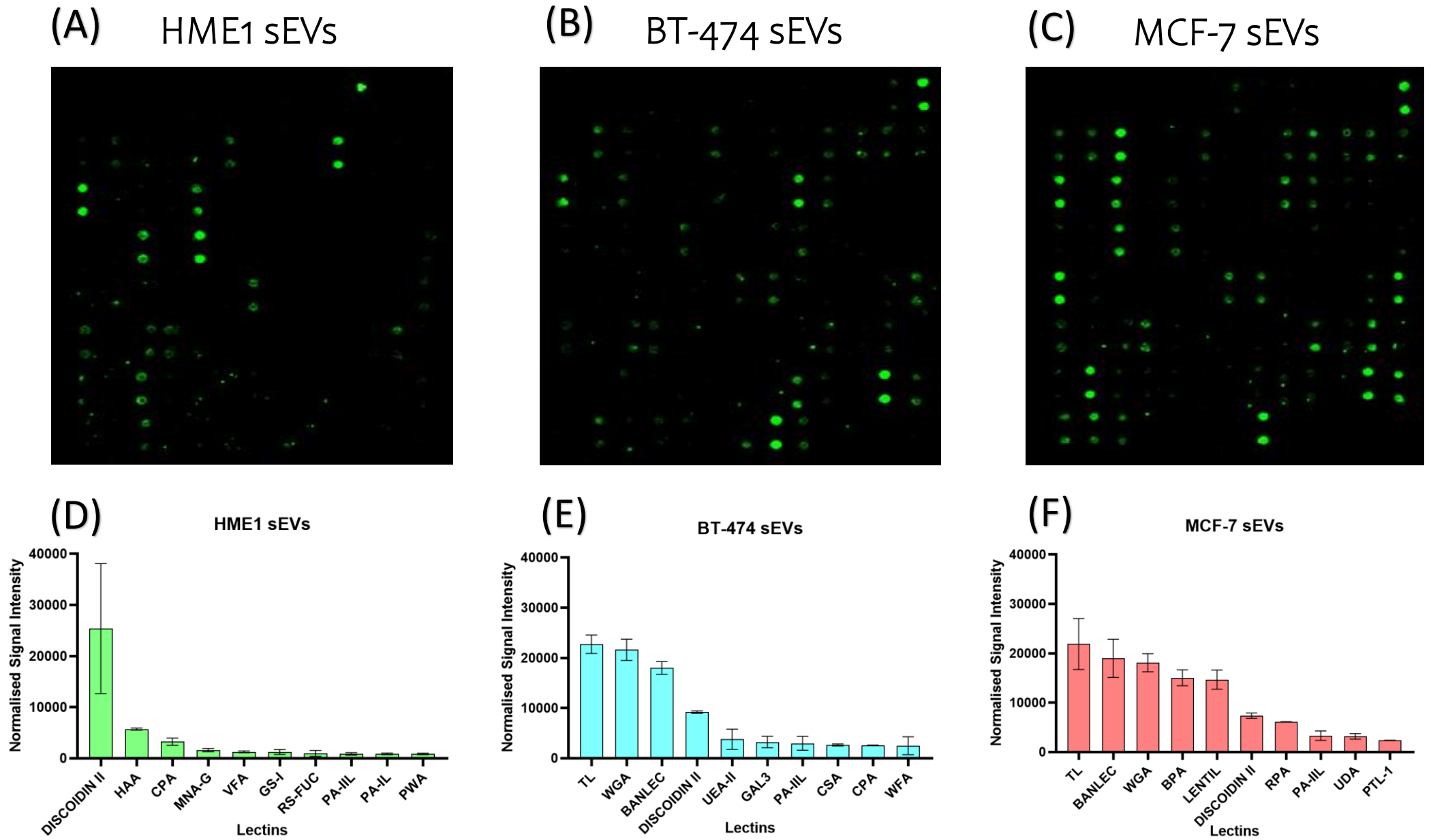
Figure 5.1. Overview of lectin microarray analysis of breast cell line derived sEVs (A-C) *Images of lectin array spots of hTERT-HME1, BT-474, and MCF-7 sEVs of CD81-positive sEVs (***D-F)** *The 10 highest-intensity lectin binding array spots after the addition of sEVs from each breast cell line.*
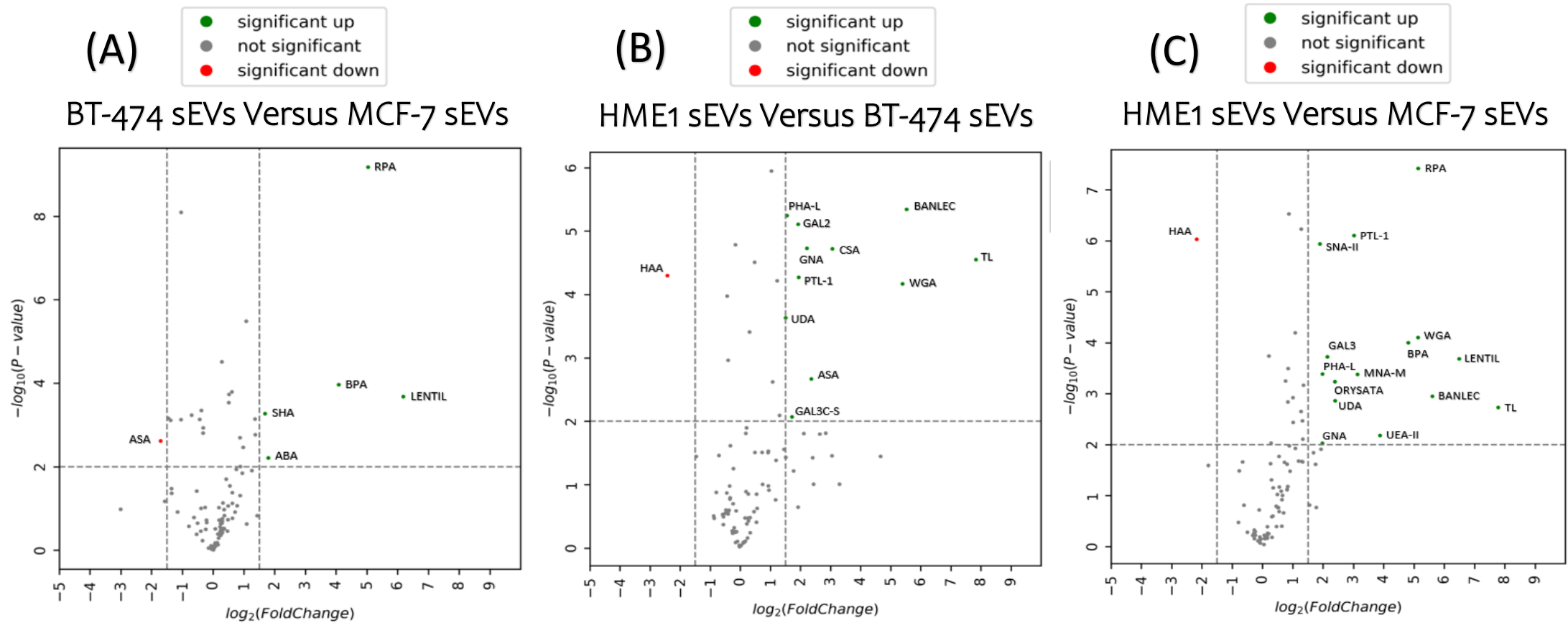
Figure 5.2. Volcano plots generated from lectin microarray (A-C) *Volcano Plot analysis of the lectin array data between BT-474 vs. MCF-7, HME vs. BT-474, and HME vs. MCF-7 sEVs. The vertical lines correspond to a 1.5-fold up- and down-regulation, whereas the horizontal lines represent a p-value of 0.05. The green dots to the left and right of the vertical lines indicate more than a 2.0-fold change, highlighting lectins that exhibit statistically significant differences in binding.*

20000

10000

MCF-7 sEVs

BT-474 sEVs

HME1 sEVs

Lectins: AAA, AAL, BC2LCN, LAL, LOTUS, PA-IIL, PA-IL, RS-FUC, UEA-I, ABA, ABL, ACL, BPA, CNL, DISCOIDIN II, ECA, EEL, GAL-7S, GHA, GSI, JACALIN, MAA, MNA-G, MOA, MPL, PHA-E, PHA-L, PHA-P, PNA, RCA120, RCA60, VRA, AMA, ASA, BANLEC, BC2L-A, CALSEPA, CONA, GNA, GRFT, HHA, LCHA, LENTIL, MALECTIN, MNA-M, NPA, ORYSATA, PALa, PSA, VFA, VVA-M, CA, CAA, CGL-2, CPA, CSA, DBA, DISCOIDIN I, GAL-2, HAA, IRA, LBA, PPL, PTL-1, PTL-2, RPA, SBA, SHA, SJA, TL, VVA, WFA, DSA, F17AG, GSII, LEA, PWA, STL, UDA, UEA-II, WGA, GAL-1, GAL-1S, GAL-3, GAL-3C-S, GAL-9, LSLN, ACG, HMA, LPA, PSL-1A, SAMB, SNA-I, SNA-II, SSA

Glycan groups: Fuc, Gal, Man, GalNAc, GlcNAc, LacNAc, NANA

Figure 5.3. Heatmap of the lectin microarray for breast cell line derived sEVs. *Normalised fluorescent intensities of the 95 lectins (rows) for breast epithelial sEVs, including hTERT-HME1, BT-474, and MCF-7 sEVs (columns). Each rectangle in the heatmap is colour-coded to represent the binding affinity of the corresponding lectin (white for no binding, blue for low binding, green for medium binding, and orange for high binding). Glycan binding specificities for the 95 lectins are indicated below: Fuc (Fucose), Gal (Galactose), Man (Mannose), GalNAc (N-Acetylgalactosamine), GlcNAc (N-Acetylglucosamine), LacNAc (N-Acetyllactosamine), NANA (N-Acetylneuraminic acid).*

5.4.2. TETRASPANIN SUBPOPULATION ANALYSIS OF BREAST AND COLORECTAL EPITHELIAL sEVs WITH FLUORESCENTLY LABELLED LCA

The lectin microarray results from section 5.4.1, utilising a CD81-only approach, highlighted a potential correlative link between LCA binding and the metastatic phenotype of breast cancer-associated sEVs. Specifically, the results revealed increased LCA binding on 'metastatic' MCF-7 sEVs compared to that of 'primary' BT-474 and 'normal' hTERT-HME1 sEVs. Therefore, to confirm the lectin microarray results, single-vesicle flow cytometry analysis was performed. The raw results, including the mean ± SD of LCA binding via CD81 or CD63 detection normalised to CFSE events/ml and MESF values, are shown in Table 5.2.

Employing fetuin as a control was important in validating the specificity of interactions between LCA and the surface glycans of breast and colorectal epithelial sEVs. Fetuin, a protein abundant in both N-linked and O-linked sugars, served as a competitive inhibition control to confirm specific binding events between LCA and the target sEVs derived from breast and colorectal epithelial cells (Figure 5.4 C-G and Figure 5.5 C-G). The histograms, as depicted in Figure 5.4 and 5.5 C-H, show a notable increase in LCA-binding fluorescence across all breast and colorectal sEVs compared to the fetuin control. Initial comparative analysis revealed diverse LCA binding patterns in breast epithelial sEVs, based on their tetraspanin composition (Figure 5.4-A). Specifically, increased LCA binding was revealed by a 1.2-fold increase in MCF-7 'CD81-positive' sEVs compared to 'CD63-positive' sEVs ($p < 0.05$). Conversely, increased LCA binding was observed in hTERT-HME1 CD63-postive sEVs in comparison to that in 'CD63-positive' sEVs ($p < 0.0001$). However, no significant differences were observed for LCA-binding BT-474 sEVs across different tetraspanin profiles. Further comparisons between the different phenotypes of sEVs revealed that MCF-7 'CD81-positive' sEVs exhibited a 2-fold and 3.5-fold increase in LCA binding compared to hTERT-HME1 and BT-474 sEVs, respectively ($p < 0.0001$ for hTERT-HME1; $p < 0.01$ for BT-474). Moreover, hTERT-HME1 'CD81-positive' sEVs showed a 1.8-fold increase in LCA binding compared to BT-474 'CD81-positive' sEVs. In contrast, 'CD63-positive' sEVs showed different binding profiles for LCA lectin. Notably, hTERT-HME1 'CD63-positive' sEVs displayed a 1.3-fold and 2.4-fold increase in LCA binding compared with MCF-7 and BT-474 'CD63-positive' sEVs, respectively. MESF analysis revealed a general enrichment of CD63-postive sEVs compared with 'CD81-positive' sEVs across all breast sEVs phenotypes. Specifically, CD63-postive sEVs of hTERT-HME1 showed a 1.7-fold increase ($p < 0.001$), BT-474 sEVs showed a 1.3-fold increase ($p < 0.05$), and MCF-7 sEVs showed a 1.2-fold increase compared to CD81-postive sEVs ($p < 0.05$) (Figure 5.4-B). Comparisons of MESF values revealed that MCF-7 CD81-postive sEVs displayed a 1.4-fold increase over BT-474 'CD81-positive' sEVs for LCA lectin binding ($p < 0.01$). However, no significant differences were observed in MESF analysis of LCA lectin binding between 'CD81-positive' sEVs of MCF-7 and hTERT-HME1 or between hTERT-HME1 and BT-474 sEVs. Additionally, hTERT-HME1 CD63-postive sEVs showed increased MESF values of 1-2-fold and 1.4-fold compared to CD63-postive sEVs of MCF-7 ($p < 0.05$) and BT-474 sEVs ($p < 0.01$), respectively, for LCA lectin binding. MCF-7 CD63-postive sEVs also showed an increase in MESF values of 1.2-fold for LCA lectin binding in comparison to BT-474 CD63-postive sEVs ($p < 0.05$).

Comparisons of colorectal sEVs revealed that LCA binding was increased in NCM460 'CD63-positive' sEVs compared to 'CD81-positive' sEVs by 1.7-fold ($p < 0.05$) (Figure 5.5-A). Comparisons of the different phenotypes between 'CD81-positive' and CD63-postive sEVs revealed no significant differences in LCA binding. In contrast, MESF analysis indicated an overall increase in LCA lectin binding in CD63-postive sEVs for all colorectal sEV phenotypes: a 1.7-fold increase in NCM460 sEVs ($p < 0.001$), a 1.7-fold increase in SW480 sEVs ($p < 0.0001$).

Table 5.2. LCA lectin binding of CD81-positive and CD63-positive sEVs on breast and colorectal epithelial sEVs by single-vesicle flow cytometry.

Organ of Origin	Cell Line sEVs	CD81-positive sEVs LCA MESF (mean ± SD)	CD63-positive sEVs LCA MESF (mean ± SD)	CD81-positive sEVs LCA events/ml % of CFSE events/ml (mean ± SD)	CD63-positive sEVs LCA events/ml % of CFSE events/ml (mean ± SD)	CD81-positive vs CD63-positive sEVs LCA	Vs HME /NCM460 sEVs	Vs MCF-7/ HT-29 sEVs
Breast	hTERT-HME1 sEVs	20.66 ± 2.83	34.34 ± 4.33	10.84 ± 0.96	22.61 ± 3.61	**Events/ml** ******** **MESF** *******		**Events/ml** **CD81= ****** **CD63=*** **MESF** **CD81= NS** **CD63= ***
	BT-474 sEVs	17.86 ± 1.41	23.62 ± 2.48.	5.72 ± 3.08	9.20 ± 6.16	**Events/ml** **NS** **MESF** *****	**Events/ml** **CD81= *** **CD63= *** **MESF** **CD81= NS** **CD63= ****	**Events/ml** **CD81= **** **CD63= NS** **MESF** **CD81= **** **CD63= ***
	MCF-7 sEVs	23.86 ± 3.37	28.43 ± 2.98	21.63 ± 3.07	18.57 ± 0.4	**Events/ml** ***** **MESF** *****	**Events/ml** **CD81= ****** **CD63= *** **MESF** **CD81= NS** **CD63= ***	
Colorectal	NCM460 sEVs	17.30 ± 4.13	30.20 ± 2.80	13.01 ± 3.01	9.44 ± 4.52	**Events/ml** ***** **MESF** *******		**Events/ml** **CD81= NS** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**
	SW480 sEVs	18.06 ± 3.49	31.03 ± 2.34	15.09 ± 2.80	12.70 ± 6.44	**Events/ml** **NS** **MESF** ********	**Events/ml** **CD81= NS** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**	**Events/ml** **CD81= NS** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**
	HT-29 sEVs	19.38 ± 5.37	32.65 ± 2.84	15.37 ± 6.31	10.90 ± 8.74	**Events/ml** **NS** **MESF** *******	**Events/ml** **CD81= NS** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**	

The asterisks (*) in the table denote significance levels.

The NS in the table indicates no significance.
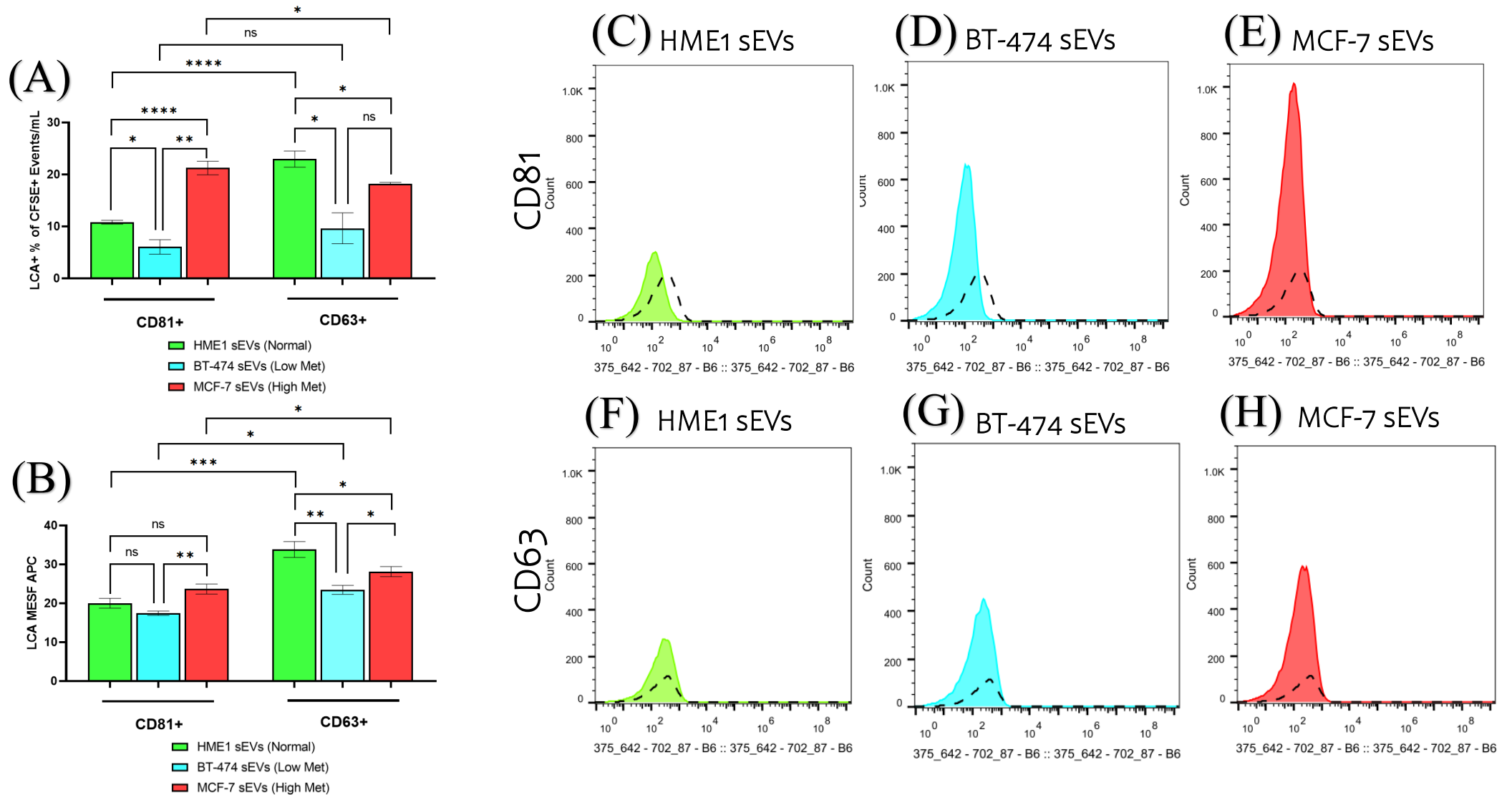
Figure 5.4. Single vesicle flow cytometry analysis of LCA lectin binding of CD81-positive or CD63-positive breast sEVs. (A) *Quantification of LCA+ events/ml normalised to the percentage of CFSE events/ml for hTERT-HME1, BT-474, and MCF-7 CD81-postive or CD63-positive sEVs.* **(B)** *APC MESF values of LCA+ for hTERT-HME1, BT-474, and MCF-7 sEVs with arbitrary units of MFI converted to standardised MESF units using vCal^TM nanorainbow beads.* **(C-E)** *Histogram profiles of the breast sEVs after gating showing LCA lectin binding of CD81-positive sEVs fluorescence alongside the sugar-specific control fetuin outlined in black.* **(F-H)** *Histogram profiles of the breast sEVs after gating showing LCA lectin binding of CD63-positive sEVs fluorescence alongside the sugar-specific control fetuin outlined in black. Each experiment consisted of biological triplicates with three technical triplicates, and the error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. Welch's correction was used for parametric t-tests, and the Mann-Whitney test was used for nonparametric t-tests. Significance levels were set at *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.*

Figure 5.5. Single vesicle flow cytometry analysis of LCA lectin binding of CD81-positive or CD63-positive colorectal sEVs. (A) *Quantification of LCA+ events/ml normalised to the percentage of CFSE events/ml for NCM460, SW480, and HT-29 CD81-postive or CD63-positiv sEVs.* **(B)** *APC MESF values of LCA+ for NCM460, SW480, and HT-29 sEVs with arbitrary units of MFI converted to standardised MESF units using vCal^TM nanorainbow beads.* **(C-E)** *Histogram profiles of the colorectal sEVs after gating showing LCA lectin binding of CD81-positive sEVs fluorescence alongside the sugar-specific control fetuin outlined in black.* **(F-H)** *Histogram profiles of the colorectal sEVs after gating showing LCA lectin binding of CD63-positive sEVs fluorescence alongside the sugar-specific control fetuin outlined in black. Each experiment consisted of biological triplicates with three technical triplicates, and error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. For parametric t-tests, Welch's correction was utilised, and for nonparametric t-tests, a Mann-Whitney test was used. Significance levels were set at *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.*

5.4.3. TETRASPANIN SUBPOPULATION ANALYSIS OF BREAST AND COLORECTAL EPITHELIAL sEVs WITH FLUORESCENTLY LABELLED TL

The lectin microarray results presented in section 5.4.1, utilising a CD81-only approach, highlighted increased binding of TL with 'primary' BT-474 and 'metastatic' MCF-7 sEVs in comparison to 'normal' hTERT-HME1 sEVs. Therefore, to confirm the lectin microarray results, single-vesicle flow cytometry analysis was performed. The raw results, including the mean ± SD of TL binding via CD81 or CD63 detection normalised to CFSE events/ml and MESF values, are shown in Table 5.3.

The histograms, as depicted in Figure 5.6 and 5.7 C-H, show a notable increase in TL-binding fluorescence across all breast and colorectal sEVs compared to the fetuin competitive inhibition control. Comparative analysis of TL binding showed a 2.2-fold increase on 'normal' hTERT-HME1 'CD63-positive' sEVs compared to 'CD81-positive' sEVs ($p < 0.05$) (Figure 5.6-A). No significant differences were observed in TL binding between 'CD63-positive' and 'CD81-positive' sEVs on MCF-7 and BT-474 cells. In-depth analysis of MCF-7 'CD81-positive' sEVs showed that TL binding was increased by 2-fold and 3.3-fold compared to hTERT-HME1 and BT-474 sEVs, respectively ($p < 0.01$). However, no significant differences were observed in TL binding between CD81-postive sEVs for hTERT-HME1 and BT-474 sEVs. hTERT-HME1 'CD63-positive' sEVs showed a 2.2-fold increase in TL binding compared to 'CD63-positive' BT-474 sEVs ($p < 0.05$). No significant differences were observed when comparing TL binding between 'CD63-positive' sEVs from hTERT-HME1 and BT-474 to those from MCF-7 cells. Analysis of MESF values revealed increased TL binding in hTERT-HME1 'CD63-postive' sEVs in comparison to their CD81-postive sEVs by 1.1-fold ($p < 0.05$) (Figure 5.6-B). Moreover, comparative analysis of MESF values also revealed that TL binding increased by 1.1-fold in BT-474 'CD81-positive' sEVs compared to that in hTERT-HME1 and MCF-7 'CD81-positive' sEVs ($p < 0.05$). TL binding of hTERT-HME1 'CD63-positive' sEVs increased by 1-fold relative to MCF-7 'CD63-positive' sEVs ($p < 0.05$), with no other significant differences observed. Comparisons of colorectal epithelial sEVs for TL binding across both CD81 and CD63 tetraspanins revealed no significant differences (Figure 5.7 A-B).

Table 5.3. TL lectin binding of CD81-positive and CD63-positive sEVs on breast and colorectal epithelial sEVs by single-vesicle flow cytometry. The asterisks (*) in the table denote significance levels.

Organ of Origin	Cell Line sEVs	CD81-positive sEVs TL MESF (mean ± SD)	CD63-positive sEVs TL MESF (mean ± SD)	CD81-positive sEVs TL events/ml % of CFSE events/ml (mean ± SD)	CD63-positive sEVs TL events/ml % of CFSE events/ml (mean ± SD)	CD81-positive vs CD63-positive sEVs TL	Vs HME /NCM460 sEVs	Vs MCF-7/ HT-29 sEVs
Breast	hTERT-HME1 sEVs	11.23 ± 0.31	11.75 ± 0.11	10.62 ± 1.72	22.32 ± 9.92	**Events/ml** * **MESF** *		**Events/ml** **CD81= **** **CD63=NS** **MESF** **CD81= NS** **CD63= ***
	BT-474 sEVs	12.03 ± 0.33	11.69 ± 0.12	6.24 ± 4.30	9.20 ± 9.23	**Events/ml** NS **MESF** NS	**Events/ml** **CD81= NS** **CD63= *** **MESF** **CD81= **** **CD63= NS**	**Events/ml** **CD81= **** **CD63= NS** **MESF** **CD81= *** **CD63= NS**
	MCF-7 sEVs	11.57 ± 0.14	11.60 ± 0.11	21.34 ± 1.51	17.16 ± 6.46	**Events/ml** NS **MESF** NS	**Events/ml** **CD81= **** **CD63= NS** **MESF** **CD81= NS** **CD63= ***	
Colorectal	NCM460 sEVs	11.63 ± 0.24	11.88 ± 0.37	15.24 ± 5.39	21.85 ± 8.72	**Events/ml** NS **MESF** NS		**Events/ml** **CD81= NS** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**
	SW480 sEVs	11.53 ± 0.29	11.76 ± 0.46	17.85 ± 5.37	25.58 ± 10.61	**Events/ml** NS **MESF** NS	**Events/ml** **CD81= NS** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**	**Events/ml** **CD81= NS** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**
	HT-29 sEVs	11.70 ± 0.39	11.72 ± 1.01	17.67 ± 6.65	25.93 ± 9.74	**Events/ml** NS **MESF** NS	**Events/ml** **CD81= NS** **CD63= NS** **MESF** **CD81= NS** **CD63= NS**	

The asterisks (*) in the table denote significance levels.

The NS in the table indicates no significance.
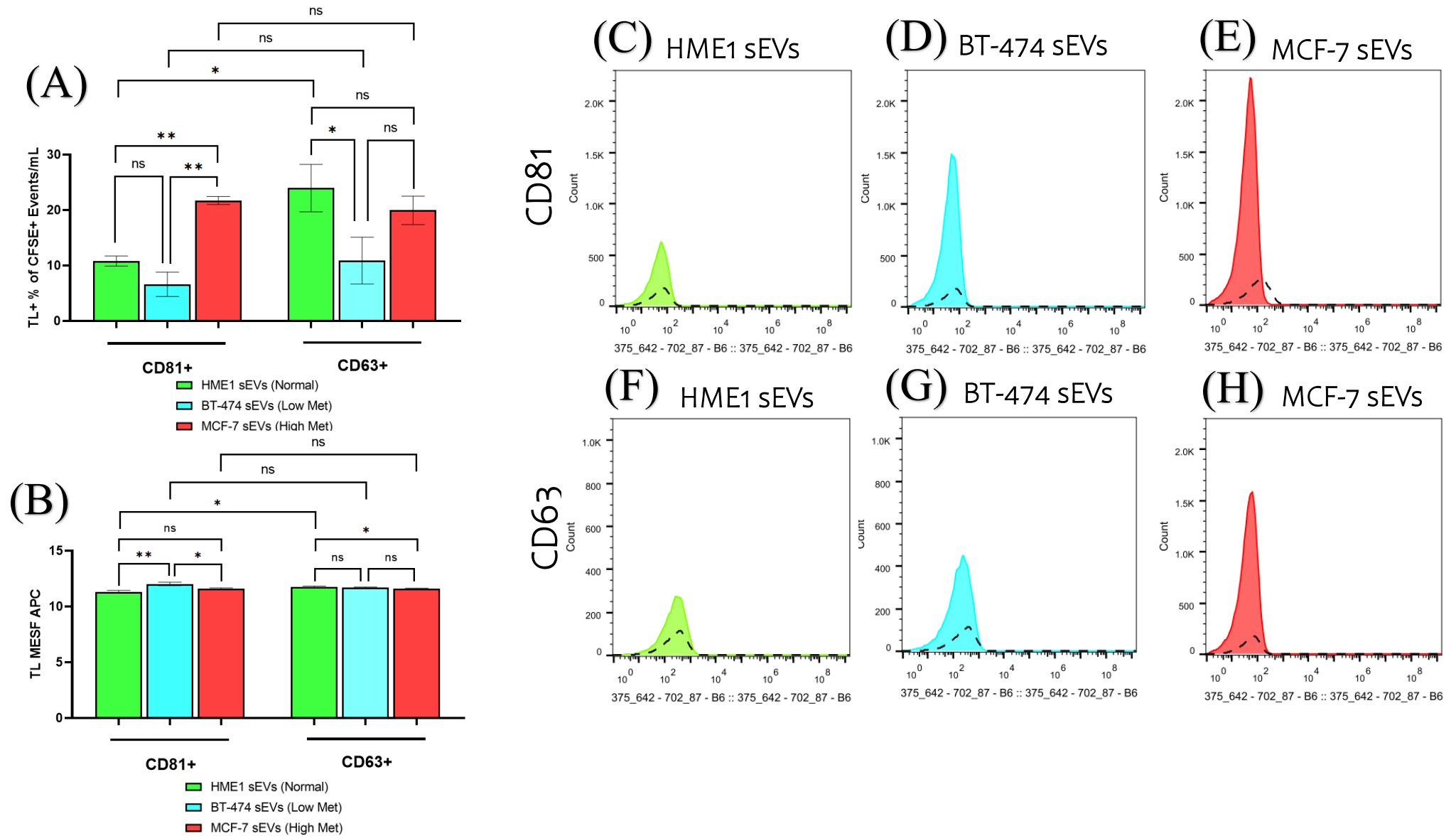
Figure 5.6. Single vesicle flow cytometry analysis of TL binding of CD81-positive or CD63-positive breast sEVs. (A) *Quantification of TL+ events/ml normalised to the percentage of CFSE events/ml for hTERT-HME1, BT-474, and MCF-7 CD81-postive or CD63-positive sEVs.* **(B)** *APC MESF values of TL+ for hTERT-HME1, BT-474, and MCF-7 sEVs with arbitrary units of MFI converted to standardised MESF units using vCal^TM nanorainbow beads.* **(C-E)** *Histogram profiles of the breast sEVs after gating showing TL lectin binding of CD81-positive sEVs fluorescence alongside the sugar-specific control fetuin outlined in black.* **(F-H)** *Histogram profiles of the breast sEVs after gating showing TL binding of CD63-positive sEVs fluorescence alongside the sugar-specific control fetuin outlined in black. Each experiment consisted of biological triplicates with three technical triplicates, and the error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. Welch's correction was used for parametric t-tests, and the Mann-Whitney test was used for nonparametric t-tests. Significance levels were set at *p < 0.05, **p < 0.01.*

Figure 5.7. Single vesicle flow cytometry analysis of TL binding of CD81-positive or CD63-positive colorectal sEVs. (A) *Quantification of TL+ events/ml normalised to the percentage of CFSE events/ml for NCM460, SW480, and HT-29 either CD81-postive or CD63-positive sEVs.* **(B)** *APC MESF values of TL+ for NCM460, SW480, and HT-29 sEVs with arbitrary units of MFI converted to standardised MESF units using vCal^TM nanorainbow beads.* **(C-E)** *Histogram profiles of the colorectal sEVs after gating showing TL binding of CD81-positive sEVs fluorescence alongside the sugar-specific control fetuin outlined in black.* **(F-H)** *Histogram profiles of the colorectal sEVs after gating showing TL lectin binding of CD63-positive sEVs fluorescence alongside the sugar-specific control fetuin outlined in black. Each experiment consisted of biological triplicates with three technical triplicates, and error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. For parametric t-tests, Welch's correction was utilised, and for nonparametric t-tests, a Mann-Whitney test was used.*

5.4.4. FLOW CYTOMETRY ANALYSIS OF LCA AND TL BINDING OF BREAST AND COLORECTAL EPITHELIAL CELLS

To identify LCA and TL binding to cell surface of the breast and colorectal epithelial cell models flow cytometry was adopted. The aim of this study was to assess whether the distinct LCA and TL binding patterns observed on sEVs, particularly those influenced by tetraspanin composition, were also evident at the cellular level.

The histograms, as depicted in Figure 5.8 A-F, show a notable increase in LCA and TL binding fluorescence across all cell lines compared to both the fetuin competitive inhibition control and the negative control. Interestingly, 'normal' hTERT-HME1 breast cells demonstrated a significant increase in LCA-binding with an average 14.7-fold and 6.3-fold MFI increase compared with BT-474 and MCF-7 breast cancer cells, respectively (Figure 5.8-G). Moreover, MCF-7 cells also demonstrated a significant increase in LCA-binding with an average 2.3-fold MFI increase compared with that of BT-474 cells. In the analysis of colorectal cell lines, SW480 cells exhibited a significant increase in LCA-binding by 1.8-fold MFI increase compared to HT-29 cells ($p<0.05$) (Figure 5.8-H).

Notably, hTERT-HME1 breast cells showed a substantial increase in TL-binding with an average 1.6-fold and 1.8-fold MFI increase compared to the cancer-associated cells BT-474 ($p<0.01$) and MCF-7 ($p<0.001$), respectively (Figure 5.9-G). Whereas the colorectal cells showed no significant differences in TL binding.
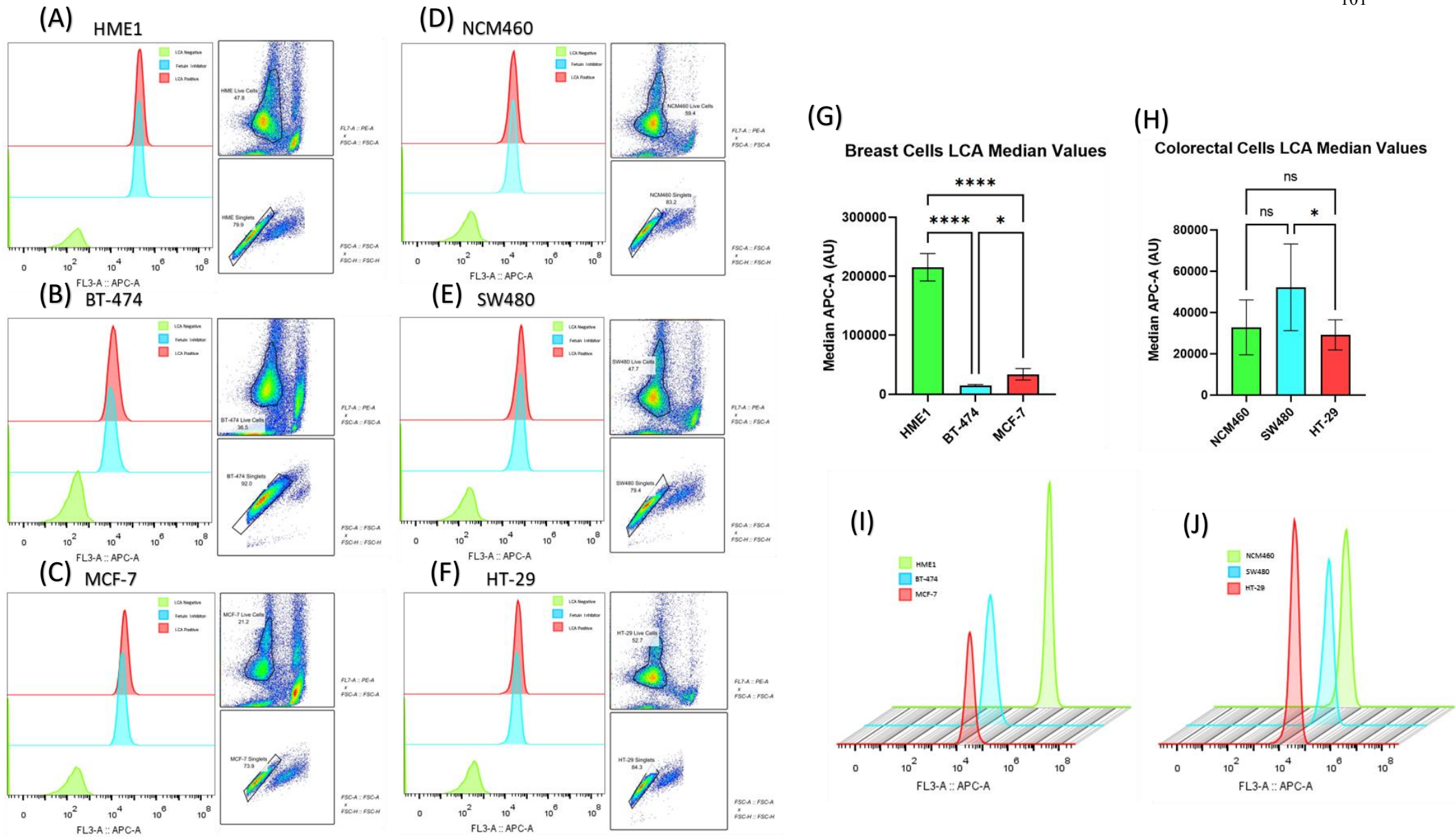
Figure 5.8. LCA lectin binding profiling in breast and colorectal epithelial cell lines using flow cytometry. (A-C) *Histogram profiles of breast cell lines showing LCA binding alongside ancestral gating of live and single-cell exclusion.* **(D-F)** *Colorectal cell lines.* **(G)** *Median fluorescence (AU) of breast cancer cell lines with LCA negative control (cells without LCA lectin), fetuin control (cells with LCA and sugar-specific control fetuin), and LCA positive control (cells with LCA to identify the presence of LCA binding glycans).* **(H)** *Median fluorescence (AU) of colorectal cell lines treated with LCA negative control, fetuin inhibitor control, and LCA positive control. A t-test followed by Welch's correction was used for statistical analysis, and significance was set at *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.* **(I)** *Staggered histogram comparison of colorectal cell lines positive for LCA.*
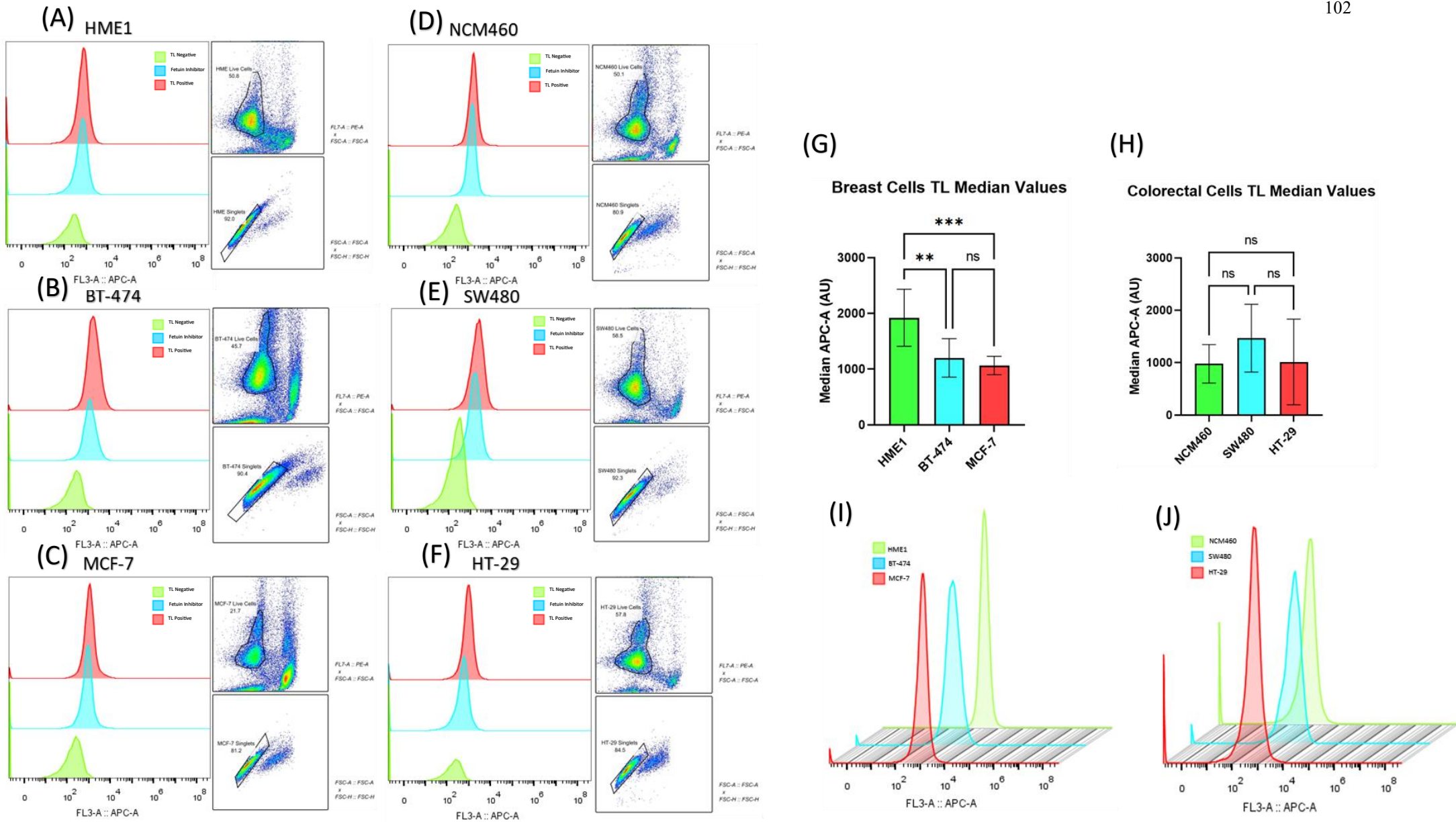
Figure 5.9. TL binding profiling in breast and colorectal epithelial cell lines by flow cytometry. (A-C) *Histogram profiles of breast cell lines showing TL binding alongside ancestral gating of live and single-cell exclusion.* **(D-F)** *Colorectal cell lines.* **(G)** *Median fluorescence (AU) of breast cancer cell lines with TL negative control (cells without TL), fetuin control (cells with TL and sugar-specific control fetuin), and LCA positive control (cells with TL to identify the presence of TL-binding glycans).* **(H)** *Median fluorescence (AU) of colorectal cell lines with TL negative, fetuin inhibitor, and TL positive controls. The t-test followed by Welch's correction was used for statistical analysis, and significance was set at *p < 0.05, **p < 0.01, and ***p < 0.001.* **(I)** *Staggered histogram comparison of breast cell lines for TL positivity.* **(J)** *Staggered histogram comparison of colorectal cell lines for TL positivity.*

5.5. DISCUSSION

5.5.1. LECTIN MICROARRAY ANALYSIS REVEALS INCREASED LECTIN BINDING OF BREAST CD81-POSITIVE CANCER-ASSOCIATED sEVs

To gain further insight into the glyco-heterogeneity of 'CD81-positive' sEVs across different phenotypes of breast sEVs, lectin microarray analysis was conducted. Initial observations of lectin binding among 'CD81-positive' breast sEVs revealed the absence of lectin binding to Sia, specifically N-acetylneuraminic acid (NANA), across all 'CD81-positive' breast sEVs. In contrast, studies have reported that Sia, namely α2,6-linked Sia, are enriched in cancer-associated sEVs (Escrevente et al., 2011; Gomes et al., 2015; Harada et al., 2019; Matsuda et al., 2020; Williams et al., 2019). A study by Matsuda et al. (2020) performed simultaneous glycomic analysis by lectin microarray analysis and lectin blotting of pancreatic cancer derived sEVs and identified unique surface glycomic signatures. The study compared lectin binding of membrane glycoproteins of lysed parental cells, as well as the supernatant taken from the pelleted sEVs denoted as the 'secreted glycoprotein fraction' before the pelleted sEVs were resuspended and further isolated by a commercially available sEVs isolation kit by precipitation. They found that the secreted glycoprotein fraction preferentially bound α2,6 Sia-binding lectins such as *Sambucus sieboldiana* agglutinin (SSA), *Sambucus nigra* agglutinin (SNA), and *Trichosanthes japonica* agglutinin. Consistent with this study, Matsuda et al. (2020) identified that the α2,6 Sia-binding lectin signals of sEVs were relatively low in pancreatic cancer-derived sEVs. In contrast, the secreted glycoproteins, which are often co-isolated with sEVs depending on the isolation method, such as ultracentrifugation, did in fact bind α2,6 Sia-binding lectins, thereby emphasising that depending on the isolation technique used, can lead to the enrichment of certain glycosylated structures.

