UNIVERSITA' VITA-SALUTE SAN RAFFAELE

CORSO DI DOTTORATO DI RICERCA INTERNAZIONALE

IN MEDICINA MOLECOLARE

CURRICULUM IN MEDICINA CLINICA E SPERIMENTALE

UTERINE FLUID EXTRACELLULAR VESICLES AS A LIQUID BIOPSY FOR THE DIAGNOSIS OF ENDOMETRIAL RECEPTIVITY

DoS: Massimo Candiani

Second Supervisor: Marie-Madeleine Dolmans

Tesi di DOTTORATO di RICERCA di Valeria Stella Vanni

matr. 013852

Ciclo di dottorato XXXIV

SSD MED/40

CONSULTAZIONE TESI DI DOTTORATO DI RICERCA

Il/la sottoscritto/*I*	Valeria Stella Vanni
Matricola / *registration number*	013852
nat_ a/ *born at*	Milano
il/*on*	27/02/1988

autore della tesi di Dottorato di ricerca dal titolo / *author of the PhD Thesis titled*

UTERINE FLUID EXTRACELLULAR VESICLES AS A LIQUID BIOPSY FOR THE DIAGNOSIS

OF ENDOMETRIAL RECEPTIVITY

X AUTORIZZA la Consultazione della tesi / AUTHORIZES *the public release of the thesis*

☐ NON AUTORIZZA la Consultazione della tesi per …… mesi /DOES NOT AUTHORIZE *the public release of the thesis for …… months*

Data/Date 05/05/2022 Firma /Signature 

DECLARATION

This thesis has been composed by myself and has not been used in any previous application for a degree. Throughout the text I use both 'I' and 'We' interchangeably.

Please note that a portion of the Results and Materials and Methods the present study and more specifically those referring to the Preliminary phase and to Aim 1 and 2 of the project have already been published in a recent article of which I am a co-author (Giacomini et al., 2021). Validation analyses (Aim 3 of the project) as well as subgroup analyses instead represent novel, unpublished material.

All the results presented here were obtained by myself, except for:

1) *Transmission Electron Microscopy (Preliminary results, Chapter 1.6, Figure 2), was performed by the Alembic, Experimental Imaging Centre, San Raffaele Scientific Institute, Milan, Italy.*
2) *Western blot analysis (Preliminary results, Chapter 1.6, Figure 3), were performed in collaboration with Dr. V.Murdica, Urological Research Institute, IRCCS San Raffaele Scientific Institute, Milan, Italy*
3) *RNA extraction and sequencing library preparation (Methods, Chapter 5.8), was performed in collaboration with Dr. E.Giacomini, Reproductive Sciences Laboratory, Obstetrics and Gynecology Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy*
4) *Nanoparticle Tracking Analysis (Results, Chapter 3.2, Figures 8 and 9), was performed by Dr. E.Giacomini in collaboration with V.Bollati and L.Cantone at the EPIGET Lab, Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy*
5) *RNA-Seq analysis and Differential Gene Expression analyses (Results, Chapters 3.3-3.5, 3.8, 3.9, 3.11, Figures 11-14, 17-21, 23-24), were performed in collaboration with Dr. G.M. Scotti and Dr. D. Lazarevic, Center for Omics Sciences, IRCCS San Raffaele Scientific Institute, Milan, Italy.*
6) *Validation of results by PCR (Results, Chapter 3.15, Figures 27 and 28), was performed by Dr. E. Giacomini, Reproductive Sciences Laboratory, Obstetrics and Gynecology Unit, IRCCS San Raffaele Scientific Institute*

All sources of information are acknowledged by means of reference.

A mia mamma Silvia Fornari

Abstract

Introduction. Implantation rates in Assisted Reproduction Technologies (ART) are still suboptimal due to a contribution of both embryonic and endometrial factors. While diagnosis of embryo euploidy can be achieved through preimplantation genetic testing, the diagnosis of a receptive endometrium remains a challenge. The main limitations of current receptivity tests based on the transcriptome of endometrial biopsies are that they: i) overlook the extracellular compartment, which appears to play a major role in the implantation-related cross-talk; ii) cannot be performed in the same cycle of the embryo transfer attempt.

Aim of the work. The aim of this project is to use non-invasively collected uterine-fluid derived extracellular vesicles (UF-EVs) as a novel source of transcriptomic markers of endometrial receptivity.

Methods. First, fertile volunteer women were included and their UF-EVs were retrieved for physical characterization and transcriptomic analysis, both in a pre-receptive (LH+2) and receptive (LH+7) phase of the cycle. Then, a cohort of ART patients and a further analogous validation cohort were included. UF-EVs were collected during the receptive (LH+7) phase of the cycle immediately preceding that of an euploid blastocyst transfer attempt, for comparison between patients with subsequent successful versus failed implantation.

Results. The transcriptome of UF-EVs markedly changes between the nonreceptive (LH+2) and receptive (LH+7) phase, with n=2247 differentially 'expressed' genes. The LH+7 transcriptomic content of ART patients who subsequently achieve implantation partly differs from that of patients who subsequently fail implantation, with n=161 differentially 'expressed' genes. UF-EVs of ART patients with failed versus successful implantation also show a higher mean size of EVs. After including the validation cohort, a significant ROC curve for prediction of implantation was calculated based on selected genes (AUC=0.86; 95% CI 0.78-0.94, $p=2.8x10^{-8}$). In the subgroup of patients with a diagnosis of recurrent implantation failure, a two-step cluster analysis model could correctly classify patients as successfully or unsuccessfully treated with a test sensitivity of 94.3% and a specificity of 80.0%.

