Characterisation of extracellular vesicles released into culture media by human embryos



Maryam Rahbar

St. Catherine's College

University of Oxford

Nuffield Department of Women's & Reproductive Health

Supervised by:

Prof Manu Vatish

Prof Dagan Wells

Dr Sofia Cerdeira

A thesis submitted for the degree of Doctor of Philosophy

Trinity Term 2023

45,000 words

Copyright © Maryam Rahbar

All rights reserved

Acknowledgements

Throughout my DPhil journey, it was my support system that helped me deal with the challenges as well as celebrate the small wins, and so I would like to take this opportunity to thank all of those who made this journey possible. First and foremost, I am deeply grateful to my supervisor (Prof Manu Vatish), whose guidance, invaluable insights, and unwavering support throughout my DPhil have been instrumental in shaping my research direction and enabling me to reach this milestone. Thank you, Manu, for being the best supervisor I could have asked for, for always being there, for all your support and for never making me feel like a burden. You made this journey possible and I will always be indebted to you for your encouragement and meaningful support. I would also like to thank Dr Sofia Cerdeira, Dr Wei Zhang, and Prof Dagan Wells for their help and support whenever I needed it.

A big thank you to all the members of team Manu (Carolina Motta Mejia, Neva Kandzija, Toluwalase Awoyemi, Shuhan Jiang, Prassana Logenthiran, Cathy Wang, William Cooke, Emily Mazey, Emily Hyde, Gabriel Jones, Faheem Seedat, Catarina Dos Palmos, Yushi Abe, Angga Lokeswara, Michelle Ma, Bryony Davies and Qian Chen) for their guidance, friendship and support. A special thank you to Tolu, Shuhan, Prassana and Neva. Tolu, you were the person who taught me my first set of experiments and I will be forever grateful for all that you taught me since that first western blot. Shuhan, you were my first friend in the lab and helped me through the entire DPhil process, both academically and personally. Prassana, you were also a significant influence and I appreciate all of your wisdom. Neva, thank you for sharing your experiences and always listening to my rants. You are all great friends and I will cherish all that you taught me academically, professionally, and personally.

I would like to thank my husband (Amir Bahador) for all his support throughout my DPhil journey. Thank you for always listening to me, motivating me, and pushing me to be the best version of myself. Thank you for being there after those long days in the lab and giving me hope and unconditional support. I could not have done it without you. I would like to thank my parents for all of their help and support. I would not be where I am today if it was not for all their sacrifices. Thank you for giving me the confidence to get where I am today and teaching me that nothing is impossible and that with hard work, I can achieve everything I set my mind to. A big thank you to my older brother (Ali), sister-in-laws (Shirine and Behnaz) and my best friend (Sara) for always

being there for me, cheering me up and showing me that life is more than just academics. Another thank you to my in-laws (Mahdis and Ali) for being my second parents in the UK and always taking care of me. I am forever grateful to each and every one of you and I dedicate this thesis to you all. I would not have been able to do it without you. Thank you all for your immense contributions towards making this DPhil journey a memorable and fulfilling experience.

Publications

During my DPhil, I contributed to experimental data, analysis and writing of manuscripts for these publications and abstracts:

- Kandzija N[1], **Rahbar M**[1], Jones GD, Motta-Mejia C, Zhang W, Couch Y, Neuhaus A A, Kishore U, Sutherland B A, Redman C, and Vatish M. Placental Capillary Pericytes release excess extracellular vesicles under hypoxic conditions inducing a pro-angiogenic profile in term pregnancy. *Biochem Biophys Res Commun*. 2023;651:20-29.
- Awoyemi T, Cerdeira S, Zhang W, Jiang S, **Rahbar M**, Logentheiran P, Redman C and Vatish M. Preeclampsia and syncytiotrophoblast membrane extracellular vesicles (STB-EVs). *Clin Sci (Lond)*. 2022;136(24):1793-1807.
- Jiang S, Zhang W, Cao Q, **Rahbar M**, Cooke W, Ono M, Fujimaki K, Saito A, Itonaga M, Oda K, Redman C, and Vatish M. ExoCounter assays identify women who may develop early-onset preeclampsia from 12.5 μL first trimester serum by characterising placental small extracellular vesicles [Under Review, *Hypertension*].
- Awoyemi T, Jiang S, Bjarkadottir B, **Rahbar M**, Logenthiran P, Collett G, Zhang W, Cribbs A, Cerdeira S, and Vatish M. Proteomic analysis of the placenta and its extracellular vesicles in normal and preeclampsia pregnancies: identification of novel biomarkers and insights into disease mechanisms [Under Review, *American journal of obstetrics and gynaecology*]
- Awoyemi T, Zhang W, **Rahbar M**, Cribbs A, Logenthiran P, Jiang S, Collett G, Cerdeira S, and Vatish M. Transcriptomic analysis of placenta and its released extracellular vesicles in normal and preeclampsia pregnancies: insights into novel biomarkers and mechanisms of disease [Under Review, *American journal of obstetrics and gynaecology*].
- Awoyemi T, Zhang W, Tanetta D, Collett G, Logenthiran P, Kandzija N, **Rahbar M**, Jiang S, Redman C, Vatish M. Investigating the roles of small STBEV's microRNAs in the pathogenesis of early-onset preeclampsia through bioinformatics [Abstract]. Placenta; 2021, 112.
- Jiang S, Awoyemi T, Logenthiran P, **Rahbar M**, Zhang W, Tuffin J, Saleem M, and Vatish M. Syncytiotrophoblast Extracellular Vesicles Can Be Internalized By Human Podocytes And Influence Their Physiological Functions [Abstract]. In: 13th International Podocyte Conference; 2021 July 27-31. Manchester, United Kingdom.
- Logenthiran P, Awoyemi T, Jiang S, **Rahbar M**, Kandzija N, Zhang W, Cerdeira S, Vatish M. Syncytiotrophoblast and syncytiotrophoblast derived extracellular vesicles express Kynurenine metabolising enzymes [Abstract]. In: International Federation of Placenta Associations; 2021 August 30 – September 2. Online.
- Awoyemi T, **Rahbar M**, Cribbs A, Zhang W, Redman C, Vatish M. Identification of Potential Aetiopathological Pathways in Early Onset Preeclampsia (EOPE) through Cluster Analysis of Medium/Large Syncytiotrophoblast Membrane Extracellular Vesicle's Protein-Protein Interaction Network [Abstract]. In: International Society of Extracellular Vesicles; 2021 May 18-21. Online.

Table of Contents

ACKNOWLEDGEMENTS **2**

PUBLICATIONS **4**

LIST OF ABBREVIATIONS **9**

CHAPTER 1

INTRODUCTION – LITERATURE REVIEW **13**

INFERTILITY 14
TREATMENT OPTIONS 14
EMBRYO DEVELOPMENT IN VITRO 17
IMPLANTATION 18
ENDOMETRIAL CHANGES 20
MORPHOLOGICAL ANALYSIS OF EMBRYOS 21
PGT AND NI-PGT 22
EXTRACELLULAR VESICLES 23
Small EVs (S-EVs) *23*
Medium/large EVs (M/L-EVs) *24*
Apoptotic bodies (Apo-EVs) *25*
Cargo of EVs *25*
Uptake of EVs by recipient cells *25*
EXTRACELLULAR VESICLES IN PREGNANCY 26
EMBRYONIC EXTRACELLULAR VESICLES (EEVS) 27
CHARACTERISATION OF EEVS USING NEW TECHNOLOGIES 31
AIMS OF THESIS 32

CHAPTER 2

CHARACTERISATION OF HUMAN EMBRYONIC EXTRACELLULAR VESICLES **33**

INTRODUCTION 34
Extracellular vesicles *34*
Embryonic EVs (EEVs) *34*
Other factors in the secretome of embryos *35*
Aims *36*
MATERIALS AND METHODS 36
Ethics *36*
Sample collection *36*
EV isolation for NTA *37*
Bicinchoninic acid protein assay *37*
Nanoparticle tracking analysis (NTA) *38*
Transmission electron microscopy *39*
HEK293 cell culture and lysate preparation as control for western blot *39*
Western blot analysis *40*
MACSPlex sample processing *41*
MACSPlex kit *42*
Flow cytometry for MACSPlex *42*
Analysis of relative quantification of small EV markers *43*
HUVEC cells culture media and MACSPlex *43*
ExoCounter *44*

ExoCounter data analysis ... 45
NanoFCM plasma sample preparation ... 45
NanoFCM ... 45
RESULTS ... 46
Particle tracking analysis, transmission electron microscopy, and western blot ... 46
MACSPlex with EEVs, control culture media and HUVECs ... 48
ExoCounter with EEVs for CD142, CD63 and CD29 ... 53
NanoFCM with EEVs and plasma ... 54
DISCUSSION ... 59
CD146 ... 60
CD29 ... 60
CD63 ... 61
CD142 ... 62
New technologies for EV characterisation ... 63
Limitations and Future work ... 64

CHAPTER 3

ISHIKAWA CELLS AND THE EFFECTS OF HUMAN RECOMBINANT CD142 PROTEIN TREATMENT ... **66**

INTRODUCTION ... 67
Implantation ... 67
Endometrial layers ... 68
Cell lines ... 69
Ishikawa cells ... 70
Role of CD142 in pregnancy ... 70
Aims ... 71
MATERIALS AND METHODS ... 71
Cell culture of Ishikawa cells ... 71
Hormonal treatment of Ishikawa cells ... 72
Brightfield microscopy ... 72
Phalloidin staining of F-actin ... 72
Cell lysate ... 73
Bicinichoninic acid protein assay ... 74
Western blot for MUC-1 expression ... 74
EEV isolation and PKH26 staining ... 75
CyQUANT Proliferation assay ... 77
Normal cellular proliferation during hormonal treatment ... 78
Dose-response and time-point for rCD142 ... 78
Tyrosine-phosphorylation western blots ... 79
RESULTS ... 80
Culture of Ishikawa cells ... 80
Uptake of EEVs by Ishikawa cells ... 83
Cellular proliferation of Ishikawa cells in response to rCD142 treatment ... 84
Tyrosine-phosphorylation of Ishikawa cellular proteins ... 88
DISCUSSION ... 90
Ishikawa cells as a model of endometrial epithelial cells ... 90
Effects of rCD142 treatment of Ishikawa cells ... 91
Limitations and Future work ... 93

CHAPTER 4

PROTEOMICS ANALYSIS OF ISHIKAWA CELLS AFTER TREATMENT WITH RECOMBINANT CD142 PROTEIN 94

INTRODUCTION 95
Proteomics *95*
Types of proteomics *95*
Process of expression proteomics *95*
Tissue factor *97*
Aims *97*
MATERIALS AND METHODS 97
Treatment of cells with recombinant tissue factor *97*
Protein preparation and digestion for mass spectrometry *98*
QExactive method TDI -DIA *99*
MS data search *99*
Data analysis using DEP package *100*
Data analysis using outlier significance A/B *101*
Western blot validation of proteomics results *103*
RESULTS 104
Missing not at random DEP analysis *107*
Missing at random DEP analysis *112*
Significance A/B data analysis *118*
STRING analysis of differentially expressed proteins *119*
Western blot validation of differentially expressed protein *126*
DISCUSSION 127
Proteomics analysis of Ishikawa cells *127*
Differentially expressed proteins *128*
Limitations and Future work *132*

CHAPTER 5

PROXIMITY BARCODING ASSAY – PILOT STUDY 133

INTRODUCTION 134
Small Extracellular vesicles *134*
Challenges with traditional methods of characterisation *134*
Proximity barcoding assay (PBA) *135*
Aims *138*
MATERIALS AND METHODS 138
Sample collection *138*
Panel of antibodies *139*
Antibody-oligonucleotide conjugates (PBA probe) *140*
Rolling circle amplification (RCA) product preparation *140*
Small EV capture *140*
Proximity barcoding *141*
Amplification and sequencing *141*
Data analysis *141*
RESULTS 142
Morphological assessment of embryos *142*
PBA data quality control and normalisation *143*
Differentially expressed proteins *144*

Cluster and STRING analysis *147*
Top 25 protein combinations *166*
DISCUSSION 168
Differentially expressed proteins using PBA *169*
Clustering analysis *171*
Limitations and Future work *173*

CHAPTER 6

FINAL DISCUSSION **175**

IMPLANTATION 176
EXTRACELLULAR VESICLES IN IMPLANTATION 176
EMBRYONIC EXTRACELLULAR VESICLES 177
SUMMARY OF CHAPTERS 178
LIMITATIONS OF STUDY 181
FUTURE WORK 182

APPENDIX **186**

REFERENCES **195**

List of abbreviations

WHO – World Health Organisation

BMI – Body mass index

IUI – Intrauterine insemination

IVF – In vitro fertilisation

ICSI – Intracytoplasmic sperm injection

PGT – Preimplantation genetic testing

ERA – Endometrial receptivity array

IMSI – Intracytoplasmic morphologic sperm injection

eSET – Elective single embryo transfer

ICM – Inner cell mass

PGT-A – Preimplantation genetic testing for aneuploidy

PGT-M – Preimplantation genetic testing for monogenetic disease

niPGT-A – Non-invasive preimplantation genetic testing for aneuploidy

EVs – Extracellular vesicles

S-EVs – Small extracellular vesicles

M/L-EVs – Medium/Large extracellular vesicles

Apo-EVs – Apoptotic extracellular vesicles

ESCRT - Endosomal complex required for transport

MVB – Multivesicular bodies

ECs – Endometrial cells

PE – Preeclampsia

GDM – Gestational diabetes mellitus

EEVs – Embryonic extracellular vesicles

TEM – Transmission electron microscopy

NTA – Nanoparticle tracking analysis

SEM – Scanning electron microscopy

PIBF – Progesterone induced blocking factor

MISEV – Minimal information for studies of extracellular vesicles

PBA – Proximity barcoding assay

fPBS – filtered phosphate buffered saline

BCA – Bicinchoninic acid

BSA – Bovine serum albumin

SEC – Size exclusion chromatography

SOP – Standard operating protocol

ECACC – European Collection of Authenticated Cell Cultures

DMEM – Dulbecco's modified eagle medium

RIPA – Radioimmunoprecipitation assay

PVDF – Polyvinylidene fluoride

HRP – Horseradish peroxidase

PMTV – Photomultiplier tube voltage

MFI – Median fluorescent intensity

nf – Normalization factor

HUVEC - Human umbilical vein endothelial cells

TNFα – Tumour necrosis factor alpha

miRNA – Micro ribonucleic acid

ITGB1 – Integrin beta 1

VEGFR 2 – Vascular endothelial growth factor receptor 2

TF – Tissue factor

PAR2 – Protease-activated receptor-2

TSG101 – Tumour susceptibility gene 101

LIF – Leukaemia inhibitory factor

M-CSF – Macrophage colony stimulating factor 1

cAMP – Cyclic adenosine monophosphate

PFA – Paraformaldehyde

MUC-1 – Mucin 1

WGA – Wheat germ agglutinin

PKCα – Protein kinase C α

NF-kB – Nuclear factor kappa B

PDGFR – Platelet-derived growth factor receptor

DEPs – Differentially expressed proteins

rCD142 – Recombinant CD142

TEAB – Triethylammonium bicarbonate

TCEP – Tris-carboxy-ethyl-phosphine

IAA – Iodoacetamide

TFA – Trifluoroacetic acid

AN – Acetonitrile

FA – Formic acid

DMSO – Dimethyl sulfoxide

DIA – Data-independent mode

FDR – False discovery rate

MAR – Missing at random

MNAR – Missing not at random

kNN – k-nearest neighbour

MLE – Maximum likelihood estimation

Man – Manual

MinProb - Probabilistic minimum

QRILC – Quantile regression imputation of left-censored data

PCA – Principal component analysis

SD – Standard deviation

WIPF2 – WAS/WASL-interacting protein family member 2

Fe-S – Iron-sulfur

RT-qPCR – Real time quantitative polymerase chain reaction

RCA – Rolling circle amplification

DTT – Dithiothreitol

EDTA – Ethylenediaminetetraacetic acid

CTB – Cholera toxin subunit B

ml – Millilitre

µl – Microlitre

µg – Microgram

Chapter 1

Introduction – literature review

Infertility

The World Health Organisation (WHO) defines infertility as the failure to achieve a pregnancy after 12 months or more of regular unprotected vaginal intercourse and it affects 1 in 6 couples worldwide[1]. Infertility can be attributed to male factors (such as azoospermia), female factors (such as tubal blockage), both female and male factors, or idiopathic causes. With changes in lifestyle such as delaying parenthood (age is directly related to fecundity), increased alcohol consumption, and increased BMI, the rates of infertility are on the rise worldwide[2,3]. In females, the effects of age are more prominent with the number and quality of oocytes reducing over time with an increase in aneuploidy[4]. With the increased costs of diagnosis and treatment as well as the lack of access to treatment, the prevalence of infertility is widely underrepresented[5]. Due to the increase in infertility and the underrepresentation of the prevalence and complexity, it is becoming important to understand and provide better care for individuals suffering from this disease.

Treatment options

Infertility treatment options are available and widely accepted. Treatment options have advanced greatly and depending on diagnosis, treatments can be tailored to meet patient needs. Options include timed intercourse with medication, intra-uterine insemination (IUI), in-vitro fertilization (IVF) with intracytoplasmic sperm injection (IVF-ICSI) or without (conventional IVF), and pre-implantation genetic testing (PGT). IVF is the most common treatment option with over 2.5 million cycles being performed annually worldwide[6]. Even though there are a variety of treatment options available, the success rates are only at best 10-50% which is highly dependent on factors such as the age of the mother[7]. Infertility treatment is expensive with costs ranging on average from £4,500 to £7,000 per IVF cycle. The low success rates and high costs have led to the emergence of adjunctive (add-on) treatment options to increase the chances of pregnancy per

cycle. These add-ons include but are not limited to embryo-glue, endometrial scratching, time-lapse imaging, endometrial receptivity array (ERA) and artificial oocyte activation (Table 1.1)[8]. There is confounding evidence as to whether these add-on treatments increase the chances of a healthy pregnancy. Nevertheless, they are routinely offered to patients which can increase the costs of treatment further.

Table 1.1: Explanation of common IVF add-on treatments offered to patients and whether they are approved for use by the human fertilisation & embryology authority (HFEA).

Add-on Treatment	**Explanation**	**Evidence based support**
Embryo glue	Transfer medium enriched with hyaluronate, low levels of albumin, essential and non-essential amino acids to aid in implantation.	Conflicting evidence for effectiveness in improving chances of a healthy livebirth[9].
Endometrial scratching	Intentional damage to the endometrium to increase the chances of pregnancy.	Conflicting evidence for effectiveness in improving chances of a healthy livebirth[10,11].
Time-lapse systems	Incubating systems with in-built cameras to take digital images at frequent time intervals allowing for continuous culture.	Conflicting evidence for effectiveness in improving the chances of a healthy livebirth but more promising results are being reported with the use of deep machine learning[12].

Endometrial receptivity array	Taking a biopsy of the endometrium and checking gene expression to find the optimal time for implantation and embryo transfer.	No evidence to show it is effective in improving the chances of a healthy livebirth[13,14].
Artificial oocyte activation	Post ICSI, using calcium ionophores to create calcium oscillations in the oocyte and start fertilisation in cases of male infertility.	Conflicting evidence for effectiveness in improving chances of a healthy livebirth[15,16].
Intrauterine culture	Fertilised oocytes are placed in an intrauterine culture device that is placed inside the uterus for a few hours for the start of embryo development and then removed at which point the embryos are cultured in incubators until transfer.	No evidence to show it is effective in improving the chances of a healthy livebirth[17].
Intracytoplasmic morphologic sperm injection (IMSI)	Sperm is viewed using very high magnification to obtain detailed images of sperm.	No evidence to show it is effective in improving the chances of a healthy livebirth[18–20].

With the laboratory and medical advances made in the field, the aim of IVF treatment is to produce more than one embryo per cycle to allow for the option of further transfers if the first transfer is unsuccessful and/or to allow couples to have more than one child without the need for multiple

IVF cycles. Not until long ago, multiple embryos would be transferred to the uterus simultaneously to increase the chances of a pregnancy per transfer. However, this resulted in the emergence of significant numbers of multiple gestations following IVF treatment. Multiple gestations are considered as high risk for pregnancy complications such as hypertensive disorders, foetal malformations, spontaneous abortion, and preterm birth[21–24]. To avoid multiple gestations, clinicians have adopted the practice of elective single embryo transfer (eSET). In this method, a single embryo is chosen from multiple embryos for transfer into the uterus; this is different from non-selective single embryo transfer, where only one embryo is available for transfer[25,26]. It is, therefore, important to choose the embryo with the highest chance of resulting in a healthy pregnancy to transfer back into the uterus. Due to the importance of embryo selection, different methods have been used to select the embryo with the highest potential for implantation.

Embryo development in vitro

Embryonic development *in vitro* mimics the *in vivo* process with the difference that the embryo is cultured outside the uterus in culture media and incubators. The process starts when the sperm enters the oocyte triggering fertilisation which can be achieved in one of two ways: either through conventional IVF, where sperm cells are introduced to the culture media which contains the oocyte and incubated until a sperm naturally enters the oocyte. In the other method, a sperm cell is chosen and injected into the oocyte using a micromanipulation apparatus and high magnification microscopy (IVF with ICSI). Once the sperm enters the oolemma, the oocyte becomes activated following a surge of cytoplasmic calcium (Ca^{2+}) followed by Ca^{2+} oscillations[27]. If the oocyte is not activated, fertilization does not occur. The first visible sign of fertilisation occurs 16 to 18 hours after the sperm enters the cell with the appearance of the pronuclei. The embryo then starts dividing and growing from the fertilised oocyte all the way through to the blastocyst stage which

occurs between days 5 and 6 of development. At blastocyst, the embryo has differentiated into the inner cell mass (which forms the foetus) and the trophectoderm layer (which is on the periphery and forms the placenta) and is ready for implantation into the uterus[28]. In IVF, at the point of a blastocyst, the embryo is either discarded (if of poor quality), transferred to the uterus (if fresh transfer in the same cycle is chosen) or cryopreserved in liquid nitrogen until future use. The culture media in which the embryo has developed is routinely discarded once the fate of the embryo has been determined. Many factors play a role in the development of the embryo and can disrupt its growth. Therefore, the culture conditions and handling of the embryo are important to ensure optimal growth. There are also benefits associated with studying the culture media in which the embryo develops as it could provide insight into the different factors involved in the growth and the implantation of the embryo. Most practices have opted for the transfer of blastocysts back into the uterus because at this stage, good quality embryos have higher potential for implantation compared to earlier stages where the embryo may be morphologically of good quality but may fail to develop to the blastocyst stage if cultured further (low predictive value)[29]. There are still some practices that choose to transfer day 3 embryos to allow the embryo to grow inside the uterus with natural factors which may not be available *in vitro*. However, the option of which embryo and when to transfer is usually decided on an individual patient basis and is decided by the expertise of the embryologist and physician. It is, therefore, important to understand the interplay of the embryo and the uterus to help in deciding which embryo to transfer to increase the chances of a healthy pregnancy per transfer and reduce the overall time to pregnancy.

Implantation

Once an embryo has been chosen for transfer, it is injected into a primed uterus which is hormonally treated to mimic the natural hormonal cycle and to ensure optimal receptivity. The

optimal priming protocol is still under debate. However, it is usually achieved through the use of progesterone which is the prominent hormone of the secretory phase of the endometrium[30] (Figure 1.1).

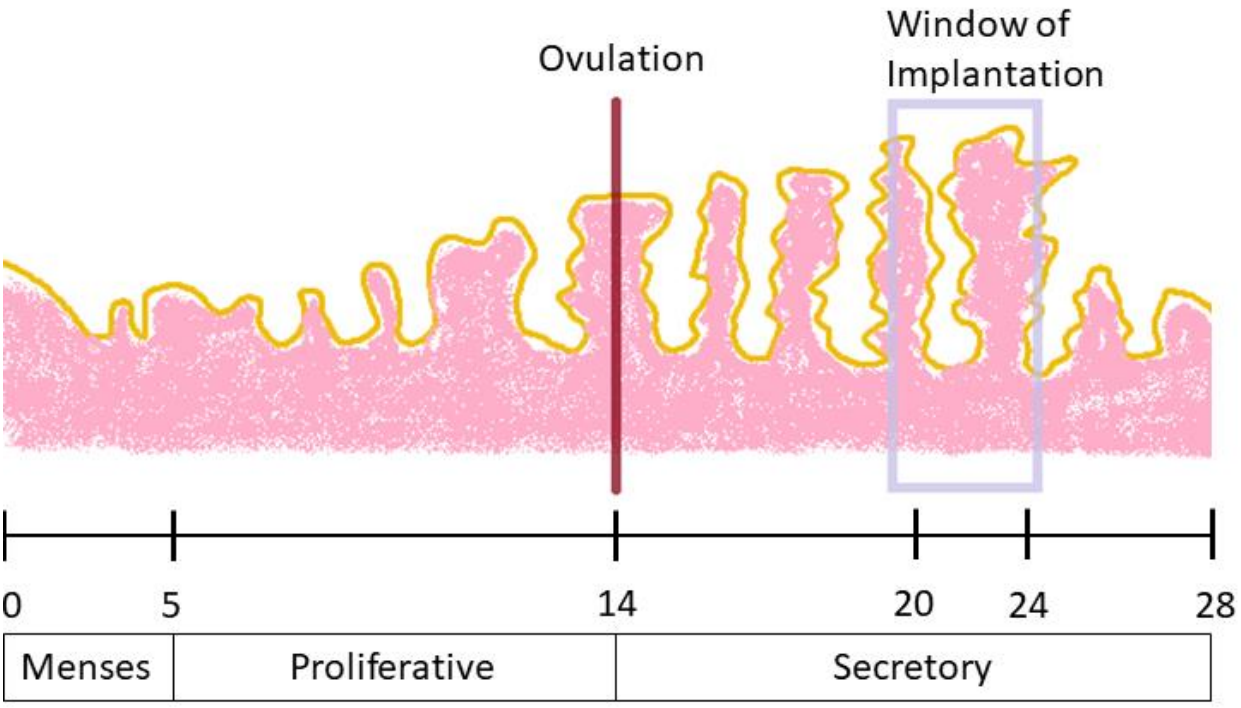

Figure 1.1: The menstrual cycle depicting the endometrium on a typical 28-day cycle. The proliferative phase is between days 5 and 14. The secretory phase is between days 14 and 28 with the window of implantation occurring between days 20 and 24 which is outlined. The figure is illustrated by the author.

There are multiple factors which play a role in the success of a transfer such as the receptiveness of the uterus, the quality of the embryo, and the interactions between the embryo and uterus. Research has shown that there is communication between the embryo and uterus prior to implantation through multiple factors such as cytokines, chemokines and growth factors[31]. The communication is important for immune regulation to allow for the implantation of the semi-allogenic embryo. The process of implantation begins when the embryo leaves its protective zona pellucida layer and attaches to the epithelial cells of the endometrium. The process of implantation consists of three steps: 1) apposition, where the embryo comes closer and orients itself with the

ICM facing the endometrium, 2) adhesion/attachment, where the embryo attaches to the luminal epithelial cells of the endometrium and 3) invasion, where the embryo penetrates through the endometrium to form the placenta and begin gestation[32] (Figure 1.2).

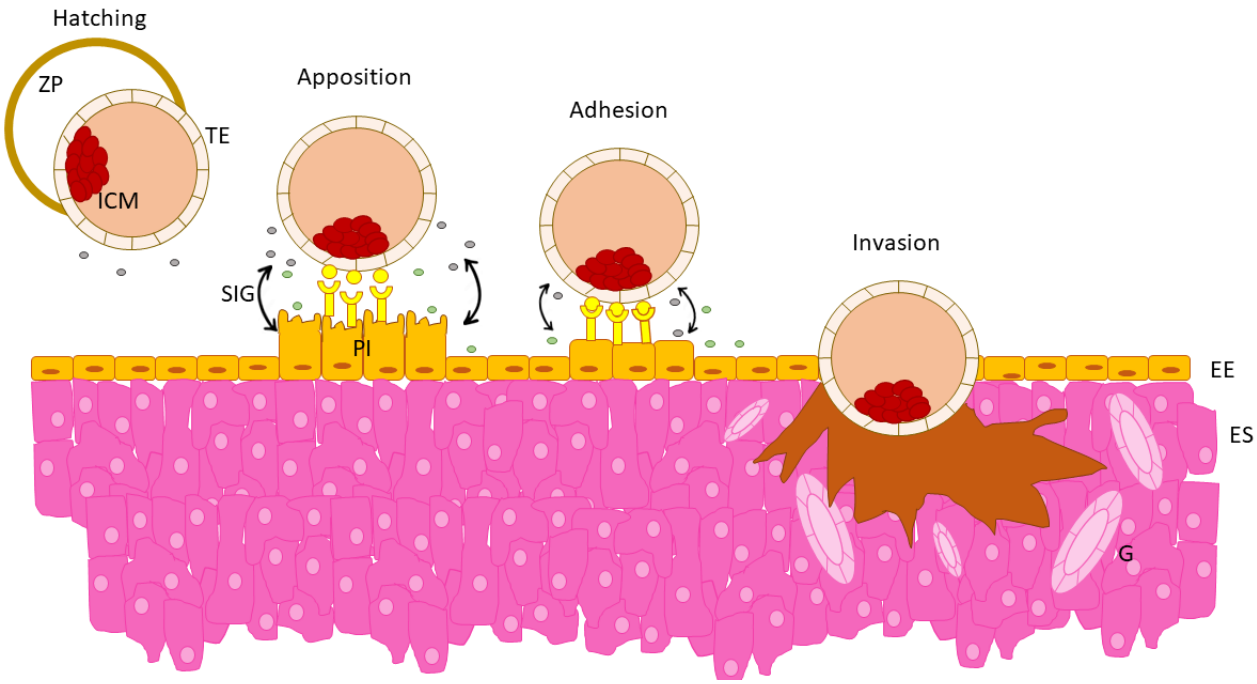

Figure 1.2: The process of implantation of an embryo into the endometrium. The embryo leaves the zona pellucida and undergoes apposition, adhesion and invasion. ZP = zona pellucida, ICM = inner cell mass, TE = trophectoderm, SIG = bidirectional signalling, PI = pinopods, EE = endometrial epithelial cells, ES = endometrial stromal cells, and G = glands. The figure is illustrated by the author.

Endometrial changes

For the endometrium to become receptive, it undergoes morphological and functional changes. During the proliferative phase of the cycle, oestrogen primes the endometrium which then becomes receptive to an embryo in response to the rise in progesterone (after ovulation). The endometrium is highly receptive during the window of implantation which occurs between day 20 to day 24 in a 28 day cycle[33,34]. The synchrony between the receptivity of the endometrium and the development of the embryo is therefore crucial for implantation. Different models of the

endometrium, such as primary endometrial cells and cell lines, have been used to study the process of implantation which will be discussed in more detail in chapter 3.

Morphological analysis of embryos

The most widely used method for embryo selection is the use of morphology-based criteria at different stages of development. These include cell count, fragmentation, pronuclei formation and grading of the trophectoderm and ICM at the blastocyst stage[35]. The morphological characteristics can be noted by an embryologist when the embryo is taken out of the incubator and analyzed under a microscope at specific time points. Typically, fertilisation is checked between 16 to 18 hours post IVF-ICSI and 24 hours after conventional IVF and is assessed by the presence and quality of the pronuclei. On day 3, the embryo is rechecked to ensure cleavage of the cells has occurred and the correct number of 8 healthy-looking cells (uniform in size with no/low levels of granulation and fragmentation) are present. On day 5 the embryo is checked to see whether a healthy-looking blastocyst has formed[36,37]. However, depending on the standard operating procedures in place for specific laboratories, these checks may be performed more or less frequently. Another way embryos are evaluated is by using time-lapse technology. With time-lapse, the embryo is always maintained in the incubator and constantly evaluated using imaging algorithms which are incorporated into the incubator and assesses the embryo based on morphokinetic parameters. Common criteria include the number, location and appearance of pronuclei, duration of first cytokinesis, level of fragmentation, blastomere movement, multinucleation, cell exclusion, compaction, and blastocyst collapse[38]. There are a variety of other parameters which can be added or deleted from the system at the embryologists' discretion. Due to the lack of standardization of which parameters should be included or excluded, there may be different criteria in different laboratories. Although the use of time-lapse technology is intriguing, there are several studies

which show no difference in fertilisation rate, embryo quality and ongoing pregnancy rates[39–41] compared to standard incubation and assessment. However, time-lapse does account for the subjectivity of the embryologist as well as allowing for a better understanding of embryonic behaviour throughout development rather than snapshot evaluations[42].

PGT and ni-PGT

Another method that has been used for embryo selection is preimplantation genetic testing for aneuploidy (PGT-A). With this method, a biopsy of a few cells (could be a single cell) is taken from the trophectoderm layer which is analyzed using next generation sequencing to determine chromosome abnormalities[43]. The embryos which are euploid are then chosen for transfer. This method increases the chances of implantation per transfer for patients with advanced maternal age and suffering from recurrent miscarriages. With PGT-A, an embryo may be identified as mosaic where some cells of the embryo are euploid and others are aneuploid. With mosaicism, there is still controversy as to whether these embryos should be transferred and it becomes more of a dilemma when a poor prognosis patient only has one mosaic embryo[44]. Mosaic/abnormal embryos may also be missed if the portion of the biopsy is taken from the euploid cells of the trophectoderm. The inner cell mass of the embryo which forms the fetus, may also have a different composition of cells compared to the trophectoderm.

Besides PGT-A, monogenetic disorders can also be identified using PGT (PGT-M). The main benefit of PGT-M is that a non-effected embryo is chosen for transfer. For specific monogenetic disorders such as cystic fibrosis and haemophilia, this method of embryo selection is highly recommended. However, there is the possibility that the viability of the embryo could be compromised due to the biopsy[45]. High technical abilities of the embryologist as well as

complicated equipment are required and the procedure is costly for the patients. Therefore, more non-invasive methods are being explored as possible indicators of embryonic health.

Due to the expenses associated with IVF, the move towards eSET, and current selection criteria, the use of non-invasive techniques could potentially aid in increasing the chances of a healthy pregnancy. For example, with non-invasive PGT-A (niPGT-A), the surrounding culture media from an IVF embryo is tested without the need for a biopsy of the trophectoderm. Although this is a promising avenue for research, there are complications such as the low quality and quantity of DNA found in the spent culture media[46]. There has also been a focus on exploring other non-invasive methods of determining the implantation potential of embryos such as examining extracellular vesicles (EVs) released by embryos.

Extracellular vesicles

EVs are small lipid membrane bound particles which are released by cells and act as autocrine and paracrine communicators in different systems. They are classified based on their biogenesis pathway and size. Small EVs (S-EVs) are smaller than 200nm, medium/large EVs (M/L-EVs) are classified as larger than 200nm and smaller than 1000nm. EVs between 1000nm and 5000nm are classified as apoptotic bodies (Apo-EVs)[47]. This nomenclature was recommended for classification by the international society of extracellular vesicles, but some studies use the old nomenclature and refer to small EVs as exosomes and medium/large EVs as microvesicles.

Small EVs (S-EVs)

The biogenesis pathway of S-EVs comprises of the endosomal pathway, where early endosomes form into multivesicular bodies which are released as S-EVs through the plasma membrane of the cell[48,49]. There are two common pathways for S-EV biogenesis: endosomal complex required for transport (ESCRT) dependent and ESCRT independent. In both processes, endocytic vesicles

merge with intracellular early endosomes forming the late endosome or multivesicular body (MVB).

In the dependent pathway, the membrane of the MVB is organized into functional membrane microdomains that are enriched for tetraspanins. Four ESCRT complexes are involved in the process. ESCRT 0 and I cluster cargos on the microdomains and recruit ESCRT II and III. ESCRT I and II are involved in the budding of the membrane. ESCRT III also aid in the budding of the membrane through interactions with ALIX[50]. The independent pathway uses alternate pathways which involve sphingomyelinase, ceramide, cholesterol, and tetraspanins[51,52]. In this process, membrane subdomains are created and tetraspanins are involved. However, ALIX and the ESCRT complexes are not involved (Figure 1.3).

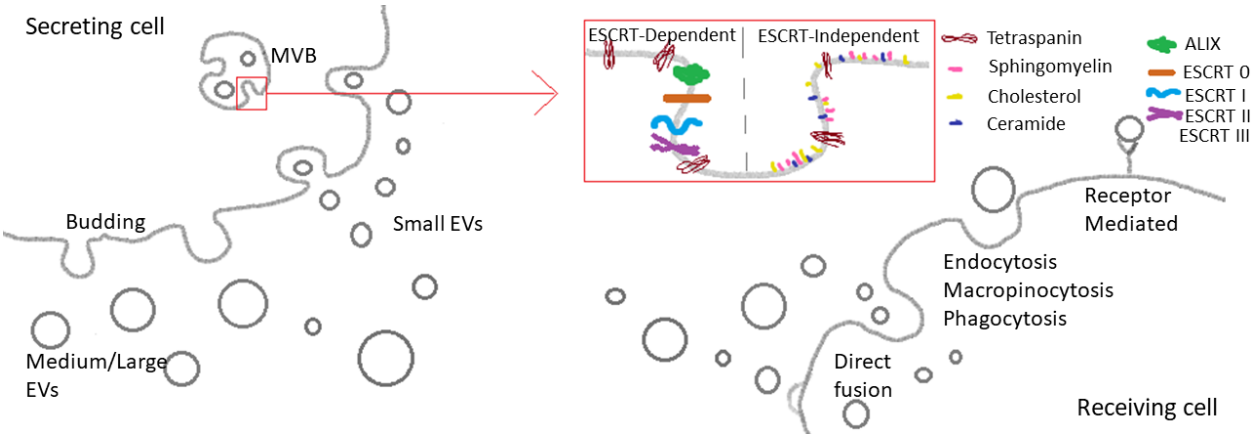

Figure 1.3: The process of medium/large and small EV production. The small EV pathway is through the multivesicular body (MVB) and the ESCRT-dependent and ESCRT-independent production are outlined in the red box. The medium/large EVs are produced through the budding of the membrane. The uptake of the EVs through direct fusion, endocytosis, macropinocytosis, phagocytosis and receptor mediated signalling are shown on the bottom right in the receiving cell. The figure is illustrated by the author.

Medium/large EVs (M/L-EVs)

M/L-EVs are released through direct outward budding of the plasma membrane. Ca^{2+} influxes into the cells signal enzymes such as flippase to translocate phosphatidylserine to the outer membrane. Exposure of phosphatidylserine causes the outward budding of the membrane and contraction of

cytoskeletal structures through the interactions of actin with myosin. The release of microvesicles occurs through the protrusion of the membrane, followed by budding and detachment[48].

Apoptotic bodies (Apo-EVs)

Apo-EVs are large vesicles which are formed when cells are undergoing programmed cell death. These vesicles contain a variety of bioactive molecules and cellular organelles[53]. Their production is apoptosis dependent and can transduce signals to other cells. In some cases, where inflammatory or immune responses are not to be activated, such as during implantation, apoptotic cells and Apo-EVs are phagocytosed by resident macrophages[54].

Cargo of EVs

The cargo of EVs includes proteins (e.g. antigen presentation, membrane transport, cytosolic, transmembrane), lipids (e.g. cholesterol, phosphatidylserine), metabolites, and nucleic acids (e.g. mRNA, noncoding RNA)[55]. EVs can alter the phenotype of the targeted cell by protecting and transferring cargo on their surface and within their membrane. EVs have been identified in a variety of different body fluids and have been shown to play a role in cellular communication and disease progression. EVs have also been found to be released by the endometrium and affect implantation through signalling to the embryo. For example, in endometrial cells (ECs) treated with oestrogen and progesterone, the small EVs released by ECs were taken up by trophoblast cells and promoted the adhesive capacity by activating focal adhesion kinase signalling pathways[56]. The release of EVs by embryos has been shown as well which could be involved in implantation.

Uptake of EVs by recipient cells

Once the EVs are packaged and secreted, they are taken up by recipient cells through different mechanisms (Figure 1.3). Some of these include receptor/ligand fusion, direct fusion, endocytosis,

phagocytosis, and macropinocytosis. In receptor/ligand mediated signalling, the proteins on the surface of the EV bind their corresponding receptor on the recipient cell and initiate intracellular signalling. For example, dendritic cells release S-EVs which contain ICAM-1 as surface proteins, which bind its receptor (LFA-1) on activated T cells[57]. In direct fusion, the plasma membrane of the EV fuses with the recipient cell, releasing its content directly into the cytoplasm. Endocytosis comprises of the engulfment of the entire EV with its intact membrane into the recipient cell. Clathrin and lipid rafts can be involved in endocytosis[58,59]. The process of phagocytosis involves enzymes such as phospholipase C and phosphatidylinositol-3-kinase[60,61]. Macropinocytosis uses proton pumps or sodium/proton exchangers as well as actin-driven lamellipodia which causes the plasma membrane to invaginate and form a macropinosome[62,63] which fuses further with a lysosome and undergoes processing. The method of uptake by the recipient cell varies depending on factors such as the proteins involved and the signalling pathways.

Extracellular vesicles in pregnancy

EVs have been characterised in many biological fluids such as urine, blood, saliva, breast milk, and cerebrospinal fluid. They have also been shown to play roles in specific processes during pregnancy, from fertilization all the way through gestation. They are well characterised in immune tolerance as immune modulators. EVs are found to be released by ovarian follicles, seminal fluid, endometrium and placental cells, specifically syncytiotrophoblast. EVs have been shown to be involved in inflammation, vascular dysfunction, and metabolic disorders contributing to preeclampsia (PE) and gestational diabetes (GDM). In PE, the syncytiotrophoblast releases more EVs into the maternal circulation compared to normal pregnancy which varies in the cargo to induce a pro-inflammatory, anti-angiogenic, and pro-coagulant response leading to inflammation, endothelial dysfunction, and coagulation[64,65]. Since EVs have been shown to be involved in

pregnancy and related to adverse outcomes, it is important to delineate their potential role in the implantation of the embryo into the uterus. Embryonic EVs (EEVs) have also been identified in the culture media from IVF embryos.

Embryonic extracellular vesicles (EEVs)

Embryonic extracellular vesicles (EEVs) have been shown to be present in culture media, from IVF embryos, throughout development. These EEVs have been identified through the processing of culture media from both animal and human embryos. The presence of EEVs in human embryo culture media was confirmed by Giacomini et al[66]. The discovery of these EEVs implied their potential as markers of embryonic health and potential communicators to prime the endometrium before the embryo starts the process of implantation. This confirms that the process of implantation begins before the attachment of the embryo and could be affected through communication via EEVs. There has been significant interest in understanding the role of EEVs in implantation and early pregnancy through signalling via their cargo and surface markers. The presence of EEVs has been confirmed by multiple studies and a reduced amount has been linked to embryos that implant[67]. The transcriptomics of spent culture media has also been studied. There is however, a lack of characterisation of the EEVs. Table 2 summarizes some of the current findings on EVs released from human and animal embryos and their potential significance in pregnancy.

Table 1.2: Summary of studies looking at extracellular vesicles released by human and animal embryos. The characterisation methods and significant findings are outlined.

Authors	Species	Characteristics of EVs	Significance
Simon et al, 2020 [68]	Murine	Using TEM and NTA confirmed the presence of EEVs between 20 and 500nm with a concentration	EVs are present in the blastocoel and culture media crossing the zona pellucida

		of $1.74x10^7 \pm 2.60x10^6$ particles/mL.	from the murine model with the possibility of detecting specific DNA sequences in EVs.
Dissanayake at al., 2020 [69]	Bovine	Using TEM, SEM and NTA confirmed the presence of nearly round shaped EEVs between 50 and 150nm with a concentration of $7.17x10^8$ particles/mL for degenerating embryos and $5.68x10^8$ particles/mL for normal quality embryos.	EEV concentration is modulated by the stage of development and quality of embryos which could be used in addition to morphological parameters for embryo selection.
Pavani et al., 2019 [70]	Bovine	Using TEM, NTA and western blot confirmed the presence of EEVs between 25 to 230nm with a concentration of $236.5\pm1.27x10^8$ particles/mL with the expression of CD63 and CD9.	Pre-labelled EEVs were taken up by zona-intact embryos which increased the blastocyst rates and decreased apoptotic cells suggesting that EEVs play a role in inter embryo communication between bovine embryos.
Pallinger et al., 2018 [71]	Murine	EEVs were identified with TEM to be between 100 to 200nm in size. The EEVs also expressed	EEVs contain PIBF which may be partly responsible for the increase in IL-10 production of CD8+ T cells, confirming the

		progesterone induced blocking factor (PIBF).	possible communication between embryo and endometrium prior to implantation.
Mellisho et al., 2017 [72]	Bovine	They identified EEVs using TEM (30 to 385nm), flow cytometry (CD63 and CD9 positive) and NTA.	The distribution and characteristics of EEVs varied depending on competence and were more highly variable in the non-competent embryos.
Vyas et al., 2019 [73]	Human	Identified EEVs between 30 to 500nm using TEM, SEM, and NTA.	Human embryos release EEVs into the perivitelline space which cross the zona pellucida into culture media.
Battaglia et al., 2019 [74]	Human	Using SEM, NTA, TEM and ELISA identified the presence of EEVs positive for CD63 with an average size of 75±3nm.	Small EEVs are present in the blastocoel fluid of embryos.
Pallinger et al., 2017 [67]	Human	Identified EEVs using propidium iodide (PI) staining of the nucleic acid content of EEVs.	The number of PI+ EEVs is lower in competent embryos and the embryo with the lowest PI+ EEV count should be transferred among all embryos from the same mother.

Abu-Halima et al., 2017 [75]	Human	Microarray analysis of miRNAs in spent culture media and NTA analysis of the EEVs.	The total number of EEVs from positive outcome embryos was 1.93-fold less than failed embryos. There was also a decrease in the miRNA repertoire of samples leading to pregnancy.
Giacomini et al., 2017 [66]	Human	Identified EEVs from day 3 and day 5 embryos with a size range between 50-200nm using TEM and NTA. Western lot analysis showed expression of CD9, CD63 and ALIX.	There was an increase in the number of EEVs from day 5 embryos compared to day 3. They also examined the mRNA content of 4 pluripotency related genes which showed similar expression between day 3 and day 5. EEVs uptake by primary endometrial cells was measured through fluorescence showing similar percent positive cells treated with day 3 and day 5 EEVs with different intensities and subcellular distribution of the fluorescence.

It is apparent that there is a lack of in-depth analysis of EEVs and their protein composition, specifically EEVs derived from human embryos. Most of the current research has focused on proving the presence of EEVs in culture media and the amount that is released rather than the characterisation of protein composition. The lack of analysis is likely due to the low access to material, low volume of sample and therefore low EV content. The cross-talk between the endometrium and embryo through EVs could be the first form of communication each of us has had with our mothers and understanding this communication in more detail will give us insight into the initial steps of pregnancy. This thesis is focused on characterising EEVs using new technologies to identify their role in implantation and pregnancy using a cellular model of the endometrial epithelial cells.

Characterisation of EEVs using new technologies

The field of EVs has grown in the past decade. The initial step in any EV study is to characterise the EVs to first confirm their presence and to further categorise them to investigate their roles in the system under study. There are traditional methods of EV characterisation which are all outlined in the MISEV guidelines published in 2018[47]. Most of the conventional methods rely on the high concentration of EVs with a higher starting volume compared to our embryonic samples. With samples similar to ours, which have a low starting volume and therefore the low concentration of EVs, besides the original characterisation methods, newer techniques have been developed to combat the difficulties in the purification and characterisation associated with conventional methods. These include the use of flow cytometry methods designed for nanoparticles (seen with NanoFCM), the use of capturing and counting EVs based on specific epitopes (seen in ExoCounter) which are outlined in chapter 2 and finally, the use of proximity barcoding assay which uses antibody-DNA conjugates and next generation sequencing to identify proteins on the

surface of individual S-EVs (discussed further in chapter 5). These new methods provide new avenues for identifying potential markers on the surface of EVs which can be further studied either *in vivo* or *in vitro*.

Aims of thesis

There are four main aims this thesis will address:

1) The aim of chapter two was to isolate and characterise EEVs generally by the use of nanoparticle tracking analysis, transmission electron microscopy and western blot. Furthermore, I aimed to identify a deeper phenotype of the EEVs based on their surface markers using MACSPlex, NanoFCM, and ExoCounter to identify putative targets for subsequent investigations.

2) Once the targets were identified, I chose the most commonly expressed surface marker (CD142/tissue factor) to study its effects on Ishikawa cells (used as a model for endometrial epithelial cells) in chapter 3.

3) The next aim was to delineate the proteomics changes (outlined in chapter 4) imposed on Ishikawa cells after treatment with recombinant CD142 protein.

4) Lastly, I wanted to use a more global characterisation technique to identify other surface proteins which could be found on EEVs using proximity barcoding assay (PBA), explained in chapter 5. PBA was used as a pilot study to delineate its potential for future experiments.

Chapter 2

Characterisation of human embryonic extracellular vesicles

Introduction

Extracellular vesicles

Extracellular vesicles (EVs) have been shown to be released by almost all cells and have been linked to disease development and progression. They have been shown to play roles in adverse pregnancy outcomes such as gestational diabetes (GDM), preeclampsia (PE) and intrauterine growth restriction. For example, in PE, the EVs driven from the syncytiotrophoblast have anti-angiogenic properties such as reduced endothelial nitric oxide synthase activity which contributes to vascular dysfunction and hypertension[76]. It has been proven that all throughout the female reproductive tract, EVs are released from cells as forms of inter- and intra-cellular communication[77,78]. The release and uptake of EVs is a novel form of communication which allows for the transfer of materials such as proteins, mRNAs, miRNAs, and lipids[79]. The role of EVs in specific biological processes can be determined through the characterisation of their surface proteins and inner cargo.

Embryonic EVs (EEVs)

Since EVs have been shown to be released by most cell types, it is not surprising that their release by embryos throughout development has also been reported. The presence of the EEVs, in culture media from IVF has been shown by multiple studies, however the in-depth characterisation has not been reported. The presence of EEVs in culture media from human embryos has been confirmed through microscopy, western blots and flow-cytometry based assays[80–85]. Both scanning and transmission electron microscopy have reported the presence of EEVs in size ranges of small and medium/large EVs (S-EVs, M/L-EVs)[81,84]. The zona pellucida is a hard glycoprotein membrane which encloses the embryo prior to implantation, and it is important for the EEVs to cross the zona pellucida if they are to affect the endometrium. The crossing of the EEVs through

the zona pellucida was demonstrated by Vyas et al., where they labelled EEVs with CD9 (a marker of EVs) and showed their presence on the plasma membrane, in the perivitelline space and through the zona pellucida[84]. Therefore, it is clear that EEVs are present and could be relaying information to the endometrium.

It has been shown that as embryonic development progresses from day 3 to day 5, the number of EEVs increases, suggesting that as the window of implantation approaches, the embryo is communicating more with the endometrium[81]. Therefore, the amount and cargo of EEVs could play a role in determining implantation success. For these EEVs to have an impact on the endometrial cells, they are required to enter the cells. Primary endometrial cells showed an uptake of EEVs with increased uptake on day 5 of development, suggesting that at this stage, the EEVs are more likely to enter the cells and modulate physiological processes[81]. Another study compared the number of EEVs from pooled samples of embryos that resulted in clinical pregnancy to pooled samples from embryos that resulted in implantation failure[82]. They revealed that those embryos which successfully implanted had a lower amount of EEVs compared to the implantation failure group[82]. This was confirmed by another group which showed that the EEV population in positive pregnancy outcomes was 50% less than those in negative outcomes[75]. The increase in EEVs could be explained by the response mechanism of the embryo where under stress conditions, more EEVs are produced.

Other factors in the secretome of embryos

Besides the presence of EEVs, studies have investigated the secretome of embryos by looking at the culture media surrounding embryos and identifying differentially expressed miRNAs. Culture media from aneuploid embryos that resulted in failed implantation showed a higher expression of miR-191 which is involved in trophoblast invasion and proliferation[86]. Blastocoel fluid from

embryos showed expression of 89 miRNAs which were specific to blastocoel fluid compared to controls[85]. These miRNAs are involved in processes such as embryo development, cell adhesion and intercellular communication. These studies confirm the communication from embryos and their potential role in implantation. Since EVs carry miRNAs as cargo, these signals may be coming from EEVs or being exacerbated by the EEVs. Even though the presence of EEVs and non-coding RNAs have been proven in embryonic samples, complete characterisation of EEV proteins has not been achieved.

Aims

The aim of this chapter is to **1)** identify and characterise EEVs using NTA, TEM and western blot **2)** to characterise the surface proteins commonly found on EEVs using MACSPlex, ExoCounter and NanoFCM (nFCM) and **3)** to identify putative targets for further exploration. Understanding the common protein markers found on the surface of EEVs could be utilised to develop non-invasive methods of identifying embryos with the best potential for implantation.

Materials and methods

Ethics

Ethical approval for this project falls under "Characterisation of extracellular vesicles released by human embryos" granted by the Oxford Ethics Committee "Medical Sciences Interdivisional Research Ethics Committee" with approval reference: R81004/RE001.

Sample collection

Samples were obtained from the IVF laboratory at IVI London. Human embryos were cultured in time-lapse Petrie dishes (using G-TL^{TM} media) after IVF-ICSI was performed up to the blastocyst stage of development. Once the fate of the embryo was decided, the culture media was collected and frozen at -80°C for storage and transportation. Due to the nature of the culturing dish, samples

were obtained from the wells (8µL) and surrounding reservoir (100µL) of each dish and pooled. The same batch of culture media was incubated in the same conditions as the embryonic culture media before collection as the negative control and stored at -80°C.

EV isolation for NTA

The samples from IVI London were used to isolate EEVs using size exclusion chromatography (SEC) with qEV original 35nm columns (SP5, IZON Science) following the manufacturer's instructions. The pooled sample was spun at 1,500 x *g* for 10 minutes, followed by 10,000 x *g* for 10 minutes. The column was flushed with 0.1µm filtered phosphate buffered saline (fPBS) prior to use. The starting volume of the sample was 800µL, from which 500µL was loaded onto the top of the column. The sample was run through and the void fraction of 3mL was collected immediately. The following 1.5mL (the EV rich fraction) was collected for analysis. The column was flushed using fPBS and the remaining 300µL was added onto the frit. The EV fraction was collected and combined with the EV fraction from the first run which was concentrated using the Amicon 100kDa filter (UFC810024, Millipore) to a final volume of 100µL. The same process was completed for the control culture media. The protein concentration of each sample was determined using the Bicinchoninic acid protein kit (Pierce BCA kit) as outlined below.

Bicinchoninic acid protein assay

Protein content was measured using the Bicinchoninic acid protein assay kit (BCA). Bovine serum albumin (BSA) stock (1mg/mL) (Sigma-Aldrich) was diluted in fPBS to create protein concentration standards (1mg/mL, 0.5mg/mL, 0.25mg/mL, 0.125mg/mL, and 0.0625mg/mL). Samples were diluted 1 in 10 in fPBS and 12.5µL of BSA standards or samples were added to a 96 well plate. All standards and samples were loaded in duplicates. To this, 100µL of BCA reagent which consists of BCA solution (23223, Thermo Scientific) and copper II sulphate solution (50:1

ratio) (23224, Thermo Scientific) was added to each well. The plate was incubated for 30 minutes at 37°C with gentle shaking. After incubation, the absorbance was read at 562nm using a FLUOstar OPTIMA plate reader. Proteins in the sample reduce Cu^{2+} to Cu^{+} in an alkaline solution, resulting in a purple colour, which is read by a spectrophotometer to extrapolate protein concentration based on protein standards.

Nanoparticle tracking analysis (NTA)

The Zetaview NTA (BASIC PMX-120, Particle Metrix) was used to determine the size distribution and particle concentration of the samples which is achieved through analysing short videos of random movement of particles in a liquid as they collide with the fast-moving particles of the liquid. It uses the displacement and the Stoke-Einstein equation to calculate the particle size and concentration. Before analysing samples, 100nm polystyrene beads were used for quality control and to check instrument performance. The beads solution was prepared by diluting the standard 1 in 1000 (using deionized water) from which 10µL was added to 2.5mL of deionized water. The solution was mixed with gentle vortexing and pipetting. This quality control sample was then loaded into the injection chamber and an auto alignment was performed by the software.

The chamber was flushed with 10mL fPBS prior to loading the first sample and in between samples. Two pooled samples and corresponding control culture media, which had undergone EV isolation using SEC, as mentioned above, were diluted using fPBS (1:20) to a final volume of 1mL and run on the ZetaView particle tracking analyser. When loading the samples, care was taken to ensure no air bubbles entered the chamber. The EV SOP was used which measures particle number and size at a shutter speed of 100 and camera sensitivity of 80 with 11 camera positions for 2 cycles. The settings were kept constant for all samples and controls. The results were presented by

the ZetaView software (version 8.05.12 SP2). However, the control culture media did not have enough particles for data to be generated.

Transmission electron microscopy

TEM imaging was completed by Shanghai GeneChem Co., Ltd. The small EEVs were separated from the collected samples using the exosome extraction kit (EXOQ5A-1, System Biosciences) according to the manufacturer's instructions. In brief, the pooled embryonic sample was centrifuged at 3,000 x *g* for 15 minutes to remove cellular debris. The supernatant was transferred to a microcentrifuge tube and 63µL of ExoQuick Exosome Precipitation solution was added. The tubes were stored in the fridge overnight with no rotation or mixing. After incubation, the sample was centrifuged at 1,500 x *g* for 30 minutes at 4°C. The S-EV precipitate was obtained after centrifugation and resuspended in 100µL PBS buffer. From the EV suspension, 5 to 10µL was allowed to drip in front of the copper carrier net and dried after 1 minute using filter paper. The staining reagent (uranyl acetate) was added for 10 to 40 seconds. The stain was then blotted off and air-dried completely. The grid was then ready for imaging using FEI Tecnai 12 TEM at 120kV with a Gatan OneView CMOS camera.