In this study, six lectins exhibited higher binding affinities for cancer-associated 'CD81-positive' sEVs than for hTERT-HME1 'CD81-positive' sEVs. These included *Musa acuminata* lectin (BANLEC), wheat germ agglutinin (WGA), *Phaseolus vulgaris* leucoagglutinin (PHA-L), *Galanthus nivalis* agglutinin (GNA), *Urtica dioica* agglutinin (UDA), and *Tulipa* lectin (TL). This indicates a broad spectrum of glycan recognition, including α1,3-glycoside bonds (BANLEC), α-Man (GNA), α-GlcNAc (WGA and UDA), α-GalNAc (TL), and 2,6-branched tri- and tetra-antennary complex-type oligosaccharides (PHA-L). This suggests that these glycans may be notably increased in 'CD81-positive' breast cancer-associated sEVs compared with their normal counterparts. In particular, Man binding is interesting because other studies have reported that tumour-derived sEVs are enriched with specific Man-containing glycoproteins identified in breast, ovarian, melanoma, and glioblastoma cancer sEVs (Costa et al., 2018; Escrevente et al., 2011; Harada et al., 2019). Kondo et al. (2022) also established WGA binding to small-cell lung carcinoma H446 sEVs by lectin blotting and identified its enrichment on sEVs compared to cell membranes, thereby suggesting a potential commonality in α-GlcNAc glycan profiles across different types of cancer associated sEVs. The binding of PHA-L is also of interest because PHA-L binding to clinical tumour samples has been widely documented to be associated with poor prognosis in colorectal and breast cancers (Dennis et al., 1987; Fernandes et al., 1991). Matsuda et al. (2020) reported a strong association of PHA-L binding with pancreatic cancer cell-derived sEVs through lectin microarray analysis. Specifically, analysis of the CD81 and CD63 sEVs subpopulations revealed reduced PHA-L lectin signals in 'CD81-positive' sEVs,

whereas increased signals correlated with 'CD63-positive' sEVs. Furthermore, Freitas et al. (2019) detected faint PHA-L binding in gastric cancer MKN45 cell EVs by lectin blotting. However, they conducted the same blotting with sEVs cultured without FBS, and observed no binding. They suggested that the observed binding of branched N-glycans by PHA-L was likely nonspecific and possibly attributable to binding to BSA. These findings may explain the discrepancies observed by Matsuda et al. (2020) and this study, highlighting the importance of experimental conditions and emphasising the necessity of considering other underlying factors, such as potential non-specific interactions when interpreting lectin-binding results. Notably, TL significantly bound 'CD81-positive' cancer-associated sEVs compared to their normal counterparts. Based on this observation, it is interesting to note that the literature suggests that TL recognises a broad range of glycans. For example, TL is characterised as a superlectin with two distinct carbohydrate-binding domains that recognises sugars of varying structures, including GalNAc, lactose, Fuc, and Gal (Cammue et al., 1986). Studies have also shown that TL contains binding sites for Man (Barre et al., 1996). A recent machine learning analysis of TL binding conducted by Bojar et al. (2022) revealed that TL also binds to biantennary N-glycans with a significantly enhanced affinity in the presence of core fucosylation. In addition, TL can tolerate binding Sia substituents, fucosylation, GlcNAc, and LacNAc; however, bisecting GlcNAc negatively affects the binding affinity. TL also demonstrates a lower affinity for triantennary N-glycans, but exhibits a preference for β1,4-branched structures over β1,6-branched structures. While previous research has indicated a broad repertoire of glycans recognised by TL, no studies have investigated its binding to cancer-associated sEVs compared to normal sEVs. Therefore, future investigation is warranted to investigate this interaction and shed light on its potential implication on breast cancer progression

Lectin binding comparisons were also conducted between the 'CD81-positive' sEVs derived from the 'primary' BT-474 and 'metastatic' MCF-7 cells to identify lectins associated with both primary and metastatic cancer associated sEVs. These included *Robinia pseudoacacia* agglutinin (RPA), *Bauhinia purpurea* agglutinin (BPA), lentil lectin, *Salivia horminum* agglutinin, (SHA), *Agaricus bisporus* agglutinin (ABA) and *Allium sativum* agglutinin (ASA). These lectins recognise glycan structures comprising α-GalNAc (RPA), Gal(β1,3)GalNAc (BPA and ABA), D-Man, D-Glu, α-Fuc (lentil), α-GalNAc (SHA), and α-Man (ASA). It is worth mentioning that lentil lectin is the only lectin among those mentioned with reported applications in cancer diagnostics. Lentil lectin comprises of two discrete proteins, LCA-A and LCA-B (Howard et al., 1971). Both proteins are described as immunochemically indistinguishable and have identical haemagglutinin activities, and the term 'LCA' if often used in the literature interchangeably for both proteins. Studies have also shown LCA specificity towards Glu, Man, and glycoproteins with branched N-glycans, wherein GlcNAc is substituted into the $Man_3GlcNAc_2$ core and core α1,6-Fuc glycans (Casset et al., 1995; Debray et al., 1981; Howard et al., 1971b; Kornfeld et al., 1981; Kornfeld et al., 1971; Loris et al., 1994). LCA has been used in affinity chromatography for the separation of alpha-fetoprotein-L3 (AFP-L3), and the FDA has approved AFP-L3 as a tumour marker in human serum samples for primary hepatic cancer because of its high specificity and sensitivity for early diagnosis, differential diagnosis, and prognosis assessment (Malaguarnera et al., 2010; Song et al., 2002; Yamamoto et al., 2010; Zhang et al., 2016). However, the observations of increased binding of LCA to breast 'metastatic' 'CD81-positive' sEVs in this study is largely unexplored making this finding particularly interesting. Therefore, future investigations are also warranted to further explore this interaction and its potential implication on cancer metastasis.

5.5.2. VALIDATION OF CD81-POSITIVE sEVs FOR TL AND LCA BINDING BY SINGLE-VESICLE ANALYSIS

Consistent with the lectin microarray findings, LCA and TL demonstrated significantly increased binding to 'metastatic' MCF-7 'CD81-positive' sEVs in comparison to 'primary' BT-474 and 'normal' hTERT-HME1 sEVs. Interestingly, Li et al. (2021) utilised LCA lectin on magnetic beads to isolate cancer-associated sEVs derived from liver HepG2 cancer cells. They demonstrated that LCA captured approximately 60% of the total EVs from liver cancer HepG2 cells, resulting in a reduction of captured EVs to less than 40% in non-tumorigenic MIHA cells compared to BSA-immobilised beads. Without isolating different tetraspanin subpopulations, this study revealed the capability of LCA to recognise liver cancer-associated sEVs. It is also particularly interesting that LCA recognises α-Fuc, given its strong association with 'metastatic' MCF-7 'CD81-positive' sEVs. Liang et al. (2018) delved into this aspect by investigating the effects of knocking out α1,6-fucosyltransferase (FUT8), an enzyme responsible for catalysing the transfer of Fuc to N-linked-type complex glycopeptides, using a mouse model. Their findings revealed downregulation of CD81 expression, suggesting a pivotal role of core fucosylation in regulating CD81 tetraspanin expression. This intriguing association, coupled with the observed increase in LCA binding to 'CD81-positive' 'metastatic' MCF-7 sEVs, suggests the dysregulation of fucosylation patterns. This dysregulation may contribute to the heightened 'CD81-positive' LCA binding observed in 'metastatic' MCF-7 breast sEVs.

Matsuda et al. (2020) also conducted a glycome analysis of pancreatic cancer cell-derived sEVs using a lectin microarray that included LCA and TL, with a focus on tetraspanin antigens such as CD81 and CD63. However, the results revealed minor associations with the CD81 subpopulations for both lectins, whereas a significant increase in the association was observed with the CD63 subpopulations. Interesting, when extending the analysis of LCA binding in this study to 'CD63-positive' sEVs it revealed an increased binding to 'normal' hTERT-HME1 sEVs compared to that of BT-474 and MCF-7 'CD63-positive' sEVs. The contrasts in findings to this study are likely attributed to the different cancer types used further highlighting the complexity of glycan enrichment of sEVs between different cancers or different cell lines. Nevertheless, the observed differences in glycomic signatures in this study, particularly when focusing on various tetraspanin subpopulations of sEVs, raises the question of a subpopulation-specific biogenesis pathway for sEVs. This suggests that distinct tetraspanin-defined subsets of sEVs may undergo disparate sorting mechanisms during their formation and warrants further investigation.

5.5.3. FLOW CYTOMETRY ANALYSIS OF BREAST AND COLORECTAL CELLS FOR LCA AND TL BINDING

Remarkably, LCA binding in the cellular models of breast cancer revealed increased LCA lectin binding on 'normal' hTERT-HME1 cells compared to that of 'primary' BT-474 and 'metastatic' MCF-7 cells, thus mirroring the observations of 'CD63-positive' LCA binding at the sEV level. Cancer cells like BT-474 and MCF-7 cells often display aberrant glycosylation patterns on their cell surfaces, as confirmed by the HPA lectin cell surface analysis (section 4.4.1). In contrast, glycan structures in normal breast epithelial cells, such as hTERT-HME1,

typically exhibit basic bi- and tri-antennary patterns (reviewed by Scott & Drake, 2019). The observed discrepancy, where LCA displayed stronger binding to 'normal' hTERT-HME1 breast cells than to 'primary' BT-474 and 'metastatic' MCF-7 cells in this study, may be attributed to the increased glycan composition of the glycocalyx in cancer-associated cells. It's plausible that LCA's binding to 'normal' hTERT-HME1 breast cells was more pronounced due to the presence of additional glycans, potentially masking the recognition sites for LCA on BT-474 and MCF-7 cells.

In contrast, colorectal cells showed increased LCA binding to 'primary' SW480 cells compared to 'metastatic' HT-29 cells. The observed variation in LCA lectin binding among different breast and colorectal cell types highlights the intricate and heterogeneous nature of the glycan composition within these distinct cell phenotypes. However, it is essential to acknowledge the limitations of using a limited number of cell lines, which may not capture the diversity of cell surface glycan profiles. Similar observations were also made for TL binding at the cellular level in breast cancer cell lines, albeit with significantly lower detected fluorescence. This suggests that although TL binding was increased in 'normal' hTERT-HME1 cells, the abundance of TL-recognising glycans remained relatively low. Interestingly, the discrepancy between the low abundance of TL-recognising glycans on the cell membranes of the parental cells and the high abundance observed at the sEV level may suggest independent cargo sorting mechanisms for these particular glycans. However, such mechanisms have largely been unexplored, owing to the limited literature on this specific lectin and warrants further investigation to support this hypothesis.

5.6. KEY FINDINGS

- Elevated lectin binding of breast cancer-associated 'CD81-positive' sEVs (MCF-7 and BT-474) in comparison to 'normal' (hTERT-HME1) sEVs, including by the lectins BANLEC, WGA, PHA-L GNA, UDA and TL.

- 'Metastatic' breast 'CD81-positive' sEVs (MCF-7) exhibit heightened lectin binding compared to 'primary' (BT-474) sEVs, including by the lectins RPA, BPA, lentil, SHA, ABA, ASA.

- Elevated LCA binding of breast 'metastatic' 'CD81-positive' sEVs (MCF-7) in comparison to 'normal' (hTERT-HME1) and 'primary' (BT-474) sEVs.

- Increased LCA binding of breast 'normal 'CD63-positive' sEVs (hTERT-HME1) in comparison to 'primary' (BT-474) and 'metastatic' (MCF-7) sEVs.

- Increased LCA cell surface binding of 'normal' breast epithelial cells (hTERT-HME1) in comparison to 'primary' (BT-474) and 'metastatic' (MCF-7) cells.

- TL binding shows a significant increase in 'metastatic' breast 'CD81-positive' sEVs (MCF-7) compared to 'normal' (hTERT-HME1) and 'primary' (BT-474) sEVs.

- Enhanced TL cell surface binding is observed in 'normal' breast epithelial cells (hTERT-HME1) relative to 'primary' (BT-474) and 'metastatic' (MCF-7) cells.

CHAPTER 6

CLINICAL SIGNIFICANCE OF LECTIN BINDING

BREAST CANCER ASSOCIATED sEVs

6. CLINICAL SIGNIFICANCE OF LECTIN BINDING BREAST CANCER ASSOCIATED sEVs

6.1. BACKGROUND

The experimental findings reported in Chapters 4 and 5 revealed an increase in HPA, LCA, and TL lectin binding on breast 'metastatic' MCF-7 'CD81-positive' sEVs in comparison to 'primary' BT-474 and 'normal' hTERT-HME1 sEVs. Because sEVs can be isolated from a variety of body fluids, including blood, it was hypothesised that probing for these specific associations in breast cancer patient-derived sEVs may offer potential biomarkers for breast cancer diagnosis. Therefore, this chapter focuses on investigating the diagnostic application of these lectins in distinguishing blood sEVs derived from cancer patients from those derived from 'healthy' individuals.

6.1.1. PLASMA-ENRICHED sEVs

Unmanipulated blood can contain sEVs originating not only from leukocytes, platelets, erythrocytes, and endothelial cells but also from various other cell types (reviewed by Alberro et al., 2021; Webber & Clayton, 2013; Yáñez-Mó et al., 2015). This diverse mixture of sEVs may pose challenges for downstream analysis especially with the aim of specifically identifying cancer-associated sEVs. Alternatively, serum and plasma are promising sources of sEVs, offering a promising avenue for identifying valuable biomarkers crucial for disease diagnosis, prognosis, and monitoring (reviewed by Hu et al., 2021; Mustapic et al., 2017). More specifically, plasma has been suggested to serve as a more appropriate source of sEVs because additional platelet-derived EVs are released during clot formation during serum preparation (reviewed by Coumans et al., 2017). Therefore, plasma was chosen as the source of sEVs in this study. Additionally, given the more promising results of lectin binding showing increased association with breast cancer in epithelial models compared to colorectal models, plasma from breast cancer patients was specifically chosen for investigation. sEVs were isolated from plasma samples derived from 'healthy' individuals and patients with 'non-metastatic' and 'metastatic' breast cancer by SEC and subjected to standardised characterisation by NTA, western blotting key markers, and TEM.

6.1.2. ADDITIONAL TARGETS OF PLASMA-ENRICHED sEVs ASSOCIATED WITH HPA LCA AND TL

The experimental findings in Chapters 4 and 5 revealed an increase in HPA, LCA, and TL lectin binding of 'CD81-positive' sEVs, with 'metastatic' MCF-7 sEVs. However, given the highly heterogeneous nature of these vesicles, concerns were raised regarding the specificity of tetraspanin detection in relation to the specific detection of plasma-enriched sEVs. This heterogeneity poses a challenge because of the variable expression of tetraspanins such as CD81 across diverse sources of sEVs. Such variability can impede the detection of specific cancer-associated sEVs of interest, thereby hindering the assessment of glycans recognised by these lectins for diagnostic capability assessment. In light of these challenges, the focus shifted to identifying additional glycosylated targets of plasma-enriched sEVs specifically associated with cancer-associated sEVs, with the aim of enhancing the diagnostic capability of these lectins. This was achieved by implementing the MACSPlex analysis exosome kit, which facilitates the identification of binding subpopulations of HPA, LCA, and TL in plasma-enriched sEVs, specifically those derived from patients with breast cancer in comparison to 'healthy' individuals.

6.1.3. THE DIAGNOSTIC APPLICATION OF PLASMA-ENRICHED sEVs AND LECTINS AS BLOOD-BASED BIOMARKERS IN CANCER

Interestingly, the majority of serum cancer biomarkers currently used in the clinic are glycoproteins for monitoring responses to treatment and disease recurrence in patients with breast cancer such as CA15-3 (section 1.6.2) (reviewed by Hashim et al., 2017). Although the exploitation of glycoproteins has proven invaluable for the detection and monitoring of cancer, studies have also shown the potential of sEVs surface glycans as an alternative biomarker source. For example, Yokose et al. (2020) identified an increase in O-glycosylated sEVs in the serum of patients with early stage pancreatic cancer, even when the samples were negative for the established CA 19-9 blood antigen test. Specifically, they observed a notable increase in the binding of lectins *Agaricus bisporus* agglutinin (ABA) and *Amaranthus caudatus* agglutinin (ACA) to O-glycans. The serum of 117 patients with pancreatic cancer exhibited a significant increase in ABA- and ACA-positive sEVs compared to 98 normal controls, with corresponding AUC values of 0.838 and 0.810, respectively, on the ROC curve. This study underscores the potential of sEVs surface glycans derived from blood samples as an alternative biomarker source for cancer diagnostics. Therefore, to investigate the diagnostic capability of HPA, LCA, and TL in distinguishing plasma-enriched sEVs derived from breast cancer patients and those derived from 'healthy' individuals, single-vesicle flow cytometry was used.

6.1.4. EpCAM DETECTION OF BREAST CANCER-ASSOCIATED sEVs

Epithelial cell adhesion molecule (EpCAM) is frequently found to be overexpressed in human breast cancer (Tandon et al., 1990). EpCAM is a transmembrane glycoprotein that facilitates cell-to-cell adhesion within epithelial tissues (reviewed by Keller et al., 2019). EpCAM overexpression has been reported to be an independent prognostic marker for poor overall survival in node-positive breast cancer (Gastl et al., 2000). Over the past decade, there has been a significant interest in the detection and molecular characterisation of EpCAM-positive CTCs. As the most extensively studied and only FDA-cleared system for CTC detection and analysis is the CellSearch® system, which isolates CTCs from whole blood by using ferrofluid coated with EpCAM antibodies (Krebs et al., 2010). Despite the FDA approval for metastatic breast cancer prognosis, its efficacy in early-stage cancer detection is hindered by the rarity and antigenic diversity of CTCs (Cohen et al., 2008; de Bono et al., 2008). Given these limitations, researchers have sought alternative targets such as plasma-enriched sEVs. Interestingly, studies have highlighted a notable increase in the abundance of 'EpCAM-positive' sEVs in plasma-enriched samples derived from patients with breast cancer compared to those from 'healthy' individuals (Tian et al., 2021; Vitale et al., 2021). As EpCAM is characterised as a glycoprotein, it raises the question of whether the introduction of lectins enhances its diagnostic capability. To further investigate the diagnostic capability of HPA, LCA, and TL in distinguishing plasma-enriched sEVs derived from patients with breast cancer from those derived from 'healthy' individuals, EpCAM was utilised as an additional target alongside these lectins.

6.2. AIMS AND OBJECTIVES

The aim of this chapter is to explore the diagnostic capability of HPA, LCA and TL in distinguishing between plasma-enriched sEVs derived from 'healthy' individuals and those from breast cancer patients.

The objectives were:

1) To isolate plasma-enriched sEVs from 'healthy' individuals and perform characterisation of each SEC fraction by NTA and BCA assay to identify fractions that most likely contain sEVs.

2) To characterise the pooled fraction that most likely contained plasma-enriched sEVs by NTA, western blotting key sEVs markers and TEM.

3) Assess the diagnostic potential of HPA, LCA, and TL binding of 'EpCAM-positive' sEVs to distinguish between plasma-enriched sEVs from 'healthy' individuals and those from cancer patients, from patients with 'non-metastatic' and 'metastatic' stages of cancer.

6.3. METHODS

6.3.1. EXTRACTION AND CHARACTERISATION OF PLASMA-ENRICHED sEVs

6.3.1.1. EXTRACTION OF PLASMA-ENRICHED sEVs

sEVs were isolated from patient samples by implementing SEC as detailed in section 2.4.2.2 and shown in Figure 2.1.

6.3.1.2. NTA AND BCA ASSAY

NTA and BCA assay was performed on the SEC fractions to confirm effective isolation of the pooled fractions to assess their concentration and size, as described in section 2.5.1 and section 2.4.

6.3.1.3. WESTERN BLOTTING KEY sEVs MARKERS

Western blotting was performed on the pooled SEC fraction to confirm the presence of plasma-enriched sEVs by blotting for key markers of sEVs (Table 2.4), as described in section 2.5.2.

6.3.1.4. TEM

TEM was performed on the pooled SEC fractions to identify the morphological and size characteristics of the plasma-enriched sEVs as described in section 2.5.3.

6.3.2. SINGLE-VESICLE FLOW CYTOMETRY

6.3.2.1. HPA, LCA AND TL PROFILING OF PLASMA-ENRICHED sEVs

Single-vesicle flow cytometry was used to quantify HPA, LCA, and TL binding of plasma-enriched sEVs derived from 'healthy' individuals and patients with 'non-metastatic' and 'metastatic' breast cancer. The methodology for isolating plasma-enriched sEVs from plasma samples is outlined in section 2.4.2. The experimental procedure (section 2.6) includes instrument settings (section 2.6.1), calibration for sizing and fluorescence before sEV acquisition (section 2.6.2), and staining and labelling steps for plasma-enriched sEVs with antibodies and lectins (section 2.6.3).

6.4. RESULTS

6.4.1. CHARACTERISATION OF PLASMA ENRICHED sEVs

To confirm the effectiveness of SEC for isolating plasma-enriched sEVs, a series of characterisations was conducted. NTA analysis of the patient plasma-enriched sEVs revealed peaks in particle number in SEC fractions 5, 6, and 7, while also demonstrating the lowest protein contamination in the BCA assay (Figure 6.1-F). Therefore, these fractions were pooled to maximise yield and obtain a single sEV sample for further characterisation by NTA, western blotting, and TEM. Further, NTA indicated that the median particle size across all donor groups fell within the 30-150 nm range, a size characteristic of sEVs (Figure 6.1 B-D). More precisely, the average size of plasma-derived sEVs varied by cohort, with 'healthy' individuals showing an average sEV size of 75 nm, whereas those with 'non-metastatic' and 'metastatic' cancer exhibited an average sEV size of 105 nm. The same key markers employed for the western blotting analysis of epithelial cell line-derived sEVs were similarly utilised for the identification of plasma-enriched sEVs. These markers included CD81, CD63 and syntenin-1 (reviewed by Jankovičová et al., 2020). Additionally, the ultralow density lipoprotein marker (LDL), ApoB, was included in the analysis (Sódar et al., 2016). To confirm the specificity of antibody binding and ensure that the observed bands were related to the proteins of interest, a H_2O control was included, revealing no bands for the antibodies used. Western blotting of the plasma-enriched sEVs revealed that CD81 and syntenin-1 were not detectable, whereas bands were only visible for CD63 and ApoB (Figure 6.1-E). Further analysis of the pooled plasma-enriched sEVs fractions by TEM revealed structures with the typical cup-shaped morphology expected for sEVs, in the size range of 30-150 nm (Figure 6.1-A). However, LDL particles were also detected, suggesting their potential to mimic sEVs and co-isolates during the isolation procedure.
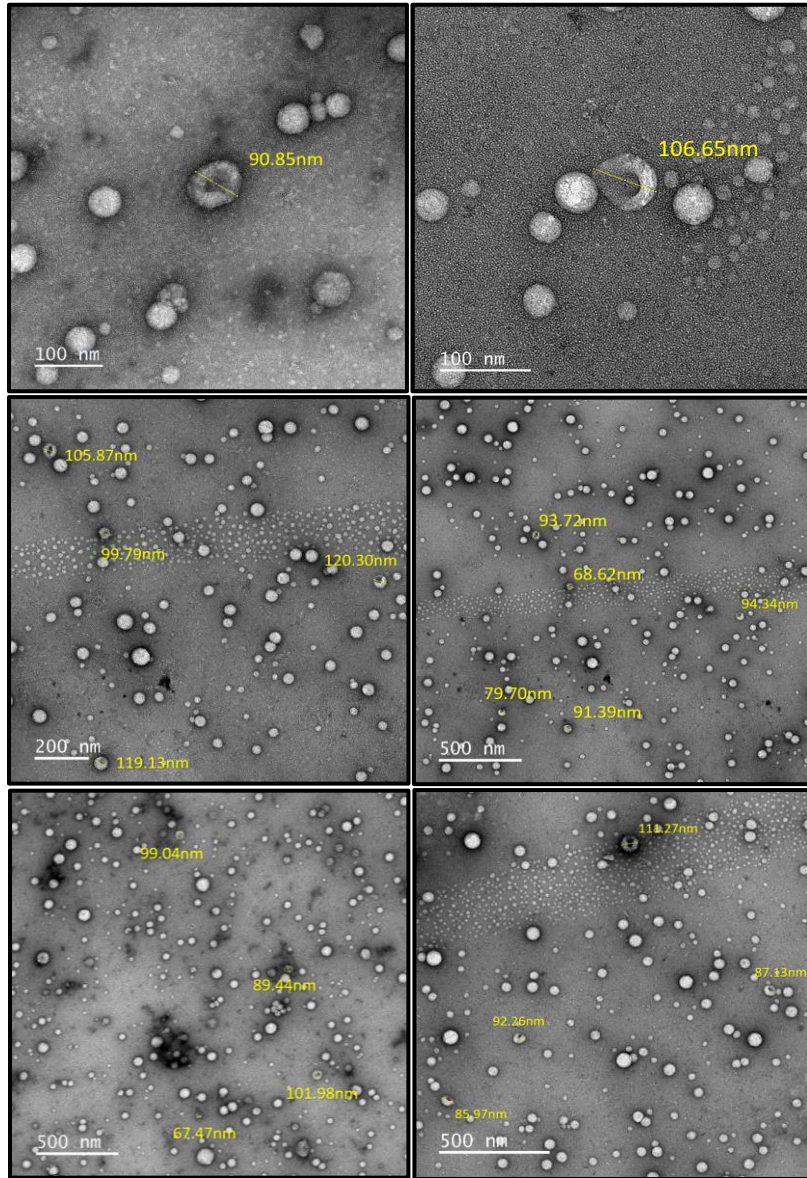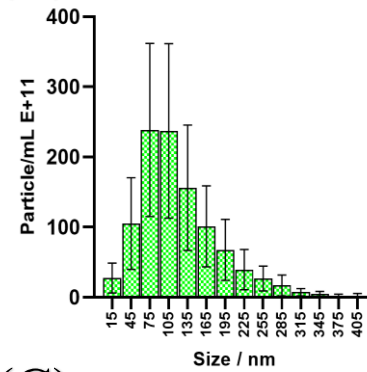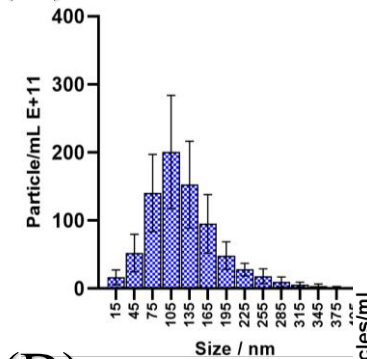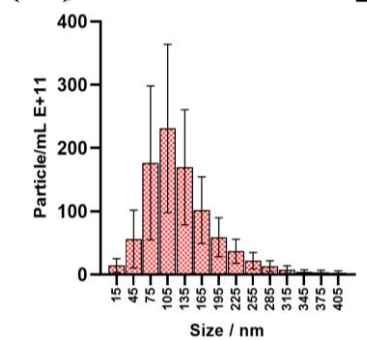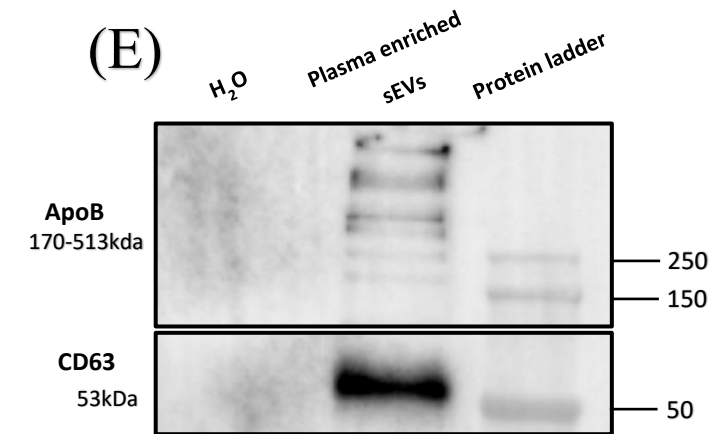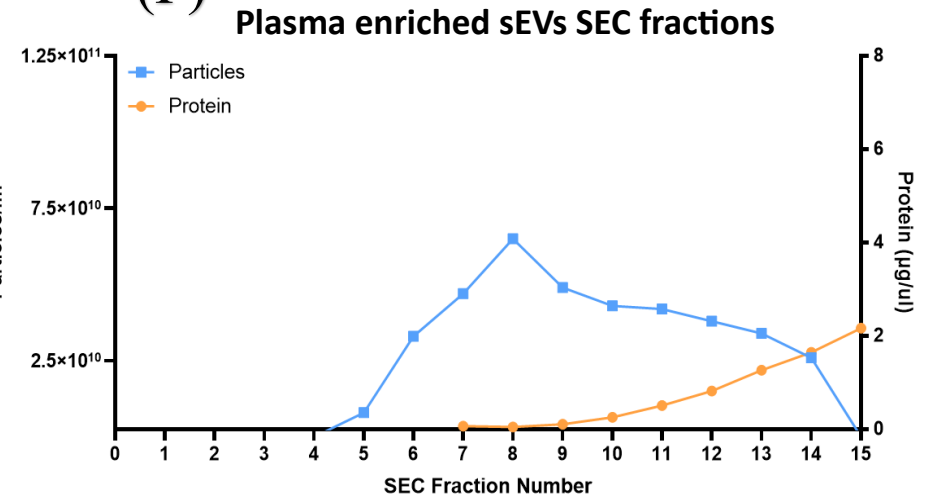
Figure 6.1. Characterisation of SEC fractions for the isolation of plasma enriched sEVs by NTA BCA TEM and western blotting. **(A)** *TEM analysis of plasma-enriched sEVs from fractions 5, 6, and 7 pooled together. (Scale bars = 100, 200, and 500 nm).* **(B-D)** *NTA size characterisation of plasma-enriched sEVs from pooled fractions 5, 6, and 7.* **(E)** *Western blotting of the sEVs protein marker CD63 and lipoprotein marker ApoB.* **(F)** *SEC fractions (500ul) containing plasma-enriched sEVs were subjected to NTA and BCA analysis to determine particle and protein concentrations to identify fractions which most likely contained sEVs.*

6.4.2. EpCAM ANALYSIS OF PLASMA-ENRICHED sEVs

In continuation of single-vesicle flow cytometry, the investigation was extended to evaluate EpCAM on the surface of plasma-enriched sEVs across the various patient cohorts (Figure 6.2). Primary data, including the mean ± SD of EpCAM binding normalised to CFSE events/ml and MESF values, are presented in Table 6.1.

The initial assessment of EpCAM abundance revealed a significant 3.3.3-fold increase in the plasma-enriched sEVs derived from breast cancer patients in comparison to 'healthy' individuals ($p < 0.0001$) (Figure 6.2 A-B). ROC curve analysis underscored the diagnostic potential of EpCAM, with an AUC value of 0.9474 for differentiating EpCAM-postive plasma-enriched sEVs derived from breast cancer patients in comparison to 'healthy' individuals (Figure 6.2-C). Analysis of EpCAM binding across the different patient cohorts revealed a 3.3-fold increase in plasma-enriched sEVs derived from 'non-metastatic' breast cancer patients ($p < 0.0001$) and a 3.4-fold increase in those from 'metastatic' breast cancer patients ($p < 0.001$) in comparison with 'healthy' individuals' samples (Figure 6.2-D). Analysis of EpCAM MESF values revealed a significant 1.3-fold increase in plasma-enriched sEVs derived from 'healthy' individuals in comparison to those of 'non-metastatic' ($p < 0.01$) and 'metastatic' breast cancer patients ($p < 0.05$) (Supplementary Figure 19). Comparative analysis of EpCAM MESF values also indicated a 1.3-fold increase in plasma-enriched sEVs derived from 'healthy' plasma-enriched sEVs in comparison to breast cancer- associated patients ($p < 0.01$) (Supplementary Figure 19). Interestingly, ROC curve analysis of EpCAM MESF values demonstrated an AUC value of 0.8842, for differentiating plasma-enriched sEVs derived from breast cancer patients in comparison to 'healthy' individuals (Supplementary Figure 19).

Table 6.1. EpCAM binding of plasma enriched sEVs derived from healthy individuals, non-metastatic and metastatic breast cancer patients.

Plasma enriched sEVs	EpCAM PE MESF (mean ± SD)	EpCAM events/ml % of CFSE events/ml (mean ± SD)	Vs healthy sEVs	Vs metastatic sEVs
Healthy individuals sEVs	210.755 ± 39.895	0.527 ± 0.170		**Events/ml** *** **MESF** *
Non-metastatic cancer patient sEVs	166.051 ± 6.218	1.773 ± 0.551	**Events/ml** **** **MESF** **	**Events/ml** NS **MESF** NS
Metastatic cancer patient sEVs	150.676 ± 38.430	1.757 ± 0.918	**Events/ml** *** **MESF** *	

The asterisks (*) in the table denote significance levels.

The NS in the table indicates no significance.

Figure 6.2. Single-vesicle analysis of EpCAM detection of plasma-enriched sEVs from metastatic breast cancer patients (n = 5), non-metastatic breast cancer patients (n = 5), and healthy individuals (n = 5). (A) *EpCAM+ events/ml normalised to % CFSE events/ml for each healthy and cancer-associated plasma enriched sEVs sample.* (B) *Comparison of EpCAM+ events/ml normalised to % CFSE events/ml between plasma-enriched sEVs derived from healthy individuals and cancer patients.* (C) *ROC curve analysis measuring EpCAM+ events/ml normalised to % CFSE events/ml, yielding an AUC of 0.9474 for discriminating patients with cancer from healthy individuals.* (D) *Comparison of EpCAM+ events/ml normalised to % CFSE events/ml between plasma-enriched sEVs derived from healthy individuals and non-metastatic and metastatic patients.* (E-G) *Histogram profiles of plasma-enriched sEVs after gating showing EpCAM fluorescence. Each experiment comprised biological triplicates with three technical triplicates; error bars represent the Stdev. The t-test followed by Welch's correction was used to determine the significance level: *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.*

6.4.3. EpCAM ANALYSIS OF PLASMA-ENRICHED sEVs WITH FLUORESCENTLY LABELLED HPA

Prior results have already established the efficacy of using EpCAM detection via single-vesicle flow cytometry on plasma-enriched sEVs for distinguishing cancer patients from 'healthy' individuals (section 6.4.2). Therefore, in continuation of adopting single-vesicle flow cytometry, the investigation shifted to explore the potential of HPA lectin binding in 'EpCAM-positive' plasma-enriched sEVs samples, with the aim of increasing the diagnostic potential for identifying samples from 'healthy' individuals and those with cancer (Figure 6.3). The primary data, including the mean ± SD of HPA lectin binding of 'EpCAM-positive' plasma-enriched sEVs normalised to CFSE events/ml and MESF values, are presented in Table 6.2.

Before diagnostically assessing HPA lectin binding to 'EpCAM-positive' sEVs, the study first explored the potential of HPA lectin alone to differentiate between plasma-enriched sEVs derived from 'healthy' individuals and those from breast cancer patients (Supplementary Figure 18). This preliminary analysis revealed a 1.2-fold increase in HPA lectin binding to plasma-enriched sEVs derived from 'healthy' individuals compared to those with cancer ($p < 0.05$) (Supplementary Figure 18). Specifically comparing HPA lectin binding of EpCAM-postive plasma-enriched sEVs revealed a 3.3-fold increase in cancer patients in comparison to 'healthy' individuals ($p < 0.0001$). ROC curve analysis revealed an AUC value of 0.9474 in distinguishing HPA lectin binding of EpCAM-postive plasma-enriched sEVs derived from 'healthy' individuals and breast cancer patients (Figure 6.3 A-C). Specifically, HPA lectin binding of 'EpCAM-positive' plasma enriched sEVs revealed a 3.1-fold and 3.4-fold increase from those derived in 'non-metastatic' and 'metastatic' breast cancer patients in comparison to 'healthy' individuals ($p < 0.0001$ and $p < 0.001$, respectively) (Figure 6.3-D). Moreover, analysis of MESF values for HPA lectin binding of 'EpCAM-positive' plasma-enriched sEVs derived from breast cancer patients and 'healthy' individuals revealed no significant difference (Supplementary Figure 20).

Table 6.2. HPA lectin binding of EpCAM-positive plasma enriched sEVs derived from healthy individuals, non-metastatic and metastatic breast cancer patients.

Plasma enriched sEVs	HPA APC MESF (mean ± SD)	HPA events/ml % of CFSE events/ml (mean ± SD)	Vs healthy sEVs	Vs metastatic sEVs
Healthy individuals sEVs	63.748 ± 26.446	0.341 ± 0.168		**Events/ml** *** **MESF** NS
Non-metastatic cancer patient sEVs	76.597 ± 9.989	1.106 ± 0.296	**Events/ml** **** **MESF** NS	**Events/ml** NS **MESF** NS
Metastatic cancer patient sEVs	79.633 ± 27.297	1.153 ± 0.624	**Events/ml** *** **MESF** NS	

The asterisks (*) in the table denote significance levels.

The NS in the table indicates no significance.
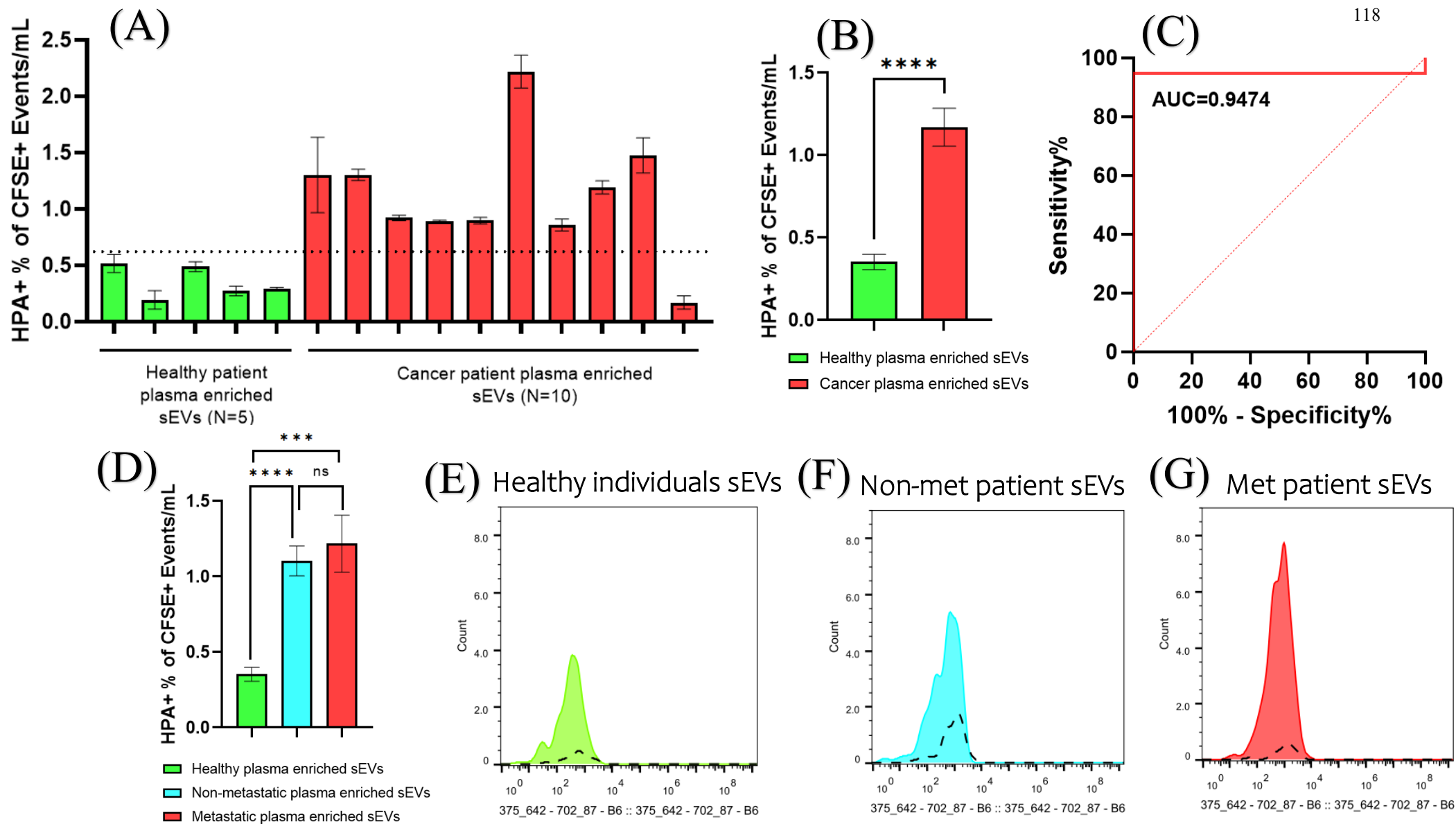
Figure 6.3. Single-vesicle analysis of HPA lectin detection of EpCAM-positive plasma-enriched sEVs from metastatic breast cancer patients (n = 5), non-metastatic breast cancer patients (n = 5), and healthy individuals (n = 5). (A) *HPA+ events/ml normalised to % CFSE events/ml for each healthy and cancer-associated plasma enriched sEVs sample.* **(B)** *Comparison of HPA+ events/ml normalised to % CFSE events/ml between plasma-enriched sEVs derived from healthy individuals and cancer patients.* **(C)** *ROC curve analysis measuring HPA+ events/ml normalised to % CFSE events/ml, yielding an AUC of 0.9474 for discriminating patients with cancer from healthy individuals.* **(D)** *Comparison of HPA+ events/ml normalised to % CFSE events/ml between plasma-enriched sEVs derived from healthy individuals and non-metastatic and metastatic cancer patients.* **(E-G)** *Histogram profiles of plasma-enriched sEVs after gating showing HPA fluorescence. Each experiment comprised biological triplicates with three technical triplicates; error bars represent the Stdev. The t-test followed by Welch's correction was used to determine the significance levels: *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.*

6.4.4. EpCAM ANALYSIS OF PLASMA-ENRICHED sEVs WITH FLUORESCENTLY LABELLED LCA

Previous findings have confirmed the effectiveness of utilising EpCAM detection through single-vesicle flow cytometry to differentiate between cancer patients and 'healthy' individuals (section 6.4.2). Building on this, the current investigation has transitioned towards examining the feasibility of LCA lectin binding in 'EpCAM-positive' plasma-enriched sEV samples. The primary data, including the mean ± SD of LCA lectin binding of 'EpCAM-positive' plasma-enriched sEVs normalised to CFSE events/ml and MESF values, are presented in Table 6.3.

Before diagnostically assessing LCA lectin binding to 'EpCAM-positive' sEVs, the study first explored the potential of LCA lectin alone to differentiate between plasma-enriched sEVs derived from 'healthy' individuals and breast cancer patients (Supplementary Figure 18). This preliminary analysis revealed a 1.1-fold increase in LCA lectin binding to plasma-enriched sEVs derived from 'healthy' individuals compared to those from cancer patients ($p < 0.05$) (Supplementary Figure 18). Specifically comparing LCA lectin binding of 'EpCAM-positive' plasma-enriched sEVs revealed a 2.9-fold increase in cancer-associated patients in comparison to 'healthy' individuals ($p < 0.01$). ROC curve analysis revealed an AUC value of 0.8 in distinguishing LCA lectin binding of 'EpCAM-positive' plasma-enriched sEVs derived from 'healthy' individuals and breast cancer patients (Figure 6.4 A-C). Specifically, LCA lectin binding of 'EpCAM-positive' plasma enriched sEVs revealed a 2.4-fold increase from those derived in 'non-metastatic' breast cancer patients in comparison to 'healthy' individuals ($p < 0.01$) (Figure 6.4-D). Whereas no significant differences were observed for LCA binding 'EPCAM-positive' plasma-enriched sEVs derived from 'metastatic' breast cancer patients in comparison to 'non-metastatic' breast cancer patients and 'healthy' individuals. Moreover, analysis of MESF values for LCA lectin binding of 'EpCAM-positive' plasma-enriched sEVs revealed a significant 1.4-fold and 1.3-fold increase in those derived from 'non-metastatic' and 'metastatic' breast cancer patients respectively, in comparison to 'healthy' individuals ($p < 0.01$). Whereas ROC curve analysis of MESF values demonstrated an AUC value of 0.6925 in distinguishing LCA lectin binding of EpCAM-postive plasma-enriched sEVs derived from 'healthy' individuals and breast cancer patients (Supplementary Figure 21).

Table 6.3. LCA lectin binding of EpCAM-positive plasma enriched sEVs derived from healthy individuals, non-metastatic and metastatic breast cancer patients.