Discussion. These results support the hypothesis that UF-EVs can be used to identify a transcriptomic signature of endometrial receptivity in ART patients. Current/future perspectives will be to validate these findings and to assess the safety of same-cycle, non-invasive sampling of UF-EVs prior to the embryo transfer attempt.

ACKNOWLEDGEMENTS

I would like to acknowledge all the medical and laboratory personnel, midwives residents and healthcare assistants at Centro Scienze Natalità, IRCCS Ospedale San Raffaele, for their cooperating attitude and collaboration in the procedures of this study.

TABLE OF CONTENTS

ACRONYMS AND ABBREVIATIONS

ART - Assisted Reproduction Technologies

AUC - Area Under the Curve

BMI - Body Mass Index

Cpm - count per million

DGE - Differential Gene Expression

ERA® - Endometrial Receptivity Analysis

FC – Fold Change

FDR - False Discovery Rate

G-CSF - Granulocyte-Colony Stimulating Factor

GO - gene Ontology

GSEA - Gene Set Enrichment Analysis

LH - Luteinizing Hormone

miRNA - micro RiboNucleic Acid

mRNA - messenger RNA

NES - Normalized Enrichment Score

PGT-A - Preimplantation Genetic Testing for Aneuploidies

PBMC - Peripheral Blood Mononuclear Cells

PCR - Polymerase Chain Reaction

PRP - Platelet Rich Plasma

QC - Quality control

RCT - Randomized Controlled Study

RIF - Recurrent Implantation Failure

RNA - RiboNucleic Acid

RPKM - Read Per Kilobase Million

SD - Standard Deviation

SEQC - Sequencing Quality Control Consortium

TEM -Transmission electron microscopy

UF-EVs - uterine fluid-derived extracellular vesicles

WGCA - weighted gene correlation network analysis

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Transcriptomic data related to endometrial receptivity

To allow blastocyst attachment, the endometrium has to undergo a complex structural and functional maturation which occurs during a self-limited time frame referred to as the window of implantation (Murphy, 2004). Such endometrial transformation is orchestrated by a large number of differentially expressed genes: a state of global endometrial transcriptional derepression has indeed been described to characterize the window of implantation (Sebastian-Leon et al., 2021). When a meta-analysis of studies investigating endometrial receptivity-associated transcripts was performed, 57 significant genes were identified, and referred to as a 'metasignature' transcriptome of implantation, which highlights the importance of immune responses and complement cascade in endometrial receptivity (Altmäe et al., 2017). The extracellular compartment also emerges as a crucial player in mid-secretory endometrial function, as 28 of these gene transcripts encode for proteins that were shown to be in extracellular vesicles (Altmäe et al., 2017). This observation was also confirmed by a more recent study using weighted gene correlation network analysis (WGCA): out of the n=33 genes that orchestrate the derepression characterizing the window of implantation, n=11 are located within extracellular vesicles (Sebastian-Leon et al., 2021).

1.2 Endometrial receptivity in ART and Recurrent Implantation Failure

Implantation rates in Assisted Reproduction Technologies (ART) are still suboptimal and generally below 30% (European IVF-Monitoring Consortium for the European Society of Human Reproduction and Embryology, 2021). Studies in this field estimate that the embryo is responsible for around one third of implantation failures and that the endometrial status instead accounts for the remaining two thirds (Craciunas et al., 2019). Strategies aimed at avoiding embryo transfer attempts in the presence of a non-receptive endometrium as well as strategies aimed at increasing endometrial receptivity thus remain a challenge.

In this context, while several cut-offs exist, the condition of recurrent implantation failure (RIF) is generally diagnosed based on failure of three or more transfer attempts with good quality embryos (Liu et al., 2020). Patients affected by RIF represent a debated and heterogeneous population that currently undergoes several investigations, empirical therapies and repeated embryo transfer attempts, with

inconclusive results. While embryo aneuploidy is recognized to play a major role in RIF and should not be disregarded when aiming at a more accurate diagnosis (Ata et al., 2021; Somigliana et al., 2022), the use of Preimplantation Genetic Testing for Aneuploidies (PGT-A) has failed to show any beneficial effect in terms of clinical pregnancy and live-birth rates in the meta-analysis (Busnelli et al., 2021) of available randomized controlled studies (RCTs) (Blockeel et al., 2008; Rubio et al., 2013) and observational studies (Yakin et al., 2008; Greco et al., 2014; Sato et al., 2019). Also in the two most recent observational studies, a persistent significant reduction in live-birth rates was confirmed for patients with more than two previous implantation failures despite the use of PGT-A (Cozzolino et al., 2020; Cimadomo et al., 2021). Among other commonly performed treatments for RIF, also endometrial scratching, intrauterine Granulocyte-Colony Stimulating Factor (G-CSF) infusion, LMWH administration, assisted hatching have failed to demonstrate a significant effect on clinical pregnancy and live-birth rates. In contrast, more recent therapies based on of a) subcutaneous G-CSF administration; b) intrauterine Platelet Rich Plasma (PRP) infusion; or c) Peripheral Blood Mononuclear Cells (PBMC) intrauterine infusion have showed a moderate-quality evidence for an increase on clinical pregnancy rates (a and b) and on both clinical pregnancy rates and live-birth rates (c) (Busnelli et al., 2021). Taken together, these data highlight impaired endometrial receptivity as the current limiting step in the successful management of RIF and prompt for further investigations in the field.