HEK293 cell culture and lysate preparation as control for western blot

HEK293 cells were obtained from Dr Wei Zhang (NDWRH, University of Oxford) and were authenticated by the European Collection of Authenticated Cell Cultures (ECACC). The cells were thawed at 37°C and cultured in T-75 flasks with Dulbecco's modified eagle medium (DMEM) supplemented with 10% fetal calf serum (FCS), 1% penicillin/streptomycin, 1% antibiotic-antimycotic, and 2mM L-glutamine at 37°C, 20% O_2, and 5% CO_2. Media was changed every 2 days until approximately 80% confluency. To produce cell lysates, the cells were first washed 3 times with PBS. Trypsin (Sigma Aldrich) was added to the cells and incubated for 5 minutes. Once

the cells had detached, complete media was added to neutralise the trypsin. The cells were transferred to a falcon tube and spun at 1,300 x *g* for 3 minutes. The supernatant was discarded and the cells were resuspended in 10mL of PBS and spun at 1,300 x *g* for 3 minutes. The supernatant was discarded, and radioimmunoprecipitation assay (RIPA) buffer, containing a cocktail of protease inhibitors, was added to the cells. The cells were incubated on ice for 15 minutes and then transferred to microcentrifuge tubes and mechanically dissociated using 21-gauge needle, followed by centrifugation at 15,000 x *g* for 15 minutes at 4°C. The supernatant was removed and placed in a new tube (this was the cell lysate). BCA was performed to determine the concentration of the cell lysates.

Western blot analysis

Prior to western blot analysis, pooled embryonic and control culture media were centrifuged at 150,000 x *g* for 2 hours at 4°C (Beckman Optima XPN ultracentrifuge and SW 55 Ti swing-out rotor); the supernatant was removed and the pellet was washed in fPBS at 150,000 x *g* for 2 hours at 4°C. The supernatant was removed and the pellet was resuspended in 15µL of fPBS. Due to the low volume and low protein concentration of the samples, the entire samples were used for western blot analysis using CD63. The culture media controls (n=2), embryonic samples (n=2), and positive control (HEK293 cells) were run in non-reducing conditions. The entire volume of the samples and control culture media was loaded with 30µg of HEK293 cells as the positive control. The samples were mixed with 5µL of 4X Laemmli buffer which was further diluted to 25µL using double distilled water. The samples were heated at 95°C for 10 minutes and 20µL of each sample was loaded into the gel (NuPAGE 4-12% Bis-Tris Gel 1.0mm x 10 well gel by Invitrogen). Precision plus proteins dual colour standards was used as a marker for protein size (Bio-Rad). Electrophoresis was performed under non-reducing conditions at 150 V for 1 hour in a Novex

minicell tank using NuPAGE MOPS SDS running buffer (Novex by Life Technologies). The proteins were then transferred to a polyvinylidene fluoride (PVDF) membrane (Bio-Rad). The transfer apparatus is described below. Four filter paper sheets pre-soaked in buffer solution (300mM Tris, 10% methanol, pH 10.4) were placed in a Novex semi-dry transfer apparatus. Two filter paper sheets pre-soaked in 25mM Tris, 20% methanol, pH 10.4 were placed on top. The gel was transferred on top of the filter papers, followed by pre-soaked (100% methanol) PVDF membrane. On top of this, three filter paper sheets pre-soaked in 25mM Tris, 40mM ε-aminohexanoic acid, and 20% methanol were placed. The transfer was run at 25 V for 45 minutes. The membrane was then blocked using 5% milk (diluted in 0.1% Tris-buffered saline and 1% Tween-20) for 1 hour followed by incubation with monoclonal CD63 mouse antibody (1:1000 dilution, Santa Cruz Biotechnology) overnight at 4°C with constant shaking. The following day, the membrane was washed with TBST, three times for 5 minutes each and then incubated for 1 hour with anti-mouse immunoglobulin horseradish peroxidase (HRP-linked) secondary antibody (1:2000 dilution, Dako UK Ltd). The membrane was washed with TBST, as mentioned previously, followed by treatment with an EZ-ECL chemiluminescence detection kit (Thermo Scientific) for 1 minute. The membrane was then developed using a gel documentation system (G-Box, Syngene) with the GeneSys software (version 1.6.7.0).

MACSPlex sample processing

The neat samples were thawed and pooled to a total of 400µL. The samples were then spun at 400 x *g* for 10 minutes, the supernatant was removed and spun at 2,000 x *g* for 10 minutes at 4°C. The supernatant was used directly for subsequent MACSPlex experiments. The control culture media was processed in the same way.

MACSPlex kit

The kit is a sandwich-based assay able to identify 37 common S-EV surface markers along with 2 isotype controls (REA and mIgG). 120µL of the processed samples was placed into a 1.5mL microcentrifuge tube and 120µL of the buffer from the kit was added to another tube and used as the control for the background signal. The capture beads provided in the kit were vortexed for 30 seconds and 15µL was added to each tube. The samples were incubated with constant rotation at room temperature overnight in the dark. The next day, 500µL of buffer was added to each tube and spun at 3,000 x *g* for 5 minutes. Subsequently, 500µL of the supernatant was removed. The detection beads (CD63, CD9 and CD81) were added (5µL of each). The sample was incubated for 1 hour at room temperature in the dark with constant rotation, followed by the addition of 500µL of the buffer. The tubes were spun at 3,000 x *g* for 5 minutes. 500µL was aspirated and replaced with 500µL of the buffer. The tubes were incubated at room temperature for 15 minutes with constant rotation and were then spun at 3,000 x *g* for 5 minutes and 500µL was aspirated from the tubes. The remainder of the pellet was resuspended by pipetting and kept in the dark until used for flow cytometry analysis.

Flow cytometry for MACSPlex

The Becton Dickinson LSRII flow cytometer (BD Biosciences) equipped with blue (488nm) and red (633nm) laser was used and calibrated using the CS&T beads before running the samples. The set-up beads provided in the MACSPlex kit were used to determine the photomultiplier tube voltage (PMTV) for each channel according to the manufacturer's instructions. The same set-up was maintained for all of the samples and controls tested. The SSC-A vs FSC-A was used to gate the singlet bead population, the Alexa-fluor 488 vs. PE plot was used to gate the individual capture

bead populations and the median fluorescent intensity (MFI) for each bead population was calculated from the Alexa-fluor 647 vs PE plot.

Analysis of relative quantification of small EV markers

Data were analysed using the FACSDiva software (BD Biosciences) and FlowJO (Tree Star Inc., Ashland, OR, version 10.8.0). The normalization factor (nf) was calculated as the mean MFI of the capture beads (CD9, CD81 and CD63). The signals of all the beads were divided by the nf of the specific sample. The nf was set at 1, and the signals from the sample were presented as relative to the nf. This was repeated for each of the samples and controls. The 4 common markers were identified and the mean and standard error of the mean (SEM) were calculated. An unpaired t-test with Welch's correction was performed to determine significance.

HUVEC cells culture media and MACSPlex

Human umbilical vein endothelial cells (HUVEC) (C2519A, Lonza) were cultured in EGM2 culture media (CC-3162, Lonza). HUVECs from passage 4 were seeded into a 24-well-plate at $1x10^4$ cells/well and cultured overnight at 37°C, 20% O_2, and 5% CO_2. The following day, culture media was replaced with EGM2 media supplemented with 10ng/mL of recombinant human TNFα (#78068.1, STEMCELL) for the treatment wells and EGM2 media supplemented with H_2O (17-724Q, water for injection, Lonza) for the control wells. The cells were cultured for a further 24 hours after which the culture media was removed and centrifuged at 1,500 x *g* for 15 minutes. The supernatant was removed and centrifuged further at 1,500 x *g* for another 15 minutes to remove all cellular debris. Following this, the supernatant was removed, placed in a new tube and frozen at -80°C. The culture media were thawed and processed in the same way as the embryonic culture media for MACSPlex. The HUVEC culture media from the control cells and the TNFα treated cells were analysed using the MACSPlex kit as mentioned above.

ExoCounter

To prove the presence of the markers identified with MACSPlex, samples were run on the ExoCounter to quantify the number of S-EVs with the markers of interest. The ExoCounter uses the same principles as an ELISA (sandwich-based assay). It is a disc which has grooves of a specific size (260nm at top and 160nm at the bottom) to allow for size selection and attachment of S-EVs only. The grooves are coated with a capture antibody, allowing for S-EVs to enter the groove and bind, which are labelled using a single labelling antibody conjugated to magnetic nanobeads that are detected by the optical pick-up which is comprised of a laser diode and photodetector[87]. The light is then converted into pulses which are counted by the pulse detector. The software is capable of calculating the number of S-EVs which express both epitopes for the capture antibody and labelling antibody.

The initial step was to coat each well of the disc with the coating antibody. Each coating antibody (CD63, CD29 or CD142) was added (90µL of 5µg/mL) in triplicate wells. No coating antibody was added into the background wells. The disc was sealed and incubated at 37°C for 1 hour. After the incubation, the wells were washed with 300µL of 0.05% Tween-20 in PBS (PBS-T) per well 3 times (program 1) on the automated washer. For the blocking stage, 200µL of 0.1µM of 1% casein in PBS-T was added to each well. The disc was sealed and incubated at room temperature for 1 hour. The disc was then washed with program 1. 50µL of the neat sample (either control culture media or spent culture media) was then added to each well (in triplicates) and the disc was sealed and incubated at 37°C for 2 hours with shaking at a speed of 100 on the disc shaker.

Meanwhile, the detection beads were prepared. The detection bead used was CD63. The beads were vortexed and diluted 1:100 using 1% casein in PBS-T. The diluted beads were then sonicated for 1 minute. Once the disc incubation was completed, the disc was washed with program 1 and

50µL of the diluted beads solution (20µg/mL) was added into each well. The disc was sealed and incubated at 37°C for 90 minutes with shaking at a speed of 100.

The plate was then washed with 3000µL of PBS-T once per well (program 2) followed by a wash using 3000µL of deionized water once per well (program 3). The discs were dried at 37°C for 15 minutes and placed into the ExoCounter with the assay side facing down. The blue-ray system read the signals and each bead was counted by the pulse sensing circuit. The raw data was obtained and analysed.

ExoCounter data analysis

For each triplicate, the corresponding background signal was subtracted. The mean and SEM was calculated from the triplicates and an unpaired t-test with Welch's correction was performed to determine significance.

NanoFCM plasma sample preparation

Normal pregnancy plasma samples were prepared for analysis with nFCM. Blood was collected from 3 women with normal pregnancies into vacutainers (BD Biosciences) with sodium citrate. The vacutainers were centrifuged at 1,500 x *g* for 15 minutes at 21°C in a swing bucket centrifuge (Beckman Coulter Avanti J-20XP centrifuge using a Beckman Coulter JS-5.3 swing-out rotor). The supernatant (platelet poor plasma) above the buffy coat was removed and spun at 13,000 x *g* for 2 minutes at 21°C in a fixed angle microfuge. The supernatant (platelet free plasma) was removed, mixed and aliquoted into 600µL and frozen at -80°C for future use.

NanoFCM

NanoAnalyzer U30 instrument (NanoFCM Inc.) was used for the detection of side-scatter (SSC) and fluorescence. The analyser had a dual 488/640nm laser and single-photon counting avalanche

photodiode detections. Measurements were taken at a pressure of 1.0 kPa for 1 minute. The sample flow rate was calibrated using 250nm silica nanoparticles of known concentration. Pooled samples were measured by nFCM to calculate the concentration in particles/ml. The concentration of samples was determined by comparison to 250nm silica nanoparticles of known concentration. From this, a volume of each sample was calculated to allow for $1x10^8$ particles to be stained.

In the case of the embryonic samples, the maximum available volume of sample was used (400µL). The samples were then incubated with FITC-conjugated anti-mouse CD63 (Abcam), and CD142-APC (Miltenyi) for 30 minutes in the dark placed on ice. Another batch of samples was stained with a cocktail of anti-mouse CD63 (Abcam), anti-mouse CD9 (Abcam), anti-mouse CD81 (Biolegend), and REA CD142-APC (Miltenyi) (1:200 dilutions for all antibodies). Control culture media and plasma samples were also stained in the same manner. The stained samples were then diluted to 1mL using fPBS and transferred into ultracentrifuge appropriate tubes. Samples were spun at 100,000 x *g* for 1 hour and the supernatant was discarded. The pellet was resuspended in 50µL of fPBS and analysed.

The processing of data was completed with the nFCM Professional Suite v1.8 software, providing a single PDF with dot plots, histograms and statistical data. The percentage positivity of the samples was obtained and the mean was compared to controls using an unpaired t-test with Welch's correction.

Results

Particle tracking analysis, transmission electron microscopy, and western blot

Once the samples were processed and the embryonic extracellular vesicles were separated using size exclusion chromatography, the samples were run using the Zetaview NTA to determine the size and concentration of the particles. Two pooled samples were run showing a mean diameter of

133.15±14.35nm with mean particle concentration of $1.45x10^7 \pm 2.5x10^6$particles/mL (Figure 2.1.A). The snapshot of the particles going through the system is shown in figure 2.1.B. The control culture media samples did not show any particles and the software did not produce a report.

I also wanted to perform characterisation of the EVs using western blot. I first isolated the EVs using ultracentrifugation. However, due to the small volume and low protein concentration of the samples, the only EV marker that showed staining in the embryonic samples (E1 and E2) was CD63 (Figure 2.1.C). CD63 is a tetraspanin that is commonly found to be expressed on the surface of EVs, specifically S-EVs. The control culture media (C1 and C2) did not show expression of this marker compared to the embryonic samples and positive control HEK293 cell line (Figure 2.1.C). The other markers were not detectable due to the low protein concentration of the samples.

For transmission electron microscopy, 4 culture media droplets containing single embryos from the same patient were pooled and analysed, the images shown in figure 2.1.D show the presence of S-EVs. The cup shaped structures are representative of what EVs would look like when processed for TEM as the structure collapses when the intra-structural liquid is removed during the drying process. The EVs outlined in the images are smaller than 200nm (Figure 2.1.D).

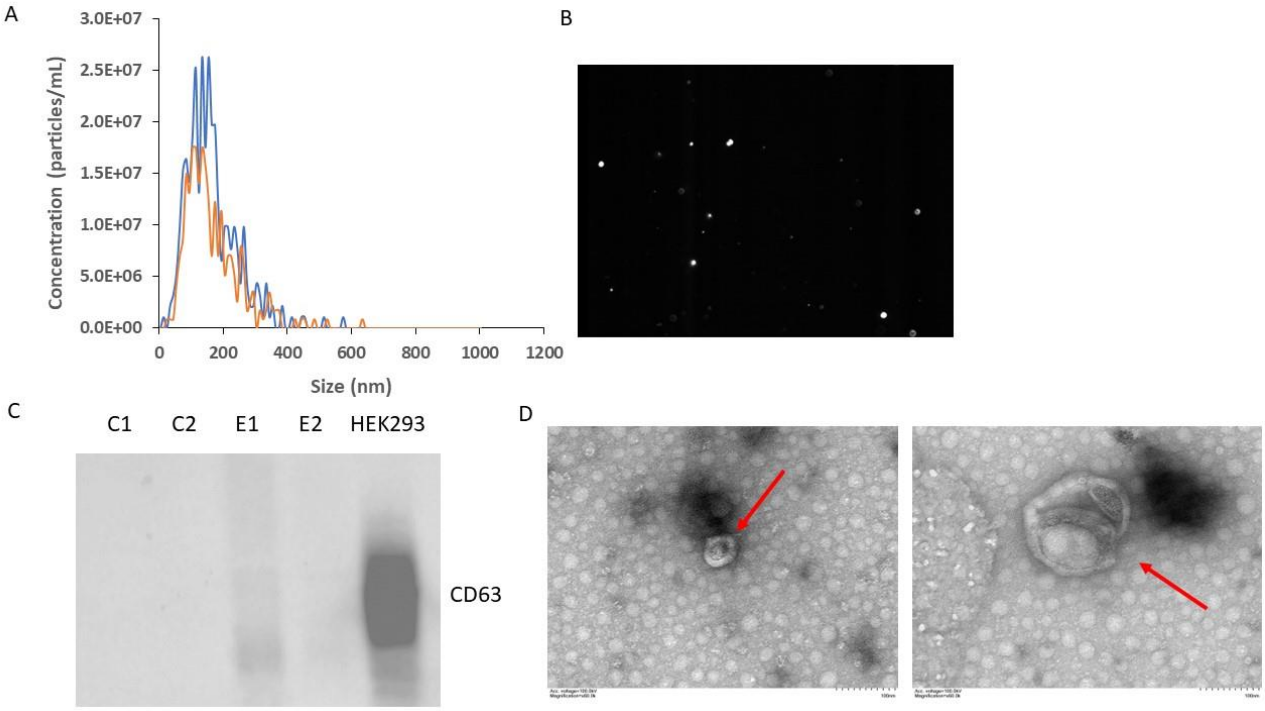

Figure 2.1: Nanoparticle tracking analysis, western blot and transmission electron microscopy of embryonic extracellular vesicles. **A)** Size distribution and concentration of two pooled EEV samples isolated using size exclusion chromatography. The Zetaview NTA (Particle Metrix, BASIC PMX-120) and ZetaView software (version 8.05.12 SP2) were used. **B)** Screenshot of the particles detected by the Zetaview system. **C)** Western blot of CD63 tetraspanin showing control culture media (C1 and C2), embryonic EEVs isolated by ultracentrifugation (E1 and E2) and positive control HEK293 cell line. **D)** Transmission electron microscopy of EEVs isolated using the exosome extraction kit (EXOQ5A-1, System Biosciences). The red arrows are pointing to small EEVs. The scale bars are 100nm and both images are at 60X magnification.

MACSPlex with EEVs, control culture media and HUVECs

With the presence of EEVs confirmed through NTA, TEM, and western blot, I next decided to look at the surface protein composition of the EEVs and compare that to control culture media as well as culture media from HUVECs. I used the MACSPlex kit which identifies 37 common surface proteins and 2 isotype controls in the sample through a sandwich-based method. The normalised median fluorescent intensity of the Alexa-Fluor 647 channel of each surface protein from each of the 8 embryonic samples is shown in figure 2.2.1. The REA and mIgG1 control beads

did not show any positive fluorescence, whereas some of the other proteins had high expression in the samples (Figure 2.2.1.D).

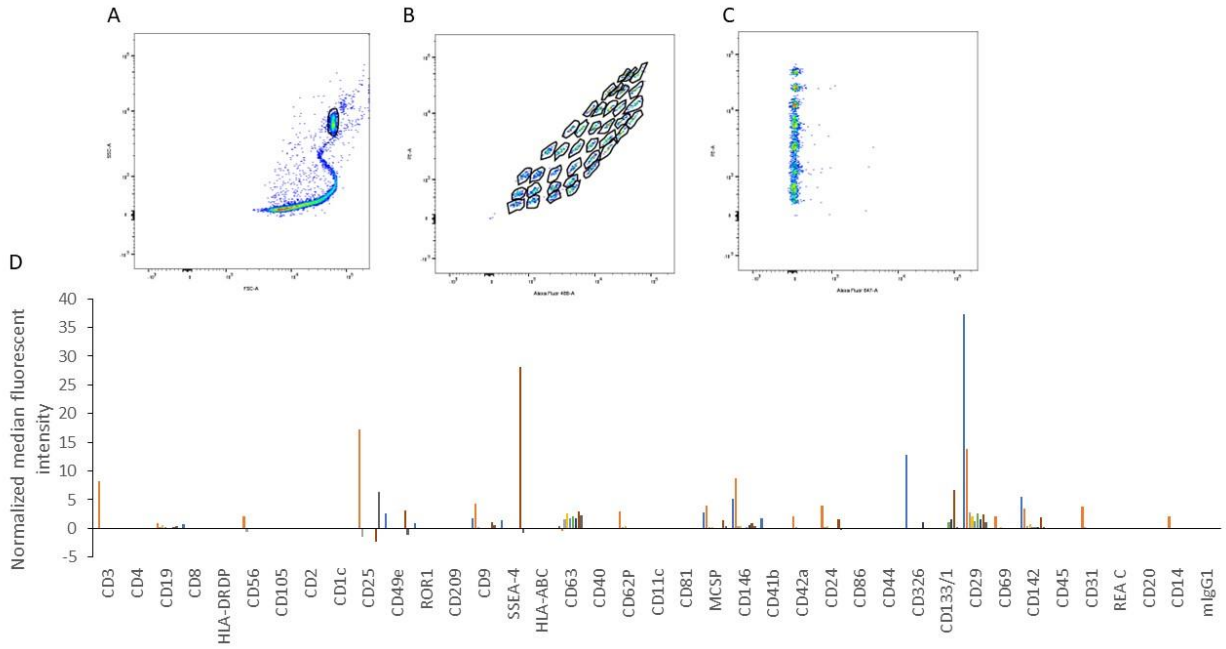

Figure 2.2.1: MACSPlex profile of embryonic extracellular vesicles showing 37 common surface proteins and 2 isotype controls. **A)** representative gating of the singlet bead population. **B)** representative gating of each of the capture beads population corresponding to each surface protein. **C)** Alexa Fluor 647 channel, where the median fluorescent intensity of each marker was calculated. **D)** Each colour bar represents the normalized median fluorescent intensity of a single sample, n=8.

From all of the surface protein markers, 4 were commonly expressed (CD142, CD29, CD63, and CD146) by all 8 samples. The average normalised MFI (nMFI) of the embryonic samples were compared to control culture media and showed higher expression in the embryonic samples. For CD142, the embryonic sample had a significantly higher nMFI of 1.197±0.361 compared to 0.105±0.015 of the control (p=0.019) (Figure 2.2.2.A). For CD29, the embryonic sample had a higher nMFI of 3.451±1.502 compared to 0.393±0.213 (p=0.083) (Figure 2.2.2.B). For CD63, the embryonic sample had a significantly higher nMFI of 1.846±0.363 compared to 0.715±0.031 of the control (p=0.017) (Figure 2.2.2.C). Lastly, CD146 showed a higher nMFI of 1.446±1.035 in

the embryonic sample compared to an nMFI of 0.101±0.129 of the control (p=0.237) (Figure 2.2.2.D).

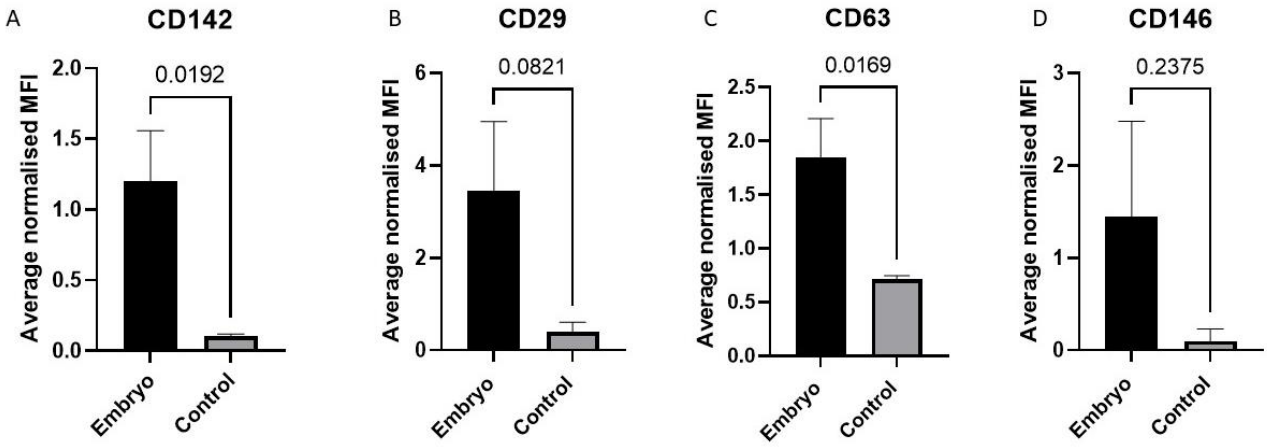

Figure 2.2.2: Four common surface markers [**A)** CD142, **B)** CD29, **C)** CD63 and **D)** CD146] expressed by embryonic samples compared to control culture media. The average normalised median fluorescent intensity is shown with SEM and an unpaired t-test with Welch's correction was performed to determine significance.

At this point, I wanted to compare the nMFI of the 4 markers in the embryonic sample with culture media from HUVECs. The rationale was to test the expression in another cell culture media which is commonly known to express CD146 and CD142 when activated. Comparing the nMFI between the embryonic samples and HUVEC culture media, CD142 (1.197±0.0361 vs -0.136±0.080, p=0.0074), CD29 (3.451±1.502 vs 1.808±0.365, p=0.319) and CD63 (1.846±0.363 vs 0.422±0.269, p=0.0132) showed a higher value for the embryonic sample which was significant for CD142 and CD63 (Figure 2.2.3.A-C). The HUVEC culture media had a significantly higher nMFI for CD146 (1.446±1.035 vs 4.917±0.856, p=0.034) which was expected as CD146 is a known marker of endothelial cell origin (Figure 2.2.3D).

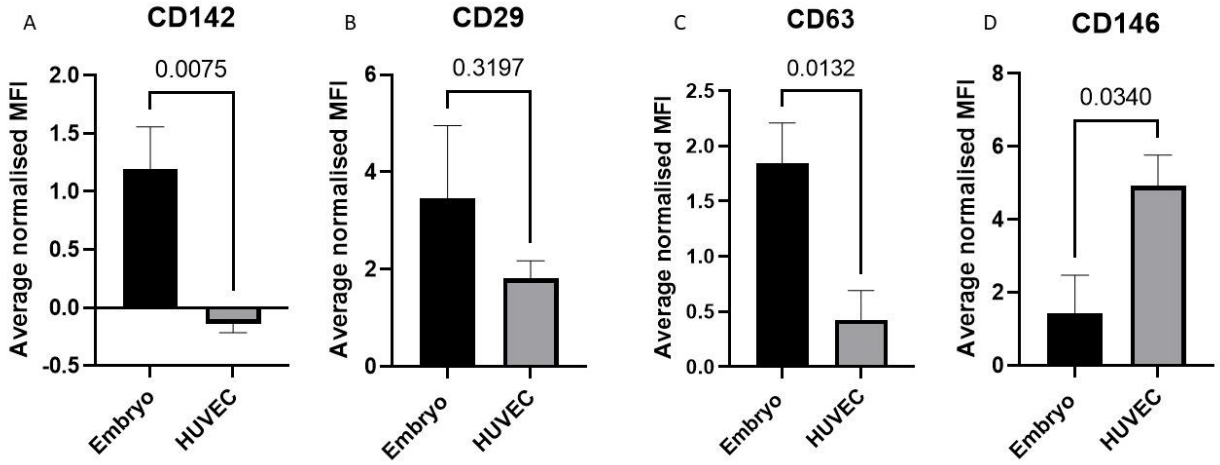

Figure 2.2.3: Four common surface markers [**A)** CD142, **B)** CD29, **C)** CD63 and **D)** CD146] expressed by embryonic samples compared to control culture media from human umbilical cord vascular endothelial cells. The average normalised median fluorescent intensity is shown with SEM and an unpaired t-test with Welch's correction was performed to determine significance.

The HUVECs were then treated with TNF-α which is used to induce a pro-inflammatory response from the cells and has been shown previously to increase the expression of CD142. The same 4 markers were tested in the same volume of culture media of TNF-α treated HUVECs and were compared to the same embryonic samples. The embryonic samples retained a significantly higher nMFI for CD142, (1.197± 0.361 vs 0.151±0.061, p=0.023), CD63 (1.846±0.363 vs 0.886±0.118, p=0.035) and a non-significant higher nMFI for CD29 (3.451±1.502 vs 1.432±0.040, p=0.221) (Figure 2.2.4A-C). Lastly, for CD146, the embryonic sample had a significantly lower nMFI compared to the activated HUVECs (1.446±1.035 vs 5.116±0.221, p=0.0091) (Figure 2.2.4D). The expected increase in the expression of CD142 as a result of treating HUVECs with TNF-α was confirmed with the nMFI of CD142 increasing from -0.136±0.080 to 0.151±0.061 in the HUVECs samples.

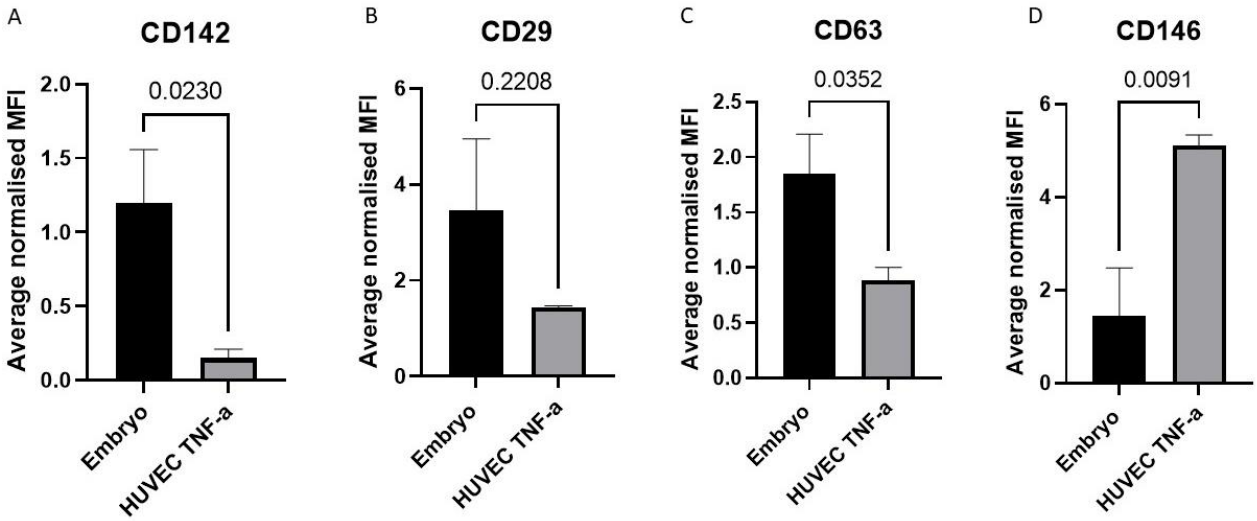

Figure 2.2.4: Four common surface markers [**A)** CD142, **B)** CD29, **C)** CD63 and **D)** CD146] expressed by embryonic samples compared to control culture media from human umbilical cord vascular endothelial cells activated with tumor necrosis factor alpha. The average normalised median fluorescent intensity is shown with SEM and an unpaired t-test with Welch's correction was performed to determine significance.

I then wanted to determine the pairings of the identified markers with the small EV tetraspanins which would be used for subsequent experiments. I used the detection beads from the MACSPlex kit (CD63, CD9 and CD81) separately on n=3 for each detection bead to identify the pairings of the tetraspanins with the surface markers. Comparing the results, CD63 capture beads showed the highest MFI for CD142 (9.467±2.347) compared to CD81 (5.15±2.082), and CD9 (3.8). CD9 capture beads showed the lowest MFI for all 4 markers, with no signal for CD146 (Figure 2.2.5). From these results, I decided to use the pairing of CD63 with CD142 and CD29 for the ExoCounter experiments since these signals were more robust. I also decided to use CD63 paired with CD63 as a control since double positive EVs have been shown to be present in all sample types.

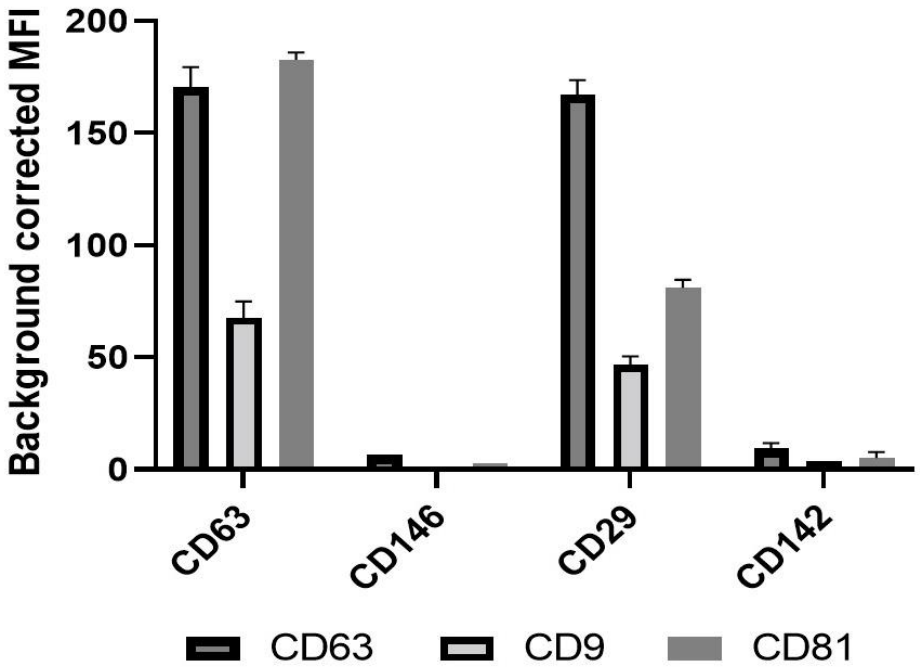

Figure 2.2.5: The four common markers found in embryonic samples with their corresponding tetraspanin pairings. Three embryonic samples were used for each tetraspanin pairing and the background corrected median fluorescent intensity from the Alexa Fluor 647 channel was calculated. The mean values with SEM are shown.

ExoCounter with EEVs for CD142, CD63 and CD29

With the ExoCounter, I had the unique opportunity to quantify a small subpopulation of S-EEVs expressing EV specific tetraspanin along with any other surface protein. The assay allows for the capture of S-EVs on a disc through size specificity and binding to a specific capture antibody, followed by labelling with another specific antibody which is quantified by the Blue-ray system. I chose to look at CD63/CD63 positive populations to confirm the presence of S-EEVs in the samples. I then looked at two pairings identified from the MACSPlex kit (CD29/CD63 and CD142/CD63). I used culture media as a negative control and was able to show that the embryonic samples had a significantly higher total S-EEV count for CD63/CD63 positive S-EEVs (33453.67±2077.792 vs 17479.33±939.815, p=0.00758), and CD29/CD63 positive S-EEVs (4818.667±453.308 vs 664.667±51.20, p=0.0109) (Figure 2.3.A-B). The embryonic samples also

had a non-significantly higher CD142/CD63 subpopulation of S-EEVs (3718±340.894 vs 1539±405.286, p=0.078) which could become significant with more samples tested (Figure 2.3.C). Also, the other subpopulations of CD142 positive EEVs which were paired with CD9 and CD81 are missed in the total count using this method.

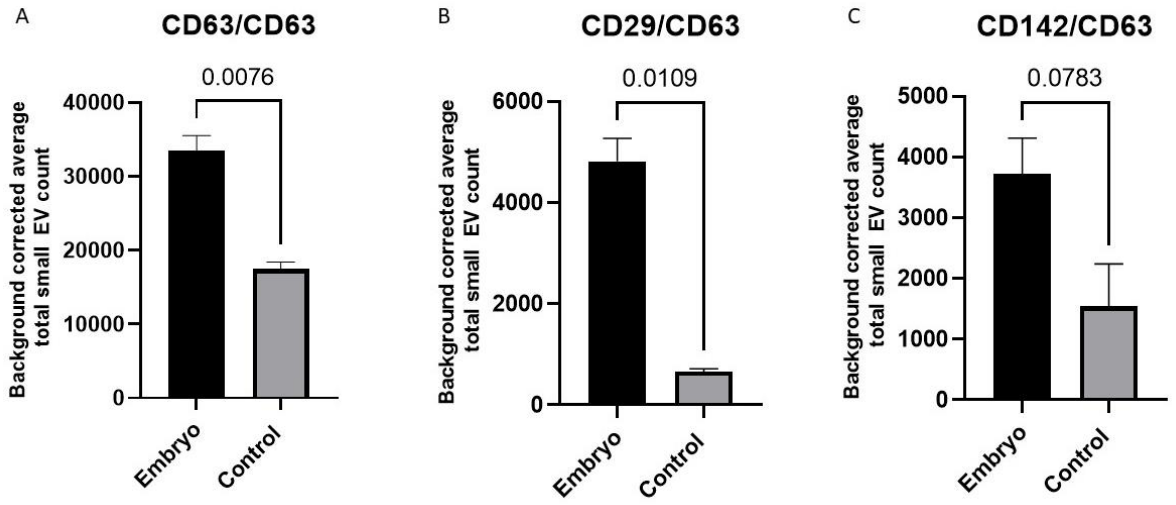

Figure 2.3: The background corrected average total small EV count for EEVs in embryonic vs control samples. **A)** CD63/CD63, **B)** CD29/CD63 and **C)** CD142/CD63 positive EEVs (n=3). The mean and SEM are shown with unpaired t-test with Welch's correction for significance.

NanoFCM with EEVs and plasma

At this point, I was interested in CD142 positive EEVs as these would be potentially physiologically relevant in the process of implantation and were not previously reported. To confirm the presence of the CD142 positive S-EEV subpopulation, I decided to use another novel method to quantify the positive percentage of EEVs in comparison to control culture media and plasma EVs. I began by using CD63-FITC conjugated antibody and CD142-APC conjugated antibody to check the positivity of each subgroup and the dual positivity for both antibodies. Comparing the CD63-FITC subpopulation, the embryonic sample had a non-significantly higher positivity (4.6%±0.00173) compared to the control culture media (3.2%±0.0035, p=0.0993) (Figure 2.4.1.A). As for the CD142-APC subpopulation, the embryonic sample had a non-

significantly higher expression (8.1%±0.0163 vs 3.8%±0.001, p=0.120) (Figure 2.4.1.B). As for the dual positive population, the embryonic sample had a non-significantly higher positive population (2.5%±0.0069 vs 1.1%±0.0005, p=0.163) (Figure 2.4.1.C).

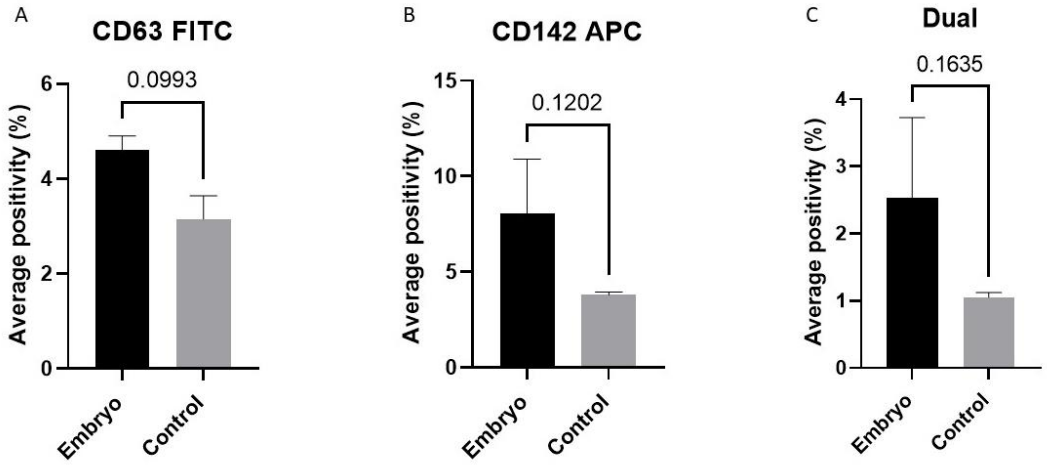

Figure 2.4.1: The average positivity of the CD63, CD142 and double positive EVs found in the embryonic sample vs the control culture media (n=3). The mean and SEM is shown with an unpaired t-test with Welch's correction for significance.

The nFCM technology was established using blood samples and so I wanted to check whether the system could pick up a difference between plasma samples from pregnant women and our embryonic samples. Plasma samples have also been shown to contain EVs, specifically CD142 positive EVs making them a good control. Using the same CD63-FITC antibody and CD142-APC antibody, the embryonic samples had a non-significant positive CD142 population (8.1%±0.0163 vs 3.7%±0.0152, p=0.125), and dual positive population (2.5%±0.0069 vs 1.2%±0.0046, p=0.193) (Figure 2.4.2.B-C). However, the plasma samples showed a non-significantly higher expression of the CD63 subpopulation (4.8%±0.00467 vs 4.6%±0.00173, p=0.0676) (Figure 2.4.2.A).

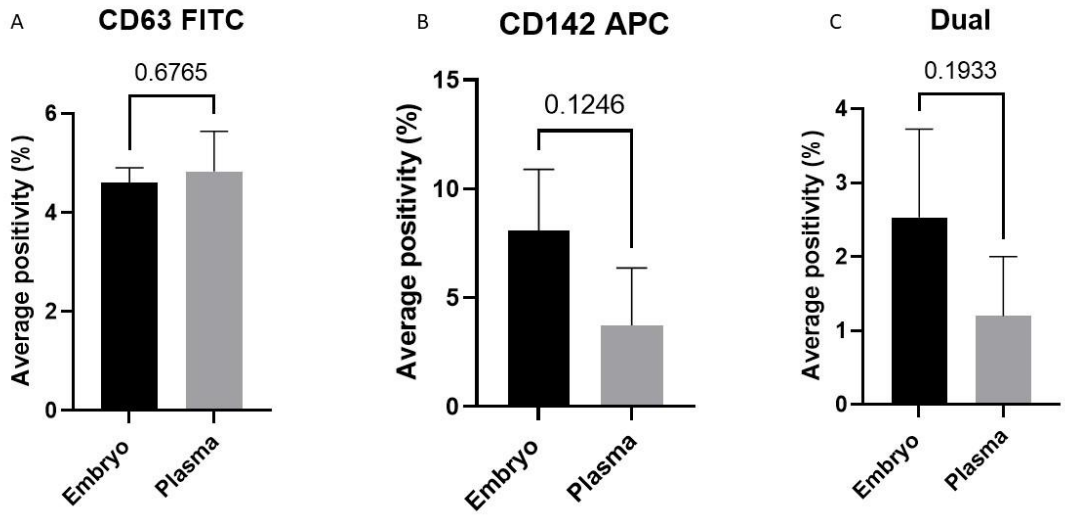

Figure 2.4.2: The average positivity of the CD63, CD142 and double positive EVs found in the embryonic sample vs plasma from normal pregnancy (n=3). The mean and SEM is shown with an unpaired t-test with Welch's correction for significance.

Similar to the ExoCounter, by using CD63-FITC conjugated antibodies, I was only identifying a portion of the CD142 positive EEVs. As shown previously by the MACSPlex kit, there are also CD142 EEVs paired with CD9 as well as CD81. To capture these populations, I decided to use a cocktail of all three tetraspanin antibodies to increase the positive populations. With this approach, the embryonic samples had a significantly higher positive cocktail-FITC population (9.6%±0.00755 vs 3.9%±0.0065, p=0.011), significantly higher positive cocktail-APC (CD142) population (11.6%±0.0048 vs 6.0%±0.001, p=0.0056), and a significantly higher dual-positive population (5.1%±0.0039 vs 1.9%±0.002, p=0.007) compared to the control culture media (Figure. 2.4.3.A-C)

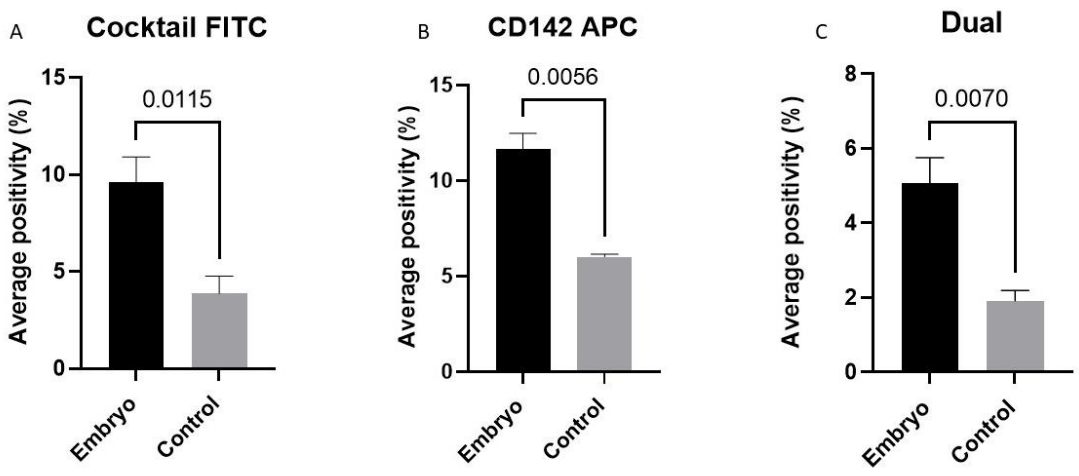

Figure 2.4.3: The average positivity of the cocktail positive (CD63, CD9, and CD81), CD142 positive and double positive EVs found in the embryonic sample vs control culture media (n=3). The mean and SEM is shown with an unpaired t-test with Welch's correction for significance.

Once again, I wanted to check the positive populations against plasma samples to ensure the signals are unique to embryonic samples and not replicated in other biological fluids. The embryonic samples showed a significantly higher cocktail-APC (CD142) positive population (11.6%±0.0048 vs 1.5%±0.0101, p=0.0035), and a significantly higher dual positive population (5.1%±0.0039 vs 0.4%±0.0033, p=0.00096) (Figure 2.4.4.B-C). The embryonic sample also had a non-significantly higher cocktail-FITC positive population (9.6%±0.00755 vs 5.5%±0.0198, p=0.166), suggesting that using the same starting volume of embryonic culture media and plasma, there are similar levels of EVs expressing the common tetraspanins CD63, CD9 and CD81 (Figure 2.4.4.A). However, there are more EEVs expressing CD142 on their surface compared to plasma, confirming that this subpopulation of EVs is likely unique to embryos.

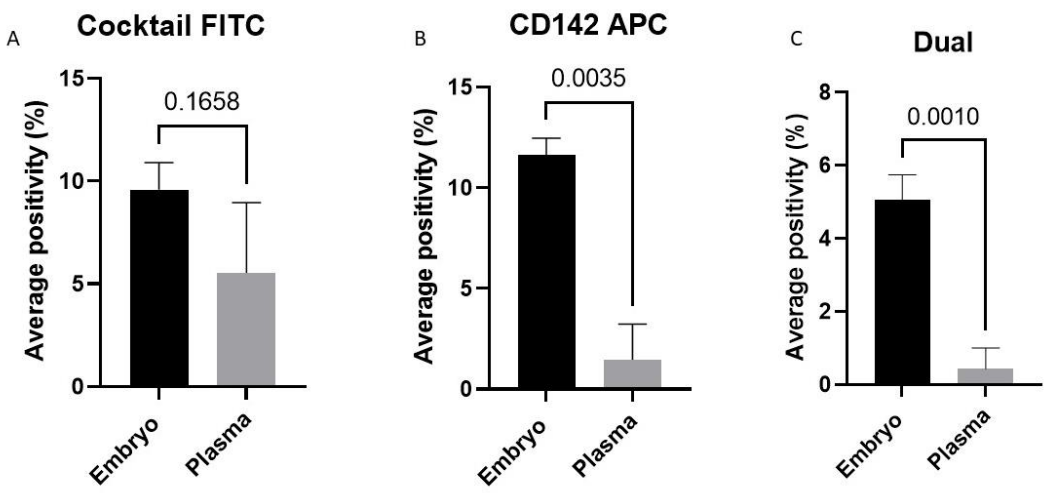

Figure 2.4.4: The average positivity of the cocktail positive (CD63, CD9, and CD81), CD142 positive and double positive EVs found in the embryonic sample vs plasma from normal pregnant women (n=3). The mean and SEM is shown with an unpaired t-test with Welch's correction for significance.

Lastly, I wanted to compare the median size of the S-EVs between the embryonic samples, control culture media as well as plasma. Using the CD63-FITC antibody the embryonic EVs showed a lower median size (85.083nm±0.927) compared to control media (160.75nm±38.5) and a significantly lower median size compared to plasma (165.583nm±2.848) (Figure 2.4.5.A). Using the cocktail-FITC antibody, the EEVs had a smaller median size (94.4167nm±4.206) compared to control media (101.0nm±1.75) and a significantly smaller median size compared to plasma EVs (125.083nm±7.219) (Figure 2.4.5.B). With the cocktail-FITC antibody, the CD9 and CD81 positive EEVs are also recognized by the technique suggesting that these EEVs may have a larger size compared to the CD63 subpopulation alone. However, this is not replicated in the plasma and control media, where the CD63 subpopulation has a larger median size compared to the cocktail. This again suggests that there are unique properties associated with EEVs which are not replicated in the controls.

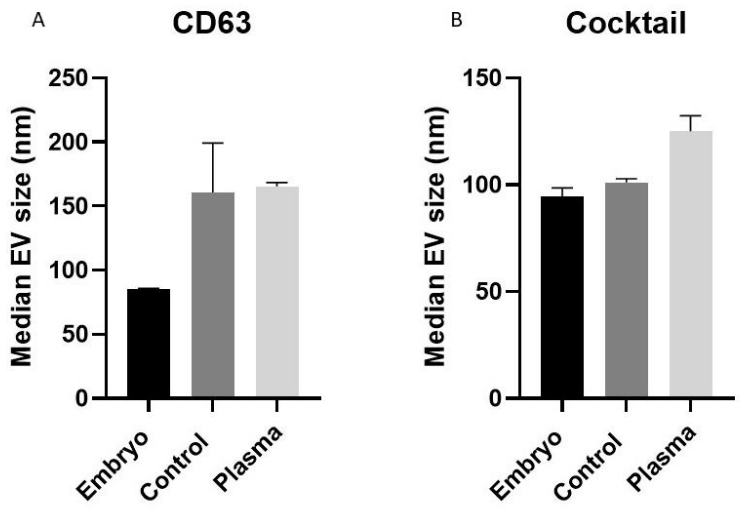

Figure 2.4.5: Median EV size for embryonic samples, control culture media and plasma samples using only CD63 tetraspanin and CD142 dual positive populations vs. cocktail of tetraspanins paired with CD142. The mean and SEM are shown for each sample type (n=3).

Discussion

EEVs are released throughout embryonic development and could be important modulators for implantation and early pregnancy. I was able to show through NTA, TEM and western blot of CD63 that in our sample type, EEVs are present. The EEVs have the characteristic size distribution and structure of S-EVs and showed higher expression of CD63 compared to control culture media (Figure 2.1). CD63 is a common tetraspanin found on the surface of EVs and is used as a marker for identifying EV populations, specifically S-EVs. The characterisation of surface proteins found on EEVs has not been reported which is why I decided to use the MACSPlex kit to identify possible common markers. The kit allows for qualitative and semi-quantitative analysis of 37 common surface markers found on S-EVs. It achieves this through a sandwich-based method of capture beads, S-EVs and detection beads which are identified through different fluorescent intensities. With the kit, I was able to identify 4 common surface proteins in the embryonic samples (CD29, CD142, CD146 and CD63) (Figure 2.2.2).

CD146

CD146, also known as MUC18 is a cell adhesion molecule and a marker of endothelial cells as part of the endothelial junction associated with actin filaments[88]. The significantly higher expression of CD146 by HUVECs, compared to embryonic samples, was therefore expected as HUVECs are established endothelial cells (Figure 2.2.3). CD146 is a 113kDa transmembrane glycoprotein belonging to the immunoglobulin superfamily. It is an angiogenic factor with its soluble form (sCD146) detectable in the plasma of pregnant women as well as in culture media from IVF embryos[89,90]. It is known to regulate the invasion of the extravillous trophoblast and is only expressed in the pregnant uterus and not in non-pregnant[91]. Blocking CD146 in the mouse uterus results in the impediment of trophoblast invasion, angiogenesis and vascularization[92]. In one study, measuring the sCD146 levels from culture media of IVF cycles using ELISA showed that there were higher levels from embryos that did not implant[90]. The soluble form can cleave and prevent the interactions of the membrane bound CD146 (mCD146) which is involved in trophoblast invasion and placentation[93]. It has therefore, been suggested that sCD146 could be used as a non-invasive method of embryo selection[90]. The expression of this protein on the EEVs could be interesting to study further as it could either compensate for the loss of function of mCD146 or introduce another source of mCD146 which is important for invasion during implantation.

CD29

CD29 is also known as integrin beta 1 (ITGB1) which is a membrane receptor involved in cell adhesion and recognition. The integrins form heterodimer integrin complexes between alpha and beta integrins to facilitate adhesion and other processes such as cellular proliferation and signalling[94]. Integrins have been shown to play a role in the implantation adhesion cascade by

binding extracellular matrix components such as collagen to facilitate adhesion, invasion and cytoskeletal reorganization[95]. ITGB1 has been shown to be expressed in the endometrium with an increased expression on days 20 to 24 of the cycle corresponding with the time of implantation[96]. It is also detected in the embryo with roles in the basement membrane formation on the inner surface of the trophectoderm and between the endoderm and ectoderm. In a mouse model, ITGB1 null mouse embryos showed peri-implantation lethality[95]. When linked with integrin alpha 4, ITGB1 plays important roles in the maternal foetal recognition and was found to be absent in the endometrium of women with unexplained infertility[97,98]. It has also been shown that in response to progesterone (the main hormone of the receptive phase), ITGB1 expression by the endometrium is increased, resulting in better efficiency of embryo attachment[99]. Therefore, integrins including ITGB1 are important for the implantation and maintenance of pregnancy. Integrins are markers found in a variety of different cells and are not specific which is why there was no significant difference in the expression of CD29 in embryonic samples and HUVECs (Figure 2.2.3). Due to the role of ITGB1 in pregnancy, its release by embryos is not surprising as it could be a marker of embryonic health and the invasive properties of the trophoblast as well as a new source to the endometrium prior to implantation.

CD63

CD63 is a tetraspanin which is commonly found to be expressed on the surface of EVs and is used as a marker for identifying EVs specifically S-EVs as it is expressed on multivesicular bodies[100]. It plays roles in the trafficking and targeting of EVs[101] with multiple CD63 proteins found on the surface of a single EV and the rationale for using double positive CD63 to confirm the presence of EVs using the ExoCounter (Figure 2.3). It has also been shown to play a role in the activation of ITGB1 leading to the activation of MAPKs, promoting cell survival, cell adhesion and invasion.

It forms a bridge between ITGB1 and vascular endothelial growth factor receptor 2 (VEGFR2) which is involved in sprouting angiogenesis of vascular cells[102]. Therefore, it is interesting that using the MACSPlex kit, CD29 was paired with CD63 at higher levels compared to CD9 and CD81 (Figure 2.2.5) suggesting a physiological reason for this pairing.

CD142

CD142, also known as tissue factor (TF) is a 47 KDa transmembrane glycoprotein and a member of the class II cytokine/haematopoietic growth factor family. It is involved in the extrinsic coagulation cascade and has been shown to be highly expressed in the placenta contributing to the hypercoagulability state of pregnancy[103]. It is mostly known for its role in the coagulation cascade. Its upregulation has been associated with antiphospholipid syndrome which can cause frequent coagulation and or miscarriage[104,105]. However, tissue factor has also been shown to play important roles in embryogenesis. It is also involved in cellular proliferation and the maintenance of blood vessels with roles in cell adhesion and migration. Its importance is seen in tissue factor null mouse embryos where the disruption of the gene resulted in 90% lethality[106–108]. Tissue factor can also transmit signals that are independent of the coagulation cascade such as, increasing cell motility, invasiveness, angiogenesis, cell proliferation and cell survival[109]. The signalling is achieved through activation of the protease-activated receptor-2 (PAR2) and integrin signalling[110]. Interestingly, the soluble form of TF was shown to induce cellular proliferation through binding to ITGB1 which was one of the other common markers we found on EEVs[111]. This suggests that there may be an interplay of the common markers to increase the receptivity of the endometrium. TF has been shown to be expressed by stromal endometrial cells, and increase the expression of pro-implantation markers such as LIF and IL-8, making it important for implantation and pregnancy[112–114]. In the stroma, it may play roles in protection from excessive bleeding during

trophoblast invasion[114]. The expression of TF is, however, much lower in endometrial epithelial cells and so its presence on the embryonic EVs is interesting because it could be affecting the epithelial cells as these are the cells that would be contacted first. The presence of the CD142 subpopulation of EEVs was confirmed further through the use of new EV characterisation technologies. Out of the four commonly expressed proteins, CD142 was chosen for further analysis because its role in implantation has not yet been fully elucidated and it could be playing interesting roles in priming the epithelial layer of the endometrium for accepting the embryo. This forms the basis of the subsequent chapters in which I treated endometrial epithelial cells with recombinant human tissue factor.

New technologies for EV characterisation

One of the main reasons why I chose to use these new technologies was due to the small sample size and small number of EEVs present in the samples. These new technologies are more sensitive compared to traditional methods such as western blot and are able to differentiate between the limited number of EVs. For example, the ExoCounter is a new equipment designed for the direct quantification of antigen-specific S-EVs in small sample volume with detection sensitivity that is 800-fold higher than traditional ELISA and 420-fold higher than flow cytometry[115]. This system confirmed the presence of CD142 positive EEVs in our samples (Figure 2.3). Another method that I used was the nFCM which is superior to traditional flow cytometry for the detection and characterisation of small sized particles. Normal flow cytometry is designed for cells and not small particles which means that it cannot distinguish between noise and small sized particles, with limit of detection being over 300nm in size[116,117]. However, the nFCM is capable of identifying particles as small as 40nm in size, making it ideal for the detection of S-EVs[118,119]. I was able to show that CD142 positive EVs were more highly expressed by embryonic samples compared to the control

culture media as well as normal pregnancy plasma samples suggesting a specific role for these subpopulations of EEVs. The size of the CD142/CD63 positive EEVs was lower than those found in the plasma samples and control media suggesting that the smaller size (Figure 2.4.5) could be important in their method of uptake and transport to the endometrial epithelial cells.

Limitations and Future work

The main limitation of this chapter was the low volume of the samples and low concentration of EVs, making it difficult to identify the EEVs using traditional methods such as western blot for other markers such as CD9, CD82, ALIX, Syntenin, and TSG101. Also, due to the nature of the culture dish, it was not possible to link samples back to a single embryo and so I could not deduce whether the embryos in the samples were of good or poor-quality or link the results to outcome data. These could be rectified by using different sources of samples. I focused on the S-EEVs which are more physiologically interesting. However, it is worth investigating the M/L-EEVs as well to delineate whether there are similarities and/or differences between the EEV subtypes.

Embryonic development is versatile and unique at each time point. For example, prior to day 3 of development, the embryonic genome is not yet activated and it is the maternal genome which is used. From day 3 onwards, the embryonic genome becomes activated and replaces the maternal genome. This suggests that there are rapid changes in the composition of the cells throughout development. Therefore, there are likely changes in the composition of the EVs being released from these cells. It would be interesting to look at the protein composition of EEVs at different time points of embryonic development to better understand the *in vivo* process from fertilisation to implantation and whether certain factors are upregulated/downregulated at certain time points. The potential clinical use would then be to possibly supplement the endometrium with the different factors at different timepoints to mimic the *in vivo* process and prime the endometrium for embryo

implantation. There is also the possibility for use in more targeted contraception at different time points after conception.