Plasma enriched sEVs	LCA APC MESF (mean ± SD)	LCA events/ml % of CFSE events/ml (mean ± SD)	Vs healthy sEVs	Vs metastatic sEVs
Healthy individuals sEVs	39.632 ± 10.676	1.427 ± 1.747		**Events/ml** NS **MESF** NS
Non-metastatic cancer patient sEVs	52.227 ± 6.824	3.584 ± 0.489	**Events/ml** ** **MESF** **	**Events/ml** NS **MESF** **
Metastatic cancer patient sEVs	42.122 ± 8.300	4.346 ± 3.096	**Events/ml** NS **MESF** NS	

The asterisks (*) in the table denote significance levels.

The NS in the table indicates no significance.
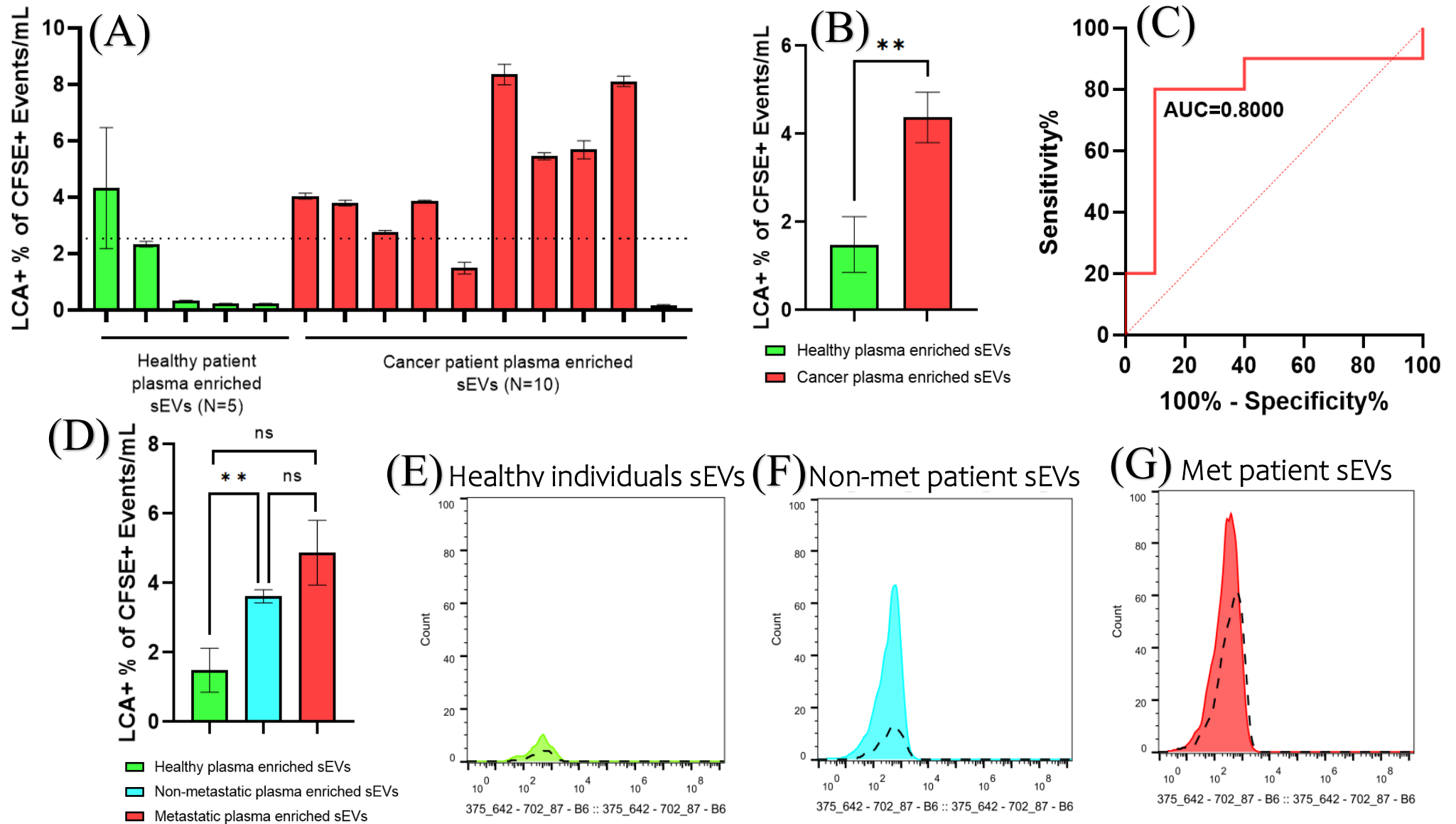
Figure 6.4. Single-vesicle analysis of LCA lectin detection of EpCAM-positive plasma-enriched sEVs from metastatic breast cancer patients (n = 5), non-metastatic breast cancer patients (n = 5), and healthy individuals (n = 5). (A) *LCA+ events/ml normalised to % CFSE events/ml for each healthy and cancer-associated plasma enriched sEVs sample.* **(B)** *Comparison of LCA+ events/ml normalised to % CFSE events/ml between plasma-enriched sEVs derived from healthy individuals and cancer patients.* **(C)** *ROC curve analysis measuring LCA+ events/ml normalised to % CFSE events/ml, yielding an AUC of 0.8 for discriminating patients with cancer from healthy individuals.* **(D)** *Comparison of LCA+ events/ml normalised to % CFSE events/ml between plasma-enriched sEVs derived from healthy individuals and non-metastatic and metastatic cancer patients.* **(E-G)** *Histogram profiles of plasma-enriched sEVs after gating showing LCA fluorescence. Each experiment comprised biological triplicates with three technical triplicates; error bars represent the Stdev. The t-test followed by Welch's correction was used to determine the significance levels. *p < 0.05, **p < 0.01.*

6.4.5. EpCAM ANALYSIS OF PLASMA-ENRICHED sEVs WITH FLUORESCENTLY LABELLED TL

Prior investigation has established the efficacy of EpCAM detection via single-vesicle flow cytometry for distinguishing between cancer patients and 'healthy' individuals (section 6.4.2). Expanding on this, the current study has shifted its focus to investigate the potential of TL binding on 'EpCAM-positive' plasma-enriched sEVs samples. The primary data, including the mean ± SD of TL binding of 'EpCAM-positive' plasma-enriched sEVs normalised to CFSE events/ml and MESF values, are presented in Table 6.4.

Before diagnostically assessing TL lectin binding to 'EpCAM-positive' sEVs, the study first explored the potential of TL alone to differentiate between plasma-enriched sEVs derived from 'healthy' individuals and those from breast cancer patients, however no significant differences were observed (Supplementary Figure 18). Specifically comparing TL binding of EpCAM-postive plasma-enriched sEVs revealed a 6.4-fold increase in cancer patients in comparison to 'healthy' individuals ($p < 0.001$). ROC curve analysis revealed an AUC value of 0.8889 in distinguishing TL binding of EpCAM-postive plasma-enriched sEVs derived from 'healthy' individuals and breast cancer patients (Figure 6.5 A-C). Specifically, TL binding of 'EpCAM-positive' plasma enriched sEVs revealed a 4.8-fold and 7.7-fold increase from those derived from 'non-metastatic' and 'metastatic' breast cancer patients in comparison to 'healthy' individuals ($p < 0.0001$ and $p < 0.001$, respectively) (Figure 6.5-D). Moreover, analysis of MESF values for TL binding of 'EpCAM-positive' plasma-enriched sEVs revealed a significant 1.1-fold increase in those derived from 'non-metastatic' and 'metastatic' breast cancer patients, in comparison to 'healthy' individuals ($p < 0.01$). Whereas ROC curve analysis of MESF values demonstrated an AUC value of 0.5278 in distinguishing TL binding of 'EpCAM-postive' plasma-enriched sEVs derived from 'healthy' individuals and breast cancer patients (Supplementary Figure 22).

Table 6.4. TL binding of EpCAM-positive plasma enriched sEVs derived from healthy individuals, non-metastatic and metastatic breast cancer patients.

Plasma enriched sEVs	TL APC MESF (mean ± SD)	TL events/ml % of CFSE events/ml (mean ± SD)	Vs healthy sEVs	Vs metastatic sEVs
Healthy individuals sEVs	13.156 ± 0.392	1.302 ± 0.805		**Events/ml** * **MESF** NS
Non-metastatic cancer patient sEVs	14.287 ± 0.869	6.034 ± 1.286	**Events/ml** **** **MESF** **	**Events/ml** NS **MESF** **
Metastatic cancer patient sEVs	12.557 ± 1.627	9.734 ± 6.047	**Events/ml** * **MESF** NS	

The asterisks (*) in the table denote significance levels.

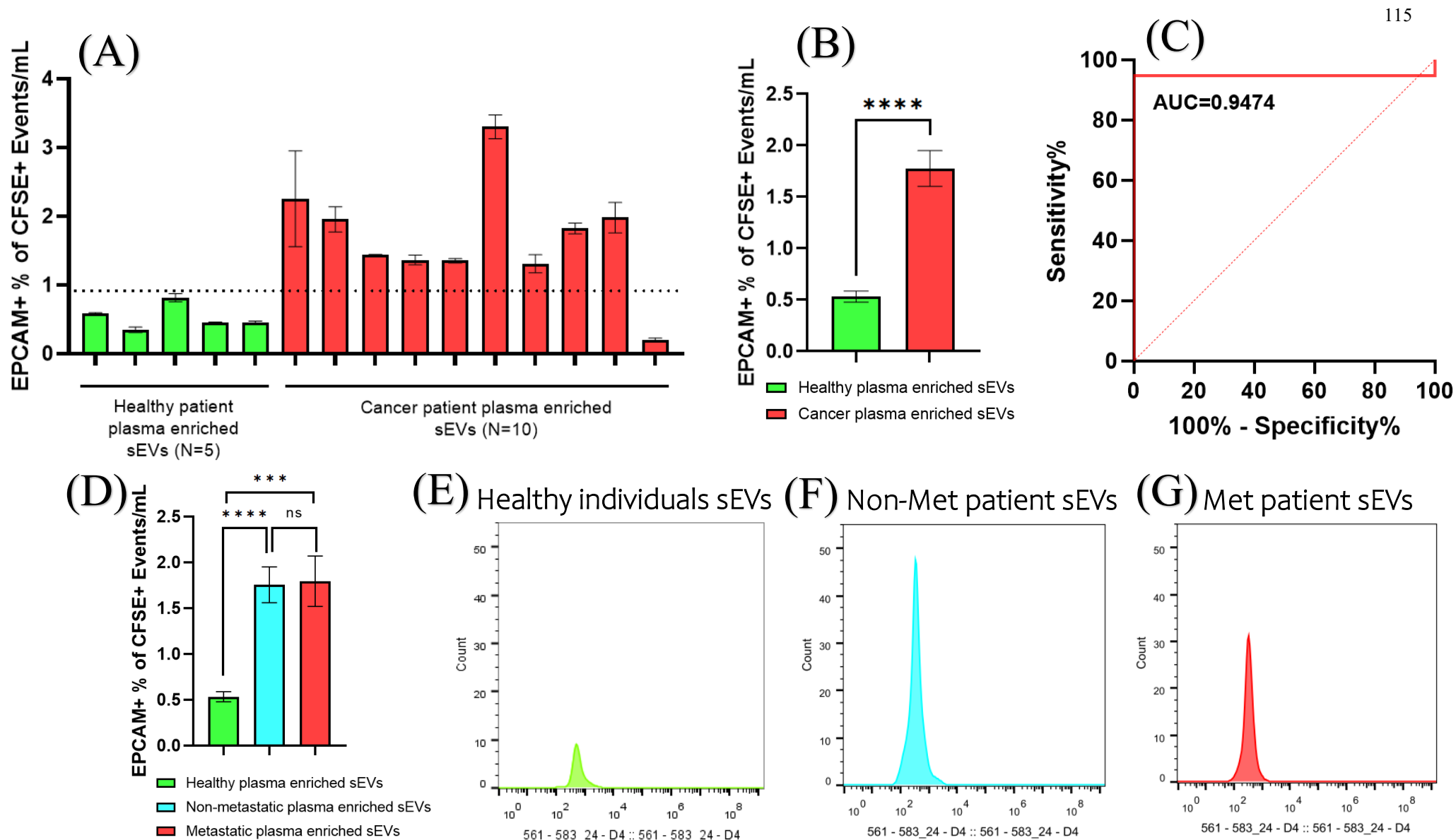**The NS in the table indicates no significance.**
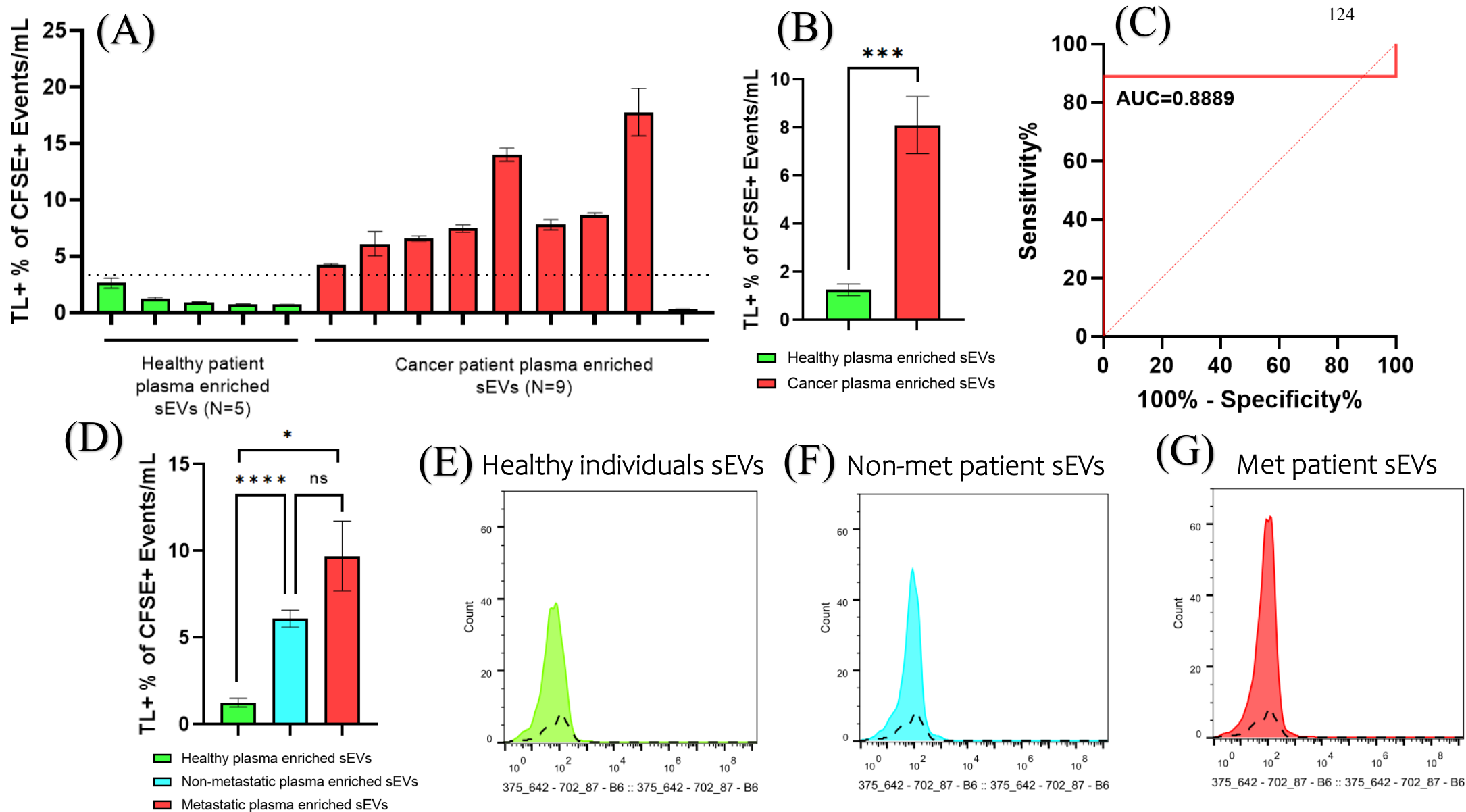
Figure 6.5. Single-vesicle analysis of TL detection of EpCAM-positive plasma-enriched sEVs from metastatic breast cancer patients (n = 5), non-metastatic breast cancer patients (n = 5), and healthy individuals (n = 5). (A) *TL+ events/ml normalised to % CFSE events/ml for each healthy and cancer-associated plasma enriched sEVs sample.* **(B)** *Comparison of TL+ events/ml normalised to % CFSE events/ml between plasma-enriched sEVs derived from healthy individuals and cancer patients.* **(C)** *ROC curve analysis measuring TL+ events/ml normalised to % CFSE events/ml, yielding an AUC of 0.8889 for discriminating patients with cancer from healthy individuals.* **(D)** *Comparison of TL+ events/ml normalised to % CFSE events/ml between plasma-enriched sEVs derived from healthy individuals and non-metastatic and metastatic cancer patients.* **(E-G)** *Histogram profiles of plasma-enriched sEVs after gating showing TL fluorescence. Each experiment comprised biological triplicates with three technical triplicates; error bars represent the Stdev. The t-test followed by Welch's correction was used to determine the significance levels: *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.*

6.5. DISCUSSION

6.5.1. LECTIN MICROARRAY ANALYSIS REVEALS INCREASED LECTIN BINDING OF BREAST CD81-POSITIVE CANCER-ASSOCIATED sEVs

Examination of SEC fractions using NTA and BCA assays indicated that fractions 5, 6, and 7 were most likely to contain plasma-enriched sEVs, based on higher particle counts and lower protein contamination. Consequently, these fractions were pooled and subjected to further characterisation. NTA revealed that sEVs derived from donor plasma samples were within the established size range of 30-150 nm (reviewed by Doyle & Wang, 2019). Interestingly, cancer-associated sEVs isolated from cancer patient plasma showed a significant increase in size compared to those from 'healthy' individuals, with average sizes of 105 nm and 75 nm, respectively. However, it is important to acknowledge that NTA has limitations in discriminating between vesicles and non-vesicular materials, potentially leading to overestimation of vesicle numbers, which could affect these findings. Isolated plasma-enriched sEVs derived from 'healthy' individuals were also characterised using key markers for sEVs, such as CD81, CD63, and syntenin-1 (reviewed by Jankovičová et al., 2020). Notably, CD81 and syntenin-1 were undetectable in western blot analysis, whereas a band for CD63 was observed, which raises intriguing questions regarding the pre-analytical variables that can affect the isolation and physiological state of individuals from whom plasma-enriched sEVs were isolated. It is established that factors such as the size of the needle used to draw blood, proper handling of blood samples, and timely and thorough separation of plasma from uncoagulated blood cells and platelets can affect the characterisation of sEVs (reviewed by Holcar et al., 2021). In addition, patient health, medication use, and genetic and environmental factors can potentially influence the expression and release of specific sEV markers. Understanding these factors is essential for isolating plasma-enriched sEVs for downstream analysis (Cesselli et al., 2018; Macia et al., 2020; Pink et al., 2022). However, it is important to note that these factors were largely unknown in this study because the plasma samples were purchased commercially. This highlights the need for further investigation of the potential impact of these pre-analytical variables on the characterisation of plasma-enriched sEVs.

Western blotting also revealed the presence of ApoB in plasma-enriched sEV samples. This indicated the presence of LDLs particles alongside plasma-enriched sEVs. This was also apparent in TEM analysis, however 'cup-shaped' particles, within the size range of 30 nm to 150 nm were also revealed and presumed to be plasma-enriched sEVs. Studies using SEC to isolate plasma-enriched sEVs have also identified the presence of LDLs (Holcar et al., 2020; Yang et al., 2021). Moreover, LDLs are approximately 10^7 times more abundant in human plasma than sEVs (Johnsen et al., 2019; Simonsen, 2017). Recent studies have focused on adopting isolation protocols to eliminate unwanted LDLs from sEVs (Iannotta et al., 2023). While such adaptations offer opportunities to improve isolation efficiency and specificity, they also present limitations, including the potential loss of sEVs of interest. Researchers must carefully weigh the advantages and disadvantages of each technique to optimise the results for downstream applications (Théry et al., 2018). Given the lack of established isolation techniques to remove lipoproteins and their effects on altering the sEV glycan composition, the original isolation protocol was retained for downstream analysis.

6.5.2. DIAGNOSTIC ASSESSMENT OF LECTINS BINDING PLASMA-ENRICHED EpCAM-POSITIVE sEVS

Initially, plasma-enriched sEVs were subjected to MACSPlex analysis alongside each lectin to identify unique colocalised marker proteins that may enhance the discriminatory capabilities of these lectins through single-vesicle analysis. Notably, HPA-positive plasma-enriched sEVs derived from breast cancer patients revealed a significant increase of co-localisation with CD9 in comparison to those derived from 'healthy' individuals. However, contrasting results emerged upon validation using single-vesicle analysis which revealed no significance of HPA lectin binding of CD9-positive sEVs in distinguishing plasma-enriched sEVs derived from 'healthy' individuals and those from breast cancer patients (Supplementary Figure 10-17). This prompted the introduction of the well-established breast cancer marker EpCAM to complement the diagnostic assessments facilitated by lectins.

The results revealed the high diagnostic capability of EpCAM to differentiate between plasma-enriched sEVs derived from patients with cancer and 'healthy' individuals. These findings align with those of other studies that have employed various technologies to identify EpCAM from breast cancer plasma-enriched sEVs compared with those from 'healthy' individuals (Tian et al., 2021; Vitale et al., 2021). To further enhance the diagnostic potential of EpCAM, lectins HPA, LCA and TL were used in conjunction. Single-vesicle flow cytometry analysis revealed that the lectins HPA, LCA, and TL were significantly associated with EpCAM in patient-plasma-enriched sEVs. However, detecting HPA, LCA, and TL on 'EpCAM-positive' sEVs did not significantly enhance the diagnostic capability compared with detecting 'EpCAM-positive' sEVs alone. Interestingly prior studies have revealed strong glycosylation of EpCAM in cancer, in contrast to its weak or absent glycosylation in healthy tissues. By inhibiting N-linked glycosylation with tunicamycin followed by western blotting Pauli et al. (2003), demonstrated a decrease in glycosylated EpCAM levels. In contrast, inhibition of O-linked glycosylation with N-acetyl galactosaminide had no effect, thereby highlighting the N-glycosylation of EpCAM. Although this study did not directly identify the exact N-linked glycans responsible for the strongly glycosylated forms of EpCAM in cancer tissues, future investigations should aim to elucidate this aspect which may enable improved detection specifically associated with cancer, particularly in the context of its abundance on sEVs. Furthermore, in the current study, all lectins were tested in the absence of EpCAM and assessed for their diagnostic capability to distinguish plasma-enriched sEVs from patients with breast cancer and 'healthy' individuals. The results revealed a decreased diagnostic capability compared to the detection of lectins with 'EpCAM-positive' sEVs, indicating the importance of EpCAM in enhancing the diagnostic accuracy of HPA, LCA, and TL detection.

6.6. KEY FINDINGS

- The extraction of plasma-enriched sEVs by SEC reveals the co-isolation of LDLs alongside sEVs.
- Increased 'EpCAM-positive' sEVs in plasma samples from cancer patients in comparison to 'healthy' individuals revealed diagnostic capability by single-vesicle flow cytometry.

- The detection of HPA, LCA, and TL lectins bound to 'EpCAM-positive' sEVs did not increase diagnostic capability compared to solely detecting 'EpCAM-positive' sEVs by single-vesicle flow cytometry.

CHAPTER 7

FINAL DISCUSSION AND FUTURE WORKS

7.   FINAL DISCUSSION AND FUTURE WORKS

7.1. COMPARISONS OF GLYCOSYLATION BETWEEN CELLS AND THEIR DERIVED sEVs

The glycosylation profiles of the model breast and colorectal epithelial cell lines and their derived sEVs were assessed using HPA, LCA, and TL binding assays to determine similarities in glycosylation patterns. Cancer cells in both breast and colorectal lines showed increased HPA binding compared to normal cells, especially in metastatic phenotypes, consistent with previous studies, as highlighted in section 4.5.1. Notably, these findings also mirrored observations established from their derived sEVs, as 'CD81-positive' sEVs from MCF-7 cells showed increased HPA lectin binding compared to those from 'normal' hTERT-HME1 and 'primary' BT-474 cells. Interestingly, Fan et al. (2023) performed a subcellular distribution analysis of CD81 and CD63 in MCF-7 cells using confocal microscopy and revealed that while CD81 was mainly localised at the plasma membrane, CD63 expression was predominantly intracellular. The localisation of HPA binding to the cell surface, as established in this study, alongside previous findings of CD81 localisation in MCF-7 cells and the observed increase in HPA lectin binding to 'CD81-positive' sEVs, is particularly intriguing. These observations suggest a potential connection between sEVs' glycosylation patterns and the surface localisation of HPA-binding glycans and CD81. Moreover, TL and LCA breast cellular flow cytometry analyses revealed increased binding of 'normal' hTERT-HME1 cells compared to that of 'primary' BT-474 and 'metastatic' MCF-7 cells. Their respective 'CD63-positive' sEVs also mirrored this observation as observed with HPA lectin analysis. However, it is important to note that the glycans recognised by LCA and TL were much broader than those of HPA, as highlighted in section 5.5.1. Specifically, the broader glycan recognition by LCA and TL encompasses both O-linked and N-linked glycans.

Other investigations comparing the glycosylation profiles of cells and their derivatives suggest that the glycosylation patterns of sEVs membranes are only partially correlated with those of cellular membranes across various cell types, as discussed in section 4.5.2. However, it is important to acknowledge that these studies did not specifically investigate the glycosylation profiles of sEVs based on tetraspanin subpopulations, in contrast to the present study. Therefore, the findings in this study may suggest a parallel glycan presentation between the cell surface and their derived sEVs, but nuanced specificity attributed to the tetraspanin composition. Interestingly, growing evidence supports the hypothesis that glycans interact with specific cellular lectins to selectively sort carrier glycoproteins into sEVs in the endosome, potentially leading to enrichment of specific membrane domains which could be the case for tetraspanin platforms (Gomes et al., 2015). According to Batista et al. (2011), the conservation of specific glycan structures, such as increased Man, polylactosamine, α-2,6 sialylation, and complex N-linked glycans, is due to the sorting mechanisms of glycoproteins and glycolipids into sEVs. For example, sorting Ewing's Sarcoma-Associated Transcript 2 (EWI-2), an immunoglobulin glycoprotein, into sEVs leads to altered protein cargo profiles in secreted sEVs (Liang et al., 2014). This may indicate a potential alternative pathway for glycoprotein sorting into sEVs, which is related to its glycan moiety and is distinct from the well-characterised ESCRT machinery pathway (Gomes et al., 2015).

Following these observations, it becomes apparent that the glycosylation patterns of sEVs, especially those identified as 'CD81-positive', likely mirror their parental cell counterparts, as recognised by HPA lectin binding. Conversely, the presence of glycan structures, as recognised by LCA and TL, may serve as indicators of cellular surface composition, particularly when the sEVs are identified as 'CD63-positive'. However, future investigations are warranted to validate and support these conclusions, such as the use of additional breast epithelial cell lines.

7.2. GLYCOSYLATION OF CANCER-ASSOCIATED sEVs

Initially, HPA lectin was used to identify the O-GalNAc glycosylation profiles of breast and colorectal epithelial cell line-derived sEVs. Notably, 'CD81-positive' sEVs derived from MCF-7 cells showed increased HPA lectin binding compared with those from 'normal' hTERT-HME1 and 'primary' BT-474 cells. Interestingly, other studies have highlighted the presence of GalNAc glycans on sEVs derived from other cancer types, as discussed in section 4.5.2. However, to date, these represent the first findings demonstrating the presence of GalNAc glycans on 'CD81-positive' sEVs, specifically attributed to those derived from a 'metastatic' cell phenotype. The interest sparked by these findings raised questions about which alternative glycosylation targets are also specifically enriched on 'CD81-positive' sEVs derived from cancer or cancer metastatic cell sources. Therefore, a lectin microarray was implemented for wider analytical coverage.

The lectin microarray revealed overall increased lectin binding to 'CD81-positive' sEVs derived from 'primary' BT-474 cells and 'metastatic' MCF-7 cells compared to 'normal' hTERT-HME1 cells. As discussed in section 5.5.1, the lectins, BANLEC, WGA, PHA-L, GNA, UDA and TL have also been shown to significantly bind other cancer-associated sEVs derived from other cell types. However, one lectin that sparked interest was TL, given the broad range of sugars it recognises and the absence of literature documenting its binding and association with cancer, especially within the context of sEVs. Other comparisons were also made between the 'CD81-positive' sEVs derived from 'primary' BT-474 and 'metastatic' MCF-7 cells to identify unique glycans associated with both 'primary' and 'metastatic' cancer-associated sEVs. Despite variations in lectin binding between cancer-associated sEVs versus 'normal' sEVs, as well as between 'metastatic' sEVs versus 'primary' sEVs, there was a notable overlap in recognising sugars such as α-GalNAc and α-Man glycans. However, it is important to note that the precise glycans recognised by these lectins often remain largely unknown, leading to ambiguous interpretation. This is also highlighted by the affinity of 'CD81-positive' 'normal' breast sEVs for *Helix aspersa* agglutinin (HAA) compared to 'CD81-positive' 'primary' and 'metastatic' sEVs. HAA is one of the five members of the H-type lectin family that is specific to GalNAc glycans along with HPA. However, it is important to acknowledge that the crystal structures of lectins such as HPA and HAA exhibit significant differences (reviewed by Jeyachandran et al., 2024). Therefore, although the agglutination of both glycans is inhibited by GalNAc, their specific binding affinity for GalNAc is dependent on a variety of other factors, including the proximity of other glycans, their charges, and other molecular characteristics. Nevertheless, the lectin microarray also revealed the increased binding of lentil lectin, or more commonly referred to LCA, which raised interest because of its already notable application in cancer diagnostics as discussed in section 5.5.1. To further validate findings of the lectin

microarray, and dismiss any non-specific interactions, TL and LCA were chosen to analyse by single-vesicle flow cytometry.

Consistent with the lectin microarray findings, LCA demonstrated a significantly higher affinity for 'metastatic' 'CD81-positive' MCF-7 sEVs than 'CD81-positive' BT-474 sEVs. In contrast, TL only exhibited increased binding to 'metastatic' 'CD81-positive' MCF-7 sEVs in comparison to 'normal' hTERT-HME1 sEVs. Collectively, these results suggest that glycans recognised by HPA, LCA, and TL were abundant on 'CD81-positive' sEVs derived from 'metastatic' MCF-7 cells compared to 'primary' BT-474 and 'normal' hTERT-HME1 cells. However, the factors underlying the enrichment of these glycans associated with breast 'metastatic' MCF-7 'CD81-positive' sEVs remain largely elusive. One possibility, as discussed in section 7.1, is the parallel of glycan presentation between the cell surface and their derived sEVs, but the nuanced specificity is attributed to the tetraspanin composition. Moreover, the presence of these distinct glycans on cancer-derived sEVs may result not only from the observed changes in the N- and O-linked glycosylation machinery within cancer cells but also from diverse regulatory factors that control glycan synthesis. These factors may include the availability, presence, and activity of glycosyltransferases, which are modulated by transcription factors, miRNAs, and DNA methylation (reviewed by Dall'Olio & Trinchera, 2017; Kasper et al., 2014; Neelamegham & Mahal, 2016; Rini et al., 2022). Moreover, the localisation of these enzymes within the ER or Golgi apparatus, availability of nucleotide donors, and local environmental factors, including pH, may also influence the enrichment of particular glycan on sEVs (Colley et al., 2022; Rudd et al., 2022). For example a pioneering study by Zheng et al. (2022) genetically engineered human embryonic kidney 293 cells to co-express fucosyltransferase VII or IX to catalyse fucosylation of GlcNAc of a distal alpha 2,3 asialylated lactosamine unit to produce sialyl Lewis-X and Lewis X antigen and demonstrated the enrichment of these specific glycan on their derived sEVs by cytometric analysis. Remarkably, another pioneering study by Hirata et al. (2023) demonstrated the enrichment of GnT-V in B16 mouse melanoma, A549 human lung cancer, and Neuro2A mouse neuroblastoma sEVs in comparison to their parental cells by implementing reverse-phase HPLC and specific substrates. The study also identified that GnT-V sEVs remodelled N-glycan profiles in recipient cells by lectin flow cytometry. Collectively these studies underscore the intricate interplay of factors governing the enrichment of specific glycans on sEVs, suggesting a need for further investigations to elucidate their role in cancer development.

7.3. THE CLINICAL APPLICATION OF GLYCAN COMPOSITION OF sEVs, BY SINGLE-VESICLE ANALYSIS

Breast and colorectal cancers represent significant public health challenges in the UK, as highlighted in sections 1.1.2.1 and 1.1.3.1, respectively. While screening programs and novel biomarkers have contributed to improved mortality rates, as outlined in section 1.6, there remains a critical need for innovative approaches to enhance early detection, prognostic accuracy, and therapeutic efficacy of these cancers. Therefore, this study aimed to determine whether the distinct glycosylation profile identified in breast epithelial cell line-derived sEVs, as discussed in section 7.2, could serve as an alternative approach and prove valuable in a diagnostic capacity when applied to plasma-enriched sEVs.

Characterisation efforts of the plasma-enriched sEVs revealed the co-isolation of LDLs in samples, which was problematic given their established glycosylation (Sukhorukov et al., 2019). Attempts have been made to develop protocols for removing unwanted LDLs from plasma-enriched sEVs samples. For example, Iannotta et al. (2023) isolated plasma-enriched sEVs using SEC and treated the samples with styrene-maleic acid (SMA) to break down lipoprotein contaminants. Through SMA treatment, they effectively removed lipoprotein from plasma-enriched sEV samples, resulting in enhanced labelling of sEVs and reduced interactions with other particles. Although this method proved effective, considerations into its adaptation in this study prompted questions regarding the potential impact of chemically induced lipoprotein removal on the glycan composition of plasma-enriched sEVs, which was the central focus of this study. In addition, implementing different techniques to improve the efficiency of isolation and removal of LDLs is problematic because it is well documented that different isolation techniques yield sEVs with varying glycan compositions (Freitas et al., 2019). Given the variability in sEV glycan composition resulting from different isolation techniques and the aim of replicating the glycosylation profile of sEVs derived from breast epithelial cell line-derived sEVs, the implementation of SEC to isolate plasma-enriched sEVs was deemed necessary to maintain consistency across observations. To overcome the challenges presented by the co-isolation of LDL with plasma-enriched sEVs, a method to selectively target cancer-associated sEVs was adopted, specifically by detecting 'EpCAM-positive' subpopulations using single-vesicle flow cytometry.

Single-vesicle flow cytometry revealed the high diagnostic capability of EpCAM to differentiate between plasma-enriched sEVs derived from cancer patients and 'healthy' individuals. These findings are consistent with previous studies that utilised different technologies to detect EpCAM on plasma-enriched sEVs. For instance, Vitale et al. (2021) quantified EpCAM expression on plasma-enriched sEVs derived from breast cancer patients using enzyme-linked immunosorbent assay (ELISA) and observed a significant increase compared to those from healthy individuals. Another intriguing aspect of EpCAM is its strong glycosylation in cancer tissues compared to weak or absent glycosylation in healthy tissues, as discussed in section 6.5.2 (Pauli et al., 2003). With this in mind, the study shifted its focus to establish whether the binding of HPA, LCA, and TL would enhance the diagnostic capability of distinguishing 'EpCAM-positive' plasma-enriched sEVs derived from breast cancer patients and those from 'healthy' individuals. While the results did not show a significant increase in diagnostic capability compared to solely detecting 'EpCAM-positive' sEVs, as outlined in section 6.5.2, the strong glycosylation of EpCAM observed in cancer is linked to N-linked glycans which may not be recognised by these lectins (Pauli et al., 2003). This raises the question of whether understanding the specific N-linked glycans abundantly expressed on EpCAM in cancers and then implementing lectins that specifically recognise them may be an area of future research. One approach would be to mirror the methodology implemented for the lectin microarray in this study, but with EpCAM detection alongside the plasma-enriched sEVs to identify unique glycans associated with each cohort.

Navigating the translation of a clinical sEV diagnostic assay presents challenges in relation to the technology adopted, specifically the use of single-vesicle flow cytometry, which was optimised and adopted throughout this study. As discussed by Nolan et al. (2022) challenges associated with measuring single sEVs include the development of appropriate concentration, size, and molecular cargo standards for single EV measurements; calibrating instrument responses in absolute rather than relative units; and creating assays that report EV number,

size, and molecular cargo in absolute units that can be compared across instruments and between laboratories. However, while the adoption of the MIFlowCyt-EV guidelines as outlined by Welsh et al. (2020), has facilitated the standardisation of experimental reporting, enabling a more robust and standardised approach to the application of single-vesicle flow cytometry, specific challenges persist in the context of this study, highlighting the need for continued refinement and improvement in the methodology. One such challenge was the technical variation between analysing plasma-enriched sEVs on different days. This emphasises the importance of running a known non-biological standard alongside each experimental run to ensure the comparability of the results and warrants further investigation. Additionally, translating the capabilities of this technology into clinical settings raises important considerations regarding affordability and expertise required to operate such advanced systems effectively. The complexity and scale of working with single-vesicle flow cytometry demands a high level of technical proficiency, which may present a barrier to its widespread adoption in clinical settings. Moreover, the challenges in achieving the high throughput necessary for clinical applications further complicate its integration into routine practice. Although the potential diagnostic benefits of single-vesicle flow cytometry are significant, careful consideration should be given to the practical and financial aspects of its implementation in clinical practice. Future investigations should focus on the continued development of methodologies to address these challenges and explore alternative technologically simpler approaches for sEV analysis. These advancements hold promise for facilitating the earlier detection of cancer and improving patient outcomes.

7.4. NOVEL CONTRIBUTIONS

In summary, this work presents several novel contributions to the understanding of the role of extracellular vesicles and glycosylation in cancer metastasis as outlined below.

- Identification of increased 'CD81-positive' sEVs in association with the 'metastatic' phenotype (MCF-7), whereas 'CD63-positive' sEVs were linked to the 'normal' phenotype (hTERT-HME1) on breast epithelial derived sEVs indicating distinct tetraspanin expression patterns.

- Demonstration that treatment of breast cells with sEVs induces changes in cellular motility.

- Observation of elevated overall lectin binding of breast cancer-associated 'CD81-positive' sEVs (MCF-7 and BT-474) compared to 'normal' (hTERT-HME1) sEVs by lectin microarray, suggesting an increased glycan composition of cancer-associated sEVs.

- Finding that breast 'metastatic' MCF-7 'CD81-positive' sEVs exhibited increased lectin binding compared to 'primary' BT-474 by lectin microarray analysis, indicating an enriched glycan composition of CD81 tetraspanin sEVs, potentially linked to metatasis.

- The correlation between increased HPA, LCA, and TL binding with breast 'metastatic' MCF-7 'CD81-positive' sEVs compared to 'normal' hTERT-HME1 and 'primary' BT-474 sEVs, suggests an elevated glycan composition recognised by these lectins, potentially associated with breast metastatic sEVs, depending on the tetraspanin subpopulation.

- The correlation between increased HPA and LCA binding with breast 'normal' hTERT-HME1 'CD63-positive' sEVs compared to 'primary' BT-474 and 'metastatic' MCF-7 sEVs suggests an elevated glycan composition recognised by these lectins potentially associated with the non-cancer phenotype dependent on the tetraspanin subpopulation.

- HPA lectin recognition mirrors breast cell surface glycosylation seen in 'CD81-positive' sEVs, while LCA and TL binding corresponds to 'CD63-positive' sEVs, thereby suggesting a partial correlation between the glycosylation patterns of sEVs membranes and their parental cells.

- Development of an adapted single-vesicle flow cytometer to analyse 'EpCAM-positive' sEVs, demonstrating diagnostic significance in distinguishing breast plasma-enriched sEVs derived from patients with cancer compared to 'healthy' individuals.

REFERENCES

Ageberg, M., & Lindmark, A. (2003). Characterisation of the biosynthesis and processing of the neutrophil granule membrane protein CD63 in myeloid cells. *Clinical & Laboratory Haematology*, *25*(5), 297–306. https://doi.org/10.1046/j.1365-2257.2003.00541.x

Alberro, A., Iparraguirre, L., Fernandes, A., & Otaegui, D. (2021). Extracellular Vesicles in Blood: Sources, Effects, and Applications. *International Journal of Molecular Sciences*, *22*(15), 8163. https://doi.org/10.3390/ijms22158163

Ali, N., Zhang, L., Taylor, S., Mironov, A., Urbé, S., & Woodman, P. (2013). Recruitment of UBPY and ESCRT Exchange Drive HD-PTP-Dependent Sorting of EGFR to the MVB. *Current Biology*, *23*(6), 453–461. https://doi.org/10.1016/j.cub.2013.02.033

Alkabban, F. M., & Ferguson, T. (2024). Breast Cancer. In *StatPearls* (Vol. 1). StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK482286/

Andreu, Z., & Yáñez-Mó, M. (2014). Tetraspanins in Extracellular Vesicle Formation and Function. *Frontiers in Immunology*, *5*, 442. https://doi.org/10.3389/fimmu.2014.00442

Anoop, T. M., Joseph P, R., Soman, S., Chacko, S., & Mathew, M. (2022). Significance of serum carcinoembryonic antigen in metastatic breast cancer patients: A prospective study. *World Journal of Clinical Oncology*, *13*(6), 529–539. https://doi.org/10.5306/wjco.v13.i6.529

Azzouz, L. L., & Sharma, S. (2023). Physiology, Large Intestine. In *StatPearls* (Vol. 1). StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK507857/

Babst, M., Katzmann, D. J., Snyder, W. B., Wendland, B., & Emr, S. D. (2002). Endosome-Associated Complex, ESCRT-II, Recruits Transport Machinery for Protein Sorting at the Multivesicular Body. *Developmental Cell*, *3*(2), 283–289. https://doi.org/10.1016/S1534-5807(02)00219-8

Bache, K. G., Brech, A., Mehlum, A., & Stenmark, H. (2003). Hrs regulates multivesicular body formation via ESCRT recruitment to endosomes. *The Journal of Cell Biology*, *162*(3), 435–442. https://doi.org/10.1083/jcb.200302131

Baietti, M. F., Zhang, Z., Mortier, E., Melchior, A., Degeest, G., Geeraerts, A., Ivarsson, Y., Depoortere, F., Coomans, C., Vermeiren, E., Zimmermann, P., & David, G. (2012). Syndecan–syntenin–ALIX regulates the biogenesis of exosomes. *Nature Cell Biology*, *14*(7), Article 7. https://doi.org/10.1038/ncb2502

Band, V., Zajchowski, D., Kulesa, V., & Sager, R. (1990). Human papilloma virus DNAs immortalize normal human mammary epithelial cells and reduce their growth factor requirements. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(1), 463–467.

Barre, A., Damme, E. J. M. V., Peumans, W. J., & Rougé, P. (1996). Structure-Function Relationship of Monocot Mannose-Binding Lectins. *Plant Physiology*, *112*(4), 1531–1540.