1.3 Endometrial Receptivity Analysis (ERA®) test

The endometrial receptivity analysis (ERA®) is currently the most commonly performed test for endometrial receptivity in women undergoing ART. For its development, the transcriptome of human endometrial biopsies was used to identify 238 genes whose expression levels characterize the window of implantation (i.e. having an absolute fold-change > 3 and a false discovery rate < 0.05) (Diaz-Gimeno et al., 2011). As a customized micro-array platform based on the 238 genes, the ERA® test describes endometrial biopsies as receptive, pre-receptive or post-receptive based on their RiboNucleic Acid (RNA) content, with an associated diagnostic probability. This result is then translated into a recommendation for 'personalized' embryo transfer in each patient on day 4, 4.5, 5, 5.5, 6 or 7 of progesterone administration. Nonetheless, the RCT published on the use of ERA®

has failed to show an advantage of personalized over standard (fresh or frozen) embryo transfer in terms of live-birth rates (Simon et al., 2020).

Based on the molecular analysis of endometrial sampling, ERA® might suffer from the limitation that - as an invasive procedure - it cannot be performed in the cycle in which the embryo transfer attempt is scheduled but only in a preceding one - potentially jeopardizing the relevance of the results. As a matter of fact, the availability of a valuable same-cycle endometrial receptivity test could become the greatest advancement for reproductive medicine in the precision molecular medicine era. As a consequence, many recent studies are also concentrating on noninvasively obtained mucosal biomarkers detected in uterine fluid and cervicovaginal mucus (Jain et al., 2022).

1.4 Uterine fluid as a 'liquid biopsy' for endometrial receptivity

For the scarce benefits provided by current diagnostics for endometrial receptivity, the uterine fluid is gaining attention as an alternative source of potential biomarkers, with the relevant advantage of potentially allowing same-cycle sampling. The components of uterine fluid have several origins: secretions from the endometrial epithelium and glands, selective transudation from blood and likely contributions from the oviductal fluid (Leese et al., 2008). Thus, it is reasonable to envisage that it should contain useful biomarkers differentially regulated between the proliferative and the secretory phases and potentially between a nonreceptive and a receptive endometrium. For the close apposition of the uterine luminal surfaces, the volume of uterine fluid is very small, and two different methods for uterine fluid collection are thus reported: i) direct aspiration; ii) flushing with around 2ml of sterile saline, with the technical drawback of the two techniques being inconsistent/scarce volume collection (1-4 µL) for the former and dilution for the latter. Uterine fluid has been used for single (Bentin-Ley et al., 2011; Florio et al., 2010; Halperin et al., 1995; Lédée-Bataiile et al., 2004) or multiple (Boomsma et al., 2009; Khadem et al., 2019; Rahiminejad et al., 2016) protein detection as well as for proteomic analysis (Casado-Vela et al, 2009) and - more recently - for microbiome analysis (Moreno et al., 2016), lipidomic analysis (Braga et al., 2019) and small RNA sequencing (von Grothusen et al., 2022).

While most of the above works performed direct uterine fluid aspiration or flushing on the same day of embryo transfer, only few studies investigated the safety of this approach in terms of subsequent implantation rates. As reported in Table, results

were reassuring across all of them but only direct uterine fluid aspiration and not flushing was performed (**Table 1**).

Table 1. Case-control studies assessing the safety of uterine fluid collection on the same day of embryo transfer, immediately before the procedure

Procedure	**Number of cases (a) and controls (b)**	**Volume of sample**	**Power calculation analysis (80%, p<0.05)**	**Implantation rates %**	**Ongoing pregnancy rates %**	**Live-Birth rates**	**Reference**
Direct UF aspiration	n=66 cases (a) n=66 controls (b)	*na*	From 30 to 10%	23.0 (a) 18.0 (b)	33 (a) 30 (b)	*na*	van der Gaast et al., 2003
Direct UF aspiration	n=210 cases (a) n=210 controls (b)	*na*	From 30 to 18%	32.0 (a) 29.0 (b)	24.3 (a) 23.8 (b)	*na*	Boomsma et al., 2009
Direct UF aspiration	n=96 (a) n=96 controls (b)	*na*	*na*	47.5(a) 39.1 (b)	50.0 (a) 39.6 (b)	47.9 (a) 37.5 (b)	Hou et al., 2022

UF: Uterine Fluid

For its intriguing role as a 'liquid biopsy' of the endometrium, the uterine fluid has also been recently investigated in the experimental setting of endometrial organoids (Simintiras et al., 2021). Endometrial organoids derived from cultured primary endometrial cells comprise both the glandular and luminal epithelium, and are polarized. For this reason, the intra- and extra- organoid uterine fluids have been separately collected and studied with both a transcriptomic and metabolomic approach, to reflect the uterine lumen and the uterine epithelial-stromal cell interface respectively (Simintiras et al., 2021). However, whether this recent experimental model will prove useful in the study of the complexity of the *in vivo* human uterine fluid composition is still to be elucidated.