In conclusion, in this chapter, I was able to show the presence of EEVs using the traditional methods of NTA, TEM and western blot. I also identified specific common surface proteins (CD146, CD29, CD63 and CD142) found on S-EEVs using the MACSPlex kit. The ExoCounter and nFCM were used as alternative methods to identify the specific subpopulations. I was able to show that there was a higher expression of TF positive EEVs compared to control culture media, HUVEC culture media and plasma samples which I decided to investigate further in the next chapter.

Chapter 3

Ishikawa cells and the effects of human recombinant CD142 protein treatment

Introduction

Implantation

One of the main bottlenecks of pregnancy is the implantation of an embryo into the uterus. The process of implantation begins with the hatching of the blastocyst from the zona pellucida and its apposition to the epithelial layer of the endometrium which is followed by the attachment and invasion of the embryo into the endometrium to begin placentation[120,121]. The interaction between the endometrium and the embryo is enhanced through the release of hormones, cytokines, and growth factors which increase the receptivity of the endometrium while attracting the embryo[122,123]. Some of these are involved in the attachment and adhesion of the embryo, such as E-cadherins, and cellular adhesion molecules[124]. Other factors include leukaemia inhibitory factor (LIF) which is involved in stromal cell proliferation and uterine receptivity[125,126]. The expression of macrophage colony stimulating factor 1 (M-CSF) is concurrent with the attachment of the embryo due to the expression of M-CSF receptor on the surface of the trophectodermal cells[127,128]. The endometrium also expresses morphological changes such as the formation of pinopods which are projections of the epithelial layer that bind their ligands on the embryo and allow for adhesion in a specific orientation facilitating further development of the embryo[129]. Once the attachment is completed, the trophectoderm cells begin to invade into the stromal layer, completing implantation and initiating placentation. The process of implantation occurs within a specific time frame (window of implantation) to provide hormonal feedback to the ovary to inhibit the next menstrual cycle and subsequent loss of the embryo[130]. During the window of implantation, the molecular and morphological changes enhance the receptiveness of the endometrium, attracting the embryo. The interplay of communication between the endometrium and the embryo is therefore crucial for

implantation and if the communication is abnormal, it could lead to adverse outcomes such as miscarriage.

The communication through extracellular vesicles released by embryos (EEVs) has been studied and in the previous chapter, I showed the presence of these EEVs in our embryonic samples along with the characterisation of common surface markers. From the identified surface markers, I decided to focus on CD142 (tissue factor) and the potential role it may play in the implantation of the embryo through its expression on EEVs. The focus of this chapter is to study the effects of recombinant CD142 protein (rCD142) on endometrial epithelial cells.

Endometrial layers

The endometrium is comprised of the epithelial cells and the underlying stromal cells and is the site of implantation. The epithelial cells are subdivided into luminal and glandular cells which both play important roles in the initial steps of implantation. The luminal cells face the lumen of the uterus, and the glandular cells form structures with secretory functions through invaginations into the stroma[121]. The epithelial cells are the first contact point of the embryo and therefore EEVs. They are columnar cells which undergo cellular differentiation to form cuboidal cells with adhesive membranes. The endometrial epithelial cells demonstrate cell polarity which typically functions as a barrier as well as regulating cell adhesion and cell signalling[131]. The polarised cells would therefore, repel the polarised trophectoderm of the blastocyst and impede implantation. For the endometrium to become receptive, the apical-basolateral cell polarity is lost and the cells undergo molecular and morphological changes[132]. The loss of cell polarity is promoted by progesterone to allow for the attachment of the apical surface of the trophectodermal cells of the embryo to the epithelial cells of the endometrium. The protein expression of the epithelial cells also changes during the loss of polarity. For example, leukaemia inhibitory factor (LIF) is

upregulated during the secretory phase and is crucial for endometrial receptivity. Some patients who have unexplained infertility and recurrent implantation failure showed a lack of upregulation of LIF, suggesting that the depolarisation of the endometrial epithelial cells may be playing a role[126]. The importance of the loss of polarity is seen in RL95-2 cell lines which are non-polarised and show higher attachment of JAR spheroids (model for blastocyst) compared to HEC-1A cells which are polarised[133].

Morphological changes of the epithelial cells include the formation of pinopods. The secretory epithelial cells which contain microvilli, form protrusions on their apical side in response to progesterone[134]. The exact role of pinopods in implantation is not understood. However, they are upregulated during the window of implantation. These structures also express proteins such as integrins, L-selectin and heparin binding epidermal growth factor which are important for the apposition and attachment of the embryo[129]. A reduction in pinopods has been linked to some cases of unexplained infertility[135]. To better understand the interactions of the endometrium and embryo as well as the changes induced peri-implantation and during implantation, cell lines have been adopted for study.

Cell lines

Primary cells are not readily available and difficult to culture with patient specific variations that may skew the results. Therefore, an efficient way to study the endometrium is through the use of cell lines which allow for reproducible studies. The focus of our study is the epithelial layer of the endometrium, as this is the first contact point for the embryo. There are a few cell lines which are commonly used, such as HEC-1A, RL95-2, AN3-CA and Ishikawa cells. All of the mentioned cell lines are responsive to hormonal treatment. However, the HEC-1A and AN3-CA cell lines have low adhesive properties and are resistant to embryo implantation. RL95-2 cell lines have high

adhesive properties, but their culture is difficult as the cells tend to pile up[136–138]. For the purposes of this study, I decided to use the Ishikawa cells (IK) as an *in vitro* model of the endometrial epithelial cells.

Ishikawa cells

IK cells were established from an endometrial adenocarcinoma of a 39-year-old Japanese patient. These cells are polygonal in shape with irregular nuclei and small nucleoli containing large vacuoles[139,140]. IK cells also express functional receptors for oestrogen, progesterone, and androgens, all of which are important for the cellular changes required for the receptiveness of the endometrium. IK cells allow for attachment and invasion which are crucial steps in implantation. These properties were confirmed in mice, where mouse embryos were able to attach and invade the cells after 48 hours of coculture[141]. The adhesiveness is seen on the apical side of the cells, like *in vivo* endometrial epithelial cells. These cells are also capable of forming dome-like structures similar to pinopods in response to hormonal treatment, making them ideal models for *in vitro* studies[142]. From the results outlined in chapter 2, we decided to investigate the effects of rCD142 on hormonally treated IK cells.

Role of CD142 in pregnancy

CD142 (tissue factor) is a 47kDa transmembrane glycoprotein and a member of the class II cytokine/haematopoietic growth factor family. It is involved in the extrinsic coagulation cascade and is highly expressed by the placenta contributing to the hypercoagulability state of pregnancy[103]. Tissue factor has also been shown to play important roles in embryogenesis and is involved in the maintenance of blood vessels with roles in cell adhesion and migration. In tissue factor null mouse embryos, the disruption of the gene resulted in 90% lethality[108]. It has been shown to be expressed by stromal endometrial cells, and important for pregnancy. It is however

not expressed by epithelial cells and so its presence on the embryonic EEVs is interesting because it could be affecting the epithelial cells as these are the cells that would be contacted first.

In our study, IK cells were chosen as the model for the endometrial epithelial layer. These cells were first hormonally treated to resemble the *in vivo* environment, followed by treatment with rCD142 to delineate the effects of the protein on the cells. The use of recombinant protein was strategic to focus on the effects of tissue factor alone rather than having other factors expressed on EEVs adding to the effects.

Aims

The aims of this chapter were to **1)** hormonally treat Ishikawa cells to ensure they are a good model to use, **2)** confirm the uptake of EEVs by Ishikawa cells using fluorescent microscopy, and **3)** treat cells with recombinant human CD142 and measure the cellular proliferation as well as tyrosine phosphorylation of cellular proteins.

Materials and methods

Cell culture of Ishikawa cells

Ishikawa (IK) cells were obtained from Ms Janet Carver (NDWRH, University of Oxford) and were authenticated by the European Collection of Authenticated Cell Cultures (ECACC) in May 2017. The cells were thawed at 37°C and cultured in T-75 flasks with Dulbecco's modified eagle medium (DMEM) (Gibco) supplemented with 10% foetal bovine serum (Sigma-Aldrich), 1% penicillin/streptomycin (Sigma-Aldrich), 1% antibiotic-antimycotic (Gibco), and 2mM L-glutamine (Sigma-Aldrich) at 37°C, 20% O_2, and 5% CO_2. Media was changed every 2 days until approximately 80% confluency, after which trypsin (Sigma-Aldrich) was used to detach the cells. Complete medium was added to neutralize the trypsin. The cells were then collected and

centrifuged for 3 minutes at 1,300 x *g*. The concentration was calculated and the cells were transferred to new flasks/plates or dishes for treatment at the required seeding density.

Hormonal treatment of Ishikawa cells

To mimic the *in vivo* process of the proliferative stage of the menstrual cycle, the IK cells were treated with dibutyryl cyclic adenosine monophosphate (cAMP) (Sigma-Aldrich), and progesterone (Sigma-Aldrich). The cells were transferred to 6-well plates (for brightfield microscopy and cell lysate) or u-type glass bottom dishes (for confocal microscopy). Prior to the addition of cell suspension to the plates/dishes, 0.5% gelatin (Sigma-Aldrich) was added and incubated for 1 hour at room temperature. The cell suspension was then added at the appropriate volume. For hormonal treatment of cells, 0.5mM dibutyryl cAMP sodium salt and 1µM progesterone was added to the complete medium and used for culture. The medium was changed every 2 days and the treatment lasted for 7 days. The control media was supplemented with PBS and cultured in the same manner as the treated cells.

Brightfield microscopy

After hormonal treatment, the cells were imaged using the Olympus CKX53 microscope to determine the changes in the morphology of the cells. The images were captured and processed using the Teledyne lumenera INFINITY ANALYZE software in brightfield mode (version 7.0.26.16).

Phalloidin staining of F-actin

F-actin is a cytoskeletal component which is found in the organisation of higher structures. In order to visualise the F-actin in IK cells, after hormonal treatment in comparison to control cells, we used phalloidin-iFluor 488 conjugate (Abcam). The cells were cultured in u-type dishes and once treatment was complete, they were washed with 1mL of PBS, 3 times. After the wash, 400µL of

4% paraformaldehyde (PFA) (Alfa Aesar) was added to each dish and incubated for 20 minutes in order to fixate the cells. After fixations, the cells were washed 3 times with 0.1% PBS-T (PBS-Triton X) and incubated with 5% BSA in PBS-T (blocking solution) for 30 minutes. The phalloidin-conjugate was diluted 1:1000 using the blocking solution. Once the blocking was complete, 200µL of the diluted phalloidin-conjugate was added into each dish and incubated for 1 hour in the dark. The cells were then washed 3 times with PBS-T. Mounting medium containing DAPI (Abcam) was used and the cells were observed using the PerkinElmer UltraVIEW spinning disk system (6 laser line, CSU-X1 confocal head, Hamamatsu C9200-50 camera, and imaging with Zeiss Axiovert 200M optics). The cells were imaged at 10X objective. The images were captured and processed in Volocity acquisition software (Quorum Technologies).

Cell lysate

To make cell lysates, the cells were first washed 3 times with PBS. Trypsin (1mL) was added to each well of the 6-well plate and incubated for 5 minutes at 37°C. Once the cells had detached, 1mL of complete media was added to neutralise the trypsin. The cells were transferred to a falcon tube and spun at 1,300 x *g* for 3 minutes. The supernatant was discarded and the cells were resuspended in 10mL of PBS and spun at 1,300 x *g* for 3 minutes. The supernatant was discarded, and 60µL of radioimmunoprecipitation assay (RIPA) buffer (Thermo Scientific) containing a cocktail of protease (Thermo Scientific) and phosphatase inhibitors (Thermo Scientific) was added to the cells. The cells were incubated on ice for 15 minutes and then transferred to microcentrifuge tubes and mechanically dissociated using a 21-gauge needle, followed by centrifugation at 15,000 x *g* for 15 minutes at 4°C. The supernatant was removed and placed in a new tube (this was the cell lysate). BCA was performed to determine the concentration of the cell lysates.

Bicinichoninic acid protein assay

Protein content was measured using the BCA. BSA stock (1mg/mL) (Sigma-Aldrich) was diluted in fPBS to create the protein concentration standards (1mg/mL, 0.5mg/mL, 0.25mg/mL, 0.125mg/mL, and 0.0625mg/mL). Samples were diluted 1 in 10 in fPBS and 12.5µL of BSA standards or samples were added to a 96 well plate. All standards and samples were loaded in duplicates. To this, 100µL of BCA reagent which consists of BCA solution and copper II sulphate solution (50:1 ratio) (Thermo Scientific) was added to each well. The plate was incubated for 30 minutes at 37°C with gentle shaking. After incubation, the absorbance was read at 562nm using a FLUOstar OPTIMA plate reader. Proteins in the sample reduce Cu^{2+} to Cu^{+} in an alkaline solution, resulting in a purple colour, which is read by a spectrophotometer to extrapolate protein concentration based on the standards.

Western blot for MUC-1 expression

Mucin-1 (MUC-1) expression is expected to increase globally in endometrial epithelial cells in response to hormonal treatment. In order to check for the expression of MUC-1, we performed western blot analysis using cell lysates from hormonally treated and control IK cells. The blots were run under reducing conditions with 30µg of protein for each sample. The samples were mixed with 5µL of 4X Laemmli buffer (Bio-Rad) with 5% β-Mercaptoethanol (Sigma-Aldrich) which was further diluted to 25µL using double distilled water. The samples were heated at 95°C for 10 minutes and 20µL of each sample was loaded into the gel (NuPAGE 4-12% Bis-Tris Gel 1.0mm x 10 well gel by Invitrogen). Precision plus proteins dual colour standards was used as a marker for protein size (Bio-Rad). Electrophoresis was performed under non-reducing conditions at 150 V for 1 hour in a Novex minicell tank using NuPAGE MOPS SDS running buffer (Novex by Life Technologies). The proteins were then transferred to a polyvinylidene (PVDF) membrane (Bio-

Rad). The transfer apparatus was as follows: four filter paper sheets pre-soaked in buffer solution (300mM Tris, 10% methanol, pH 10.4) were placed in a Novex semi-dry transfer apparatus. Two filter paper sheets pre-soaked in 25mM Tris, 20% methanol, pH 10.4) were placed on top. The gel was transferred on top of the filter papers, followed by pre-soaked (100% methanol) PVDF membrane. On top of this, three filter paper sheets pre-soaked in 25mM Tris, 40mM ε-aminohexanoic acid, and 20% methanol was placed. The transfer was run at 25 V for 45 minutes. The membrane was then blocked using 5% milk (diluted in 0.1% Tris-buffered saline and 1% Tween-20) for 1 hour. The membrane was then incubated with monoclonal MUC-1 mouse antibody (1:1000 dilution, Invitrogen) overnight at 4°C with constant shaking. The following day, the membrane was washed with TBST, 3 times for 5 minutes each. The membrane was then incubated for 1 hour with anti-mouse immunoglobulin horseradish peroxidase (HRP-linked) secondary antibody (1:2000 dilution, Dako UK Ltd). The membrane was washed with TBST, as mentioned previously, followed by treatment with EZ-ECL chemiluminescence detection kit (Thermo Scientific) for 1 minute. The membrane was then developed using a gel documentation system (G-Box, Syngene) with the GeneSys software (version 1.6.7.0). Immunoblot normalization was carried out against total protein and band intensities were calculated using ImageJ (Java 1.8.0_172). The mean was calculated for the triplicates with SEM and an unpaired t-test with Welch's correction was used for calculating significance.

EEV isolation and PKH26 staining

Prior to staining of the EEVs, pooled embryonic samples were centrifuged at 150,000 x *g* for 2 hours (Beckman Optima XPN ultracentrifuge and SW 55 Ti swing-out rotor); the supernatant was removed and the pellet was washed in fPBS at 150,000 x *g* for 2 hours. The supernatant was removed and the pellet was resuspended in 15µL of fPBS. Due to the low volume and low protein

concentration of the samples, the entire samples were used for staining. The EEVs were stained using the PKH26 dye kit (Sigma-Aldrich). The samples were thawed at room temperature and Diluent C from the kit was added to the samples until a final volume of 1mL was reached. To each sample, 6µL of PKH26 dye was added and mixed continuously for 30 seconds by gentle pipetting. The mixture was left for 5 minutes at room temperature in the dark. The mixture was transferred to a 15mL falcon tube and quenched by adding 10% BSA. Serum-free cell culture medium was added to a final volume of 8.5mL and the solution was transferred to an ultracentrifuge tube. To this, 1.5mL of sucrose solution was added to the bottom of the tube and 3mL of serum-free medium was added to the top of the tube. The sample was centrifuged at 190,000 x *g* for 3 hours at 4°C (Beckman Optima XPN ultracentrifuge and SW 41 Ti swing-out rotor). The supernatant and interface layer were aspirated and discarded. The labelled S-EEVs were present in the pellet which was resuspended in 1mL of fPBS. This was transferred to an Amicon 100kDa filter column (Merck Millipore) and topped up with 4mL of fPBS. The column was spun at 3,000 x *g* until the volume of the upper chamber reached 50µL (this was the labelled S-EEVs). The labelled S-EEVs were then transferred to a microcentrifuge tube and stored at -80°C.

The stained EEVs were thawed and 1mL was added to a u-type glass-bottom dish containing hormonally treated Ishikawa cells and was incubated for 1 hour and 3 hours. After the incubation, the media was removed and the cells were washed in 4% heparin/PBS and gently shaken for a few seconds. The heparin was removed and the cells were washed with PBS, 3 times. Then 400µL of 4% PFA was added to the cells and incubated for 20 minutes in the dark at room temperature. The fixation solution was aspirated and the cells were washed with PBS, 3 times. Four hundred microlitres of 5µg/ml of WGA Alexa-Fluor 488 (Invitrogen) was added to the cells for 10 minutes at room temperature in the dark. The cells were rinsed with PBS 2 times, each for 5 minutes. One

drop of mounting medium containing DAPI was added onto the dish and the cells were observed using the confocal microscope PerkinElmer UltraVIEW spinning disk system (6 laser line, CSU-X1 confocal head, Hamamatsu C9200-50 camera and imaging with Zeiss Axiovert 200M optics) at 20X objective. Images were captured and processed in Volocity acquisition software (Quorum Technologies). Once I was able to prove that the IK cells were capable of up taking vesicles, I wanted to look further into the effects of recombinant tissue factor on the IK cells.

CyQUANT Proliferation assay

I started with first determining the standard proliferation of the cells without the addition of recombinant protein to produce a standard curve for the normal IK cellular proliferation using the CyQUANT proliferation assay (Invitrogen, 2397778). The CyQUANT proliferation assay measures the density of cells in culture. The kit uses a green dye (CyQUANT) which has high fluorescence when bound to cellular nucleic acids. To perform the proliferation assay, the reagents were prepared according to the manufacturer's instructions. In brief, the cell lysis buffer stock solution was diluted 20X using distilled water to get 1X lysis buffer. The CyQUANT GR solution was then diluted 400-fold in the 1X lysis buffer (working solution).

For the reference standard curve, trypsin was added and a cell suspension of $6.25x10^5$ cells was made; the cells were centrifuged at 200 x *g* for 5 minutes. The supernatant was discarded and the pellet was frozen overnight. The cells were thawed the next day and resuspended in 2.5mL of the working solution. A dilution series was created with 8 cell concentrations ($5x10^4$, $4x10^4$, $3x10^4$, $2x10^4$, $1x10^4$, $5x10^3$, $2.5x10^3$, and 0) using the working solution. To each well, 200µL of the cell suspensions were added (triplicates) and incubated for 3 minutes in the dark. The fluorescence was then measured using the OPTIMA-FLUOR plate reader with 480nm excitation and 530nm emission and a standard curve was created to convert fluorescence to cell number.

Normal cellular proliferation during hormonal treatment

I wanted to check whether the kit could measure the proliferation of cells over the period of hormonal treatment and whether it was suitable for experiments using rCD142 treatment extending beyond the 7 days of hormonal treatment. Also, I wanted to check whether the seeding concentration was appropriate and would stay within the detection limit of the kit. For this part of the experiment, cells were seeded at $0.1x10^4$ in four 96 well plates in triplicate wells. The background wells were also cultured the same way but did not have any cells. The cells were hormonally treated with 0.5mM cAMP and 1µM progesterone as described above. On day 1, the media was removed completely, and the wells were washed with 200µL of PBS. The PBS was removed completely, and the plate was frozen. The same procedure was repeated for the other plates on day 3, day 5 and day 7 of hormonal treatment. The proliferation was measured by adding 200µL of the working solution to each well, incubating in the dark for 3 minutes and measuring the fluorescence as described above. The background was subtracted from the experimental wells, and the mean fluorescence and cell number were calculated with SEM.

Dose-response and time-point for rCD142

To find the optimal conditions for treating the cells with rCD142, I began by using different concentrations of rCD142 (0.001µg/ml, 0.01µg/ml, 0.1µg/ml, and 1µg/ml) (Biolegend, 784406) in a dose-response experiment followed by a time-course experiment. The cells were first seeded onto two 96-well plate at $0.1x10^4$ and hormonally treated for 7 days as described previously. Control wells were cultured in the same way without the addition of rCD142. The background wells were cultured in the same way but did not have any cells. On day 7, one of the plates was gently washed with 200µL of PBS. The media was removed completely, and the plate was frozen. For the other plate, we treated the cells with the different concentrations of rCD142 (triplicates)

for 24 hours, after which the plate was washed gently with 200µL of PBS. The PBS was removed completely, and the plate was frozen overnight. The proliferation was measured by adding 200µL of the working solution to each well, incubating in the dark for 3 minutes and measuring the fluorescence. The background was subtracted from the experimental wells, and the mean fluorescence and cell number (using the standard curve) were calculated with SEM. An unpaired t-test with Welch's correction was used for calculating significance.

We then moved on to performing a time-course study looking at the proliferation at 0 hours, 6 hours, 12 hours, 24 hours and 48 hours after the addition of 0.1µg/ml of rCD142 (the concentration which caused the highest significant difference compared to control). Cells were seeded at $0.1x10^4$ in 5 plates in triplicate wells. Control wells were cultured in the same way without the addition of rCD142. The background wells were cultured in the same way but did not have any cells. After the 7 days of hormonal treatment, the first plate was gently washed with PBS. The media was removed completely, and the plate was frozen overnight. The same was repeated for the other plates at the corresponding time points after incubation with rCD142. The signal from the background wells was subtracted from the experimental wells. The mean fluorescence and cell number (using the standard curve) were calculated with SEM. The change in fluorescence and cell number was calculated relative to 0 hours of rCD142 treatment. An unpaired t-test with Welch's correction was used for calculating significance.

Tyrosine-phosphorylation western blots

I treated the cells with 0.1µg/ml for 10 minutes, 6 hours, 24 hours, and 48 hours. These time-points were chosen to ensure that I would not miss the tyrosine phosphorylation of proteins. Cells were seeded at $0.2x10^6$ into 6 well plates in triplicates for each timepoint. The cells were hormonally treated as described previously for 7 days followed by treatment using 0.1µg/ml of rCD142 for the

specified time points. Control cells were incubated for the same time without the addition of rCD142. Another set of control cells was removed and lysed right after decidualisation to account for the initial phosphorylation status of the cells after hormonal treatment, and this would be the baseline.

At the specified time, cells were washed with PBS, 3 times. The RIPA lysis buffer containing protease and phosphorylation inhibitors was added at 150µL per well and the plate was incubated on ice for 5 minutes. The cells were then scraped off and transferred to 1.5mL microcentrifuge tubes. They were incubated on ice for 15 minutes and resuspended using a 21-gauge needle and syringe. The tubes were then centrifuged at 15,000 x *g* for 15 minutes. The supernatant was removed and stored at -20°C for future experiments.

Western blot analysis was performed using anti-tyrosine phosphorylation mouse antibody clone 4G10 (EMD Millipore). Twenty micrograms of cell lysate were added to each lane and a non-reducing western blot was completed as described above. Having confirmed that rCD142 induced proliferation and tyrosine phosphorylation, I proceeded to a more global proteomic experiment outlined in the next chapter.

Results

Culture of Ishikawa cells

Ishikawa cells were chosen as a model of the epithelial cells of the endometrium. The cells were hormonally treated using dibutyryl cAMP and progesterone for 7 days to mimic the *in vivo* environment prior to implantation. The cells continued to proliferate during the 7 days with the treated cells forming dome-like structures similar to pinopods (Figure 3.1.1.A). Multiple structures could be seen in the lower magnification images, compared to the control cells which show none

of these structures (Figure 3.1.1.B). These results are similar to *in vivo* processes which occur prior to implantation with the emergence of pinopods.

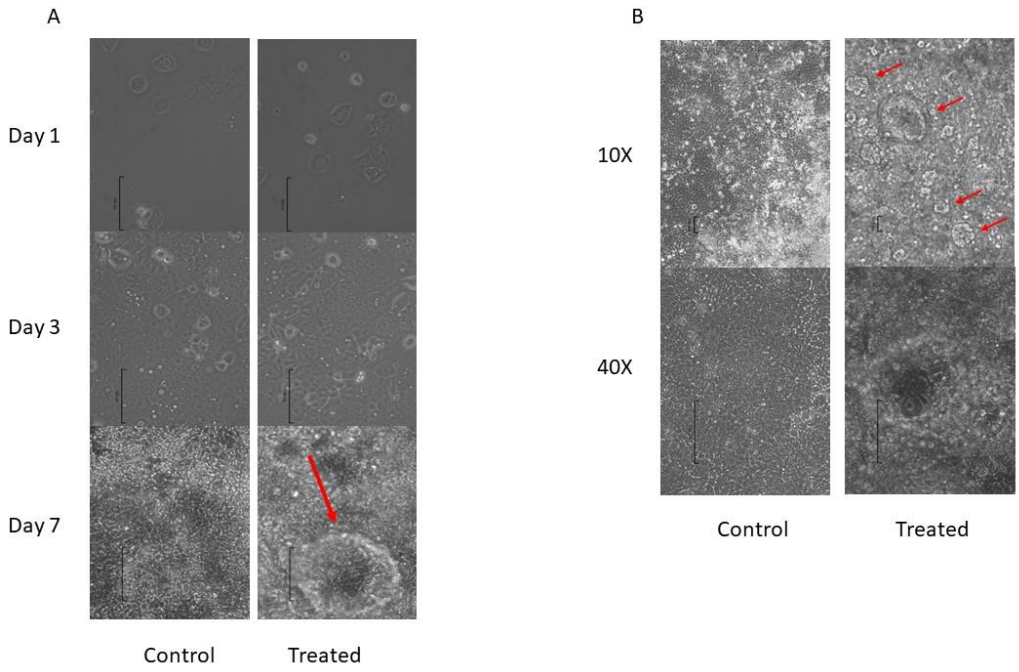

Figure 3.1.1: Representative brightfield images of Ishikawa cells treated with 0.5mM dibutyryl cAMP and 1µM progesterone compared to control cells. **A)** Images taken on day 1, day 3, and day 7 of culture comparing control cells to hormonally treated cells. **B)** Control cells vs treated cells imaged on day 7 of culture, comparing 10X magnification to 40X. The red arrows are pointing to the dome-like structures. The scale bars are 100µm in length. Cells were imaged using the Olympus CKX53 microscope. The images were captured and processed using the Teledyne lumenera INFINITY ANALYZE software in brightfield mode (version 7.0.26.16).

I then wanted to investigate the cells using fluorescent microscopy to have a better visual of the dome-like structures in comparison to the control cells. The cytoskeletal components of the cells were stained with phalloidin (green in colour), which stains F-actin specifically, and the nuclei were labelled with DAPI (blue in colour). The dome-like structures showed a higher aggregation of cells on the periphery and less cells in the middle with the dome-like structures being visible (Figure 3.1.2).

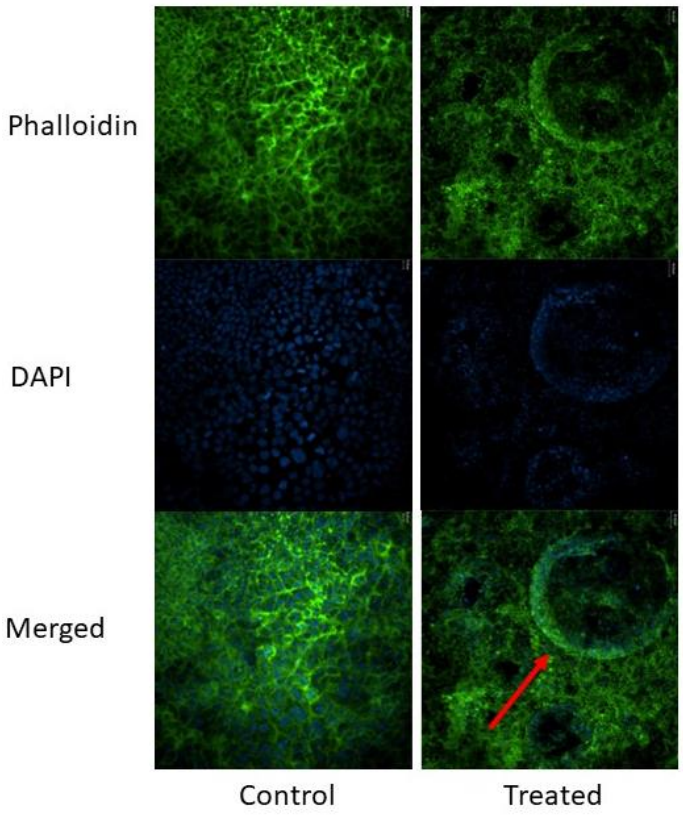

Figure 3.1.2: Representative confocal images of Ishikawa cells treated with 0.5mM dibutyrul cAMP and 1µM progesterone compared to control cells. Images were taken on day 7 of culture. The F-actin of the cytoskeleton is labelled with phalloidin in green. The nuclei of the cells are labelled with DAPI in blue. The merged images show both stains for each cell. The scale bars are 40µm in length. The microscope used was PerkinElmer UltraVIEW spinning disk system (6 laser line, CSU-X1 confocal head, Hamamatsu C9200-50 camera and imaging with Zeiss Axiovert 200M optics). The cells were imaged at 10X objective and the scale bar measures 21µm. The images were captured and processed in Volocity acquisition software (Quorum Technologies).

Another way to assess whether the cells were responding to the hormonal treatment, other than morphological changes, was to quantify the amount of MUC-1 expression by the cells. MUC-1 is expressed by endometrial epithelial cells and is upregulated at the time of implantation. However, at the site of implantation, MUC-1 expression is reduced. Since we were using an immortalized cell line and were not interested in identifying the site of implantation, we hypothesised that the hormonally treated cells would have an increase in the global expression of MUC-1. We performed western blot analysis and showed that the cells had a significantly higher protein fold change compared to the control cells (1.186±0.058 vs 0.809±0.103, p=0.0464) (Figure 3.1.3.B).

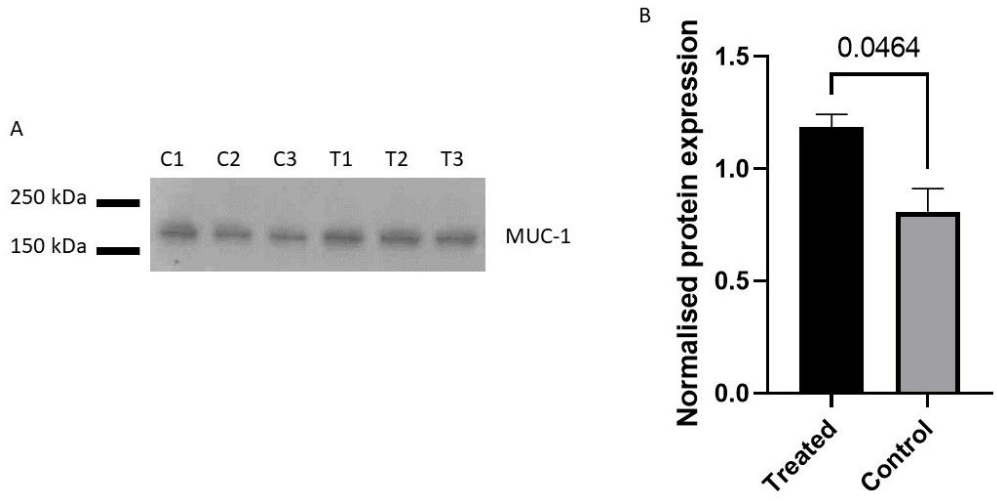

Figure 3.1.3: The expression of MUC-1 protein by hormonally treated cell lysates vs control cell lysates (n=3). **A)** Western blot image of the expression between the cells, C1-C3 represent the control cells and T1-T3 represent the treated cells. MUC-1 has a molecular weight of 160kDa. **B)** Normalised protein expression to total protein loaded, the mean and SEM are shown for each sample type. An unpaired t-test with Welch's correction was used for calculating significance.

Uptake of EEVs by Ishikawa cells

While studying the effect of EVs or EV proteins on cells, it is important to show that the cells are able to uptake EVs. Due to the low supply of embryonic samples, I decided to use all of the isolated EEVs. I stained the EEVs with PKH26 (red), the cell periphery with WGA (green) and the cell nucleus with DAPI (blue). The control cells were treated with PBS which is the carrier medium of the EEVs. The cells were incubated for 1 hour and 3 hours with the stained EEVs and fluorescent microscopy showed that the red EEVs were taken up by the cells at both time points (Figure 3.2). The overlay of the three channels shows that the red EEVs are mostly found within the cells, compared to the control, which did not have any red staining, suggesting that there was no autofluorescence in the red channel and that the red colour in the treated cells was coming from the stained EEVs (Figure 3.2). More vesicles were present after 3 hours of incubation compared to 1 hour (Figure 3.2). This confirmed that the cells are capable of up taking EEVs which increases over time.

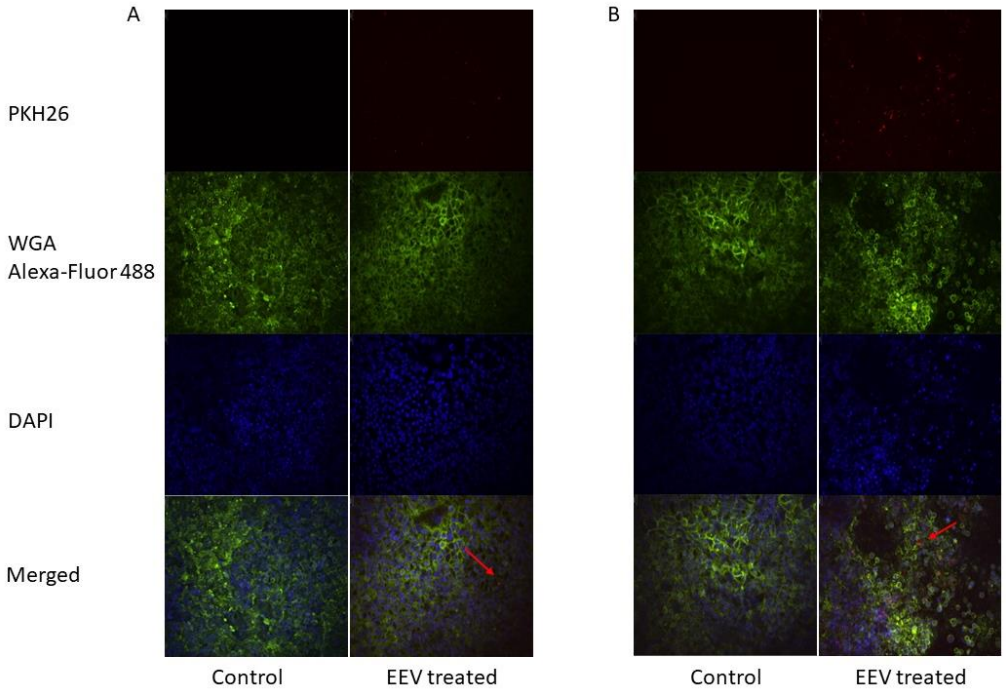

Figure 3.2: Representative confocal images of Ishikawa cells treated with PKH26 stained EEVs. **A)** Cells treated with PKH26 stained EEVs and PBS (Control) for 1 hour. **B)** Cells treated with PKH26 stained EEVs and PBS (Control) for 3 hours. Cells were observed using the confocal microscope PerkinElmer UltraVIEW spinning disk system (6 laser line, CSU-X1 confocal head, Hamamatsu C9200-50 camera and imaging with Zeiss Axiovert 200M optics) at 20X objective. The scale bars are 21µm in length. Images were captured and processed in Volocity acquisition software (Quorum Technologies).

Cellular proliferation of Ishikawa cells in response to rCD142 treatment

Since I was able to show that EEVs are able to enter Ishikawa cells, and CD142 was found to be commonly expressed on the EEVs (chapter 2), I wanted to check what effects rCD142 has on the cells. Tissue factor was identified as being commonly expressed by embryonic S-EEVs in the previous chapter and so we wanted to test the functionality of the protein. Tissue factor is involved in the coagulation cascade but interestingly also plays roles in proliferation, cell motility and angiogenesis. I wanted to investigate whether rCD142 had an impact on the proliferation of the cells, as this could be important during the process of implantation. To do this, I first looked at the normal proliferation rate of the cells within the 7 days of hormonal treatment using the CyQUANT assay to develop a standard curve for cell number measurements. A standard curve was produced

using serial dilutions of cells and was used to convert fluorescent values to cell number (Figure 3.3.1).

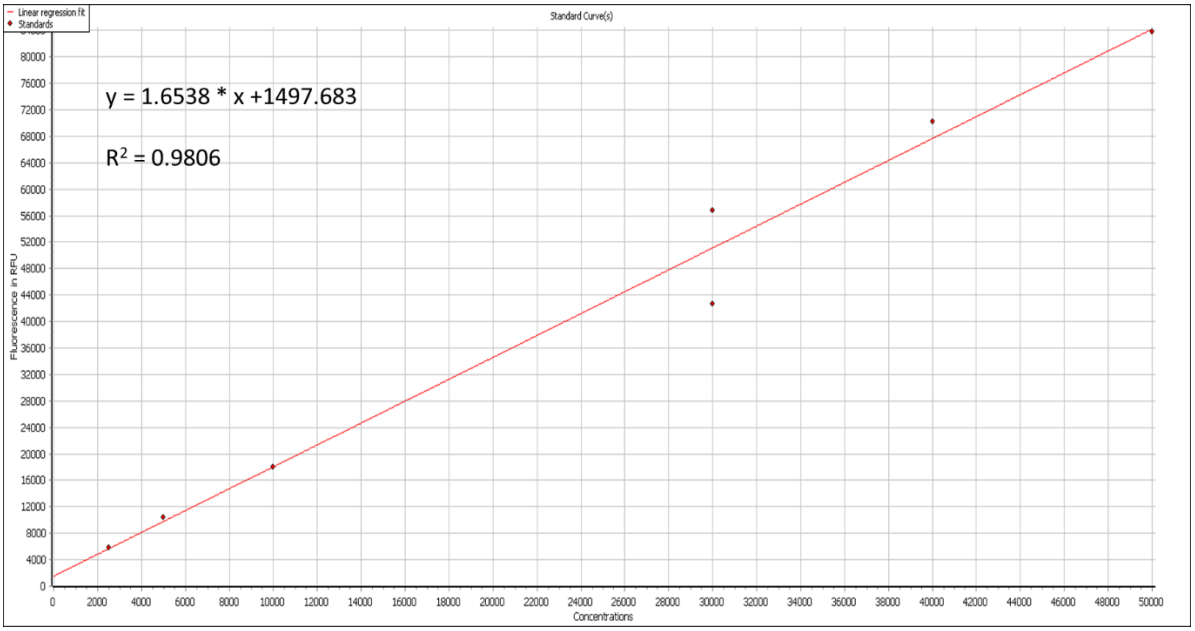

Figure 3.3.1: Standard curve correlating fluorescence of CyQUANT dye to the number of cells using the line of best fit. Eight known cell concentrations ($5x10^4$, $4x10^4$, $3x10^4$, $2x10^4$, $1x10^4$, $5x10^3$, $2.5x10^3$, and 0) were used to measure the fluorescence. The fluorescence was measured using the OPTIMA-FLUOR plate reader with 480nm excitation and 530nm emission wavelengths.

I was able to show that within a 7-day period, cellular proliferation could be measured by the kit and was within the limit of detection, suggesting that the seeding concentration was suitable for subsequent experiments (Figure 3.3.2.A). Over a 7-day period, there was an increase of 10,255.37±2892.14 cells compared to day 1 (Figure 3.3.2.B).

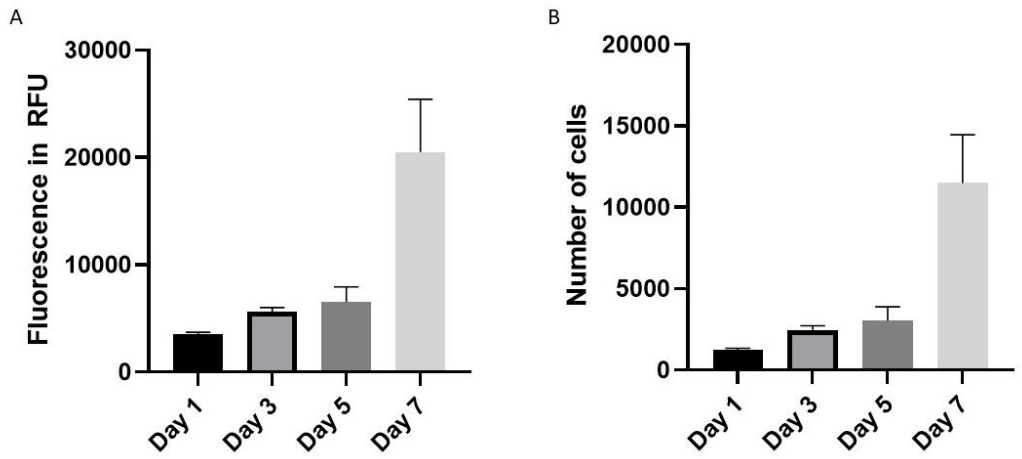

Figure 3.3.2: The proliferation of Ishikawa cells over the period of 7 days measured using the CyQUANT proliferation kit. **A)** The fluorescence of the cells was measured using the OPTIMA-FLUOR plate reader with 480nm excitation and 530nm emission wavelengths. **B)** The corresponding number of cells was calculated using the standard curve in figure 3.3.1. The mean and SEM are shown. Graphs were generated using GraphPad Prism (Version 9.3.1).

From this point, I decided to treat the cells with rCD142 and measure their proliferation in response to treatment. First, as already described, I hormonally treated the cells for 7 days using cAMP and progesterone. I then treated the cells with 4 different concentrations (0.001µg/ml, 0.01µg/ml, 0.1µg/ml, and 1µg/ml) of rCD142 to determine the optimal concentration for treatment. I treated the cells for 24 hours and showed that all of the concentrations increased cellular proliferation compared to the control (no rCD142 added) (Figure 3.4.1). This increase was significant for 0.001µg/ml ($p=0.0223$), 0.01µg/ml ($p=0.0012$) and 0.1µg/ml ($p=5.017x10^{-5}$). The other concentration of 1µg/ml did not show a significant increase in cell proliferation. Out of the three significant concentrations, 0.1µg/ml had the highest impact, with an increase of 3660.626±1639.305 cells compared to the control (Figure 3.4.1.B). Therefore, 0.1µg/ml was chosen for subsequent experiments.

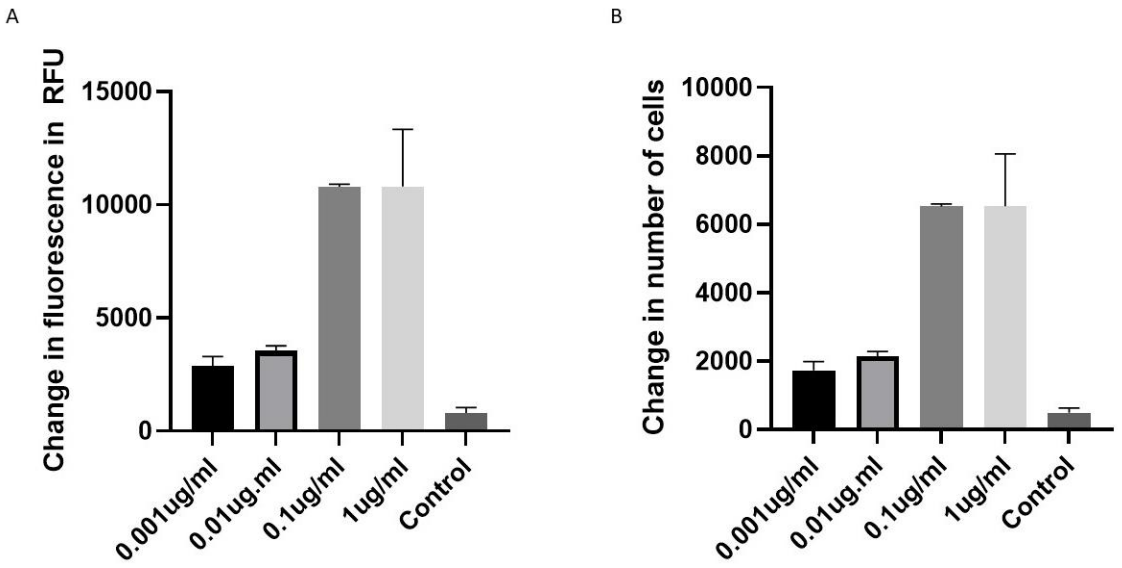

Figure 3.4.1: Dose-response of Ishikawa cells to recombinant CD142 measured by cellular proliferation. Four different concentrations of rCD142 were used. Control wells were cultured in the same way without the addition of rCD142. **A)** The fluorescence of the cells was measured using the OPTIMA-FLUOR plate reader with 480nm excitation and 530nm emission wavelengths. **B)** The corresponding number of cells was calculated using the standard curve. An unpaired t-test with Welch's correction was used for calculating significance. Graphs were generated using GraphPad Prism (Version 9.3.1).

I then wanted to find the optimal time point for the treatment of the cells. I treated the cells with rCD142 for 6 hours, 12 hours, 24 hours and 48 hours. The fluorescence and cell number were calculated and showed an increase in the treated cells compared to the controls beyond 0 hours (Figure 3.4.2.A-B). Although, 6 hours and 12 hours of treatment resulted in a decrease in cell number compared to 0 hours of treatment which suggested that the cells needed more time for the rCD142 to affect the rate of proliferation (Figure 3.4.2.C-D). After 24 hours of treatment, the change in cell number was significantly higher than the control with a difference of 7614.988±1722.16, p=0.0448 (Figure 3.4.2.D). After 48 hours of treatment, the change in cell number was also significantly higher in the treated cells compared to the control with a difference of 11685.02±1872.101, p=0.0271 (Figure 3.4.2.D). These results suggested that the cells are beginning to respond to the treatment after 24 hours which could be a good time point to use for the proteomics experiment outlined in the next chapter.

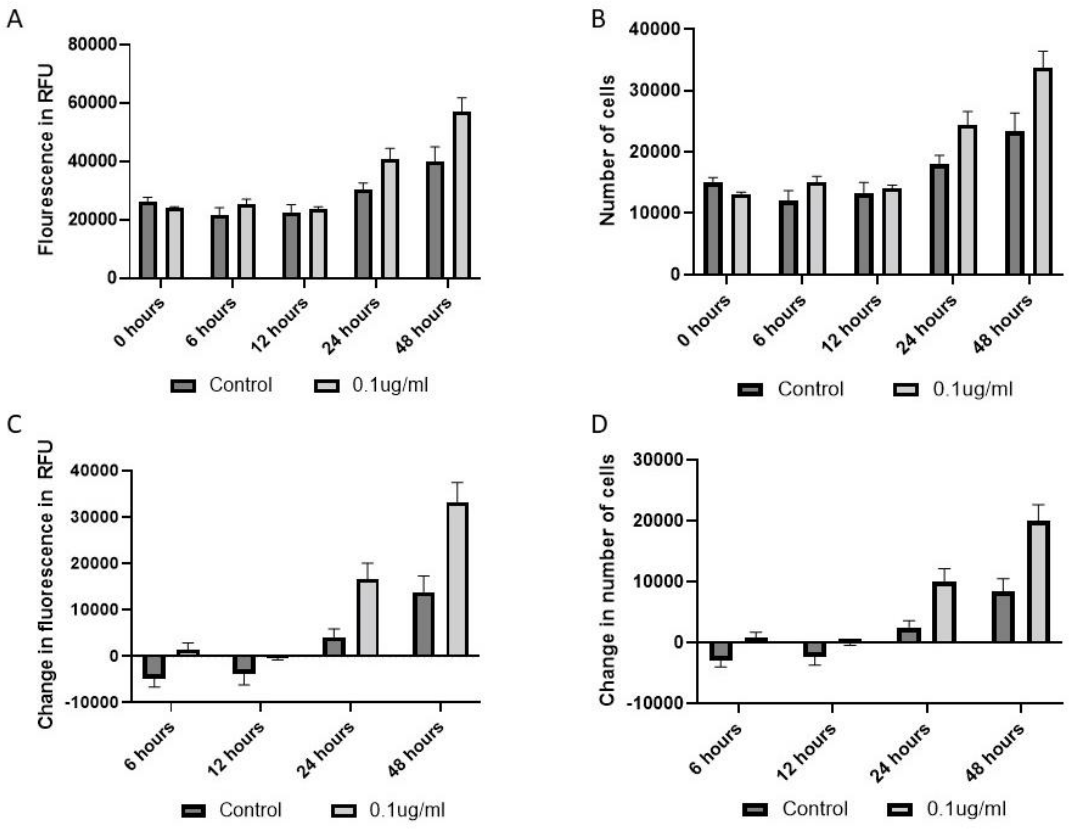

Figure 3.4.2: Treatment time course of Ishikawa cells treated with rCD142 and control. Cells were treated for 6, 12, 24 and 48 hours with 0.1µg/ml of rCD142 or PBS (control). The fluorescence and number of cells for each time point is shown in panel **A** and **B**. The change in fluorescence relative to 0 hours of treatment was calculated and is shown in panel **C** and **D**. The mean and SEM are shown. An unpaired t-test with Welch's correction was used to determine the significance of differences between the control and treated cells. Graphs were generated using GraphPad Prism (Version 9.3.1).

Tyrosine-phosphorylation of Ishikawa cellular proteins

When it comes to proliferation assays, to account for the loss of cells during washing steps and differences in the seeding of the cells, other methods are recommended to show that the cells are responding and becoming activated before further investigations. I decided to look at the tyrosine phosphorylation of multiple protein kinases in the sample. This was to corroborate the cell proliferation results. Cells were treated for 10 minutes, 6 hours, 24 hours and 48 hours with 0.1µg/ml of rCD142 and cell lysates were made for western blot analysis. Cells that were not exposed to rCD142 were collected at the same time points and used as controls to ensure the results

were caused by the protein treatment and not by the amount of time in culture. Anti-phosphotyrosine antibody was used to look at the global tyrosine phosphorylation of kinase proteins in the samples. The phosphorylation western blot showed a clear increased phosphorylation of the top band between the 24 hours treated cells compared to the 24 hours controls (C2) and 0 hours controls (C1) (Figure 3.5.A). This was normalised to the total protein loaded in each lane and showed an increase in the fold change for treated cells compared to 0 hours control (1.653±0.103 vs 1.041±0.052, p=0.0137), and 24 hours control (1.653±0.103 vs 0.941±0.034, p=0.0136) (Figure 3.5.B). However, the other time points did not show any differences in the tyrosine phosphorylation of the kinase proteins in the cells which again confirmed that 24 hours would be an optimal time point to investigate further.

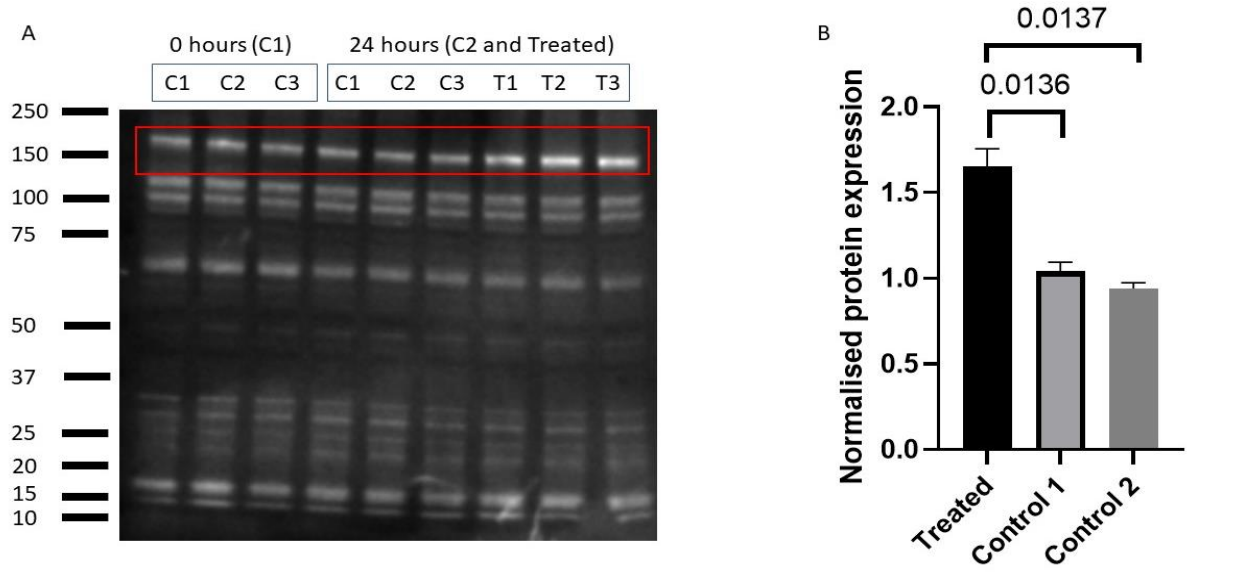

Figure 3.5: Western blot analysis of tyrosine phosphorylation of proteins in Ishikawa cells, treated with rCD142 for 24 hours and controls. **A)** Western blot comparing the untreated controls at 0 hours with cells treated for 24 hours with rCD142 and cells treated for 24 hours treated with PBS. The gel was developed using a gel documentation system (G-Box, Syngene) with the GeneSys software (version 1.6.7.0). **B)** Immunoblot normalization was carried out against total protein and band intensities were calculated using ImageJ (Java 1.8.0_172). The mean was calculated for the triplicates with SEM. Graphs were generated using GraphPad Prism (Version 9.3.1).

Nevertheless, the 24-hour phosphorylation blot was another proof that the cells were responding to the rCD142 treatment through specific pathways. From here, we decided to look at the global

proteome changes between the treated cells and control cells. The same cell lysates were sent for mass spectrometry at the Target Discovery Institute which is explained in detail in the next chapter.

Discussion

Ishikawa cells as a model of endometrial epithelial cells

With the identification and characterisation of CD142 on embryonic extracellular vesicles, my next aim was to investigate the functionalities of the protein and determine whether there are any physiologically relevant associations which could be impacting the endometrium. To study effects on the endometrium, cell lines are commonly used as models. Standard cell lines that are used for *in vitro* studies of the epithelial layer of the endometrium are HEC-1C, RL95 and Ishikawa (IK) cells. I wanted to focus on the epithelial layer as this would be the first contact point for the EEVs and where CD142 would have its first impact. The IK cells were ideal models because they have both oestrogen and progesterone receptors and respond to the hormone in a similar manner as *in vivo*, with the formation of dome-like structures similar to pinopod formation[140]. The IK cells also possess adhesive properties. I was able to confirm the formation of the dome-like structures as well as an increase in MUC-1 expression in response to hormonal (dibutyryl cAMP and progesterone) treatment. This finding was important to show that the cells resembled the *in vivo* environment and could be a good model to use. *In vivo*, MUC-1 plays an important role in the attachment of the embryo. MUC-1 hydrates cellular surfaces and acts as a protective barrier[143]. Locally, at the site of implantation, MUC-1 expression is reduced due to its anti-adhesive properties[144]. However, the overall expression of MUC-1 in the endometrium is increased, which I observed with our IK cell model. The importance of MUC-1 is shown in patients with recurrent implantation failure and idiopathic recurrent pregnancy loss where MUC-1 protein levels were significantly reduced[145].

The potential of EVs to communicate with the target cells is seen only if the EVs reach and are taken up by the target cells. I was able to show that the IK cells are capable of up taking EEVs through PKH26 staining of the EEVs and fluorescent microscopy. With PKH26 staining, removing free dye from the sample is important to ensure that the signal is coming from the EVs and not the dye. In our experiment, I used differential centrifugation to stain the sample, removing most of the free dye. I also used filters to remove any remaining free dye. However, it would be ideal to corroborate this in other ways. For example, when using other sources of EVs such as syncytiotrophoblast EVs from the placenta, a lot of research has been done to identify the mRNA cargo of the EVs and so one way of ensuring those EVs have entered the cells is by looking at the expression of specific mRNA that are unique to the EVs after treatment of cells with the EVs. However, due to a lack of knowledge on the cargo of EEVs, it is not yet plausible to use this method.

Effects of rCD142 treatment of Ishikawa cells

I was able to show that the addition of rCD142 resulted in an increase in proliferation in IK cells when compared to controls. CD142 is known for its function in the extrinsic coagulation cascade. It has also been shown to be involved in cellular proliferation, cell motility and angiogenesis. The proliferation of cells is simple to measure and more accurate than other physiological assays. One way that tissue factor is involved in proliferation of cells is through signalling through the protease activated receptor-2 (PAR2). The activation of PAR-2 results in the phosphorylation of protein kinase C α (PKCα) which then promotes the ERK1/2 and nuclear factor kappa B phosphorylation[110,146]. This regulates gene expression which would be linked to the proliferation of the cells. Phosphorylated PKCα could also activate MAPK which regulates gene expression and processes such as proliferation[146]. Our results confirmed that the addition of rCD142 caused an

increase in the proliferation of the IK cells. Proliferation would be beneficial in the process of implantation in several ways. Prior to implantation, the epithelial cells of the endometrium proliferate to allow for the *functionalis* layer to become responsive to the embryo. During the window of implantation, the luminal epithelial cells of the endometrium do not undergo proliferation, however the glandular cells proliferate to prepare for the decidualisation of the stromal cell layer underneath. Since IK cells are a mixture of luminal and glandular cells, and CD142 is known to increase proliferation, we expected the cells to proliferate at a higher rate in response to treatment. To the best of our knowledge, we are the first to show that the addition of rCD142 results in the proliferation of endometrial epithelial cells, specifically IK cells.