Basharat, S., & Horton, J. (2022). *Emerging Multi-Cancer Early Detection Technologies: CADTH Horizon Scan*. Canadian Agency for Drugs and Technologies in Health. http://www.ncbi.nlm.nih.gov/books/NBK598992/

Batista, B. S., Eng, W. S., Pilobello, K. T., Hendricks-Muñoz, K. D., & Mahal, L. K. (2011). Identification of a Conserved Glycan Signature for Microvesicles. *Journal of Proteome Research*, *10*(10), 4624–4633. https://doi.org/10.1021/pr200434y

Bebelman, M. P., Janssen, E., Pegtel, D. M., & Crudden, C. (2021). The forces driving cancer extracellular vesicle secretion. *Neoplasia*, *23*(1), 149–157. https://doi.org/10.1016/j.neo.2020.11.011

Bebelman, M. P., Smit, M. J., Pegtel, D. M., & Baglio, S. R. (2018). Biogenesis and function of extracellular vesicles in cancer. *Pharmacology & Therapeutics*, *188*, 1–11. https://doi.org/10.1016/j.pharmthera.2018.02.013

Becker, A., Thakur, B. K., Weiss, J. M., Kim, H. S., Peinado, H., & Lyden, D. (2016). Extracellular vesicles in cancer: Cell-to-cell mediators of metastasis. *Cancer Cell*, *30*(6), 836–848. https://doi.org/10.1016/j.ccell.2016.10.009

Berditchevski, F. (2001). Complexes of tetraspanins with integrins: More than meets the eye. *Journal of Cell Science*, *114*(23), 4143–4151. https://doi.org/10.1242/jcs.114.23.4143

Berditchevski, F., Bazzoni, G., & Hemler, M. E. (1995). Specific Association Of CD63 with the VLA-3 and VLA-6 Integrins. *Journal of Biological Chemistry*, *270*(30), 17784–17790. https://doi.org/10.1074/jbc.270.30.17784

Bertolini, I., Ghosh, J. C., Kossenkov, A. V., Mulugu, S., Krishn, S. R., Vaira, V., Qin, J., Plow, E. F., Languino, L. R., & Altieri, D. C. (2020). Small Extracellular Vesicle Regulation of Mitochondrial Dynamics Reprograms a Hypoxic Tumor Microenvironment. *Developmental Cell*, *55*(2), 163-177.e6. https://doi.org/10.1016/j.devcel.2020.07.014

Biswas, S. K., Banerjee, S., Baker, G. W., Kuo, C.-Y., & Chowdhury, I. (2022). The Mammary Gland: Basic Structure and Molecular Signaling during Development. *International Journal of Molecular Sciences*, *23*(7), 3883. https://doi.org/10.3390/ijms23073883

Bojar, D., Meche, L., Meng, G., Eng, W., Smith, D. F., Cummings, R. D., & Mahal, L. K. (2022). A Useful Guide to Lectin Binding: Machine-Learning Directed Annotation of 57 Unique Lectin Specificities. *ACS Chemical Biology*, *17*(11), 2993–3012. https://doi.org/10.1021/acschembio.1c00689

Boucheix, C., Duc, G. H. T., Jasmin, C., & Rubinstein, E. (2001). Tetraspanins and malignancy. *Expert Reviews in Molecular Medicine*, *3*(4), 1–17. https://doi.org/10.1017/S1462399401002381

Boucheix, C., & Rubinstein, E. (2001). Tetraspanins. *Cellular and Molecular Life Sciences CMLS*, *58*(9), 1189–1205. https://doi.org/10.1007/PL00000933

Breitwieser, K., Koch, L. F., Tertel, T., Proestler, E., Burgers, L. D., Lipps, C., Adjaye, J., Fürst, R., Giebel, B., & Saul, M. J. (2022). Detailed Characterization of Small Extracellular Vesicles from Different Cell Types Based on Tetraspanin Composition by ExoView R100 Platform. *International Journal of Molecular Sciences*, *23*(15), Article 15. https://doi.org/10.3390/ijms23158544

Brennan, K., Martin, K., FitzGerald, S. P., O'Sullivan, J., Wu, Y., Blanco, A., Richardson, C., & Mc Gee, M. M. (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), 1039. https://doi.org/10.1038/s41598-020-57497-7

Brito-Rocha, T., Constâncio, V., Henrique, R., & Jerónimo, C. (2023). Shifting the Cancer Screening Paradigm: The Rising Potential of Blood-Based Multi-Cancer Early Detection Tests. *Cells*, *12*(6), 935. https://doi.org/10.3390/cells12060935

Brockhausen, I. (2006). Mucin-type O-glycans in human colon and breast cancer: Glycodynamics and functions. *EMBO Reports*, *7*(6), 599–604. https://doi.org/10.1038/sj.embor.7400705

Brooks, S. (2024). Lectins as versatile tools to explore cellular glycosylation. *European Journal of Histochemistry*, *68*(1), Article 1. https://doi.org/10.4081/ejh.2024.3959

Brooks, S. A., Hall, D. M., & Buley, I. (2001). GalNAc glycoprotein expression by breast cell lines, primary breast cancer and normal breast epithelial membrane. *British Journal of Cancer*, *85*(7), 1014–1022. https://doi.org/10.1054/bjoc.2001.2028

Brooks, S. A., & Leathem, A. J. C. (1991). Prediction of lymph node involvement in breast cancer by detection of altered glycosylation in the primary tumour. *The Lancet*, *338*(8759), 71–74. https://doi.org/10.1016/0140-6736(91)90071-V

Cammue, B. P. A., Peeters, B., & Peumans, W. J. (1986). A new lectin from tulip (Tulipa) bulbs. *Planta*, *169*(4), 583–588. https://doi.org/10.1007/BF00392110

Cancer Research U.K. (2015, May 14). Cancer Research UK. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer

Cao, Y., Stosiek, P., Springer, G. F., & Karsten, U. (1996). Thomsen-Friedenreich-related carbohydrate antigens in normal adult human tissues: A systematic and comparative study. *Histochemistry and Cell Biology*, *106*(2), 197–207. https://doi.org/10.1007/BF02484401

Casset, F., Hamelryck, T., Loris, R., Brisson, J.-R., Tellier, C., Dao-Thi, M.-H., Wyns, L., Poortmans, F., Pérez, S., & Imberty, A. (1995). NMR, Molecular Modeling, and Crystallographic Studies of Lentil Lectin-Sucrose Interaction (∗). *Journal of Biological Chemistry*, *270*(43), 25619–25628. https://doi.org/10.1074/jbc.270.43.25619

Catalano, M., & O'Driscoll, L. (2019). Inhibiting extracellular vesicles formation and release: A review of EV inhibitors. *Journal of Extracellular Vesicles*, *9*(1), 1703244. https://doi.org/10.1080/20013078.2019.1703244

Čaval, T., Alisson-Silva, F., & Schwarz, F. (2023). Roles of glycosylation at the cancer cell surface: Opportunities for large scale glycoproteomics. *Theranostics*, *13*(8), 2605–2615. https://doi.org/10.7150/thno.81760

Cesselli, D., Parisse, P., Aleksova, A., Veneziano, C., Cervellin, C., Zanello, A., & Beltrami, A. P. (2018). Extracellular Vesicles: How Drug and Pathology Interfere With Their Biogenesis and Function. *Frontiers in Physiology*, *9*, 1394. https://doi.org/10.3389/fphys.2018.01394

Charrin, S., le Naour, F., Silvie, O., Milhiet, P.-E., Boucheix, C., & Rubinstein, E. (2009). Lateral organization of membrane proteins: Tetraspanins spin their web. *Biochemical Journal*, *420*(2), 133–154. https://doi.org/10.1042/BJ20082422

Chirco, R., Liu, X.-W., Jung, K.-K., & Kim, H.-R. C. (2006). Novel functions of TIMPs in cell signaling. *Cancer and Metastasis Reviews*, *25*(1), 99–113. https://doi.org/10.1007/s10555-006-7893-x

Chuo, S. T.-Y., Chien, J. C.-Y., & Lai, C. P.-K. (2018). Imaging extracellular vesicles: Current and emerging methods. *Journal of Biomedical Science*, *25*(1), 91. https://doi.org/10.1186/s12929-018-0494-5

Colley, K. J., Varki, A., Haltiwanger, R. S., & Kinoshita, T. (2022). Cellular Organization of Glycosylation. In A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, D. Mohnen, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar, & P. H. Seeberger (Eds.), *Essentials of Glycobiology* (4th ed., Vol. 1). Cold Spring Harbor Laboratory Press. http://www.ncbi.nlm.nih.gov/books/NBK579926/

Colombo, M., Raposo, G., & Théry, C. (2014). Biogenesis, Secretion, and Intercellular Interactions of Exosomes and Other Extracellular Vesicles. *Annual Review of Cell and Developmental Biology*, *30*(1), 255–289. https://doi.org/10.1146/annurev-cellbio-101512-122326

Costa, J., Gatermann, M., Nimtz, M., Kandzia, S., Glatzel, M., & Conradt, H. S. (2018). N-Glycosylation of Extracellular Vesicles from HEK-293 and Glioma Cell Lines. *Analytical Chemistry*, *90*(13), 7871–7879. https://doi.org/10.1021/acs.analchem.7b05455

Coumans, F. A. W., Brisson, A. R., Buzas, E. I., Dignat-George, F., Drees, E. E. E., El-Andaloussi, S., Emanueli, C., Gasecka, A., Hendrix, A., Hill, A. F., Lacroix, R., Lee, Y., van Leeuwen, T. G., Mackman, N., Mäger, I., Nolan, J. P., van der Pol, E., Pegtel, D. M., Sahoo, S., … Nieuwland, R. (2017). Methodological Guidelines to Study Extracellular Vesicles. *Circulation Research*, *120*(10), 1632–1648. https://doi.org/10.1161/CIRCRESAHA.117.309417

Cowell, C. F., Weigelt, B., Sakr, R. A., Ng, C. K. Y., Hicks, J., King, T. A., & Reis-Filho, J. S. (2013). Progression from ductal carcinoma in situ to invasive breast cancer: Revisited. *Molecular Oncology*, *7*(5), 859–869. https://doi.org/10.1016/j.molonc.2013.07.005

Cummings, R. D. (2009). The repertoire of glycan determinants in the human glycome. *Molecular BioSystems*, *5*(10), 1087–1104. https://doi.org/10.1039/B907931A

Dalla, P. V., Santos, J., Milthorpe, B. K., & Padula, M. P. (2020). Selectively-Packaged Proteins in Breast Cancer Extracellular Vesicles Involved in Metastasis. *International Journal of Molecular Sciences*, *21*(14), 4990. https://doi.org/10.3390/ijms21144990

Dall'Olio, F., & Trinchera, M. (2017). Epigenetic Bases of Aberrant Glycosylation in Cancer. *International Journal of Molecular Sciences*, *18*(5), 998. https://doi.org/10.3390/ijms18050998

Dar, G. H., Mendes, C. C., Kuan, W.-L., Speciale, A. A., Conceição, M., Görgens, A., Uliyakina, I., Lobo, M. J., Lim, W. F., EL Andaloussi, S., Mäger, I., Roberts, T. C., Barker, R. A., Goberdhan, D. C. I., Wilson, C., & Wood, M. J. A. (2021). GAPDH controls extracellular vesicle biogenesis and enhances the therapeutic potential of EV mediated siRNA delivery to the brain. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-27056-3

Debray, H., Decout, D., Strecker, G., Spik, G., & Montreuil, J. (1981). Specificity of Twelve Lectins Towards Oligosaccharides and Glycopeptides Related to N-Glycosylproteins. *European Journal of Biochemistry*, *117*(1), 41–51. https://doi.org/10.1111/j.1432-1033.1981.tb06300.x

Dennis, J. W., Laferté, S., Waghorne, C., Breitman, M. L., & Kerbel, R. S. (1987). B1-6 Branching of Asn-Linked Oligosaccharides Is Directly Associated with Metastasis. *Science*, *236*(4801), 582–585. https://doi.org/10.1126/science.2953071

Desai, P. R. (2000). Immunoreactive T and Tn antigens in malignancy: Role in carcinoma diagnosis, prognosis, and immunotherapy. *Transfusion Medicine Reviews*, *14*(4), 312–325. https://doi.org/10.1053/tmrv.2000.16229

DeSantis, C. E., Ma, J., Gaudet, M. M., Newman, L. A., Miller, K. D., Goding Sauer, A., Jemal, A., & Siegel, R. L. (2019). Breast cancer statistics, 2019. *CA: A Cancer Journal for Clinicians*, *69*(6), 438–451. https://doi.org/10.3322/caac.21583

Detilleux, D., Spill, Y. G., Balaramane, D., Weber, M., & Bardet, A. F. (2022). Pan-cancer predictions of transcription factors mediating aberrant DNA methylation. *Epigenetics & Chromatin*, *15*(1), 10. https://doi.org/10.1186/s13072-022-00443-w

Di Gioia, D., Dresse, M., Mayr, D., Nagel, D., Heinemann, V., & Stieber, P. (2015). Serum HER2 in combination with CA 15-3 as a parameter for prognosis in patients with early breast cancer. *Clinica Chimica Acta*, *440*, 16–22. https://doi.org/10.1016/j.cca.2014.11.001

Di Vizio, D., Morello, M., Dudley, A. C., Schow, P. W., Adam, R. M., Morley, S., Mulholland, D., Rotinen, M., Hager, M. H., Insabato, L., Moses, M. A., Demichelis, F., Lisanti, M. P., Wu, H., Klagsbrun, M., Bhowmick, N. A., Rubin, M. A., D'Souza-Schorey, C., & Freeman, M. R. (2012). Large Oncosomes in Human Prostate Cancer Tissues and in the Circulation of Mice with Metastatic Disease. *The American Journal of Pathology*, *181*(5), 1573–1584. https://doi.org/10.1016/j.ajpath.2012.07.030

Dores, M. R., Chen, B., Lin, H., Soh, U. J. K., Paing, M. M., Montagne, W. A., Meerloo, T., & Trejo, J. (2012). ALIX binds a YPX3L motif of the GPCR PAR1 and mediates ubiquitin-independent ESCRT-III/MVB sorting. *The Journal of Cell Biology*, *197*(3), 407–419. https://doi.org/10.1083/jcb.201110031

Doyle, L. M., & Wang, M. Z. (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

Doyotte, A., Mironov, A., McKenzie, E., & Woodman, P. (2008). The Bro1-related protein HD-PTP/PTPN23 is required for endosomal cargo sorting and multivesicular body morphogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(17), 6308–6313. https://doi.org/10.1073/pnas.0707601105

Duan, B., Zhao, Y., Bai, J., Wang, J., Duan, X., Luo, X., Zhang, R., Pu, Y., Kou, M., Lei, J., & Yang, S. (2022). Colorectal Cancer: An Overview. In J. A. Morgado-Diaz (Ed.), *Gastrointestinal Cancers* (Vol. 1). Exon Publications. http://www.ncbi.nlm.nih.gov/books/NBK586003/

Duffy, M. J. (2001). Carcinoembryonic Antigen as a Marker for Colorectal Cancer: Is It Clinically Useful? *Clinical Chemistry*, *47*(4), 624–630. https://doi.org/10.1093/clinchem/47.4.624

Duffy, M. J., Diamandis, E. P., & Crown, J. (2021). Circulating tumor DNA (ctDNA) as a pan-cancer screening test: Is it finally on the horizon? *Clinical Chemistry and Laboratory Medicine (CCLM)*, *59*(8), 1353–1361. https://doi.org/10.1515/cclm-2021-0171

Duffy, M. J., Shering, S., Sherry, F., McDermott, E., & O'Higgins, N. (2000). CA 15–3: A Prognostic Marker in Breast Cancer. *The International Journal of Biological Markers*, *15*(4), 330–333. https://doi.org/10.1177/172460080001500410

Dusoswa, S. A., Horrevorts, S. K., Ambrosini, M., Kalay, H., Paauw, N. J., Nieuwland, R., Pegtel, M. D., Würdinger, T., Van Kooyk, Y., & Garcia-Vallejo, J. J. (2019). Glycan modification of glioblastoma-derived extracellular vesicles enhances receptor-mediated targeting of dendritic cells. *Journal of Extracellular Vesicles*, *8*(1), 1648995. https://doi.org/10.1080/20013078.2019.1648995

Ender, F., Zamzow, P., von Bubnoff, N., & Gieseler, F. (2019). Detection and Quantification of Extracellular Vesicles via FACS: Membrane Labeling Matters! *International Journal of Molecular Sciences*, *21*(1), 291. https://doi.org/10.3390/ijms21010291

Escrevente, C., Keller, S., Altevogt, P., & Costa, J. (2011). Interaction and uptake of exosomes by ovarian cancer cells. *BMC Cancer*, *11*(1), 108. https://doi.org/10.1186/1471-2407-11-108

Esmail, S., & Manolson, M. F. (2021). Advances in understanding N-glycosylation structure, function, and regulation in health and disease. *European Journal of Cell Biology*, *100*(7), 151186. https://doi.org/10.1016/j.ejcb.2021.151186

Fan, Y., Pionneau, C., Cocozza, F., Boëlle, P.-Y., Chardonnet, S., Charrin, S., Théry, C., Zimmermann, P., & Rubinstein, E. (2023). Differential proteomics argues against a general role for CD9, CD81 or CD63 in the sorting of proteins into extracellular vesicles. *Journal of Extracellular Vesicles*, *12*(8), 12352. https://doi.org/10.1002/jev2.12352

Fares, J., Fares, M. Y., Khachfe, H. H., Salhab, H. A., & Fares, Y. (2020). Molecular principles of metastasis: A hallmark of cancer revisited. *Signal Transduction and Targeted Therapy*, *5*(1), Article 1. https://doi.org/10.1038/s41392-020-0134-x

Feng, Y., Guo, Y., Li, Y., Tao, J., Ding, L., Wu, J., & Ju, H. (2018). Lectin-mediated in situ rolling circle amplification on exosomes for probing cancer-related glycan pattern. *Analytica Chimica Acta*, *1039*, 108–115. https://doi.org/10.1016/j.aca.2018.07.040

Ferguson, K., Yadav, A., Morey, S., Abdullah, J., Hrysenko, G., Eng, J. Y., Sajjad, M., Koury, S., & Rittenhouse-Olson, K. (2014). Preclinical studies with JAA-F11 anti-Thomsen-Friedenreich monoclonal antibody for human breast cancer. *Future Oncology*, *10*(3), 385–399. https://doi.org/10.2217/fon.13.209

Fernandes, B., Sagman, U., Auger, M., Demetrio, M., & Dennis, J. W. (1991). B1–6 Branched Oligosaccharides as a Marker of Tumor Progression in Human Breast and Colon Neoplasia1. *Cancer Research*, *51*(2), 718–723.

Fogh, J., & Trempe, G. (1975). New Human Tumor Cell Lines. In J. Fogh (Ed.), *Human Tumor Cells in Vitro* (Vol. 1, pp. 115–159). Springer US. https://doi.org/10.1007/978-1-4757-1647-4_5

Freitas, D., Balmaña, M., Poças, J., Campos, D., Osório, H., Konstantinidi, A., Vakhrushev, S. Y., Magalhães, A., & Reis, C. A. (2019). Different isolation approaches lead to diverse glycosylated extracellular vesicle populations. *Journal of Extracellular Vesicles*, *8*(1), 1621131. https://doi.org/10.1080/20013078.2019.1621131

Fu, C., Zhao, H., Wang, Y., Cai, H., Xiao, Y., Zeng, Y., & Chen, H. (2016). Tumor-associated antigens: Tn antigen, sTn antigen, and T antigen. *HLA*, *88*(6), 275–286. https://doi.org/10.1111/tan.12900

Fu, Y., & Li, H. (2016). Assessing Clinical Significance of Serum CA15-3 and Carcinoembryonic Antigen (CEA) Levels in Breast Cancer Patients: A Meta-Analysis. *Medical Science Monitor*, *22*, 3154–3162. https://doi.org/10.12659/MSM.896563

Gahloth, D., Heaven, G., Jowitt, T. A., Mould, A. P., Bella, J., Baldock, C., Woodman, P., & Tabernero, L. (2017). The open architecture of HD-PTP phosphatase provides new insights into the mechanism of regulation of ESCRT function. *Scientific Reports*, *7*, 9151. https://doi.org/10.1038/s41598-017-09467-9

Galindo-Hernandez, O., Gonzales-Vazquez, C., Cortes-Reynosa, P., Reyes-Uribe, E., Chavez-Ocaña, S., Reyes-Hernandez, O., Sierra-Martinez, M., & Salazar, E. P. (2015). Extracellular vesicles from women with breast cancer promote an epithelial-mesenchymal transition-like process in mammary epithelial cells MCF10A. *Tumor Biology*, *36*(12), 9649–9659. https://doi.org/10.1007/s13277-015-3711-9

Gastl, G., Spizzo, G., Obrist, P., Dünser, M., & Mikuz, G. (2000). Ep-CAM overexpression in breast cancer as a predictor of survival. *The Lancet*, *356*(9246), 1981–1982. https://doi.org/10.1016/S0140-6736(00)03312-2

Gill, D. J., Teo, H., Sun, J., Perisic, O., Veprintsev, D. B., Emr, S. D., & Williams, R. L. (2007). Structural insight into the ESCRT-I/-II link and its role in MVB trafficking. *The EMBO Journal*, *26*(2), 600–612. https://doi.org/10.1038/sj.emboj.7601501

Gill, D. J., Tham, K. M., Chia, J., Wang, S. C., Steentoft, C., Clausen, H., Bard-Chapeau, E. A., & Bard, F. A. (2013). Initiation of GalNAc-type O-glycosylation in the endoplasmic reticulum promotes cancer cell invasiveness. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(34), E3152–E3161. https://doi.org/10.1073/pnas.1305269110

Godley, K. C., Gladwell, C., Murray, P. J., & Denton, E. (2017). The UK breast screening program – what you need to know. *Climacteric*, *20*(4), 313–320. https://doi.org/10.1080/13697137.2017.1342619

Goldstein, I. J., Hughes, R. C., Monsigny, M., Osawa, T., & Sharon, N. (1980). What should be called a lectin? *Nature*, *285*(5760), 66–66. https://doi.org/10.1038/285066b0

Gomes, J., Gomes-Alves, P., Carvalho, S. B., Peixoto, C., Alves, P. M., Altevogt, P., & Costa, J. (2015). Extracellular Vesicles from Ovarian Carcinoma Cells Display Specific Glycosignatures. *Biomolecules*, *5*(3), 1741–1761. https://doi.org/10.3390/biom5031741

González-King, H., Tejedor, S., Ciria, M., Gil-Barrachina, M., Soriano-Navarro, M., Sánchez-Sánchez, R., Sepúlveda, P., & García, N. A. (2022). Non-classical Notch signaling by MDA-MB-231 breast cancer cell-derived small extracellular vesicles promotes malignancy in poorly invasive MCF-7 cells. *Cancer Gene Therapy*, *29*(7), Article 7. https://doi.org/10.1038/s41417-021-00411-8

Görgens, A., Bremer, M., Ferrer-Tur, R., Murke, F., Tertel, T., Horn, P. A., Thalmann, S., Welsh, J. A., Probst, C., Guerin, C., Boulanger, C. M., Jones, J. C., Hanenberg, H., Erdbrügger, U., Lannigan, J., Ricklefs, F. L., El-Andaloussi, S., & Giebel, B. (2019). Optimisation of imaging flow cytometry for the analysis of single extracellular vesicles by using fluorescence-tagged vesicles as biological reference material. *Journal of Extracellular Vesicles*, *8*(1), 1587567. https://doi.org/10.1080/20013078.2019.1587567

Görgens, A., Corso, G., Hagey, D. W., Jawad Wiklander, R., Gustafsson, M. O., Felldin, U., Lee, Y., Bostancioglu, R. B., Sork, H., Liang, X., Zheng, W., Mohammad, D. K., van de Wakker, S. I., Vader, P., Zickler, A. M., Mamand, D. R., Ma, L., Holme, M. N., Stevens, M. M., … EL Andaloussi, S. (2022). Identification of storage conditions stabilizing extracellular vesicles preparations. *Journal of Extracellular Vesicles*, *11*(6), e12238. https://doi.org/10.1002/jev2.12238

Gudjonsson, T., Adriance, M. C., Sternlicht, M. D., Petersen, O. W., & Bissell, M. J. (2005). Myoepithelial Cells: Their Origin and Function in Breast Morphogenesis and Neoplasia. *Journal of Mammary Gland Biology and Neoplasia*, *10*(3), 261–272. https://doi.org/10.1007/s10911-005-9586-4

Gurung, S., Perocheau, D., Touramanidou, L., & Baruteau, J. (2021). The exosome journey: From biogenesis to uptake and intracellular signalling. *Cell Communication and Signaling : CCS*, *19*, 47. https://doi.org/10.1186/s12964-021-00730-1

Han, Y., Azuma, K., Watanabe, S., Semba, K., & Nakayama, J. (2022). Metastatic profiling of HER2-positive breast cancer cell lines in xenograft models. *Clinical & Experimental Metastasis*, *39*(3), 467–477. https://doi.org/10.1007/s10585-022-10150-1

Hanahan, D., & Weinberg, R. A. (2000). The Hallmarks of Cancer. *Cell*, *100*(1), 57–70. https://doi.org/10.1016/S0092-8674(00)81683-9

Harada, Y., Kizuka, Y., Tokoro, Y., Kondo, K., Yagi, H., Kato, K., Inoue, H., Taniguchi, N., & Maruyama, I. (2019). N-glycome inheritance from cells to extracellular vesicles in B16 melanomas. *Febs Letters*, *593*(9), 942–951. https://doi.org/10.1002/1873-3468.13377

Hashim, O. H., Jayapalan, J. J., & Lee, C.-S. (2017). Lectins: An effective tool for screening of potential cancer biomarkers. *PeerJ*, *5*, e3784. https://doi.org/10.7717/peerj.3784

Häuselmann, I., & Borsig, L. (2014). Altered Tumor-Cell Glycosylation Promotes Metastasis. *Frontiers in Oncology*, *4*, 28. https://doi.org/10.3389/fonc.2014.00028

Hemler, M. E. (1998). Integrin associated proteins. *Current Opinion in Cell Biology*, *10*(5), 578–585. https://doi.org/10.1016/S0955-0674(98)80032-X

Hemler, M. E. (2001). Specific tetraspanin functions. *The Journal of Cell Biology*, *155*(7), 1103–1108. https://doi.org/10.1083/jcb.200108061

Hemler, M. E. (2005). Tetraspanin functions and associated microdomains. *Nature Reviews Molecular Cell Biology*, *6*(10), Article 10. https://doi.org/10.1038/nrm1736

Hemler, M. E. (2008). Targeting of tetraspanin proteins—Potential benefits and strategies. *Nature Reviews. Drug Discovery*, *7*(9), 747–758. https://doi.org/10.1038/nrd2659

Hemler, M. E. (2014). Tetraspanin proteins promote multiple cancer stages. *Nature Reviews Cancer*, *14*(1), Article 1. https://doi.org/10.1038/nrc3640

Hirata, T., Harada, Y., Hirosawa, K. M., Tokoro, Y., Suzuki, K. G. N., & Kizuka, Y. (2023). *N*-acetylglucosaminyltransferase-V (GnT-V)-enriched small extracellular vesicles mediate *N*-glycan remodeling in recipient cells. *iScience*, *26*(1), 105747. https://doi.org/10.1016/j.isci.2022.105747

Ho, W.-L., Hsu, W.-M., Huang, M.-C., Kadomatsu, K., & Nakagawara, A. (2016). Protein glycosylation in cancers and its potential therapeutic applications in neuroblastoma. *Journal of Hematology & Oncology*, *9*(1), 100. https://doi.org/10.1186/s13045-016-0334-6

Holcar, M., Ferdin, J., Sitar, S., Tušek-Žnidarič, M., Dolžan, V., Plemenitaš, A., Žagar, E., & Lenassi, M. (2020). Enrichment of plasma extracellular vesicles for reliable quantification of their size and concentration for biomarker discovery. *Scientific Reports*, *10*(1), 21346. https://doi.org/10.1038/s41598-020-78422-y

Holcar, M., Kandušer, M., & Lenassi, M. (2021). Blood Nanoparticles – Influence on Extracellular Vesicle Isolation and Characterization. *Frontiers in Pharmacology*, *12*. https://doi.org/10.3389/fphar.2021.773844

Howard, I. K., Sage, H. J., Stein, M. D., Young, N. M., Leon, M. A., & Dyckes, D. F. (1971). Studies on a Phytohemagglutinin from the Lentil. *Journal of Biological Chemistry*, *246*(6), 1590–1595. https://doi.org/10.1016/S0021-9258(18)62353-1

Hsu, J. L., & Hung, M.-C. (2016). The role of HER2, EGFR, and other receptor tyrosine kinases in breast cancer. *Cancer Metastasis Reviews*, *35*(4), 575–588. https://doi.org/10.1007/s10555-016-9649-6

Hu, T., Wolfram, J., & Srivastava, S. (2021). Extracellular Vesicles in Cancer Detection: Hopes and Hypes. *Trends in Cancer*, *7*(2), 122–133. https://doi.org/10.1016/j.trecan.2020.09.003

Huang, C., & Wen, B. (2016). Phenotype transformation of immortalized NCM460 colon epithelial cell line by TGF-β1 is associated with chromosome instability. *Molecular Biology Reports*, *43*(10), 1069–1078. https://doi.org/10.1007/s11033-016-4038-3

Huang, L.-K., & Wang, M.-J. J. (1995). Image thresholding by minimizing the measures of fuzziness. *Pattern Recognition*, *28*(1), 41–51. https://doi.org/10.1016/0031-3203(94)E0043-K

Huttenlocher, A., & Horwitz, A. R. (2011). Integrins in Cell Migration. *Cold Spring Harbor Perspectives in Biology*, *3*(9), a005074. https://doi.org/10.1101/cshperspect.a005074

Iannotta, D., A., A., Lai, A., Nair, S., Koifman, N., Lappas, M., Salomon, C., & Wolfram, J. (2023). Chemically-Induced Lipoprotein Breakdown for Improved Extracellular Vesicle Purification. *Small*, *1*(n/a), 2307240. https://doi.org/10.1002/smll.202307240

Jankovičová, J., Sečová, P., Michalková, K., & Antalíková, J. (2020). Tetraspanins, More than Markers of Extracellular Vesicles in Reproduction. *International Journal of Molecular Sciences*, *21*(20), 7568. https://doi.org/10.3390/ijms21207568

Jayaprakash, N. G., & Surolia, A. (2017). Role of glycosylation in nucleating protein folding and stability. *Biochemical Journal*, *474*(14), 2333–2347. https://doi.org/10.1042/BCJ20170111

Jeyachandran, S., Radhakrishnan, A., & Ragavendran, C. (2024). Harnessing the power of mollusc lectins as immuno-protective biomolecules. *Molecular Biology Reports*, *51*(1), 182. https://doi.org/10.1007/s11033-023-09018-8

Johnsen, K. B., Gudbergsson, J. M., Andresen, T. L., & Simonsen, J. B. (2019). What is the blood concentration of extracellular vesicles? Implications for the use of extracellular vesicles as blood-borne biomarkers of cancer. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, *1871*(1), 109–116. https://doi.org/10.1016/j.bbcan.2018.11.006

Jung, K., Liu, X., Chirco, R., Fridman, R., & Kim, H. C. (2006). Identification of CD63 as a tissue inhibitor of metalloproteinase-1 interacting cell surface protein. *The EMBO Journal*, *25*(17), 3934–3942. https://doi.org/10.1038/sj.emboj.7601281

Kabel, A. M. (2017). Tumor markers of breast cancer: New prospectives. *Journal of Oncological Sciences*, *3*(1), 5–11. https://doi.org/10.1016/j.jons.2017.01.001

Kakeji, Y., Tsujitani, S., Mori, M., Maehara, Y., & Sugimachi, K. (1991). Helix pomatia agglutinin binding activity is a predictor of survival time for patients with gastric carcinoma. *Cancer*, *68*(11), 2438–2442. https://doi.org/10.1002/1097-0142(19911201)68:11<2438::AID-CNCR2820681119>3.0.CO;2-#

Kasper, B. T., Koppolu, S., & Mahal, L. K. (2014). Insights into MiRNA Regulation of the Human Glycome. *Biochemical and Biophysical Research Communications*, *445*(4), 774–779. https://doi.org/10.1016/j.bbrc.2014.01.034

Kaszuba, K., Grzybek, M., Orłowski, A., Danne, R., Róg, T., Simons, K., Coskun, Ü., & Vattulainen, I. (2015). N-Glycosylation as determinant of epidermal growth factor receptor conformation in membranes. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(14), 4334–4339. https://doi.org/10.1073/pnas.1503262112

Katzmann, D. J., Stefan, C. J., Babst, M., & Emr, S. D. (2003). Vps27 recruits ESCRT machinery to endosomes during MVB sorting. *The Journal of Cell Biology*, *162*(3), 413–423. https://doi.org/10.1083/jcb.200302136

Kekelidze, M., D'Errico, L., Pansini, M., Tyndall, A., & Hohmann, J. (2013). Colorectal cancer: Current imaging methods and future perspectives for the diagnosis, staging and therapeutic response evaluation. *World Journal of Gastroenterology : WJG*, *19*(46), 8502–8514. https://doi.org/10.3748/wjg.v19.i46.8502

Keller, L., Werner, S., & Pantel, K. (2019). Biology and clinical relevance of EpCAM. *Cell Stress*, *3*(6), 165–180. https://doi.org/10.15698/cst2019.06.188

Kenific, C. M., Zhang, H., & Lyden, D. (2021). An exosome pathway without an ESCRT. *Cell Research*, *31*(2), Article 2. https://doi.org/10.1038/s41422-020-00418-0

Khaldoyanidi, S. K., Glinsky, V. V., Sikora, L., Glinskii, A. B., Mossine, V. V., Quinn, T. P., Glinsky, G. V., & Sriramarao, P. (2003). MDA-MB-435 Human Breast Carcinoma Cell Homo- and Heterotypic Adhesion under Flow Conditions Is Mediated in Part by Thomsen-Friedenreich Antigen-Galectin-3 Interactions *. *Journal of Biological Chemistry*, *278*(6), 4127–4134. https://doi.org/10.1074/jbc.M209590200

Khosrowabadi, E., Wenta, T., Keskitalo, S., Manninen, A., & Kellokumpu, S. (2022). Altered glycosylation of several metastasis-associated glycoproteins with terminal GalNAc defines the highly invasive cancer cell phenotype. *Oncotarget*, *13*, 73–89. https://doi.org/10.18632/oncotarget.28167

Kim, Y.-W., Park, J., Lee, H.-J., Lee, S.-Y., & Kim, S.-J. (2012). TGF-β sensitivity is determined by N-linked glycosylation of the type II TGF-β receptor. *Biochemical Journal*, *445*(Pt 3), 403–411. https://doi.org/10.1042/BJ20111923

Kobayashi, T., Stang, E., Fang, K. S., de Moerloose, P., Parton, R. G., & Gruenberg, J. (1998). A lipid associated with the antiphospholipid syndrome regulates endosome structure and function. *Nature*, *392*(6672), Article 6672. https://doi.org/10.1038/32440

Kondo, K., Harada, Y., Nakano, M., Suzuki, T., Fukushige, T., Hanzawa, K., Yagi, H., Takagi, K., Mizuno, K., Miyamoto, Y., Taniguchi, N., Kato, K., Kanekura, T., Dohmae, N., Machida, K., Maruyama, I., & Inoue, H. (2022). Identification of distinct N-glycosylation patterns on extracellular vesicles from small-cell and non-small-cell lung cancer cells. *The Journal of Biological Chemistry*, *298*(6), 101950. https://doi.org/10.1016/j.jbc.2022.101950

Kong, S., Zhang, Y. H., & Zhang, W. (2018). Regulation of Intestinal Epithelial Cells Properties and Functions by Amino Acids. *BioMed Research International*, *2018*, 2819154. https://doi.org/10.1155/2018/2819154

Koo, S., Neilson, L. J., Von Wagner, C., & Rees, C. J. (2017). The NHS Bowel Cancer Screening Program: Current perspectives on strategies for improvement. *Risk Management and Healthcare Policy*, *10*, 177–187. https://doi.org/10.2147/RMHP.S109116

Kormelink, T. G., Arkesteijn, G. J. A., Nauwelaers, F. A., van den Engh, G., Nolte-'t Hoen, E. N. M., & Wauben, M. H. M. (2016). Prerequisites for the analysis and sorting of extracellular vesicle subpopulations by high-resolution flow cytometry. *Cytometry Part A*, *89*(2), 135–147. https://doi.org/10.1002/cyto.a.22644

Kornfeld, K., Reitman, M. L., & Kornfeld, R. (1981). The carbohydrate-binding specificity of pea and lentil lectins. Fucose is an important determinant. *Journal of Biological Chemistry*, *256*(13), 6633–6640. https://doi.org/10.1016/S0021-9258(19)69037-X

Kornfeld, S., Rogers, J., & Gregory, W. (1971). The Nature of the Cell Surface Receptor Site for Lens culinaris Phytohemagglutinin. *Journal of Biological Chemistry*, *246*(21), 6581–6586. https://doi.org/10.1016/S0021-9258(19)34153-5

Kramer, R. H., & Nicolson, G. L. (1979). Interactions of tumor cells with vascular endothelial cell monolayers: A model for metastatic invasion. *Proceedings of the National Academy of Sciences*, *76*(11), 5704–5708. https://doi.org/10.1073/pnas.76.11.5704

Krebs, M. G., Hou, J.-M., Ward, T. H., Blackhall, F. H., & Dive, C. (2010). Circulating tumour cells: Their utility in cancer management and predicting outcomes. *Therapeutic Advances in Medical Oncology*, *2*(6), 351–365. https://doi.org/10.1177/1758834010378414

Krishnamoorthy, L., Bess, J. W., Preston, A. B., Nagashima, K., & Mahal, L. K. (2009). HIV-1 and microvesicles from T-cells share a common glycome, arguing for a common origin. *Nature Chemical Biology*, *5*(4), 244–250. https://doi.org/10.1038/nchembio.151

Kugeratski, F. G., Hodge, K., Lilla, S., McAndrews, K. M., Zhou, X., Hwang, R. F., Zanivan, S., & Kalluri, R. (2021). Quantitative proteomics identifies the core proteome of exosomes with syntenin-1 as the highest abundant protein and a putative universal biomarker. *Nature Cell Biology*, *23*(6), Article 6. https://doi.org/10.1038/s41556-021-00693-y

Laack, E., Nikbakht, H., Peters, A., Kugler, C., Jasiewicz, Y., Edler, L., Hossfeld, D. K., & Schumacher, U. (2002). Lectin histochemistry of resected adenocarcinoma of the lung: Helix pomatia agglutinin binding is an independent prognostic factor. *The American Journal of Pathology*, *160*(3), 1001–1008. https://doi.org/10.1016/S0002-9440(10)64921-8

Lam, S. K., & Ng, T. B. (2011). Lectins: Production and practical applications. *Applied Microbiology and Biotechnology*, *89*(1), 45–55. https://doi.org/10.1007/s00253-010-2892-9

Landras, A., Reger de Moura, C., Jouenne, F., Lebbe, C., Menashi, S., & Mourah, S. (2019). CD147 Is a Promising Target of Tumor Progression and a Prognostic Biomarker. *Cancers*, *11*(11), 1803. https://doi.org/10.3390/cancers11111803

Lasfargues, E. Y., Coutinho, W. G., & Redfield, E. S. (1978). Isolation of two human tumor epithelial cell lines from solid breast carcinomas. *Journal of the National Cancer Institute*, *61*(4), 967–978.

Laulagnier, K., Motta, C., Hamdi, S., Roy, S., Fauvelle, F., Pageaux, J.-F., Kobayashi, T., Salles, J.-P., Perret, B., Bonnerot, C., & Record, M. (2004). Mast cell- and dendritic cell-derived exosomes display a specific lipid composition and an unusual membrane organization. *Biochemical Journal*, *380*(Pt 1), 161–171. https://doi.org/10.1042/BJ20031594

Lee, K.-M., Seo, E.-C., Lee, J.-H., Kim, H.-J., & Hwangbo, C. (2023). The Multifunctional Protein Syntenin-1: Regulator of Exosome Biogenesis, Cellular Function, and Tumor Progression. *International Journal of Molecular Sciences*, *24*(11), 9418. https://doi.org/10.3390/ijms24119418

Leibovltz, A., Stlnson, J. C., McCombs, W. B., McCoy, C. E., Mazur, K. C., & Mabry, N. D. (1976). *Classification of Human Colorectal Adenocarcinoma Cell Lines*.

Li, B., Hao, K., Ma, C., Li, Z., Li, H., Du, W., Sun, L., Jia, T., Liu, A., Li, Y., Xu, L., Gao, Q., Yang, R., & Lin, C. (2021). *Isolation and characterization of fucosylated extracellular vesicles based on a novel high-throughput GlyExo-Capture techniqu*. https://doi.org/10.1101/2021.12.09.471505

Li, C., Sun, C., Lohcharoenkal, W., Ali, M. M., Xing, P., Zheng, W., Görgens, A., Gustafsson, M. O., EL Andaloussi, S., Sonkoly, E., & Pivarcsi, A. (2023). Cutaneous squamous cell carcinoma-derived extracellular vesicles exert an oncogenic role by activating cancer-associated fibroblasts. *Cell Death Discovery*, *9*, 260. https://doi.org/10.1038/s41420-023-01555-2

Li, K., Chen, Y., Li, A., Tan, C., & Liu, X. (2019). Exosomes play roles in sequential processes of tumor metastasis. *International Journal of Cancer*, *144*(7), 1486–1495. https://doi.org/10.1002/ijc.31774

Li, W., Shao, B., Liu, C., Wang, H., Zheng, W., Kong, W., Liu, X., Xu, G., Wang, C., Li, H., Zhu, L., & Yang, Y. (2018). Noninvasive Diagnosis and Molecular Phenotyping of Breast Cancer through Microbead-Assisted Flow Cytometry Detection of Tumor-Derived Extracellular Vesicles. *Small Methods*, *2*(11), 1800122. https://doi.org/10.1002/smtd.201800122

Liang, W., Mao, S., Sun, S., Li, M., Li, Z., Yu, R., Ma, T., Gu, J., Zhang, J., Taniguchi, N., & Li, W. (2018). Core Fucosylation of the T Cell Receptor Is Required for T Cell Activation. *Frontiers in Immunology*, *9*, 78. https://doi.org/10.3389/fimmu.2018.00078

Liang, Y., Eng, W. S., Colquhoun, D. R., Dinglasan, R. R., Graham, D. R., & Mahal, L. K. (2014). Complex N-linked glycans serve as a determinant for exosome/microvesicle cargo recruitment. *The Journal of Biological Chemistry*, *289*(47), 32526–32537. https://doi.org/10.1074/jbc.M114.606269

Lianidou, E. (2021). Detection and relevance of epigenetic markers on ctDNA: Recent advances and future outlook. *Molecular Oncology*, *15*(6), 1683–1700. https://doi.org/10.1002/1878-0261.12978

Lin, S., Zhou, S., & Yuan, T. (2020). The "sugar-coated bullets" of cancer: Tumor-derived exosome surface glycosylation from basic knowledge to applications. *Clinical and Translational Medicine*, *10*(6), e204. https://doi.org/10.1002/ctm2.204

Lin, Y., & Lubman, D. M. (2024). The role of N-glycosylation in cancer. *Acta Pharmaceutica Sinica. B*, *14*(3), 1098–1110. https://doi.org/10.1016/j.apsb.2023.10.014

Liu, M. C., Oxnard, G. R., Klein, E. A., Swanton, C., & Seiden, M. V. (2020). Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. *Annals of Oncology : Official Journal of the European Society for Medical Oncology*, *31*(6), 745–759. https://doi.org/10.1016/j.annonc.2020.02.011

Loris, R., Casset, F., Bouckaert, J., Pletinckx, J., Dao-Thi, M.-H., Poortmans, F., Imberty, A., Perez, S., & Wyns, L. (1994). The monosaccharide binding site of lentil lectin: An X-ray and molecular modelling study. *Glycoconjugate Journal*, *11*(6), 507–517. https://doi.org/10.1007/BF00731301

Lu, Q., Hope, L. W., Brasch, M., Reinhard, C., & Cohen, S. N. (2003). TSG101 interaction with HRS mediates endosomal trafficking and receptor down-regulation. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(13), 7626–7631. https://doi.org/10.1073/pnas.0932599100

Luga, V., Zhang, L., Viloria-Petit, A. M., Ogunjimi, A. A., Inanlou, M. R., Chiu, E., Buchanan, M., Hosein, A. N., Basik, M., & Wrana, J. L. (2012). Exosomes Mediate Stromal Mobilization of Autocrine Wnt-PCP Signaling in Breast Cancer Cell Migration. *Cell*, *151*(7), 1542–1556. https://doi.org/10.1016/j.cell.2012.11.024

Ma, L., Li, Y., Peng, J., Wu, D., Zhao, X., Cui, Y., Chen, L., Yan, X., Du, Y., & Yu, L. (2015). Discovery of the migrasome, an organelle mediating release of cytoplasmic contents during cell migration. *Cell Research*, *25*(1), 24–38. https://doi.org/10.1038/cr.2014.135

Macia, L., Nanan, R., Hosseini-Beheshti, E., & Grau, G. E. (2020). Host- and Microbiota-Derived Extracellular Vesicles, Immune Function, and Disease Development. *International Journal of Molecular Sciences*, *21*(1), Article 1. https://doi.org/10.3390/ijms21010107

Maia, J., Batista, S., Couto, N., Gregório, A. C., Bodo, C., Elzanowska, J., Strano Moraes, M. C., & Costa-Silva, B. (2020). Employing Flow Cytometry to Extracellular Vesicles Sample Microvolume Analysis and Quality Control. *Frontiers in Cell and Developmental Biology*, *8*. https://doi.org/10.3389/fcell.2020.593750

Makki, J. (2015). Diversity of Breast Carcinoma: Histological Subtypes and Clinical Relevance. *Clinical Medicine Insights. Pathology*, *8*, 23–31. https://doi.org/10.4137/CPath.S31563

Malaguarnera, G., Giordano, M., Paladina, I., Berretta, M., Cappellani, A., & Malaguarnera, M. (2010). Serum Markers of Hepatocellular Carcinoma. *Digestive Diseases and Sciences*, *55*(10), 2744–2755. https://doi.org/10.1007/s10620-010-1184-7

Martins, Á. M., Ramos, C. C., Freitas, D., & Reis, C. A. (2021). Glycosylation of Cancer Extracellular Vesicles: Capture Strategies, Functional Roles and Potential Clinical Applications. *Cells*, *10*(1), 109. https://doi.org/10.3390/cells10010109

Mastoridis, S., Bertolino, G. M., Whitehouse, G., Dazzi, F., Sanchez-Fueyo, A., & Martinez-Llordella, M. (2018). Multiparametric Analysis of Circulating Exosomes and Other Small Extracellular Vesicles by Advanced Imaging Flow Cytometry. *Frontiers in Immunology*, *9*, 1583. https://doi.org/10.3389/fimmu.2018.01583

Matsuda, A., Kuno, A., Yoshida, M., Wagatsuma, T., Sato, T., Miyagishi, M., Zhao, J., Suematsu, M., Kabe, Y., & Narimatsu, H. (2020). Comparative Glycomic Analysis of Exosome Subpopulations Derived from Pancreatic Cancer Cell Lines. *Journal of Proteome Research*, *19*(6), 2516–2524. https://doi.org/10.1021/acs.jproteome.0c00200

McSherry, E. A., Donatello, S., Hopkins, A. M., & McDonnell, S. (2007). Common Molecular Mechanisms of Mammary Gland Development and Breast Cancer. *Cellular and Molecular Life Sciences*, *64*(24), 3201–3218. https://doi.org/10.1007/s00018-007-7388-0

Mezu-Ndubuisi, O. J., & Maheshwari, A. (2021). The role of integrins in inflammation and angiogenesis. *Pediatric Research*, *89*(7), 1619–1626. https://doi.org/10.1038/s41390-020-01177-9

Molina, R., Jo, J., Filella, X., Zanon, G., Pahisa, J., Muñoz, M., Farrus, B., Latre, M. L., Escriche, C., Estape, J., & Ballesta, A. M. (1998). c-erbB-2 oncoprotein, CEA, and CA 15.3 in patients with breast cancer: Prognostic value. *Breast Cancer Research and Treatment*, *51*(2), 109–119. https://doi.org/10.1023/A:1005734429304

Morales-Kastresana, A., Telford, B., Musich, T. A., McKinnon, K., Clayborne, C., Braig, Z., Rosner, A., Demberg, T., Watson, D. C., Karpova, T. S., Freeman, G. J., DeKruyff, R. H., Pavlakis, G. N., Terabe,

M., Robert-Guroff, M., Berzofsky, J. A., & Jones, J. C. (2017). Labeling Extracellular Vesicles for Nanoscale Flow Cytometry. *Scientific Reports*, *7*(1), Article 1. https://doi.org/10.1038/s41598-017-01731-2

Moyer, M. P., Manzano, L. A., Merriman, R. L., Stauffer, J. S., & Tanzer, L. R. (1996). NCM460, a normal human colon mucosal epithelial cell line. *In Vitro Cellular & Developmental Biology - Animal*, *32*(6), 315–317. https://doi.org/10.1007/BF02722955

Munkley, J. (2016). The Role of Sialyl-Tn in Cancer. *International Journal of Molecular Sciences*, *17*(3), 275. https://doi.org/10.3390/ijms17030275

Mustapic, M., Eitan, E., Werner, J. K., Berkowitz, S. T., Lazaropoulos, M. P., Tran, J., Goetzl, E. J., & Kapogiannis, D. (2017). Plasma Extracellular Vesicles Enriched for Neuronal Origin: A Potential Window into Brain Pathologic Processes. *Frontiers in Neuroscience*, *11*, 278. https://doi.org/10.3389/fnins.2017.00278

Na, T.-Y., Schecterson, L., Mendonsa, A. M., & Gumbiner, B. M. (2020). The functional activity of E-cadherin controls tumor cell metastasis at multiple steps. *Proceedings of the National Academy of Sciences*, *117*(11), 5931–5937. https://doi.org/10.1073/pnas.1918167117

Nagao, K., Maeda, K., Hosomi, K., Morioka, K., Inuzuka, T., & Ohtsubo, K. (2022). Sialyl-Tn antigen facilitates extracellular vesicle-mediated transfer of FAK and enhances motility of recipient cells. *The Journal of Biochemistry*, *171*(5), 543–554. https://doi.org/10.1093/jb/mvac008

Neelamegham, S., & Mahal, L. K. (2016). Multi-level regulation of cellular glycosylation: From genes to transcript to enzyme to structure. *Current Opinion in Structural Biology*, *40*, 145–152. https://doi.org/10.1016/j.sbi.2016.09.013

Nicolás-Ávila, J. A., Lechuga-Vieco, A. V., Esteban-Martínez, L., Sánchez-Díaz, M., Díaz-García, E., Santiago, D. J., Rubio-Ponce, A., Li, J. L., Balachander, A., Quintana, J. A., Martínez-de-Mena, R., Castejón-Vega, B., Pun-García, A., Través, P. G., Bonzón-Kulichenko, E., García-Marqués, F., Cussó, L., A-González, N., González-Guerra, A., … Hidalgo, A. (2020). A Network of Macrophages Supports Mitochondrial Homeostasis in the Heart. *Cell*, *183*(1), 94-109.e23. https://doi.org/10.1016/j.cell.2020.08.031

Nishida-Aoki, N., Tominaga, N., Kosaka, N., & Ochiya, T. (2020). Altered biodistribution of deglycosylated extracellular vesicles through enhanced cellular uptake. *Journal of Extracellular Vesicles*, *9*(1), 1713527. https://doi.org/10.1080/20013078.2020.1713527

Nolan, J. P., Chiu, D. T., & Welsh, J. A. (2022). Rigor and reproducibility: Status and challenges for single vesicle analysis. *Extracellular Vesicles and Circulating Nucleic Acids*, *3*(3), 244–248. https://doi.org/10.20517/evcna.2022.28

O'Brien, K., Rani, S., Corcoran, C., Wallace, R., Hughes, L., Friel, A. M., McDonnell, S., Crown, J., Radomski, M. W., & O'Driscoll, L. (2013). Exosomes from triple-negative breast cancer cells can transfer phenotypic traits representing their cells of origin to secondary cells. *European Journal of Cancer*, *49*(8), 1845–1859. https://doi.org/10.1016/j.ejca.2013.01.017

O'Brien, M. J., & Gibbons, D. (1996). The Adenoma-carcinoma Sequence in Colorectal Neoplasia. *Surgical Oncology Clinics of North America*, *5*(3), 513–530. https://doi.org/10.1016/S1055-3207(18)30361-2

Ogata, S., Maimonis, P. J., & Itzkowitz, S. H. (1992). Mucins Bearing the Cancer-associated Sialosyl-Tn Antigen Mediate Inhibition of Natural Killer Cell Cytotoxicity1. *Cancer Research*, *52*(17), 4741–4746.