1.5 Uterine fluid-derived extracellular vesicles

In addition to its soluble secretome, the uterine fluid has been found to be characterized by the presence of extracellular vesicles (Ng et al., 2013; Vilella et al., 2015; Campoy et al., 2016; Luddi et al., 2019) that have been demonstrated to be uptaken by human trophectoderm and to regulate its adhesive and invasive capacity (Greening et al., 2016; Evans et al., 2019). The identification of markers of receptivity in extracellular vesicles from uterine fluid may thus have important implications. Compared to the soluble secretome, extracellular vesicles are characterized by having a phospholipid bilayer membrane and by constituting a molecular fingerprint representative of the cell of origin. In fact, the differential origin of extracellular vesicles determines their specific cargos composed of lipids, proteins and nucleic acids (Yanez-Mo et al., 2015). The transfer of this material can regulate gene expression and affect the function of target cells, which may be activated or become differentiated or dedifferentiated according to the information received. Based on their release, extracellular vesicles are distinguished into: (a) exosomes (characterized by their small size, 30–100 nm), that are formed within the endosomal compartment and released upon fusion of multi-vesicular bodies with the plasma membrane; (b) microvesicles (100– 1000 nm), formed by external bulging of the plasma membrane; and (c) apoptotic bodies (the largest type of EVs, 50–5000 nm), that are remnants of cells undergoing apoptosis.

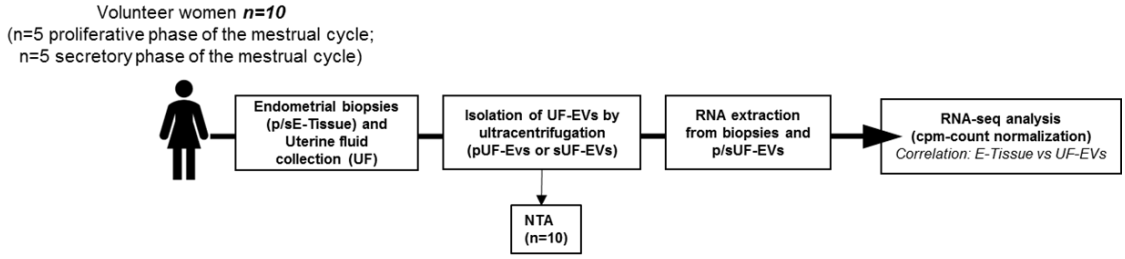
Ng and colleagues (2013) were able to identify for the first time exosomes containing microRNAs (miRNAs), proteins and messenger RNAs (mRNAs) in the uterine fluid (Ng et al., 2013). Compared to studies on cell-free nucleic acid in the uterine fluid (von Grothusen et al., 2022), a more abundant transcriptional content was reported for uterine fluid-derived extracellular vesicles (UF-EVs) (Ng et al., 2013; Li et al., 2021). In contrast, the proteomic characterization of uterine fluid-derived extracellular vesicles has only been achieved recently, due to technical difficulties, and a smaller amount of protein content in UF-EVs than in the soluble uterine fluid was described (Rai et al., 2021). Thus, the transcriptomic approach currently appears as the most promising for exploitation of UF-EVs in the study of molecular mechanisms and biomarkers of implantation. In line, a recent pilot study has also proposed the use of small RNAs contained in UF-EVs as a marker for the early diagnosis of ovarian cancer (Skryabin et al., 2022).

1.6 Preliminary results of the project

Declaration of the author: please note that the preliminary results of the project have already been published in a recent article of which I am a co-author (Giacomini et al., 2021).

As a preliminary finding to this study (Giacomini et al., 2021), we sought to demonstrate that transcriptomic changes occur in UF-EVs and closely resemble those of the corresponding endometrial tissue. To this aim, we included a group of n=10 regularly cycling women [with an age of 31.4 ± 4.8 , a Body Mass Index (BMI) of 21.3 ± 3.9 and a menstrual cycle duration of 29.1 ± 2.7 , Mean ± Standard Deviation (SD)] who underwent paired endometrial biopsy and uterine fluid sampling either in the proliferative (n=5) or in the secretory (n=5) phase (**Figure 1**).

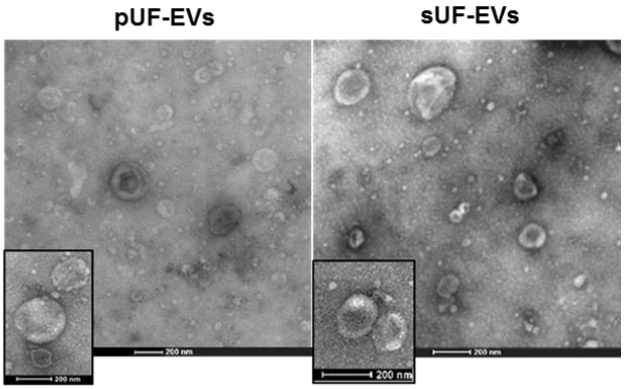
Figure 4. Schematic representation of the preliminary phase of the study



CPM: count per million; E-:endometrial; NTA: Nanoparticle Tracking Analysis; p: proliferative phase; s: secretory phase. Transcriptomes of paired endometrial tissue and UF-EVs samples were studied.