Other than the increased rate of proliferation, I wanted to delineate the biochemical marker which could be responsible for the proliferation of the cells after the addition of rCD142 and I decided to look at the phosphorylation of the tyrosine kinase receptor proteins. I was able to show that after 24 hours of treatment with rCD142, there was an increase in the phosphorylation of the western blot band at approximately 170kDa which is representative of where platelet-derived growth factor receptor (PDGFR) would be. PDGFR is a member of the tyrosine kinase receptor family which are known for their roles in biological processes such as proliferation. PDGFR induces proliferation through PLCγ pathway which causes an increase in intracellular calcium and ultimately proliferation and cell mobility[147]. PDGFR also activates the Ras and MAPK pathway leading to the regulation of transcription factors, proliferation, differentiation and survival of cells. One possible link between the addition of rCD142 and activation of PDGFR would be that tissue factor has been shown to increase the levels of PDGF-CC which is a ligand that activates PDGFR resulting in the protein tyrosine phosphorylation of PDGFR and subsequent intracellular signalling[148]. Tissue factor could also increase the levels of PDGFR itself through the activation

of NF-kappa B and subsequent genes[146]. The activation of PDGFR and MAPK have also been shown to increase the production of tissue factor which confirms that there is a link between the two proteins and that this pathway could explain the biochemical response of the cells leading to proliferation[149]. Due to the observed changes, I decided to focus on the cells' total proteome change, which is discussed in detail in the next chapter.

Limitations and Future work

There were a few limitations with our cellular study. I only used the endometrial epithelial cells, whereas it would be interesting to look at a coculture system including both epithelial cells as well as stromal cells to represent a more physiologically relevant model. The stromal cells may also be changing in response to the rCD142 since it is an important player in coagulation and angiogenesis which are both processes involved in the invasion of the embryo into the stroma. It would also be interesting to look at the effects of rCD142 on primary endometrial cells. Another limitation was that the concentration of rCD142 that was used, although optimal for the cellular model, is not representative of the *in vivo* concentration. Although, it would be difficult to replicate the *in vivo* concentration using an *in vitro* cell line model. Our findings confirm that CD142 would be a good candidate to consider in taking forward and investigating the proteomic changes that are induced as a result of rCD142 addition to Ishikawa cells (outlined in chapter 4).

Chapter 4

Proteomics analysis of Ishikawa cells after treatment with recombinant CD142 protein

Introduction

Proteomics

Proteomics analysis allows for the identification and quantification of proteins found in a biological system such as cells and investigating the proteome also allows for the analysis of post-transcriptional and post-translational modifications[150]. Whereas analysing the genome of a system may be useful, the proteome gives a more in depth understanding of the conditions of the system at a specific time by identifying the proteins expressed at that time in response to a specific treatment[151]. Therefore, proteomics is more dynamic and is a good starting point for understanding cellular responses.

Types of proteomics

Proteomic studies are commonly performed on cellular models to represent the *in vivo* conditions. The most common form of proteomics is through mass spectrometry with subsequent bioinformatics analysis of the results. There are different types of proteomics which focus on identifying different aspects of proteins. Expression proteomics looks at the differential expression of proteins between two conditions[152]. Structural proteomics looks at the 3-D structure of functional proteins[153]. Lastly, functional proteomics determines the protein function and interaction with other proteins[154]. For this experimental chapter, expression proteomics was used to determine the differences of protein expression between Ishikawa cells treated with rCD142 and control cells.

Process of expression proteomics

The process begins with the extraction of proteins, followed by protein digestion, protein separation, mass spectrometry and data analysis. Once the proteins have been extracted from the cells through enzymatic and/or mechanical methods, they are enzymatically digested into peptides

which are then passed through a mass spectrometer. The spectra identify the mass and charge of the peptides along with peptide intensity which is used to determine the identity of the protein by matching the spectra to a sequence database[150]. The results are then configured into a list of proteins with corresponding intensities which can be used to perform downstream analysis. Using established methods and packages such as microarray analysis packages, the differentially expressed proteins (DEPs) between treatment and control groups can be determined. These proteins can then be investigated further to understand their role in biological processes with the potential for disease diagnosis/treatment options. The DEPs can then be validated, in the cellular system under study, through methods such as western blot analysis. Once the DEPs are identified using mass spectrometry and confirmed to be differentially expressed using another method, they can be studied further and potentially incorporated into clinical care.

From the previous experiments in this thesis, described in detail in chapters 2 and 3, I identified CD142 (tissue factor) as being commonly expressed by embryonic extracellular vesicles (EEVs). To understand whether tissue factor (TF) could be playing a role in endometrial receptivity in preparation for embryo implantation, I chose Ishikawa cells as a cellular model. The addition of recombinant TF (rCD142) resulted in an increase in the proliferative ability of the cells as well as changes in the tyrosine phosphorylation of specific kinase proteins (outlined in chapter 3). To better understand the cellular pathways that may be altered in response to the treatment of the cells with rCD142, I decided to perform proteomics analysis of the cells. Analysing the proteins found in the cells and identifying the differences between treated and control cells gives a unique insight into the regulatory network of the proteins in response to rCD142. This would be a good starting point to investigate the effects of TF on primary endometrial cells and its mode of action with

possible roles in implantation of the embryo. It would also pave the way for looking at other factors that could work in conjunction with TF to enhance or ameliorate its effects.

Tissue factor

TF is a glycoprotein with an extracellular, transmembrane, and cytoplasmic domain. The extracellular domain is involved in the extrinsic coagulation cascade through binding FVIIa[155]. The transmembrane portion anchors the protein into the membrane of the cell. The cytoplasmic domain is involved in signalling which affects cellular processes such as angiogenesis and cell proliferation[156]. It is common practice to use recombinant TF for functional studies. Due to the challenges associated with isolating TF expressing EEVs (as a result of the low starting volume), and wanting to understand the specific role of TF protein (and to exclude the contribution of other components of EEVs), recombinant TF was used for the treatment of Ishikawa cells and proteomics analysis.

Aims

The aims of this chapter are to **1)** determine the differentially expressed proteins (DEPs) between cells treated with rCD142 and control cells, **2)** identify potential roles the DEPs may be involved in which could be important for implantation/pregnancy, and **3)** compare the DEP package to significance A/B analysis to determine which method is best suited for the analysis of the obtained dataset to be used for future studies.

Materials and methods

Treatment of cells with recombinant tissue factor

Cells were seeded at $0.2x10^6$/well into a 6 well plate. The cells in all 6 wells were hormonally treated as described in chapter 3 for 7 days using 0.5mM cAMP and 1uM progesterone. Three wells of the plate were treated with 0.1µg/mL of rCD142 protein for 24 hours which was

determined to be the optimal concentration and time point for treatment as outlined in chapter 3. The other three wells were used as control cells which were incubated for the same time with complete media with fPBS (carrier media for the rCD142). After 24 hours, the cells were washed with PBS a total of 3 times. The radioimmunoprecipitation assay (RIPA) lysis buffer containing protease and phosphorylation inhibitors was added at 150µL per well and the plate was incubated on ice for 5 minutes. The cells were then scraped off and transferred to 1.5mL microcentrifuge tubes. They were incubated on ice for 15 minutes and resuspended using a 21-gauge needle and syringe. The samples were then centrifuged at 15,000 x *g* for 15 minutes. The supernatant was removed and the protein concentrations were determined using bicinchoninic acid assay (BCA) after which the samples were stored at -20°C until transfer to the Target Discovery Institute (TDI) for proteomics analysis. For each sample, a minimum of 150µg of protein was sent to TDI.

Protein preparation and digestion for mass spectrometry

The samples were prepared at TDI with the filter aided sample preparation (FASP) kit from Sartorius (Vicacon500, VN01H22 30kD) according to the manufacturer's instructions. In brief, 20µg of the sample was diluted in 200µL of 8M urea and loaded onto the filter. To denature the proteins, 100µL of 8M urea in 100mM triethylammonium bicarbonate (TEAB) was added and the samples were incubated for 30 minutes at room temperature. To this 10mM tris-carboxy-ethyl-phosphine (TCEP) was added and incubated for a further 30 minutes at room temperature. After the incubation, 50mM of iodoacetamide (IAA) was added and the samples were incubated for 30 minutes at room temperature in the dark. The columns were then centrifuged at 14,300 x g for 10 minutes and washed twice with 200µL of 50mM TEAB. The sample was again spun at 14,300 x g for 10 minutes.

The protein digestion was achieved by using trypsin. The sample was moved to a new tube and 200µL of 50mM TEAB containing 2µg of trypsin was added and incubated at 37°C. The samples were spun and the flow-through was kept. The filter was washed with 200µL 0.1% trifluoroacetic acid (TFA). The flow-through was combined with the previous one. The filter was rewashed with 200µL of 50% acetonitrile (AN) in 0.1% TFA. The flow-through was collected and combined. The samples were then dried using speedvac and resuspended in 21µL of 5% formic acid (FA) and 5% DMSO.

QExactive method TDI -DIA

The resuspended peptides were trapped on an Acclaim™ PepMap™ 100 C18 HPLC Columns using 0.1% FA in water at a pressure of 60 bar and separated on an Ultimate 3000 UHPLC system (Thermo Fisher Scientific) coupled to a QExactive mass spectrometer (Thermo Fisher Scientific). The peptides were separated on an Easy Spray PepMap RSLC column (Thermo Fisher) and then electro-sprayed directly into a QExactive mass spectrometer (Thermo Fisher Scientific) through an EASY-Spray nano-electrospray ion source (Thermo Fisher Scientific) using a linear gradient. The raw data was acquired in the mass spectrometer in a data-independent mode (DIA). Full scan MS spectra were acquired in the Orbitrap.

MS data search

The DIA-NN 1.8 software was used for identifying the proteins in the dataset. The database used was UPR_Homo sapiens_9606_UP000005640_20200803.fasta with the following settings: Library-free search, deep learning-based spectra, RTs and IMs prediction, maximum number of protease (trypsin) missed cleavages set to 1, PTMs: Carbamidomethyl (C), Oxidation (M), false discovery rate (FDR) set to 0.01 (1%), and MBR (match between runs). The data were collated

into excel files used for downstream analysis to identify DEPs between the control and treatment samples.

Data analysis using DEP package

The data were first analysed using the DEP package on R (version 4.2.1) which has been developed for proteomics analysis of mass spectrometry for differential enrichment based on linear models and empirical Bayes statistics. The package was used according to the analysis tutorial. First, the number of proteins in each sample was identified, followed by the filtering of the data based on missing values. The threshold was set as proteins that were identified in all replicates of at least one group. The distribution of the samples was checked to determine whether they were normally distributed and were normalised to remove any outlier effects before further analysis.

The missing values in the data set needed to be imputed for further analysis using either the missing at random (MAR) or missing not at random (MNAR) imputation methods. MAR refers to when the missing data is dependent on the observed measurements and the data are missing at random from all samples. MNAR is when the missing data is related to events that are not measured by the researcher and specific proteins are missing from specific samples. For my data set, I decided to look at both MAR and MNAR with different imputation methods within each category to explore the differentially expressed proteins and find commonalities between all the different methods. Within the MAR method, both k-nearest neighbour (kNN) and maximum likelihood estimation (MLE) imputations were analysed.

Within the MNAR method, manual (man), probabilistic minimum (MinProb), and quantile regression imputation of left-censored data (QRILC) were analysed. For each input method, principal component analysis (PCA) plots were generated to determine the batch effect between the two groups and whether there were consistencies between the different input methods of the

MAR and MNAR. The correlation matrix was used to identify the similarities between each sample type using the different imputation methods.

The next step was to produce a heatmap for the differentially expressed proteins using the different imputation methods. The similarly behaving proteins were clustered together using k-clustering. The volcano plots were generated to show the differentially expressed proteins between the control and treated samples with a cut-off for the protein fold change of 1.5 and a false discovery rate (FDR) of less than 0.05. The plots were generated for each of the imputation methods.

The p-value and adjusted p-value were identified for each protein, comparing the expression in the control to the treated samples. The adjusted p-value when calculated using the DEP package was showing values smaller than the p-values which is unexpected, suggesting that the DEP package was not suitable for this dataset. After investigating the DEP package further, it was concluded that the package uses the t-statistics-derived FDR rather than the conventional p-value-derived FDR. The calculations for the adjusted p-values were then changed to properly calculate the adjusted p-value using the Benjamin-Hochberg p-value-derived FDR. With this adjustment to the computation, the new list of differentially expressed proteins was identified.

Data analysis using outlier significance A/B

Another way to analyse the data is by using the significance A/B method of analysis. With the DEP analysis, the results are generated by comparing 'within replicate group variance' to 'between replicate group variance' using an approach similar to the t-test. This approach compares if a single protein in the treatment group is up or down regulated compared to the control group. It does not take into consideration if the protein is differentially expressed among the entire protein data set. The new technique using significance A/B determines the protein ratio compared to the distribution of all the protein ratios by taking signal intensities into account. Using this method,

highly abundant proteins have a more focused statistical spread compared to the low abundant proteins. Therefore, a protein with a ratio of 2 will be very significant if it is highly abundant and only marginally significant if present at low abundance. On the plot, the spread of the points is lower at high abundance which indicates that the quantification is more precise. The DEPs were calculated by checking whether a protein acted as an outlier using a normal distribution. Log protein ratios were calculated followed by the outlier significance score (significance A). The estimate of standard deviation (SD) from the main distribution was calculated using the 15.87, 50 and 84.13 percentiles represented by r_{-1}, r_0, and r_1. As a result, r_1-r_0 and r_0-r_{-1} are right side and left side SD, respectively. To measure the significant outlier ratio (3), the distance of r in relation to r_0 and r_{-1} is calculated in terms of the right SD (1) and left SD (2) respectively. Significance B was then calculated in the same manner as significance A with the addition that the proteins were divided into 10 equal groups before calculating the significance. I used the significance B to generate the figure of the DEPs with a Benjamin-Hochberg corrected p-value and threshold of 0.05.

$$z = \frac{r - r_0}{r_1 - r_0} \quad (1)$$

$$z = \frac{r_0 - r}{r_0 - r_{-1}} \quad (2)$$

$$significance\ A = \frac{1}{2} erfc\left(\frac{z}{\sqrt{2}}\right) \quad (3)$$

The DEPs with a log fold change of 1 and adjusted p value of less than 0.0001 were chosen for STRING analysis (online, version 11.5).

Western blot validation of proteomics results

In order to validate the proteomics results, I performed western blot analysis to check the expression of BRD2 of cell lysates from IK cells treated with rCD142, control cells (not treated with rCD142) and HEK293 cell lysates (positive control). The blot was run under reducing conditions with 30µg of protein for each sample. The samples were mixed with 5µL of 4X Laemmli buffer (Bio-Rad) with 5% β-Mercaptoethanol (Sigma-Aldrich) which was further diluted to 25µL using double distilled water. The samples were heated at 95°C for 10 minutes and 20µL of each sample was loaded into the gel (NuPAGE 4-12% Bis-Tris Gel 1.0mm x 10 well gel by Invitrogen). Precision plus proteins dual colour standards was used as a marker for protein size (Bio-Rad). Electrophoresis was performed under non-reducing conditions at 150 V for 1 hour in a Novex minicell tank using NuPAGE MOPS SDS running buffer (Novex by Life Technologies). The proteins were then transferred to a polyvinylidene (PVDF) membrane (Bio-Rad). The transfer apparatus was as follows: four filter paper sheets pre-soaked in buffer solution (300mM Tris, 10% methanol, pH 10.4) were placed in a Novex semi-dry transfer apparatus. Two filter paper sheets pre-soaked in 25mM Tris, 20% methanol, pH 10.4) were placed on top. The gel was transferred on top of the filter papers, followed by pre-soaked (100% methanol) PVDF membrane. On top of this, three filter paper sheets pre-soaked in 25mM Tris, 40mM ε-aminohexanoic acid, and 20% methanol were placed. The transfer was run at 25 V for 45 minutes. The membrane was then blocked using 5% milk (diluted in 0.1% Tris-buffered saline and 1% Tween-20) for 1 hour. The membrane was then incubated with monoclonal BRD2 mouse antibody (1:1000 dilution, Invitrogen) overnight at 4°C with constant shaking. The following day, the membrane was washed with TBST, 3 times for 5 minutes each. The membrane was then incubated for 1 hour with anti-mouse immunoglobulin horseradish peroxidase (HRP-linked) secondary antibody (1:2000

dilution, Dako UK Ltd). The membrane was washed with TBST, as mentioned previously, followed by treatment with EZ-ECL chemiluminescence detection kit (Thermo Scientific) for 1 minute. The membrane was then developed using a gel documentation system (G-Box, Syngene) with the GeneSys software (version 1.6.7.0). Immunoblot normalization was carried out against total protein and band intensities were calculated using ImageJ (Java 1.8.0_172). The mean was calculated for the triplicates with SEM and an unpaired t-test with Welch's correction was used for calculating significance.

Results

The DEP package was first used for the analysis of the data. The first step was to identify the number of proteins that were detected in each sample which was over 6,000 proteins in all 6 samples (Figure 4.1.1.A). We then wanted to identify how many missing values were present in each sample and the data showed that almost all proteins were identified in all 6 samples (Figure 4.1.1.B). The total number of proteins before filtering the data was 6,636. From here, the missing values were filtered out to keep the proteins that were identified in all the replicates of at least the control or treatment group. The data quality was good and as is shown in figure 4.1.1.C, only a small portion of the proteins were filtered out, with 6,507 proteins remaining after the filtration. The data distribution within all samples was normal (Figure 4.1.1.D top panel). However, I normalised the data regardless, to ensure that any outlier effects were disregarded for the rest of the analysis (Figure 4.1.1.D, bottom panel).

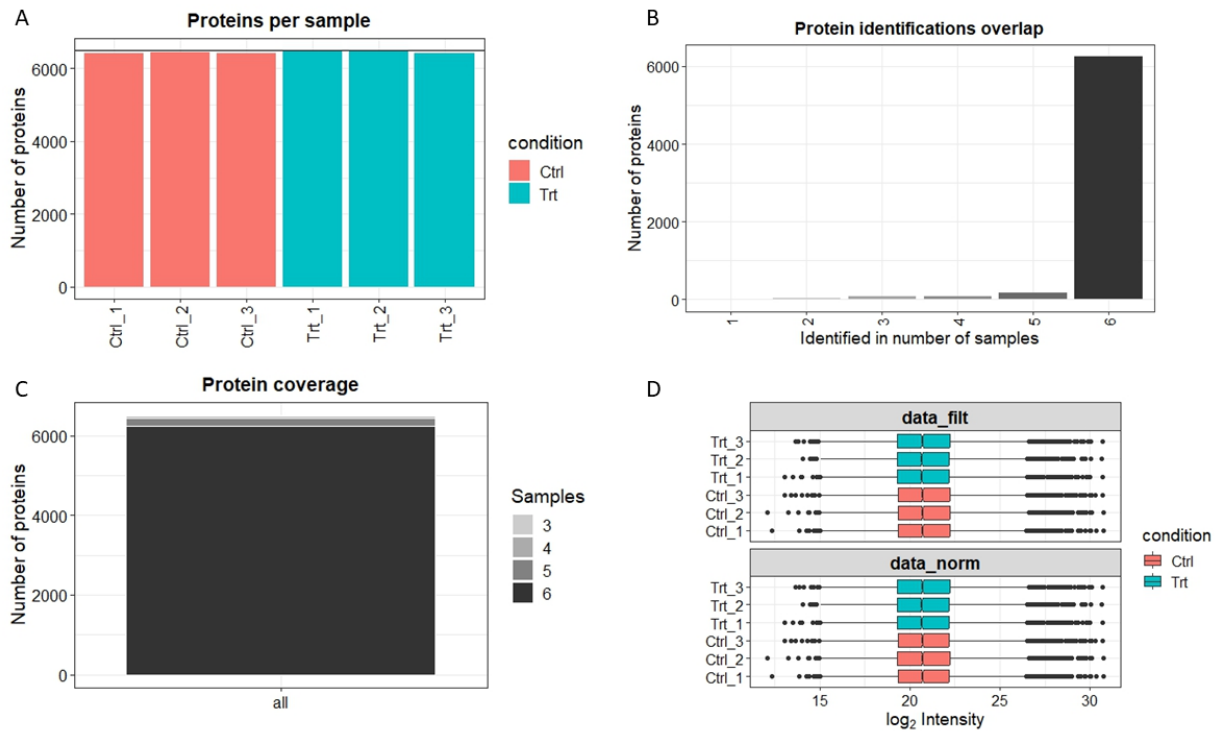

Figure 4.1.1: Data representation using the DEP package on R (version 4.2.1). **A)** The number of proteins present in each sample type with pink representing control samples and blue representing treatment samples. **B)** The number of proteins identified in the number of samples to determine missing values in each sample. **C)** The number of proteins that were present in all samples after filtering the missing values. **D)** Data normalisation to account for outliers within the samples. Data_filt is the data distribution after filtering the missing values. Data_norm is the data after normalisation.

Next, I needed to determine whether the missing data were missing at random (MAR) or missing not at random (MNAR). Looking at the pattern of the missing values (Figure 4.1.2.A), it is apparent that there are only a few proteins missing in some samples and even fewer proteins missing in all 3 replicates of either control or treatment samples suggesting that maybe the results are missing at random. On the other hand, when looking at the densities and cumulative fractions for proteins with and without missing values (Figure 4.1.2.B), the figure shows that proteins with missing values have lower intensities which suggests that the data may possibly be missing not at random. From the data distributions, I decided to look at both MAR and MNAR with the different imputation methods for each.

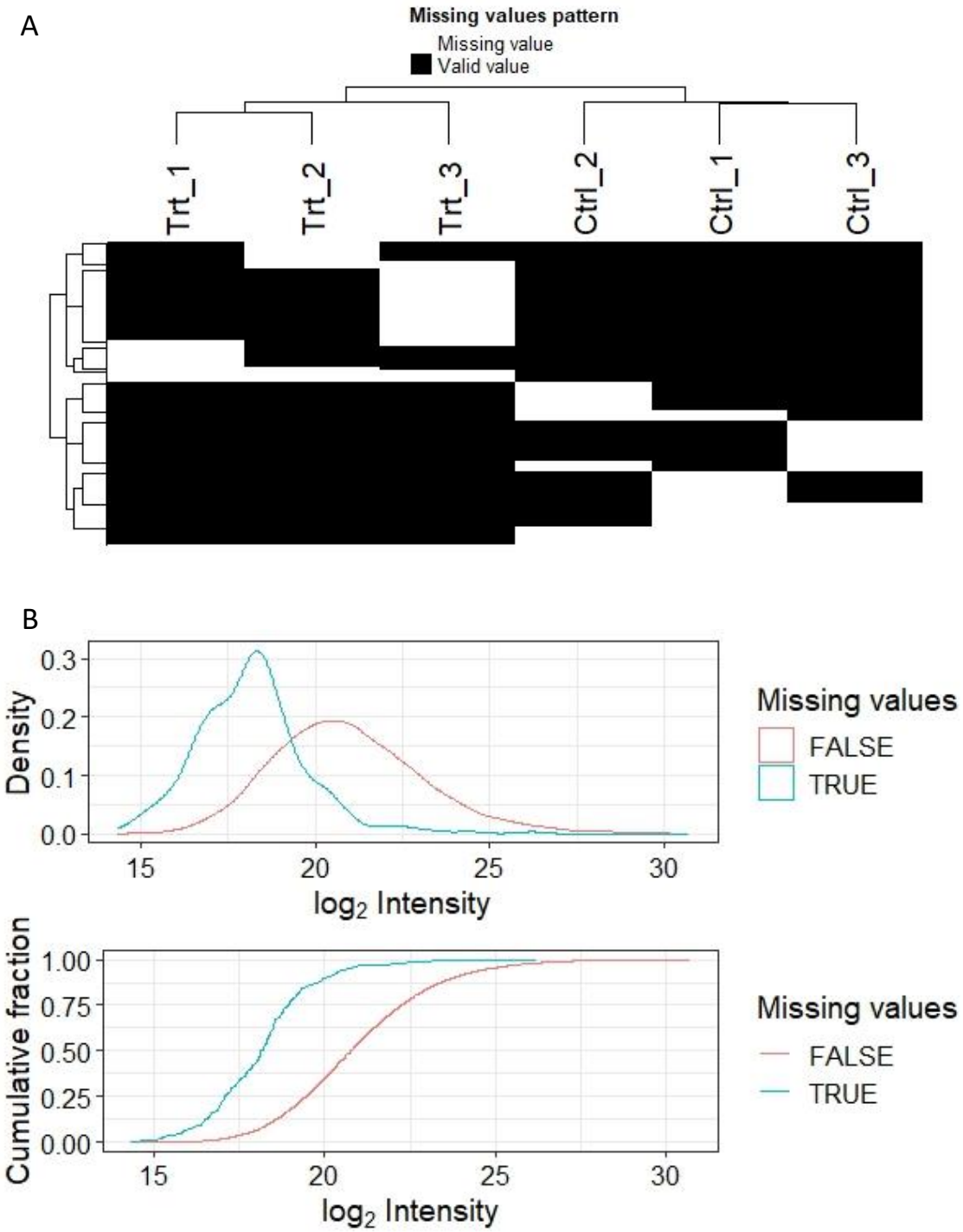

Figure 4.1.2: Pattern of missing values, density and cumulative fraction of the dataset. **A)** The pattern of missing values within the sample types. The white spaces correspond to that sample type having missing values and the black bars correspond to no missing values within that sample type. **B)** The intensity of protein expression for those with missing values and those present in all sample types. Figures produced using DEP package in R (version 4.2.1).

Missing not at random DEP analysis

Three different imputation methods were used for MNAR analysis, 'man', 'QRILC' and 'MinProb'. The PCA plots and correlation matrix for all three methods show that the control and treatment groups can be separated nicely along PC1 (Figure 4.1.3). Slightly different proteins were identified using the three methods, as seen on the heatmap (Figure 4.1.4) and volcano plots (Figure 4.1.5). The 'man' method showed the smallest number of DEPs (40 proteins) (Table 4.1); 'MinProb' identified 51 DEPs (Table S4.2) and 'QRILC' identified 47 DEPs (Table S4.3).

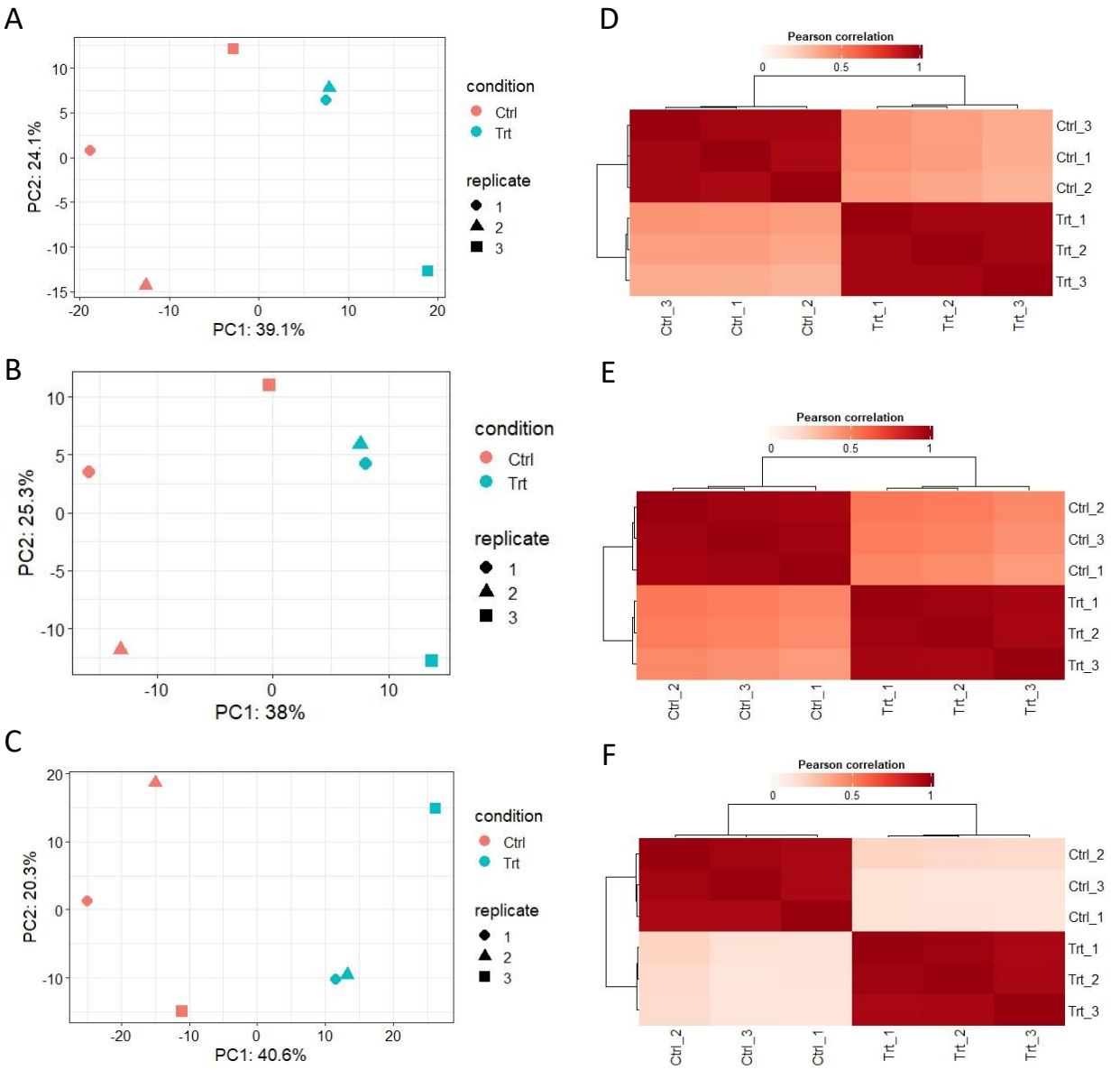

Figure 4.1.3: Missing not at random imputation methods. **A)** Principal component analysis (PCA) plot using the 'man' imputation method of analysis. **B)** PCA plot using the 'MinProb' imputation method. **C)** PCA plot using the 'QRILC' imputation method. **D-F)** the correlation matrix of the three imputation methods 'man', 'MinProb' and 'QRILC', respectively. The figures were generated using DEP package on R (version 4.2.1).

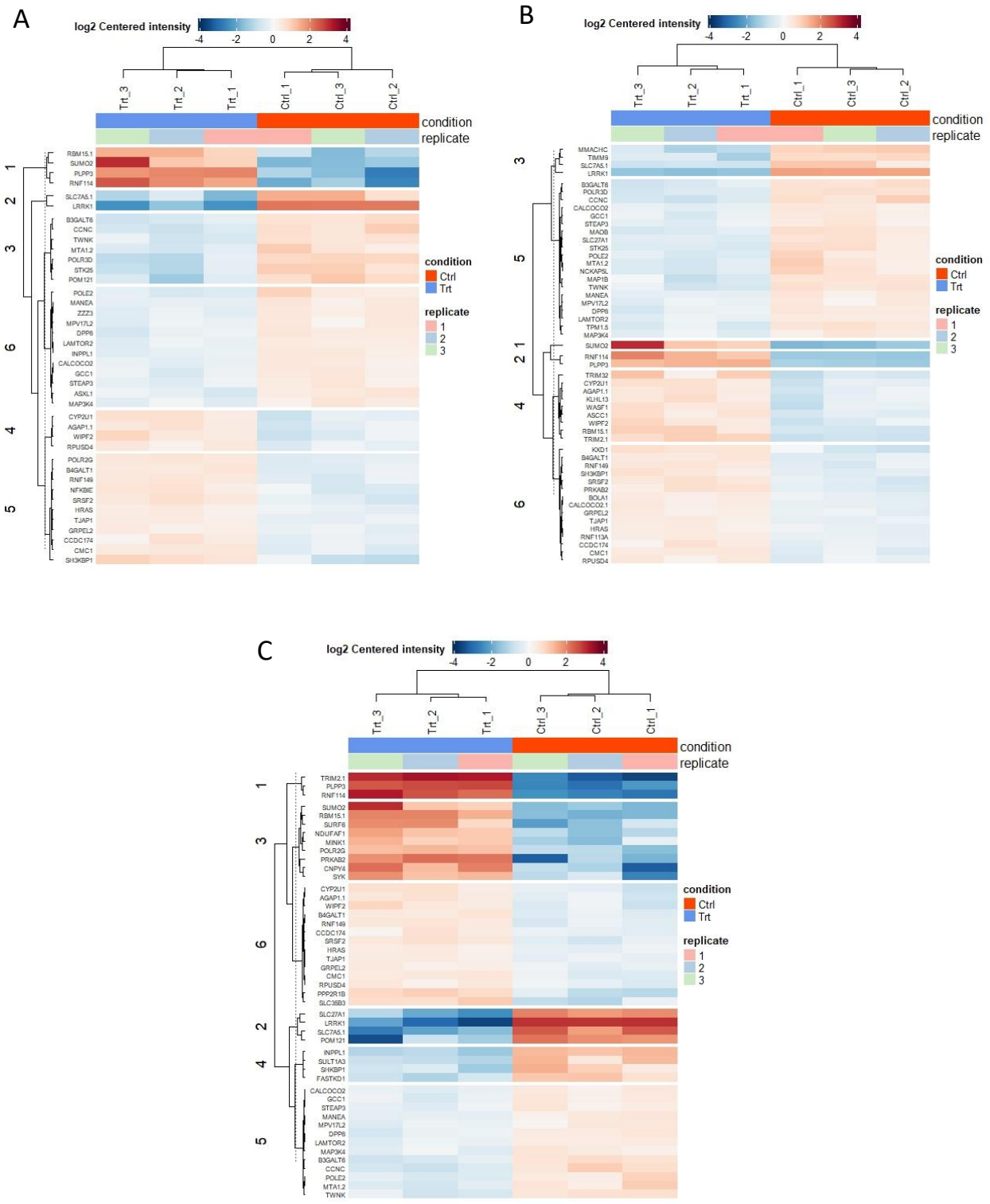

Figure 4.1.4: Heatmap of the differentially expressed proteins using missing not at random imputation methods. **A)** heatmap using the 'man' imputation method of analysis. **B)** heatmap using the 'MinProb' imputation method. **C)** heatmap using the 'QRILC' imputation method. Each column represents a sample type and each row represents one of the identified differentially expressed proteins. The proteins were also clustered using k-means clustering. The figures were produced using the DEP package on R (version 4.2.1).

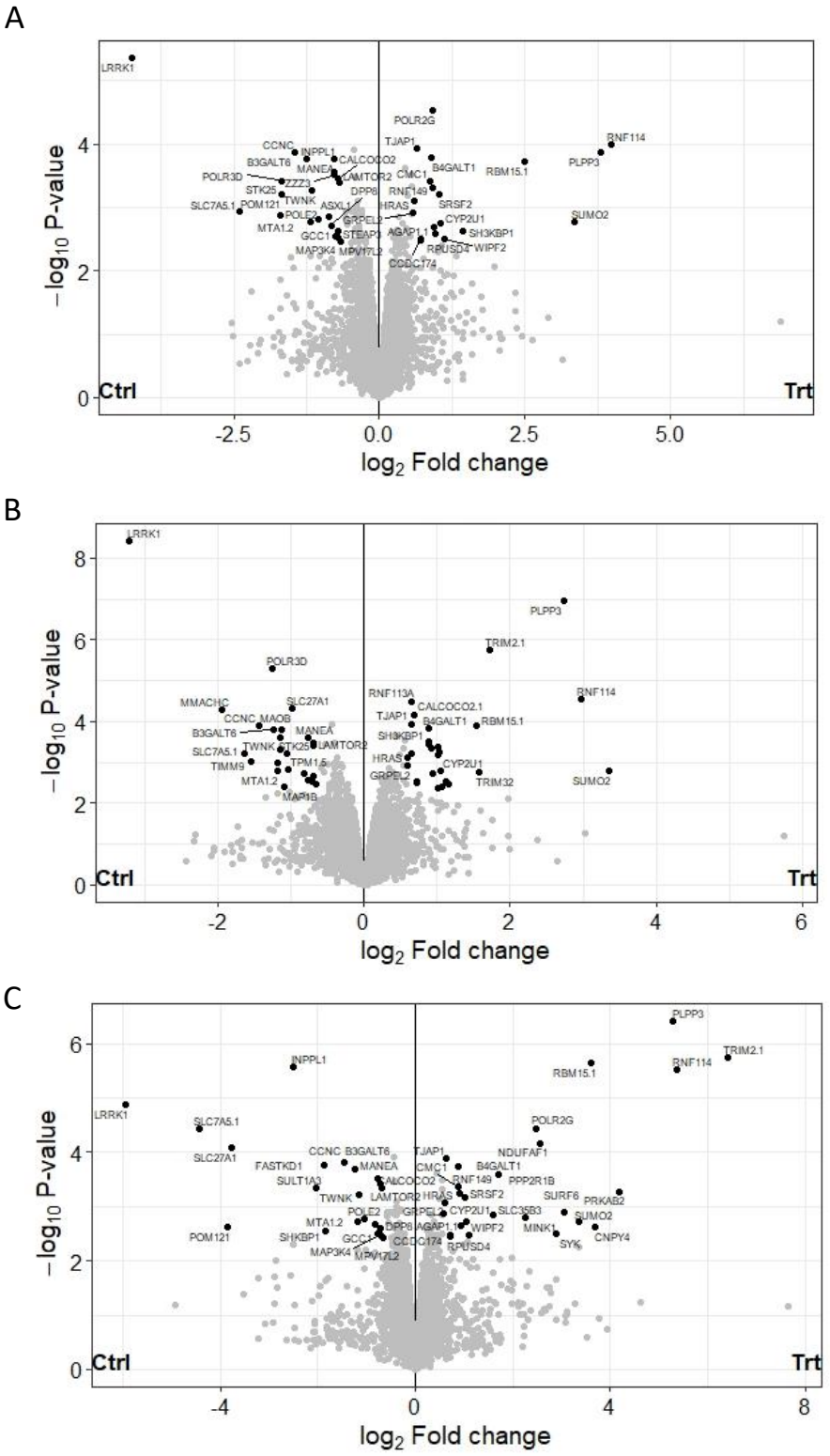

Figure 4.1.5: Volcano plot of the differentially expressed proteins using missing not at random imputation methods. **A)** Volcano plot using the 'man' imputation method of analysis. **B)** Volcano plot using the 'MinProb' imputation method. **C)** Volcano plot using the 'QRILC' imputation method. The cut off values were 1.5 $\log_2$ fold change and false detection rate of less than 0.05. The figures were produced using the DEP package on R (version 4.2.1).

Table 4.1: The proteins identified as differentially expressed between the treated cells and control cells using the 'man' imputation method of the missing not at random technique. The p-value, and adjusted p-value are shown. The calculations were completed by the DEP package on R (version 4.2.1). The highlighted rows outline the proteins where the adjusted p-value is smaller than the p-value.

Protein name	Trt_vs_Ctrl_p.val	Trt_vs_Ctrl_p.adj	Significant	Trt_vs_Ctrl_ratio
RNF114	0.000102	3.67E-09	TRUE	3.98
PLPP3	0.000138	2.45E-08	TRUE	3.8
SUMO2	0.001713	0.00722	TRUE	3.35
RBM15.1	0.000192	2.36E-07	TRUE	2.49
SH3KBP1	0.002404	0.0173	TRUE	1.43
WIPF2	0.003129	0.0308	TRUE	1.11
CYP2U1	0.001754	0.00774	TRUE	1.05
SRSF2	0.000619	0.000172	TRUE	1.03
NFKBIE	0.00265	0.0216	TRUE	0.948
AGAP1.1	0.002028	0.0116	TRUE	0.938
RNF149	0.000499	6.21E-05	TRUE	0.923
POLR2G	2.88E-05	7.05E-13	TRUE	0.912
B4GALT1	0.000162	7.78E-08	TRUE	0.884
CMC1	0.00038	1.51E-05	TRUE	0.875
RPUSD4	0.003231	0.0327	TRUE	0.722
CCDC174	0.003159	0.0314	TRUE	0.714
TJAP1	0.000119	9.48E-09	TRUE	0.648
HRAS	0.000772	0.000459	TRUE	0.601
GRPEL2	0.001216	0.00261	TRUE	0.587
MPV17L2	0.003447	0.0367	TRUE	-0.664
LAMTOR2	0.000412	2.28E-05	TRUE	-0.686
CALCOCO2	0.000346	9.34E-06	TRUE	-0.702
STEAP3	0.002353	0.0165	TRUE	-0.703
MAP3K4	0.003053	0.0294	TRUE	-0.717
GCC1	0.002878	0.026	TRUE	-0.762
ZZZ3	0.000311	5.27E-06	TRUE	-0.766
INPPL1	0.000172	1.13E-07	TRUE	-0.767
MANEA	0.000272	2.49E-06	TRUE	-0.768
DPP8	0.001905	0.00981	TRUE	-0.829
ASXL1	0.001382	0.00371	TRUE	-0.859
POLE2	0.001569	0.00548	TRUE	-1.05
TWNK	0.000532	8.45E-05	TRUE	-1.15
MTA1.2	0.00172	0.0073	TRUE	-1.19
B3GALT6	0.000174	1.21E-07	TRUE	-1.24
CCNC	0.000135	2.13E-08	TRUE	-1.45
STK25	0.00062	0.000173	TRUE	-1.67
POLR3D	0.000384	1.59E-05	TRUE	-1.68

POM121	0.001332	0.00337	TRUE	-1.7
SLC7A5.1	0.001174	0.00234	TRUE	-2.39
LRRK1	4.35E-06	7.05E-13	TRUE	-4.25

Missing at random DEP analysis

For the MAR analysis, two imputation methods were used. The PCA plots and correlation matrix shows that there are more similarities between the control and treatment groups compared to the MNAR methods (Figure 4.1.6).

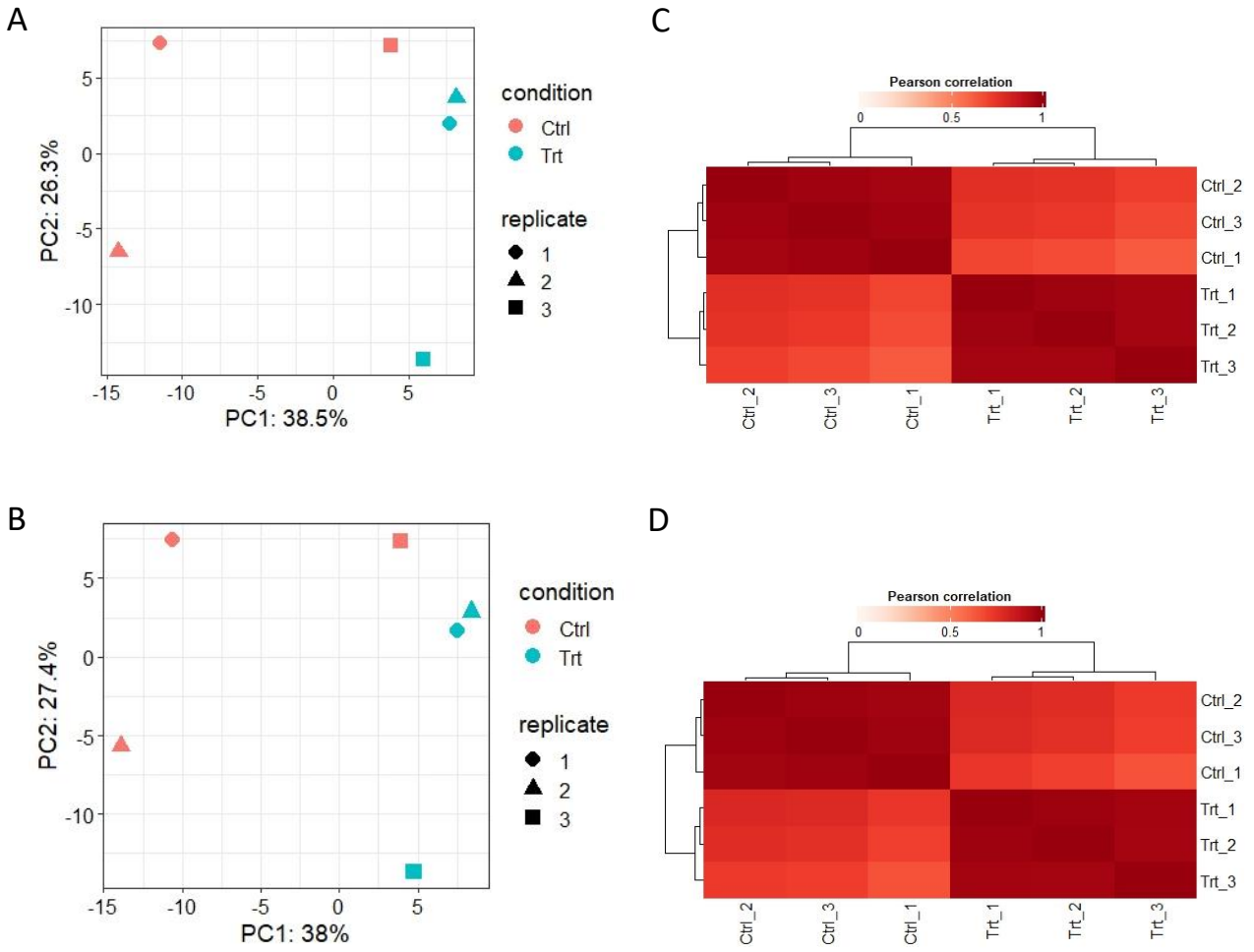

Figure 4.1.6: Missing at random imputation methods. **A)** Principal component analysis (PCA) plot using the 'knn' imputation method of analysis. **B)** PCA plot using the 'MLE' imputation method. **C-D)** the correlation matrix of the two imputation methods 'kNN', and 'MLE', respectively. The figures were generated using DEP package on R (version 4.2.1).

The heatmap (Figure 4.1.7) and volcano plots (Figure 4.1.8) also confirm that there were less DEPs identified with MAR analysis compared to MNAR, with ‘kNN’ identifying 38 DEPs (Table 4.4) and ‘MLE’ identifying 36 DEPs (Table S4.5).

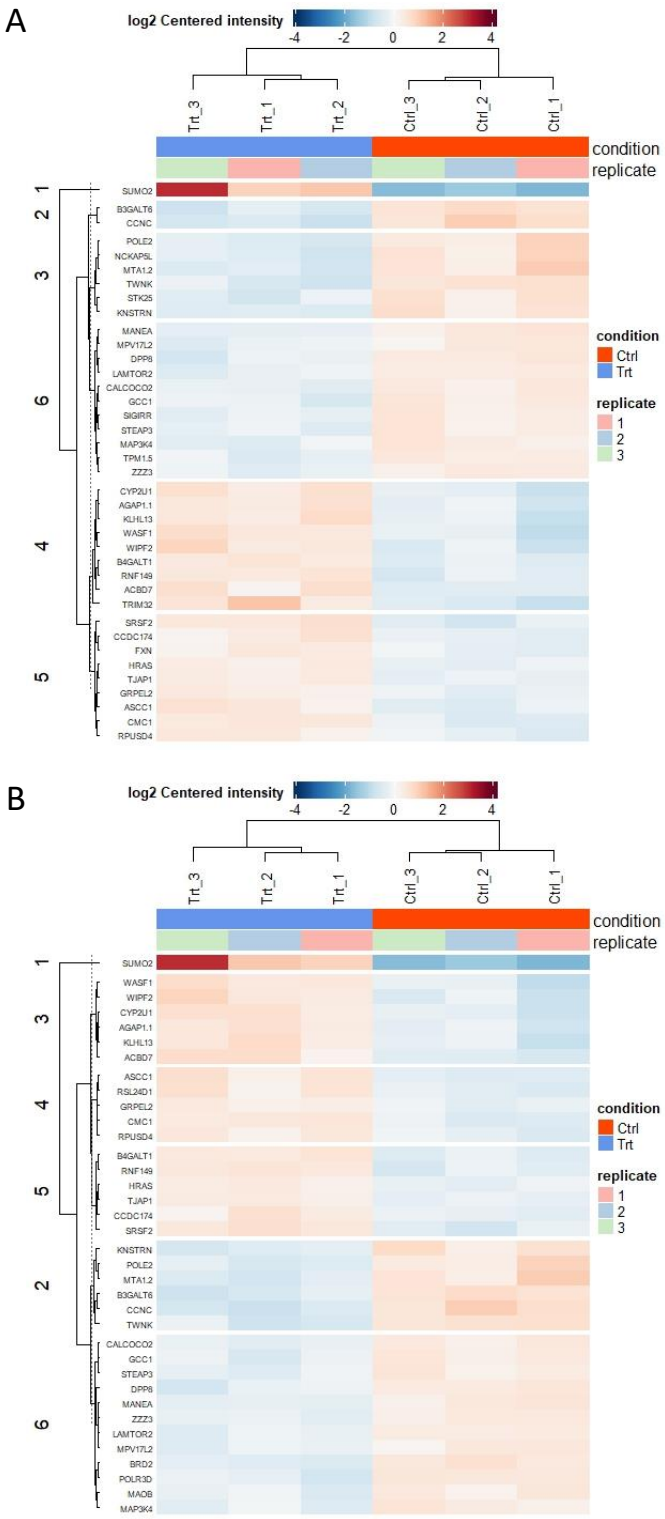

Figure 4.1.7: Heatmap of the differentially expressed proteins using missing at random imputation methods. **A)** Heatmap using the 'kNN' imputation method of analysis. **B)** Heatmap using the 'MLE' imputation method. Each column represents a sample type and each row represents one of the differentially expressed proteins. The proteins were also clustered using k-means clustering. The figures were produced using the DEP package on R (version 4.2.1).

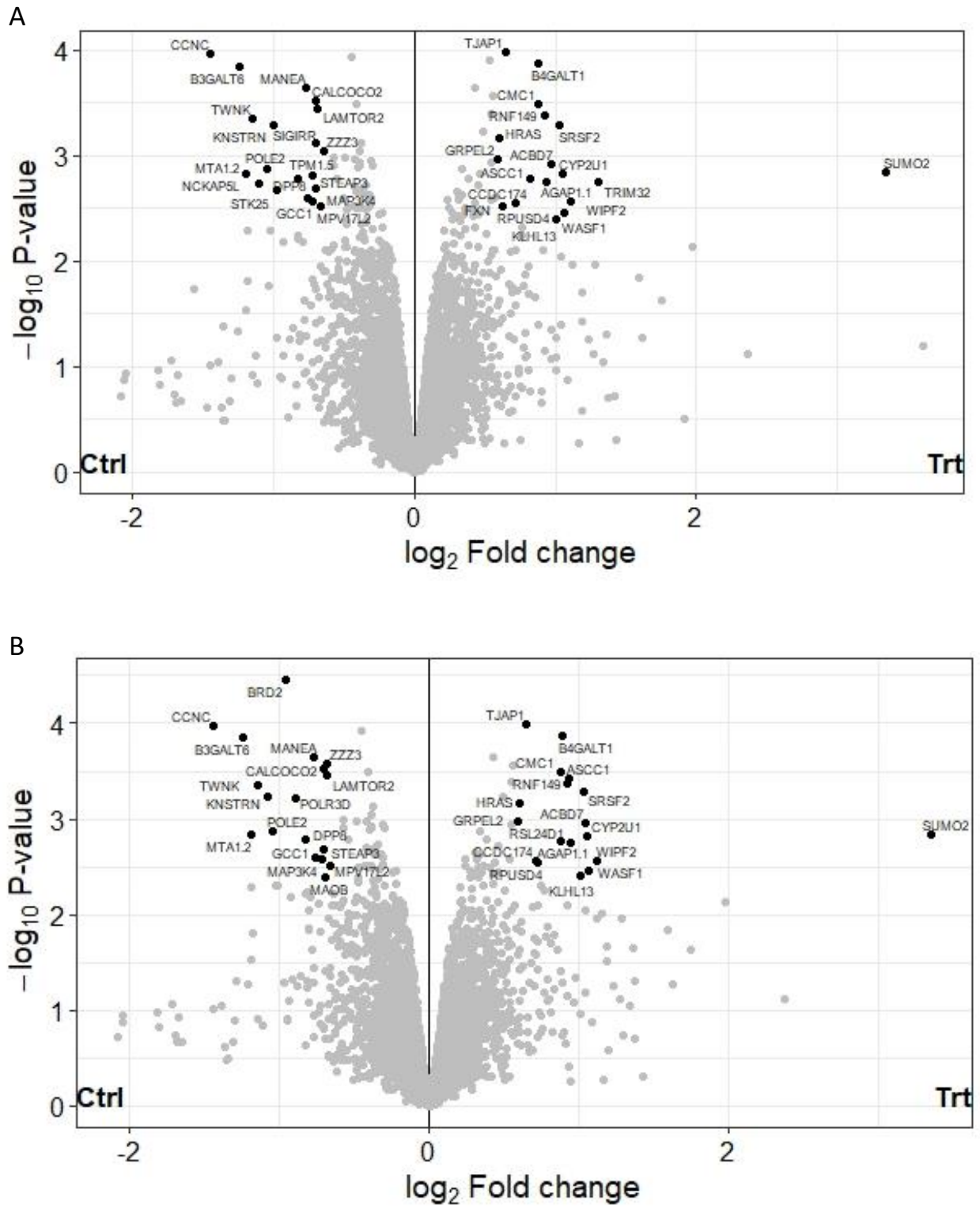

Figure 4.1.8: Volcano plot of the differentially expressed proteins using missing at random imputation methods. **A)** Volcano plot using the 'kNN' imputation method of analysis. **B)** Volcano plot using the 'MLE' imputation method. The cut off values were 1.5 $\log_2$ fold change and false detection rate of less than 0.05. The figures were produced using the DEP package on R (version 4.2.1).

Table 4.4: The proteins identified as differentially expressed between the treated cells and control cells using the 'kNN' imputation method of the missing at random technique. The p-value, and adjusted p-value are shown. The calculations were completed by the DEP package on R (version 4.2.1). The highlighted rows outline the proteins where the adjusted p-value is smaller than the p-value.

Protein name	Trt_vs_Ctrl_p.val	Trt_vs_Ctrl_p.adj	significant	Trt_vs_Ctrl_ratio
SUMO2	0.001455	0.00381	TRUE	3.35
TRIM32	0.001765	0.00642	TRUE	1.31
WIPF2	0.002724	0.0184	TRUE	1.11
WASF1	0.003513	0.0306	TRUE	1.07
CYP2U1	0.001509	0.00418	TRUE	1.05
SRSF2	0.000521	8.09E-05	TRUE	1.03
KLHL13	0.004	0.0408	TRUE	1
ACBD7	0.001222	0.00236	TRUE	0.97
AGAP1.1	0.001755	0.00632	TRUE	0.938
RNF149	0.000421	2.95E-05	TRUE	0.923
B4GALT1	0.000135	2.78E-08	TRUE	0.884
CMC1	0.00032	7.17E-06	TRUE	0.875
ASCC1	0.001635	0.00518	TRUE	0.825
RPUSD4	0.002848	0.0201	TRUE	0.722
CCDC174	0.002785	0.0192	TRUE	0.714
TJAP1	0.000104	9.82E-09	TRUE	0.648
FXN	0.003026	0.0225	TRUE	0.627
HRAS	0.000678	0.000281	TRUE	0.601
GRPEL2	0.001073	0.00164	TRUE	0.587
ZZZ3	0.000894	0.000875	TRUE	-0.642
MPV17L2	0.003053	0.0229	TRUE	-0.664
LAMTOR2	0.000355	1.24E-05	TRUE	-0.686
SIGIRR	0.000773	0.000484	TRUE	-0.7
CALCOCO2	0.000298	5.18E-06	TRUE	-0.702
STEAP3	0.002065	0.00969	TRUE	-0.703
MAP3K4	0.002689	0.0179	TRUE	-0.717
TPM1.5	0.001525	0.00429	TRUE	-0.721
GCC1	0.002526	0.0157	TRUE	-0.762
MANEA	0.000231	1.17E-06	TRUE	-0.768
DPP8	0.001654	0.00534	TRUE	-0.829
STK25	0.00213	0.0105	TRUE	-0.97
KNSTRN	0.000512	7.49E-05	TRUE	-1
POLE2	0.001346	0.00306	TRUE	-1.05
NCKAP5L	0.001846	0.00725	TRUE	-1.1
TWNK	0.000445	3.80E-05	TRUE	-1.15
MTA1.2	0.001474	0.00394	TRUE	-1.19
B3GALT6	0.000142	3.60E-08	TRUE	-1.24

CCNC	0.000108	9.82E-09	TRUE	-1.45

Delving into more details within the DEPs tables, it was interesting that the adjusted p-value for some of the DEPs (highlighted in tables) were in fact smaller than the p-values which is unexpected. Whereas a p-value of 0.05 suggests that 5% of all proteins will be false positive, an adjusted p-value suggests that 5% of significant proteins will be false positives. This means that the adjusted p-value will result in fewer false positives and should as a result have a larger value compared to the p-value. I therefore, identified that the calculation needed to be altered and once this was computed (outlined in methods section above), the resulting number of DEPs identified were much smaller. For the 'man' method, the new calculation only identified 2 proteins as being significantly different compared to the previous calculation which identified 40 proteins (Table 4.6). The tables for the other imputation methods are included in the supplementary data (Tables S4.7 and S4.8.). The 'kNN' and 'MLE' method showed no significant DEPs using the new Benjamini-Hochberg adjusted p-value calculation.

Table 4.6: The proteins identified as differentially expressed between the treated cells and control cells using the 'man' imputation method for missing not at random analysis. The adjusted p-value was calculated using the Benjamini-Hochberg calculation.

Protein name	Trt_vs_Ctrl_p.val	Trt_vs_Ctrl_p.adj (new)	Significant	Trt_vs_Ctrl_ratio
LRRK1	1.39E-06	0.00681	TRUE	-3.78
RNF114	1.74E-05	0.042554	TRUE	3.77

It was also interesting that the imputations methods identified some different proteins as being differentially expressed. Although some proteins were found by all 5 methods (such as CCDC174, CYP2U1, DPP8, and GCC1), there were some that were only identified by some of the methods and not others (such as ASXL1, BOLA1, FASTKD1, and FXN).

Significance A/B data analysis

From this point, I decided to use the significance A/B method, which determines outlier protein ratios compared to the distribution of all protein ratios. With previous studies, this method was able to identify more DEPs compared to the microarray methods (such as the DEP package). The DEPs identified by the significance A/B method are shown on the volcano plot with the positive fold change showing the proteins upregulated in the treatment group and the negative fold change showing the proteins downregulated in the treatment group (Figure 4.2.1). The list of differentially expressed proteins identified by the significance A/B method are shown in the supplementary results (Table S4.9). Using a fold change cut off value of 1 and FDR of 0.05, there were 76 proteins significantly upregulated and 77 proteins significantly downregulated in the treatment group compared to the control.