Ogawa, T., Hirohashi, Y., Murai, A., Nishidate, T., Okita, K., Wang, L., Ikehara, Y., Satoyoshi, T., Usui, A., Kubo, T., Nakastugawa, M., Kanaseki, T., Tsukahara, T., Kutomi, G., Furuhata, T., Hirata, K., Sato, N., Mizuguchi, T., Takemasa, I., & Torigoe, T. (2017). ST6GALNAC1 plays important roles in enhancing cancer stem phenotypes of colorectal cancer via the Akt pathway. *Oncotarget*, *8*(68), 112550–112564. https://doi.org/10.18632/oncotarget.22545

Osaki, M., & Okada, F. (2019). Exosomes and Their Role in Cancer Progression. *Yonago Acta Medica*, *62*(2), 182–190. https://doi.org/10.33160/yam.2019.06.002

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66. https://doi.org/10.1109/TSMC.1979.4310076

Pauli, C., Münz, M., Kieu, C., Mack, B., Breinl, P., Wollenberg, B., Lang, S., Zeidler, R., & Gires, O. (2003). Tumor-specific glycosylation of the carcinoma-associated epithelial cell adhesion molecule EpCAM in head and neck carcinomas. *Cancer Letters*, *193*(1), 25–32. https://doi.org/10.1016/S0304-3835(03)00003-X

Pinho, S. S., Reis, C. A., Paredes, J., Magalhães, A. M., Ferreira, A. C., Figueiredo, J., Xiaogang, W., Carneiro, F., Gärtner, F., & Seruca, R. (2009). The role of N-acetylglucosaminyltransferase III and V in the post-transcriptional modifications of E-cadherin. *Human Molecular Genetics*, *18*(14), 2599–2608. https://doi.org/10.1093/hmg/ddp194

Pink, R. C., Beaman, E.-M., Samuel, P., Brooks, S. A., & Carter, D. R. F. (2022). Utilising extracellular vesicles for early cancer diagnostics: Benefits, challenges and recommendations for the future. *British Journal of Cancer*, *126*(3), 323–330. https://doi.org/10.1038/s41416-021-01668-4

Pinto, D., & Parameswaran, R. (2023). Role of Truncated O-GalNAc Glycans in Cancer Progression and Metastasis in Endocrine Cancers. *Cancers*, *15*(13), 3266. https://doi.org/10.3390/cancers15133266

Pishas, K. I., Adwal, A., Neuhaus, S. J., Clayer, M. T., Farshid, G., Staudacher, A. H., & Callen, D. F. (2015). XI-006 induces potent p53-independent apoptosis in Ewing sarcoma. *Scientific Reports*, *5*(1), Article 1. https://doi.org/10.1038/srep11465

Pols, M. S., & Klumperman, J. (2009). Trafficking and function of the tetraspanin CD63. *Experimental Cell Research*, *315*(9), 1584–1592. https://doi.org/10.1016/j.yexcr.2008.09.020

Pons-Belda, O. D., Fernandez-Uriarte, A., & Diamandis, E. P. (2021). Can Circulating Tumor DNA Support a Successful Screening Test for Early Cancer Detection? The Grail Paradigm. *Diagnostics*, *11*(12), 2171. https://doi.org/10.3390/diagnostics11122171

Poon, I. K. H., Parkes, M. A. F., Jiang, L., Atkin-Smith, G. K., Tixeira, R., Gregory, C. D., Ozkocak, D. C., Rutter, S. F., Caruso, S., Santavanond, J. P., Paone, S., Shi, B., Hodge, A. L., Hulett, M. D., Chow, J. D. Y., Phan, T. K., & Baxter, A. A. (2019). Moving beyond size and phosphatidylserine exposure: Evidence for a diversity of apoptotic cell-derived extracellular vesicles in vitro. *Journal of Extracellular Vesicles*, *8*(1), 1608786. https://doi.org/10.1080/20013078.2019.1608786

Pužar Dominkuš, P., Stenovec, M., Sitar, S., Lasič, E., Zorec, R., Plemenitaš, A., Žagar, E., Kreft, M., & Lenassi, M. (2018). PKH26 labeling of extracellular vesicles: Characterization and cellular internalization of contaminating PKH26 nanoparticles. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *1860*(6), 1350–1361. https://doi.org/10.1016/j.bbamem.2018.03.013

Rachmilewitz, J. (2010). Glycosylation. *Self Nonself*, *1*(3), 250–254. https://doi.org/10.4161/self.1.3.12330

Radford, K. J., Thorne, R. F., & Hersey, P. (1997). Regulation of tumor cell motility and migration by CD63 in a human melanoma cell line. *The Journal of Immunology*, *158*(7), 3353–3358. https://doi.org/10.4049/jimmunol.158.7.3353

Ragni, E., Palombella, S., Lopa, S., Talò, G., Perucca Orfei, C., De Luca, P., Moretti, M., & de Girolamo, L. (2020). Innovative Visualization and Quantification of Extracellular Vesicles Interaction with and Incorporation in Target Cells in 3D Microenvironments. *Cells*, *9*(5), Article 5. https://doi.org/10.3390/cells9051180

Rajwar, Y. C., Jain, N., Bhatia, G., Sikka, N., Garg, B., & Walia, E. (2015). Expression and Significance of Cadherins and Its Subtypes in Development and Progression of Oral Cancers: A Review. *Journal of Clinical and Diagnostic Research : JCDR*, *9*(5), ZE05–ZE07. https://doi.org/10.7860/JCDR/2015/11964.5907

Raman, J., Guan, Y., Perrine, C. L., Gerken, T. A., & Tabak, L. A. (2012). UDP-N-acetyl-α-d-galactosamine:polypeptide N-acetylgalactosaminyltransferases: Completion of the family tree. *Glycobiology*, *22*(6), 768–777. https://doi.org/10.1093/glycob/cwr183

Ramos, E. K., Tsai, C.-F., Jia, Y., Cao, Y., Manu, M., Taftaf, R., Hoffmann, A. D., El-Shennawy, L., Gritsenko, M. A., Adorno-Cruz, V., Schuster, E. J., Scholten, D., Patel, D., Liu, X., Patel, P., Wray, B., Zhang, Y., Zhang, S., Moore, R. J., … Liu, H. (2022). Machine learning-assisted elucidation of CD81–CD44 interactions in promoting cancer stemness and extracellular vesicle integrity. *eLife*, *11*, e82669. https://doi.org/10.7554/eLife.82669

Rana, S., Yue, S., Stadel, D., & Zöller, M. (2012). Toward tailored exosomes: The exosomal tetraspanin web contributes to target cell selection. *The International Journal of Biochemistry & Cell Biology*, *44*(9), 1574–1584. https://doi.org/10.1016/j.biocel.2012.06.018

Riches, A., Campbell, E., Borger, E., & Powis, S. (2014). Regulation of exosome release from mammary epithelial and breast cancer cells – A new regulatory pathway. *European Journal of Cancer*, *50*(5), 1025–1034. https://doi.org/10.1016/j.ejca.2013.12.019

Rini, J. M., Moremen, K. W., Davis, B. G., & Esko, J. D. (2022). Glycosyltransferases and Glycan-Processing Enzymes. In A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, D. Mohnen, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar, & P. H. Seeberger (Eds.), *Essentials of Glycobiology* (4th ed., Vol. 1). Cold Spring Harbor Laboratory Press. http://www.ncbi.nlm.nih.gov/books/NBK579908/

Romanowska, J., Kokh, D. B., & Wade, R. C. (2015). When the Label Matters: Adsorption of Labeled and Unlabeled Proteins on Charged Surfaces. *Nano Letters*, *15*(11), 7508–7513. https://doi.org/10.1021/acs.nanolett.5b03168

Roth, J. (2011). Lectins for histochemical demonstration of glycans. *Histochemistry and Cell Biology*, *136*(2), 117–130. https://doi.org/10.1007/s00418-011-0848-5

Rudd, P. M., Karlsson, N. G., Khoo, K.-H., Thaysen-Andersen, M., Wells, L., & Packer, N. H. (2022). Glycomics and Glycoproteomics. In A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, D.

Mohnen, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar, & P. H. Seeberger (Eds.), *Essentials of Glycobiology* (4th ed., Vol. 1). Cold Spring Harbor Laboratory Press. http://www.ncbi.nlm.nih.gov/books/NBK579904/

Rush, J. S. (2016). Role of Flippases in Protein Glycosylation in the Endoplasmic Reticulum. *Lipid Insights*, *8*(Suppl 1), 45–53. https://doi.org/10.4137/LPI.S31784

Sanchez, J.-F., Lescar, J., Chazalet, V., Audfray, A., Gagnon, J., Alvarez, R., Breton, C., Imberty, A., & Mitchell, E. P. (2006). Biochemical and Structural Analysis of Helix pomatia Agglutinin. *Journal of Biological Chemistry*, *281*(29), 20171–20180. https://doi.org/10.1074/jbc.M603452200

Santos, A. F. S., Silva, M. D. C. da, Napoleão, T. H., Paiva, P. M. G., Correia, M. T. S., & Coelho, L. C. B. B. (2014). *Lectins: Function, structure, biological properties and potential applications*. *1*(1). https://repositorium.sdum.uminho.pt/handle/1822/43440

Schindlbeck, C., Jeschke, U., Schulze, S., Karsten, U., Janni, W., Rack, B., Sommer, H., & Friese, K. (2005). Characterisation of disseminated tumor cells in the bone marrow of breast cancer patients by the Thomsen–Friedenreich tumor antigen. *Histochemistry and Cell Biology*, *123*(6), 631–637. https://doi.org/10.1007/s00418-005-0781-6

Schjoldager, K. T.-B. G., & Clausen, H. (2012). Site-specific protein O-glycosylation modulates proprotein processing—Deciphering specific functions of the large polypeptide GalNAc-transferase gene family. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1820*(12), 2079–2094. https://doi.org/10.1016/j.bbagen.2012.09.014

Schumacher, U., & Adam, E. (1997). Lectin histochemical HPA-binding pattern of human breast and colon cancers is associated with metastases formation in severe combined immunodeficient mice. *The Histochemical Journal*, *29*(9), 677–684. https://doi.org/10.1023/a:1026404832394

Schumacher, U., Higgs, D., Loizidou, M., Pickering, R., Leathem, A., & Taylor, I. (1994). Helix pomatia agglutinin binding is a useful prognostic indicator in colorectal carcinoma. *Cancer*, *74*(12), 3104–3107. https://doi.org/10.1002/1097-0142(19941215)74:12<3104::AID-CNCR2820741207>3.0.CO;2-0

Schwarz, F., & Aebi, M. (2011). Mechanisms and principles of N-linked protein glycosylation. *Current Opinion in Structural Biology*, *21*(5), 576–582. https://doi.org/10.1016/j.sbi.2011.08.005

Scott, D. A., & Drake, R. R. (2019). Glycosylation and its implications in breast cancer. *Expert Review of Proteomics*, *16*(8), 665–680. https://doi.org/10.1080/14789450.2019.1645604

Seale, K. N., & Tkaczuk, K. H. R. (2022). Circulating Biomarkers in Breast Cancer. *Clinical Breast Cancer*, *22*(3), e319–e331. https://doi.org/10.1016/j.clbc.2021.09.006

Seeberger, P. H., Finney, N., Rabuka, D., & Bertozzi, C. R. (2009). Chemical and Enzymatic Synthesis of Glycans and Glycoconjugates. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart, & M. E. Etzler (Eds.), *Essentials of Glycobiology* (2nd ed., Vol. 1). Cold Spring Harbor Laboratory Press. http://www.ncbi.nlm.nih.gov/books/NBK1956/

Shah, R., Sabir, S., & Alhawaj, A. F. (2023). Physiology, Breast Milk. In *StatPearls* (Vol. 1). StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK539790/

Shao, Y., Sun, X., He, Y., Liu, C., & Liu, H. (2015). Elevated Levels of Serum Tumor Markers CEA and CA15-3 Are Prognostic Parameters for Different Molecular Subtypes of Breast Cancer. *PLOS ONE*, *10*(7), e0133830. https://doi.org/10.1371/journal.pone.0133830

Shimoda, A., Miura, R., Tateno, H., Seo, N., Shiku, H., Sawada, S., Sasaki, Y., & Akiyoshi, K. (2022). Assessment of Surface Glycan Diversity on Extracellular Vesicles by Lectin Microarray and Glycoengineering Strategies for Drug Delivery Applications. *Small Methods*, *6*(2), 2100785. https://doi.org/10.1002/smtd.202100785

Shiraishi, T., Atsumi, S., & Yatani, R. (1992). Comparative Study of Prostatic Carcinoma Bone Metastasis among Japanese in Japan and Japanese Americans and Whites in Hawaii. In J. P. Karr & H. Yamanaka (Eds.), *Prostate Cancer and Bone Metastasis* (pp. 7–16). Springer US. https://doi.org/10.1007/978-1-4615-3398-6_2

Simonsen, J. B. (2017). What Are We Looking At? Extracellular Vesicles, Lipoproteins, or Both? *Circulation Research*, *121*(8), 920–922. https://doi.org/10.1161/CIRCRESAHA.117.311767

Sódar, B. W., Kittel, Á., Pálóczi, K., Vukman, K. V., Osteikoetxea, X., Szabó-Taylor, K., Németh, A., Sperlágh, B., Baranyai, T., Giricz, Z., Wiener, Z., Turiák, L., Drahos, L., Pállinger, É., Vékey, K., Ferdinandy, P., Falus, A., & Buzás, E. I. (2016). Low-density lipoprotein mimics blood plasma-derived exosomes and microvesicles during isolation and detection. *Scientific Reports*, *6*, 24316. https://doi.org/10.1038/srep24316

Song, B.-C., Suh, D. J., Yang, S. H., Lee, H. C., Chung, Y.-H., Sung, K.-B., & Lee, Y. S. (2002). Lens culinaris Agglutinin-Reactive Alpha-Fetoprotein as a Prognostic Marker in Patients With Hepatocellular Carcinoma Undergoing Transcatheter Arterial Chemoembolization. *Journal of Clinical Gastroenterology*, *35*(5), 398.

Soule, H. D., Vazquez, J., Long, A., Albert, S., & Brennan, M. (1973). A Human Cell Line From a Pleural Effusion Derived From a Breast Carcinoma2. *JNCI: Journal of the National Cancer Institute*, *51*(5), 1409–1416. https://doi.org/10.1093/jnci/51.5.1409

Springer, G. F. (1984). T and Tn, General Carcinoma Autoantigens. *Science*, *224*(4654), 1198–1206. https://doi.org/10.1126/science.6729450

Springer, G. F. (1989). Tn epitope (N-acetyl-d-galactosamineα -O -serine/threonine) density in primary breast carcinoma: A functional predictor of aggressiveness. *Molecular Immunology*, *26*(1), 1–5. https://doi.org/10.1016/0161-5890(89)90013-8

Springer, G. F. (1997). Immunoreactive T and Tn epitopes in cancer diagnosis, prognosis, and immunotherapy. *Journal of Molecular Medicine*, *75*(8), 594–602. https://doi.org/10.1007/s001090050144

Stefani, F., Zhang, L., Taylor, S., Donovan, J., Rollinson, S., Doyotte, A., Brownhill, K., Bennion, J., Pickering-Brown, S., & Woodman, P. (2011). UBAP1 Is a Component of an Endosome-Specific ESCRT-I Complex that Is Essential for MVB Sorting. *Current Biology*, *21*(14), 1245–1250. https://doi.org/10.1016/j.cub.2011.06.028

Stipp, C. S., Kolesnikova, T. V., & Hemler, M. E. (2003). Functional domains in tetraspanin proteins. *Trends in Biochemical Sciences*, *28*(2), 106–112. https://doi.org/10.1016/S0968-0004(02)00014-2

Stowell, S. R., Ju, T., & Cummings, R. D. (2015). Protein Glycosylation in Cancer. *Annual Review of Pathology*, *10*, 473–510. https://doi.org/10.1146/annurev-pathol-012414-040438

Stuffers, S., Sem Wegner, C., Stenmark, H., & Brech, A. (2009). Multivesicular Endosome Biogenesis in the Absence of ESCRTs. *Traffic*, *10*(7), 925–937. https://doi.org/10.1111/j.1600-0854.2009.00920.x

Sukhorukov, V., Gudelj, I., Pučić-Baković, M., Zakiev, E., Orekhov, A., Kontush, A., & Lauc, G. (2019). Glycosylation of human plasma lipoproteins reveals a high level of diversity, which directly impacts their functional properties. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1864*(5), 643–653. https://doi.org/10.1016/j.bbalip.2019.01.005

Sun, X., Zhan, M., Sun, X., Liu, W., & Meng, X. (2021). C1GALT1 in health and disease. *Oncology Letters*, *22*(2), 589. https://doi.org/10.3892/ol.2021.12850

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, *71*(3), 209–249. https://doi.org/10.3322/caac.21660

Tandon, A. K., Clark, G. M., Chamness, G. C., & McGuire, W. L. (1990). Association of the 323/A3 Surface Glycoprotein with Tumor Characteristics and Behavior in Human Breast Cancer1. *Cancer Research*, *50*(11), 3317–3321.

Taylor, M. E., Drickamer, K., Taylor, M. E., & Drickamer, K. (2011). *Introduction to Glycobiology* (Third Edition, Third Edition, Vol. 1). Oxford University Press.

Teng, F., & Fussenegger, M. (2021). Shedding Light on Extracellular Vesicle Biogenesis and Bioengineering. *Advanced Science*, *8*(1), 2003505. https://doi.org/10.1002/advs.202003505

Teo, H., Perisic, O., González, B., & Williams, R. L. (2004). ESCRT-II, an Endosome-Associated Complex Required for Protein Sorting: Crystal Structure and Interactions with ESCRT-III and Membranes. *Developmental Cell*, *7*(4), 559–569. https://doi.org/10.1016/j.devcel.2004.09.003

Terävä, J., Verhassel, A., Botti, O., Islam, Md. K., Leivo, J., Wittfooth, S., Härkönen, P., Pettersson, K., & Gidwani, K. (2022). Primary breast cancer biomarkers based on glycosylation and extracellular vesicles detected from human serum. *Cancer Reports*, *5*(8), e1540. https://doi.org/10.1002/cnr2.1540

Théry, C., Witwer, K. W., Aikawa, E., Alcaraz, M. J., Anderson, J. D., Andriantsitohaina, R., Antoniou, A., Arab, T., Archer, F., Atkin-Smith, G. K., Ayre, D. C., Bach, J.-M., Bachurski, D., Baharvand, H., Balaj, L., Baldacchino, S., Bauer, N. N., Baxter, A. A., Bebawy, M., … Zuba-Surma, E. K. (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

Thorens, B., & Vassalli, P. (1986). Chloroquine and ammonium chloride prevent terminal glycosylation of immunoglobulins in plasma cells without affecting secretion. *Nature*, *321*(6070), 618–620. https://doi.org/10.1038/321618a0

Tian, F., Zhang, S., Liu, C., Han, Z., Liu, Y., Deng, J., Li, Y., Wu, X., Cai, L., Qin, L., Chen, Q., Yuan, Y., Liu, Y., Cong, Y., Ding, B., Jiang, Z., & Sun, J. (2021). Protein analysis of extracellular vesicles to monitor and predict therapeutic response in metastatic breast cancer. *Nature Communications*, *12*, 2536. https://doi.org/10.1038/s41467-021-22913-7

Trajkovic, K., Hsu, C., Chiantia, S., Rajendran, L., Wenzel, D., Wieland, F., Schwille, P., Brügger, B., & Simons, M. (2008). Ceramide Triggers Budding of Exosome Vesicles into Multivesicular Endosomes. *Science*, *319*(5867), 1244–1247. https://doi.org/10.1126/science.1153124

Tran, D. T., & Ten Hagen, K. G. (2013). Mucin-type O-Glycosylation during Development. *The Journal of Biological Chemistry*, *288*(10), 6921–6929. https://doi.org/10.1074/jbc.R112.418558

Tricarico, C., Clancy, J., & D'Souza-Schorey, C. (2016). Biology and biogenesis of shed microvesicles. *Small GTPases*, *8*(4), 220–232. https://doi.org/10.1080/21541248.2016.1215283

Uehara, M., Kinoshita, T., Hojo, T., Akashi-Tanaka, S., Iwamoto, E., & Fukutomi, T. (2008). Long-term prognostic study of carcinoembryonic antigen (CEA) and carbohydrate antigen 15-3 (CA 15-3) in breast cancer. *International Journal of Clinical Oncology*, *13*(5), 447–451. https://doi.org/10.1007/s10147-008-0773-3

van de Wakker, S. I., van Oudheusden, J., Mol, E. A., Roefs, M. T., Zheng, W., Görgens, A., El Andaloussi, S., Sluijter, J. P. G., & Vader, P. (2022). Influence of short-term storage conditions, concentration methodsand excipients on extracellular vesicle recovery and function. *European Journal of Pharmaceutics and Biopharmaceutics*, *170*, 59–69. https://doi.org/10.1016/j.ejpb.2021.11.012

van Niel, G., Charrin, S., Simoes, S., Romao, M., Rochin, L., Saftig, P., Marks, M. S., Rubinstein, E., & Raposo, G. (2011). The tetraspanin CD63 regulates ESCRT-independent and dependent endosomal sorting during melanogenesis. *Developmental Cell*, *21*(4), 708–721. https://doi.org/10.1016/j.devcel.2011.08.019

van Niel, G., D'Angelo, G., & Raposo, G. (2018). Shedding light on the cell biology of extracellular vesicles. *Nature Reviews Molecular Cell Biology*, *19*(4), Article 4. https://doi.org/10.1038/nrm.2017.125

Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., & Seeberger, P. H. (Eds.). (2022). *Essentials of Glycobiology* (4th ed., Vol. 1). Cold Spring Harbor Laboratory Press. http://www.ncbi.nlm.nih.gov/books/NBK579918/

Vences-Catalán, F., Duault, C., Kuo, C.-C., Rajapaksa, R., Levy, R., & Levy, S. (2017). CD81 as a tumor target. *Biochemical Society Transactions*, *45*(2), 531–535. https://doi.org/10.1042/BST20160478

Vitale, S. R., Helmijr, J. A., Gerritsen, M., Coban, H., van Dessel, L. F., Beije, N., van der Vlugt-Daane, M., Vigneri, P., Sieuwerts, A. M., Dits, N., van Royen, M. E., Jenster, G., Sleijfer, S., Lolkema, M., Martens, J. W. M., & Jansen, M. P. H. M. (2021). Detection of tumor-derived extracellular vesicles in plasma from patients with solid cancer. *BMC Cancer*, *21*(1), 315. https://doi.org/10.1186/s12885-021-08007-z

Wandall, H. H., Nielsen, M. A. I., King-Smith, S., de Haan, N., & Bagdonaite, I. (2021). Global functions of O-glycosylation: Promises and challenges in O-glycobiology. *The FEBS Journal*, *288*(24), 7183–7212. https://doi.org/10.1111/febs.16148

Wang, Y., Xiao, T., Zhao, C., & Li, G. (2023). The Regulation of Exosome Generation and Function in Physiological and Pathological Processes. *International Journal of Molecular Sciences*, *25*(1), 255. https://doi.org/10.3390/ijms25010255

Webber, J., & Clayton, A. (2013). How pure are your vesicles? *Journal of Extracellular Vesicles*, *2*, 10.3402/jev.v2i0.19861. https://doi.org/10.3402/jev.v2i0.19861

Wei, D., Zhan, W., Gao, Y., Huang, L., Gong, R., Wang, W., Zhang, R., Wu, Y., Gao, S., & Kang, T. (2021). RAB31 marks and controls an ESCRT-independent exosome pathway. *Cell Research*, *31*(2), Article 2. https://doi.org/10.1038/s41422-020-00409-1

Welsh, J. A., Killingsworth, B., Kepley, J., Traynor, T., McKinnon, K., Savage, J., Appel, D., Aldape, K., Camphausen, K., Berzofsky, J. A., Ivanov, A. R., Ghiran, I. H., & Jones, J. C. (2021). A simple, high-throughput method of protein and label removal from extracellular vesicle samples. *Nanoscale*, *13*(6), 3737–3745. https://doi.org/10.1039/d0nr07830a

Welsh, J. A., Van Der Pol, E., Arkesteijn, G. J. A., Bremer, M., Brisson, A., Coumans, F., Dignat-George, F., Duggan, E., Ghiran, I., Giebel, B., Görgens, A., Hendrix, A., Lacroix, R., Lannigan, J., Libregts, S. F. W. M., Lozano-Andrés, E., Morales-Kastresana, A., Robert, S., De Rond, L., … Jones, J. C. (2020). MIFlowCyt-EV: A framework for standardized reporting of extracellular vesicle flow cytometry experiments. *Journal of Extracellular Vesicles*, *9*(1), 1713526. https://doi.org/10.1080/20013078.2020.1713526

Wenzel, E. M., Schultz, S. W., Schink, K. O., Pedersen, N. M., Nähse, V., Carlson, A., Brech, A., Stenmark, H., & Raiborg, C. (2018). Concerted ESCRT and clathrin recruitment waves define the timing and morphology of intraluminal vesicle formation. *Nature Communications*, *9*, 2932. https://doi.org/10.1038/s41467-018-05345-8

Williams, C., Pazos, R., Royo, F., González, E., Roura-Ferrer, M., Martinez, A., Gamiz, J., Reichardt, N.-C., & Falcón-Pérez, J. M. (2019). Assessing the role of surface glycans of extracellular vesicles on cellular uptake. *Scientific Reports*, *9*, 11920. https://doi.org/10.1038/s41598-019-48499-1

Williams, C., Royo, F., Aizpurua-Olaizola, O., Pazos, R., Boons, G.-J., Reichardt, N.-C., & Falcon-Perez, J. M. (2018). Glycosylation of extracellular vesicles: Current knowledge, tools and clinical perspectives. *Journal of Extracellular Vesicles*, *7*(1), 1442985. https://doi.org/10.1080/20013078.2018.1442985

Wollert, T., & Hurley, J. H. (2010). Molecular Mechanism of Multivesicular Body Biogenesis by ESCRT Complexes. *Nature*, *464*(7290), 864–869. https://doi.org/10.1038/nature08849

Wollert, T., Wunder, C., Lippincott-Schwartz, J., & Hurley, J. H. (2009). Membrane Scission by the ESCRT-III Complex. *Nature*, *458*(7235), 172–177. https://doi.org/10.1038/nature07836

Wu, A. M., & Sugii, S. (1991). Coding and classification of d-galactose, N-acetyl-d-galactosamine, and β-d-Gal-[1→3(4)]-β-d-GlcNAc, specificities of applied lectins. *Carbohydrate Research*, *213*, 127–143. https://doi.org/10.1016/S0008-6215(00)90604-9

Wu, L., & Gao, C. (2023). Comprehensive Overview the Role of Glycosylation of Extracellular Vesicles in Cancers. *ACS Omega*, *8*(50), 47380–47392. https://doi.org/10.1021/acsomega.3c07441

Xi, Y., & Xu, P. (2021). Global colorectal cancer burden in 2020 and projections to 2040. *Translational Oncology*, *14*(10), 101174. https://doi.org/10.1016/j.tranon.2021.101174

Xu, M., Ji, J., Jin, D., Wu, Y., Wu, T., Lin, R., Zhu, S., Jiang, F., Ji, Y., Bao, B., Li, M., Xu, W., & Xiao, M. (2022). The biogenesis and secretion of exosomes and multivesicular bodies (MVBs): Intercellular shuttles and implications in human diseases. *Genes & Diseases*, *10*(5), 1894–1907. https://doi.org/10.1016/j.gendis.2022.03.021

Yamamoto, K., Imamura, H., Matsuyama, Y., Kume, Y., Ikeda, H., Norman, G. L., Shums, Z., Aoki, T., Hasegawa, K., Beck, Y., Sugawara, Y., & Kokudo, N. (2010). AFP, AFP-L3, DCP, and GP73 as markers for monitoring treatment response and recurrence and as surrogate markers of clinicopathological variables of HCC. *Journal of Gastroenterology*, *45*(12), 1272–1282. https://doi.org/10.1007/s00535-010-0278-5

Yáñez-Mó, M., Mitᵀᴱᴸbrunn, M., & Sánchez-Madrid, F. (2001). Tetraspanins and Intercellular Interactions. *Microcirculation*, *8*(3), 153–168. https://doi.org/10.1111/j.1549-8719.2001.tb00166.x

Yáñez-Mó, M., Siljander, P. R.-M., Andreu, Z., Zavec, A. B., Borràs, F. E., Buzas, E. I., Buzas, K., Casal, E., Cappello, F., Carvalho, J., Colás, E., Silva, A. C., Fais, S., Falcon-Perez, J. M., Ghobrial, I. M., Giebel, B., Gimona, M., Graner, M., Gursel, I., … De Wever, O. (2015). Biological properties of extracellular vesicles and their physiological functions. *Journal of Extracellular Vesicles*, *4*, 10.3402/jev.v4.27066. https://doi.org/10.3402/jev.v4.27066

Yang, Y., Wang, Y., Wei, S., Zhou, C., Yu, J., Wang, G., Wang, W., & Zhao, L. (2021). Extracellular vesicles isolated by size-exclusion chromatography present suitability for RNomics analysis in plasma. *Journal of Translational Medicine*, *19*(1), 104. https://doi.org/10.1186/s12967-021-02775-9

Yates, A. G., Pink, R. C., Erdbrügger, U., Siljander, P. R.-M., Dellar, E. R., Pantazi, P., Akbar, N., Cooke, W. R., Vatish, M., Dias-Neto, E., Anthony, D. C., & Couch, Y. (2022). In sickness and in health: The functional

role of extracellular vesicles in physiology and pathology in vivo: Part I: Health and Normal Physiology: Part I: Health and Normal Physiology. *Journal of Extracellular Vesicles*, *11*(1), e12151. https://doi.org/10.1002/jev2.12151

Yokose, T., Kabe, Y., Matsuda, A., Kitago, M., Matsuda, S., Hirai, M., Nakagawa, T., Masugi, Y., Hishiki, T., Nakamura, Y., Shinoda, M., Yagi, H., Abe, Y., Oshima, G., Hori, S., Nakano, Y., Honda, K., Kashiro, A., Morizane, C., … Kitagawa, Y. (2020). O-Glycan-Altered Extracellular Vesicles: A Specific Serum Marker Elevated in Pancreatic Cancer. *Cancers*, *12*(9), 2469. https://doi.org/10.3390/cancers12092469

Yoshida, Y., Okamura, T., & Shirakusa, T. (1993). An immunohistochemical study of helix pomatia agglutinin binding on carcinomas of the esophagus. *Surgery, Gynecology & Obstetrics*, *177*(3), 299–302.

Zaia, J. (2010). Mass Spectrometry and Glycomics. *OMICS: A Journal of Integrative Biology*, *14*(4), 401–418. https://doi.org/10.1089/omi.2009.0146

Zhang, A., Skog, S., Wang, S., Ke, Y., Zhang, Y., Li, K., He, E., & Li, N. (2016). A Chemiluminescent Protein Microarray Method for Determining the Seroglycoid Fucosylation Index. *Scientific Reports*, *6*, 31132. https://doi.org/10.1038/srep31132

Zhang, H., Freitas, D., Kim, H. S., Fabijanic, K., Li, Z., Chen, H., Mark, M. T., Molina, H., Martin, A. B., Bojmar, L., Fang, J., Rampersaud, S., Hoshino, A., Matei, I., Kenific, C. M., Nakajima, M., Mutvei, A. P., Sansone, P., Buehring, W., … Lyden, D. (2018). Identification of distinct nanoparticles and subsets of extracellular vesicles by asymmetric-flow field-flow fractionation. *Nature Cell Biology*, *20*(3), 332–343. https://doi.org/10.1038/s41556-018-0040-4

Zhang, N., Zuo, L., Zheng, H., Li, G., & Hu, X. (2018). Increased Expression of CD81 in Breast Cancer Tissue is Associated with Reduced Patient Prognosis and Increased Cell Migration and Proliferation in MDA-MB-231 and MDA-MB-435S Human Breast Cancer Cell Lines In Vitro. *Medical Science Monitor*, *24*, 5739–5747. https://doi.org/10.12659/MSM.911612

Zhang, Q., Jeppesen, D. K., Higginbotham, J. N., Graves-Deal, R., Trinh, V. Q., Ramirez, M. A., Sohn, Y., Neininger, A. C., Taneja, N., McKinley, E. T., Niitsu, H., Cao, Z., Evans, R., Glass, S. E., Ray, K. C., Fissell, W. H., Hill, S., Rose, K. L., Huh, W. J., … Coffey, R. J. (2021). Supermeres are functional extracellular nanoparticles replete with disease biomarkers and therapeutic targets. *Nature Cell Biology*, *23*(12), 1240–1254. https://doi.org/10.1038/s41556-021-00805-8

Zhang, Y., Liu, Y., Liu, H., & Tang, W. H. (2019). Exosomes: Biogenesis, biologic function and clinical potential. *Cell & Bioscience*, *9*(1), 19. https://doi.org/10.1186/s13578-019-0282-2

Zhang, Y., Sun, L., Lei, C., Li, W., Han, J., Zhang, J., & Zhang, Y. (2022). A Sweet Warning: Mucin-Type O-Glycans in Cancer. *Cells*, *11*(22), 3666. https://doi.org/10.3390/cells11223666

Zheng, W., He, R., Liang, X., Roudi, S., Bost, J., Coly, P.-M., van Niel, G., & Andaloussi, S. E. L. (2022). Cell-specific targeting of extracellular vesicles through engineering the glycocalyx. *Journal of Extracellular Vesicles*, *11*(12), e12290. https://doi.org/10.1002/jev2.12290

APPENDIX 1 – ETHICAL APPROVAL

UNIVERSITY RESEARCH ETHICS COMMITTEE

REQUEST FOR STUDY AMENDMENT APPROVAL BY CHAIR'S ACTION

This application form should only be completed for a UREC Ethics Application that has previously received **full ethics approval**.

UREC Registration Number:	
Project Title:	**Cell Surface Proteins on Breast Cancer Cells**

RESEARCH TEAM: *Research projects should include all members of the research team and indicate any changes to the team from the original approval. Doctoral students SHOULD be listed after their supervisors, identifying the Director of Studies.*

TITLE & NAME	POST	DEPT & FACULTY / DIRECTORATE	PHONE	EMAIL
Dr Ryan Pink	Senior Lecturer	HLS/BMS	3607	rpink@brookes.ac.uk
Prof Susan Brooks	Professor of Cell Biology	HLS/BMS	3285	sbrooks@brookes.ac.uk
Dr Ellie Beaman	Post-Doc Researcher	HLS/BMS	8635	ebeaman@brookes.ac.uk
Jamie Cooper	PhD researcher	HLS/BMS	N/A	19125250@brookes.ac.uk

DATE FULL APPROVAL RECEIVED:	1ST JULY 2020	DATE APPROVAL EXPIRES:	1ST JULY 2023

OXFORD
BROOKES
UNIVERSITY

Dr Ellie-May Beaman
Faculty of Health and Life Sciences
Oxford Brookes University

1st July 2020

Dear Dr Beaman,

Study Title: Cell surface proteins on breast cancer cells

Thank you for your email of 30th June 2020 outlining your study and attaching the necessary documents. Based on the supplied information, I am pleased to inform you that I give Chair's approval for this study.

The UREC approval period for the data collection phase of the study is two years from the date of this letter / six months after the expected completion date, so until 1st July 2022. If you need the approval to be extended please do contact me nearer the time of expiry.

Should the recruitment, methodology or data storage change from your original plans, or should any study participants experience adverse physical, psychological, social, legal or economic effects from the research, please inform me with full details as soon as possible.

Yours sincerely,

Dr Sarah Quinton
Chair of the University Research Ethics Committee

cc Prof. Susan Brookes, Co-investigator
   Ms Kellie Tune, Research Ethics Officer

**OXFORD
BROOKES
UNIVERSITY**

Dr Ellie-May Beaman
Faculty of Health and Life Sciences
Oxford Brookes University

8th September 2021

Dear Dr Beaman,

Study Title: Cell surface proteins on breast cancer cells

Thank you for your email of 7th September 2021 requesting an amendment to the original study approved by UREC on 1st July 2020.
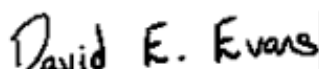
You would like to include human plasma samples and amend the research team to due to a change of personnel. You have provided the relevant updated documentation.

There are no other changes to the study and on this basis I give Chair's approval for this change. In addition, you have requested an extension to the UREC approval. A further year has now been given, and UREC approval is now valid until 1st July 2023.

Should the recruitment, methodology or data storage change from your original plans, or should any study participants experience adverse physical, psychological, social, legal or economic effects from the research, please inform me with full details as soon as possible.

I wish you continued success with your research.

Yours sincerely,

David E. Evans

Prof. David Evans
Chair of the University Research Ethics Committee

cc Prof. Susan Brookes, Co-investigator
    Dr Ryan Pink, Co-investigator
    Dr Adam White, Research Ethics Officer
    Dr Robyn Curtis, Research Ethics & Integrity Officer

**INVESTORS
IN PEOPLE** | Gold

www.brookes.ac.uk

PRECISION
for medicine

Human Subjects Protection Statement of Compliance

Order:SO-1-45303

Through the signature of its appointed representative below, Precision For Medicine certifies that the biospecimens contained within this order have been collected under a clinical study that has been reviewed by an Institutional/Independent Review Board (IRB) and/or Independent Ethics Committee (IEC) in accordance with requirements of local governing regulatory agencies, including the Department of Health and Human Services (DHHS) and Food and Drug Administration (FDA) Codes of Federal Regulations, on the Protection of Human Subjects (45 CFR Part 46 and 21 CFR Part 56, respectively).

Precision For Medicine also certifies that these biospecimens have been collected in compliance with, as applicable: i) 21 CFR 50, through obtaining informed consent from human research subjects for use of their biological materials in medical research, or ii) the FDA guidance document "Guidance on Informed Consent for In Vitro Diagnostic Device Studies Using Leftover Human Specimens that are Not Individually Identifiable" (April 25, 2006), through the biospecimens meeting all criteria of "human remnant specimens" making them exempt from the requirements of informed consent.

To the extent defined above, these biospecimens meet requirements for their use in existing and future investigational research, quality control, and clinical studies to support applications for marketing permits for products, including in vitro diagnostic test kits, regulated by the FDA.