Transmission electron microscopy (TEM) confirmed detection of extracellular vesicles in samples collected from both proliferative and secretory phase, with diameters ranging from less than 100 nm to bigger than 200 nm (**Figure 2**).

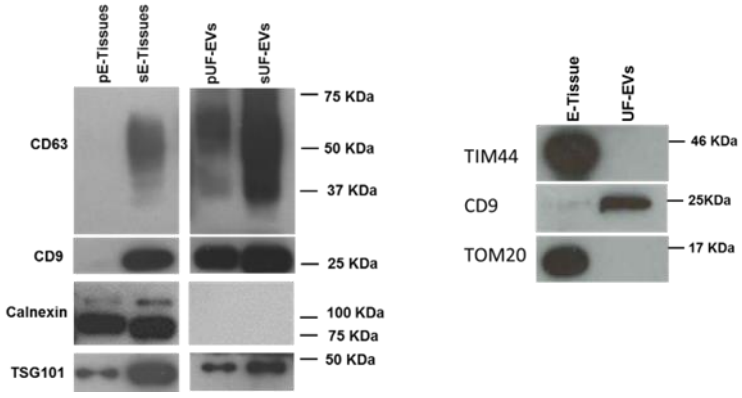
Figure 2. TEM images of pUF-EVs (left panel) and sUF-EV samples (right panel).



p: proliferative phase; s: secretory phase. Scale bar column: 200 nm TEM confirmed detection of EV-like particles in samples collected from both the proliferative and secretory phase, with diameters ranging from less than 100 nm to bigger than 200 nm.

Presence of commonly used extracellular vesicle positive protein markers TSG101, CD63 and CD9 was confirmed, while western blot of negative extracellular vesicle markers using endometrial tissue samples as controls (**Figure 3**) confirmed the lack of calnexin (CNX), TOM20 and TIM44 expression in EVs, providing proof for the absence cellular debris by exclusion of endoplasmic reticulum (CNX) and mitochondrial (TOM 20, TIM44) contamination respectively.

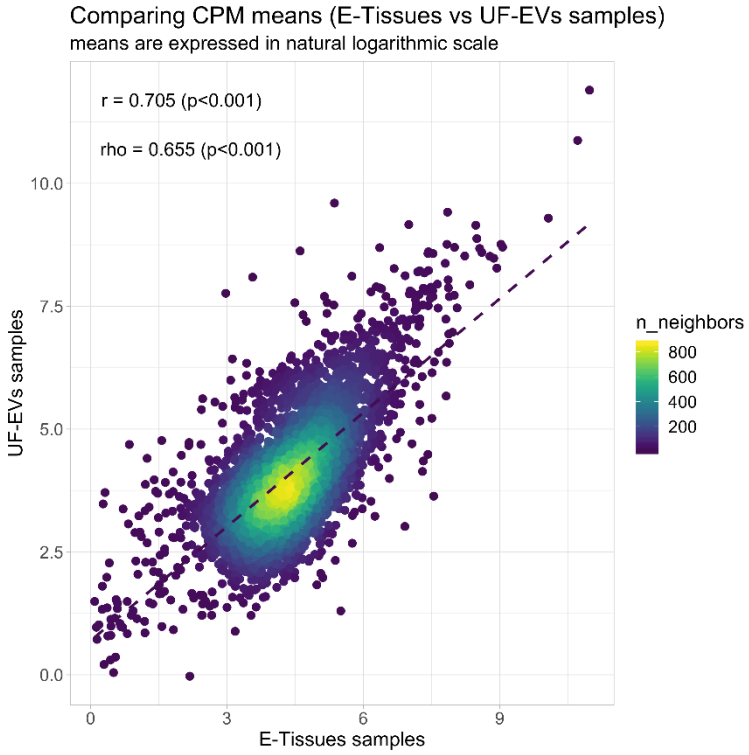
Figure 3. Western blots for extracellular vesicle positive and negative markers



E-:endometrial; p: proliferative phase; s: secretory phase. The presence of different canonical extracellular vesicles' markers (CD63, CD9, TSG101) as well as the absence of negative extracellular vesicles' markers (Calnexin, TIM44, TOM20) are shown in both pUF-EVs and sUF-EVs.

Results of the transcriptomic analysis showed a generally lower gene 'expression' in UF-EVs than in tissues, but the transcriptomic pattern was strongly conserved between the endometrial tissue and the corresponding UF-EVs (Pearson's $r = 0.70$, $P < 0.0001$; Spearman's $\rho = 0.65$, $P < 0.0001$) (**Figure 4**) and the very significant correlation between the transcriptome of endometrial biopsies and of corresponding UF-EVs was confirmed both in the proliferative ($r = 0.64$ $P < 0.001$; $\rho = 0.60$ $P < 0.00001$) and in the secretory ($r = 0.64$ $P < 0.001$; $\rho = 0.61$ $P < 0.00001$) phase.