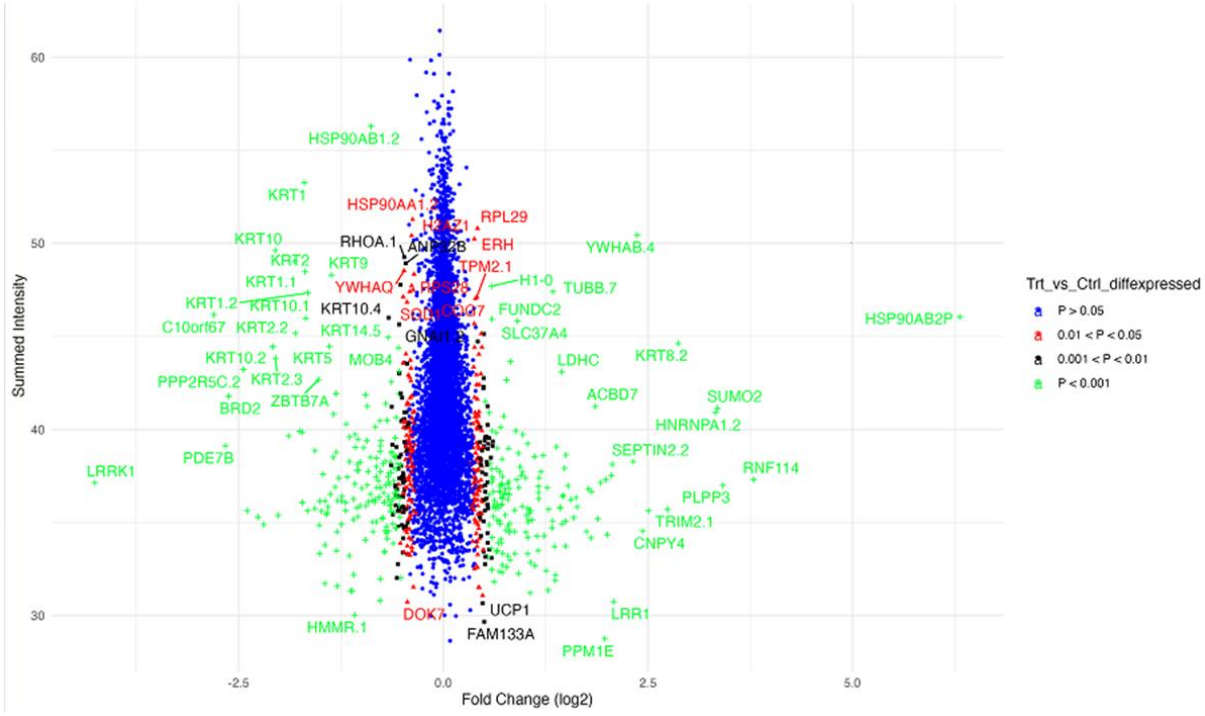

Figure 4.2.1: The volcano plot looking at the differentially expressed proteins in the treated cells vs the control cells. The positive fold change corresponds to proteins upregulated in the treated cells. The negative fold change corresponds to proteins downregulated in the treated cells. The p-values corresponding to the different colours are outlined in the legend.

Using the significance A/B analysis, the protein that was the most significantly upregulated in the treatment samples was HSP90AB2P and the protein that was most significantly downregulated in the treated samples was LRRK1 (Figure 4.2.1). With HSP90AB2P, the p-value is noted as 0 (Table S4.9), this is because the p-value is very close to 0 and it exceeds the limit of a normal computer's ability to calculate the p-value. Comparing the microarray DEP method to the significance A/B methods, the significance A/B method identified many more DEPs with the correct p-value and adjusted p-value calculations and so was the superior analysis method for this dataset.

STRING analysis of differentially expressed proteins

Using the STRING analysis platforms, protein networks were identified. The two networks upregulated in the treatment samples that were relevant to our study are shown in figure 4.2.2.

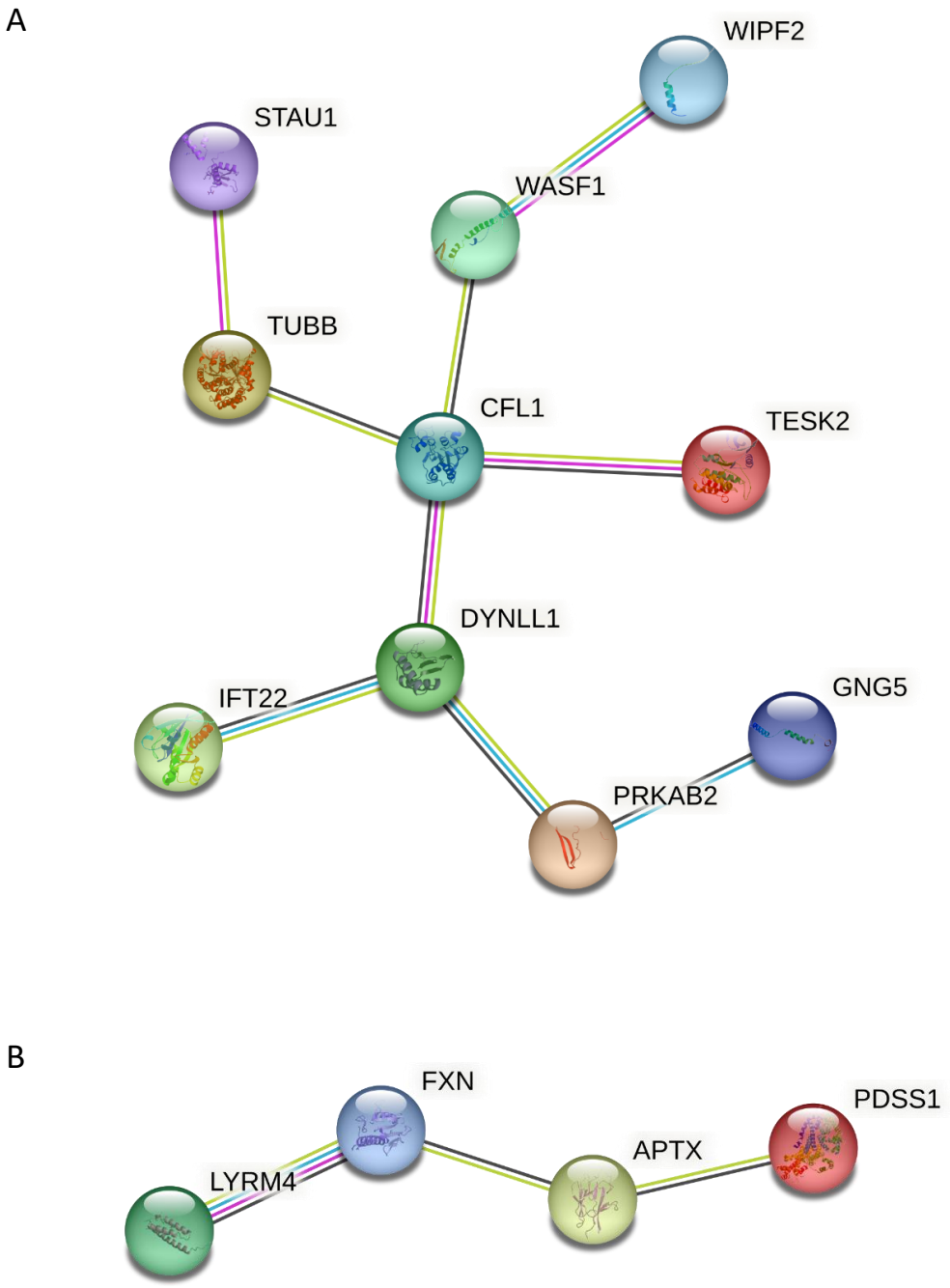

Figure 4.2.2: STRING analysis network of the differentially expressed proteins significantly upregulated in cells treated with recombinant CD142. The protein interactions are outlined within the **A)** first cluster and **B)** second cluster. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighbourhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for text mining evidence, light blue line is for database evidence, and black line represents co-expression evidence.

Within the first network (Figure 4.2.2.A), the biological process that was enriched was the 'cytoskeleton-dependent intracellular transport'. The cellular component term was the 'cytoskeleton' and the reactome pathways included the 'RHO GTPase Activate WASPs and WAVEs' and 'Organelle biogenesis and maintenance' (Figure 4.2.3).

Biological Process (Gene Ontology)

GO-term	description	count in network	strength	false discovery rate
GO:0030705	Cytoskeleton-dependent intracellular transport	4 of 195	1.6	0.0267

Cellular Component (Gene Ontology)

GO-term	description	count in network	strength	false discovery rate
GO:0005856	Cytoskeleton	7 of 2221	0.79	0.0366

Local network cluster (STRING)

cluster	description	count in network	strength	false discovery rate
CL:17257	Mixed, incl. slingshot, n-terminal, and adf/cofilin	2 of 13	2.48	0.0279
CL:17055	SCAR complex, and actin cortical patch localization	2 of 14	2.45	0.0279
CL:17047	Mixed, incl. regulation of actin polymerization or depolymeriza...	4 of 118	1.82	0.0013

Reactome Pathways

pathway	description	count in network	strength	false discovery rate
HSA-5663213	RHO GTPases Activate WASPs and WAVEs	2 of 36	2.04	0.0395
HSA-2029482	Regulation of actin dynamics for phagocytic cup formation	3 of 62	1.98	0.0090
HSA-5617833	Cilium Assembly	3 of 197	1.47	0.0325
HSA-195258	RHO GTPase Effectors	4 of 289	1.43	0.0104
HSA-1852241	Organelle biogenesis and maintenance	4 of 290	1.43	0.0104

Figure 4.2.3: Terms associated with the first protein cluster upregulated in the cells treated with recombinant CD142. The biological process terms, cellular component terms, local network clusters and reactome pathways are shown with their corresponding p-values (false discovery rates).

Within the second cluster (Figure 4.2.2.B), the proteins were involved in the cellular component 'mitochondrial matrix' and 'L-cysteine desulfurase complex'. The other terms can be seen in figure 4.2.4.

Cellular Component (Gene Ontology)				
GO-term	*description*	*count in network*	*strength*	*false discovery rate*
GO:1990221	L-cysteine desulfurase complex	2 of 2	3.69	0.00032
GO:0005759	Mitochondrial matrix	3 of 479	1.49	0.0495

Local network cluster (STRING)				
cluster	*description*	*count in network*	*strength*	*false discovery rate*
CL:14030	Mitochondrial iron-sulfur cluster biogenesis	2 of 6	3.21	0.0040

Reactome Pathways				
pathway	*description*	*count in network*	*strength*	*false discovery rate*
HSA-1362409	Mitochondrial iron-sulfur cluster biogenesis	2 of 13	2.88	0.0071

Figure 4.2.4: Terms associated with the second protein cluster upregulated in the cells treated with recombinant CD142. The cellular component terms, local network clusters and reactome pathways are shown with their corresponding false discovery rates.

Three protein networks were identified by the proteins that were significantly downregulated in the treated samples (Figure 4.2.5).

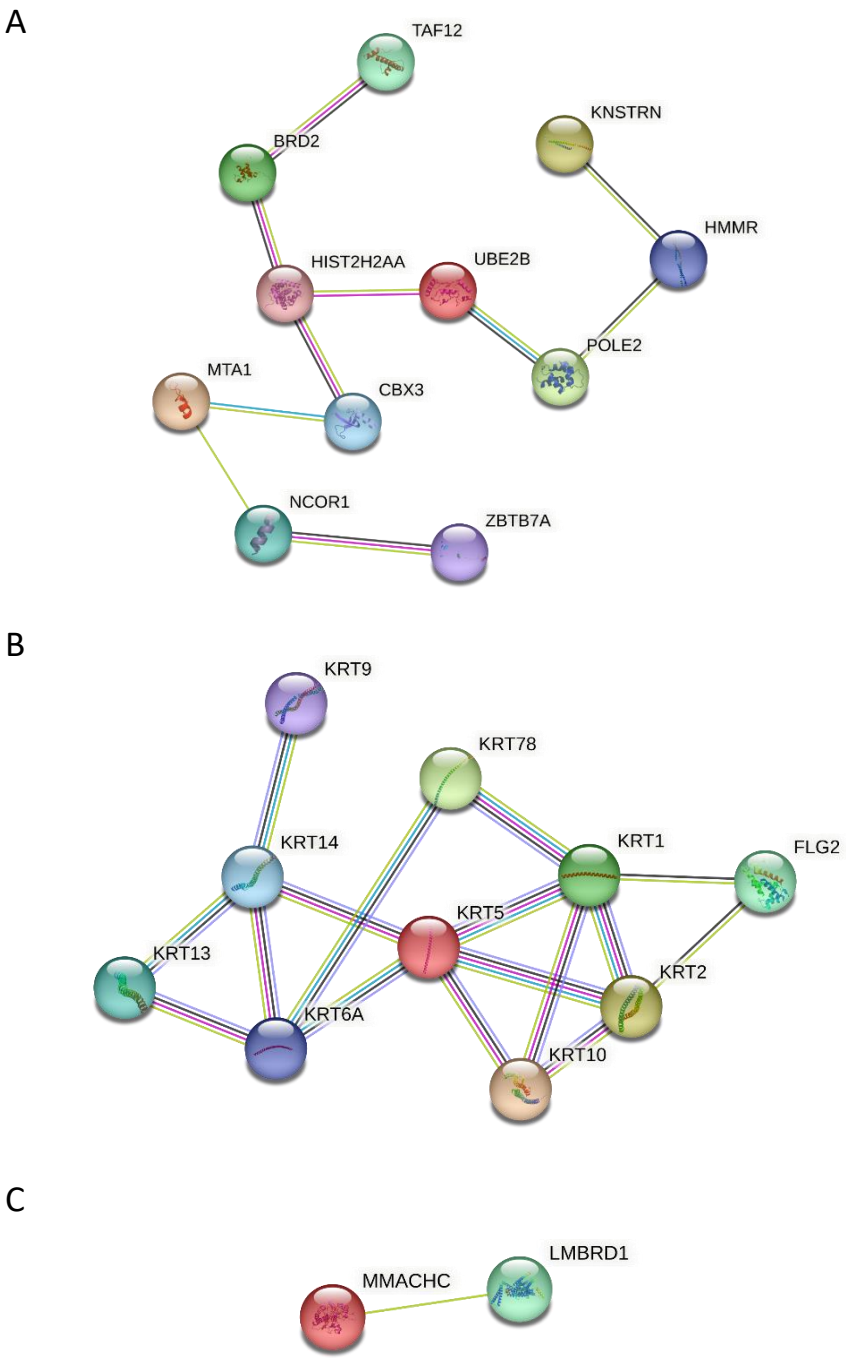

Figure 4.2.5: STRING analysis network of the differentially expressed proteins significantly downregulated in cells treated with recombinant CD142. The protein interactions are outlined within the **A)** first cluster, **B)** second cluster and **C)** third cluster. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighbourhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for text mining evidence, light blue line is for database evidence, and black line represents co-expression evidence.

The first network showed enrichment of the biological process terms ‘chromatin organization’ and ‘chromosome organization’. Some of the cellular component terms were ‘nuclear chromosome’ and ‘catalytic complex’. The rest of the enriched terms are shown in figure 4.2.6.

Biological Process (Gene Ontology)

GO-term	description	count in network	strength	false discovery rate
GO:0006325	Chromatin organization	8 of 713	1.3	3.13e-06
GO:0051276	Chromosome organization	10 of 1066	1.22	3.25e-08

Cellular Component (Gene Ontology)

GO-term	description	count in network	strength	false discovery rate
GO:0000228	Nuclear chromosome	6 of 1256	0.93	0.0210
GO:0005694	Chromosome	8 of 1712	0.92	0.00077
GO:1902494	Catalytic complex	6 of 1328	0.91	0.0210

Annotated Keywords (UniProt)

keyword	description	count in network	strength	false discovery rate
KW-0090	Biological rhythms	3 of 145	1.57	0.0226
KW-0678	Repressor	4 of 605	1.07	0.0434
KW-1017	Isopeptide bond	6 of 1706	0.8	0.0314
KW-0539	Nucleus	10 of 5208	0.53	0.0101

Figure 4.2.6: Terms associated with the first protein cluster downregulated in the cells treated with recombinant CD142. The biological processes terms, cellular component terms, and annotated keywords are shown with their corresponding false discovery rates.

The second protein network contained majority of the keratin proteins that were upregulated in the control samples. Some of the biological process terms were ‘hemidesmosome assembly’, ‘protein heterotetramization’, and ‘peptide crosslinking. The rest of the terms are shown in figure 4.2.7.

Biological Process (Gene Ontology)				
GO-term	description	count in network	strength	false discovery rate
GO:0031581	Hemidesmosome assembly	2 of 12	2.51	0.0114
GO:0051290	Protein heterotetramerization	2 of 15	2.42	0.0161
GO:0045109	Intermediate filament organization	3 of 25	2.37	0.00031
GO:0018149	Peptide cross-linking	3 of 34	2.24	0.00063
GO:0070268	Cornification	9 of 113	2.19	1.34e-15

Molecular Function (Gene Ontology)				
GO-term	description	count in network	strength	false discovery rate
GO:0030280	Structural constituent of skin epidermis	3 of 14	2.62	7.02e-05
GO:0005200	Structural constituent of cytoskeleton	5 of 104	1.97	1.95e-06
GO:0005198	Structural molecule activity	9 of 635	1.44	1.34e-09

Cellular Component (Gene Ontology)				
GO-term	description	count in network	strength	false discovery rate
GO:0001533	Cornified envelope	4 of 45	2.24	1.53e-06
GO:0045095	Keratin filament	7 of 92	2.17	4.58e-12
GO:0005882	Intermediate filament	9 of 199	1.95	2.44e-14
GO:0070062	Extracellular exosome	9 of 2099	0.92	3.27e-06
GO:0031982	Vesicle	10 of 3879	0.7	1.24e-05

Figure 4.2.7: Terms associated with the second protein cluster downregulated in the cells treated with recombinant CD142. The biological processes terms, molecular function terms and cellular component terms are shown with their corresponding false discovery rates.

The third and last protein interaction network was the interaction between MMACHC and LMBRD1. The enriched biological process was ‘cobalamine metabolic process’. The remainder of the terms for the protein pathway can be seen in figure 4.2.8.

Biological Process (Gene Ontology)				
GO-term	*description*	*count in network*	*strength*	*false discovery rate*
GO:0009235	Cobalamin metabolic process	2 of 21	2.97	0.0170

Molecular Function (Gene Ontology)				
GO-term	*description*	*count in network*	*strength*	*false discovery rate*
GO:0031419	Cobalamin binding	2 of 10	3.29	0.0011

Local network cluster (STRING)				
cluster	*description*	*count in network*	*strength*	*false discovery rate*
CL:20083	Defects in cobalamin (B12) metabolism	2 of 12	3.21	0.0022

KEGG Pathways				
pathway	*description*	*count in network*	*strength*	*false discovery rate*
hsa04977	Vitamin digestion and absorption	2 of 24	2.91	0.00057

Reactome Pathways				
pathway	*description*	*count in network*	*strength*	*false discovery rate*
HSA-3296469	Defects in cobalamin (B12) metabolism	2 of 14	3.15	0.0014
HSA-196741	Cobalamin (Cbl, vitamin B12) transport and metabolism	2 of 21	2.97	0.0014

Figure 4.2.8: Terms associated with the third protein cluster downregulated in the cells treated with recombinant CD142. The biological processes terms, molecular function terms and cellular component terms are shown with their corresponding false discovery rates.

Western blot validation of differentially expressed protein

One of the proteins downregulated in the treated cells (BRD2) was chosen to validate the proteomics results. The cell lysates from the cells treated with rCD142 (Figure 4.3.A T1-T3) show a lower expression of BRD2 compared to the control cells not treated with rCD142 (Figure 4.3.A C1-C3). The HEK293 cell lysate was used as a positive control, showing a band at the 88kDa molecular weight which corresponds to BRD2. The protein expression was normalised to total protein loaded with normalised fold change of the treated cells being 0.932±0.012 and the control cells had a normalized fold change of 1.218±0.027 with a p-value of 0.0290 (Figure 4.3.B).

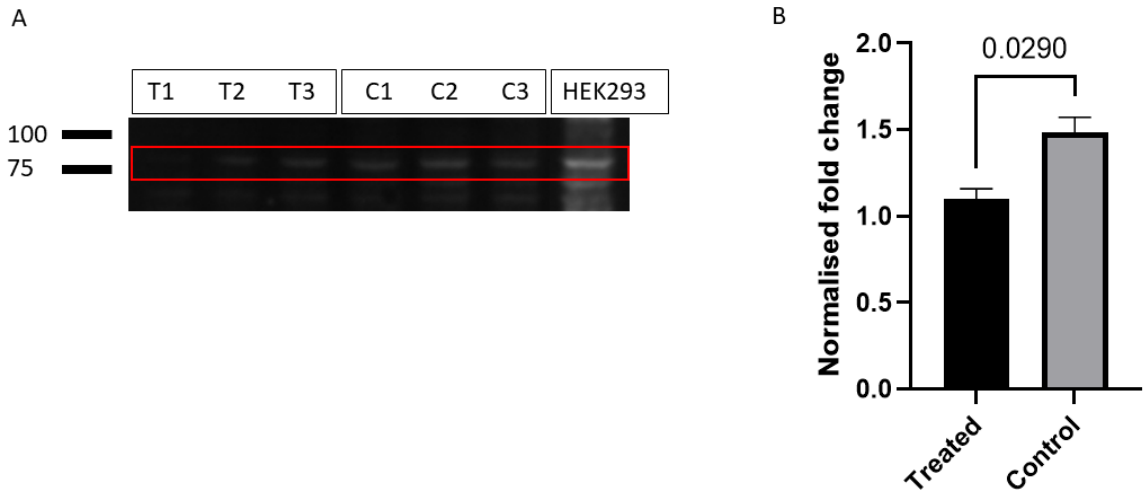

Figure 4.3: Western blot validation using BRD2 antibody. **A)** Western blot comparing the Ishikawa cell lysates treated with recombinant CD142 (T1-T3), control Ishikawa cell lysates (not treated with rCD142) and HEK293 cell lysate (positive control). The gel was developed using a gel documentation system (G-Box, Syngene) with the GeneSys software (version 1.6.7.0). **B)** Immunoblot normalization was carried out against total protein and band intensities were calculated using ImageJ (Java 1.8.0_172). The mean was calculated for the triplicates with SEM. The graph was generated using GraphPad Prism.

Discussion

Proteomics analysis of Ishikawa cells

Ishikawa cells are commonly used as an *in vitro* model of the endometrial epithelium. These cells are a good model because under hormonal treatment (oestrogen and progesterone) they mimic the endometrial epithelium at the window of implantation[140]. Therefore, it is possible to test the effects of specific treatments on the cells as a model of the changes that would occur *in vivo* in response to treatment. A comprehensive way to investigate the effects of a treatment on cellular models is through proteomics analysis. Proteomics allows for the study of the proteomes or the proteins expressed in a system[150]. The difference in the proteome between different conditions can allow for the identification of different pathways and interactions that could result in metabolic changes and effect downstream processes[152]. For this study, I treated cells with recombinant CD142 (identified to be commonly expressed by embryonic extracellular vesicles as outlined in chapter 2) and followed a standard proteomics pipeline to understand the effects of CD142 on the

expression of cellular proteins and identify interesting cellular processes that are altered as a result. To the best of my knowledge, this study is the first study to investigate the effects of recombinant CD142 on Ishikawa cells. The dataset was analysed using 2 methods, out of which the significance A/B method proved to be superior in identifying more differentially expressed proteins between the experimental groups. In this study, I identified 76 proteins that were significantly upregulated and 77 proteins that were significantly downregulated in the treatment groups compared to the control.

Differentially expressed proteins

Using the significance A/B method, my analysis showed that heat shock protein 90 alpha class B member 2 pseudogene (HSP90AB2P) was the most highly upregulated protein in the treated cells. HSP90AB2P has been involved in ATP binding, protein folding, protein stabilization and promotes structural maintenance and regulation of target proteins[157]. Heat shock proteins (HSPs) are important in maintaining protein homeostasis in endometrial cells. The HSP90 family acts as a chaperone which enhances the folding of new or misfolded proteins[158]. As well, they play a role in the degradation of disordered proteins. HSPs have been shown to play roles in decidualisation, implantation and placentation. Focusing on the HSP90 family, during decidualisation, the proteins may be involved in the maturation and recognition of steroid (progesterone and oestrogen) receptors found on the endometrial cells[159]. It has been suggested that during the process of implantation, HSPs may destabilise the epithelial barrier and aid invasion of the embryo. The role of HSPs in placentation and throughout pregnancy has also been studied[157]. For the purposes of our experiment, the initial stages of pregnancy such as decidualisation and implantation, would be the stages that could therefore be affected by the introduction of TF as seen with the increase of HSP90AB2P expression.

The next highly significant protein upregulated in the rCD142 treated cells was ring finger protein 114 (RNF114), a member of the E3 ubiquitin-protein ligase[160]. The E3 ligase family of proteins are involved in ubiquitinating other proteins and regulating cellular processes such as the cell cycle. Although no direct links have been identified between RNF114 and implantation of the embryo, increased ubiquitination of proteins was seen in the uterus during implantation suggesting that E3 ligase proteins may be playing a role in initiating pregnancy[160].

The most significantly downregulated protein in the rCD142 treated cells was leucine rich repeat kinase 1 (LRRK1) which is involved in intracellular signal transduction pathways[161]. The link between LRRK1 and implantation has not been studied. However, in a study by Jacobo-Baca et al., proteomics analysis of serum samples from pregnant women showed that LRRK1 was significantly seen in preeclampsia. The serum samples were taken at 10 to 12 weeks and the development of PE was noted as the pregnancy progressed, suggesting that LRRK1 could be playing a role in the early pathophysiology of PE[162]. Therefore, there is a potential that LRRK1 could be involved in other adverse pregnancy outcomes and should be investigated further. The downregulation of the protein could be beneficial in the success of the pregnancy.

Two relevant clusters were identified after the STRING analysis of the upregulated proteins in the treated samples. In the first cluster, WAS/WASL-interacting protein family member 2 (WIPF2) acts to induce the reorganization of actin filaments involved in cellular proliferation. WIPF2 is activated downstream of platelet derived growth factor receptor (PDGFR)[163]. Although PDGFR itself was not found to be differentially expressed, the activated (phosphorylated) form was more highly expressed in the rCD142 treated cells (shown in Chapter 3). This suggests that rCD142 activates PDGFR, increasing the expression of downstream proteins such as WIPF2.

The second cluster contained 4 proteins: LYRM4, FXN, APTX, and PDSS1. Out of these, Both LYRM4 and FXN are involved in mitochondrial physiology through iron-sulfur (Fe-S) cluster assembly[164]. The Fe-S cluster then incorporates into proteins and acts as catalysts or act as redox proteins in the mitochondria. The proteins that contain Fe-S clusters have also been shown to play roles in gene regulation. FXN is involved in the early stages of the Fe-S cluster incorporation into proteins and mutations of the gene cause iron build up in the mitochondria. The effects are seen in FXN null mutations in mice which causes embryonic lethality after implantation[165]. APTX (aprataxin) is a protein involved in single stranded DNA (ssDNA) repair. It does so through DNA deadenylation of the 5' end of a break allowing the repair to take place[166]. If left untreated, ssDNA breaks can lead to cellular dysfunction, cellular instability, and cellular death. APTX has also been shown to be involved in mitochondrial DNA repair playing a role in mitochondrial stability. PDSS1 is an enzyme and involved in the respiratory chain of the mitochondria, specifically in the coenzyme Q biosynthesis[167]. These proteins are therefore important for the proper functioning of the mitochondria and, if downregulated, could lead to cellular dysfunction.

Looking at the clusters of proteins which were downregulated in the treated cells, within the first cluster, BRD2 was identified which is bromodomain containing protein 2. It was shown that the inhibition of BRD proteins suppresses mediators which are induced by inflammation involved in the rupture of foetal membranes[168]. Suppressing the proteins could act as a target for reducing inflammation associated contractions[169]. Therefore, the downregulation of BRD2 could be a route of suppressing inflammation induced events associated with early pregnancy loss. The downregulation of BRD2 in the treated cells was confirmed using western blot analysis. Another interesting protein which was part of the same cluster was UBE2B which is a ubiquitin conjugating enzyme of the E2 family. Although the direct link of UBE2B in implantation has not been studied,

the association of UBE2Q2 with site-specific implantation in rabbits has been shown[170]. UBE2Q2 is closely related to UBE2B and falls within the same family of E2 enzymes. The expression of UBE2Q2 was found only at the site of implantation on the luminal epithelium with no expression on adjacent sites[170]. Therefore, if this site-specific regulation is altered, the implantation of the embryo may be affected and having regulation on the expression of the protein may play an important role[171]. The specific action of UBE2B should be investigated further to determine whether this protein is involved in determining the site of implantation in humans.

The second cluster of proteins included the keratin network of proteins which are important in signalling in epithelial cells. Keratins are a subgroup of intermediate filament proteins which are categorised based on their molecular weight and isoelectric point. Keratins play roles in a variety of signalling pathways through the formation of complexes with other proteins such as kinases, and receptors[172]. Their effects can be seen in cellular processes such as protein synthesis, cellular development and differentiation[173]. Due to the multitude of cellular actions, the expression of keratins needs to be tightly regulated. It would be interesting to test the effects of each keratin protein on the process of implantation. It is apparent that the presence of keratins is important but the regulation of expression needs to be examined further.

Within the third downregulated cluster, the 2 proteins were involved in the cobalamin metabolic pathway which has been studied in a cohort of Korean women with respect to recurrent implantation failure[174]. Cobalamin is also known as vitamin B12 and is involved in folate recycling which is required for cell proliferation, embryogenesis and implantation. Within the study, polymorphisms of genes associated with cobalamin transport to cells and binding were investigated which showed that specific genotypes were linked to recurrent implantation failure. The finding was confirmed in women undergoing IVF treatment and shows that differential

expression of the proteins involved in this pathway could lead to adverse outcomes in implantation of the embryo[174].

Limitations and Future work

The main limitation with this experimental setup was the small sample size for both the control and treated cells (n=3). The identified DEPs will need to be investigated further using real time quantitative polymerase chain reaction (RT-qPCR). Due to time constraints, limited samples and limited resources, the validation of a single protein (BRD2) was completed which warrants the need to validate the expression of more proteins. It would also be interesting to investigate the activity of TF found on embryonic extracellular vesicles as this would be the natural source of TF *in vivo*. To do this, the EEVs would need to be isolated and those expressing TF would be captured and removed using bead capture. The TF expressing EEVs could then be used to treat cells or investigated further to identify other markers which are co-expressed to begin to unravel the role of the specific subtype of EEVs in implantation.

The effects of rCD142 on the Ishikawa cells show that the proteome was changed and could therefore play important roles in cellular processes and possibly in the implantation of the embryo. The potential routes that could be investigated further to improve infertility treatments with regards to my findings are discussed in more detail in chapter 6.

Chapter 5

Proximity barcoding assay – pilot study

Introduction

Small Extracellular vesicles

Small extracellular vesicles (S-EVs) are a class of EVs which are smaller than 200nm in size and are produced by cells through the endosomal pathways as a means of cellular communication. This category of EVs is typically more difficult to characterise due to their small size and limitations with traditional methods of EV characterisation. However, S-EVs are of growing interest because of their biogenesis pathway, proteins and other targeted cargo for communication. The role of these EVs has been confirmed in disease development and progression in a multitude of diseases including pregnancy related diseases. S-EVs also have a role in regular cell development and are being investigated further in different fields. This chapter will focus on introducing a new technique for characterising embryonic S-EVs (S-EEVs) based on multiple surface proteins.

Challenges with traditional methods of characterisation

The traditional methods of characterising EVs include electron microscopy, nanoparticle tracking analysis, western blot, flow cytometry, proteomics, transcriptomics, and metabolomics analysis[175]. For the characterisation of EVs using traditional methods, the limitations that arise are related to the starting volume of the sample, the amount of EVs that are present, the size of the EVs that are being investigated and the proteins of interest. The initial step for EV characterisation would be to isolate the EVs which can be achieved through multiple techniques including ultracentrifugation, gradients, chromatography, and immunoaffinity[47]. The traditional isolation methods are particularly hindered when the starting volume of the sample is low and there is a low concentration of EVs present. Investigating the S-EVs in a low volume sample is even more challenging due to the size of the EVs. For these samples, new methods have been developed for

characterisation. Such methods include flow cytometry methods, and sandwich-based assays some of which were explained in chapter 2. With traditional flow cytometry methods, the lower size detection limit is 300nm which is not sufficient for S-EVs[175]. However, new techniques have been developed which allow for the detection of small particles as small as 50nm[119]. The limiting factor for flow cytometry and the sandwich-based assays is that only a maximum of 2 markers can be simultaneously investigated on the same EV. One of the other new methods that has been recently introduced into the field is proximity barcoding assay (PBA) which uses antibody DNA conjugates and next generation sequencing to identify surface proteins found on individual S-EVs[176]. The characterisation of EVs based on surface proteins is typically focused on identifying a single protein or, in some cases, using a pair of proteins. However, PBA allows for the simultaneous identification and characterisation of multiple surface proteins found on an individual S-EV using small volumes of samples.

Proximity barcoding assay (PBA)

In the initial PBA experiments, 38 surface proteins were simultaneously profiled on S-EVs from 18 different sample types which were then clustered using a combination of 3 or more proteins (found on individual S-EVs) into distinct clusters that could be used to categorise the S-EVs into the sample types[176]. The experiments proved that PBA can be used to identify multiple proteins on S-EVs and categories S-EVs based on their signature protein markers. PBA achieves the high-throughput data by using single stranded DNA sequences created through rolling circle amplification to code each S-EV. In detail, oligonucleotides (containing unique moleculeTags and proteinTags) are first conjugated to antibodies (corresponding to the proteins of interest) to create PBA probes. Rolling circle amplification (RCA) products are created using circularized oligonucleotides containing complexTags. The S-EVs in the sample are allowed to bind the PBA

probes through their surface proteins. The S-EVs are then captured on microtitre wells and the unbound EVs as well as contaminants, are washed away. The RCA products are added and allowed to anneal to the PBA probes through oligonucleotides brought into proximity which then hybridise and allow for extension using DNA polymerase. Each RCA product has a unique complexTag which is incorporated into the sequence after DNA polymerization. Through polymerase chain reaction and sequencing, the protein sequences of each S-EV are identified. The sequence reads are then sorted based on the Tags. The proteinTags determine which proteins are found on the S-EVs. The random moleculeTags associated with each PBA can be used to identify the presence of more than one of the same proteins on the same S-EV. The complexTag can be used to identify all the different proteins found on an individual S-EV. Counting the number of moleculeTags for each of the proteinTags which share the same complexTags can be used to detect the proteins found on an individual S-EV (Figure 5.1).

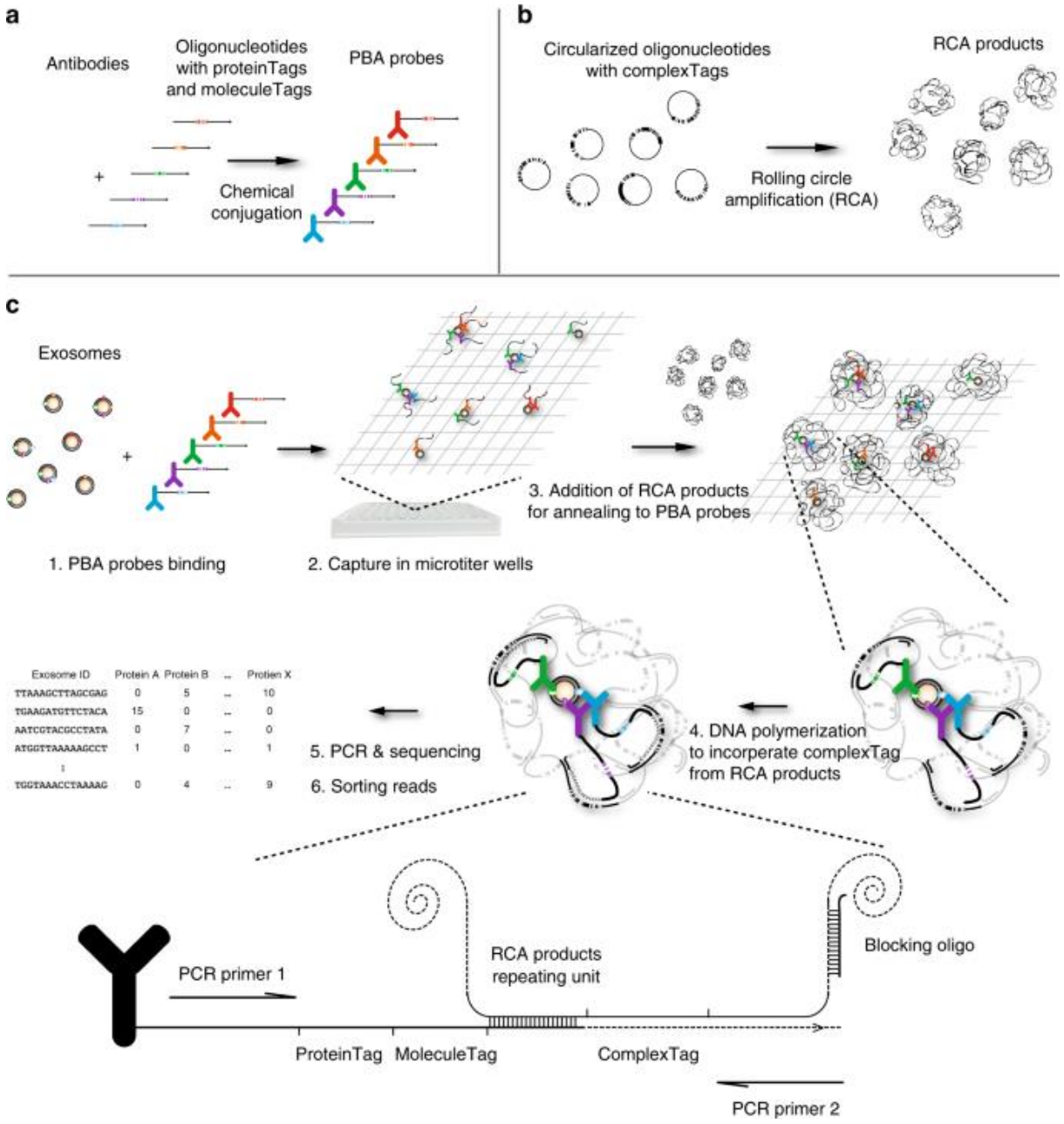

Figure 5.1: Schematic representation of proximity barcoding assay (PBA). **A)** Preparation of PBA probes with proteinTag and random sequence moleculeTag. **B)** RCA products containing random sequence complexTags. **C)** Small EVs (S-EV) are incubated with PBA probes, followed by capture in microtiter wells. Oligonucleotides on PBA probes brought together by binding the same S-EV bind to a unique RCA product, followed by enzymatic extension, incorporating the complexTag and standard sequence motif, later used for amplification. Successfully extended DNA molecules are amplified by PCR. DNA sequencing is performed to record the numbers of proteins with specific tag combinations, to identify all the proteins on an individual S-EV.

PBA allows for the simultaneous analysis of a high number of proteins, using targeted antibodies, found on an individual S-EV. The specificity and simultaneous labelling of multiple proteins on the same EV was the reason the PBA method was used for the analysis of our S-EEVs. Compared to PBA, traditional immunoassays are not designed for multiplexing due to the high cross-reactivity of antibodies. Using PCR in PBA analysis, allows for exponential amplification of signals increasing the assay sensitivity while using extremely small sample volumes which is beneficial for our EEV samples. The benefits of PBA would be identifying multiple proteins commonly expressed on the same S-EEV, which could be used to understand the cross-talk to the endometrium and S-EEVs for downstream analysis.

Aims

The aim of the chapter is to **1)** explore a more global characterisation of S-EEVs using 210 surface antibodies to identify specific surface proteins on S-EEVs comparing samples from good-quality embryos to poor-quality embryos, **2)** to determine whether clustering of the S-EEVs based on their expressed proteins can be achieved and **3)** to identify combinations of 2 proteins found on the same S-EEV for future investigations. This chapter aims to confirm a new proof-of-concept study for characterising S-EEVs which could be taken further in future studies.

Materials and methods

Sample collection

The samples used for this pilot study were collected and stored with our collaborators in Nanfeng Hospital, China. Thirty-five microliters of culture media were collected from a single culture media droplet containing one day 3 embryo after the embryo was either transferred to the uterus

or deemed not suitable for transfer. The collected media was centrifuged at 300 x g for 10 minutes at 4°C to remove larger cells and debris. The supernatant was removed and stored at -80°C.

For this pilot study, 6 pooled samples were used. Each pooled sample contained culture media from 2 embryos (either good quality embryos which were transferred to the uterus or poor quality embryos deemed not suitable for transfer). The samples were sent to vesicode China for proximity barcoding assay analysis. Due to the low sample volume and limited number of small vesicles present in the samples, neat samples were used without further preparation.

Panel of antibodies

For this experiment, the list of the 210 surface antibodies used are shown in table 5.1.1. The antibodies used were all in the cell adhesion molecule panel due to their importance in targeting and adhesion of S-EVs.

Table 5.1.1: List of the antibodies used for interrogation of surface proteins, using proximity barcoding assay, found on human embryonic small extracellular vesicles.

5T4	Cadherin-13	CD73	Claudin-3	DSG3	ICAM-2	ITGAL	L1CAM/CD171	Nectin-4	Protocadherin-8	TIMP-2
A33	cadherin-17	CD73_m	Claudin-6	DSG4	ICAM-3	ITGAM	LAG-3	nectin1	Siglec-1	TrkC
ABCG2	Cadherin-4	CD81	Claudin-8	E-Cadherin	ICAM-4	ITGAM_m	LAMP1	Nestin	Siglec-10	TROP-2
ADAM10	Cadherin-6	CD9	CLDN19	E-Selectin	IGSF4B	ITGAV	LAMP2	Neurofascin	Siglec-11	ULBP-1
Adiponectin	Cadhherin	CD90/Thy1	CLDN4	Endomucin	IGSF4C	ITGAVB3	Lgr5/GPR49	Neuroligin 1	Siglec-14	ULBP-2/5/6
ADSF	Calsyntenin-1	CDCP-1	CLEC-2	EpCAM	IL-1/RAPL1	ITGAVB5	LYN	NG2/MCSP	Siglec-2	ULBP-3
ALCAM_m	Carboxypeptidase M	CDH2	CLEC-2A	EphA2	IL-6	ITGAX	MAdCAM-1	NGFR/TNFRSF16	Siglec-3	uPA
ALDH1A1	Caveolin-1	CDO	CLP24	Ephrin-B2	IL-8/CXCL8	ITGB1	MCAM	P-Cadherin	Siglec-5/14	uPAR
AMGO1	CCR-2	CEACAM-1	CTLA-4	ErbB2/Her2	ILK	ITGB2	MDL-1	PAR2	Siglec-5/CD170	VCAM-1
AMICA/JAML	CCR-6	CEACAM-3	CXADR	ESAM	ITGA1	ITGB3	MICA	PCDH15	Siglec-6	VE-Cadherin
AMIGO2	CD117/c-kit	CEACAM-4	CXCL16	FN	ITGA11	ITGB4	MICB	PCDHα1	Siglec-7	
Annexin A1	CD133	CEACAM-5	CXCR4	FOLH1	ITGA2	ITGB5	MMP-9	PD-1	Siglec-8	
APN	CD151	CEACAM-6/CD66c	DEP-1/CD148	Galectin-9	ITGA2B	ITGB6	MUC-1	PD-L1	Siglec-9	
ASAM	CD20	CEACAM-7	Desmocollin-1	Gelsolin	ITGA3	ITGB7	MUC-19	PECAM-1	SynCAM3/IGSF4B	
Axl	CD24	CEACAM-8/CD66b	Desmocollin-2	Glypican 1	ITGA4	ITGB8	MUC-4	Plexin B1	SynCAM4	
BCAM	CD26	Claudin-1	Desmocollin-3	GPR56	ITGA4B7	JAM-A	MUCDHL	Podoplanin	Teneurin-1	
BOC	CD27	Claudin-10b	DSCAM	HepaCAM	ITGA5	JAM-B	N-Cadherin	Protocadherin γC3	Teneurin-2	
CA125	CD36	Claudin-11	DSCAM-L1	HLA-DR	ITGA6	JAM-C	NCAM-1	Protocadherin-1	Teneurin-4	
Cadherin-11	CD44	Claudin-12	DSG1	HLAA	ITGA8	L-Selectin	Nectin-2	Protocadherin-17	TICAM2	
Cadherin-12	CD63	Claudin-17	DSG2	ICAM-1	ITGA9	L1CAM	Nectin-3	Protocadherin-19	TIM-3	

The PBA probes and rolling circle amplification (RCA) products were prepared according to the methods outlined in Wu et al., 2019 and briefly explained below[176].

Antibody-oligonucleotide conjugates (PBA probe)

The antibodies were activated using sulfosuccinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate (sulfo-SMCC) in dimethyl sulfoxide for 2 hours at room temperature. After 1 hour of activation, the modified oligonucleotides were reduced using dithiothreitol (DTT) and ethylenediaminetetraacetic acid (EDTA) and were incubated for 1 hour at 37°C. The activated antibodies and reduced oligonucleotides were separately purified. Each antibody was mixed with one type of oligonucleotide followed by dialysis at 4°C overnight. The conjugates were then stored in phosphate buffered saline (PBS) with 0.1% bovine serum albumin (BSA) at 4°C.

Rolling circle amplification (RCA) product preparation

Oligonucleotides were ligated into circular DNA at 37°C for 30 minutes. Un-ligated oligonucleotides were then removed using exonuclease I and III. The RCA primers were added and allowed to re-anneal to the circular DNA. RCA was performed using phi29 polymerase. The RCA products were then kept at -20°C until use.

Small EV capture

In PBA, S-EVs are first mixed with PBA probes in solution then captured on a 96-well microtiter wells coated with cholera toxin subunit B (CTB) which binds GM1 gangliosides on the membrane of S-EVs. The S-EEVs were first mixed with all antibody conjugates in PBA buffer at 4°C overnight. After incubation, 1µL of the S-EEV antibody conjugate complexes were diluted in PBA buffer and 20µL was incubated in CTB coated wells for 15 minutes at room temperature. Replicate measurements were done in duplicate wells.

Proximity barcoding

The captured S-EV antibody conjugates were washed twice in washing buffer. The RCA products were added at a concentration of 1nM and incubated at 37°C for 15 minutes. This was again washed 2 times with washing buffer followed by the addition of T4 DNA polymerase and dNTPs and incubated at room temperature for 15 minutes. The wells were washed once.

Amplification and sequencing

PCR mix containing 2mM MgCl2, 0.2mM dNTPs, 1x SYBR Green I, 1% DMSO, 500nM PCR primers, 0.02U/µL Phusion polymerase and 0.02 U/µL Uracil-DNA glycosylase was added at 30µL to each well. The thermos cycler (MX3005; Stratagene) was used for polymerase chain reaction with initial incubation at 37°C for 15 minutes followed by 95°C for 2 minutes and 30 cycles of 95°C for 15 seconds, 60°C for 30 seconds and 72°C for 30 seconds.

The amplified reaction was added to index PCR mix containing forward and reverse index primer pairs and the reaction was incubated for 2 min at 95°C then 2 cycles of 95°C for 15 seconds, 60°C for 60 seconds and 72°C for 60 seconds. The indexed PCR products were diluted into new PCR mixes containing dNTPs, SYBR Green I, PCR primers, phusion buffer and programmed for an initial 2 minutes at 95°C, followed by 15 cycles of 95°C for 15 seconds, 60°C for 30 seconds, and 72°C for 30 seconds. The PCR products were pooled and purified using QIAquick PCR purification kit. The purified DNA was sequenced using MiSeq Reagent kit v2, 300 cycles by NextSeq Reagent Kit 75SE, NextSeq (Illumina).

Data analysis

BCL files were converted to fastq format using bcl2fastq (Illumina) with pair indexes. The different Tags were removed from three fixed segments within each read. Reads with only one

count were removed. The Tags were sorted into their corresponding complex Tag, protein Tag and moleculeTag using a Vesicode in-house Perl Script. The number of S-EEVs with given combinations were calculated using a Vesicode in-house developed R-script. The t-SNE algorithm was implemented in the R package Rtsne and all parameters were set to default.

A more detailed protocol of the steps can be found in the paper by Wu et al., 2019[176]. The proprietary scripts and pipelines used for the data analysis were developed by Vesicode and are not accessible.

STRING analysis was used to determine the connection between the top 10 expressed proteins in each cluster. The proteins were inserted into the online tool (version 11.5) and checked against the *homo sapiens* genes. The figures representing the protein connections were used.

Results

Morphological assessment of embryos

Samples were chosen based on the morphology of the day 3 embryos as either good-quality embryos that were transferred to the uterus or poor-quality embryos that were not transferred. For the good-quality embryos, 2 embryos were simultaneously transferred into the uterus to increase chances of pregnancy. Each sample analysed for PBA was the pooled culture media from 2 embryos. The good-quality embryos had uniform cells, low levels of fragmentation, low levels of granulation and the cells seemed to be healthy and alive (representative images in Figure 5.2). The poor-quality embryos had high levels of fragmentation with non-uniform cells, and high cellular degradation (representative images in Figure 5.2).

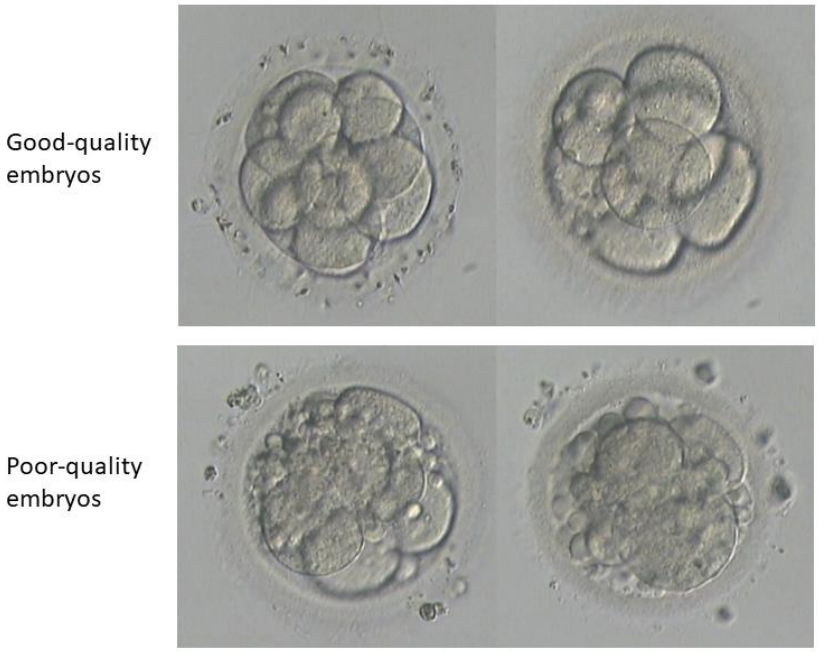

Figure 5.2: Representative images of good-quality and poor-quality embryos on Day 3 of development. The top row is one pair of the good-quality embryos whose culture media was pooled and used. The bottom row is one pair of the poor-quality embryos used for PBA.

PBA data quality control and normalisation

The samples were then analysed using PBA as outlined above. Once the data was obtained, the sequences underwent quality control using fastQC. The quality control measures are listed in table 5.2.1.

Table 5.2.1: Quality control measures of all the samples and controls using fastQC.

SampleID	Clean Q20(%)	Total Reads	Mapped Reads	Unique Reads	EV Count	Unique Reads:Mapped Reads	EV Count:Protein
IVF_CON1	99.01%	151703	129341	50482	28326	01:02.6	01:01.8
IVF_CON2	99.13%	172469	146296	47700	26646	01:03.1	01:01.8
PBS_G11	99.14%	1029496	872773	77530	42967	01:11.3	01:01.8
I_1	99.16%	1504819	1287191	148442	64584	01:08.7	01:02.3
I_2	99.17%	470162	402653	66325	34936	01:06.1	01:01.9
I_3	99.18%	1630844	1371365	1011225	133543	01:01.4	01:07.6
II_4	99.21%	1513115	1276082	178544	74794	01:07.1	01:02.4
II_5	97.99%	159305	123162	68953	43240	01:01.8	01:01.6
II_6	99.21%	634384	534711	130578	60453	01:04.1	01:02.2

The good quality samples (I_1, I_2, and I_3) showed a slightly higher average EV count, protein count and protein to EV ratio compared to the controls (IVF_CON1, IVF_CON2, and PBS_G11) and poor-quality samples (II_4, II_5, and II_6) (Table 5.2.1). Although, these were not significantly different. The samples were put through normalisation and the data distribution prior to and after normalisation is shown in figure 5.3.

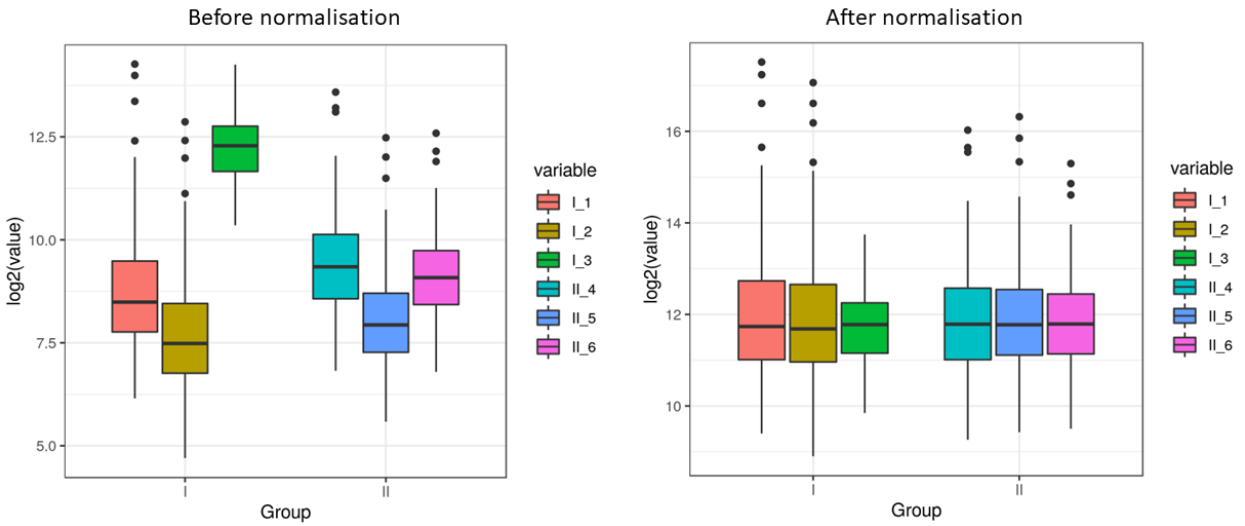

Figure 5.3: The normalisation of the dataset prior to further analysis. Normalisation was completed using the FastQC quality control tool.

Differentially expressed proteins

All the identified proteins in the samples are shown using a dot plot. All the identified proteins were expressed in all samples with some proteins having higher expressions which are outlined with a larger dot and a redder colour on the plot (Figure 5.4).

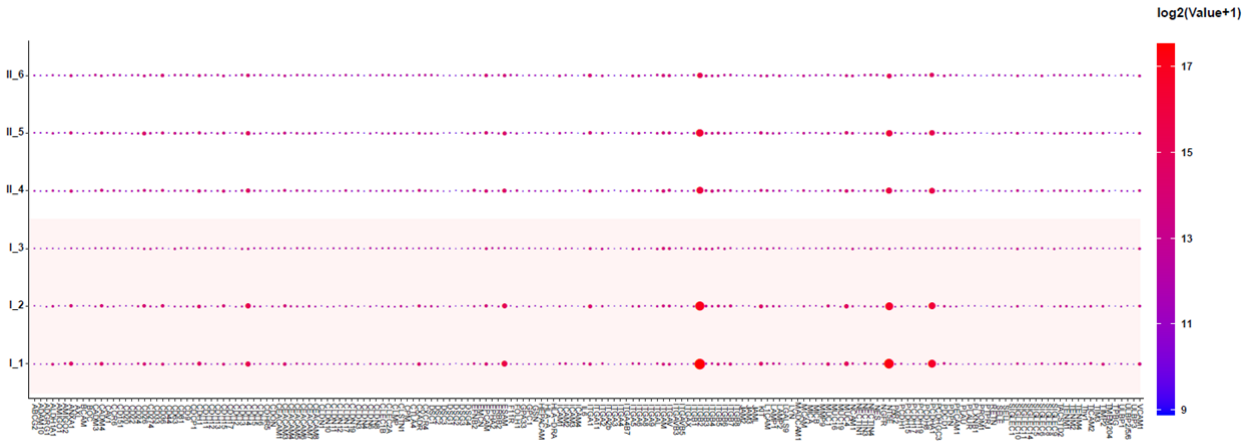

Figure 5.4: Dot plot representing all the proteins detected on the S-EEVs in all 6 samples. I_1, I_2, and I_3 are from good-quality embryos. II_4, II_5, and II_6 are from poor-quality embryos. The proteins are listed on the y-axis. The colour from the log2 scale corresponds to the expression levels and the larger the dot, the higher the expression in that specific sample.

Comparing the good-quality embryos to poor-quality embryos, 7 proteins were identified, using a heatmap, that were differentially expressed (Figure 5.5.A). CDON, GPA33, and CD20 were more highly expressed by the good-quality embryos. CD9, PCDH8, CCR6 and CEACAM1 were more highly expressed by the poor-quality embryos. For CDON, samples I_1 and I_2 showed a higher expression out of the 3 pooled good-quality embryos (Figure 5.5.A). As for GPA33 and CD20, sample I_2 had higher expression compared to the other 2 samples (Figure 5.5.A). The poor-quality samples had a more uniform expression across the samples for the 4 identified proteins (Figure 5.5.B). The volcano plot shows the log fold change of the 7 identified proteins within the entire proteins that were detected with respect to the good-quality embryos (Figure 5.5.B).