Signature: Marina Wright

Date: 7 / 30 / 2021

APPENDIX 2 – SUPPLEMENTARY MATERIALS

Supplementary Figure 1. **Flow cytometry optimisation of lectins using MCF-7 cells with concentrations ranging from 1 to 5 μg/ml at 10, 30, and 60-minute incubation times, employing a stain index calculation. Methods outlined in section 2.3. (A)** *HPA lectin optimisation,* **(B)** *TL optimisation,* **(C)** *LCA lectin optimisation.*
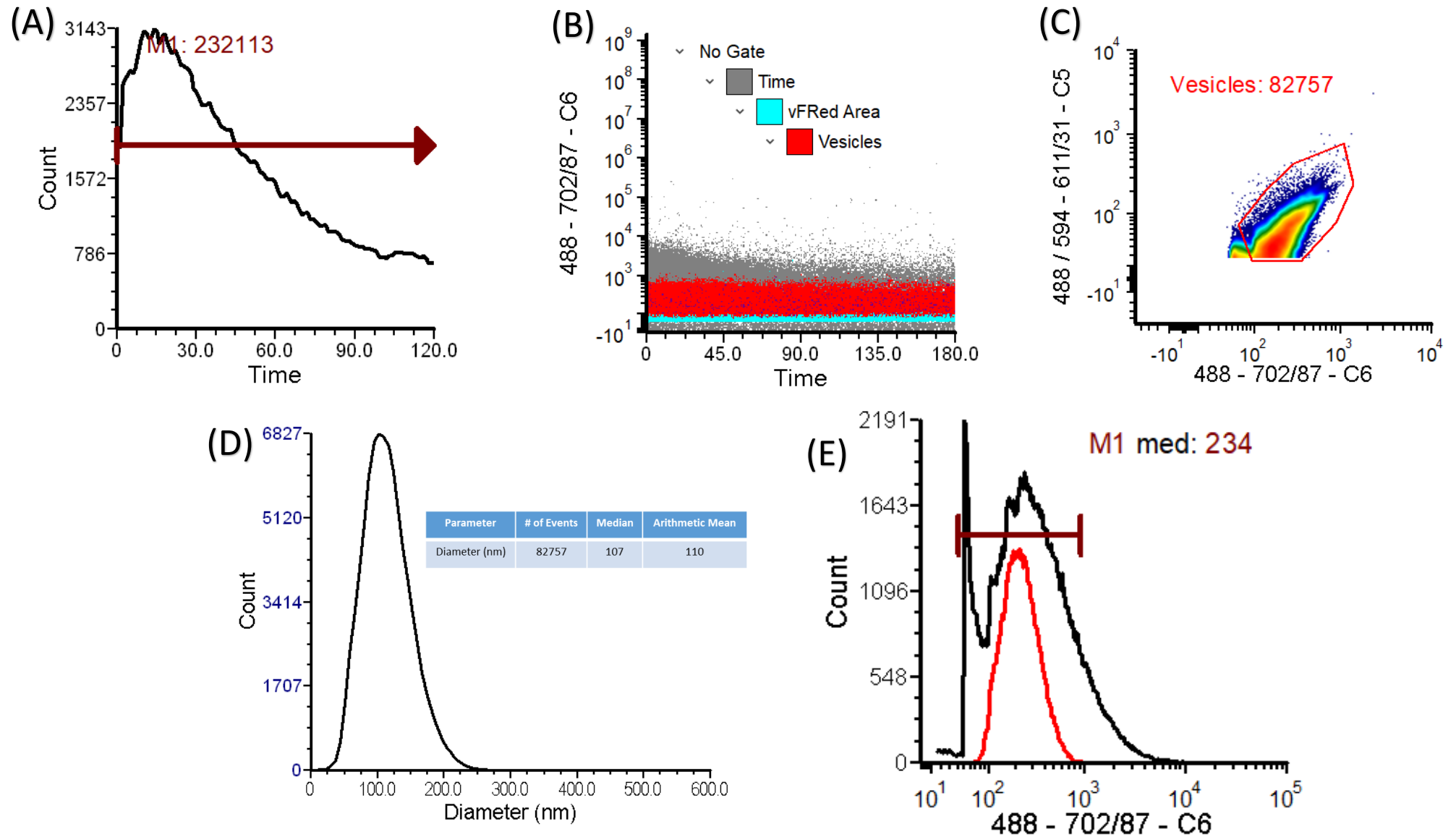
Supplementary Figure 2. **CellStream sEVs size calibration. Methods outlined in section 2.6.2. (A-B)** *Vesicle flow cytometry time gate of vFRed ^{TM} stained Lipo100 ^{TM} vesicle size standard to eliminate background from fluidic disturbances* **(C)** *The selection of events with vFRed ^{TM} stained Lipo100 ^{TM}, pulse height and area that can eliminate low intensity background events* **(D)** *Vesicle size calibration of Lipo100 ^{TM} with a measured diameter distribution measured by NTA* **(E)** *The fluorescence distribution of vFRed ^{TM} stained Lipo100 ^{TM} vesicle standards which is proportional to the surface area distribution.*
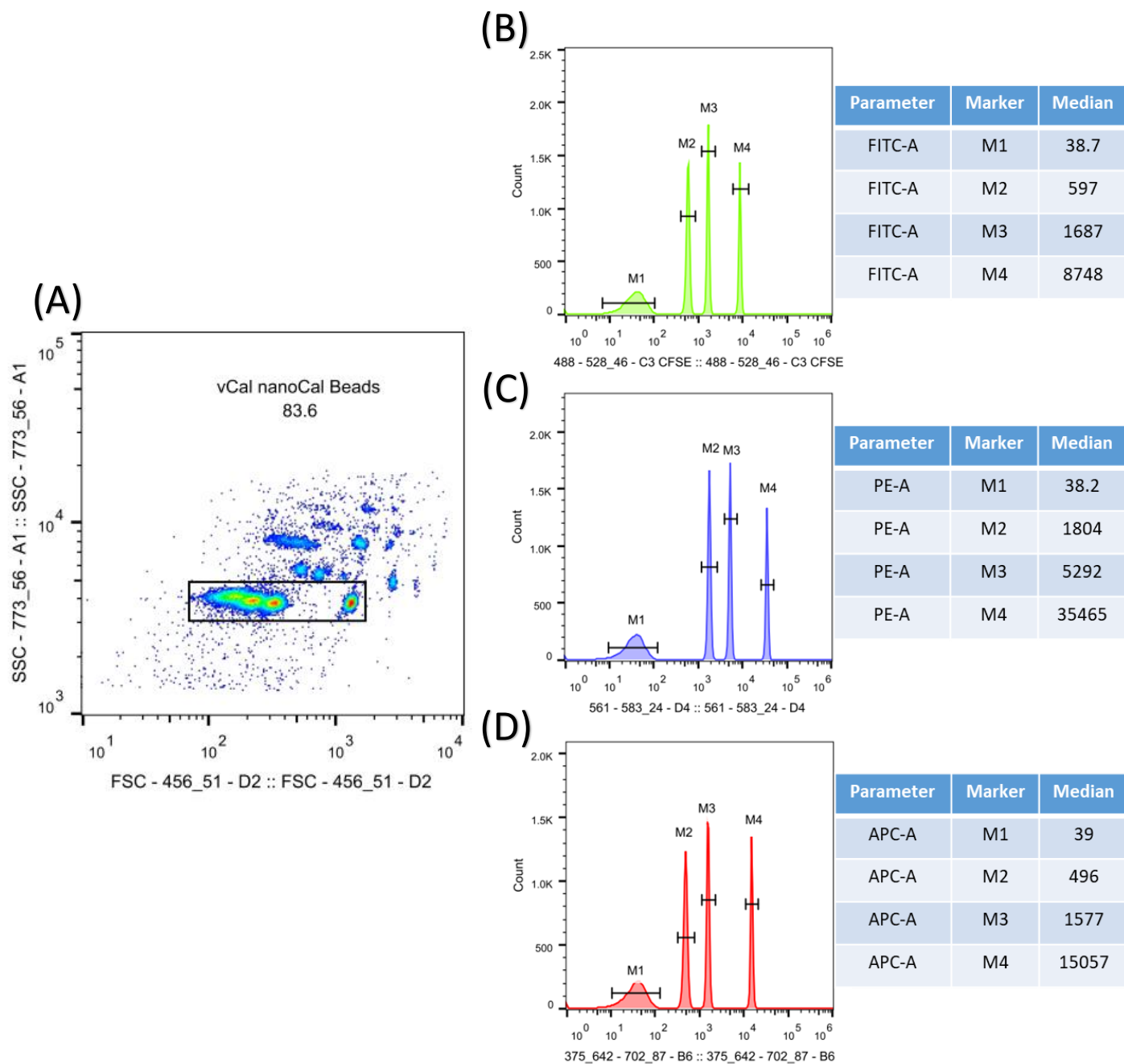
Supplementary Figure 3. **CellStream sEVs fluorescence calibration employing vCal^TM nanorainbow beads to calculate MESF standardised units from median values. Methods outlined in section 2.6.2. (A)** *FSC and SSC gate to identify vCal^TM nanorainbow bead population* **(B)** *Histogram gating of vCal™ nanorainbow beads at 488 nm wavelength to determine median values corresponding to known MESF values* **(C)** *561 nm wavelength* **(D)** *647 nm wavelength.*
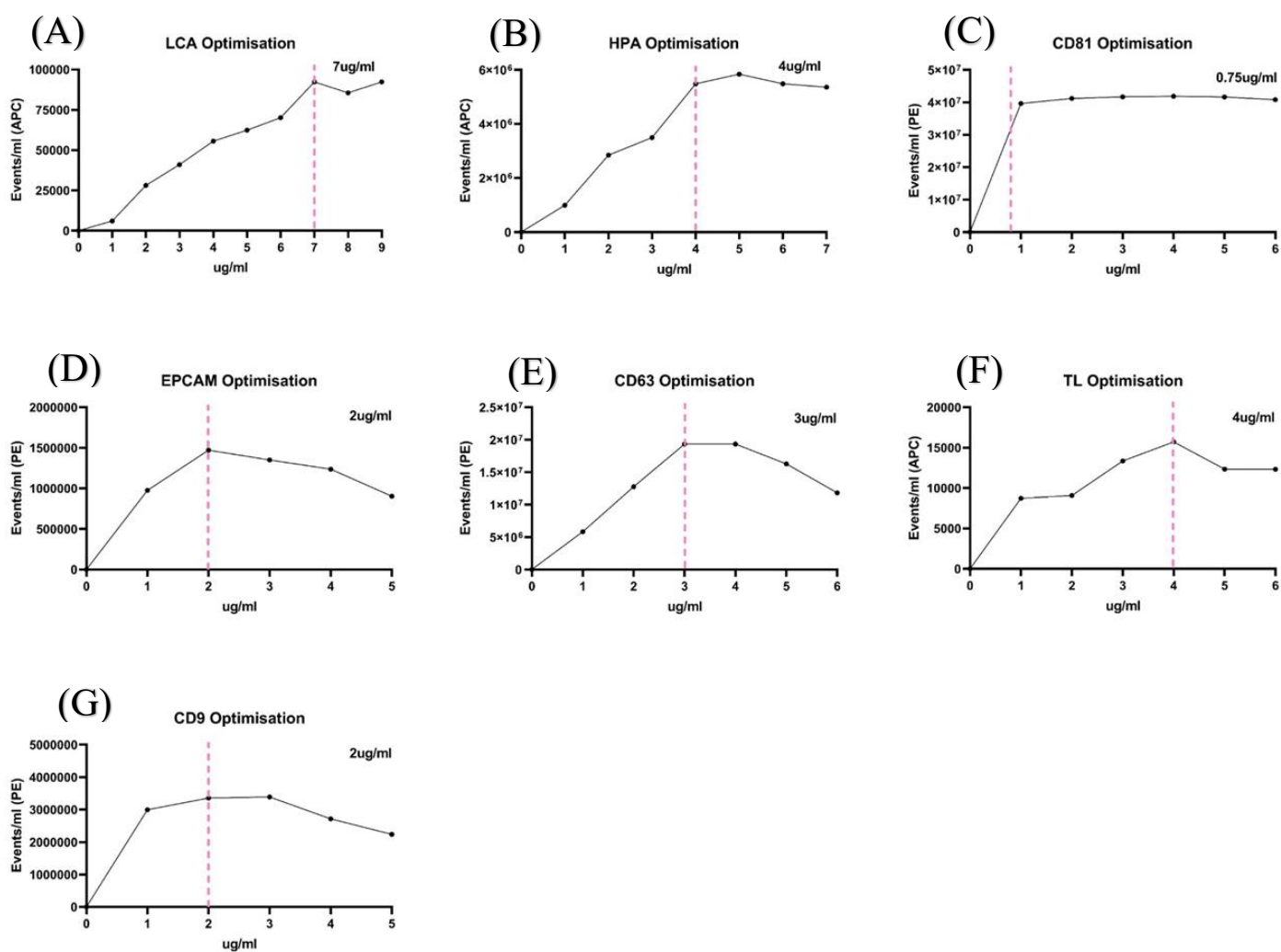
Supplementary Figure 4. **Optimisation of the antibodies and lectins used for single-vesicle analysis by determining the optimum events/ml achieved in MCF-7 sEVs. Methods outlined in section 2.6.3. (A)** *LCA DyLight 649 optimisation 8µg/ml* **(B)** *HPA Alexa Fluor 647 optimisation 4µg/ml* **(C)** *CD81 PE 561 optimisation 0.75µg/ml* **(D)** *EpCAM PE 561 optimisation 2µg/ml* **(E)** *CD63 PE 561 optimisation 3µg/ml* **(F)** *TL Cy5 650 optimisation 4µg/ml* **(G)** *CD9 PE 561 optimisation 2µg/ml.*
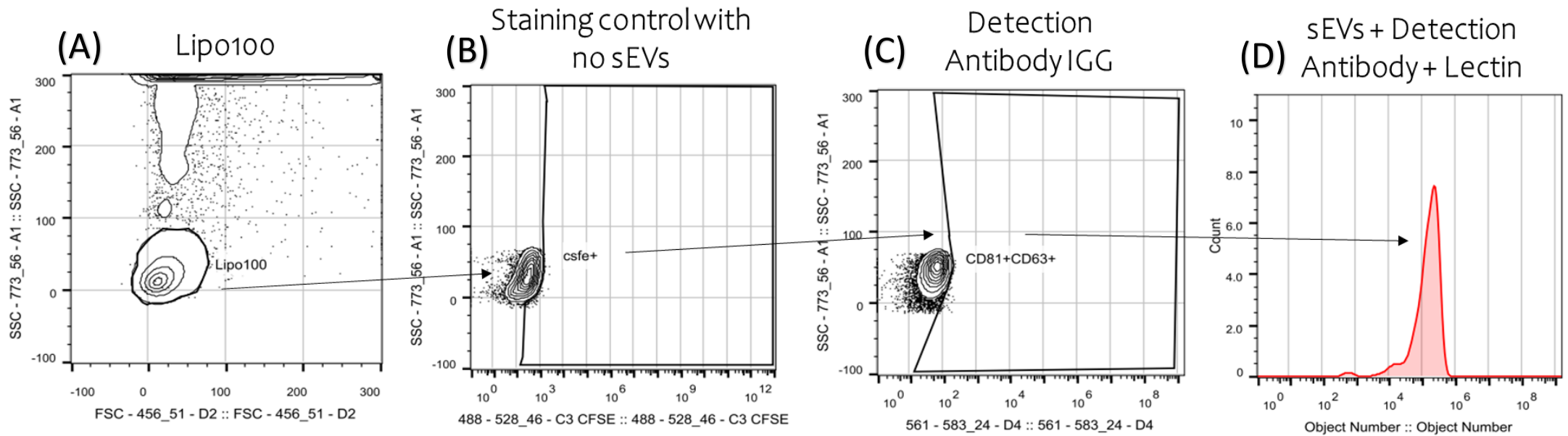
Supplementary Figure 5. **Gating strategy adopted to identify sEVs. Methods outlined in section 2.6.3. (A)** *SSC and FSC gate to identify vFRed ^{TM} stained Lipo100 ^{TM} vesicle size standards used to size calibrate instrument.* **(B)** *Control samples comprised of staining reagents used in experiment without sEVs to omit background signal.* **(C)** *Isotype control of detection antibodies used in experiment to omit background signal* **(D)** *The quantified signal of interest after gating comprising of staining reagents used in experiment with sEVs.*
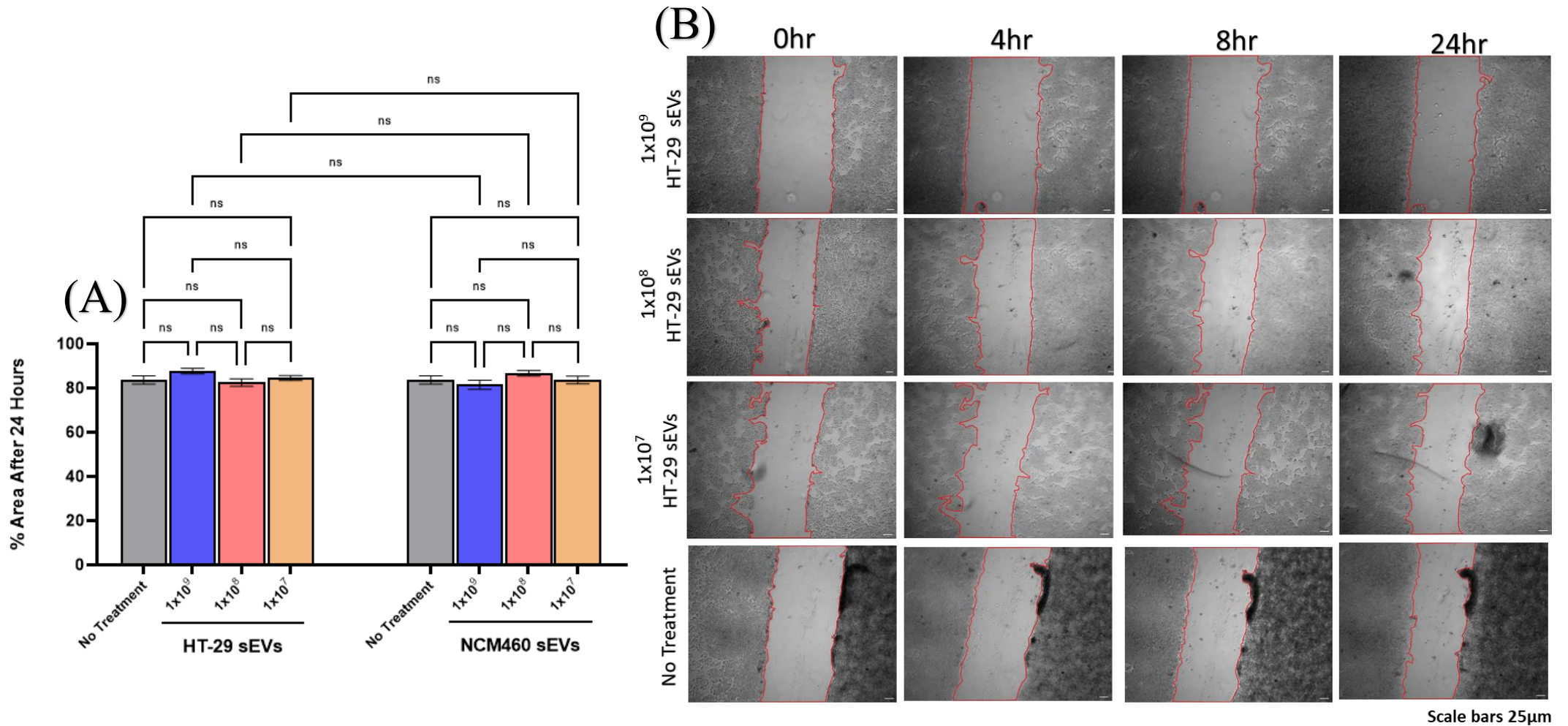
Supplementary Figure 6. Wound healing assay assessment of colorectal NCM460 cells. Methods outlined in section 2.7.2. (A) *The migration area of NCM460 cells after treatment with different densities of HT-29 and NCM460 sEVs (1×10^9, 1×10^8, and 1×10^7 particles/ml, as determined by NTA) alongside a no-treatment control comprising growth medium.* **(B)** *Representative images of the NCM460 scratch assay used to quantify cell migration at 0, 4, 8, and 24 h time points. Quantified area shown in red. Scale bars are 25μm. Each experiment consisted of biological triplicates with three technical triplicates, and error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. For parametric t-tests, Welch's correction was utilised, and for nonparametric t-tests, a Mann-Whitney test was used.*
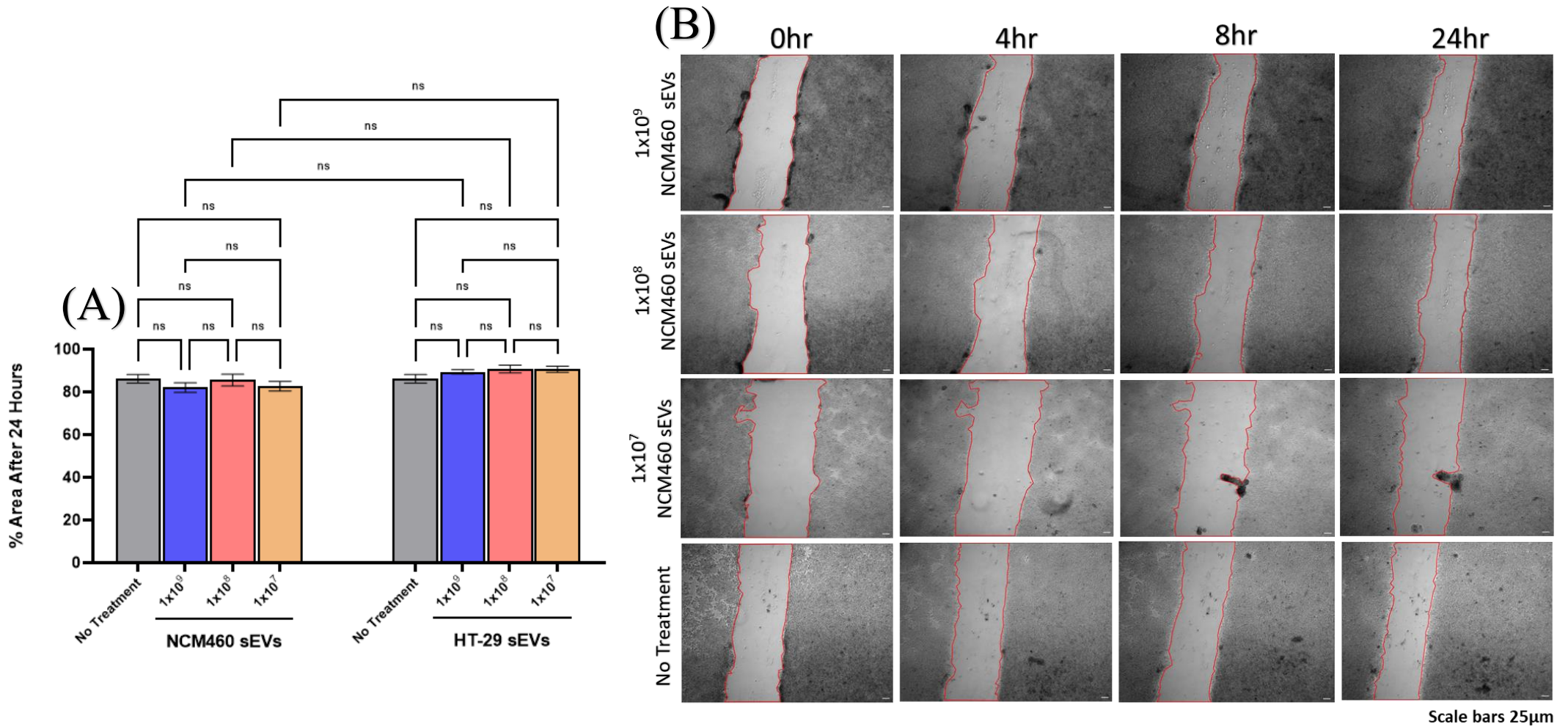
Supplementary Figure 7. Wound healing assay assessment of colorectal HT-29 cells. Methods outlined in section 2.7.2. (A) *The migration area of HT-29 cells after treatment with different densities of NCM460 and HT-29 sEVs (1×10^9, 1×10^8, and 1×10^8 particles/ml, as determined by NTA) alongside a no-treatment control comprising growth medium. **(B)** Representative images of the HT-29 scratch assay used to quantify cell migration at 0, 4, 8, and 24 h time points. Quantified area shown in red. Scale bars are 25μm. Each experiment consisted of biological triplicates with three technical triplicates, and error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. For parametric t-tests, Welch's correction was utilised, and for nonparametric t-tests, a Mann-Whitney test was used.*
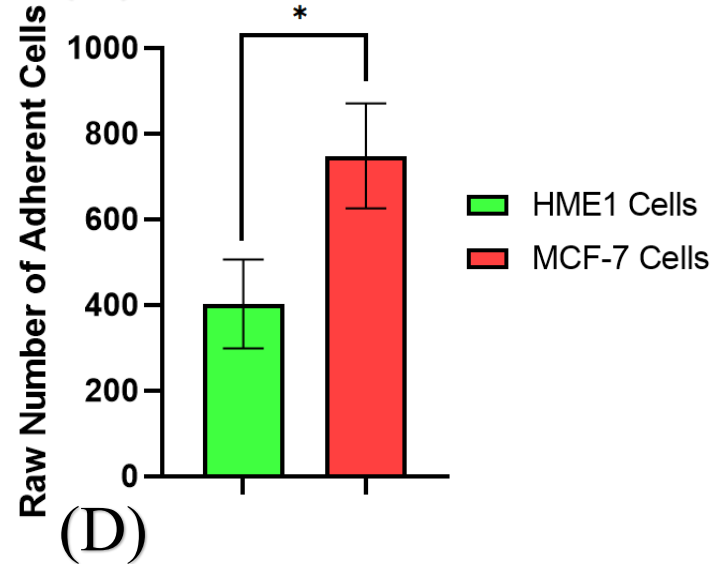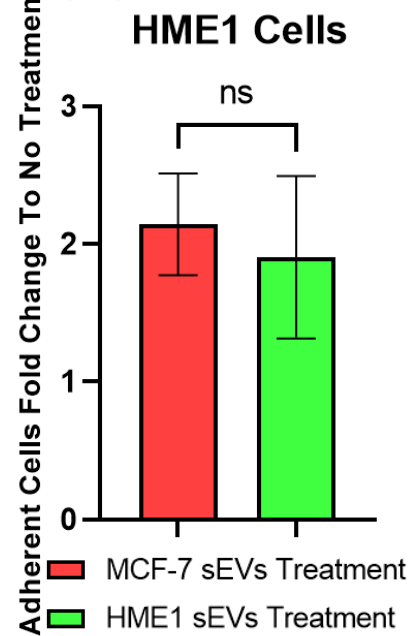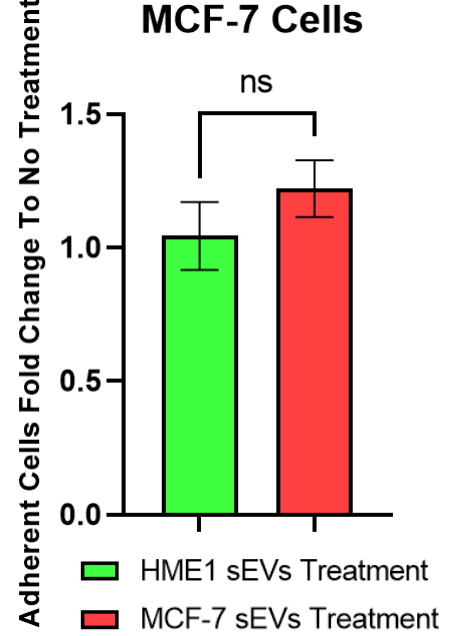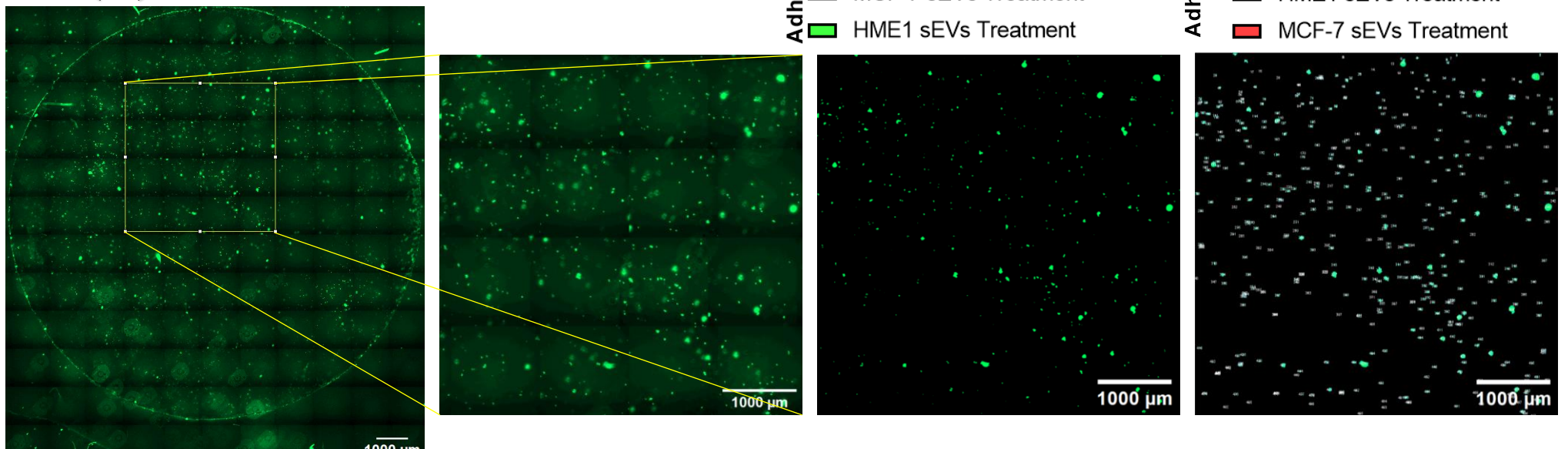
Supplementary Figure 8. Static adhesion assay of breast epithelial cells. Methods outlined section 2.7.3. (A) *hTERT-HME1 and MCF-7 breast epithelial cells were quantified for the number of cells bound to the HUVEC after 10 min of incubation.* **(B)** *hTERT-HME1 cells were treated for 1 h with metastatic MCF-7 derived sEVs and normal hTERT-HME1 derived sEVs and quantified for the number of cells bound to HUVEC alongside a no-treatment control containing only growth media. Raw counts of the adhered cells were normalised to those of the untreated control group.* **(C)** *MCF-7 cells* **(D)** *Image representation of the tiling images taken of coverslips containing HUVEC and bound breast epithelial cells (green); 16 images were taken to another window and subjected to Otsu auto thresholding, and bound cells were quantified by the analysis of particle function. (scale bar = 1000 µm). Each experiment consisted of biological triplicates with three technical triplicates, and error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. For parametric t-tests, Welch's correction was utilised, and for nonparametric t-tests, a Mann-Whitney test was used. The significance level was set at * < 0.05.*
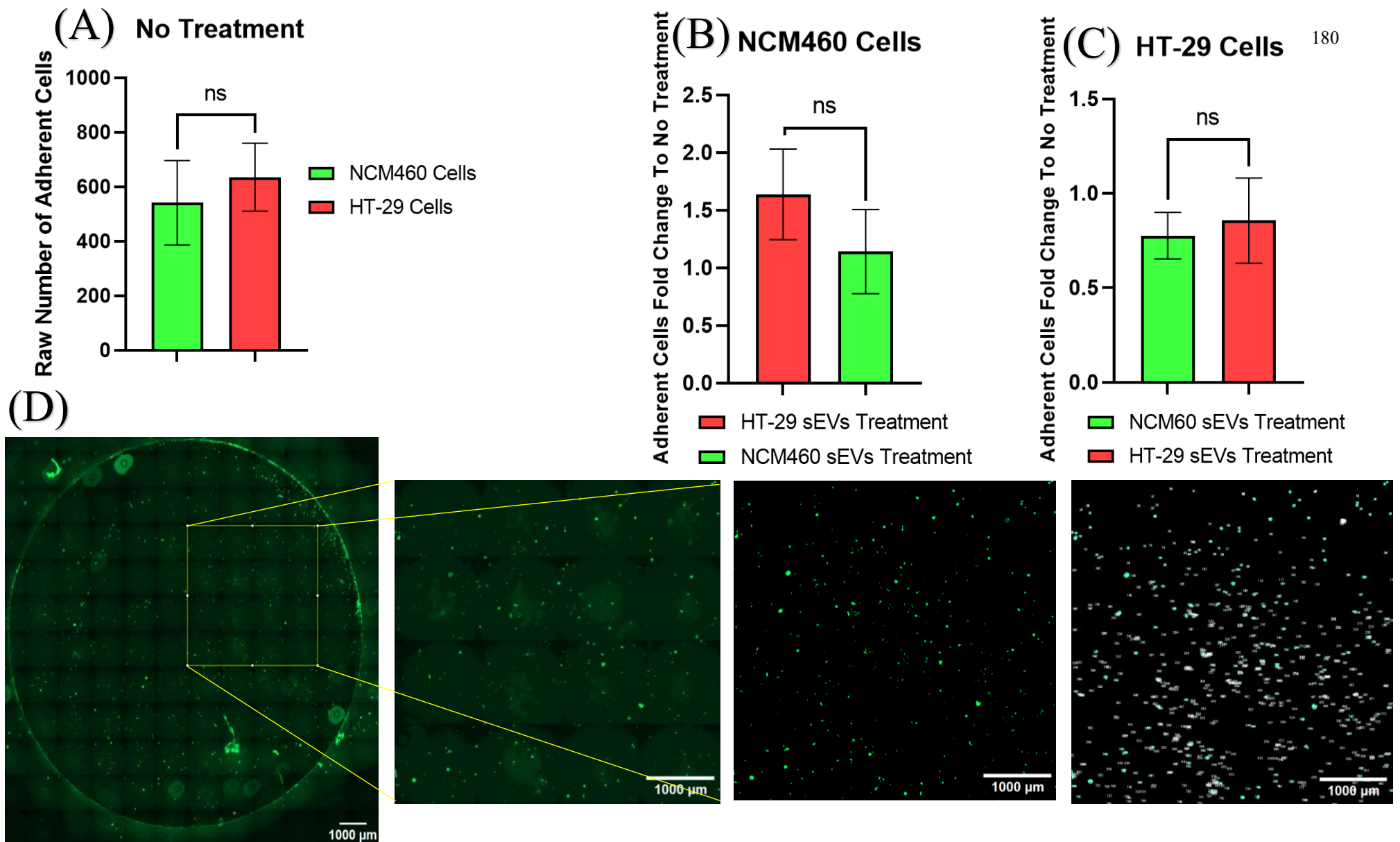
Supplementary Figure 9. Static adhesion assay of colorectal epithelial cells. Methods outlined in section 2.7.3. (A) *Colorectal epithelial cells NCM460 and HT-29 were quantified for the number of cells bound to HUVEC after 10 min of incubation.* **(B)** *NCM460 cells were treated for 1 h with metastatic HT-29 dervived sEVs and normal NCM460 derived sEVs and quantified for the number of cells bound to HUVEC alongside a no-treatment control containing only growth media. Raw counts of the adhered cells were normalised to the no treatment control group.* **(C)** *HT-29 cells* **(D)** *Image representation of the tiling images taken of coverslips containing HUVEC and bound colorectal epithelial cells (green); 16 images were taken to another window and subjected to Otsu auto thresholding, and bound cells were quantified by the analysis of particle function (scale bars = 1000 µm). Each experiment consisted of biological triplicates with three technical triplicates, and error bars indicate the Stdev. Normality of datasets was assessed using the Shapiro-Wilk test, and parametric or non-parametric tests were applied accordingly. For parametric t-tests, Welch's correction was utilised, and for nonparametric t-tests, a Mann-Whitney test was used.*
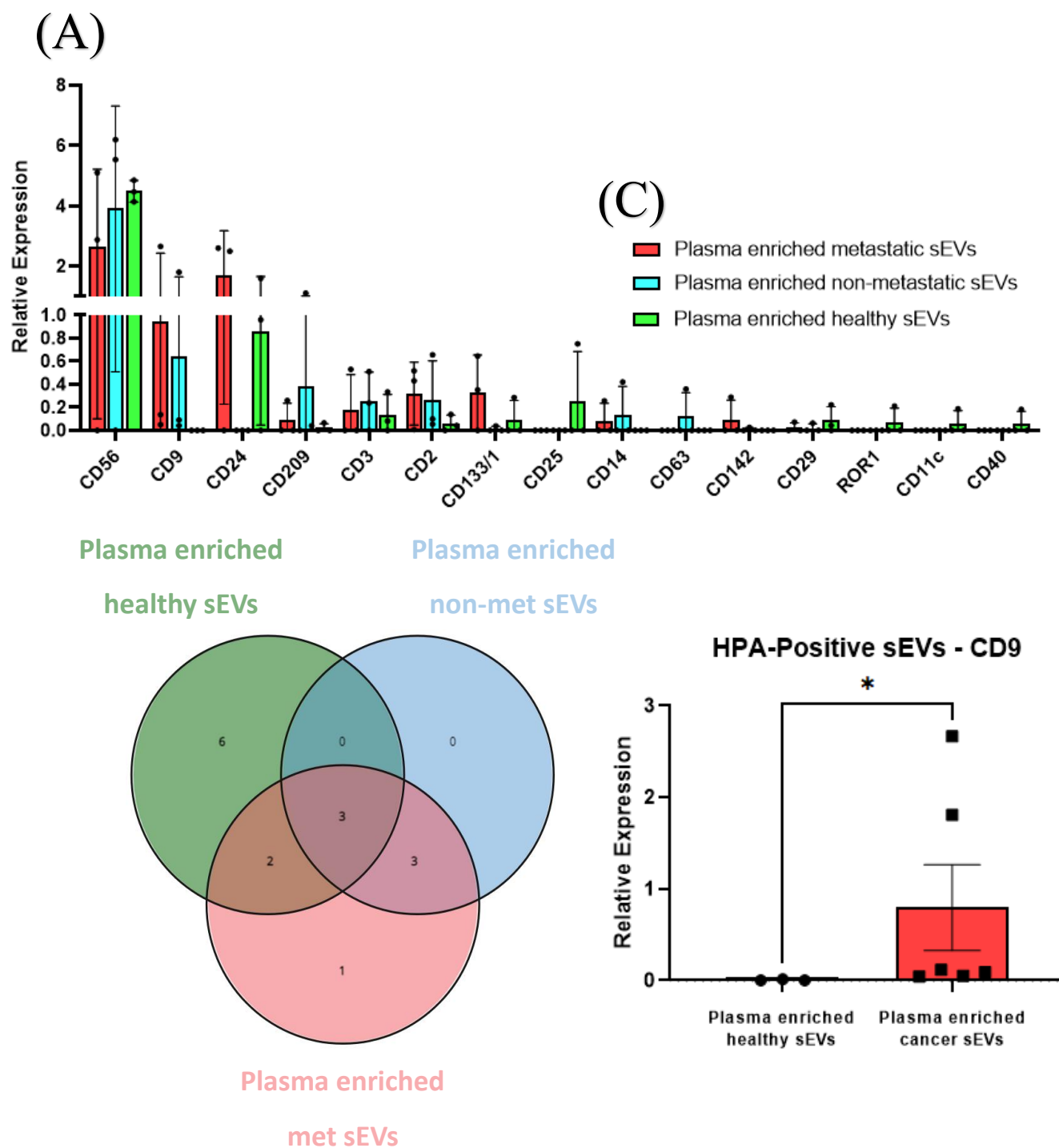
Supplementary Figure 10. MACSPlex analysis of plasma-enriched HPA-positive sEVs. Methods outlined in section 2.9.
(A) *Relative expression of each marker bound to plasma-enriched sEVs in each patient group, normalised to mean fluorescence intensity (red = plasma enriched sEVs derived from metastatic patients, blue = plasma enriched sEVs derived from non-metastatic patients, green = plasma enriched sEVs derived from healthy individuals).* **(B)** *Venn diagram showing the number of unique markers associated with HPA-positive sEVs in each patient cohort.* **(C)** *CD9 marker's relative expression of HPA-positive sEVs binding significantly more to cancer-associated plasma-enriched sEVs than to healthy individuals' plasma-enriched sEVs. Experiments were performed in triplicate, and error bars represent the Stdev. The t-test followed by Welch's correction was used to determine the significance levels: $*p < 0.05$.*

Marker	Plasma enriched met sEVs			Plasma enriched non-met sEVs			Plasma enriched healthy sEVs		
CD81	-261.00	-418.60	-323.00	-284.00	-310.00	-284.00	-77.00	-242.00	-299.00
CD63	-16.60	-12.50	-33.70	-4.10	-36.90	-20.50	78.30	-4.70	-35.40
CD9	19.00	67.30	8.70	9.30	14.80	14.60	-63.40	0.80	-1.40
CD41b	-34.00	-112.60	-11.00	-29.00	-62.00	-40.00	0	-18.00	-32.00
CD42a	-72.40	-29.20	-92.90	-88.30	-66.40	-60.30	-14.00	-64.30	-58.20
CD62P	-31.20	-62.40	-61.50	-25.40	-21.10	-43.80	-32.90	-49.40	-44.50
CD29	-8.60	-27.10	4.60	-9.70	10.40	-7.40	-54.30	14.20	10.90
CD45	-51.00	-107.70	-62.00	-84.00	-76.00	-108.60	-16.00	-40.00	-43.00
CD142	-22.90	-49.90	18.40	5.70	-17.70	-43.90	-88.60	-27.10	-24.70
CD24	923.60	-22.60	165.60	-10.20	0.60	-38.30	209.60	-31.50	393.60
CD14	-3.00	6.50	-47.70	-22.80	-17.10	3.40	-13.20	-7.20	-15.00
HLA-DRDPDQ	-70.00	-120.40	-124.00	-106.50	-129.60	-89.00	4.00	-62.00	-62.00
CD40	-33.00	-91.90	-22.00	-50.90	-32.40	-6.00	40.00	-22.00	-26.10
CD86	-30.10	-65.70	-59.90	-47.50	-49.60	-34.10	-64.60	-46.00	-43.10
CD1c	9.60	-17.70	-9.00	11.90	-28.10	-11.20	-1.90	-13.50	12.30
CD11c	-29.70	-45.20	-48.10	-58.00	-46.60	-47.60	41.00	-34.60	-76.00
CD209	-9.10	-37.80	16.50	-16.20	6.80	9.00	-56.90	-14.70	16.00
CD2	3.20	10.90	32.90	21.70	9.00	5.30	-16.90	8.90	10.60
CD3	196.00	-130.30	-87.30	52.00	84.00	-80.00	18.00	-85.00	82.00
CD4	-35.40	-48.70	-30.80	-44.80	-53.70	-28.80	-36.70	-34.10	-39.80
CD8	-39.40	-19.90	-25.90	-41.50	-48.70	-21.30	-46.80	-36.80	-31.30
CD56	1063.00	-103.00	324.00	1211.00	1018.00	-51.00	976.00	265.00	1194.00
CD69	-53.30	-2.90	-63.71	-21.60	-9.90	-15.60	-110.40	-9.30	-17.70
CD25	-123.00	-342.81	-161.00	-188.00	-151.00	-249.20	164.00	-149.00	-138.00
CD19	-56.50	-63.80	-114.00	-53.20	-42.00	-87.00	-19.70	-53.20	-72.60
CD20	-173.00	-272.70	-162.00	-207.20	-198.80	-211.00	-74.00	-151.00	-187.00
CD31	-112.90	-165.20	-87.00	-108.20	-113.90	-99.00	-33.00	-90.00	-91.00
CD146	-58.00	-78.40	-53.80	-49.40	-53.60	-63.70	-93.00	-52.80	-44.60
CD105	-96.40	-118.00	-155.20	-112.90	-113.20	-125.70	-84.00	-89.30	-86.90
CD326	-295.00	-430.40	-336.00	-317.00	-379.00	-317.00	-34.00	-245.00	-313.00
SSEA-4	-1120.00	-1813.00	-1349.00	-1230.00	-1318.00	-1152.00	-153.00	-913.00	-1507.00
CD133/1	-9.10	16.30	22.30	-19.80	6.30	-15.90	-93.70	18.30	-14.50
HLA-ABC	-590.00	-936.90	-825.00	-675.00	-712.00	-634.00	-83.00	-554.00	-627.00
CD44	-58.00	-155.00	-110.00	-96.00	-93.00	-99.00	-38.00	-76.00	-94.00
CD49e	-22.30	-60.10	-54.90	-46.90	-43.10	-20.50	-8.90	-15.40	-4.90
ROR1	-24.20	-2.60	-3.70	-21.80	-17.60	-10.90	-108.80	13.60	-2.50
MCSP	-81.60	-73.20	-101.10	-63.10	-66.30	-78.60	-19.00	-87.70	-54.30
REA ctrl	-204.00	-285.50	-240.00	-212.00	-233.00	-234.00	129.00	-198.00	-214.00
mIgG1 ctrl	-239.00	-316.00	-271.00	-283.00	-268.00	-257.00	114.00	-223.00	-252.00

Marker group annotations (left axis): Exosome Markers — CD81, CD63, CD9; Platelets — CD41b, CD42a, CD62P, CD29; Neutrophils/Platelets — CD45, CD142, CD24; Monocytes / Macrophages/DC — CD14, HLA-DRDPDQ, CD40, CD86, CD1c, CD11c, CD209; T Cells — CD2, CD3, CD4, CD8; NK Cells — CD56; T Cells — CD69, CD25; B Cells — CD19, CD20; Platelets / Endothelium — CD31, CD146, CD105; Epithelium — CD326; Stem Cells — SSEA-4, CD133/1; Multiple Cells — HLA-ABC, CD44, CD49e; Adipocytes/Parathyroid/Cancer — ROR1; Melanocytes/Smooth muscle cells/Cancer — MCSP; Controls — REA ctrl, mIgG1 ctrl. (Leukocytes bracket spans Neutrophils/Platelets through B Cells.)

Colour scale: 0 – 500 – 1000.