Figure 4. Correlation analysis of gene transcripts detected in endometrial tissue and paired UF-EVs samples

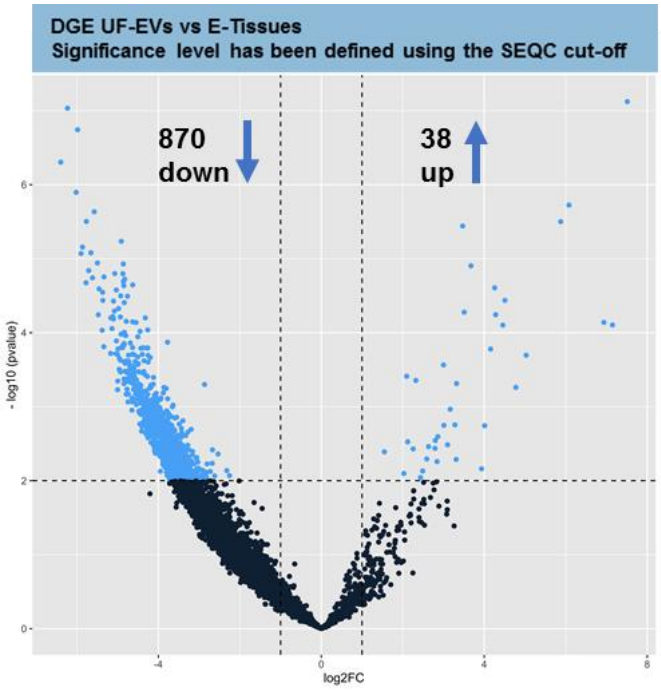


CPM: count per million; E-:endometrial; p: proliferative phase; s: secretory phaseEach dot represents a gene, and values on x and y axes are the average gene expression values in tissue and UF-EVs samples, respectively. Colour code represents the number of overlapping dots in a specific position of the graph. This number is quantified by the n-neighbors statistic that is the number of dots in the neighbourhood of the considered dot.

Differential Gene Expression (DGE) analysis between endometrial biopsies and corresponding UF-EVs showed n=908 genes to be differentially 'expressed', with the vast majority (n=870) being - as expected - over represented in tissues compared to UF-EVs and only n=38 being more abundant in UF-EVs compared to biopsies,

suggesting selective RNA packaging in endometrium-derived extracellular vesicles (**Figure 5**).

Figure 5. Volcano plot of significantly differentially detected RNAs in UF-EVs and paired endometrial tissue samples



DGE: Differential Gene Expression; FC: Fold Change; UF: Uterine Fluid; E: Endometrial; SEQC: Sequencing Quality Control Consortium.

The transcripts that are enriched in UF-EVs compared to paired endometrial biopsies are reported in **Table 2**.

Table 2. Genes that are more abundant in UF-EVs compared to endometrial biopsies.

GeneID	log2Fold Change	pvalue	padj	description
HNRNPLP1	7,51	7,5E-08	1,9E-04	heterogeneous nuclear ribonucleoprotein L pseudogene 1 [HGNC:48748]
RP11-197P3.1	6,08	1,9E-06	1,3E-03	
IGF2	5,88	3,2E-06	1,4E-03	insulin like growth factor 2 [HGNC:5466]
ANP32B	3,47	3,6E-06	1,5E-03	acidic nuclear phosphoprotein 32 family member B [HGNC:16677]

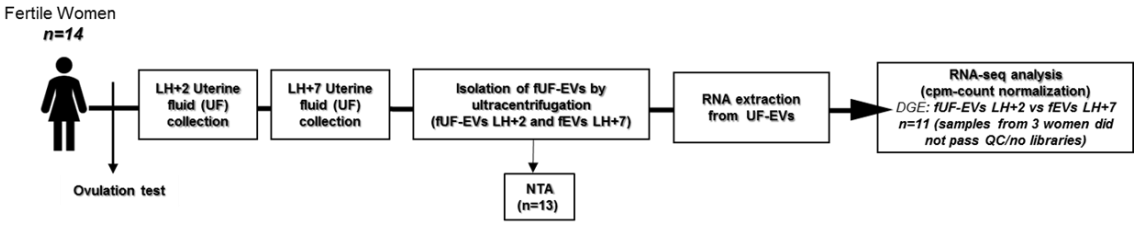
RAB13	3,67	1,2E-05	3,0E-03	RAB13, member RAS oncogene family [HGNC:9762]
HSP90AA5P	4,26	2,5E-05	3,6E-03	heat shock protein 90 alpha family class A member 5, pseudogene [HGNC:32535]
RPL23AP77	4,51	3,7E-05	4,3E-03	ribosomal protein L23a pseudogene 77 [HGNC:36354]
SAMHD1	3,51	5,3E-05	5,5E-03	SAM and HD domain containing deoxynucleoside triphosphate triphosphohydrolase 1 [HGNC:15925]
RP11-274H24.1	4,28	5,7E-05	5,6E-03	
CTD-2017F17.2	6,94	7,3E-05	5,9E-03	
PLEKHA4	4,46	7,9E-05	6,2E-03	pleckstrin homology domain containing A4 [HGNC:14339]
RP11-214O1.3	7,15	7,9E-05	6,2E-03	
RP11-130F10.1	4,16	1,7E-04	8,8E-03	
CD160	5,03	2,0E-04	9,6E-03	CD160 molecule [HGNC:17013]
TCEAL4	3,00	2,7E-04	1,1E-02	transcription elongation factor A like 4 [HGNC:26121]
RP11-941H19.1	3,32	4,9E-04	1,4E-02	
PTMA	2,10	3,9E-04	1,3E-02	prothymosin alpha [HGNC:9623]
RPL26	2,32	4,4E-04	1,4E-02	ribosomal protein L26 [HGNC:10327]
RP11-113I24.1	4,78	5,5E-04	1,5E-02	
RP11-167P22.4	3,17	1,1E-03	1,9E-02	
RP3-417G15.1	3,01	1,8E-03	2,1E-02	
SOX2-OT	4,01	1,8E-03	2,1E-02	SOX2 overlapping transcript [HGNC:20209]