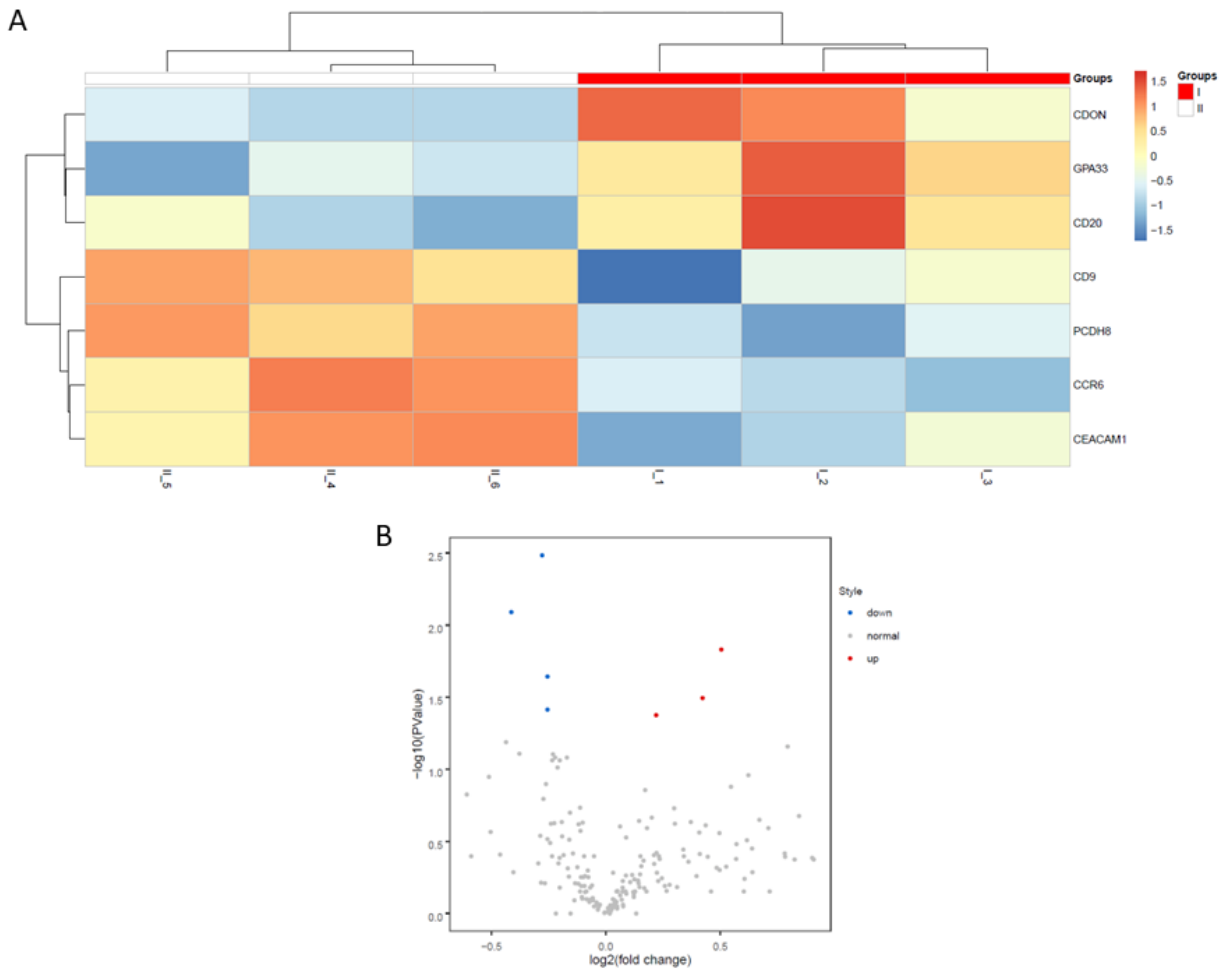

Figure 5.5: The differentially expressed proteins between the good-quality embryos and poor-quality embryos. **A)** Heatmap of the 7 proteins with the colours corresponding to log2 fold change. **B)** The volcano plot of the proteins detected in the good-quality embryos showing the 7 proteins and the corresponding fold change.

The good-quality embryos showed a significantly higher protein expression (log2 fold change) for CD20 (10.73 vs 10.51, p=0.040), CDON (12.22 vs 11.81, p=0.026), and GPA33 (10.56 vs 10.05, p=0.015) compared to the poor-quality embryos (Figure 5.6). The poor-quality embryos showed a significantly higher protein expression for 4 proteins (log2 fold change). The identified proteins were PCDH8 (12.01 vs 11.73, p=0.0041), CEACAM1 (11.94 vs 11.68, p=0.0235), CD9 (12.57 vs 12.31 p=0.0492) and CCR6 (12.51 vs 12.11, p=0.0068) compared to the good-quality embryos (Figure 5.6).

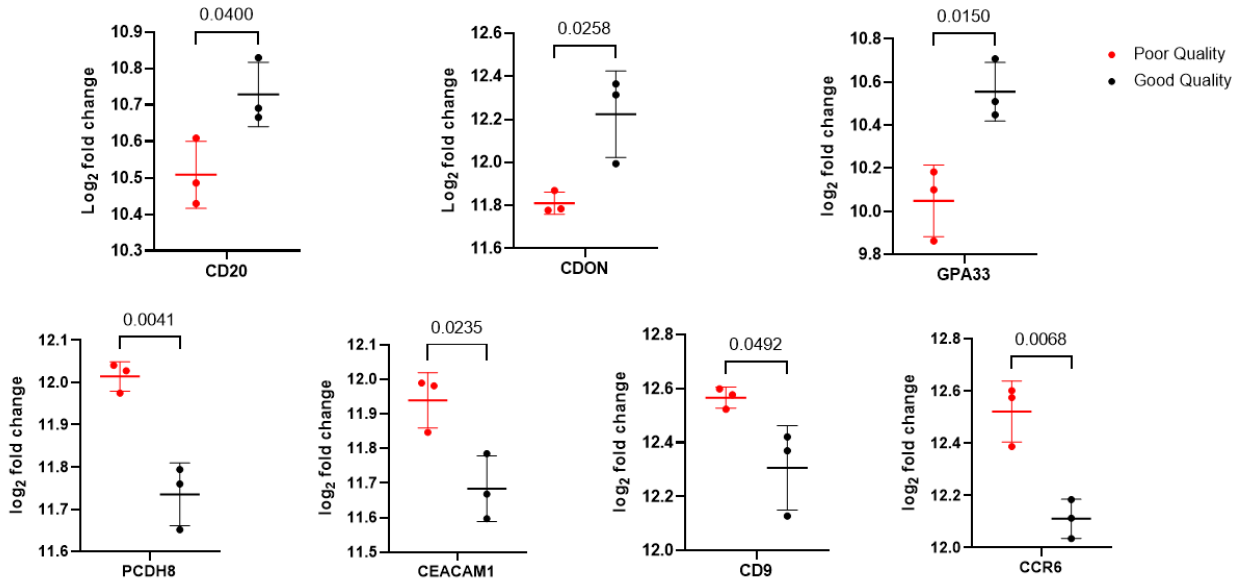

Figure 5.6: The fold change between good-quality and poor-quality embryos for the 7 differentially expressed proteins. The top row are upregulated in good-quality embryos and the bottom row are upregulated in poor-quality embryos. n=3 biological replicates for each group. The p-value is indicated on top of the line of each graph. Graphs were generated using GraphPad Prism.

Cluster and STRING analysis

From here, the aim was to identify subpopulations of S-EEVs based on protein expression. Clustering analysis was completed using the FLOWSOM algorithm using tSNE plots to reduce the dimensionality of the results. The clustering of the S-EEVs can be seen for all samples with sample I_3 showing very different results with more S-EEVs identified compared to the other samples but it is difficult to conclude if I_3 is an outlier or if there is an underlying reason for the difference due to the small sample size (Figure 5.7).

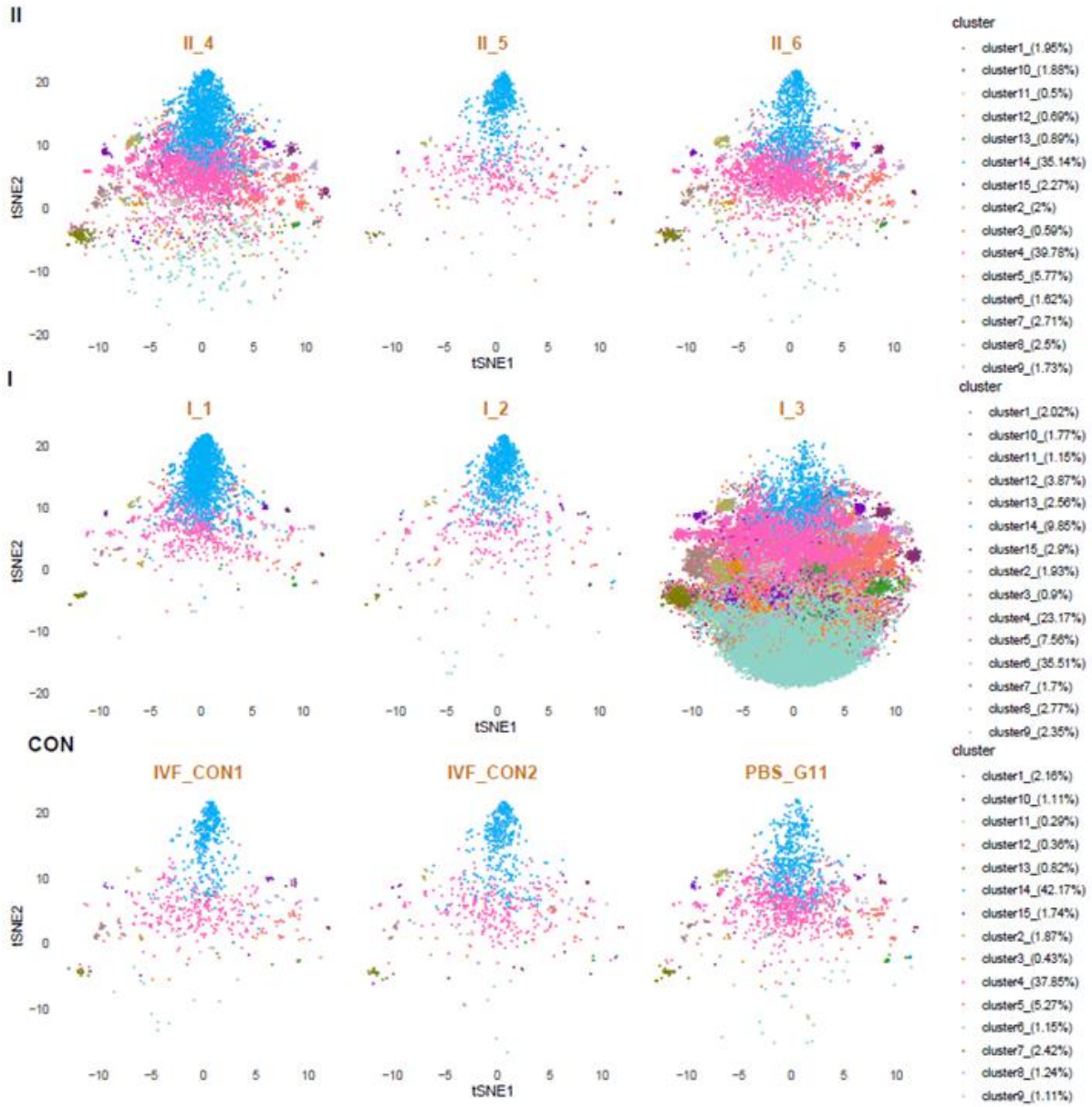

Figure 5.7: FLOWSOM algorithm was used to produce tSNE plots and show the identified clusters within each sample. The corresponding EV percentages in each cluster within the sample types are shown on the right. The top row corresponds with the poor-quality embryonic samples, the middle row is the good-quality and the bottom row are the controls. Each dot on the tSNE is one S-EEV and each S-EEV has its featured proteins which were used to assign that S-EEV to a specific cluster.

The clustering data is not conclusive because there was a low concentration of S-EEVs in all samples making it challenging to visualise clear clusters, suggesting that the detected proteins were shared between the samples and not exclusive to a subpopulation in any one sample type. There are two possibilities for the results, one is that this method of analysis may not be appropriate for

these samples or alternatively, that more samples need to be tested to reach a conclusive result. Nevertheless, the rest of this pilot analysis was completed to provide a framework of what the analysis would look like for future studies using either more samples or different samples. From the clustering analysis, 15 clusters were identified. For each cluster, the distribution of each sample type within the cluster was determined and the top 10 proteins were identified; STRING analysis was performed on the top 10 proteins to better understand their protein-protein interactions.

In cluster 1, 77.51% of the S-EEVs in the cluster belonged to group I (good-quality embryos) compared to the control (5.71%) and group II (16.78%, poor-quality embryos) (Figure 5.8.1.B). The top 10 proteins found on the S-EEVs in cluster 1 are shown with the two top proteins being identified as ITGA3 and CDH2 with approximately a frequency of 0.4 (Figure 5.8.1.C). The STRING analysis showed high degrees of interactions with the ITG proteins and no interactions with SIGLEC8 (Figure 5.8.1.D).

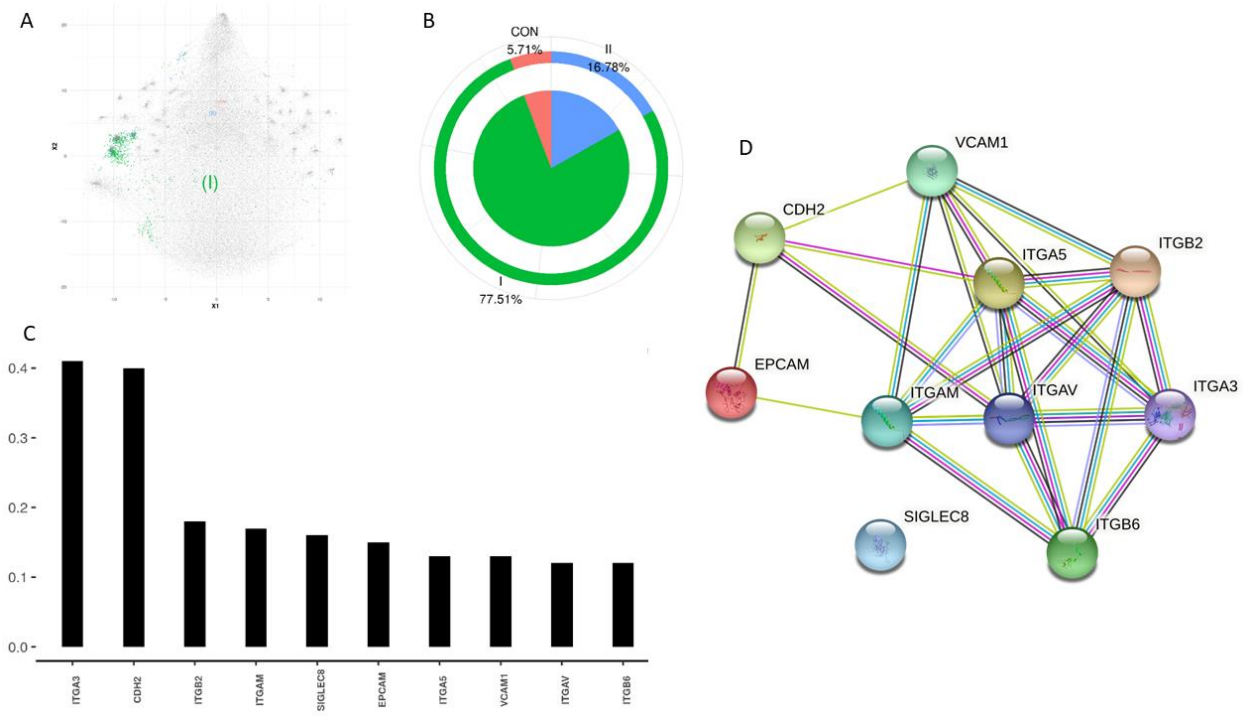

Figure 5.8.1: Cluster 1 from the tSNE analysis. **A**) The position of cluster 1 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

In cluster 2, 77.02% of the S-EEVs in the cluster are found in samples from the good-quality group. The most highly abundant protein identified on the S-EEVs is DSG1 (frequency of 0.3) and all of the top 10 proteins interact with each other with high interactions between the ITG proteins (Figure 5.8.2).

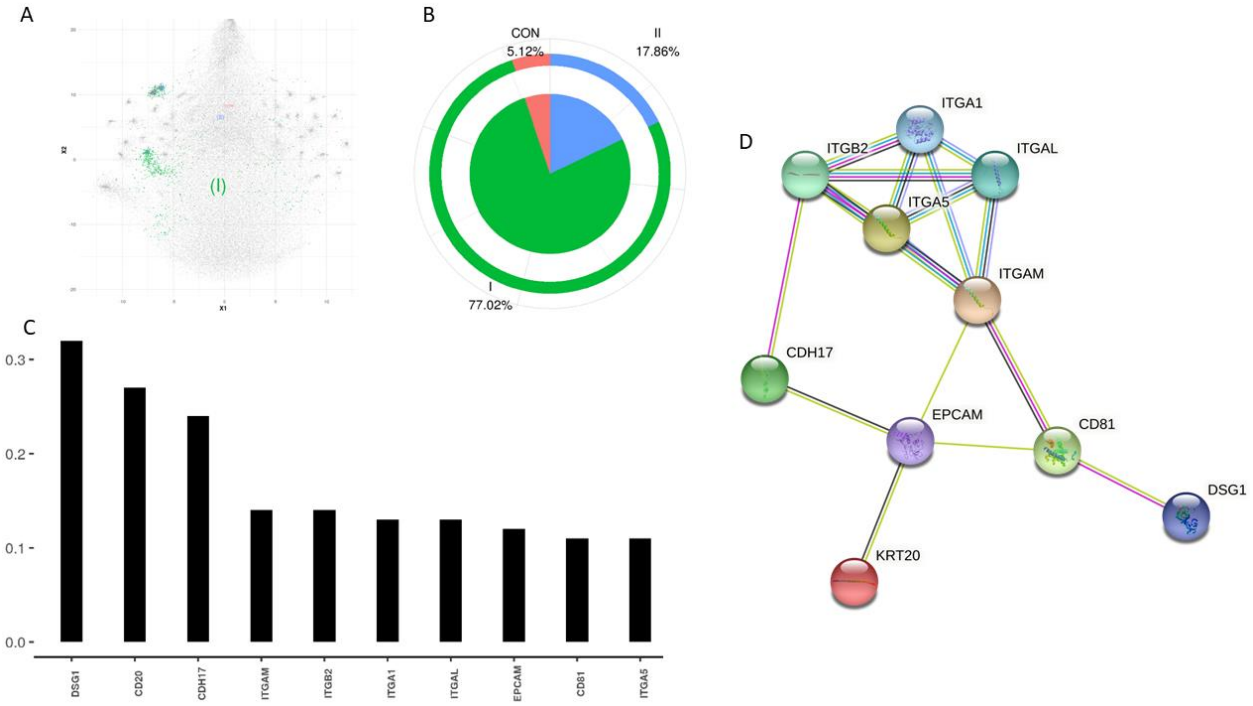

Figure 5.8.2: Cluster 2 from the tSNE analysis. **A**) The position of cluster 2 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

In cluster 3, 84.78% of the identified S-EEVs were found in group I. The most abundant protein found on the S-EEVs in this cluster was DSG2 with a frequency of over 0.6. From the STRING analysis, the ITG proteins had high interactions with each other. DSG2 only interacts with EPCAM which interacts with ITGAM, connecting it to the ITG family of proteins (Figure 5.8.3).

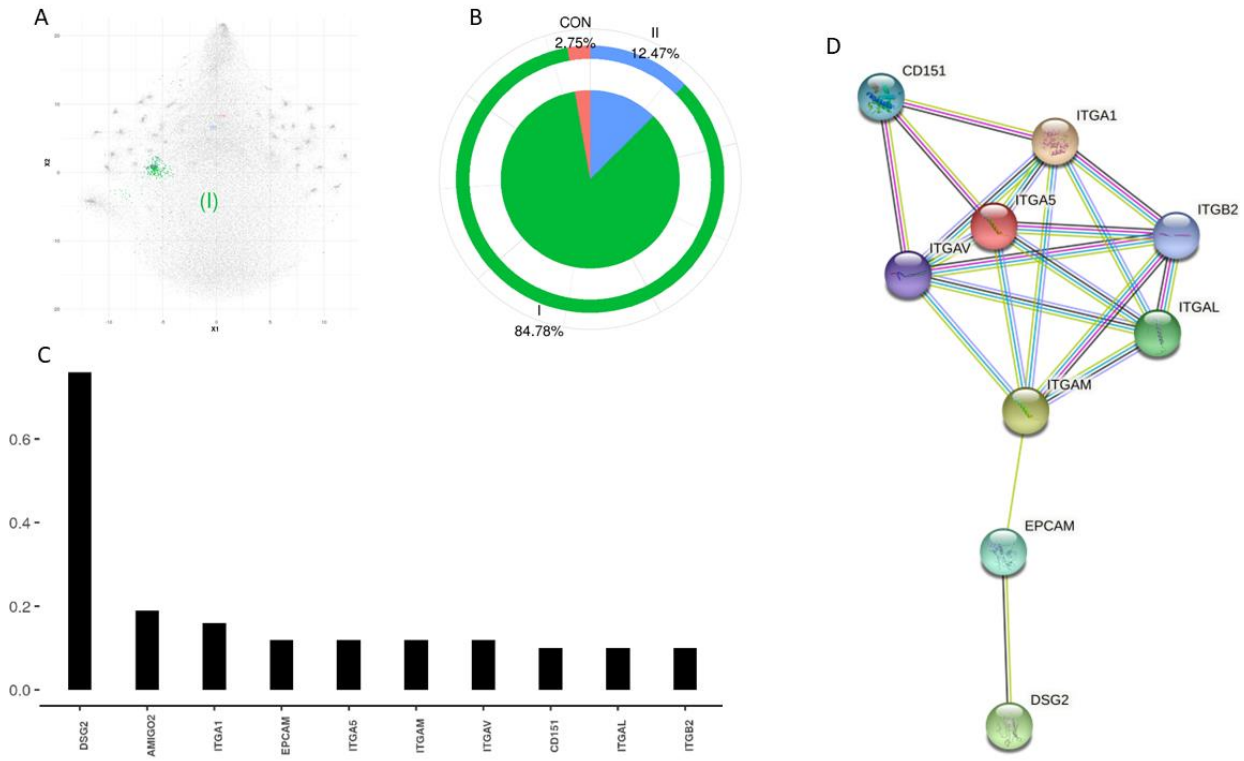

Figure 5.8.3: Cluster 3 from the tSNE analysis. **A**) The position of cluster 3 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

Cluster 4 showed that 66.79% of the S-EEVs were found in group I. The top 10 proteins in this cluster were more uniformly expressed with the highest expressed being ITGAM. The ITG proteins again showed high interactions with each other. SIGLEC 10 and CDH showed no interactions with the other top proteins in the cluster (Figure 5.8.4).

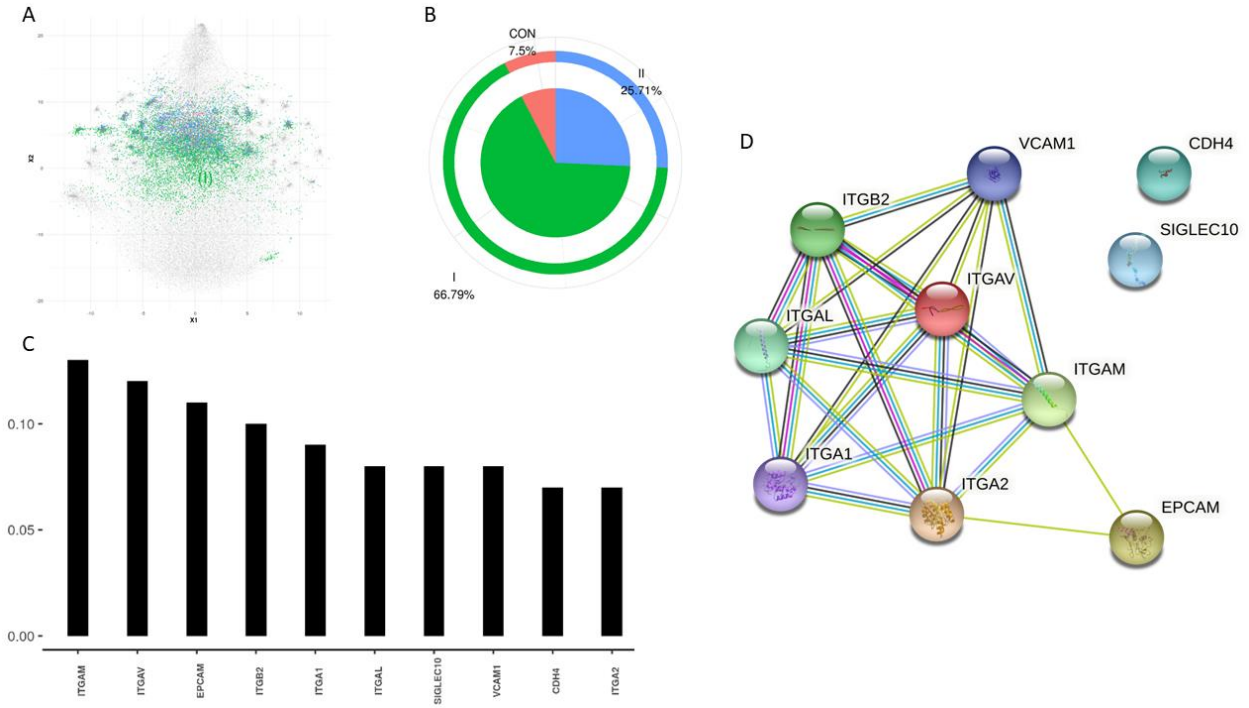

Figure 5.8.4: Cluster 4 from the tSNE analysis. **A**) The position of cluster 4 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

The good-quality embryos account for 82.03% of the S-EEVs found in cluster 5. The top 10 proteins expressed on the S-EEVs in this cluster are uniformly present with ITGAM and MICA having slightly higher frequency. The STRING analysis again confirms the strong interactions between the ITG proteins; whereas, MICA, GSN and TPBG showed no interactions (Figure 5.8.5).

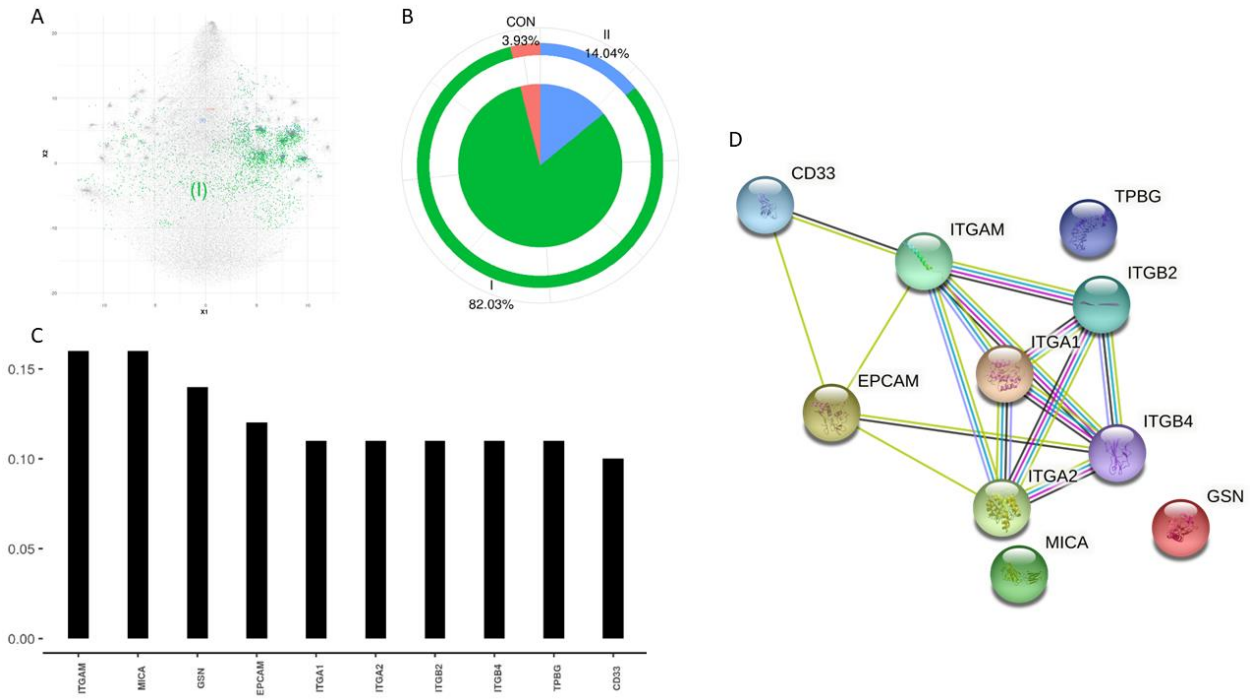

Figure 5.8.5: Cluster 5 from the tSNE analysis. **A**) The position of cluster 5 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

In cluster 6, the S-EEVs are mostly represented by the good-quality samples with 98.77% of the S-EEVs belonging to this group. The top 10 proteins have uniform expression in the cluster and have high degrees of protein interactions with the exception of SIGLEC10 (Figure 5.8.6).

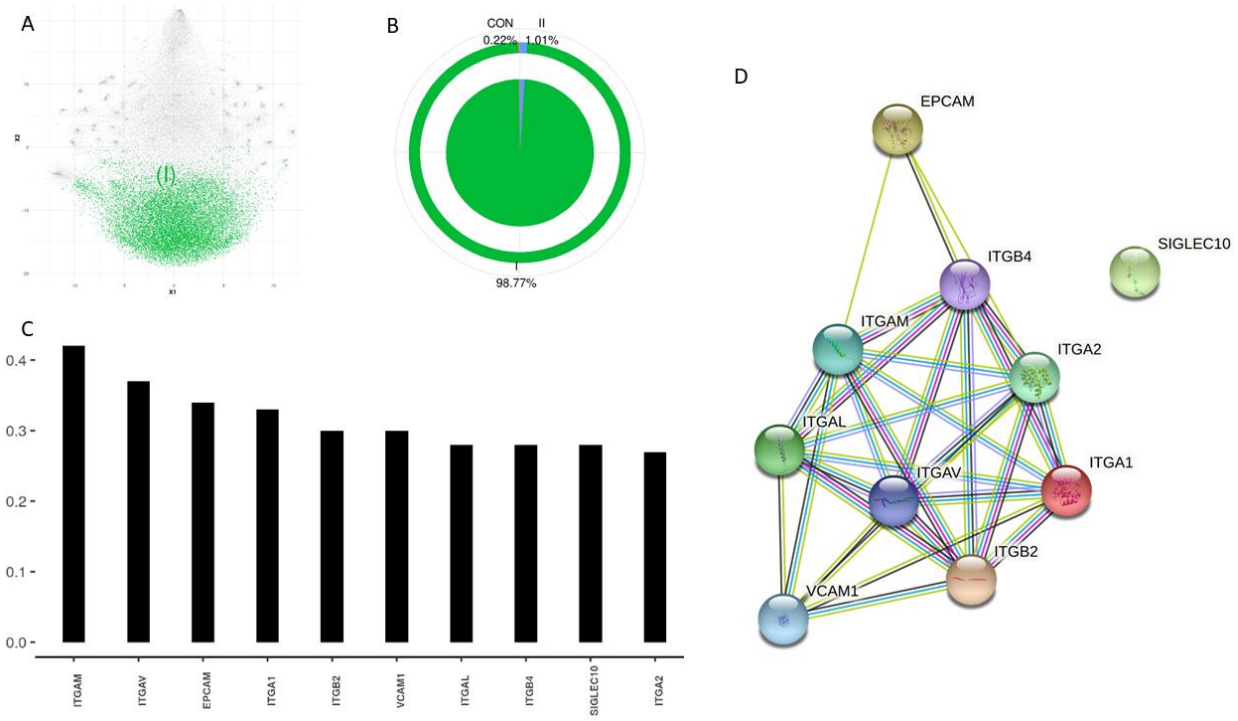

Figure 5.8.6: Cluster 6 from the tSNE analysis. **A**) The position of cluster 6 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

From the sample groups, the good-quality embryos account for 68.7% of the S-EEVs in cluster 7.

The main protein expressed on the EEVs in this cluster is PCDH17 with an approximate frequency of 1. However, PCDH17 does not have any interactions with the other top 10 proteins of the cluster (Figure 5.8.7).

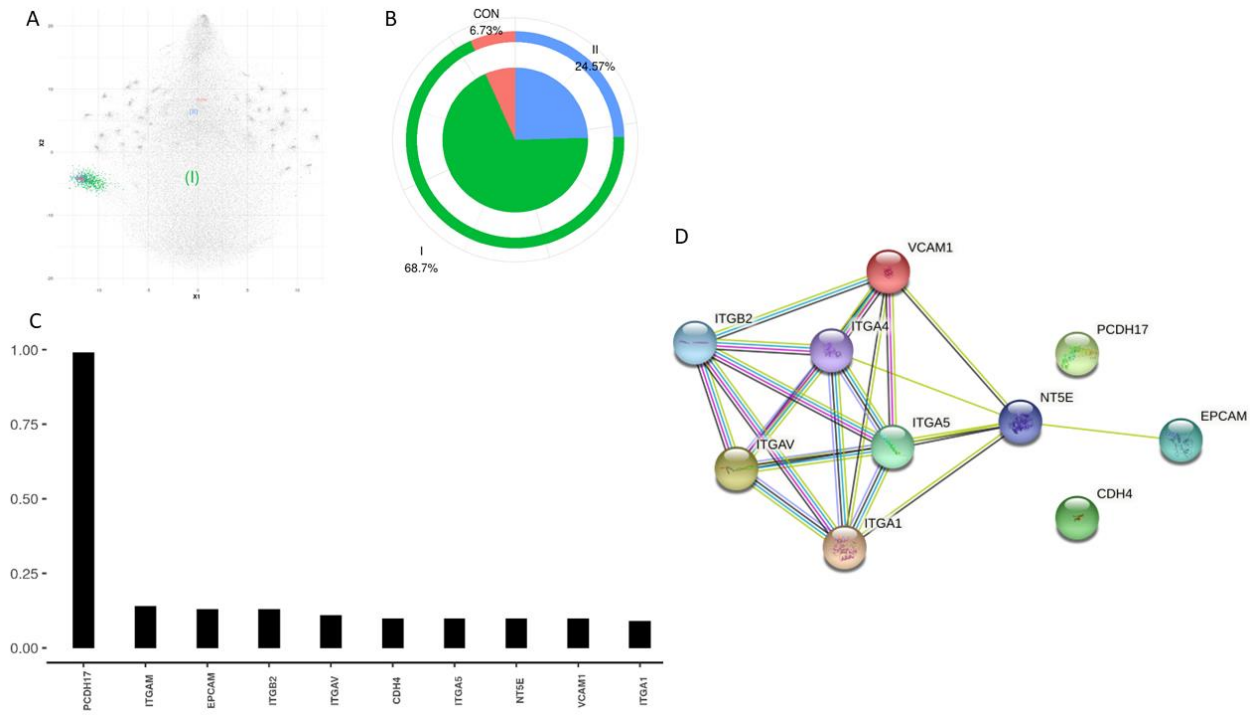

Figure 5.8.7: Cluster 7 from the tSNE analysis. **A**) The position of cluster 7 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

In cluster 8, 81.08% of the S-EEVs were from good-quality samples with ADAM10 having the highest frequency between the top 10 proteins identified on the S-EEVs in this cluster (Figure 5.8.8).

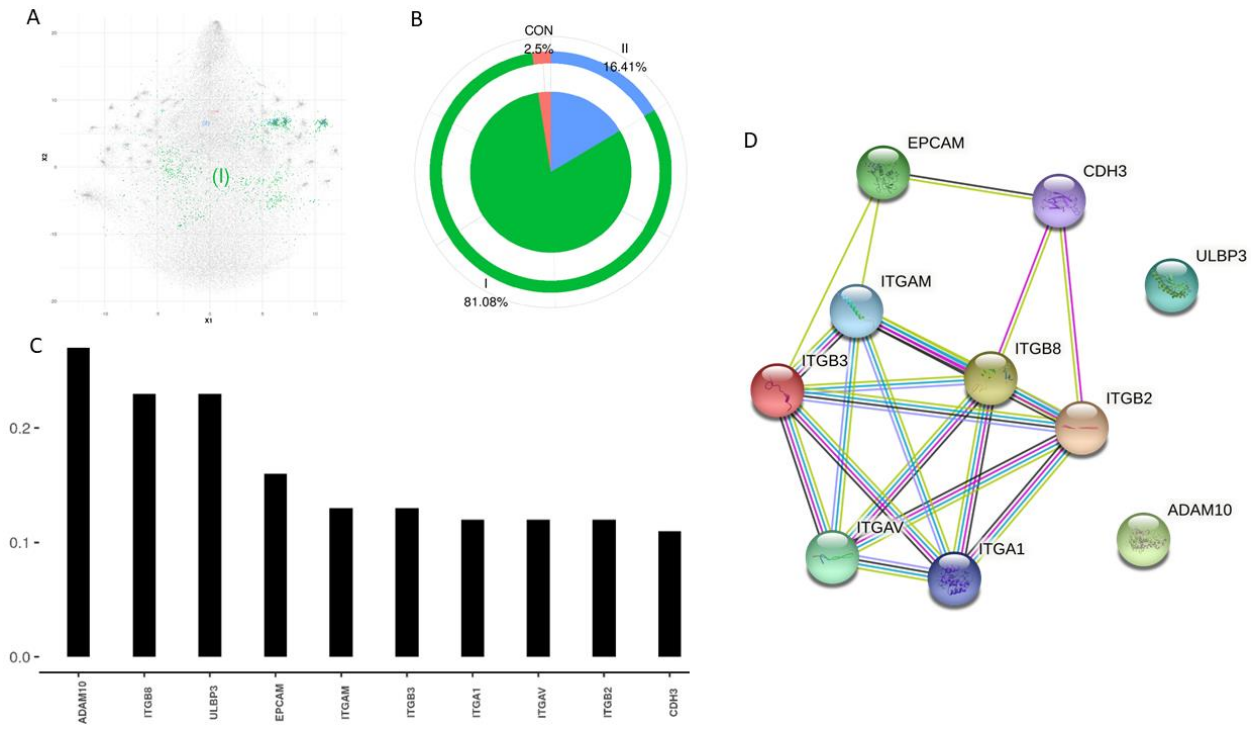

Figure 5.8.8: Cluster 8 from the tSNE analysis. **A**) The position of cluster 8 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

Cluster 9 showed a more concentrated distribution on the tSNE plot, with 83.52% of the S-EEVs belonging to the good-quality embryonic samples. TENM2 had the highest frequency of 0.3 compared to the other 10 proteins and had protein-proteins interactions with ITGAV, connecting it to the ITG protein network (Figure 5.8.9).

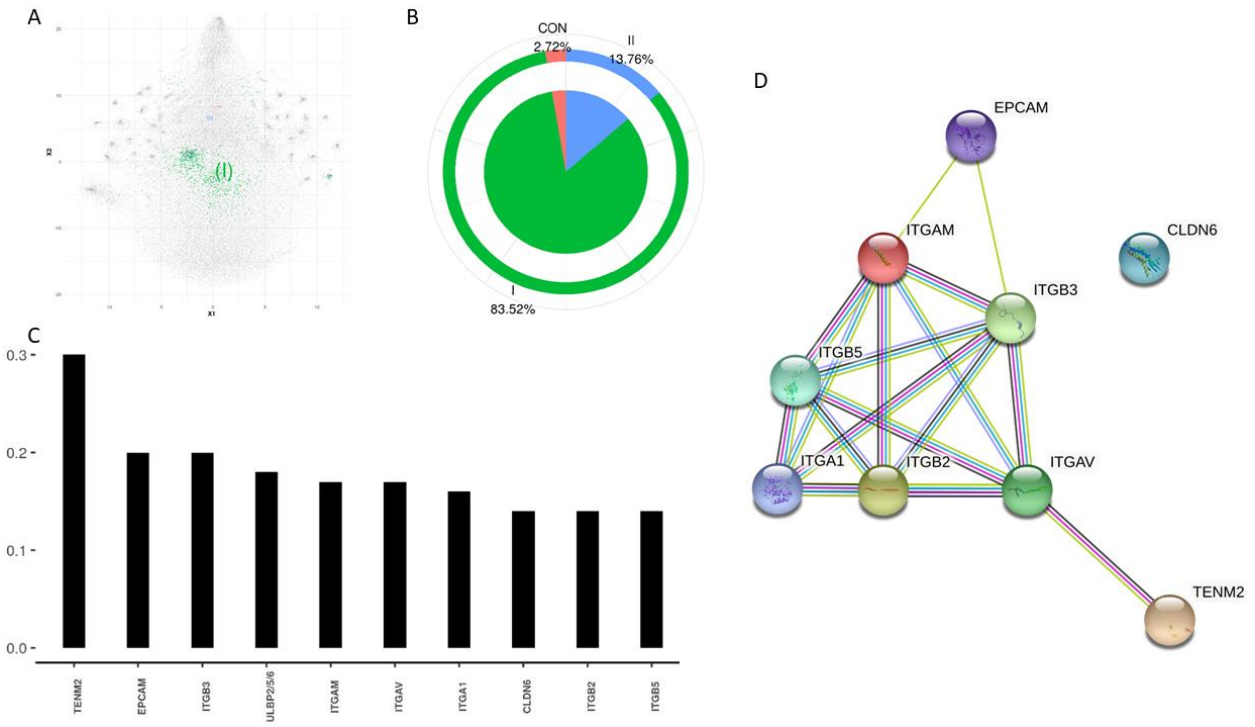

Figure 5.8.9: Cluster 9 from the tSNE analysis. **A**) The position of cluster 9 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

Cluster 10 showed localised dispersed clusters on the tSNE, with 78.05% of the S-EEVs belonging to the good-quality group. RETN had the highest frequency of approximately 0.5 in this cluster and had interactions with ITGAM (Figure 5.8.10).

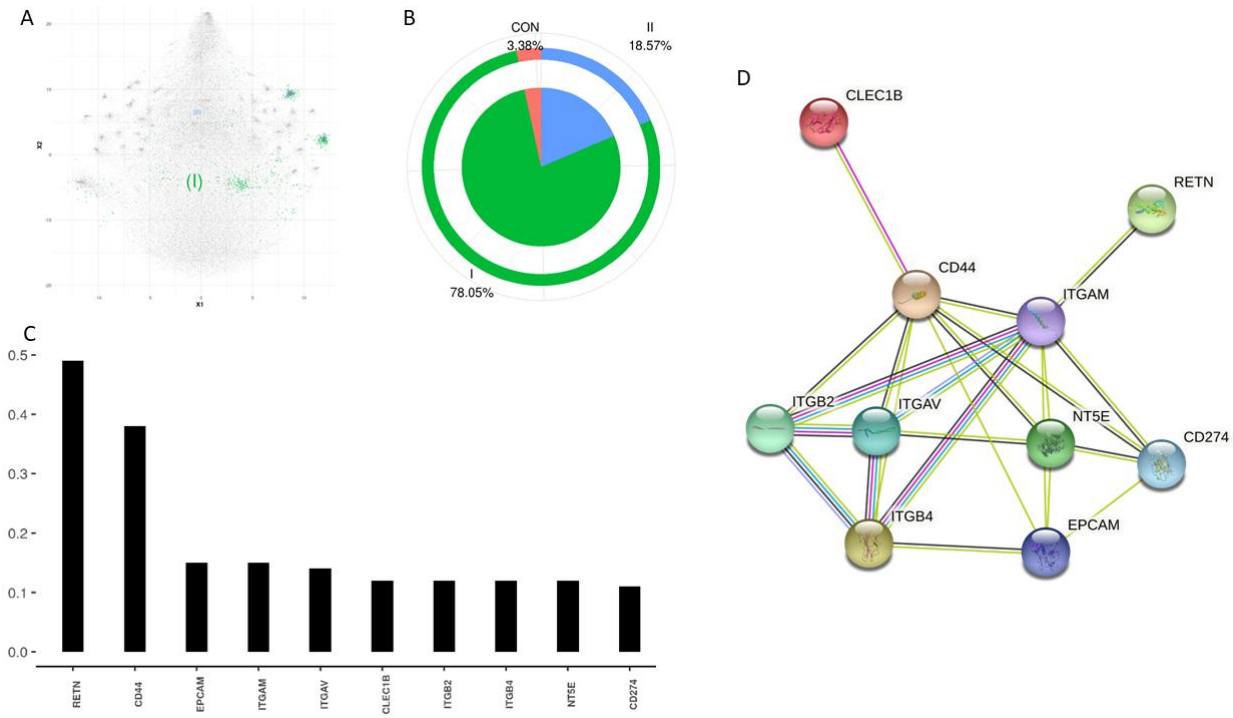

Figure 5.8.10: Cluster 10 from the tSNE analysis. **A**) The position of cluster 10 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

In cluster 11, 89.67% of the S-EEVs belong to the good-quality group with 8.76% from the poor quality and 1.58% from the controls. The most frequent protein on the S-EEVs in this cluster was JAM2 (frequency of approximately 0.45) which was highly connected to the ITG network (Figure 5.8.11).

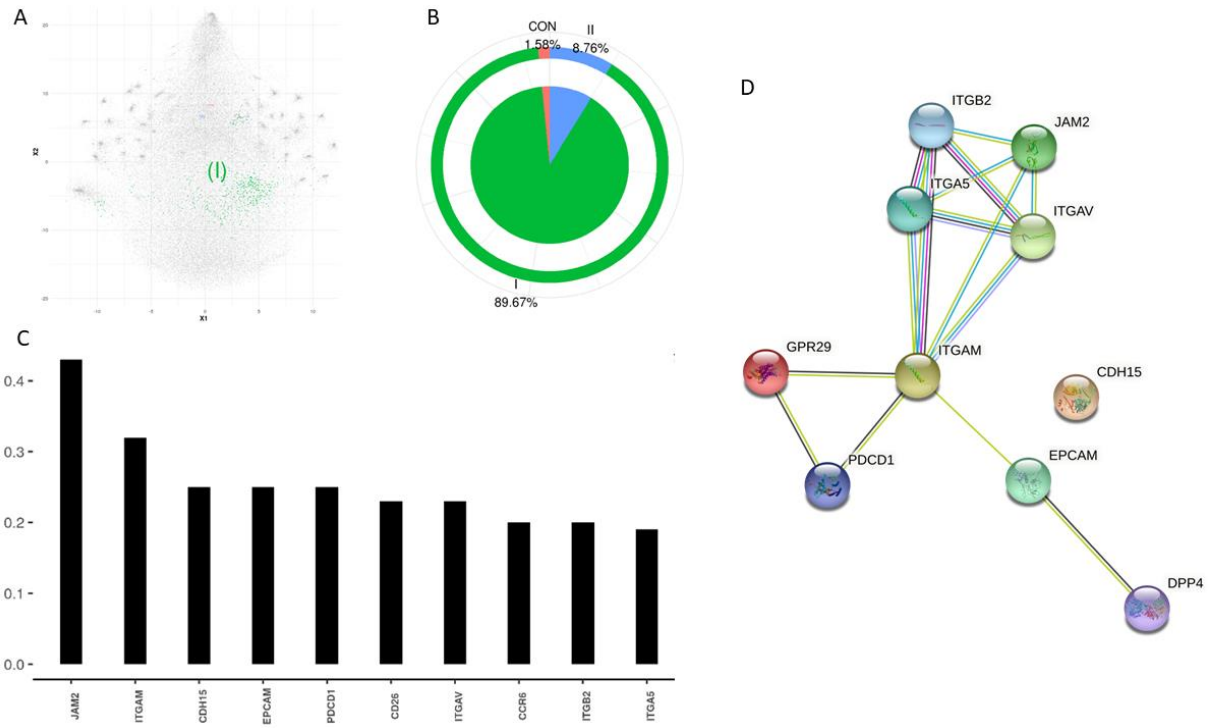

Figure 5.8.11: Cluster 11 from the tSNE analysis. **A**) The position of cluster 11 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

The good-quality group made up 95.56% of the S-EEVs found in cluster 12. ITGAM had the highest frequency of the proteins found on the S-EEVs in this cluster and was connected to the other ITG proteins identified in this cluster. SIGLEC10 and EMCN did now show any interactions with the other top 10 proteins in the cluster (Figure 5.8.12).

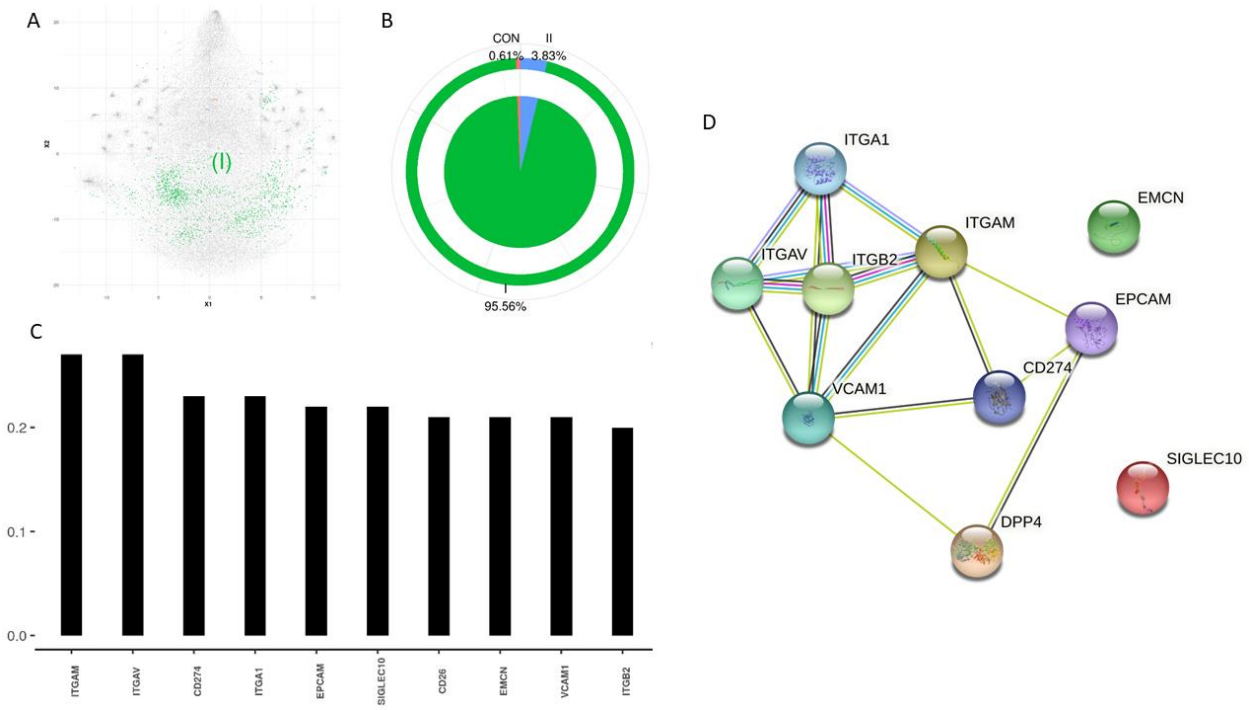

Figure 5.8.12: Cluster 12 from the tSNE analysis. **A**) The position of cluster 12 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

The majority of the S-EEVs found in cluster 13 were also represented by the good-quality group with 90.9% of the S-EEV population. LYN had the highest frequency of approximately 0.3 and was found to have interactions with the ITG proteins in this cluster (Figure 5.8.13).

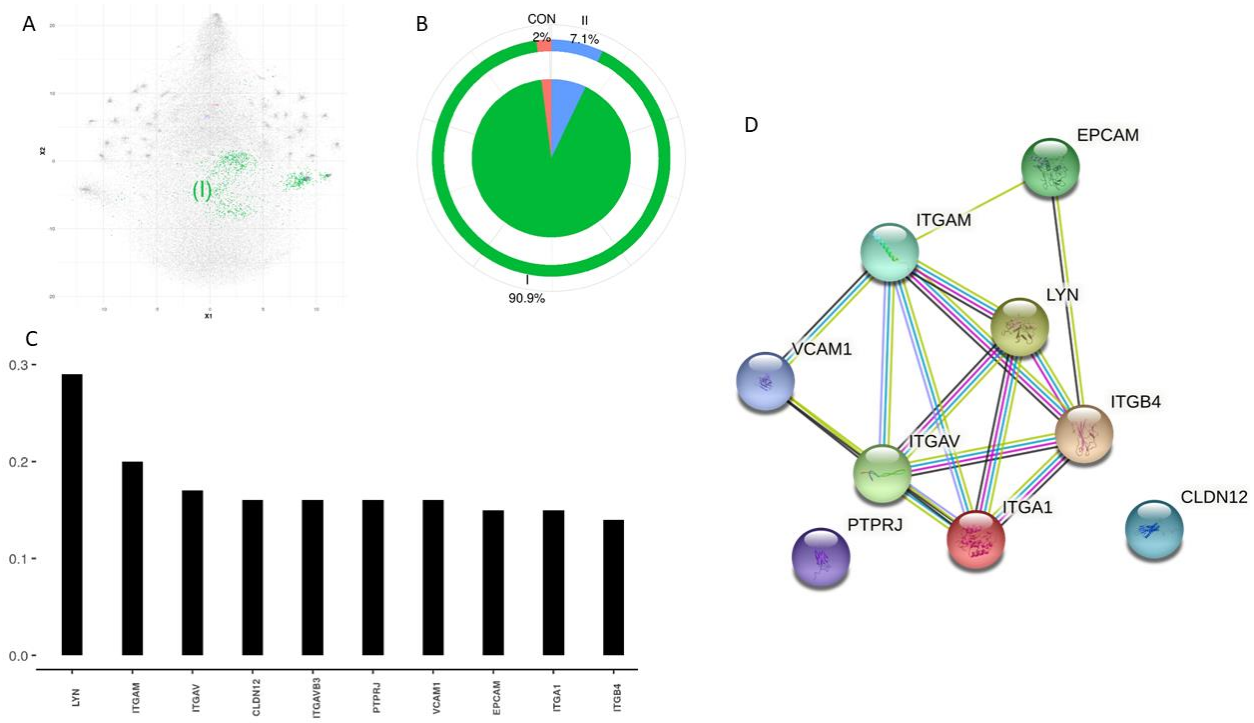

Figure 5.8.13: Cluster 13 from the tSNE analysis. **A**) The position of cluster 13 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

Cluster 14 had a more uniform expression between the sample types. The good-quality group made up 47.75% of the S-EEV population, 38.2% was expressed by the poor-quality group and the control group made up 14.05% of the S-EVs in this cluster. ITGB2 had the highest frequency out of the other proteins. However, 5 of the top 10 proteins had no interactions with each other (Figure 5.8.14). The fact that there were no interactions between the proteins is interesting because of the more dispersed expression of the proteins between the sample types compared to the other clusters. The top 10 proteins could be expressed on S-EEVs from different sample types and therefore, would not be expected to interact with each other.

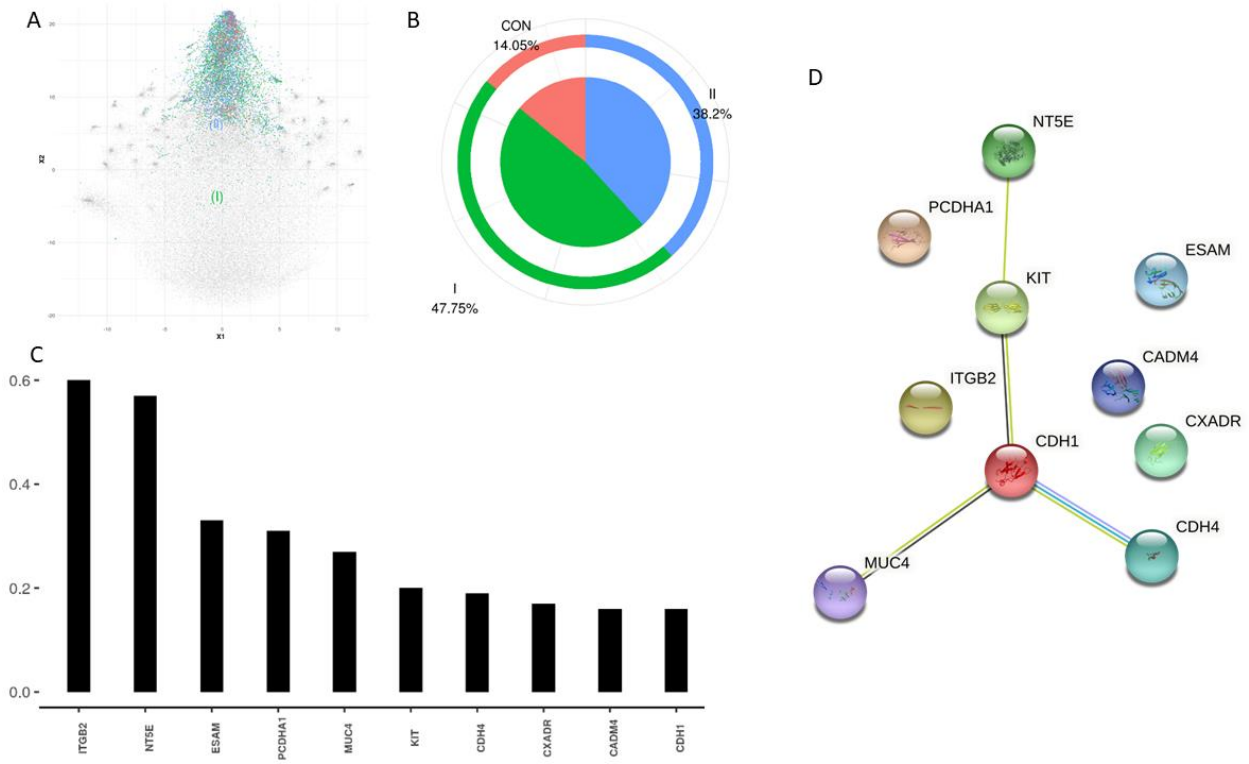

Figure 5.8.14: Cluster 14 from the tSNE analysis. **A**) The position of cluster 14 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

Lastly, the S-EEVs found in cluster 15 were again mostly from the good-quality group with 82.2% of the S-EEV population in this cluster. CLDN10 had the highest frequency and was found to have interactions with CLDN19 which was also in the top 10 proteins of this cluster (Figure 5.8.15).