Supplementary Figure 11. Heatmap of the MACSPlex analysis of plasma-enriched HPA-positive sEVs. Methods outlined in section 2.9. *Heatmap of MACSPlex analysis of plasma-enriched HPA-positive sEVs. Heatmap displaying fluorescence intensities of plasma-enriched sEVs after background subtraction of controls. Each rectangle is colour-coded to indicate lectin association with each group (white, no binding; purple, low binding; blue, medium binding; green/yellow, high binding). The proposed marker groups are listed: Experiments were performed in biological triplicates.*
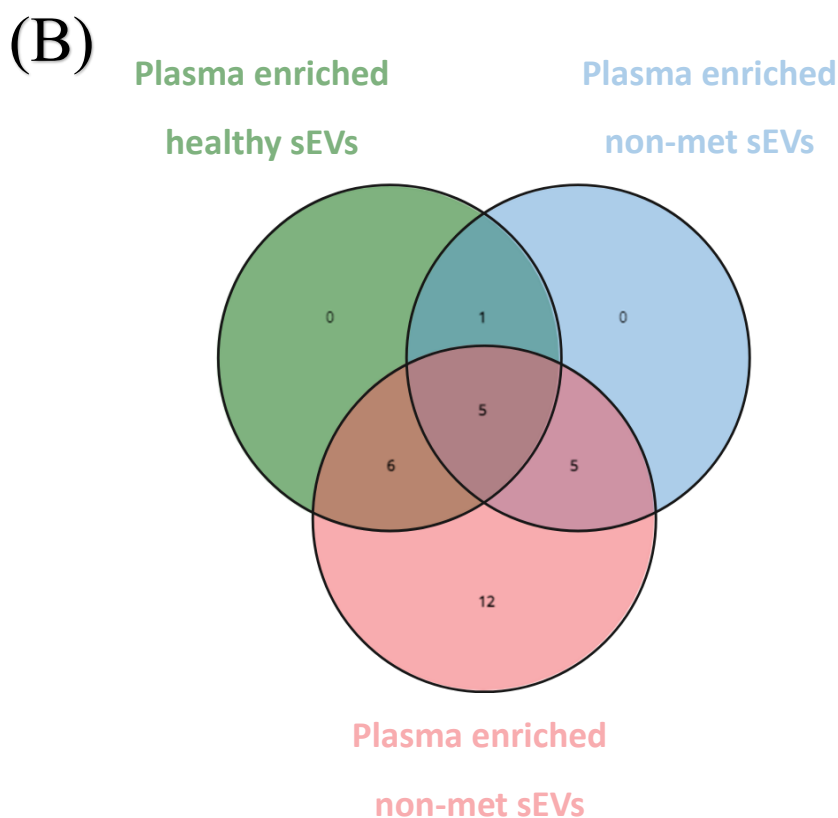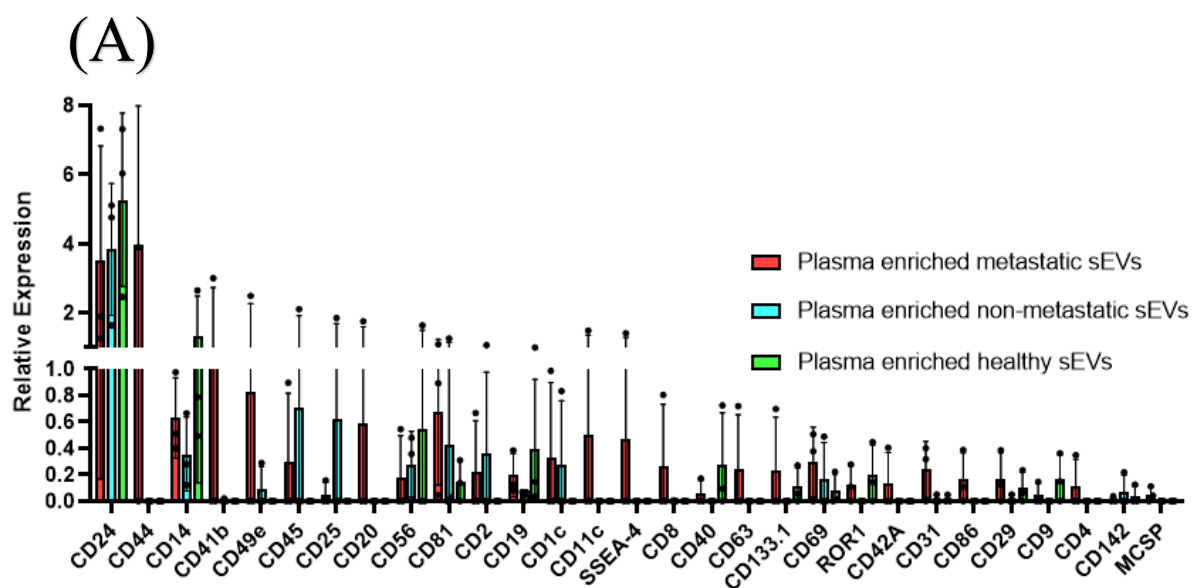
Supplementary Figure 12. MACSPlex analysis of plasma-enriched LCA-positive sEVs. Methods outlined in section 2.9.
(A) *Relative expression of each marker bound to plasma-enriched sEVs in each patient group, normalised to mean fluorescence intensity (red = plasma enriched sEVs derived from metastatic patients, blue = plasma enriched sEVs derived from non-metastatic patients, green = plasma enriched sEVs derived from healthy individuals).* **(B)** *Venn diagram showing the number of unique markers associated with LCA-positive sEVs in each patient cohort. Experiments were performed in triplicate, and error bars represent the Stdev. The t-test followed by Welch's correction was used to determine the significance levels.*

Marker group	Marker	Plasma enriched met sEVs			Plasma enriched non-met sEVs			Plasma enriched healthy sEVs		
Exosome Markers	CD81	5.2	33.4	255.4	-0.5	8.2	74.4	16.4	-3.4	16.4
	CD63	-49.0	-82.0	167.0	-67.6	-70.0	-19.0	-43.0	-14.0	-65.8
	CD9	-5.0	-30.8	34.0	-12.2	-1.0	-20.7	8.0	8.0	-10.9
Platelets	CD41b	-44.0	-39.0	698.0	-43.0	6.0	-33.0	-24.0	-46.0	-16.0
	CD42a	-49.7	-29.4	94.0	-31.7	-29.7	-53.2	-39.8	-29.7	-8.0
	CD62P	-1120.0	-1672.0	-784.0	-1225.0	-1040.0	-924.0	-789.0	-1125.0	-1019.0
	CD29	14.0	-54.6	89.6	9.4	-11.3	-40.0	12.4	1.4	-14.3
Neutrophils/Platelets	CD45	-11.0	-183.4	208.0	-114.0	-81.0	124.0	-98.0	-117.0	-101.0
	CD142	-3.8	-7.7	8.9	-28.5	-15.9	12.9	-7.9	-17.9	14.9
	CD24	834.6	47.6	440.6	924.6	1368.6	96.6	386.6	54.6	717.6
Monocytes	CD14	45.7	36.5	118.7	50.7	35.0	39.1	41.7	58.7	58.7
Macrophages/DC	HLA-DRDPDQ	-4273.0	-5342.0	-7012.0	-4615.0	-3450.0	-4107.0	-2488.0	-3785.0	-3450.0
	CD40	-2.5	-17.9	40.0	-29.0	-2.5	-7.2	5.0	16.0	-1.7
	CD86	-31.7	14.5	25.5	-41.2	-31.0	-31.2	-8.7	-39.5	-17.9
	CD1c	-47.6	-78.0	229.0	-81.9	-57.0	49.0	-23.0	-37.5	-37.7
	CD11c	-139.0	-221.8	347.0	-127.0	-114.0	-13.0	-41.0	-101.0	-105.0
	CD209	1.6	-17.6	-7.3	-20.5	-28.6	-52.8	-18.3	-18.5	-44.3
T Cells	CD2	-61.0	-123.8	155.0	-103.7	-94.0	63.0	-72.0	-64.0	-91.0
	CD3	130.0	-90.0	208.0	-61.0	-7.0	126.0	-60.0	-86.0	-76.0
	CD4	-6.5	-10.4	81.0	-29.2	-23.6	-3.3	-8.4	-13.7	-12.2
	CD8	-38.0	-7.8	186.6	1.3	-25.7	-10.3	-9.4	-10.3	-25.7
NK Cells	CD56	62.0	-131.0	-44.0	87.0	-37.0	21.0	87.0	-169.0	-14.0
T Cells	CD69	-8.0	19.0	87.6	-43.2	-13.3	28.7	11.8	-6.9	1.5
	CD25	18.0	-54.0	-117.7	-57.0	-52.0	109.0	-65.0	-61.0	-24.0
B Cells	CD19	14.4	14.4	18.6	12.7	19.2	-0.5	2.0	22.1	17.5
	CD20	-10.0	-41.5	409.0	-33.2	-40.8	-38.3	-17.0	-14.0	-61.2
Platelets	CD31	-1.0	15.0	74.0	-2.0	-10.7	3.0	-17.7	-14.4	6.0
Endothelium	CD146	-1997.0	-1704.0	-88.0	-2292.0	-1933.0	-1782.0	-1407.0	-2060.0	-1887.0
	CD105	-1165.0	0	-2425.0	-1223.0	-1677.0	-689.0	-659.0	-1223.0	-930.0
Epithelium	CD326	-257.0	-236.0	1.0	-270.0	-265.0	-255.0	-240.0	-237.0	-249.0
Stem Cells	SSEA-4	-69.0	-133.0	329.0	-115.0	-76.0	-72.0	-50.0	-89.0	-34.0
	CD133/1	-6.0	-53.9	162.0	-2.0	-39.0	-18.2	3.0	6.0	-11.0
Multiple Cells	HLA-ABC	-93.0	-185.0	17.0	-127.0	-95.0	-118.0	-95.0	-97.0	-96.0
	CD44	-69.0	146.0	1863.0	-98.0	-61.0	-39.0	-57.0	-36.0	-65.0
	CD49e	-41.0	-91.6	579.0	-53.8	-40.0	17.0	-27.0	-16.0	-61.2
Adipocytes/Parathyroid/Cancer	ROR1	-0.2	10.4	25.2	-14.7	-0.8	-8.8	7.6	9.9	-10.6
Melanocytes/Smooth muscle cells/Cancer	MCSP	12.9	-17.4	9.8	-17.3	-16.1	-1.8	-19.9	-10.2	-13.5
Controls	REA ctrl	-663.0	-1314.0	-162.0	-676.0	-622.0	-510.0	-339.0	-740.0	-525.0
	mIgG1 ctrl	-33.1	-12.0	88.0	-7.0	-10.0	-14.9	-6.0	-13.1	-20.0

Supplementary Figure 13. Heatmap of the MACSPlex analysis of plasma-enriched LCA-positive sEVs. Methods outlined in section 2.9. *Heatmap of MACSPlex analysis of plasma-enriched LCA-positive sEVs. Heatmap displaying fluorescence intensities of plasma-enriched sEVs after background subtraction of controls. Each rectangle is colour-coded to indicate lectin association with each group (white, no binding; purple, low binding; blue, medium binding; green/yellow, high binding). The proposed marker groups are listed: Experiments were performed in biological triplicates.*
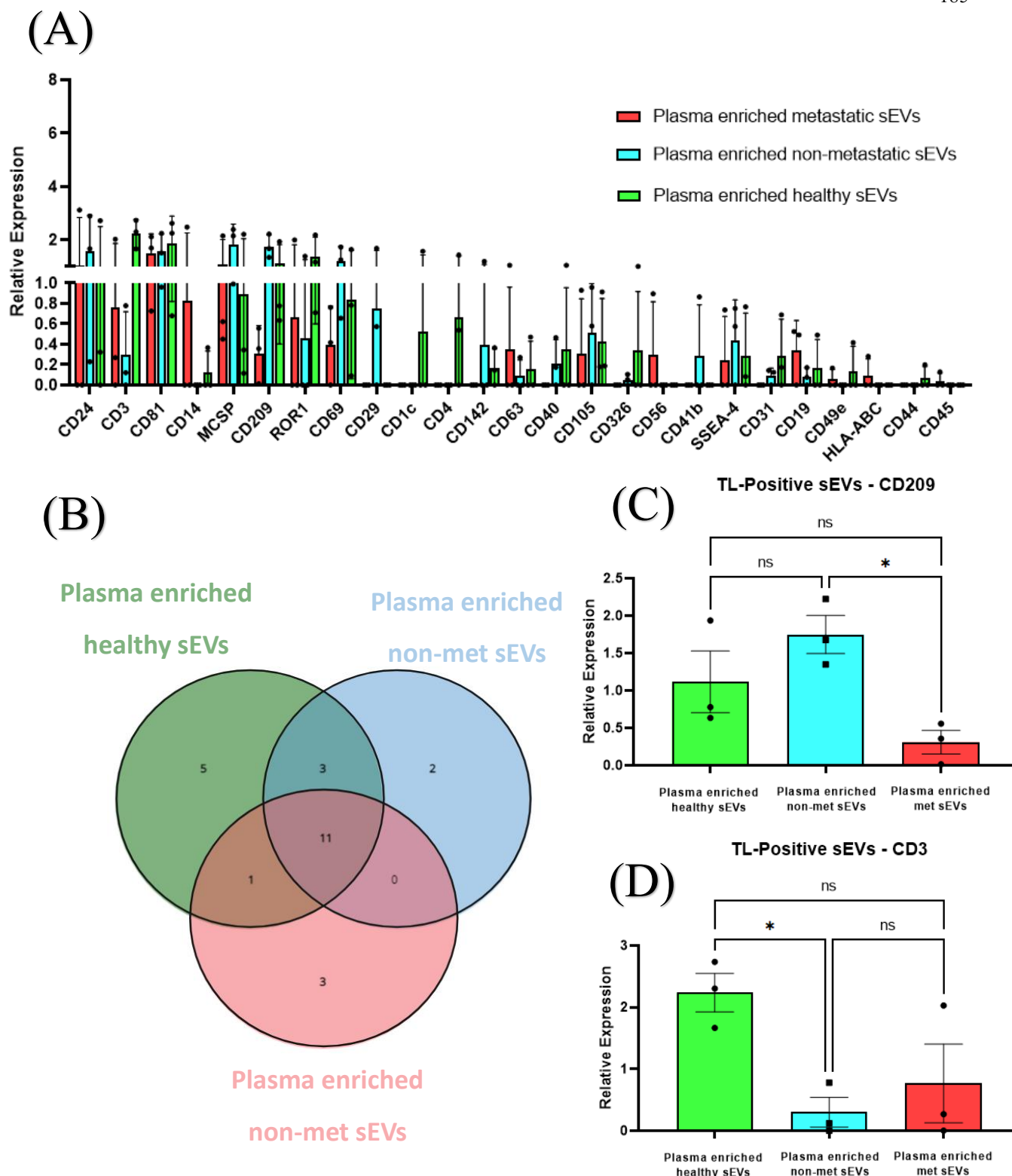
Supplementary Figure 14. MACSPlex analysis of plasma-enriched TL-positive sEVs. Methods outlined in section 2.9.
(A) *Relative expression of each marker bound to plasma-enriched sEVs in each patient group, normalised to mean fluorescence intensity (red = plasma enriched sEVs derived from metastatic patients, blue = plasma enriched sEVs derived from non-metastatic patients, green = plasma enriched sEVs derived from healthy individuals).* **(B)** *Venn diagram showing the number of unique markers associated with TL-positive sEVs in each patient cohort.* **(C)** *CD209 marker's relative expression of TL-positive sEVs binding significantly more to non-metastatic cancer-derived plasma-enriched sEVs than to metastatic cancer-derived plasma enriched sEVs.* **(D)** *CD3 marker's relative expression of TL-positive sEVs binding significantly more to healthy individuals derived plasma-enriched sEVs in comparison to non-metastatic cancer-derived plasma-enriched sEVs. Experiments were performed in triplicate, and error bars represent the Stdev. The t-test followed by Welch's correction was used to determine the significance levels: *p < 0.05.*

Marker	Plasma enriched met sEVs			Plasma enriched non-met sEVs			Plasma enriched healthy sEVs		
CD81	22.71	38.61	65.41	41.61	19.01	24.41	37.61	13.01	74.71
CD63	-2.10	24.00	-27.50	5.00	-18.70	-18.00	-3.10	-14.20	13.40
CD9	-42.51	-25.30	-4.80	-32.90	-29.80	-26.70	-16.70	-23.40	-32.20
CD41b	-6.00	-33.30	-12.60	16.00	-14.80	-6.60	-6.20	-15.00	-23.40
CD42a	-13.60	3.60	-45.01	-11.60	-16.40	-2.20	-10.90	-17.20	2.50
CD62P	-5.20	-23.30	-27.70	-10.20	-40.61	-27.80	-0.30	-17.90	-33.70
CD29	-22.42	-18.20	-15.80	10.60	33.30	-14.10	-8.00	-24.11	-1.80
CD45	4.00	-104.50	-2.00	-8.00	-53.70	-15.00	-30.00	-44.90	-5.00
CD142	-3.60	-1.00	-49.41	-31.00	-25.50	19.30	6.10	-17.40	4.10
CD24	97.50	-15.00	-53.20	53.50	33.10	3.70	5.40	-8.40	77.50
CD14	-41.90	-41.80	76.60	-5.70	-14.70	-36.80	6.10	-21.60	0
HLA--DRDPDQ	-74.00	-211.70	-179.30	-88.00	-126.00	-69.00	-24.00	-57.00	-113.00
CD40	-12.10	-11.30	-36.00	8.60	-14.10	2.80	17.50	-19.70	-8.70
CD86	-55.50	-64.00	-76.10	-30.30	-15.30	-33.00	-61.50	-26.80	-47.80
CD1c	-25.90	-36.60	-4.80	-7.10	-18.20	-17.80	-21.30	30.20	-2.80
CD11c	-193.00	-221.30	-94.00	-126.00	-149.00	-129.00	-59.00	-71.00	-128.00
CD209	11.19	0.30	17.20	31.00	26.70	35.90	10.59	37.10	22.10
CD2	-39.70	-28.50	-0.20	-49.00	-44.70	-22.10	-39.70	-56.30	-43.30
CD3	63.40	-45.90	8.30	14.40	2.40	-10.60	38.40	52.40	47.40
CD4	-31.30	-14.50	-5.60	-9.30	-15.10	-21.10	9.00	-25.90	40.90
CD8	-36.70	-39.10	-16.80	-39.70	-33.90	-17.70	-50.40	-21.00	-31.50
CD56	28.00	-55.70	-5.20	-21.20	-25.80	-22.50	-13.80	-6.50	-21.20
CD69	23.80	9.50	-29.41	23.10	13.00	28.00	27.40	15.00	2.60
CD25	-23.10	-49.50	-16.80	-43.10	-42.20	-51.00	-48.00	-62.00	-52.80
CD19	16.40	11.20	-18.30	1.40	3.40	0	-7.30	9.40	-3.00
CD20	11.19	-11.50	35.30	31.00	26.70	35.90	10.59	37.10	22.10
CD31	-12.80	-15.60	-19.70	2.30	-11.50	2.30	2.90	13.20	-19.50
CD146	-45.30	-65.60	-72.70	-41.80	-48.90	-49.60	-55.30	-28.10	-60.30
CD105	-11.80	-30.20	28.70	10.70	-19.60	15.50	3.00	3.60	26.00
CD326	-5.40	-9.10	-13.30	1.00	-9.40	1.70	16.80	-22.60	-17.70
SSEA-4	-15.20	-22.50	22.80	13.90	-5.30	9.30	12.80	-13.50	2.40
CD133/1	-30.50	-26.10	-38.20	-42.00	-24.20	-35.50	-3.10	-37.90	-21.60
HLA-ABC	-31.60	-62.00	8.70	-35.30	-32.00	-13.60	-1.70	-23.40	-16.20
CD44	-2.00	-21.10	-13.30	-8.60	-8.30	-23.40	1.30	3.80	-24.80
CD49e	-18.00	3.80	-15.50	-6.60	-15.00	-14.30	-29.90	8.00	-16.00
ROR1	-6.20	45.60	-7.00	-5.00	27.50	-22.80	19.50	41.90	20.20
MCSP	14.11	49.01	19.21	44.21	19.61	34.81	36.91	2.21	9.81
REA ctrl	-63.10	-98.60	-44.90	-73.20	-72.70	-19.00	-36.90	-36.10	-63.90
mIgG1 ctrl	-12.10	-23.00	18.10	-12.90	-5.10	-1.30	-13.50	-21.30	-5.00

Marker groups (left annotations): Exosome Markers; Platelets; Neutrophils/Platelets; Monocytes; Macrophages/DC; Leukocytes; T Cells; NK Cells; B Cells T Cells; Platelets; Endothelium; Epithelium; Stem Cells; Multiple Cells; Adipocytes/Parathyroid/Cancer; Melanocytes/Smooth muscle cells/Cancer; Controls.
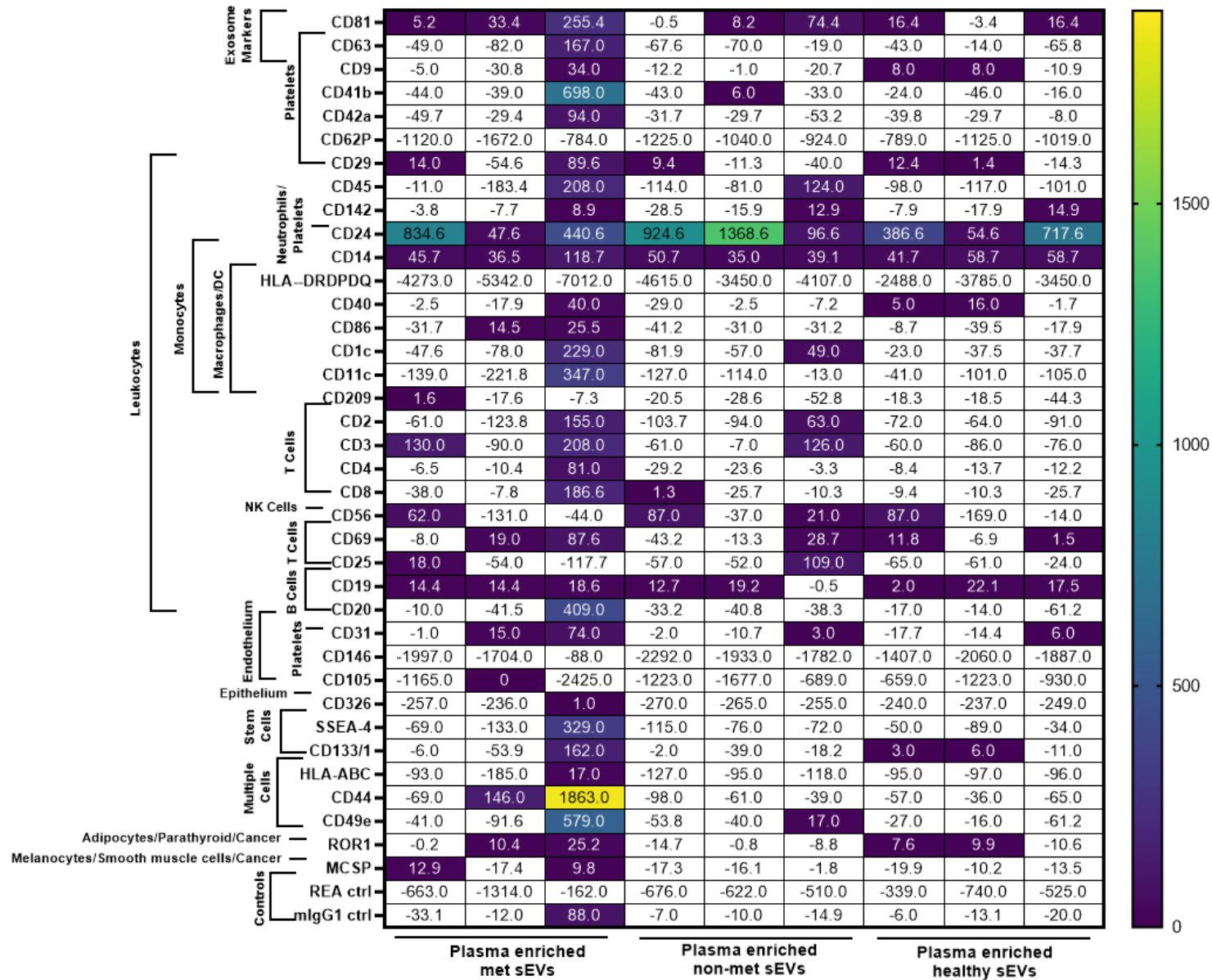
Supplementary Figure 15. Heatmap of the MACSPlex analysis of plasma-enriched TL-positive sEVs. Methods outlined in section 2.9. *Heatmap displaying fluorescent intensities of plasma-enriched sEVs after background subtraction of controls. Each rectangle is color-coded to indicate lectin association to each group (white = no binding, purple = low binding, blue = medium binding, green/yellow = high binding). Proposed marker groups are listed. Experiments were performed with biological triplicates.*
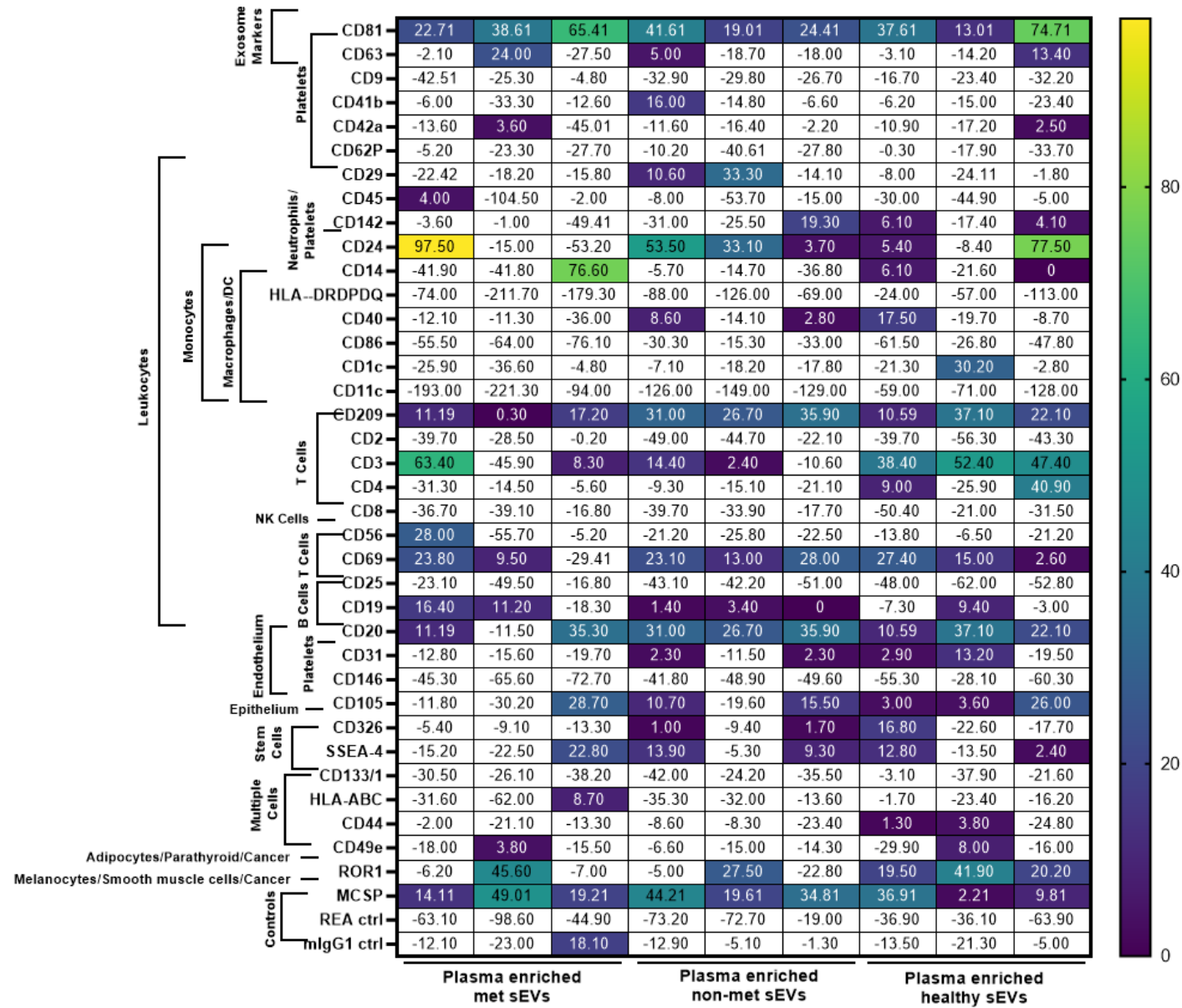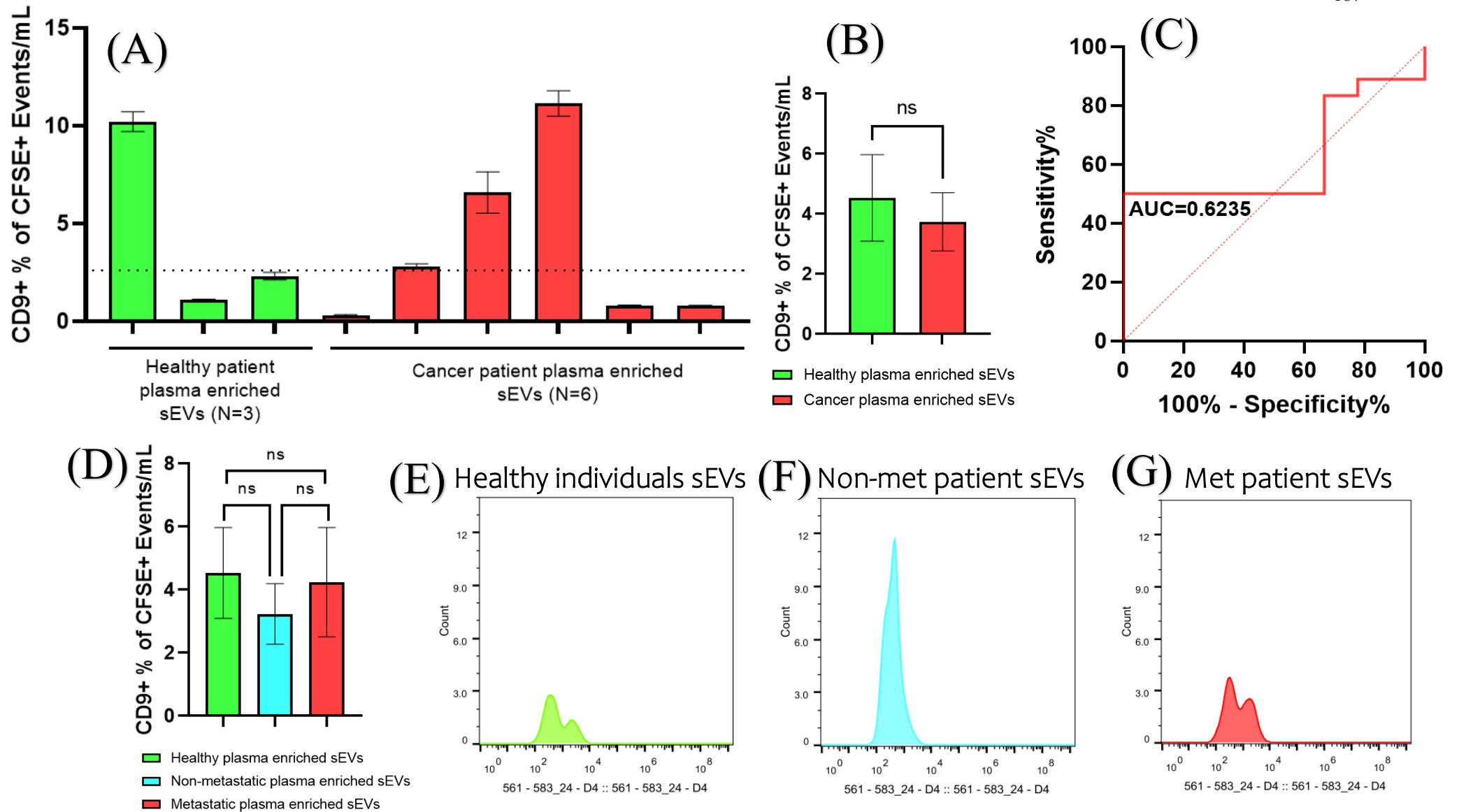
Supplementary Figure 16. Single-vesicle analysis of CD9 detection of plasma-enriched sEVs from breast metastatic patients (n = 3), non-metastatic patients (n = 3), and healthy individuals (n = 3). Methods outlined in section 2.6.3. (A) *CD9+ events/ml normalised to % CFSE events/ml for each healthy and cancer-associated plasma enriched sEVs sample.* **(B)** *Comparison of CD9+ events/ml normalised to % CFSE events/ml between plasma-enriched sEVs derived from healthy individuals and cancer patients.* **(C)** *ROC curve analysis measuring CD9+ events/ml normalised to % CFSE events/ml, yielding an AUC of 0.6235 for discriminating patients with cancer from healthy individuals.* **(D)** *Comparison of CD9+ events/ml normalised to % CFSE events/ml between plasma-enriched sEVs derived from healthy individuals and non-metastatic and metastatic patients.* **(E-G)** *Histogram profiles of plasma-enriched sEVs after gating showing CD9 fluorescence. Each experiment comprised biological triplicates with three technical triplicates; error bars represent the Stdev. The t-test followed by Welch's correction was used to determine the significance levels.*
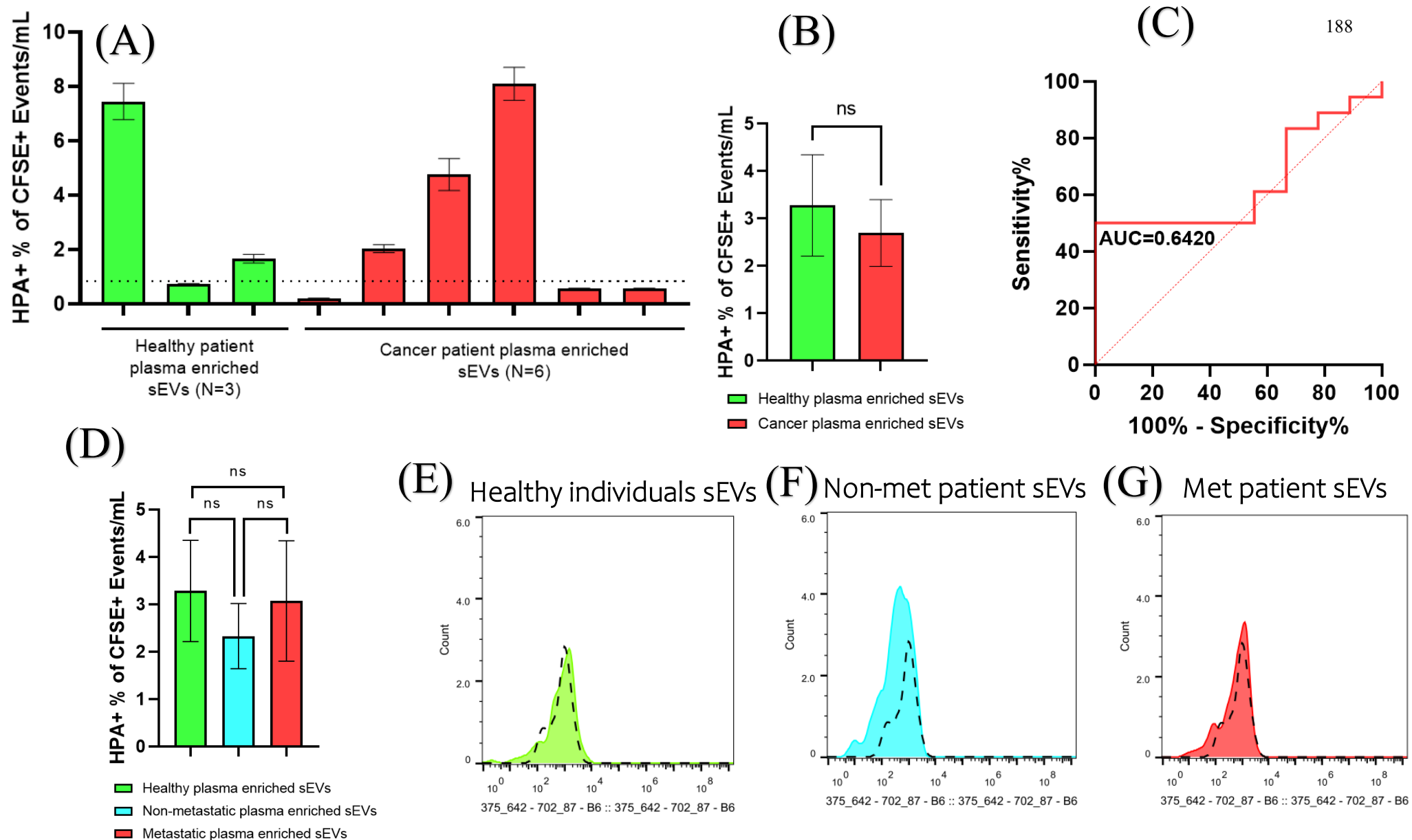
Supplementary Figure 17. Single-vesicle analysis of HPA lectin detection of CD9-positive plasma-enriched sEVs from breast metastatic patients (n = 3), non-metastatic patients (n = 3), and healthy individuals (n = 3). Methods outlined in section 2.6.3. *(A) HPA+ events/ml normalised to % CFSE events/ml for each healthy and cancer-associated plasma enriched sEVs sample.* **(B)** *Comparison of HPA+ events/ml normalised to % CFSE events/ml between plasma-enriched sEVs derived from healthy individuals and cancer patients.* **(C)** *ROC curve analysis measuring HPA+ events/ml normalised to % CFSE events/ml, yielding an AUC of 0.6420 for discriminating patients with cancer from healthy individuals.* **(D)** *Comparison of HPA+ events/ml normalised to % CFSE events/ml between plasma-enriched sEVs derived from healthy individuals and non-metastatic and metastatic patients.* **(E-G)** *Histogram profiles of plasma-enriched sEVs after gating showing HPA fluorescence. Each experiment comprised biological triplicates with three technical triplicates; error bars represent the Stdev. The t-test followed by Welch's correction was used to determine the significance levels.*
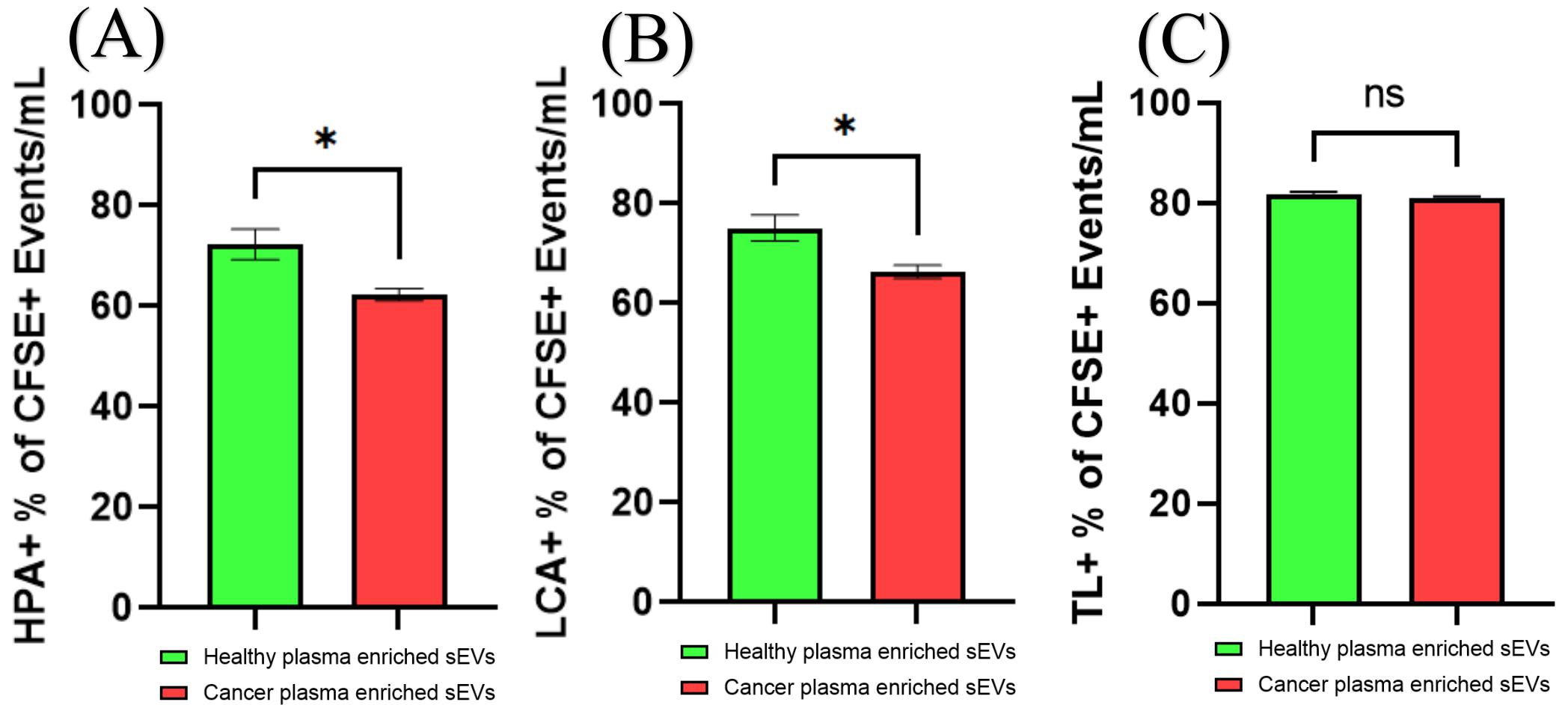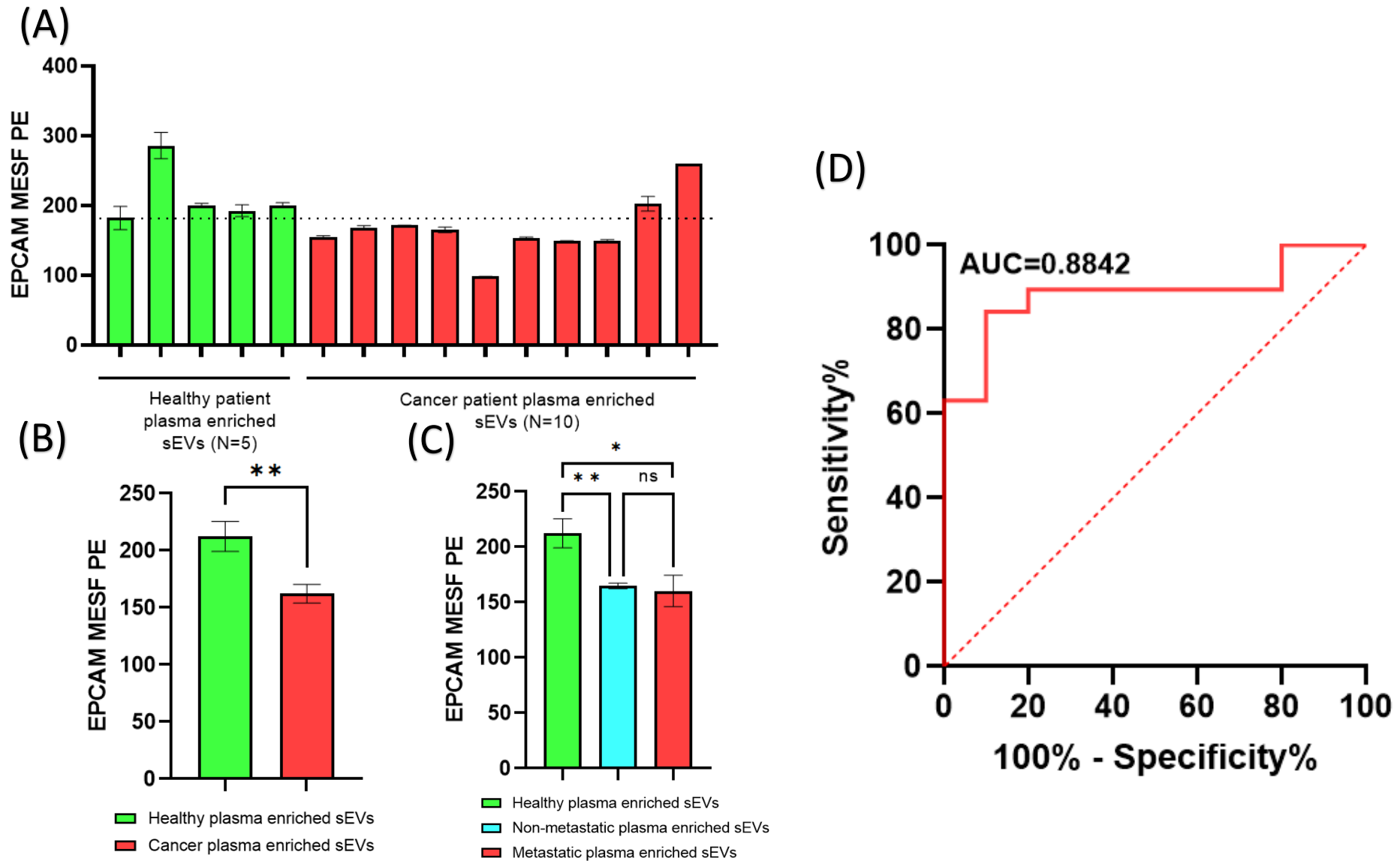
Supplementary Figure 18. Single-vesicle analysis of HPA, LCA and TL of plasma-enriched sEVs derived from breast metastatic patients (n = 3), non-metastatic patients (n = 3), and healthy individuals (n = 3). Methods outlined in section 2.6.3. **(A)** *HPA+ events/ml normalised to % CFSE events/ml for each healthy and cancer-associated plasma enriched sEVs sample.* **(B)** *LCA+ events/ml normalised to % CFSE events/ml for each healthy and cancer-associated plasma enriched sEVs sample.* **(C) TL**+ *events/ml normalised to % CFSE events/ml for each healthy and cancer-associated plasma enriched sEVs sample. Each experiment comprised biological triplicates with three technical triplicates; error bars represent Stdev. T-test followed by Welch's correction determined significance levels: *p < 0.05.*