RP11-44K6.3	3,28	1,8E-03	2,1E-02	
RPLP2P1	2,86	2,5E-03	2,4E-02	ribosomal protein lateral stalk subunit P2 pseudogene 1 [HGNC:18724]
RPL23AP65	2,80	2,9E-03	2,5E-02	ribosomal protein L23a pseudogene 65 [HGNC:36289]
RP11-417N10.5	3,10	3,3E-03	2,6E-02	
CTB-47B8.1	2,63	3,5E-03	2,7E-02	
RP11-159H22.1	2,79	3,7E-03	2,8E-02	
RBM25	2,26	3,7E-03	2,8E-02	RNA binding motif protein 25 [HGNC:23244]
RPS29P16	3,31	5,2E-03	3,3E-02	ribosomal protein S29 pseudogene 16 [HGNC:35560]
RP11-253E3.1	2,59	5,1E-03	3,3E-02	
KAT6B	2,84	5,5E-03	3,4E-02	lysine acetyltransferase 6B [HGNC:17582]
RPS24	1,55	4,1E-03	2,9E-02	ribosomal protein S24 [HGNC:10411]
RP11-51L5.2	2,49	7,4E-03	3,9E-02	
PTMS	3,93	6,9E-03	3,7E-02	parathymosin [HGNC:9629]
MT-RNR2	2,12	3,0E-03	2,6E-02	mitochondrially encoded 16S rRNA [HGNC:7471]
NSA2	2,03	8,0E-03	4,0E-02	NSA2 ribosome biogenesis factor [HGNC:30728]
RPS28P7	2,43	9,1E-03	4,3E-02	ribosomal protein S28 pseudogene 7 [HGNC:35787]

2. AIM OF THE WORK

Aim 1. To explore the physiological UF-EVs of the receptive versus nonreceptive phase. Uterine fluid samples for extraction and analysis of UF-EVs were collected from n=14 women with proven fertility both on day Luteinizing Hormone (LH) +2 (nonreceptive) and on day LH+7 (receptive phase) (**Figure 6**).

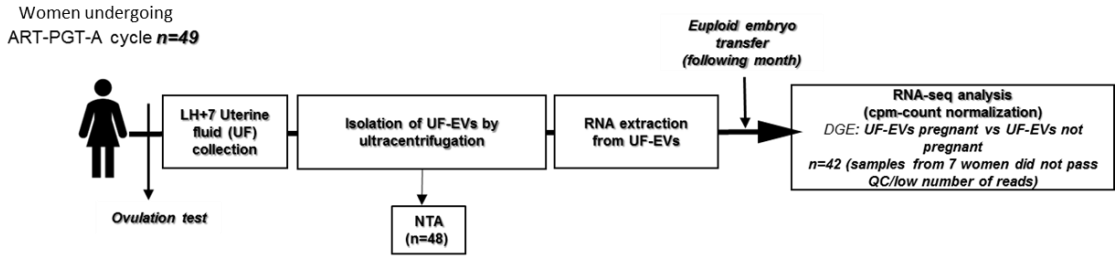
Figure 6. Schematic representation of the first phase of the study.



CPM: count per million; QC: Quality control; NTA: Nanoparticle Tracking Analysis Transcriptomes of LH+2 and LH+7 UF-EVs samples were studied.

Aim 2. To investigate whether UF-EVs content can predict endometrial receptivity in n=49 ART patients undergoing their first transfer with PGT-A (**Figure 7**). This category of women has been chosen since, transferring an euploid blastocyst, their implantation potential estimated at ≈ 50% is more likely dependent on endometrial characteristics. In this population, extracellular vesicles were purified from LH+7 uterine fluid samples.

Figure 7. Schematic representation of the second phase of the study.



CPM: count per million; QC: Quality control; NTA: Nanoparticle Tracking Analysis. Transcriptomes of UF-EVs of women with successful versus failed implantation following ART with PGT-A were studied.

Aim 3. To validate results from Aim 2 in a new cohort of n=52 patients undergoing ART and PGT-A, selected with the same criteria as in Aim 2 and in whom UF-EVs were analogously collected on day LH+7.

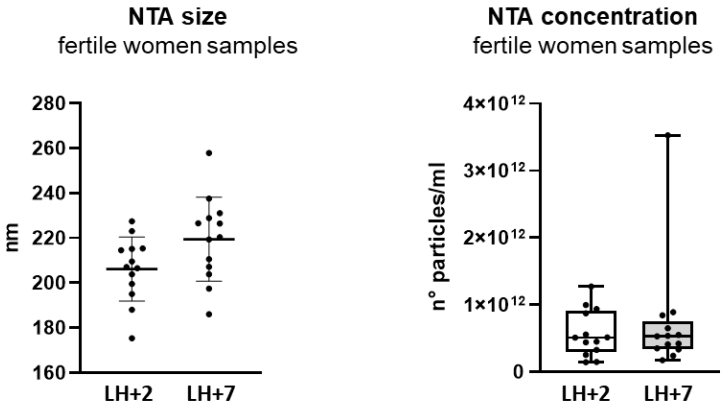
3. RESULTS

Declaration of the author: please note that results referring to Aim 1 and Aim 2 of the project have already been published in a recent article of which I am a co-author (Giacomini et al., 2021).