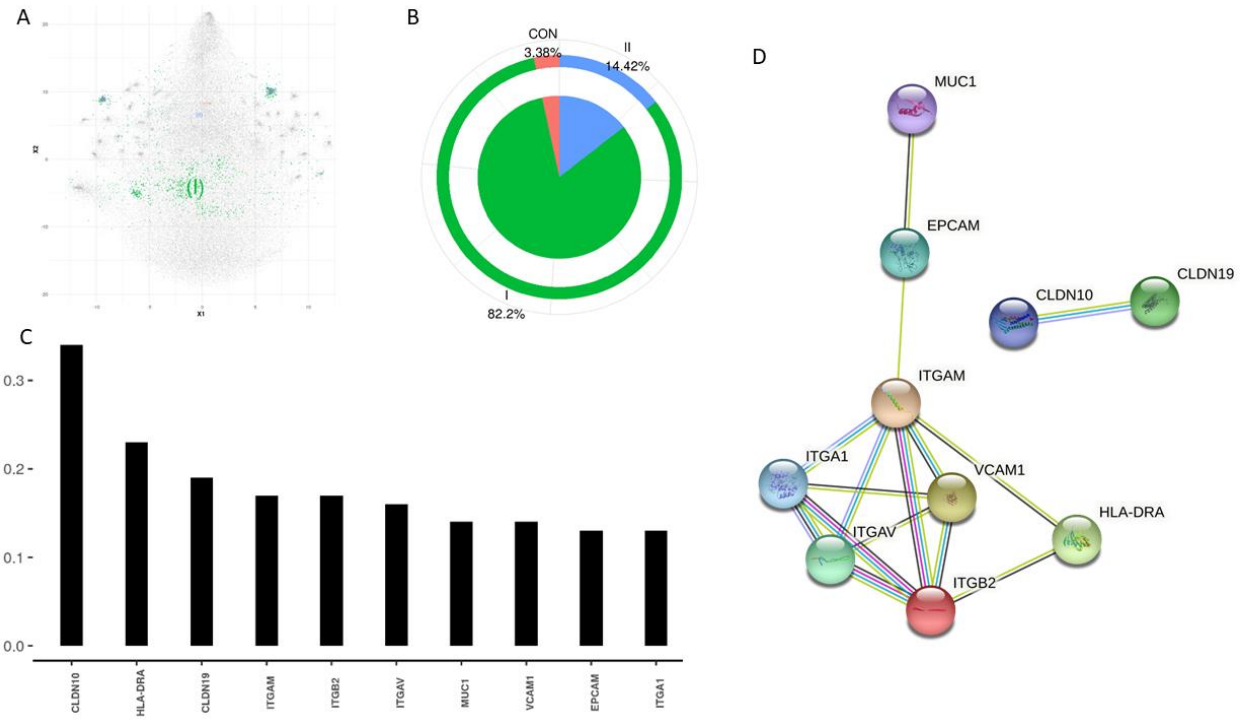

Figure 5.8.15: Cluster 15 from the tSNE analysis. **A**) The position of cluster 15 is outlined in the tSNE plot. **B**) The percentage of each sample group in the cluster is shown with the pie chart. Green is the percentage of the cluster from good-quality embryos, blue being the poor-quality and pink being the controls. **C**) The frequency of the top 10 proteins found on the S-EEVs in this cluster are shown by the black bars. **D**) On the right we have the STRING analysis network of the top 10 proteins showing the connections within the proteins. The lines of the STRING analysis each correspond to a different form of evidence for the interactions between the 2 proteins. Red line indicates the presence of fusion evidence, green line corresponds with neighborhood evidence, blue line is representative of cooccurrence evidence, purple line represents experimental evidence, yellow line is for textmining evidence, light blue line is for database evidence, and black line represents coexpression evidence.

To have a better visual of the top proteins found in all 15 clusters, a heatmap was produced. The proteins which had the highest expressions in each cluster were chosen for a total of 40 proteins. As can be seen on the heatmap, majority of the proteins are dispersed throughout the clusters with a few exceptions such as LYN (expressed mostly in cluster 13), HLA-DRA (mostly in cluster 15), ULBP3 (cluster 8), PCDH17 (cluster 7), RETN and CD44 (both in cluster 10). The ITG family of proteins were dispersed in most clusters and were the least concentrated which correlates to the wide spread of the clusters on the tSNE plot (Figure 5.9).

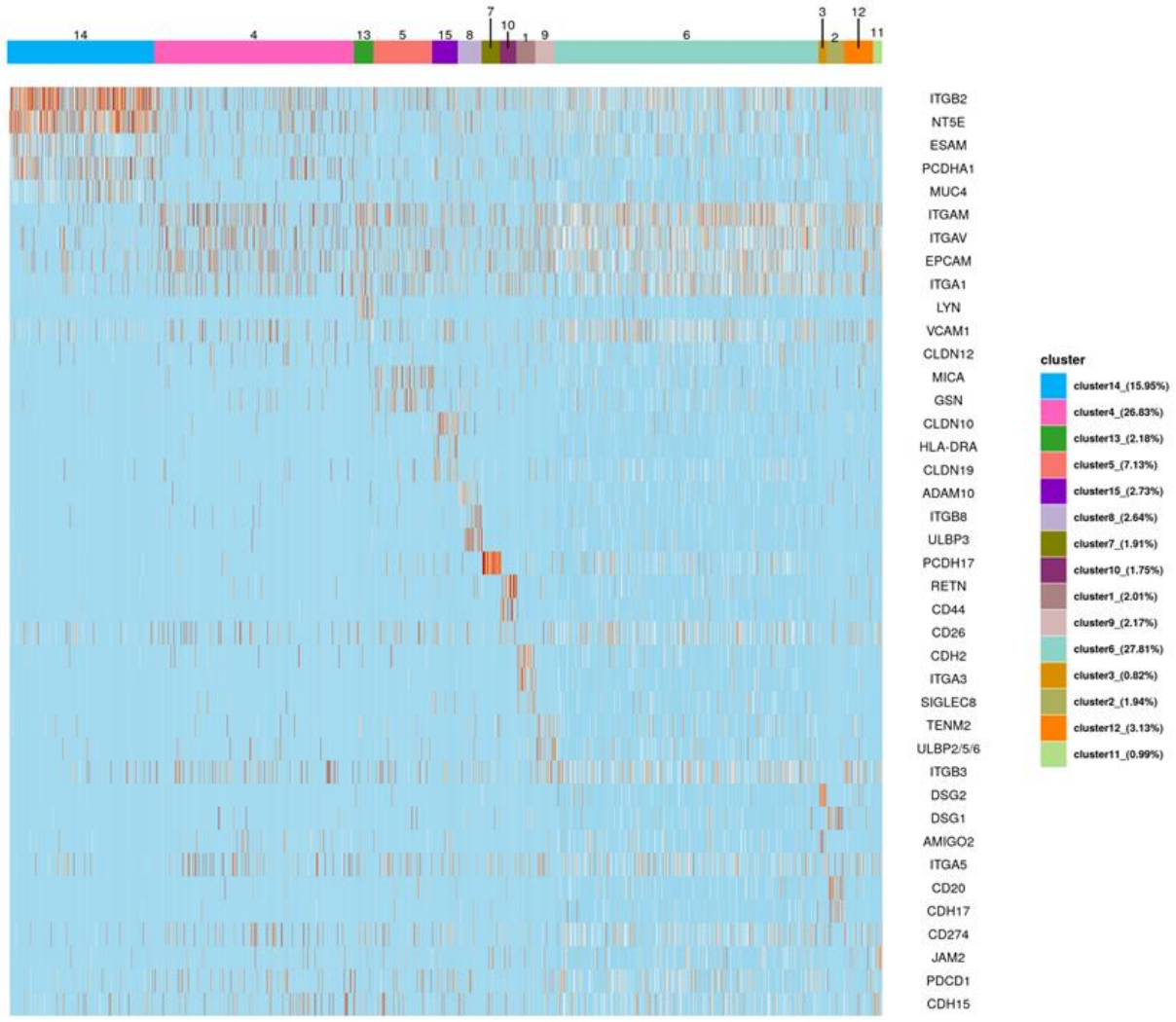

Figure 5.9: The expression of the top 40 proteins on the S-EEVs in the clusters. Each vertical blue line represents an S-EEV and if that S-EEV is expressing the protein outlined on the right, it is coloured orange. The intensity of the orange colour corresponds with the frequency of the protein in the cluster. The clusters are outlined on the top of the graph and their corresponding percentage representing their contribution to the tSNE plot of all samples is outlined on the right.

The high percentage of the S-EEVs in all clusters being associated with samples from good-quality embryos (group I) could be a result of sample I_3 skewing the results due to its high expression of S-EEVs. To account for this, sample three was removed from the differentially expressed proteins identified in the first part of the analysis to determine whether the 7 identified proteins would still be significantly different between the good-quality and poor-quality groups. Only CD20 was no longer significantly different, with all other proteins remaining significantly different between the

two groups. Therefore, I_3 does not seem to be the driving force for the differentially expressed proteins (Figure 5.10). However, more data is needed to corroborate these findings.

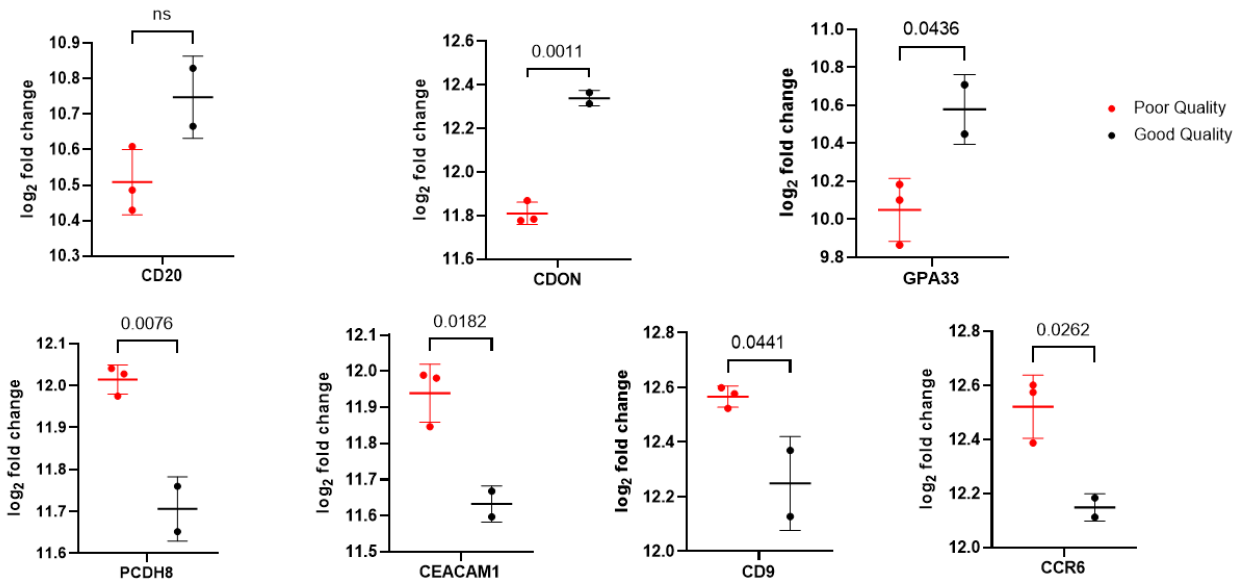

Figure 5.10: The differentially expressed proteins identified with the removal of sample I_3. The red represents the poor-quality embryos and the black represent the good-quality embryos. The ns stands for non-significant.

Top 25 protein combinations

The subsequent analysis was to analyse the possibility of the combination of 2 proteins being found on the same S-EEV comparing the two sample types. The top 25 combinations are shown using a heatmap with all the combinations being more differentially expressed by the good-quality group (Figure 5.11).

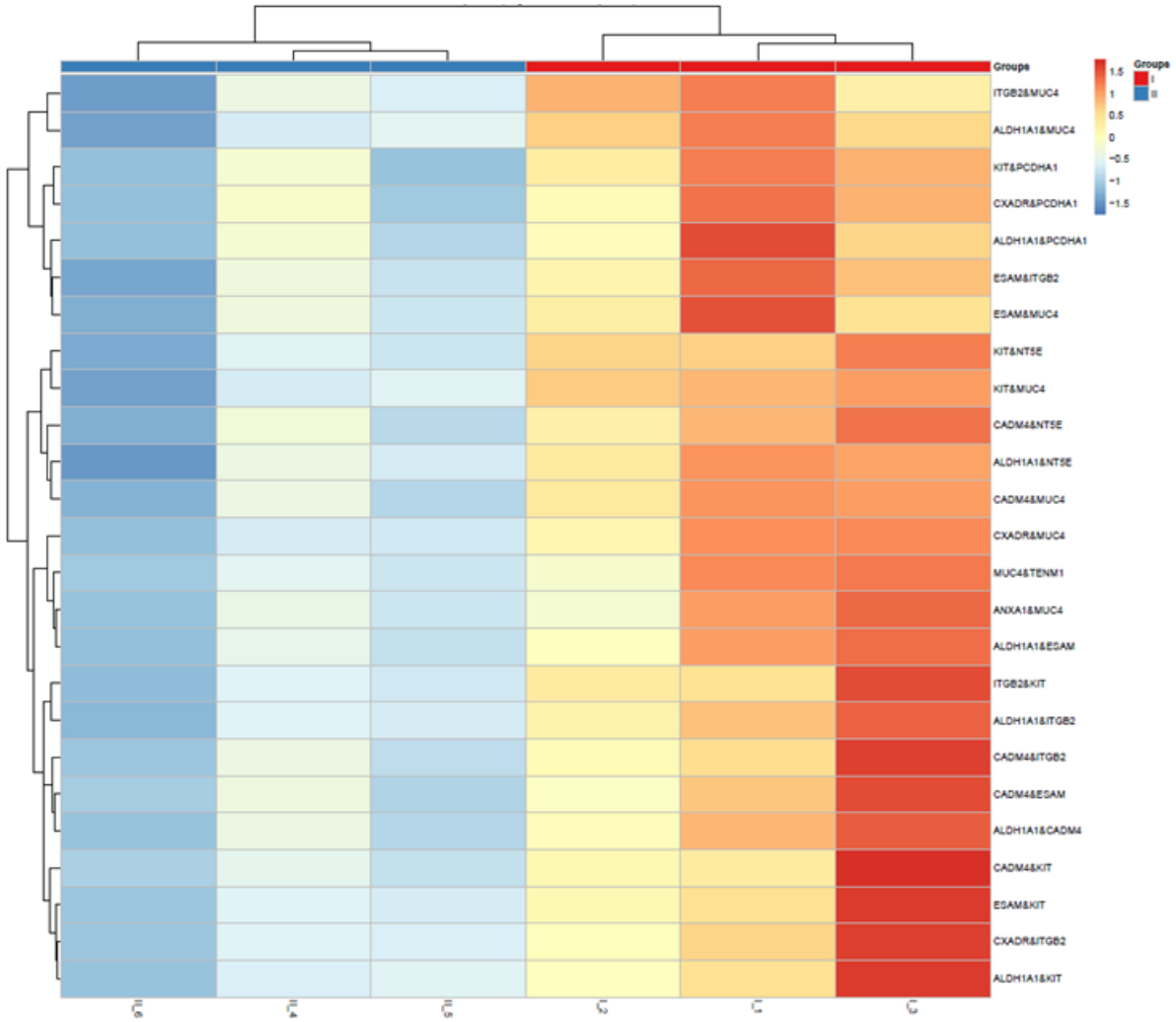

Figure 5.11: The heatmap of the top 25 protein combinations found on the S-EEVs. The more intense the red colour of the bar, the more expressed that protein combination is in that specific sample. The samples are denoted on the bottom of the graph.

In this analysis, sample I_3 again showed high expressions for some of the protein combinations and from our suspicion that I_3 could be possibly skewing the results, the fold change between the sample types was calculated from the original data and after removal of sample I_3. Again, for the 4 protein combinations where I_3 seemed to be driving the difference, all 4 combinations remained statistically significant after the removal of sample I_3 from the dataset (Figure 5.12). The role of I_3 can be investigated further in future experiments to understand whether it is truly an outlier.

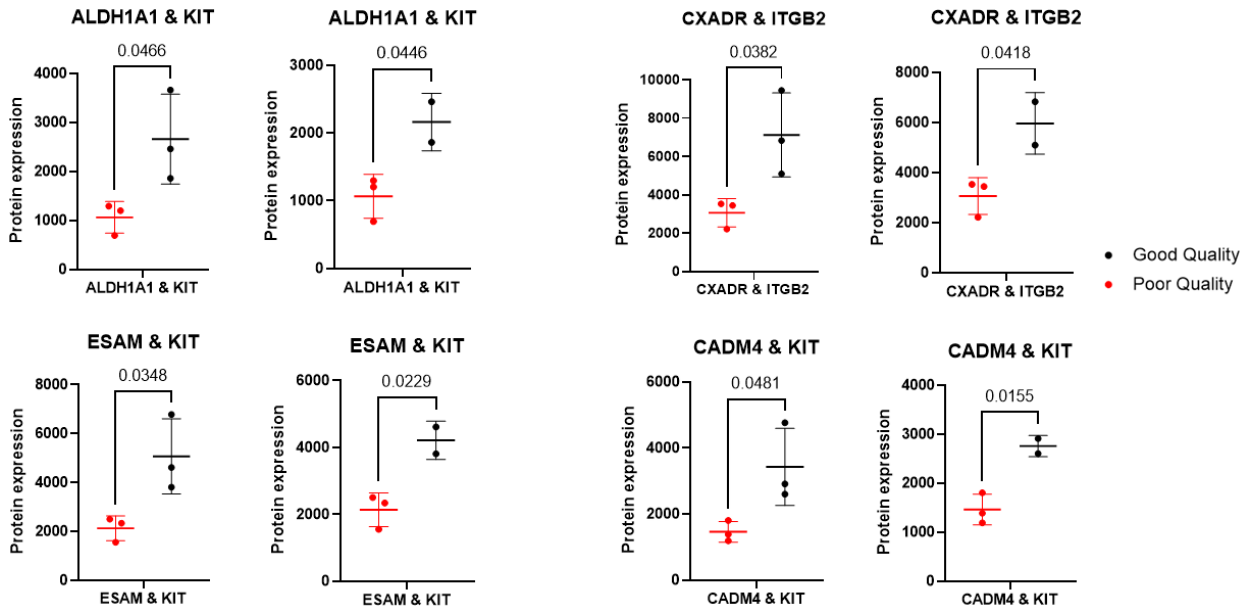

Figure 5.12: The protein expression level of 4 protein combinations which had the highest signal from sample I_3. The left panels represent the results from all samples, the right panels represent the result after the removal of sample I_3. The red represents the poor-quality embryos and the black represent the good-quality embryos. The ns stands for non-significant. The significance is represented by the p-value between the sample types.

Discussion

Characterising S-EVs has been challenging which has led to new avenues with development of new technologies to provide a more comprehensive analysis of the surface proteins as well as cargo of these S-EVs. With EEVs, the main challenge is the small starting volume of samples which limits the type of analysis that can be achieved. With the advances made in identifying multiple proteins at once and reducing the starting volume of samples, the future of EV analysis and their potential role in disease progression and diagnosis is promising. One of these new methods is proximity barcoding assay (PBA) which simultaneously identifies multiple proteins found on the surface of an individual S-EV[176]. We used this method as a pilot study for our sample type to identify proteins (other than the proteins identified in chapter 2) to understand their potential role

in implantation of the embryo. Furthermore, to test whether this technique could be used for more in-depth future analysis using more clinically relevant samples.

Differentially expressed proteins using PBA

Through the sequencing analysis, 7 proteins were identified that were differentially expressed between the good-quality embryos and poor-quality embryos. The three that were upregulated for good-quality embryos were CD20 (membrane spanning 4-domains A1), CDON (cell adhesion associated, oncogene regulated), and GPA33 (glycoprotein A33). The protein that was most significantly upregulated in the good-quality embryos (p=0.0150) was GPA33 which is a cell surface glycoprotein. The molecular functions are not well known but it has been shown to be associated with the immune response. It has been found to be expressed on $CD4^+$ T cell, B cells, dendritic cells and natural killer T cells which are all involved in the initial steps of pregnancy and avoiding the rejection of the semi-allogeneic embryo[177,178]. The next protein was CDON (p=0.0258) which is a co-receptor of Ptch1 involved in the hedgehog (HH) signalling pathway[179]. HH signalling is important for vasculogenesis and angiogenesis in embryonic development; it also is required for placental development and pregnancy[180,181]. Research has shown that the dysregulation of HH signalling can result in recurrent miscarriage by impairing trophoblast invasion and angiogenesis[182]. The last protein which was upregulated in good-quality embryos was CD20 (p=0.040) which is a marker of B-lymphocytes and part of the adaptive immune response, playing a role in B cell activation and proliferation[183,184]. The exact role of CD20 in pregnancy is not well understood but due to the high immune activity during the initial processes of establishing pregnancy, it could be playing an important role and should be investigated further.

The four proteins which were upregulated in the poor-quality samples were PCDH8 (protocadherin 8, p=0.0041), CCR6 (C-C motif chemokine receptor 6, p=0.0068), CEACAM1 (carcinoembryonic

antigen-related cell adhesion molecule 1, p=0.0235). and CD9 (tetraspanin 9, p=0.0492). PCDH8 is a protocadherin involved in cell-cell adhesion and the epithelial-mesenchymal transition which needs to be regulated for implantation[185]. The regulation of PCDH8 is achieved through microRNAs which is seen when miR-429 suppresses the expression of PCDH8 with its upregulation causing a suppression of the invasive and migratory capabilities of endometrial cells in the murine model[186,187]. Therefore, the interplay of PCDH8 and microRNAs are important in implantation and need to be tightly regulated. CCR6 is the receptor for CC-motif chemokine 20 (CCL20) and beta defensins. In the porcine model, CCR6-CCL20 signalling results in an increase in the cellular proliferation of the luminal epithelial cells of the endometrium suggesting that it may improve endometrial receptivity[188]. However, due to its role in the immune and inflammatory response, its upregulation could have negative effects. In one study, there was an increased expression of CCR6 in women suffering from recurrent implantation failure[189]. The increase in CCR6 resulted in an increase in Th17 signalling, thereby increasing the inflammatory signals in the endometrium. Therefore, an increase in CCR6 could have a negative impact on the implantation of the embryo. CEACAM1 is a member of the carcinoembryonic antigen family and is involved in the adhesion of the trophoblast with the maternal endometrium through homophilic and heterophilic binding (with cadherins and integrins)[190]. The presence of CEACAM1 is deemed to be important for implantation, however there is no consensus as to how much CEACAM1 is needed at the maternal-foetal interface and whether the overexpression could impede implantation. CD9 is a tetraspanin which is commonly found on EVs and a previous study suggested that embryos which did not result in a pregnancy had higher levels of EVs compared to embryos that did implant[75]. Therefore, the increased expression of CD9 on poor-quality embryos could reinforce the previous finding.

Due to the potential masking of the results by sample I_3 the sample was removed from the analysis which showed that all the differentially expressed proteins remained significantly different besides CD20. The results therefore, suggest that these proteins could be important for the development of day 3 embryos and their interaction with the endometrium at this point. It would be interesting to test these proteins further and corroborate their expression in other sample types.

Clustering analysis

Another way to characterise the EEVs was to investigate their clustering based on their protein expressions. Again, with the clustering analysis, sample I_3 seemed to contribute the most to the results and mask the clustering from the other samples. Whether I_3 is an outlier, needs to be confirmed through analysis of more samples. Nevertheless, the integrin (ITG) family of proteins seemed to be expressed at high amounts by most of the clusters. Integrins are cell surface proteins which comprise of an alpha chain and a beta chain and different combinations of the chains can result in multiple integrin proteins. The integrin subfamily plays an important role in the implantation of embryos though their roles in cell-cell adhesion and changing the endometrial phenotype[191,192]. The expression of integrins by the trophectoderm as well as the endometrium have been confirmed. The interactions of the integrins on the embryo and the integrins on the endometrium can aid implantation. For example, the reduction of ITGαV on the epithelium was shown in infertile women[193]. There are other proteins (such as fibronectin) expressed by the trophectoderm which can act as ligands for integrins found on the endometrium (such as ITGαVβ3). ITGαVβ3 forms one of the first blastocyst ligands found on the endometrium with its reduction being linked to impairment of cell-cell adhesion[194]. It is also interesting to note that in chapter 2, one of the markers that was consistently found to be highly expressed in the embryonic

samples was CD29 or ITGβ1 which has been linked to implantation. The next step forward would be to investigate this group of proteins further using other samples and endometrial models to investigate their role in implantation.

Another way that the PBA method of identifying proteins expressed on S-EEVs can be utilised is through the analysis of the combination of proteins found together on the same S-EEV. The identification of these combinations could be used in future studies to isolate S-EEVs based on 2 proteins that are simultaneously expressed and delineate their specific roles on endometrial cells. The use of 2 proteins to isolate the EEVs would allow for a more specific isolation. From the analysis of the combination of 2 proteins, the top 25 combinations were all more highly expressed by the good-quality embryos. Even after the removal of sample I_3, which may have been an outlier, the 4 combinations that had the highest difference remained significantly higher in the good-quality group. Looking at the proteins within these combinations, ALDH1A1 (aldehyde dehydrogenase 1 family member A1), KIT (proto-oncogene, receptor tyrosine kinase), CXADR (coxsackievirus and adenovirus receptor), ITGB2 (integrin beta 2), ESAM (endothelial cell adhesion molecule) and CADM4 (cell adhesion molecule 4) made up the most combinations of the significantly different proteins. ALDH1A1 is the enzyme responsible for converting retinol to retinoic acid and was shown to increase during the initial processes of pregnancy resulting in increased retinoic acid and increased progesterone[195,196]. The retinoid pathway plays important roles in endometrial differentiation and blastocyst implantation[197]. ALDH1A1 is also a marker of mesenchymal stem cells and in a previous study it was shown to be expressed at lower levels in non-receptive endometrium, suggesting that there may be a depletion of stem cells in endometrial tissue from implantation failure samples[198]. KIT is a receptor tyrosine kinase which is found to be expressed by the trophoblast layer of the placenta[199]. Stem cell factor is the cytokine ligand which

binds to KIT, resulting in the phosphorylation of intracellular proteins causing proliferation, differentiation, migration and apoptosis[200]. CXADR is involved in maintaining the integrity of tight junctions and is found to be expressed by the trophectoderm of blastocysts. The knockdown of CXADR in the mouse model resulted in reduced implantation[201]. Kai et al., identified 11 differentially expressed genes which were higher in the trophectoderm of blastocysts with a higher expectation of pregnancy (determined by the diameter of the blastocyst and the time taken to reach the blastocyst stage); one of these genes was CXADR[202]. Our finding of increased CXADR on S-EEVs could be an indicator of the integrity of the trophectoderm. ITGB2 was one of the other identified proteins and when it forms heterodimers with alpha chains, it is known to play roles in cellular adhesion, cell signalling and immune responses[203,204]. ESAM and CADM4 both mediate cell-cell adhesion and could be involved in the initial interactions of the trophectoderm with the endometrium although their specific signalling needs to be investigated further. These protein combinations and their expression by S-EEVs need to be further investigated using blastocyst culture media and stricter criteria of which embryos are considered good-quality and poor-quality with respect to pregnancy outcome data.

Limitations and Future work

The main limitation for this chapter was the sample types that were used. Day 3 embryonic samples have less EEVs present compared to day 5 embryos. They are also less clinically relevant in that they may not continue developing until the blastocyst stage. These samples were used to achieve a proof of principle protocol which could then be taken forward with more clinically relevant samples (culture media from day 5 embryos with clinical data such as outcome of pregnancy as well as any other relevant information such as chromosome ploidy results). Another limitation is that the pipeline/antibodies used were developed for identifying different sources of the EVs.

Using day 5 culture media, a new pipeline and antibody panel (to include other markers that could be relevant to implantation as well as looking at the markers identified in chapter 2 (CD142, CD29 and CD146) should be created for the analysis of EEVs specifically to provide more relevant information. The method of PBA proved to be successful in detecting and characterising S-EEVs from embryonic samples which was advantageous due to the small starting volume and limited concentration of EEVs. The technique and analysis could be taken further using more rigorously chosen samples and more specific PBA pipelines. It could be beneficial by simultaneously characterising multiple proteins on the same S-EEV which could be taken forward to understand their role on implantation.

Chapter 6

Final discussion

Implantation

The process of implantation is one of the main bottlenecks in pregnancy in both natural conception and cycles associated with assisted reproductive technologies (ART). Implantation begins when the blastocyst leaves its protective zona pellucida and starts attachment and invasion into the endometrium. The interaction of the embryo and endometrium is enhanced through the release of hormones, cytokine and growth factors. The endometrial epithelial cells also express pinopods which are projections of the epithelial layer that allow for adhesion of the embryo in a specific orientation. Once the embryo attaches to the endometrium, it undergoes invasion into the stromal layer and begins the process of placentation. During this process, there is continuous communication between the embryo and endometrium to ensure proper implantation occurs to avoid adverse outcomes such as miscarriage. One of these sources of communication is via extracellular vesicles.

Extracellular vesicles in implantation

Extracellular vesicles are lipid bound particles which are released by cells for inter and intra-cellular communication. The role of EVs in pregnancy has been under investigation for the past decade and has yielded some interesting insights in terms of implantation. The role of endometrial EVs in receptivity has been demonstrated in a few studies which have identified different transcripts in the EVs from uterine fluid that was collected at different stages of the menstrual cycle[205,206]. More research is needed to understand whether these EVs act in an autocrine fashion to prime the endometrium for implantation and possibly determine the site of implantation. Another mechanism by which EVs have been linked to implantation is through the role of embryonic EVs (EEVs). The culturing of embryos in groups has been shown to enhance the developmental capacity of the embryos suggesting that the secretome which contains EEVs could

be a method of communication between embryos[207]. The role of specific EEV cargo (e.g. miR-378a-3p released by bovine EEVs) has been shown to be involved in the hatching of embryos[208]. A few animal studies have also shown that EEVs could be involved in priming the endometrium for implantation as seen in the study by Su et al., where they showed that EVs from trophectoderm cells of cow embryos, decreased the expression of MUC-1 at the site of implantation, which has anti-adhesive properties, by endometrial cells[209]. These findings suggest that EEVs are involved in embryonic cross-talk as well as potential messengers for priming the endometrium. Taking these findings into consideration, it is plausible that EEVs from human embryos also affect the endometrium pre-implantation. It is therefore important to investigate EEVs from human embryos further for their potential paracrine function in implantation and pregnancy.

Embryonic extracellular vesicles

The presence of EEVs has been confirmed through western blots, transmission electron microscopy (TEM), scanning electron microscopy (SEM) and nanoparticle tracking analysis (NTA) but their protein composition has not been fully characterised due to low availability of samples and low starting volumes[66,73,75]. EEVs could be the first source of communication between us and our mothers and being able to understand the components of this communication will enhance our knowledge of implantation and overall pregnancy. The rationale for the experiment in this thesis was to delineate some aspect of the communication from the embryo to the endometrium through EEVs. The aims of this thesis were to:

1) Isolate and characterize EEVs using NTA, TEM, western blot, and to identify a deeper phenotype of the surface protein markers using MACSPlex, nanoFCM and ExoCounter.

2) The most expressed surface marker (CD142/tissue factor) was selected to study its effects on the proliferation and tyrosine activation of Ishikawa cells.

3) Once the initial response of the cells was identified, the proteomics changes were analysed.
4) In the last chapter, a more global characterisation technique (proximity barcoding assay) was used to identify other surface proteins found on EEVs and determine whether the technique was suitable for the sample type.

Summary of chapters

In chapter 2, I was able to characterise EEVs using conventional methods such as NTA, western blot and TEM. The NTA results confirmed the size distribution of the EVs as being 133.15±14.35nm which is indicative of small EVs. The western blot analysis confirmed the expression of CD63 which is a common marker used to prove the presence of EVs in samples. Once the presence of EEVs was confirmed in the sample, I wanted to characterise the surface markers in more detail. I used an antigen characterisation method and was able to characterise the surface proteins found on EEVs to identify 4 markers that were commonly expressed in all of the tested samples. The 4 identified markers were CD142, CD29, CD146 and CD63. From the identified markers, CD142 was chosen for further investigation because it showed the most significant difference, with a p-value of 0.0192, when compared to control culture media. The presence of CD142 positive EEVs and their higher expression in embryonic samples was confirmed using two additional methods of ExoCounter and NanoFCM. To delineate whether CD142 could be playing a role in the implantation of the embryo, recombinant CD142 protein (rCD142) was added to an endometrial epithelial cellular model. The other 3 markers (CD146, CD29 and CD63) were not investigated further due to limited time and resources. It would be interesting to perform more in-depth analysis of these other markers in a similar fashion as was completed for CD142. Being able to understand the role of multiple proteins which are commonly found on the surface of EEVs, may open new avenues in increasing our knowledge of the

implantation process. Using multiple proteins adds more confidence in understanding the role of the interactions of the embryo with the endometrium. To be able to understand the role of multiple proteins, each protein should be investigated individually, which is the reasoning behind choosing rCD142 as a starting point. The recombinant protein was chosen to reduce the effects of any other components from the EEVs.

When investigating the *in vivo* environment, it is commonly accepted to use cellular models which represent the environment under study. In the uterus, the first point of contact between the embryo and the endometrium is through the endometrial epithelial cells. These cells are hormonally primed to receive the embryo and allow for implantation. There are multiple cellular models which are currently used. For the purposes of this project, I chose Ishikawa cells as they are **a)** representative of endometrial epithelial cells **b)** can be hormonally primed to resemble luminal and glandular cells and **c)** are a widely used model for this tissue. I was able to show that the cells could be hormonally primed (through morphological and protein expression changes). After treatment with progesterone and dibutyryl cAMP for 7 days, the cells expressed dome-like structures similar to pinopod formation *in vivo*. The cells also showed a significantly higher expression of MUC-1 when compared to untreated cells using western blot analysis (1.186±0.058 vs 0.809±0.103, p=0.0464). MUC-1 plays an important role in the attachment of the embryo with the overall expression in the endometrium increasing in response to hormonal treatment and prior to implantation, which is what I observed with our Ishikawa cell model. Once the cells were confirmed to become hormonally primed, I wanted to confirm the uptake of EEVs by the cells to prove that the EEVs had the potential of entering endometrial cells and affecting them prior to implantation. The EEVs were stained with PKH26 and were shown to be up taken by Ishikawa cells using confocal microscopy. The cells were then treated with rCD142, which at 0.1µg/ml

caused the highest change in proliferation with an increase of 3660.626±1639.305 cells compared to the control cells ($p=5.017x10^{-5}$). Prior to proteomics investigations, to ensure that the cells were responding to the rCD142 treatment, global tyrosine proliferation of tyrosine kinases was tested using western blot analysis. The phosphorylation western blot showed an increased phosphorylation of platelet derived growth factor receptor of the treated cells compared to control cells (1.653±0.103 vs 0.941±0.034, p=0.0136). The details of these experiments are outlined in chapter 3.

Once I was able to show that the cells were responding to rCD142 treatment, the treated cells were analysed using proteomics in an attempt to identify proteins and pathways which were differentially expressed in the treated cells compared to controls. The Ishikawa cells were treated with rCD142 or fPBS (control) and a standard proteomics protocol was used. Multiple bioinformatics analyses were tested with the significance A/B test identifying multiple differentially expressed proteins as well as upregulated pathways which could be involved in implantation and pregnancy. Using a fold change cut of value of 1 and FDR of 0.05, there were 76 proteins significantly upregulated and 77 proteins significantly downregulated in the treatment group compared to the control. The proteomics analysis allowed for a more in-depth analysis of the mechanistic pathways which is discussed in full detail in chapter 4.

In an attempt to identify other methods which could be used for characterization of the EEVs, I used proximity barcoding assay which is outlined in chapter 5. With this technique, I was able to show that the EEVs could be characterized based on multiple proteins found on the surface of a single EEV. Through the sequencing analysis, 7 proteins were identified that were differentially expressed between the good-quality embryos and poor-quality embryos. Although, definitive results were not obtained, the experiment was a proof of concept which has the potential of

revealing specific protein-protein interactions on EEVs. Understanding the protein interactions of EEVs as the initial source of communication from the embryo to the endometrium would be beneficial in identifying new ways of embryo selection prior to transfer into the uterus which could improve treatment outcomes in infertility cycles.

Limitations of study

There were some limitations that were identified in the experiments. Some of the experiments would benefit from being repeated using a larger sample size as well as more clinically relevant embryonic samples. The samples used for the initial characterization of the EEVs (chapter 2) had to be pooled due to the nature of the culturing dish which also meant that the outcome of the embryos was not known. In order to better understand the role of the markers in pregnancy, it would be best to repeat the experiments using samples from single culture droplets which would allow for direct comparison of the data to pregnancy outcome. Due to limited time and resources, I was only able to test the effects of rCD142 on the proteome of a single cell model. The next step in unravelling the exact role of CD142 would be to test the effects of rCD142 using a coculture of epithelial cells with stromal cells, primary endometrial cell model, and *in vivo* animal models. Some of the other identified markers could be investigated further as well.

Further investigation of the proteomics results using qRT-PCR would be useful in confirming the proteomics findings. The bioinformatics analysis could also be completed using a more powerful computer which would be able to compute the very small p-values. Since considerable protein changes were seen with the treatment of the cells, transcriptomic analysis would also provide important information for delineating the roles of CD142 in the process of implantation and pregnancy.

For the PBA analysis, more clinically relevant samples which could be linked to outcome data would provide more appropriate results that could be important in diagnosis and treatment options for assisted reproductive technology (ART) cycles. Although obtaining these samples may be challenging, the results outlined in chapter 5 confirm that the PBA technique can characterise the EEVs and warrants the need for collection of such samples. Combining the different characterisation techniques used throughout this thesis can provide an in-depth view of the surface proteins on EEVs as the first form of communication to the endometrium.

It is important to note that while these limitations were present, major findings were identified and explored. I was able to, for the first time, characterise human EEVs using multiple technologies which could be repeated on more clinically relevant EEV samples. I was able to provide an analysis of the mechanism of how CD142 could be involved in the communication between the embryo and endometrium and could be important for implantation. I was also able to show that the new technique of PBA could be used on EEVs and explored further.

Future work

The experiments outlined in this thesis pave the way for more investigations into the potential role of EEVs in implantation and pregnancy. There are potential avenues which could be explored further including clinical applications both in the field of ART and reproductive health in general. Some of these clinical applications include: embryo competency assessment, enhanced embryo transfer medium, and non-hormonal methods of contraception.

EEVs have the potential of being used in conjunction with current methods of embryo selection (such as morphological assessment and preimplantation genetic testing) for identifying the embryo with the highest potential for a healthy pregnancy. The majority of IVF embryos when transferred to the uterus do not produce a pregnancy. Preimplantation genetic testing (PGT) and

morphological assessments are associated with success but do not necessarily guarantee a viable pregnancy. This can be seen in cases where a euploid embryo of high morphological grade results in failure to achieve a pregnancy[2μ]. There is a need for more comprehensive selection criteria for identifying which embryo to transfer to increase chances of a healthy pregnancy. It is likely that EEVs may be involved in modulating the endometrium to communicate the presence of the embryo and prime for implantation as is suggested by the data presented in my thesis. It is therefore, possible that EEVs could be used as a marker of embryo potential in conjunction with current selection methods. One way of using the EEVs to distinguish and select between embryos is to identify which markers are commonly expressed by embryos with a good outcome (healthy pregnancy) compared to those with a poor outcome (implantation failure). Once these markers are identified, they can be used in embryo selection using equipment such as the ExoCounter. CD142 has the potential of being one such marker, however further testing is required to confirm this. One approach would be to collect culture media from single embryos cultured to the blastocyst stage and measure the amount of CD142 positive EEVs using the ExoCounter. The samples would then be divided into 2 groups: culture media from those embryos which resulted in an ongoing pregnancy (>12 weeks) and those which were transferred but did not result in a pregnancy or were associated with early pregnancy loss. It would be beneficial to include PGT results in the grouping of the embryos. The amount of CD142-positive EEVs could then be determined using the ExoCounter which would delineate whether there is a difference between embryos deemed to be 'good quality' and resulted in a pregnancy and embryos with similar morphological characteristics which did not result in a pregnancy or were associated with early pregnancy loss. If a difference is identified, then measuring the amount of CD142-positive EEVs could be used in conjunction to other selection methods to decide which embryo has a better chance of implanting and should be

transferred to the uterus. Other experiments could also be carried out to validate this as well as other markers that are commonly found on the surface of EEVs.

While being able to choose the embryo with the highest potential for implantation is valuable, it does not alter embryonic or maternal physiology. Another possible application of using EEVs in the field of reproductive health is through priming the endometrium. I was able to show that rCD142 increased the rate of proliferation and activated specific protein pathways which could be important for implantation. To determine whether this could be applied clinically, the same set of experiments will need to be completed using primary endometrial epithelial cells. If similar results are found, the toxicity of the recombinant protein would need to be tested on the embryos as well as the endometrium. If the protein is deemed suitable and non-toxic, a specified dose could be added to transfer culture media to improve the chances of implantation. The addition of the recombinant protein could also be applied in a topical treatment in natural cycles to help aid implantation. Being able to support the process of implantation through a modified transfer media may be helpful in improving the chances of pregnancy per transfer and per cycle of treatment. Based on the Ishikawa cell study in this thesis, it is reasonable to assume that CD142 can enhance cellular responses in primary endometrial cells and therefore could enhance embryo implantation *in vivo*. The timing of exposing the endometrium to the protein will need to be determined, however, the time course studies from the Ishikawa cell model confirm that the cells respond within 24 hours of treatment suggesting that the rCD142 could be introduced simultaneously as the embryo during the transfer process. It is therefore, possible that once the action of rCD142 is fully understood, it could be added to the embryo transfer medium to improve the endometrium receptivity to the embryo.

Another interesting avenue which could be explored further is the use of blocking specific proteins and their subsequent signaling pathways as a means of contraception. Current methods of contraception for women are mostly hormonal which although reliable, carry a myriad of symptoms and are becoming less popular in younger generations[211]. If the role of the identified proteins such as CD142 are studied further, there is a possibility of being able to block these on endometrial cells to prevent an embryo from implanting. The applicability of this needs to be studied. It is important to note that there are commercially available CD142 inhibitors which could be researched further in terms of contraception.

All of the possible applications mentioned above could be investigated in terms of any of the other surface proteins which were identified through MACSPlex analysis and PBA. The use of multiple proteins could enhance the effectiveness of the treatments, once the initial experiments have been completed. Although it is clear that many different pathways are involved in the process of implantation and it is not solely reliant on EEVs, being able to identify the exact role of CD142 and other proteins expressed on EEVs could be beneficial in understanding some aspects of the process of implantation and improving treatment options for those struggling with infertility as well as other areas of reproductive health. The work outlined in my thesis could pave the way for investigating some of these possible routes further and providing better options for patients.

Appendix

Table S4.2: The proteins identified as differentially expressed between the treated cells and control cells using the 'MinProb' imputation method of the missing not at random technique. The p-value, and adjusted p-value are shown. The calculations were completed by the DEP package on R. The highlighted rows outline the proteins where the adjusted p-value is smaller than the p-value.

Protein name	Trt_vs_Ctrl_p.val	Trt_vs_Ctrl_p.adj	Significant	Trt_vs_Ctrl_ratio
SUMO2	0.001604	0.00415	TRUE	3.35
RNF114	2.75E-05	2.33E-13	TRUE	2.97
PLPP3	1.11E-07	2.33E-13	TRUE	2.73
TRIM2.1	1.72E-06	2.33E-13	TRUE	1.71
TRIM32	0.001762	0.0055	TRUE	1.57
RBM15.1	0.00013	6.55E-09	TRUE	1.53
KXD1	0.003504	0.0297	TRUE	1.15
WIPF2	0.002959	0.0214	TRUE	1.11
WASF1	0.003802	0.0357	TRUE	1.07
CYP2U1	0.001651	0.00453	TRUE	1.05
SRSF2	0.000577	7.78E-05	TRUE	1.03
ASCC1	0.000656	0.000136	TRUE	1.01
PRKAB2	0.000428	1.76E-05	TRUE	1.01
KLHL13	0.004318	0.0471	TRUE	1
AGAP1.1	0.001914	0.00703	TRUE	0.938
RNF149	0.000466	2.71E-05	TRUE	0.923
B4GALT1	0.00015	2.07E-08	TRUE	0.884
SH3KBP1	0.00032	3.62E-06	TRUE	0.883
CMC1	0.000355	6.42E-06	TRUE	0.875
RPUSD4	0.003071	0.0231	TRUE	0.722
CCDC174	0.003003	0.022	TRUE	0.714
CALCOCO2.1	7.05E-05	2.16E-11	TRUE	0.679
RNF113A	3.17E-05	2.33E-13	TRUE	0.65
TJAP1	0.000113	2.30E-09	TRUE	0.648
BOLA1	0.000609	9.82E-05	TRUE	0.646
HRAS	0.000732	0.000226	TRUE	0.601
GRPEL2	0.001156	0.00142	TRUE	0.587
MPV17L2	0.003283	0.0263	TRUE	-0.664
LAMTOR2	0.000388	1.04E-05	TRUE	-0.686
CALCOCO2	0.000326	4.02E-06	TRUE	-0.702
STEAP3	0.002232	0.0108	TRUE	-0.703
MAP3K4	0.002901	0.0205	TRUE	-0.717
GCC1	0.002731	0.0179	TRUE	-0.762
MANEA	0.000255	9.09E-07	TRUE	-0.768
DPP8	0.0018	0.00584	TRUE	-0.829
SLC27A1	4.83E-05	4.00E-13	TRUE	-0.989

POLE2	0.001475	0.00316	TRUE	-1.05
TPM1.5	0.000591	8.64E-05	TRUE	-1.05
MAP1B	0.003903	0.0378	TRUE	-1.1
MAOB	0.000158	2.93E-08	TRUE	-1.13
STK25	0.000239	6.06E-07	TRUE	-1.15
TWNK	0.000495	3.68E-05	TRUE	-1.15
NCKAP5L	0.001	0.000847	TRUE	-1.18
MTA1.2	0.001616	0.00425	TRUE	-1.19
B3GALT6	0.000161	3.25E-08	TRUE	-1.24
POLR3D	5.04E-06	2.33E-13	TRUE	-1.26
CCNC	0.000123	4.43E-09	TRUE	-1.45
TIMM9	0.00097	0.000752	TRUE	-1.55
SLC7A5.1	0.000591	8.64E-05	TRUE	-1.63
MMACHC	5.25E-05	8.75E-13	TRUE	-1.94
LRRK1	3.76E-09	2.33E-13	TRUE	-3.22

Table S4.3: The proteins identified as differentially expressed between the treated cells and control cells using the 'QRILC' imputation method of the missing not at random technique. The p-value, and adjusted p-value are shown. The calculations were completed by the DEP package on R. The highlighted rows outline the proteins where the adjusted p-value is smaller than the p-value.

Protein name	Trt_vs_Ctrl_p.val	Trt_vs_Ctrl_p.adj	Significant	Trt_vs_Ctrl_ratio
TRIM2.1	1.81E-06	1.77E-13	TRUE	6.4
RNF114	3.07E-06	1.77E-13	TRUE	5.36
PLPP3	3.82E-07	1.77E-13	TRUE	5.28
PRKAB2	0.000539	7.78E-05	TRUE	4.18
CNPY4	0.002395	0.0174	TRUE	3.7
RBM15.1	2.26E-06	1.77E-13	TRUE	3.61
SUMO2	0.001908	0.00997	TRUE	3.35
SURF6	0.001278	0.00298	TRUE	3.05
SYK	0.003234	0.0339	TRUE	2.89
NDUFAF1	6.83E-05	1.27E-11	TRUE	2.55
POLR2G	3.59E-05	1.77E-13	TRUE	2.48
MINK1	0.001567	0.00545	TRUE	2.25
PPP2R1B	0.000262	1.35E-06	TRUE	1.71
SLC35B3	0.00144	0.0042	TRUE	1.61
WIPF2	0.003428	0.038	TRUE	1.11
CYP2U1	0.001937	0.0104	TRUE	1.05
SRSF2	0.000692	0.000254	TRUE	1.03
AGAP1.1	0.00223	0.0149	TRUE	0.938
RNF149	0.000557	9.10E-05	TRUE	0.923
B4GALT1	0.000182	1.00E-07	TRUE	0.884
CMC1	0.000424	2.33E-05	TRUE	0.875

RPUSD4	0.003508	0.0396	TRUE	0.722
CCDC174	0.003429	0.038	TRUE	0.714
TJAP1	0.000129	6.66E-09	TRUE	0.648
HRAS	0.000838	0.000595	TRUE	0.601
GRPEL2	0.001316	0.00326	TRUE	0.587
MPV17L2	0.00373	0.0446	TRUE	-0.664
LAMTOR2	0.000452	3.21E-05	TRUE	-0.686
CALCOCO2	0.000381	1.29E-05	TRUE	-0.702
STEAP3	0.002561	0.0199	TRUE	-0.703
MAP3K4	0.003316	0.0357	TRUE	-0.717
GCC1	0.003134	0.0318	TRUE	-0.762
MANEA	0.000301	3.29E-06	TRUE	-0.768
DPP8	0.00209	0.0127	TRUE	-0.829
POLE2	0.001735	0.00754	TRUE	-1.05
TWNK	0.000598	0.000128	TRUE	-1.15
MTA1.2	0.001903	0.0099	TRUE	-1.19
B3GALT6	0.000199	1.97E-07	TRUE	-1.24
CCNC	0.000155	3.11E-08	TRUE	-1.45
SHKBP1	0.002745	0.0237	TRUE	-1.85
FASTKD1	0.000168	5.67E-08	TRUE	-1.88
SULT1A3	0.000457	3.36E-05	TRUE	-2.04
INPPL1	2.66E-06	1.77E-13	TRUE	-2.51
SLC27A1	8.18E-05	8.61E-11	TRUE	-3.76
POM121	0.002393	0.0174	TRUE	-3.86
SLC7A5.1	3.72E-05	1.77E-13	TRUE	-4.42
LRRK1	1.34E-05	1.77E-13	TRUE	-5.94

Table S4.5: The proteins identified as differentially expressed between the treated cells and control cells using the 'MLE' imputation method of the missing at random technique. The p-value, and adjusted p-value are shown. The calculations were completed by the DEP package on R. The highlighted rows outline the proteins where the adjusted p-value is smaller than the p-value.

Protein name	Trt_vs_Ctrl_p.val	Trt_vs_Ctrl_p.adj	Significant	Trt_vs_Ctrl_ratio
SUMO2	0.001443	0.00346	TRUE	3.35
WIPF2	0.002705	0.0186	TRUE	1.11
WASF1	0.00349	0.0307	TRUE	1.07
CYP2U1	0.001497	0.00385	TRUE	1.05
ACBD7	0.0011	0.00155	TRUE	1.04
SRSF2	0.000517	6.59E-05	TRUE	1.03
KLHL13	0.003974	0.0406	TRUE	1
AGAP1.1	0.001743	0.00603	TRUE	0.938
ASCC1	0.000376	1.38E-05	TRUE	0.927
RNF149	0.000418	2.26E-05	TRUE	0.923

B4GALT1	0.000134	2.22E-08	TRUE	0.884
CMC1	0.000318	5.78E-06	TRUE	0.875
RSL24D1	0.001688	0.00548	TRUE	0.874
RPUSD4	0.002831	0.0205	TRUE	0.722
CCDC174	0.002768	0.0196	TRUE	0.714
TJAP1	0.000104	4.59E-09	TRUE	0.648
HRAS	0.000674	0.00022	TRUE	0.601
GRPEL2	0.001067	0.00141	TRUE	0.587
MPV17L2	0.003036	0.0234	TRUE	-0.664
LAMTOR2	0.000353	1.00E-05	TRUE	-0.686
ZZZ3	0.000259	2.02E-06	TRUE	-0.687
MAOB	0.004011	0.0414	TRUE	-0.701
CALCOCO2	0.000296	4.11E-06	TRUE	-0.702
STEAP3	0.002052	0.00958	TRUE	-0.703
MAP3K4	0.002673	0.0182	TRUE	-0.717
GCC1	0.00251	0.0159	TRUE	-0.762
MANEA	0.000229	9.91E-07	TRUE	-0.768
DPP8	0.001643	0.00508	TRUE	-0.829
POLR3D	0.000597	0.000125	TRUE	-0.897
BRD2	3.49E-05	1.39E-12	TRUE	-0.961
POLE2	0.001336	0.00273	TRUE	-1.05
KNSTRN	0.00058	0.000111	TRUE	-1.08
TWNK	0.000441	2.93E-05	TRUE	-1.15
MTA1.2	0.001463	0.0036	TRUE	-1.19
B3GALT6	0.000141	2.90E-08	TRUE	-1.24
CCNC	0.000107	5.15E-09	TRUE	-1.45

Table S4.7: The proteins identified as differentially expressed between the treated cells and control cells using the 'MinProb' imputation method of the missing not at random technique. The p-value, and adjusted p-value using the Benjamini-Hochberg method are shown. The calculations were completed by the DEP package on R.

Protein name	Trt_vs_Ctrl_p.val	Trt_vs_Ctrl_p.adj (new)	Significant	Trt_vs_Ctrl_ratio
PLPP3	6.19E-08	0.000304	TRUE	2.67
LRRK1	1.31E-07	0.000323	TRUE	-3.27
POLR3D	9.25E-07	0.001416	TRUE	-1.23
TRIM2.1	1.44E-06	0.00177	TRUE	1.75
RNF114	2.14E-05	0.019254	TRUE	3.19
SLC27A1	4.05E-05	0.028948	TRUE	-1.17
MMACHC	8.27E-05	0.040658	TRUE	-1.97
CALCOCO2.1	9.57E-05	0.042943	TRUE	0.664
TJAP1	0.000113	0.04541	TRUE	0.648
RBM15.1	0.000119	0.046139	TRUE	1.68

CD47	0.000119	0.046185	FALSE	-0.449
RNF113A	0.000123	0.046616	TRUE	0.607
CCNC	0.000123	0.04665	TRUE	-1.45

Table S4.8: The proteins identified as differentially expressed between the treated cells and control cells using the 'QRILC' imputation method of the missing not at random technique. The p-value, and adjusted p-value using the Benjamini-Hochberg method are shown. The calculations were completed by the DEP package on R.

Protein name	Trt_vs_Ctrl_p.val	Trt_vs_Ctrl_p.adj (new)	Significant	Trt_vs_Ctrl_ratio
INPPL1	6.85E-07	0.001335	TRUE	-3.31
LRRK1	1.11E-06	0.001335	TRUE	-6.11
PPP2R1B	6.12E-07	0.001335	TRUE	1.98
TRIM2.1	9.68E-07	0.001335	TRUE	4.12
SLC27A1	3.15E-06	0.003038	TRUE	-3.71
RBM15.1	7.71E-06	0.005506	TRUE	3.64
PLPP3	8.38E-06	0.005765	TRUE	5.12
POM121	2.61E-05	0.015214	TRUE	-3.38
SLC7A5.1	3.63E-05	0.019454	TRUE	-4.01
FASTKD1	9.64E-05	0.039266	TRUE	-2.8
CNPY4	0.000118	0.044275	TRUE	4.87
CD47	0.000119	0.044489	FALSE	-0.449
TJAP1	0.000129	0.046585	TRUE	0.648

Table S4.9: The proteins identified as differentially expressed between the treated cells and control cells using the significance A/B method of calculation.