Supplementary Figure 19. Single-vesicle analysis of EpCAM detection of plasma-enriched sEVs derived from breast metastatic patients (n = 3), non-metastatic patients (n = 3), and healthy individuals (n = 3). **Methods outlined in section 2.6.3. (A)** *EpCAM+ MESF PE for each healthy and cancer-associated plasma enriched sEVs sample.* **(B)** *Comparison of EpCAM+ MESF PE between plasma enriched sEVs derived from healthy individuals and cancer-associated patients.* **(C)** *Comparison of EpCAM+ MESF PE between plasma enriched sEVs derived from healthy individuals, non-metastatic and metastatic patients.* **(D)** *ROC curve analysis measuring EpCAM+ MESF PE, yielding an AUC of 0.8842 for discriminating cancer patients from healthy individuals. Each experiment comprised biological triplicates with three technical triplicates; error bars represent Stdev. T-test followed by Welch's correction determined significance levels: *p < 0.05, **p < 0.01.*
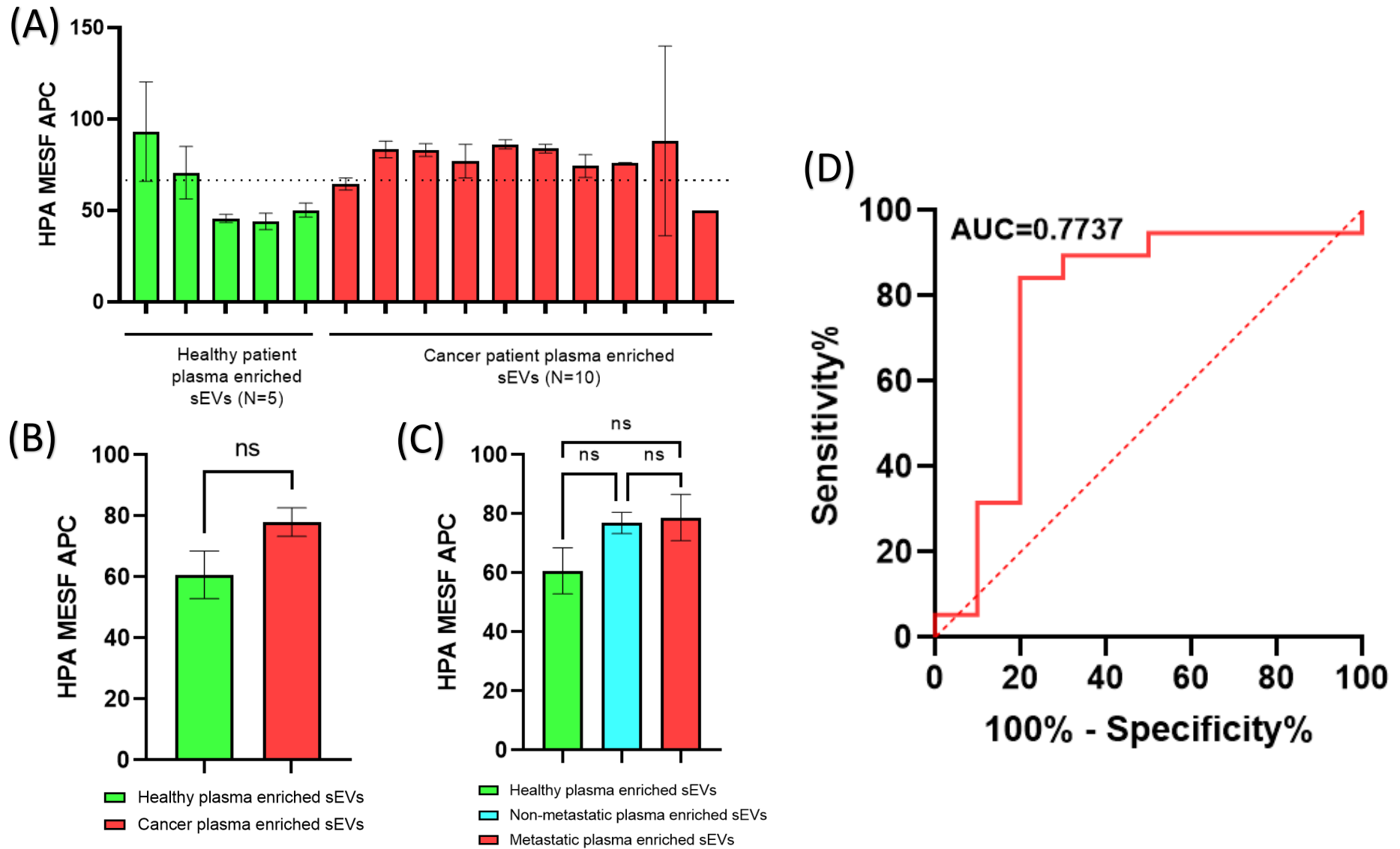
Supplementary Figure 20. Single-vesicle analysis of HPA lectin binding of plasma-enriched EpCAM-postive sEVs derived from breast metastatic patients (n = 3), non-metastatic patients (n = 3), and healthy individuals (n = 3). Methods outlined in section 2.6.3. **(A)** *HPA+ MESF APC for each healthy and cancer-associated plasma enriched sEVs sample.* **(B)** *Comparison of HPA+ MESF APC between plasma enriched sEVs derived from healthy individuals and cancer-associated patients.* **(C)** *Comparison of HPA+ APC between plasma enriched sEVs derived from healthy individuals, non-metastatic and metastatic patients.* **(D)** *ROC curve analysis measuring HPA+ MESF APC, yielding an AUC of 0.7737 for discriminating cancer patients from healthy individuals. Each experiment comprised biological triplicates with three technical triplicates; error bars represent Stdev. T-test followed by Welch's correction determined significance levels.*
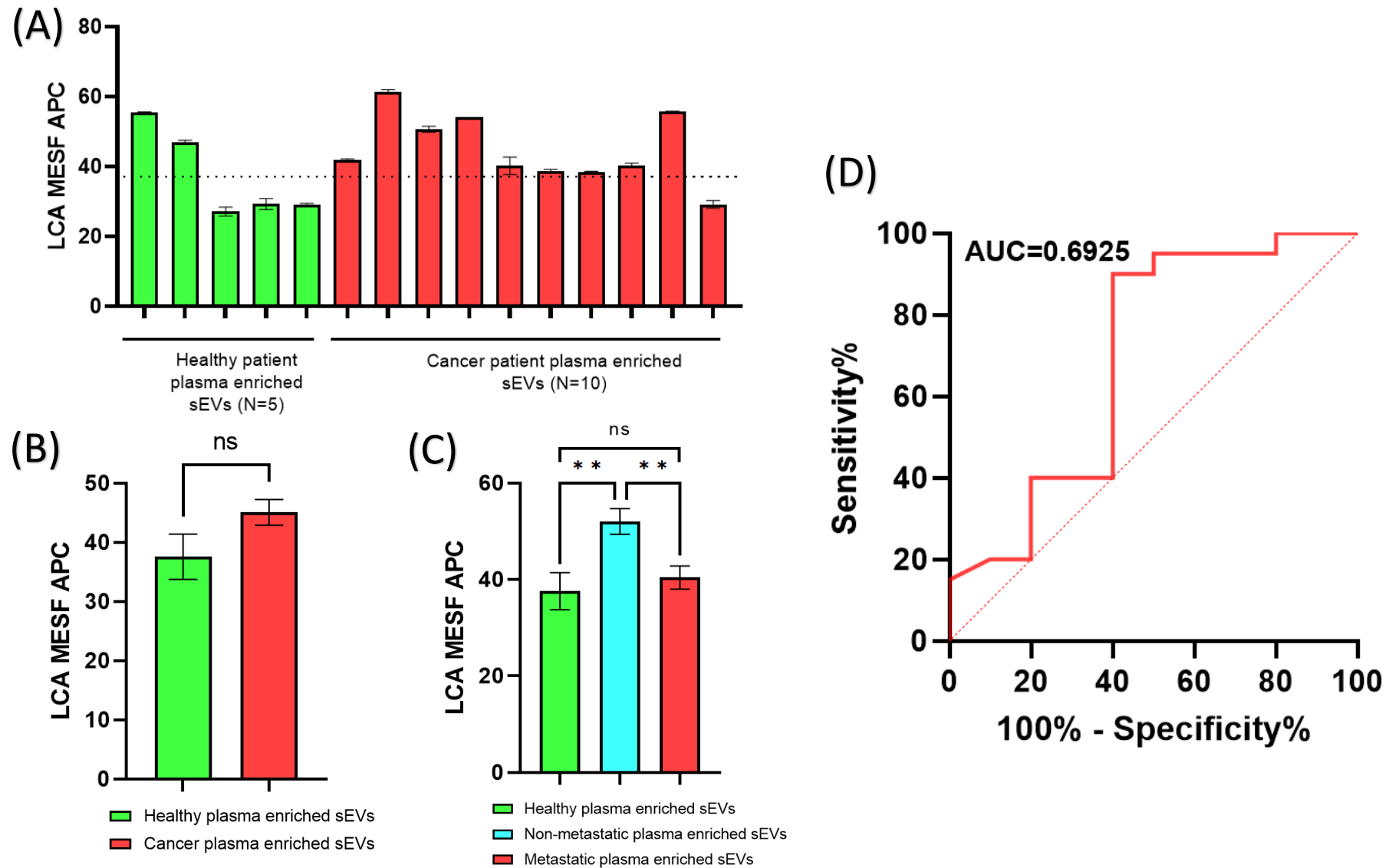
Supplementary Figure 21. **Single-vesicle analysis of LCA lectin binding of plasma-enriched EpCAM-postive sEVs derived from breast metastatic patients (n = 3), non-metastatic patients (n = 3), and healthy individuals (n = 3). Methods outlined in section 2.6.3. (A)** *LCA+ MESF APC for each healthy and cancer-associated plasma enriched sEVs sample.* **(B)** *Comparison of LCA+ MESF APC between plasma enriched sEVs derived from healthy individuals and cancer-associated patients.* **(C)** *Comparison of LCA+ APC between plasma enriched sEVs derived from healthy individuals, non-metastatic and metastatic patients.* **(D)** *ROC curve analysis measuring LCA+ MESF APC, yielding an AUC of 0.6925 for discriminating cancer patients from healthy individuals. Each experiment comprised biological triplicates with three technical triplicates; error bars represent Stdev. T-test followed by Welch's correction determined significance levels: *p < 0.05, **p < 0.01.*
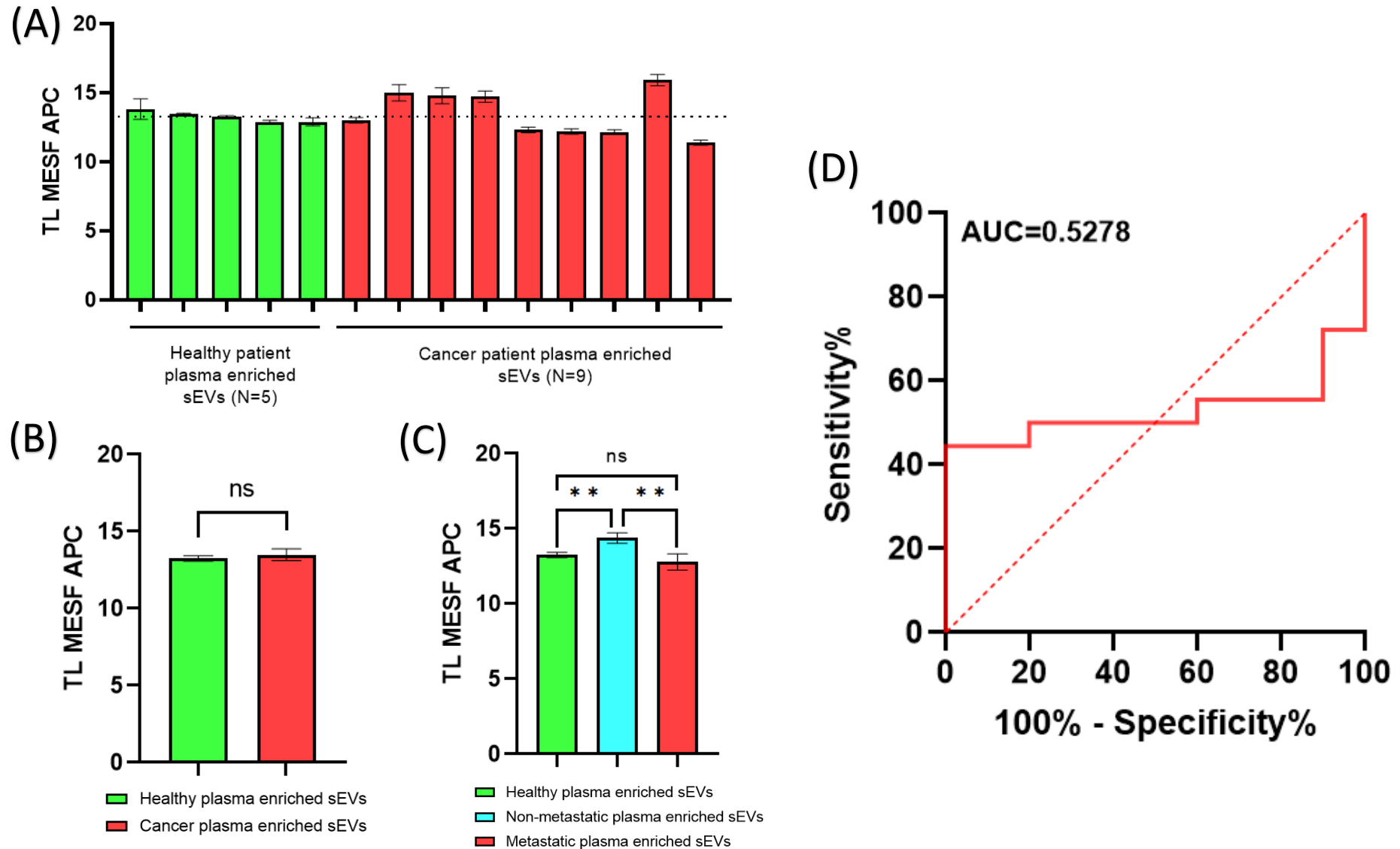
Supplementary Figure 22. Single-vesicle analysis of TL binding of plasma-enriched EpCAM-postive sEVs derived from breast metastatic patients (n = 3), non-metastatic patients (n = 3), and healthy individuals (n = 3). Methods outlined in section 2.6.3. **(A)** *TL+ MESF APC for each healthy and cancer-associated plasma enriched sEVs sample.* **(B)** *Comparison of TL+ MESF APC between plasma enriched sEVs derived from healthy individuals and cancer-associated patients.* **(C)** *Comparison of TL+ APC between plasma enriched sEVs derived from healthy individuals, non-metastatic and metastatic patients.* **(D)** *ROC curve analysis measuring TL+ MESF APC, yielding an AUC of 0.5278 for discriminating cancer patients from healthy individuals. Each experiment comprised biological triplicates with three technical triplicates; error bars represent Stdev. T-test followed by Welch's correction determined significance levels: *p < 0.05, **p < 0.01.*

Supplementary Table 1. List of cleared or approved diagnostic tests for breast cancer applications

Diagnostic Name Manufacturer	Sample Type	Biomarker details	Clinical Application	Assay	FDA Approval (PMA or 510(k))
Bond Oracle HER2 IHC System (Leica Biosystems)	Breast Cancer - Tissue	HER-2 protein overexpression	Predication of therapy	Immunohistochemistry	P090015 (04/18/2012)
BRACAnalysis CDx (Myriad Genetic Laboratories, Inc.)	Breast Cancer - Whole Blood	BRCA1 and BRCA2 Mutations	Predication of therapy	PCR (Polymerase Chain Reaction) and Sanger sequencing	P140020/S012 (01/12/2018)
ADVIA Centaur CA 15-3	Breast Cancer - Serum	CA 15-3	Predication of therapy, Monitoring	Sandwich immunoassay	K012357 02/28/2002
ADVIA Centaur CA 27.29	Breast Cancer - Serum	CA 27.29	Predication of therapy, Monitoring	Sandwich immunoassay	K193489 02/13/2020
VIDAS CEA (S) ASSAY (BIOMERIEUX, INC.)	Breast Cancer – Serum / Plasma	Carcinoembryonic Antigen (CEA)	Prognosis, Monitoring	Sandwich immunoassay	K080194 10/09/2008
FoundationOne CDx (Foundation Medicine, Inc.)	Breast Cancer - Tissue	ERBB2 (HER2) amplification	Predictive of therapy, Monitoring, Staging	Next-generation sequencing (NGS)-based assay	P170019 (11/30/2017)
FoundationOne CDx (Foundation Medicine, Inc.)	Breast Cancer - Tissue	PIK3CA by C420R, E542K, E545A, E545D [1635G>T only], E545G, E545K, Q546E, Q546R, H1047L, H1047R, and H1047Y mutations	Predication of therapy	NGS-based assay	P170019/S006 (12/03/2019)
FoundationOne CDx (Foundation Medicine, Inc.)	Breast Cancer - Tissue	PIK3CA by AKT1/PTEN alterations	Predication of therapy	NGS-based assay	P170019/S048 (11/16/2023)
FoundationOne Liquid CDx (Foundation Medicine, Inc.)	Breast Cancer - Plasma	PIK3CA by C420R, E542K, E545A, E545D [1635G>T only], E545G, E545K, Q546E, Q546R, H1047L, H1047R, and H1047Y	Predication of therapy	NGS-based assay	P200006 (10/26/2020)
Guardant360 CDx (Guardant Health, Inc.)	Breast Cancer - Plasma	ESR1 missense mutations between codons 310 and 547	Predication of therapy	NGS-based assay	P200010/S010 (01/27/2023)
HER2 CISH pharmDx Kit (Dako Denmark A/S)	Breast Cancer - Tissue	HER-2/neu (ERBB2) gene amplification	Prognosis	chromogenic in situ hybridization (CISH)	P100024 (11/30/2011)
HER2 FISH pharmDx Kit (Dako Denmark A/S)	Breast Cancer - Tissue	HER-2/neu (ERBB2) gene amplification	Prognosis, predication of therapy	Fluorescence in situ hybridization (FISH)	P040005/S009 (02/22/2013)
HercepTest (Dako Denmark A/S)	Breast Cancer - Tissue	HER-2 protein overexpression	Predication of therapy	Immunohistochemistry	P980018/S016 (02/22/2013)
INFORM HER2 Dual ISH DNA Probe Cocktail (Ventana Medical Systems, Inc.)	Breast Cancer - Tissue	HER-2/neu (ERBB2) gene amplification	Prognosis, Predication of therapy	FISH	P100027/S030 (05/03/2019)
InSite Her-2/neu (CB11) Monoclonal Antibody (Biogenex Laboratories, Inc.)	Breast Cancer - Tissue	HER2 protein overexpression	Predication of therapy	Immunohistochemistry	P040030 (12/22/2004)
Ki-67 IHC MIB-1 pharmDx (Dako Omnis) (Agilent Technologies)	Breast Cancer - Tissue	Ki-67 protein expression	Predication of therapy	Immunohistochemistry	P210026 (10/12/2021)
PATHWAY anti-Her2/neu (4B5) Rabbit Monoclonal Primary Antibody (Ventana Medical Systems, Inc.)	Breast Cancer - Tissue	HER-2 protein overexpression	Predication of therapy	Immunohistochemistry	P990081/S047 (09/30/2022)
PD-L1 IHC 22C3 pharmDx (Dako North America, Inc.)	Triple-Negative Breast Cancer (TNBC) - Tissue	PD-L1 protein expression	Predication of therapy	Immunohistochemistry	P150013/S020 (11/13/2020)
SPOT-LIGHT HER2 CISH Kit (Life Technologies Corporation)	Breast Cancer - Tissue	HER-2/neu (ERBB2) gene amplification	Predication of therapy	CISH	P050040 (07/01/2008)
therascreen PIK3CA RGQ PCR Kit (QIAGEN GmbH)	Breast Cancer - Tissue or Plasma	PIK3CA by C420R, E542K, E545A, E545D [1635G>T only], E545G, E545K, Q546E, Q546R, H1047L, H1047R, and H1047Y	Predication of therapy	PCR	P190001 (05/24/2019) P190004 (05/24/2019)
Ventana HER2 Dual ISH DNA Probe Cocktail (Ventana	Breast Cancer - Tissue	HER-2/neu (ERBB2) gene amplification	Predication of therapy	CISH	P190031 (07/28/2020)

Medical Systems, Inc.)					

Supplementary Table 2. List of cleared or approved diagnostic tests for colorectal cancer applications

Diagnostic Name Manufacturer	Sample Type	Biomarker details	Clinical Application	Assay	FDA Approval (PMA or 510(k))
cobas KRAS Mutation Test (Roche Molecular Systems, Inc.)	Colorectal Cancer - Tissue	KRAS by Mutations in codons 12 and 13 of KRAS gene	prediction of therapy	PCR	P140023 (05/07/2015)
CRCDx RAS Mutation Detection Assay Kit (EntroGen, Inc.)	Colorectal Cancer - Tissue	KRAS wild-type biomarkers (the absence of mutations in exons 2, 3, or 4) and NRAS wild-type biomarkers (the absence of mutations in exons 2, 3, or 4)	prediction of therapy	PCR	P220005 (09/29/2023)
Dako EGFR pharmDx Kit (Dako North America, Inc.)	Colorectal Cancer - Tissue	EGFR (HER1) protein expression	prediction of therapy	Immunohistochemistry	P030044/S002 (09/27/2006)
Carcinoembryonic antigen flex reagent cartridge (Dade Behring, Inc)	Colorectal Cancer - Serum	CEA	Monitoring, Prognosis	Sandwich immunoassay	k071603 06/25/2008
Epi proColon (Epigenomics AG)	Colorectal Cancer - Plasma	Methylated Septin 9 DNA	Screening and prognosis	PCR	P130001 04/12/2016
FoundationOne CDx (Foundation Medicine, Inc.)	Colorectal Cancer - Tissue	KRAS wild-type (absence of mutations in codons 12 and 13)	prediction of therapy	NGS-based assay	P170019 (11/30/2017)
FoundationOne CDx (Foundation Medicine, Inc.)	Metastatic Colorectal Cancer - Plasma	BRAF by V600E alteration	prediction of therapy	NGS-based assay	P190032/S010 (06/08/2023)
Hemoccult® ICT (Beckman Coulter, Inc)	Faeces	Hemoglobin	Screening	Immunochemical assay	K080812 06/25/2008
OC-Light S FIT (EIKEN CHEMICAL CO. LTD)	Faeces	Hemoglobin	Screening	Immunoassay	K143325 08/20/2015
ONCO/Reveal Dx Lung & Colon Cancer Assay (O/RDx-LCCA) (Pillar Biosciences, Inc.)	Colorectal Cancer - Tissue	KRAS wild-type (absence of mutations in codons 12 and 13)	prediction of therapy	NGS-based assay	P200011 (07/30/2021)
Praxis Extended RAS Panel (Illumina, Inc.)	Colorectal Cancer - Tissue	KRAS wild-type (absence of mutations in exons 2, 3, and 4) and NRAS wild type (absence of mutations in exons 2, 3, and 4)	prediction of therapy	NGS-based assay	P160038 (06/29/2017)
therascreen BRAF V600E RGQ PCR Kit (QIAGEN GmbH)	Colorectal Cancer - Tissue	BRAF by V600E alteration	prediction of therapy	PCR	P190026 (04/15/2020)
therascreen KRAS RGQ PCR Kit (Qiagen Manchester, Ltd.)	Colorectal Cancer - Tissue	KRAS by G12A, G12D, G12R, G12C, G12S, G12V, G13D	prediction of therapy	PCR	P110027 (05/23/2014)
therascreen KRAS RGQ PCR Kit (Qiagen Manchester, Ltd.)	Colorectal Cancer - Tissue	KRAS wild-type (absence of mutations in codons 12 and 13)	prediction of therapy	PCR	P110027/S013 (12/02/2022)
xT CDx (Tempus Labs, Inc.)	Colorectal Cancer - Tissue (Matching Blood/Saliva)	KRAS wild-type (absence of mutations in exons 2, 3, or 4) and NRAS wild-type (absence of mutations in exons 2, 3, or 4)	Prognosis, predication of therapy	NGS-based assay	P210011 (04/28/2023)

Supplementary Table 3. Overview of the 95 lectins used on the lectin microarrays, forms, abbreviations, specificity groups, and specificities as provided by manufacturer (RayBiotech Life).

Lectin form	Abbreviation	Specificity Group	Specificity
Anguilla anguilla	AAA	fucose	αFuc
Aleuria aurantia	AAL	fucose	Fucα6GlcNAc
Burkholderia cenocepacia	BC2LCN	fucose	Fucα1-2Galβ1-3GalNAc, Fucα1-2Galβ1-3GlcNAc
Laburnum anagyroides	LAL	fucose	LALa-Me-L-Fuc
Lotus tetragonolobus	LOTUS	fucose	αFuc
Pseudomonas aeruginosa II	PA-IIL	fucose	Fuc and Fuc containing oligosaccharides
Pseudomonas aeruginosa I	PA-IL	fucose	Fuc and Fuc containing oligosaccharides
Ralstonia solanacearum	RS-FUC	fucose	Fuc
Ulex europaeus I	UEA-I	fucose	αFuc
Agaricus bisporus	ABA	galactose	Gal (β1,3) GalNAc
Agaricus bisporus	ABL	galactose	Gal (β-1,3) GalNAc Gal (β-1,3) GlcNAc
Amaranthus caudatus	ACL	galactose	Gal (β1,3) GalNAc
Bauhinia purpurea	BPA	galactose	Gal (β1,3) GalNAc
Clitocybe nebularis	CNL	galactose	Gal (β1-4) GlcNAc
Dictyostelium discoideum	DISCOIDIN II	galactose	Gal, LacNAc, asialoglycans, Gal/GalNAcβ1-4GlcNAcβ1-6Gal/GalNAc
Erythrina cristagalli	ECA	galactose	Galβ4GlcNAc
Eunonymus europaeus	EEL	galactose	Galα3Gal
Human galectin 7-S	GAL-7S	galactose	Galβ1-3GlcNAc
Glechoma hederacea	GHA	galactose	Gal, methyl a-D-galactopyranoside, GalNAc
Griffonia (Banderaea) simplicifolia I	GSI	galactose	αGal, α3GalNAc
Jacalin	JACALIN	galactose	Galβ3GalNAc
Maackia amurensis I	MAA	galactose	Galβ4GlcNAc
Pure Morniga G	MNA-G	galactose	Gal
Marasmius oreades	MOA	galactose	3Galβ1-4GlcNAc, Galα1-3Gal
Maclura pomifera	MPL	galactose	Galβ3GalNAc
Phaseolus vulgaris erythroagglutinin	PHA-E	galactose	Bisected bi- and tri-antennary complex-type oligosaccharides
Phaseolus Vulgaris leucoagglutinin	PHA-L	galactose	2,6-Branched tri- and tetra-antennary complex-type oligosaccharides
Phaseolus Vulgaris phytohemagglutinin	PHA-P	galactose	Branched complex-type poly-oligosaccharides
Peanut	PNA	galactose	Galβ3GalNAc
Ricinus communis agglutinin I	RCA120	galactose	Gal, Lac
Ricinus communis agglutinin II	RCA60	galactose	Gal, αGalNAc, Lac
Vigna radiata	VRA	galactose	α-Gal
Arum maculatum	AMA	mannose	αMan
Allium sativum	ASA	mannose	αMan
Musa acuminata	BANLEC	mannose	containing α1,3-glycoside bond
Burkholderia cenocepacia	BC2L-A	mannose	αMan
Calystegia sepium	CALSEPA	mannose	αMan

Coanavalin A	CONA	mannose	αMan, αGlc
Galanthus nivalis	GNA	mannose	αMan
Griffithia sp.	GRFT	mannose	αMan
Hippeastrum hybrid	HHA	mannose	αMan
Lentil	LENTIL	mannose	D-Man, D-Glu
Human malectin	MALECTIN	mannose	Glc2-N-biose
Morniga M	MNA-M	mannose	αMan
Narcissus pseudonarcissus	NPA	mannose	αMan
Oryza sative	ORYSATA	mannose	αMan
Phlebodium aureum	PALa	mannose	αMan
Pisum sativum	PSA	mannose	αMan, αGlc, αFuc
Vicia faba	VFA	mannose	αMan
Vicia villosa	VVA-M	mannose	αMan
Colchicum autumnale	CA	N-acetylgalactosamine	Lac > αGalNAc > Gal
Caragana arborescens	CAA	N-acetylgalactosamine	αGalNAc
Coprinopsis cinerea	CGL-2	N-acetylgalactosamine	GalNAcα1-3Gal, Galα1-3Gal
Cicer arietinum	CPA	N-acetylgalactosamine	fetuin
Cytisus scoparius	CSA	N-acetylgalactosamine	αGalNAc
Dolichos biflorus	DBA	N-acetylgalactosamine	αGalNAc
Dictyostelium discoideum	DISCOIDIN I	N-acetylgalactosamine	αGalNAc, LacNAc
Human galectin 2	GAL-2	N-acetylgalactosamine	GalNAcα1-3Gal, branched LacNAc
Helix aspersa	HAA	N-acetylgalactosamine	αGalNAc
Iris hybrid	IRA	N-acetylgalactosamine	αGalNAc
Phaseolus lunatus	LBA	N-acetylgalactosamine	GalNAcα(1,3)[αFuc(1,2)]Gal
Pleurocybella porrigens	PPL	N-acetylgalactosamine	α/βGalNAc
Psophocarpus tetragonolobus I	PTL-1	N-acetylgalactosamine	αGalNAc, Gal
Psophocarpus tetragonolobus II	PTL-2	N-acetylgalactosamine	blood group H structures and the T-antigen
Robinia pseudoacacia	RPA	N-acetylgalactosamine	αGalNAc,
Soybean	SBA	N-acetylgalactosamine	α > βGalNAc
Salivia horminum	SHA	N-acetylgalactosamine	αGalNAc
Sophora japonica	SJA	N-acetylgalactosamine	α > βGalNAc
Tulipa	TL	N-acetylgalactosamine	αGalNAc
Vicia villosa	VVA	N-acetylgalactosamine	αGalNAc
Wisteria floribunda	WFA	N-acetylgalactosamine	αGalNAc
Datura stramonium	DSA	N-acetylglucosamine	αGlcNAc
e. coli	F17AG	N-acetylglucosamine	αGlcNAc
Griffonia (Banderaea) simplicifolia II	GSII	N-acetylglucosamine	α or βGlcNAc
Lycopersicon esculentum	LEA	N-acetylglucosamine	αGlcNAc
Phytolacca americana	PWA	N-acetylglucosamine	GlcNAc(β1,4) GlcNAc oligomers, [Gal-(β1,4)
Solanum tuberosum	STL	N-acetylglucosamine	αGlcNAc
Urtica dioica	UDA	N-acetylglucosamine	αGlcNAc

Ulex europaeus II	UEA-II	N-acetylglucosamine	αGlcNAc
Wheat germ agglutinin	WGA	N-acetylglucosamine	αGlcNAc
Human galectin 1	GAL-1	N-acetyllactosamine	branched LacNAc, Gal
Human galectin 1-S	GAL-1S	N-acetyllactosamine	branched LacNAc
Human galectin 3	GAL-3	N-acetyllactosamine	poly LacNAc
Human galectin 3C-S	GAL-3C-S	N-acetyllactosamine	poly LacNAc
Human galectin 9	GAL-9	N-acetyllactosamine	poly LacNAc, GalNAcα1-3Gal
Laetiporus sulphureus	LSLN	N-acetyllactosamine	LacNAc, poly LacNAc
Agrocybe cylindracea	ACG	sialic acid	αSia
Homarus americanus	HMA	sialic acid	αSia, GalNAc
Limulus polyphemus	LPA	sialic acid	αSia
Polyporus squamosus	PSL-1A	sialic acid	α2-6 Sia
Sambucus sieboldiana	SAMB	sialic acid	NeuAcα2-6Gal/GalNAc
Sambucus nigra I	SNA-I	sialic acid	αSia (2,6)GalNAc > GalNAc = Lac >
Sambucus nigra II	SNA-II	sialic acid	αSia (2,6)GalNAc > GalNAc = Lac >
Salvia sclarea	SSA	sialic acid	αSia(2,6)GalNAc > GalNAc = Lac >

Supplementary Table 4. The MACSPlex EV kit comprising of 37 surface epitopes that are present on sEVs plus two isotype control beads

Antibody	Isotype	Marker Groups
CD81	Recombinant human IgG1	Exosome marker
CD63	Mouse IgG1k	Exosome marker
CD9	Mouse IgG1	Exosome and platelet markers
CD41b	Recombinant human IgG1	Platelet marker
CD42a	Recombinant human IgG1	Platelet marker
CD62P	Recombinant human IgG1	Platelet marker
CD29	Mouse IgG1k	Platelet and leukocyte marker
CD45	Mouse IgG2a	Platelet and leukocyte marker
CD142	Mouse IgG1k	Leukocyte marker
CD24	Mouse IgG1	Neutrophil, platelet and leukocyte marker
CD14	Mouse IgG2a	Monocyte and leukocyte marker
HLA--DRDPDQ	Recombinant human IgG1	Leukocyte, monocyte and macrophages marker
CD40	Mouse IgG1k	Leukocyte, monocyte and macrophages marker
CD86	Mouse IgG1	Leukocyte, monocyte and macrophages marker
CD1c	Mouse IgG2a	Leukocyte, monocyte and macrophages marker
CD11c	Mouse IgG2b	Leukocyte, monocyte and macrophages marker
CD209	Mouse IgG1	Leukocyte, monocyte and macrophages marker
CD2	Mouse IgG2b	T cell and leukocyte marker
CD3	Mouse IgG2a	T cell and leukocyte marker
CD4	Mouse IgG2a	T cell and leukocyte marker
CD8	Mouse IgG2a	T cell and leukocyte marker

CD56	Recombinant human IgG1	Natural killer cell and leukocyte marker
CD69	Mouse IgG1k	T cell and leukocyte marker
CD25	Mouse IgG1	T cell and leukocyte marker
CD19	Mouse IgG1	B cell and leukocyte marker
CD20	Mouse IgG1	B cell and leukocyte marker
CD31	Mouse IgG1	Platelet, endothelium and leukocyte marker
CD146	Mouse IgG1	Endothelium marker
CD105	Recombinant human IgG1	Endothelium marker
CD326	Mouse IgG1	Epithelium marker
SSEA-4	Recombinant human IgG1	Stem cell marker
CD133/1	Mouse IgG1	Stem cell marker
HLA-ABC	Recombinant human IgG1	Multiple cell marker
CD44	Mouse IgG1	Multiple cell marker
CD49e	Recombinant human IgG1	Multiple cell marker
ROR1	Mouse IgG1k	Adipocytes, parathyroid and cancer marker
MCSP	Mouse IgG1	Melanocytes, smooth muscle cell and cancer marker
REA ctrl	Recombinant human IgG1	Isotype control
mIgG1 ctrl	Mouse IgG1	Isotype control

Supplementary Table 5. MIFlowCyt-EV checklist for epithelial cell line sEVs and human plasma-enriched sEVs experiments

Framework criteria	What to report	Epithelial cell line sEVs experiments	Human plasma enriched sEVs experiments
1.1 Preanalytical variables conforming to MISEV guidelines.	Preanalytical variables relating to EV sample including source, collection, isolation, storage, and any others relevant and available in the performed study.	Epithelial cell line derived sEVs were isolated as described in section 2.4.1.	Human plasma enriched derived sEVs were isolated as described in section 2.4.2.2.
1.2 Experimental design according to MIFlowCyt guidelines.	EV-FC manuscripts should provide a brief description of the experimental aim for the performed FC experiment(s) using MIFlowCyt checklist criteria: 1.1, 1.2, and 1.3, respectively.	<u>Aims</u> To identify CD81 and CD63 tetraspanin composition of epithelial cell line sEVs (section 3.4.3). To identify HPA, LCA and TL binding to 'CD81-positive' and 'CD63-positive' epithelial cell line sEVs (section 4.4,3, 5.4.2, 5.4.4).	<u>Aims</u> To identify the diagnostic potential of HPA, LCA and TL lectin to 'EpCAM-positive' plasma-enriched derived sEVs (section 6.4.3, 6.4.6, 6.4.7, 6.4.8, 6.4.9).
2.1 Sample staining details	State any steps relating to the staining of samples. Along with the method used for staining, provide relevant reagent descriptions as listed in MIFlowCyt guidelines (section 2.6 fluorescence reagent(s) descriptions).	Methods used for the staining of sEVs are outlined in section 2.6. The methods for staining optimisation are described in section 3.3.3.	Methods used for the staining of sEVs are outlined in section 2.6. The methods for staining optimisation are described in section 3.3.3.
2.2 Sample washing details	State any steps relating to the washing of samples.	Epithelial cell line sEVs were stained with CFDA-SE and washed using NAP-5 SEC columns before further labelling as outlined in section 2.6.	Human plasma enriched sEVs were stained with CFDA-SE and washed using NAP-5 SEC columns before further labelling as outlined in section 2.6.
2.3 Sample dilution details	All methods and steps relating to sample dilution.	Samples were diluted 1:20 to a final volume of 500 µL prior to acquisition.	Samples were diluted 1:20 to a final volume of 500 µL prior to acquisition.
3.1 Buffer alone controls.	State whether a buffer-only control was analysed at the same settings and during the same experiment as the samples of interest. If utilized it is recommended that all samples be recorded for a consistent set period of time e.g. 5 minutes, rather than stopping analysis at a set recorded event count e.g. 100,000 events. This allows comparisons of total particle count between controls and samples.	Dilution buffer only control (PBS) (pH 7.4) was run alongside each experiment	Dilution buffer only control (PBS) (pH 7.4) was run alongside each experiment
3.2 Buffer with reagent controls.	State whether a buffer with reagent control was analysed at the same settings, same concentrations, and during the same experiment as the samples of interest. If used state what the results were.	PBS dye, antibody and lectin control samples were run for each experiment and in combination when dual labelling was performed.	PBS dye, antibody and lectin control samples were run for each experiment and in combination when dual labelling was performed.

3.3 Unstained controls	State whether unstained control samples were analysed at the same settings and during the same experiment as stained samples. If used, state what the results were, preferably in standard units.	Unstained epithelial cell line sEVs samples were analysed using the same settings as stained samples.	Unstained epithelial cell line sEVs samples were analysed using the same settings as stained samples.
3.4 Isotype controls	The use of isotype controls is applicable to immunofluorescence labelling only. State whether isotype controls were analyzed at the same settings and during the same experiment as stained samples. If utilized, state which antibody they are matched to, the concentration used, and what the results were (section 4.2, 4.3, 4.4). Due to conjugation differences between manufacturers if should be stated if the isotype controls are from the same manufacturer as the matched antibodies.	Isotypes were used for detection antibodies at the same concentration and used for gating purposes (Table 2.5).	Isotypes were used for detection antibodies at the same concentration and used for gating purposes (Table 2.5).
3.5 Single-stained controls.	State whether single-stained controls were included. If used state whether the single-stained controls were recorded using the same settings, dilutions, and during the same experiment as stained samples and state what the results were, preferably in standard units (section 4.2, 4.3, 4.4).	Single stained controls were included, recorded in the same settings, and dilutions.	Single stained controls were included, recorded in the same settings, and dilutions.
3.6 Procedural controls.	State whether procedural controls were included. If used, state the procedure and if procedural controls were acquired at the same settings and during the same experiment as stained samples.	N/A	N/A
3.7 Serial dilutions.	State whether serial dilutions were performed on samples and note the dilution range and manner of testing. The fluorescence and/or scatter signal intensity would ideally be reported in standard units (see section 4.3, 4.4) but arbitrary units can also be used. This data is best reported by plotting the recorded number events/concentration over a set period of time at different sample dilution. The MFI at each of the dilutions should also ideally be plotted on the same or a separate plot	Serial dilution controls for each experiment were included and recorded using the same settings	Serial dilution controls for each experiment were included and recorded using the same settings
3.8. Detergent treated EVsamples	State whether samples were detergent treated to assess lability. If utilized, state what detergent was used, the end concentration of the detergent, and what the results were of the lysis.	Detergent treated controls for each experiment were included and recorded using the same settings.	Detergent treated controls for each experiment were included and recorded using the same settings.
4.1 Trigger Channel(s) and Threshold(s).	The trigger channel(s) and threshold(s) used for event detection. Preferably, the fluorescence calibration (section 4.3) and/or scatter calibration (section 4.4) should be used in order to report the trigger channel(s) and threshold(s) in standardized units.	The Amnis Cellstream instrument triggers automatically on all channels for any signal/pixel above background levels. No additional thresholds were defined	The Amnis Cellstream instrument triggers automatically on all channels for any signal/pixel above background levels. No additional thresholds were defined
4.2 Flow Rate / Volumetric quantification.	State if the flow rate was quantified/validated and if so, report the result and how they were obtained.	The volumetric flow rate was 3.66 µL/min and was otherwise not investigated.	The volumetric flow rate was 3.66 µL/min and was otherwise not investigated.
4.3 Fluorescence Calibration.	State whether fluorescence calibration was implemented, and if so, report the materials and methods used, catalogue numbers, lot numbers, and supplied reference units for the standards. Fluorescence parameters may be reported in standardized units of MESF, ERF, or ABC beads. The type of regression used, and the resulting scatter plot of arbitrary data vs standard data for the reference particles should be supplied.	Fluorescence calibration was performed (Supplementary figure 3) and all data is provided in PE or APC MESF units. Details regarding calibration beads used are provided in section 2.6.2.	Fluorescence calibration was performed (Supplementary figure 3) and all data is provided in PE or APC MESF units. Details regarding calibration beads used are provided in section 2.6.2.
4.4 Light Scatter Calibration.	State whether and how light scatter calibration was implemented. Light scatter parameters may be reported in standardized units of nm2, along with information required to reproduce the model	Light scatter calibration method is described in section 2.6.2 and illustrated in supplementary figure 2.	Light scatter calibration method is described in section 2.6.2 and illustrated in supplementary figure 2.
5.1 EV diameter/surface area/volume approximation.	State whether and how EV diameter, surface area, and/or volume has been calculated using FC measurements.	Diameter of particles has been estimated by NTA throughout. EV diameter has been qualitatively evaluated by TEM	Diameter of particles has been estimated by NTA throughout. EV diameter has been qualitatively evaluated by TEM
5.2 EV refractive index approximation.	State whether the EV refractive index has been approximated and how this was done.	EV refractive index has not been assessed in this study.	EV refractive index has not been assessed in this study.
5.3 EV epitope number approximation.	State whether EV epitope number has been approximated, and if so, how it was approximated.	N/A	N/A
6.1 Completion of MIFlowCyt checklist.	Complete MIFlowCyt checklist criteria 1 to 4 using the MIFlowCyt guidelines. Template found at www.evflowcytometry.org.	Table S2	Table S2
6.2 Calibrated channel detection range	If fluorescence or scatter calibration has been carried out, authors should state whether the upper and lower limits of a calibrated detection channel was calculated in standardized units. This can be done by converting the arbitrary unit scale to a calibrated scaled, as discussed in section 4.3 and 4.4, and providing the highest unit on this scale and the lowest detectable unit above	Fluorescence calibrated data based on usage of PE and APC calibration beads as shown in supplementary figure 3. Unit scale conversion was used to convert median values to MESF.	Fluorescence calibrated data based on usage of PE and APC calibration beads as shown in supplementary figure 3. Unit scale conversion was used to convert median values to MESF.

	the unstained population. The lowest unit at which a population is deemed 'positive' can be determined a variety of ways, including reporting the 99th percentile measurement unit of the unstained population for fluorescence. The chosen method for determining at what unit an event was deemed positive should be clearly outlined.		
6.3 EV number/concentration.	State whether EV number/concentration has been reported. If calculated, it is preferable to report EV number/concentration in a standardized manner, stating the number/concentration between a set detection range.	Cellstream enables output for events/ml. To facilitate comparison and normalisation for the varying amounts of loaded sEVs in each sample, the events/ml for the final gated lectin+ signal was calculated as a percentage of the CFSE+ events/ml.	Cellstream enables output for events/ml. To facilitate comparison and normalisation for the varying amounts of loaded sEVs in each sample, the events/ml for the final gated lectin+ signal was calculated as a percentage of the CFSE+ events/ml.
6.4 EV brightness.	When applicable, state the method by which the brightness of EVs is reported in standardized units of scatter and/or fluorescence.	All experimental runs data is reported in PE or APC MESF depending on label used.	All experimental runs data is reported in PE or APC MESF depending on label used.