Protein name	Fold change (log_2)	Adjusted P value	P value
HSP90AB2P	6.311396	0	0
LRRK1	-4.25654	4.8E-171	1.5E-174
RNF114	3.792446	3.3E-133	1.5E-136
PLPP3	3.415633	4.5E-130	2.7E-133
HNRNPA1.2	3.323755	1.1E-115	8.2E-119
SUMO2	3.348331	4E-101	3.7E-104
C10orf67	-2.802	1.9E-86	2.05E-89
TRIM2.1	2.737181	1.87E-83	2.3E-86
BRD2	-2.62078	4.55E-80	6.29E-83
KRT8.2	2.86791	3.35E-76	5.14E-79
RBM15.1	2.505318	1.12E-75	1.9E-78
CNPY4	2.435993	4.05E-73	7.48E-76
YWHAB.4	2.367862	1.22E-67	2.43E-70
PDE7B	-2.662	8.66E-67	1.86E-69
SLC7A5.1	-2.39617	8.97E-54	2.07E-56

SLC27A1	-2.19284	6.9E-53	1.7E-55
KRT2.3	-2.05013	2.27E-51	5.93E-54
MAOB	-2.01737	9.46E-50	2.62E-52
KRT10	-2.05382	3.2E-49	9.35E-52
SEPTIN2.2	2.319033	6.15E-48	1.89E-50
PPP2R5C.2	-2.44208	8.71E-46	2.81E-48
LRR1	2.080117	6.89E-45	2.33E-47
PRKAB2	1.885368	5.09E-43	1.82E-45
KRT10.2	-2.0834	5.09E-43	1.88E-45
KRT10.3	2.062441	4E-42	1.54E-44
SYK	2.002504	1.21E-41	4.83E-44
PPM1E	1.971699	2.23E-40	9.24E-43
OR6N2	-1.72315	4.61E-38	1.98E-40
ST6GALNAC2	-1.75789	7.47E-38	3.33E-40
SUMO3	2.053188	1.18E-37	5.45E-40
POLR3D	-2.26342	9.84E-37	4.69E-39
IFT22	1.982441	1.63E-36	7.99E-39
ACBD7	1.857969	2.97E-34	1.5E-36
DYNLL1	1.748961	5.2E-34	2.72E-36
KRT1	-1.69648	1.07E-33	5.78E-36
POLR1D	1.923251	1.16E-33	6.41E-36
KRT13.3	-1.72216	1.09E-32	6.17E-35
KRT1.2	-1.65541	4.17E-32	2.43E-34
NCKAP5L	-1.56928	1.17E-31	7E-34
MFAP3	1.810494	1.7E-30	1.04E-32
PDSS1	1.801989	1.35E-29	8.51E-32
KRT1.1	-1.68894	2.19E-28	1.41E-30
HNF4A	1.773361	3.68E-28	2.44E-30
ASCC1	1.633414	8.28E-28	5.6E-30
RSL24D1	1.497011	2.51E-27	1.73E-29
AZGP1	-1.50654	2.81E-26	1.99E-28
GLRX5	-1.65903	5.7E-26	4.12E-28
ARL6IP1	-1.88985	1.13E-25	8.35E-28
KRT2	-1.81779	1.62E-25	1.22E-27
ZDHHC3	-1.51504	2.49E-25	1.91E-27
TIMM9	-1.51181	3.13E-25	2.46E-27
PHF23	1.619853	1.81E-24	1.45E-26
STK25	-1.40717	2.25E-24	1.84E-26
DGCR2	1.453866	1.05E-23	8.71E-26
KRT2.2	-1.80764	1.59E-23	1.35E-25
GPR107	-1.56818	3.04E-23	2.62E-25
TMA7	1.490641	3.06E-23	2.71E-25
ZBTB7A	-1.52486	3.06E-23	2.72E-25

TIMM10B	1.590927	3.32E-23	3.01E-25
FXN	1.521695	3.7E-23	3.41E-25
GFAP.1	1.426469	7.25E-23	6.8E-25
KRT5	-1.39099	8.4E-23	8E-25
PITPNA.1	1.558584	1.01E-22	9.79E-25
CALCOCO2.1	1.451229	1.81E-22	1.78E-24
PPT2	-1.37844	2.07E-22	2.06E-24
EMC9	1.334646	3.23E-22	3.28E-24
KRT9	-1.36745	4.56E-22	4.7E-24
RPE	1.478917	6.29E-22	6.58E-24
DCAF16	1.467082	1.39E-21	1.47E-23
ZNF428	1.381286	1.72E-21	1.85E-23
KRT78	-1.44647	5.43E-21	5.93E-23
KRT10.1	-1.67987	2.11E-20	2.33E-22
RECQL4	-1.31952	2.89E-20	3.24E-22
TUBB.7	1.339045	3.25E-20	3.7E-22
KBTBD11	1.494765	3.55E-20	4.09E-22
GNG5	1.270092	3.55E-20	4.15E-22
MYO10	-1.4484	5.63E-20	6.66E-22
ZFC3H1	-1.29889	5.73E-20	6.87E-22
LDHC	1.444674	2.92E-19	3.54E-21
GPSM1	1.232975	4.91E-19	6.03E-21
SLC50A1	1.286393	1.32E-18	1.64E-20
COA3	1.327405	1.49E-18	1.88E-20
ATG14	-1.19308	1.62E-18	2.06E-20
KXD1	1.430135	1.63E-18	2.1E-20
H2AC18	-1.18758	2.34E-18	3.05E-20
SLC35A2	-1.3401	4.28E-18	5.65E-20
TRIM32	1.256773	8.67E-18	1.16E-19
MTA1.2	-1.19234	1.17E-17	1.58E-19
NDUFAF1	1.35777	2.34E-17	3.2E-19
KRT14	-1.30951	2.67E-17	3.7E-19
FLG2	-1.48776	2.74E-17	3.83E-19
TWNK	-1.15026	2.94E-17	4.16E-19
B3GALT6	-1.24347	3.23E-17	4.62E-19
ZG16B	-1.30391	3.59E-17	5.19E-19
MRPS16	-1.1437	4.48E-17	6.54E-19
ZFYVE26	1.370229	4.72E-17	6.97E-19
APTX	1.366323	5.84E-17	8.71E-19
BCAP29	-1.17237	4.46E-16	6.72E-18
PDS5A.1	1.191808	4.75E-16	7.24E-18
ELOVL7	-1.29657	4.82E-16	7.4E-18
WEE1	-1.43029	5.04E-16	7.83E-18

GRIP1	1.220613	8.88E-16	1.39E-17
RSPRY1	-1.41765	9.45E-16	1.5E-17
KNSTRN	-1.46917	9.88E-16	1.58E-17
SPRY1	1.170899	1.62E-15	2.61E-17
NCOR1.1	-1.23722	1.62E-15	2.65E-17
ZBTB24	-1.13202	2.72E-15	4.47E-17
TESK2	1.279218	2.98E-15	4.94E-17
KNL1	1.157556	3.02E-15	5.08E-17
CCNC	-1.44517	3.02E-15	5.11E-17
PTH1R	-1.07642	3.38E-15	5.76E-17
HECA	-1.09716	4.73E-15	8.15E-17
TPM1.5	-1.24386	6.82E-15	1.18E-16
MBP	1.162066	1.18E-14	2.06E-16
FGG	-1.05502	1.28E-14	2.27E-16
POLE2	-1.04711	2.1E-14	3.74E-16
WASF1	1.065642	2.24E-14	4.02E-16
SLC25A21	1.118272	2.91E-14	5.27E-16
KRT6A	-1.34534	2.99E-14	5.47E-16
LSM14A.1	-1.05428	6.25E-14	1.15E-15
SIGIRR	-1.08172	9.28E-14	1.73E-15
LMBRD1	-1.03799	1.64E-13	3.07E-15
TLCD1	1.086717	1.71E-13	3.26E-15
NFKBIE	1.031195	1.71E-13	3.25E-15
RAB33B	-1.30742	1.75E-13	3.35E-15
MMACHC	-1.09593	1.78E-13	3.45E-15
SLC39A9	1.142676	2.43E-13	4.74E-15
OTUD7B	-1.04236	4.75E-13	9.34E-15
STAU1.1	1.059209	8.66E-13	1.72E-14
PTBP1.1	-1.04105	9.03E-13	1.8E-14
CAV1	1.164377	1.9E-12	3.83E-14
CFL1.2	1.161827	2.14E-12	4.34E-14
SLC25A32	1.036408	2.62E-12	5.36E-14
ZNF3	-1.2917	2.9E-12	5.98E-14
GAB1	-1.27709	5.43E-12	1.13E-13
C2orf49	1.040053	7.54E-12	1.58E-13
CLIC3	-1.01901	1.1E-11	2.33E-13
KLC3.1	1.029277	1.29E-11	2.74E-13
WIPF2	1.113769	2.01E-11	4.3E-13
RAB36	1.052683	2.19E-11	4.71E-13
HMMR.1	-1.08458	2.42E-11	5.25E-13
LYRM4	1.008521	6.41E-11	1.41E-12
NKIRAS2	1.066674	9.02E-11	2.01E-12
PARS2	-1.14588	1.89E-10	4.33E-12

CYP2U1	1.050054	3.39E-10	7.82E-12
MDH1B	1.026275	4.11E-10	9.59E-12
SRSF2	1.028782	4.64E-10	1.1E-11
YPEL5	1.015124	6.59E-10	1.58E-11
KLHL13	1.004946	1.26E-09	3.11E-11
TAF12	-1.0963	1.31E-09	3.25E-11
STAT6	-1.11048	3.69E-09	9.31E-11
UBE2B	-1.07314	1.4E-08	3.71E-10
CBX3.1	-1.01972	8.56E-08	2.47E-09

References

1. ICD-11 for Mortality and Morbidity Statistics. 2022; Available from: https://icd.who.int/browse11/l-m/en#/http://id.who.int/icd/entity/1237004558

2. Homan GF, Davies M, Norman R. The impact of lifestyle factors on reproductive performance in the general population and those undergoing infertility treatment: a review. *Hum. Reprod. Update*. 2007;13:209–223.

3. Petraglia F, Serour GI, Chapron C. The changing prevalence of infertility. *Int. J. Gynecol. Obstet.* 2013;123:S4–S8.

4. Baird DT, Collins J, Egozcue J, Evers LH, Gianaroli L, Leridon H, Sunde A, Templeton A, Van Steirteghem A, Cohen J, et al. Fertility and ageing. *Hum. Reprod. Update*. 2005;11:261–276.

5. Chambers GM, Adamson GD, Eijkemans MJC. Acceptable cost for the patient and society. *Fertil. Steril.* 2013;100:319–327.

6. Fauser BC. Towards the global coverage of a unified registry of IVF outcomes. *Reprod. Biomed. Online*. 2019;38:133–137.

7. National Summary Report; Available from: https://www.sartcorsonline.com/rptCSR_PublicMultYear.aspx?reportingYear=2018

8. Treatment add-ons | Human Fertilisation and Embryology Authority; Available from: https://www.hfea.gov.uk/treatments/treatment-add-ons/

9. Yung SSF, Lai SF, Lam MT, Lui EMW, Ko JKY, Li HWR, Wong JYY, Lau EYL, Yeung WSB, Ng EHY. Hyaluronic acid-enriched transfer medium for frozen embryo transfer: a randomized, double-blind, controlled trial. *Fertil. Steril*. 2021;116:1001–1009.

10. Hoogenhuijze NE van, Kasius JC, Broekmans FJM, Bosteels J, Torrance HL. Endometrial scratching prior to IVF; does it help and for whom? A systematic review and meta-analysis. *Hum. Reprod. Open*. 2019;2019:1–18.

11. Metwally M, Chatters R, Dimairo M, Walters S, Pye C, White D, Bhide P, Chater T, Cheong Y, Choudhary M, et al. A randomised controlled trial to assess the clinical effectiveness and

safety of the endometrial scratch procedure prior to first-time IVF, with or without ICSI. *Hum. Reprod*. 2021;36:1841–1853.

12. Lundin K, Park H. Time-lapse technology for embryo culture and selection. *Ups. J. Med. Sci*. 2020;125:77.

13. Simón C, Gómez C, Cabanillas S, Vladimirov I, Castillón G, Giles J, Boynukalin K, Findikli N, Bahçeci M, Ortega I, et al. A 5-year multicentre randomized controlled trial comparing personalized, frozen and fresh blastocyst transfer in IVF. *Reprod. Biomed. Online*. 2020;41.

14. Riestenberg C, Kroener L, Quinn M, Ching K, Ambartsumyan G. Routine endometrial receptivity array in first embryo transfer cycles does not improve live birth rate. *Fertil. Steril.* 2021;115:1001–1006.

15. Sfontouris IA, Nastri CO, Lima MLS, Tahmasbpourmarzouni E, Raine-Fenning N, Martins WP. Artificial oocyte activation to improve reproductive outcomes in women with previous fertilization failure: a systematic review and meta-analysis of RCTs. *Hum. Reprod*. 2015;30:1831–1841.

16. Ebner T, Montag M. Artificial oocyte activation: evidence for clinical readiness. *Reprod. Biomed. Online.* 2016;32:271–273.

17. Victoria Askew. Background to the novel process application for Anecova AneVivo Intrauterine device. 2020; Available from: https://www.hfea.gov.uk/media/k0wk3dih/2020-06-08-scaac-background-to-the-novel-process-application-for-anecova-anevivo-intrauterine-device.pdf

18. Souza Setti A, Ferreira RC, Paes De Almeida Ferreira Braga D, De Cássia Sávio Figueira R, Iaconelli A, Borges E. Intracytoplasmic sperm injection outcome versus intracytoplasmic morphologically selected sperm injection outcome: a meta-analysis. *Reprod. Biomed. Online*. 2010;21:450–455.

19. Leandri RD, Gachet A, Pfeffer J, Celebi C, Rives N, Carre-Pigeon F, Kulski O, Mitchell V, Parinaud J. Is intracytoplasmic morphologically selected sperm injection (IMSI) beneficial in the first ART cycle? A multicentric randomized controlled trial. *Andrology*. 2013;1:692–697.

20. Klement AH, Koren-Morag N, Itsykson P, Berkovitz A. Intracytoplasmic morphologically selected sperm injection versus intracytoplasmic sperm injection: a step toward a clinical algorithm. *Fertil. Steril.* 2013;99:1290–1293.

21. Collins J. Global epidemiology of multiple birth. *Reprod. Biomed. Online*. 2007;15 Suppl 3:45–52.

22. Stock S, Norman J. Preterm and term labour in multiple pregnancies. *Semin. Fetal Neonatal Med*. 2010;15:336–341.

23. Wei J, Wu QJ, Zhang TN, Shen ZQ, Liu H, Zheng DM, Cui H, Liu CX, He J, Li X, et al. Complications in multiple gestation pregnancy: A cross-sectional study of ten maternal-fetal medicine centers in China. *Oncotarget*. 2016;7:30797–30803.

24. Sibai BM, Hauth J, Caritis S, Lindheimer MD, MacPherson C, Klebanoff M, VanDorsten JP, Landon M, Miodovnik M, Paul R, et al. Hypertensive disorders in twin versus singleton gestations. *American Journal of Obstetrics and Gynecology*. 2000. p. 938–942.

25. Martikainen H, Tiitinen A, Tomás C, Tapanainen J, Orava M, Tuomivaara L, Vilska S, Hydén-Granskog C, Hovatta O, Morin-Papunen L, et al. One versus two embryo transfer after IVF and ICSI: a randomized study. *Hum. Reprod*. 2001;16:1900–1903.

26. Lee AM, Connell MT, Csokmay JM, Styer AK. Elective single embryo transfer- the power of one. *Contracept. Reprod. Med*. 2016;1.

27. Neri Q V., Lee B, Rosenwaks Z, Machaca K, Palermo GD. Understanding fertilization through intracytoplasmic sperm injection (ICSI). *Cell Calcium*. 2014;55:24.

28. Wang H, Zhang S, Lin H, Kong S, Wang S, Wang H, Armant DR. Physiological and molecular determinants of embryo implantation. *Mol. Aspects Med*. 2013;34:939.

29. Hatırnaz Ş, Pektaş MK. Day 3 embryo transfer versus day 5 blastocyst transfers: A prospective randomized controlled trial. *Turkish J. Obstet. Gynecol*. 2017;14:82.

30. Mumusoglu S, Polat M, Ozbek IY, Bozdag G, Papanikolaou EG, Esteves SC, Humaidan P, Yarali H. Preparation of the Endometrium for Frozen Embryo Transfer: A Systematic Review. *Front. Endocrinol. (Lausanne)*. 2021;12.

31. Idelevich A, Vilella F. Mother and Embryo Cross-Communication. *Genes (Basel)*. 2020;11.

32. Mor G, Aldo P, Alvero AB. The unique immunological and microbial aspects of pregnancy. *Nat. Rev. Immunol.* 2017;17:469–482.

33. Navot D, Scott RT, Droesch K, Veeck LL, Liu HC, Rosenwaks Z. The window of embryo transfer and the efficiency of human conception in vitro. *Fertil. Steril.* 1991;55:114–118.

34. Enciso M, Aizpurua J, Rodríguez-Estrada B, Jurado I, Ferrández-Rives M, Rodríguez E, Pérez-Larrea E, Climent AB, Marron K, Sarasa J. The precise determination of the window of implantation significantly improves ART outcomes. *Sci. Reports*. 2021;11:1–8.

35. Gardner DK, Balaban B. Assessment of human embryo development using morphological criteria in an era of time-lapse, algorithms and "OMICS": Is looking good still important? *Mol. Hum. Reprod*. 2016;22:704–718.

36. Sjöblom P, Menezes J, Cummins L, Mathiyalagan B, Costello MF. Prediction of embryo developmental potential and pregnancy based on early stage morphological characteristics. *Fertil. Steril*. 2006;86:848–861.

37. Nasiri N, Eftekhari-Yazdi P. An Overview of The Available Methods for Morphological Scoring of Pre-Implantation Embryos in In VitroFertilization. *Cell J*. 2015;16:392.

38. technology EW group on T, Apter S, Ebner T, Freour T, Guns Y, Kovacic B, Le Clef N, Marques M, Meseguer M, Montjean D, et al. Good practice recommendations for the use of time-lapse technology. *Hum. Reprod. Open*. 2020;2020:1–26.

39. Nakahara T, Iwase A, Goto M, Harata T, Suzuki M, Ienaga M, Kobayashi H, Takikawa S, Manabe S, Kikkawa F, et al. Evaluation of the safety of time-lapse observations for human embryos. *J. Assist. Reprod. Genet*. 2010;27:93–96.

40. Cruz M, Gadea B, Garrido N, Pedersen KS, Martínez M, Pérez-Cano I, Muñoz M, Meseguer M. Embryo quality, blastocyst and ongoing pregnancy rates in oocyte donation patients whose embryos were monitored by time-lapse imaging. *J. Assist. Reprod. Genet*. 2011;28:569–573.

41. Kirkegaard K, Hindkjaer JJ, Grøndahl ML, Kesmodel US, Ingerslev HJ. A randomized

clinical trial comparing embryo culture in a conventional incubator with a time-lapse incubator. *J. Assist. Reprod. Genet*. 2012;29:565–572.

42. Lundin K, Park H. Time-lapse technology for embryo culture and selection. *Ups. J. Med. Sci*. 2020;125:77–84.

43. Dang TT, Phung TM, Le H, Nguyen TB Van, Nguyen TS, Nguyen TLH, Nga VT, Chu DT, Hoang VL, Nguyen DB. Preimplantation genetic testing of aneuploidy by next generation sequencing: Association of maternal age and chromosomal abnormalities of blastocyst. *Open Access Maced. J. Med. Sci*. 2019;7:4427–4431.

44. Sachdev NM, Maxwell SM, Besser AG, Grifo JA. Diagnosis and clinical management of embryonic mosaicism. *Fertil. Steril*. 2017;107:6–11.

45. De Rycke M, Berckmoes V. Preimplantation Genetic Testing for Monogenic Disorders. *Genes (Basel)*. 2020;11:1–15.

46. Leaver M, Wells D. Non-invasive preimplantation genetic testing (niPGT): the next revolution in reproductive genetics? *Hum. Reprod. Update*. 2020;26:16–42.

47. Théry C, Witwer KW, Aikawa E, Alcaraz MJ, Anderson JD, Andriantsitohaina R, Antoniou A, Arab T, Archer F, Atkin-Smith GK, et al. 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. 2018;7.

48. Van Niel G, D'Angelo G, Raposo G. Shedding light on the cell biology of extracellular vesicles. *Nat. Rev. Mol. Cell Biol*. 2018;19:213–228.

49. van Niel G, Carter DRF, Clayton A, Lambert DW, Raposo G, Vader P. Challenges and directions in studying cell–cell communication by extracellular vesicles. *Nat. Rev. Mol. Cell Biol*. 2022;23:369–382.

50. Colombo M, Moita C, Van Niel G, Kowal J, Vigneron J, Benaroch P, Manel N, Moita LF, Théry C, Raposo G. Analysis of ESCRT functions in exosome biogenesis, composition and secretion highlights the heterogeneity of extracellular vesicles. *J. Cell Sci*. 2013;126:5553–5565.

51. Trajkovic K, Hsu C, Chiantia S, Rajendran L, Wenzel D, Wieland F, Schwille P, Brügger B, Simons M. Ceramide triggers budding of exosome vesicles into multivesicular endosomes. *Science*. 2008;319:1244–1247.

52. van Niel G, Charrin S, Simoes S, Romao M, Rochin L, Saftig P, Marks MS, Rubinstein E, Raposo G. The tetraspanin CD63 regulates ESCRT-independent and -dependent endosomal sorting during melanogenesis. *Dev. Cell*. 2011;21:708–721.

53. Lynch C, Panagopoulou M, Gregory CD. Extracellular vesicles arising from apoptotic cells in tumors: Roles in cancer pathogenesis and potential clinical applications. *Front. Immunol.* 2017;8:1174.

54. Abrahams VM, Kim YM, Straszewski SL, Romero R, Mor G. Macrophages and Apoptotic Cell Clearance During Pregnancy. *Am. J. Reprod. Immunol*. 2004;51:275–282.

55. Mir B, Goettsch C. Extracellular Vesicles as Delivery Vehicles of Specific Cellular Cargo. *Cells*. 2020;9:1601.

56. Greening DW, Nguyen HPT, Elgass K, Simpson RJ, Salamonsen LA. Human endometrial exosomes contain hormone-specific cargo modulating trophoblast adhesive capacity: Insights into endometrial-embryo interactions. *Biol. Reprod*. 2016;94:38–39.

57. Nolte-'t Hoen ENM, Buschow SI, Anderton SM, Stoorvogel W, Wauben MHM. Activated T cells recruit exosomes secreted by dendritic cells via LFA-1. *Blood*. 2009;113:1977–1981.

58. Svensson KJ, Christianson HC, Wittrup A, Bourseau-Guilmain E, Lindqvist E, Svensson LM, Mörgelin M, Belting M. Exosome uptake depends on ERK1/2-heat shock protein 27 signaling and lipid raft-mediated endocytosis negatively regulated by caveolin-1. *J. Biol. Chem*. 2013;288:17713–17724.

59. Tian T, Zhu YL, Zhou YY, Liang GF, Wang YY, Hu FH, Xiao ZD. Exosome uptake through clathrin-mediated endocytosis and macropinocytosis and mediating miR-21 delivery. *J. Biol. Chem*. 2014;289:22258–22267.

60. Feng D, Zhao WL, Ye YY, Bai XC, Liu RQ, Chang LF, Zhou Q, Sui SF. Cellular internalization of exosomes occurs through phagocytosis. *Traffic*. 2010;11:675–687.

61. Gurung S, Perocheau D, Touramanidou L, Baruteau J. The exosome journey: from biogenesis to uptake and intracellular signalling. *Cell Commun. Signal.* 2021;19:1–19.

62. Fitzner D, Schnaars M, Van Rossum D, Krishnamoorthy G, Dibaj P, Bakhti M, Regen T, Hanisch UK, Simons M. Selective transfer of exosomes from oligodendrocytes to microglia by macropinocytosis. *J. Cell Sci*. 2011;124:447–458.

63. Xiao F, Li J, Huang K, Li X, Xiong Y, Wu M, Wu L, Kuang W, Lv S, Wu L, et al. Macropinocytosis: mechanism and targeted therapy in cancers. *Am. J. Cancer Res*. 2021;11:14.

64. Goswamia D, Tannetta DS, Magee LA, Fuchisawa A, Redman CWG, Sargent IL, von Dadelszen P. Excess syncytiotrophoblast microparticle shedding is a feature of early-onset pre-eclampsia, but not normotensive intrauterine growth restriction. *Placenta*. 2006;27:56–61.

65. Tannetta D, Masliukaite I, Vatish M, Redman C, Sargent I. Update of syncytiotrophoblast derived extracellular vesicles in normal pregnancy and preeclampsia. *J. Reprod. Immunol.* 2017;119:98–106.

66. Giacomini E, Vago R, Sanchez AM, Podini P, Zarovni N, Murdica V, Rizzo R, Bortolotti D, Candiani M, Viganò P. Secretome of in vitro cultured human embryos contains extracellular vesicles that are uptaken by the maternal side. *Sci. Reports*. 2017;7:1–13.

67. Pallinger E, Bognar Z, Bodis J, Csabai T, Farkas N, Godony K, Varnagy A, Buzas E, Szekeres-Bartho J. A simple and rapid flow cytometry-based assay to identify a competent embryo prior to embryo transfer. *Sci. Rep.* 2017;7.

68. Simon B, Bolumar D, Amadoz A, Jimenez-Almazán J, Valbuena D, Vilella F, Moreno I. Identification and Characterization of Extracellular Vesicles and Its DNA Cargo Secreted During Murine Embryo Development. *Genes*. 2020;11:203.

69. Dissanayake K, Nõmm M, Lättekivi F, Ressaissi Y, Godakumara K, Lavrits A, Midekessa G, Viil J, Bæk R, Jørgensen MM, et al. Individually cultured bovine embryos produce extracellular vesicles that have the potential to be used as non-invasive embryo quality markers. *Theriogenology*. 2020;149:104–116.

70. Pavani KC, Hendrix A, Van Den Broeck W, Couck L, Szymanska K, Lin X, De Koster J, Van Soom A, Leemans B. Isolation and Characterization of Functionally Active Extracellular Vesicles from Culture Medium Conditioned by Bovine Embryos In Vitro. *Int. J. Mol. Sci*. 2018;20:38.

71. Pallinger E, Bognar Z, Bogdan A, Csabai T, Abraham H, Szekeres-Bartho J. PIBF+ extracellular vesicles from mouse embryos affect IL-10 production by CD8+ cells. *Sci. Reports*. 2018;8:1–9.

72. Mellisho EA, Velásquez AE, Nuñez MJ, Cabezas JG, Cueto JA, Fader C, Castro FO, Rodríguez-Álvarez L. Identification and characteristics of extracellular vesicles from bovine blastocysts produced in vitro. *PLoS One*. 2017;12:e0178306.

73. Vyas P, Balakier H, Librach CL. Ultrastructural identification of CD9 positive extracellular vesicles released from human embryos and transported through the zona pellucida. *Syst Biol Reprod Med*. 2019;65:273–280.

74. Battaglia R, Palini S, Vento ME, La Ferlita A, Lo Faro MJ, Caroppo E, Borzì P, Falzone L, Barbagallo D, Ragusa M, et al. Identification of extracellular vesicles and characterization of miRNA expression profiles in human blastocoel fluid. *Sci. Rep*. 2019;9.

75. Abu-Halima M, Häusler S, Backes C, Fehlmann T, Staib C, Nestel S, Nazarenko I, Meese E, Keller A. Micro-ribonucleic acids and extracellular vesicles repertoire in the spent culture media is altered in women undergoing In Vitro Fertilization. *Sci. Rep*. 2017;7.

76. Motta-Mejia C, Kandzija N, Zhang W, Mhlomi V, Cerdeira AS, Burdujan A, Tannetta D, Dragovic R, Sargent IL, Redman CW, et al. Placental Vesicles Carry Active Endothelial Nitric Oxide Synthase and Their Activity is Reduced in Preeclampsia. *Hypertension*. 2017;70:372–381.

77. Tannetta D, Dragovic R, Alyahyaei Z, Southcombe J. Extracellular vesicles and reproduction–promotion of successful pregnancy. *Cell. Mol. Immunol*. 2014;11:548–563.

78. Machtinger R, Laurent LC, Baccarelli AA. Extracellular vesicles: roles in gamete maturation, fertilization and embryo implantation. *Hum. Reprod. Update*. 2016;22:182–193.

79. Dellar ER, Hill C, Melling GE, Carter DR, Alberto Baena-Lopez L, Elizabeth Dellar CR. Unpacking extracellular vesicles: RNA cargo loading and function. *J. Extracell. Biol*. 2022;1:e40.

80. Ferreira YJ, Gardiner C, Poli M, Turner K, Child T, Sargent IL, Theofanakis C, Dinopoulou V, Mavrogianni D, Anagnostou E, et al. Session 66: Embryo quality: does it predict pregnancy? *Hum. Reprod*. 2013;28:i104–i106.

81. Giacomini E, Vago R, Sanchez AM, Podini P, Zarovni N, Murdica V, Rizzo R, Bortolotti D, Candiani M, Viganò P. Secretome of in vitro cultured human embryos contains extracellular vesicles that are uptaken by the maternal side. *Sci. Reports*. 2017;7:1–13.

82. Pallinger E, Bognar Z, Bodis J, Csabai T, Farkas N, Godony K, Varnagy A, Buzas E, Szekeres-Bartho J. A simple and rapid flow cytometry-based assay to identify a competent embryo prior to embryo transfer. *Sci. Rep.* 2017;7.

83. Abu-Halima M, Häusler S, Backes C, Fehlmann T, Staib C, Nestel S, Nazarenko I, Meese E, Keller A. Micro-ribonucleic acids and extracellular vesicles repertoire in the spent culture media is altered in women undergoing In Vitro Fertilization. *Sci. Rep*. 2017;7.

84. Vyas P, Balakier H, Librach CL. Ultrastructural identification of CD9 positive extracellular vesicles released from human embryos and transported through the zona pellucida. *Syst Biol Reprod Med*. 2019;65:273–280.

85. Battaglia R, Palini S, Vento ME, La Ferlita A, Lo Faro MJ, Caroppo E, Borzì P, Falzone L, Barbagallo D, Ragusa M, et al. Identification of extracellular vesicles and characterization of miRNA expression profiles in human blastocoel fluid. *Sci. Rep*. 2019;9.

86. Rosenbluth EM, Shelton DN, Wells LM, Sparks AET, Van Voorhis BJ. Human embryos secrete microRNAs into culture media--a potential biomarker for implantation. *Fertil. Steril*. 2014;101:1493–1500.

87. Kabe Y, Sakamoto S, Hatakeyama M, Yamaguchi Y, Suematsu M, Itonaga M, Handa H. Application of high-performance magnetic nanobeads to biological sensing devices. *Anal. Bioanal. Chem*. 2019;411:1825–1837.

88. Bardin N, Anfosso F, Massé JM, Cramer E, Sabatier F, Bivic A Le, Sampol J, Dignat-

George F. Identification of CD146 as a component of the endothelial junction involved in the control of cell-cell cohesion. *Blood*. 2001;98:3677–3684.

89. Bardin N, Francès V, Combes V, Sampol J, Dignat-George F. CD146: biosynthesis and production of a soluble form in human cultured endothelial cells. *FEBS Lett.* 1998;421:12–14.

90. Bouvier S, Paulmyer-Lacroix O, Molinari N, Bertaud A, Paci M, Leroyer A, Robert S, George FD, Blot-Chabaud M, Bardin N. Soluble CD146, an innovative and non-invasive biomarker of embryo selection for in vitro fertilization. *PLoS One.* 2017;12:e0173724.

91. Liu Q, Yan X, Li Y, Zhang Y, Zhao X, Shen Y. Pre-eclampsia is associated with the failure of melanoma cell adhesion molecule (MCAM/CD146) expression by intermediate trophoblast. *Lab. Investig*. 2003;84:221–228.

92. Liu Q, Zhang B, Zhao X, Zhang Y, Liu Y, Yan X. Blockade of adhesion molecule CD146 causes pregnancy failure in mice. *J. Cell. Physiol*. 2008;215:621–626.

93. Bouvier S, Kaspi E, Joshkon A, Paulmyer-Lacroix O, Piercecchi-Marti MD, Sharma A, Leroyer AS, Bertaud A, Gris JC, Dignat-George F, et al. The Role of the Adhesion Receptor CD146 and Its Soluble Form in Human Embryo Implantation and Pregnancy. *Front. Immunol.* 2021;12:3227.

94. Park EJ, Myint PK, Ito A, Appiah MG, Darkwah S, Kawamoto E, Shimaoka M. Integrin-Ligand Interactions in Inflammation, Cancer, and Metabolic Disease: Insights Into the Multifaceted Roles of an Emerging Ligand Irisin. *Front. Cell Dev. Biol.* 2020;8:1196.

95. Singh H, Aplin JD. Adhesion molecules in endometrial epithelium: tissue integrity and embryo implantation. *J. Anat.* 2009;215:3.

96. Quenby S, Anim-Somuah M, Kalumbi C, Farquharson R, Aplin JD. Different types of recurrent miscarriage are associated with varying patterns of adhesion molecule expression in endometrium. *Reprod. Biomed. Online*. 2007;14:224–234.

97. Klentzeris LD, Bulmer JN, Trejdosiewicz LK, Morrison L, Cooke ID. Infertility: Beta-1 integrin cell adhesion molecules in the endometrium of fertile and infertile women. *Hum. Reprod*. 1993;8:1223–1230.

98. Lessey BA, Castelbaum AJ, Sawin SW, Sun J. Integrins as markers of uterine receptivity in women with primary unexplained infertility. *Fertil. Steril.* 1995;63:535–542.

99. Chen G, Xin A, Liu Y, Shi C, Chen J, Tang X, Chen Y, Yu M, Peng X, Li L, et al. Integrins β1 and β3 are biomarkers of uterine condition for embryo transfer. *J. Transl. Med*. 2016;14:1–10.

100. Andreu Z, Yáñez-Mó M. Tetraspanins in extracellular vesicle formation and function. *Front. Immunol.* 2014;5:442.

101. Pols MS, Klumperman J. Trafficking and function of the tetraspanin CD63. *Exp. Cell Res.* 2009;315:1584–1592.

102. Tugues S, Honjo S, König C, Padhan N, Kroon J, Gualandi L, Li X, Barkefors I, Thijssen VL, Griffioen AW, et al. Tetraspanin CD63 promotes vascular endothelial growth factor receptor 2-β1 integrin complex formation, thereby regulating activation and downstream signaling in endothelial cells in Vitro and in Vivo. *J. Biol. Chem*. 2013;288:19060–19071.

103. Prochazkova J, Slavik L, Ulehlova J, Prochazka M. The role of tissue factor in normal pregnancy and in the development of preeclampsia: A review. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*. 2015;159:192–196.

104. Kinev A V., Roubey RAS. Tissue factor in the antiphospholipid syndrome. *Lupus*. 2008;17:952–958.

105. Vikerfors A, Mobarrez F, Bremme K, Holmström M, Ågren A, Eelde A, Bruzelius M, Antovic A, Wallén H, Svenungsson E. Studies of microparticles in patients with the antiphospholipid syndrome (APS). *Lupus*. 2012;21:802–805.

106. Bugge TH, Xiao Q, Kombrinck KW, Flick MJ, Holmbäck K, Danton MJS, Colbert MC, Witte DP, Fujikawa K, Davie EW, et al. Fatal embryonic bleeding events in mice lacking tissue factor, the cell-associated initiator of blood coagulation. *Proc. Natl. Acad. Sci. U. S. A*. 1996;93:6258–6263.

107. Toomey JR, Kratzer KE, Lasky NM, Stanton JJ, Broze GJ. Targeted Disruption of the Murine Tissue Factor Gene Results in Embryonic Lethality. *Blood*. 1996;88:1583–1587.

108. Parry GCN, Mackman N. Mouse embryogenesis requires the tissue factor extracellular domain but not the cytoplasmic domain. *J. Clin. Invest.* 2000;105:1547–1554.

109. Koizume S, Miyagi Y. Tissue Factor-Factor VII Complex As a Key Regulator of Ovarian Cancer Phenotypes. *Biomark. Cancer*. 2015;7:1–13.

110. Han X, Guo B, Li Y, Zhu B. Tissue factor in tumor microenvironment: a systematic review. *J. Hematol. Oncol*. 2014;7:54.

111. Kocatürk B, Van Den Berg YW, Tieken C, Mieog JSD, De Kruijf EM, Engels CC, Van Der Ent MA, Kuppen PJ, Van De Velde CJ, Ruf W, et al. Alternatively spliced tissue factor promotes breast cancer growth in a β1 integrin-dependent manner. *Proc. Natl. Acad. Sci. U. S. A.* 2013;110:11517–11522.

112. Lockwood CJ, Nemerson Y, Guller S, Krikun G, Alvarez M, Hausknecht V, Gurpide E, Schatz F. Progestational regulation of human endometrial stromal cell tissue factor expression during decidualization. *J. Clin. Endocrinol. Metab.* 1993;76:231–236.

113. Schatz F, Krikun G, Caze R, Rahman M, Lockwood CJ. Progestin-regulated expression of tissue factor in decidual cells: implications in endometrial hemostasis, menstruation and angiogenesis. *Steroids*. 2003;68:849–860.

114. Krikun G, Lockwood CJ, Paidas MJ. Tissue factor and the endometrium: From physiology to pathology. *Thromb. Res.* 2009;124:393–396.

115. Kabe Y, Suematsu M, Sakamoto S, Hirai M, Koike I, Hishiki T, Matsuda A, Hasegawa Y, Tsujita K, Ono M, et al. Development of a Highly Sensitive Device for Counting the Number of Disease-Specific Exosomes in Human Sera. *Clin. Chem.* 2018;64:1463–1473.

116. Van Der Pol E, Hoekstra AG, Sturk A, Otto C, Van Leeuwen TG, Nieuwland R. Optical and non-optical methods for detection and characterization of microparticles and exosomes. *J. Thromb. Haemost*. 2010;8:2596–2607.

117. Erdbrügger U, Rudy CK, E. Etter M, Dryden KA, Yeager M, Klibanov AL, Lannigan J. Imaging flow cytometry elucidates limitations of microparticle analysis by conventional flow cytometry. *Cytom. Part A*. 2014;85:756–770.

118. Tian Y, Gong M, Hu Y, Liu H, Zhang W, Zhang M, Hu X, Aubert D, Zhu S, Wu L, et al. Quality and efficiency assessment of six extracellular vesicle isolation methods by nano-flow cytometry. *J. Extracell. vesicles*. 2019;9.

119. Vogel R, Savage J, Muzard J, Camera G Della, Vella G, Law A, Marchioni M, Mehn D, Geiss O, Peacock B, et al. Measuring particle concentration of multimodal synthetic reference materials and extracellular vesicles with orthogonal techniques: Who is up to the challenge? *J. Extracell. Vesicles*. 2021;10.

120. Cha J, Sun X, Dey SK. Mechanisms of implantation: strategies for successful pregnancy. *Nat. Med*. 2012;18:1754–1767.

121. Ochoa-Bernal MA, Fazleabas AT. Physiologic Events of Embryo Implantation and Decidualization in Human and Non-Human Primates. *Int. J. Mol. Sci*. 2020;21:1973.

122. Kodaman PH, Taylor HS. Hormonal regulation of implantation. *Obstet. Gynecol. Clin. North Am.* 2004;31:745–766.

123. Simón C, Martín JC, Pellicer A. Paracrine regulators of implantation. *Best Pract. Res. Clin. Obstet. Gynaecol.* 2000;14:815–826.

124. Yang Y, Chen X, Saravelos SH, Liu Y, Huang J, Zhang J, Li TC. HOXA-10 and E-cadherin expression in the endometrium of women with recurrent implantation failure and recurrent miscarriage. *Fertil. Steril*. 2017;107:136-143.e2.

125. Stewart CL, Kaspar P, Brunet LJ, Bhatt H, Gadi I, Köntgen F, Abbondanzo SJ. Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. *Nat.* 1992;359:76–79.

126. Shuya LL, Menkhorst EM, Yap J, Li P, Lane N, Dimitriadis E. Leukemia Inhibitory Factor Enhances Endometrial Stromal Cell Decidualization in Humans and Mice. *PLoS One*. 2011;6:e25288.

127. Salmassi A, Mettler L, Jonat W, Buck S, Koch K, Schmutzler AG. Circulating level of macrophage colony-stimulating factor can be predictive for human in vitro fertilization outcome. *Fertil. Steril.* 2010;93:116–123.

128. Othman R, Omar MH, Shan LP, Shafiee MN, Jamal R, Mokhtar NM. Microarray profiling of secretory-phase endometrium from patients with recurrent miscarriage. *Reprod. Biol.* 2012;12:183–199.

129. Nikas G, Aghajanova L. Endometrial pinopodes: some more understanding on human implantation? *Reprod. Biomed. Online*. 2002;4:18–23.

130. Enciso M, Aizpurua J, Rodríguez-Estrada B, Jurado I, Ferrández-Rives M, Rodríguez E, Pérez-Larrea E, Climent AB, Marron K, Sarasa J. The precise determination of the window of implantation significantly improves ART outcomes. *Sci. Reports*. 2021;11:1–8.

131. Rodriguez-Boulan E, Nelson WJ. Morphogenesis of the polarized epithelial cell phenotype. *Science*. 1989;245:718–725.

132. Thie M, Fuchs P, Denker HW. Epithelial cell polarity and embryo implantation in mammals. *Int. J. Dev. Biol*. 2003;40:389–393.

133. Heneweer C, Kruse LH, Kindháuser F, Schmidt M, Jakobs KH, Denker HW, Thie M. Adhesiveness of human uterine epithelial RL95-2 cells to trophoblast: Rho protein regulation. *Mol. Hum. Reprod*. 2002;8:1014–1022.

134. Stavreus-Evers A, Nikas G, Sahlin L, Eriksson H, Landgren BM. Formation of pinopodes in human endometrium is associated with the concentrations of progesterone and progesterone receptors. *Fertil. Steril.* 2001;76:782–791.

135. Usadi RS, Murray MJ, Bagnell RC, Fritz MA, Kowalik AI, Meyer WR, Lessey BA. Temporal and morphologic characteristics of pinopod expression across the secretory phase of the endometrial cycle in normally cycling women with proven fertility. *Fertil. Steril.* 2003;79:970–974.

136. Harduf H, Goldman S, Shalev E. Human uterine epithelial RL95-2 and HEC-1A cell-line adhesiveness: the role of plexin B1. *Fertil. Steril*. 2007;87:1419–1427.

137. Hevir-Kene N, Rižner TL. The endometrial cancer cell lines Ishikawa and HEC-1A, and the control cell line HIEEC, differ in expression of estrogen biosynthetic and metabolic genes, and in androstenedione and estrone-sulfate metabolism. *Chem. Biol. Interact.* 2015;234:309–319.

138. Nyen T Van, Moiola CP, Colas E, Annibali D, Amant F. Modeling Endometrial Cancer: Past, Present, and Future. *Int. J. Mol. Sci*. 2018;19.

139. Nishida M, Kasahara K, Kaneko M, Iwasaki H, Hayashi K. [Establishment of a new human endometrial adenocarcinoma cell line, Ishikawa cells, containing estrogen and progesterone receptors]. *Nihon Sanka Fujinka Gakkai Zasshi*. 1985;37:1103–1111.

140. Nishida M. The Ishikawa cells from birth to the present. *Hum. Cell*. 2002;15:104–117.

141. Zhang D, Pingping L, Zhang R, Luo Q, Ding G, Yin L, Li J, Xu G, Qu F, Sheng J, et al. A new model for embryo implantation: coculture of blastocysts and Ishikawa cells. *Gynecol. Endocrinol*. 2012;28:288–292.

142. Fleming H. Differentiation in human endometrial cells in monolayer culture: dependence on a factor in fetal bovine serum. *J. Cell. Biochem*. 1995;57:262–270.

143. Dharmaraj N, Gendler SJ, Carson DD. Expression of Human MUC1 During Early Pregnancy in the Human MUC1 Transgenic Mouse Model. *Biol. Reprod*. 2009;81:1182.

144. Meseguer M, Aplin JD, Caballero-Campo P, O'Connor JE, Martín JC, Remohí J, Pellicer A, Simón C. Human Endometrial Mucin MUC1 Is Up-Regulated by Progesterone and Down-Regulated In Vitro by the Human Blastocyst. *Biol. Reprod*. 2001;64:590–601.

145. Horne AW, Lalani EN, Margara RA, Ryder TA, Mobberley MA, White JO. The expression pattern of MUC1 glycoforms and other biomarkers of endometrial receptivity in fertile and infertile women. *Mol. Reprod. Dev*. 2005;72:216–229.

146. Wu B, Zhou H, Hu L, Mu Y, Wu Y. Involvement of PKCα activation in TF/VIIa/PAR2-induced proliferation, migration, and survival of colon cancer cell SW620. *Tumor Biol.* 2013;34:837–846.

147. Evrova O, Buschmann J. In vitro and in vivo effects of PDGF-BB delivery strategies on tendon healing: A review. *Eur. Cells Mater.* 2017;34:15–39.

148. Gebhard C, Akhmedov A, Mocharla P, Angstenberger J, Sahbai S, Camici GG, Lüscher TF, Tanner FC. PDGF-CC induces tissue factor expression: role of PDGF receptor alpha/beta. *Basic Res. Cardiol*. 2010;105:349–356.

149. Kamimura M, Bea F, Akizawa T, Katus HA, Kreuzer J, Viedt C. Platelet-derived growth factor induces tissue factor expression in vascular smooth muscle cells via activation of Egr-1. *Hypertension*. 2004;44:944–951.

150. Al-Amrani S, Al-Jabri Z, Al-Zaabi A, Alshekaili J, Al-Khabori M. Proteomics: Concepts and applications in human medicine. *World J. Biol. Chem*. 2021;12:57.

151. Holman JD, Dasari S, Tabb DL. Informatics of protein and posttranslational modification detection via shotgun proteomics. *Methods Mol. Biol*. 2013;1002:167–179.

152. Banks RE, Dunn MJ, Hochstrasser DF, Sanchez JC, Blackstock W, Pappin DJ, Selby PJ. Proteomics: new perspectives, new biomedical opportunities. *Lancet*. 2000;356:1749–1756.

153. Serpa JJ, Parker CE, Petrotchenko E V., Han J, Pana J, Borchers CH. Mass spectrometry-based structural proteomics. *Eur. J. Mass Spectrom*. 2012;18:251–267.

154. Monti M, Orrù S, Pagnozzi D, Pucci P. Functional proteomics. *Clin. Chim. Acta*. 2005;357:140–150.

155. Krudysz-Amblo J, Jennings ME, Mann KG, Butenas S. Carbohydrates and activity of natural and recombinant tissue factor. *J. Biol. Chem*. 2010;285:3371–3382.

156. Belting M, Dorrell MI, Sandgren S, Aguilar E, Ahamed J, Dorfleutner A, Carmeliet P, Mueller BM, Friedlander M, Ruf W. Regulation of angiogenesis by tissue factor cytoplasmic domain signaling. *Nat. Med*. 2004;10:502–509.

157. Jee B, Dhar R, Singh S, Karmakar S. Heat Shock Proteins and Their Role in Pregnancy: Redefining the Function of "Old Rum in a New Bottle." *Front. Cell Dev. Biol*. 2021;9.

158. Retzlaff M, Stahl M, Eberl HC, Lagleder S, Beck J, Kessler H, Buchner J. Hsp90 is regulated by a switch point in the C-terminal domain. *EMBO Rep*. 2009;10:1147–1153.

159. Chadli A, Graham JD, Abel MG, Jackson TA, Gordon DF, Wood WM, Felts SJ, Horwitz KB, Toft D. GCUNC-45 is a novel regulator for the progesterone receptor/hsp90 chaperoning pathway. *Mol. Cell. Biol*. 2006;26:1722–1730.

160. Rodriguez MS, Egaña I, Lopitz-Otsoa F, Aillet F, Lopez-Mato MP, Dorronroso A, Lobato-

Gil S, Sutherland JD, Barrio R, Trigueros C, et al. The RING ubiquitin E3 RNF114 interacts with A20 and modulates NF-κB activity and T-cell activation. *Cell Death Dis*. 2014;5.

161. Hanafusa H, Ishikawa K, Kedashiro S, Saigo T, Iemura SI, Natsume T, Komada M, Shibuya H, Nara A, Matsumoto K. Leucine-rich repeat kinase LRRK1 regulates endosomal trafficking of the EGF receptor. *Nat. Commun*. 2011;2:1–12.

162. Jacobo-Baca G, Salazar-Ybarra RA, Torres-de-la-Cruz V, Guzmán-López S, Elizondo-Omaña RE, Guzmán-López A, Vázquez-Barragán MÁ, Martínez-de-Villarreal LE. Proteomic profile of preeclampsia in the first trimester of pregnancy. *J Matern Fetal Neonatal Med*. 2020;35:3446–3452.

163. Aspenström P. The WASP-Binding Protein WIRE Has a Role in the Regulation of the Actin Filament System Downstream of the Platelet-Derived Growth Factor Receptor. *Exp. Cell Res.* 2002;279:21–33.

164. Lill R, Dutkiewicz R, Freibert SA, Heidenreich T, Mascarenhas J, Netz DJ, Paul VD, Pierik AJ, Richter N, Stümpfig M, et al. The role of mitochondria and the CIA machinery in the maturation of cytosolic and nuclear iron-sulfur proteins. *Eur. J. Cell Biol*. 2015;94:280–291.

165. Cossée M, Puccio H, Gansmuller A, Koutnikova H, Dierich A, LeMeur M, Fischbeck K, Dollé P, Kœnig M. Inactivation of the Friedreich ataxia mouse gene leads to early embryonic lethality without iron accumulation. *Hum. Mol. Genet.* 2000;9:1219–1226.

166. Iyama T, Wilson DM. DNA repair mechanisms in dividing and non-dividing cells. *DNA Repair (Amst).* 2013;12:620–636.

167. Mollet J, Giurgea I, Schlemmer D, Dallner G, Chretien D, Delahodde A, Bacq D, De Lonlay P, Munnich A, Rötig A. Prenyldiphosphate synthase, subunit 1 (PDSS1) and OH-benzoate polyprenyltransferase (COQ2) mutations in ubiquinone deficiency and oxidative phosphorylation disorders. *J. Clin. Invest*. 2007;117:765–772.

168. Lim R, Nguyen-Ngo C, Lappas M. Targeting bromodomain-containing proteins to prevent spontaneous preterm birth. *Clin. Sci*. 2019;133:2379–2400.

169. Liong S, Barker G, Lappas M. Bromodomain protein BRD4 is increased in human placentas

from women with early-onset preeclampsia. *Reproduction*. 2018;155:573–582.

170. Melner MH, Haas AL, Klein JM, Brash AR, Boeglin WE, NagDas SK, Winfrey VP, Olson GE. Demonstration of Ubiquitin Thiolester Formation of UBE2Q2 (UBCi), a Novel Ubiquitin-Conjugating Enzyme with Implantation Site-Specific Expression. *Biol. Reprod*. 2006;75:395–406.

171. Bebington C, Doherty FJ, Ndukwe G, Fleming SD. The progesterone receptor and ubiquitin are differentially regulated within the endometrial glands of the natural and stimulated cycle. *Mol. Hum. Reprod*. 2000;6:264–268.

172. Hesse M, Zimek A, Weber K, Magin TM. Comprehensive analysis of keratin gene clusters in humans and rodents. *Eur. J. Cell Biol.* 2004;83:19–26.

173. Alam H, Sehgal L, Kundu ST, Dalal SN, Vaidya MM. Novel function of keratins 5 and 14 in proliferation and differentiation of stratified epithelial cells. *Mol. Biol. Cell*. 2011;22:4068–4078.

174. Park HS, Kim JO, An HJ, Ryu CS, Ko EJ, Kim YR, Ahn EH, Lee WS, Kim JH, Kim NK. Genetic polymorphisms of the cobalamin transport system are associated with idiopathic recurrent implantation failure. *J. Assist. Reprod. Genet*. 2019;36:1513.

175. Ramirez MI, Amorim MG, Gadelha C, Milic I, Welsh JA, Freitas VM, Nawaz M, Akbar N, Couch Y, Makin L, et al. Technical challenges of working with extracellular vesicles. *Nanoscale*. 2018;10:881–906.

176. Wu D, Yan J, Shen X, Sun Y, Thulin M, Cai Y, Wik L, Shen Q, Oelrich J, Qian X, et al. Profiling surface proteins on individual exosomes using a proximity barcoding assay. *Nat. Commun*. 2019;10:1–10.

177. Enquobahrie DA, Qiu C, Muhie SY, Williams MA. Maternal peripheral blood gene expression in early pregnancy and preeclampsia. *Int. J. Mol. Epidemiol. Genet*. 2011;2:78.

178. Opstelten R, Suwandi JS, Slot MC, Morgana F, Scott AM, Laban S, Nikolic T, Turksma AW, Kroeze A, Voermans C, et al. GPA33 is expressed on multiple human blood cell types and distinguishes CD4+ central memory T cells with and without effector function. *Eur. J. Immunol*. 2021;51:1377–1389.

179. Chapouly C, Guimbal S, Hollier PL, Renault MA. Role of Hedgehog Signaling in Vasculature Development, Differentiation, and Maintenance. *Int. J. Mol. Sci*. 2019;20:3076.

180. Lawson ND, Vogel AM, Weinstein BM. sonic hedgehog and vascular endothelial growth factor Act Upstream of the Notch Pathway during Arterial Endothelial Differentiation. *Dev. Cell*. 2002;3:127–136.

181. Huang Z, Huang S, Song T, Yin Y, Tan C. Placental Angiogenesis in Mammals: A Review of the Regulatory Effects of Signaling Pathways and Functional Nutrients. *Adv. Nutr*. 2021;12:2415.

182. Pan Y, Yan L, Chen Q, Wei C, Dai Y, Tong X, Zhu H, Lu M, Zhang Y, Jin X, et al. Dysfunction of Shh signaling activates autophagy to inhibit trophoblast motility in recurrent miscarriage. *Exp. Mol. Med*. 2021;53:52.

183. Deans JP, Kalt L, Ledbetter JA, Schieven GL, Bolen JB, Johnson P. Association of 75/80-kDa Phosphoproteins and the Tyrosine Kinases Lyn, Fyn, and Lck with the B Cell Molecule CD20. *J. Biol. Chem*. 1995;270:22632–22638.

184. Popoff IJ, Savage JA, Blake J, Johnson P, Deans JP. The association between CD20 and Src-family Tyrosine kinases requires an additional factor. *Mol. Immunol.* 1998;35:207–214.

185. Makarenkova H, Sugiura H, Yamagata K, Owens G. Alternatively spliced variants of protocadherin 8 exhibit distinct patterns of expression during mouse development. *Biochim. Biophys. Acta - Gene Struct. Expr.* 2005;1681:150–156.

186. Li Z, Gou J, Jia J, Zhao X. MicroRNA-429 functions as a regulator of epithelial–mesenchymal transition by targeting Pcdh8 during murine embryo implantation. *Hum. Reprod*. 2015;30:507–518.

187. Shi S, Tan Q, Liang J, Cao D, Wang S, Liang J, Chen K, Wang Z. Placental trophoblast cell-derived exosomal microRNA-1290 promotes the interaction between endometrium and embryo by targeting LHX6. *Mol. Ther. Nucleic Acids*. 2021;26:760–772.

188. Park C, Bae H, Bazer FW, Song G, Lim W. Activation of CCL20 and its receptor CCR6 promotes endometrium preparation for implantation and placenta development during the

early pregnancy period in pigs. *Dev. Comp. Immunol.* 2019;92:35–42.

189. Amjadi F, Zandieh Z, Mehdizadeh M, Aghajanpour S, Raoufi E, Aghamajidi A, Aflatoonian R. The uterine immunological changes may be responsible for repeated implantation failure. *J. Reprod. Immunol.* 2020;138:103080.

190. Bamberger AM, Sudahl S, Löning T, Wagener C, Bamberger CM, Drakakis P, Coutifaris C, Makrigiannakis A. The Adhesion Molecule CEACAM1 (CD66a, C-CAM, BGP) Is Specifically Expressed by the Extravillous Intermediate Trophoblast. *Am. J. Pathol.* 2000;156:1165–1170.

191. Damsky CH, Fitzgerald ML, Fisher SJ. Distribution patterns of extracellular matrix components and adhesion receptors are intricately modulated during first trimester cytotrophoblast differentiation along the invasive pathway, in vivo. *J. Clin. Invest*. 1992;89:210.

192. Kilburn BA, Wang J, Duniec-Dmuchkowski ZM, Leach RE, Romero R, Armant DR. Extracellular Matrix Composition and Hypoxia Regulate the Expression of HLA-G and Integrins in a Human Trophoblast Cell Line. *Biol. Reprod*. 2000;62:739–747.

193. Elnaggar A, Farag AH, Gaber ME, Hafeez MA, Ali MS, Atef AM. AlphaVBeta3 Integrin expression within uterine endometrium in unexplained infertility: a prospective cohort study. *BMC Womens. Health.* 2017;17.

194. Ordi J, Creus M, Ferrer B, Fábregues F, Carmona F, Casamitjana R, Vanrell JA, Balasch J. Midluteal endometrial biopsy and alphavbeta3 integrin expression in the evaluation of the endometrium in infertility: implications for fecundity. *Int. J. Gynecol. Pathol.* 2002;21:231–238.

195. Jiang Y, Chen L, Taylor RN, Li C, Zhou X. Physiological and pathological implications of retinoid action in the endometrium. *J. Endocrinol*. 2018;236:R169–R188.

196. Hughes CHK, Inskeep EK, Pate JL. Temporal changes in the corpus luteum during early pregnancy reveal regulation of pathways that enhance steroidogenesis and suppress luteolytic mechanisms. *Biol. Reprod*. 2020;103:70–84.

197. Han BC, Xia HF, Sun J, Yang Y, Peng JP. Retinoic acid-metabolizing enzyme cytochrome

P450 26a1 (cyp26a1) is essential for implantation: functional study of its role in early pregnancy. *J. Cell. Physiol*. 2010;223:471–479.

198. Tomari H, Kawamura T, Asanoma K, Egashira K, Kawamura K, Honjo K, Nagata Y, Kato K. Contribution of senescence in human endometrial stromal cells during proliferative phase to embryo receptivity. *Biol. Reprod*. 2020;103:104–113.

199. Kauma S, Huff T, Krystal G, Ryan J, Takacs P, Turner T. The expression of stem cell factor and its receptor, c-kit in human endometrium and placental tissues during pregnancy. *J. Clin. Endocrinol. Metab*. 1996;81:1261–1266.

200. Liang J, Wu YL, Chen BJ, Zhang W, Tanaka Y, Sugiyama H. The C-Kit Receptor-Mediated Signal Transduction and Tumor-Related Diseases. *Int. J. Biol. Sci*. 2013;9:435.

201. Jeong Y, Ock SA, Yoo JG, Yu DY, Choi I. The Cxadr–Adam10 complex plays pivotal roles in tight junction integrity and early trophoblast development in mice. *Mol. Reprod. Dev.* 2019;86:1628–1638.

202. Kai Y, Shonan Y, Clinic Y, Mei H, Kawano H, Nakajima N, Takai A, Kumon M, Yamashita N. Transcriptomic signatures in trophectoderm and inner cell mass of human blastocysts with different expectations of pregnancy. *PLoS One*. 2022; 17: e0278663

203. Amin Arnaout M. Structure and function of the leukocyte adhesion molecules CD11/CD18. *Blood*. 1990;75:1037–1050.

204. Bouhouche A, Tabache Y, Askander O, Charoute H, Mesnaoui N, Belayachi L, El Hafidi N, Hardizi H, El Fahime E, Erreimi N, et al. Novel ITGB2 Mutation Is Responsible for a Severe Form of Leucocyte Adhesion Deficiency Type 1. *Biomed Res. Int.* 2022.

205. Braundmeier AG, Dayger CA, Mehrotra P, Belton RJ, Nowak RA. EMMPRIN is secreted by human uterine epithelial cells in microvesicles and stimulates metalloproteinase production by human uterine fibroblast cells. *Reprod. Sci*. 2012;19:1292–1301.

206. Segura-Benítez M, Carbajo-García MC, Corachán A, Faus A, Pellicer A, Ferrero H. Proteomic analysis of extracellular vesicles secreted by primary human epithelial endometrial cells reveals key proteins related to embryo implantation. *Reprod. Biol. Endocrinol*. 2022;20:1–16.

207. Saadeldin IM, Kim SJ, Choi Y Bin, Lee BC. Improvement of Cloned Embryos Development by Co-Culturing with Parthenotes: A Possible Role of Exosomes/Microvesicles for Embryos Paracrine Communication. *Cell Reprogram*. 2014;16:223–234.

208. Pavani KC, Meese T, Pascottini OB, Guan XF, Lin X, Peelman L, Hamacher J, van Nieuwerburgh F, Deforce D, Boel A, et al. Hatching is modulated by microRNA-378a-3p derived from extracellular vesicles secreted by blastocysts. *Proc. Natl. Acad. Sci. U. S. A.* 2022;119:e2122708119.

209. Su Y, Li Q, Zhang Q, Li Z, Yao X, Guo Y, Xiao L, Wang X, Ni H. Exosomes derived from placental trophoblast cells regulate endometrial epithelial receptivity in dairy cows during pregnancy. *J. Reprod. Dev*. 2022;68:21–29.

210. Treff NR, Franasiak JM. Detection of segmental aneuploidy and mosaicism in the human preimplantation embryo: technical considerations and limitations. *Fertil. Steril.* 2017;107:27–31.

211. Le Guen M, Schantz C, Régnier-Loilier A, de La Rochebrochard E. Reasons for rejecting hormonal contraception in Western countries: A systematic review. *Soc. Sci. Med.* 2021;284:114247.