3.1 Physical characterization of UF-EVs collected during the window of implantation

The size of UF-EVs collected from LH + 7 samples (i.e. during the window of implantation) was not statistically different compared to UF-EVs of LH + 2 (i.e. pre-receptive) samples, with a diameter of 219.5 ± 18.71 nm (Mean ± SD) and of 206.21 ± 14.26 nm (Mean ± SD) in the two groups respectively. Particle concentrations were also similar in the two groups, with 5.28E + 11 particles/ml; IQR= 3.41E + 11–7.45E + 11 particles/ml in LH+7 samples and 5.09E + 11 particles/ml; IQR = 2.88E + 11–9.04E + 11 particles/ml in LH+2 samples (**Figure 8**).

Figure 8. Particle size distributions and concentration of UF-EVs of fertile women

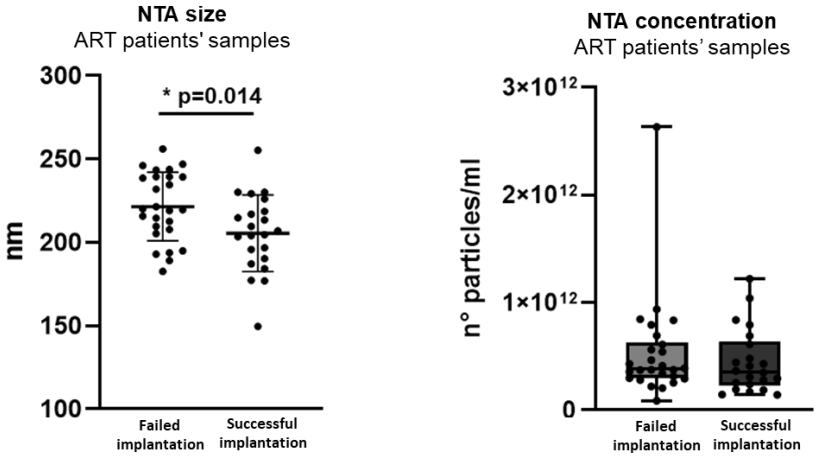


	Mode (IQR) nm	D10 (IQR) nm	D50 (IQR) nm	D90 (IQR) nm
UF-EVs LH+2	154.9 (149.7-163.8)	120.0 (99.8-126.6)	184.7 (178.4-197.3)	322.8 (303.9-340.8)
UF-EVs LH+7	176.2 (155.3-185.7)	114.8 (101.4-131.9)	202.4 (101.4-131,9)	336.6 (312.9-371.4)

3.2 Physical characterization of UF-EVs of patients undergoing ART

We observed a significant difference in the size of UF-EVs of women with successful implantation compared to women with failed implantation (205.5 ± 22.97 nm versus 221.5 ± 20.57 nm, Mean diameter ± SD, $P = 0.014$), whereas particle concentrations were comparable between the two groups [3.58E + 11 particles/ml (IQR= 2.29E + 11–6.33E + 11) and 3.84E + 11 particles/ml (IQR = 2.97E + 11–6.31E + 11) particles/ml, respectively] (**Figure 9**).

Figure 9. Particle size distributions and concentration of UF-EVs of ART patients

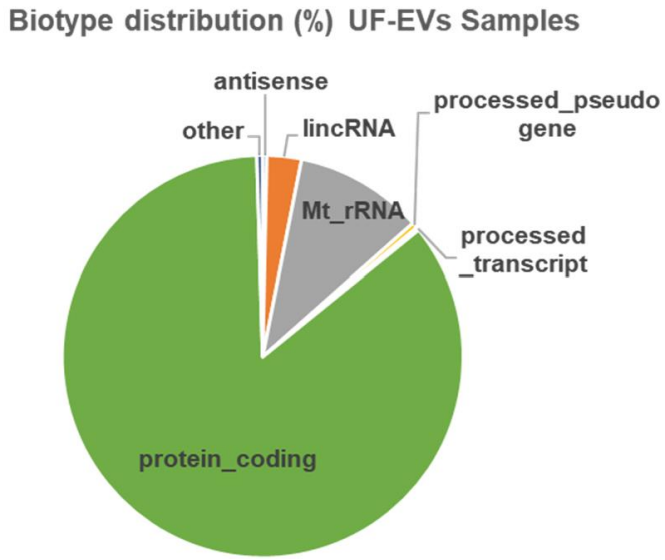


	Mode (IQR) nm	D10 (IQR) nm	D50 (IQR) nm	D90 (IQR) nm
UF-EVs failed implantation	164.3 (148.8-184.9)	121.3 (114.1-131.5)	198.7 (177.8-217.1)	351.4 (322.5-388.7)
UF-EVs successful implantation	159.9 (139.1-166.6)	113.3 (104.9-126.9)	188.9 (169.9-204.1)	325.5 (295.5-352.4)

3.3 UF-EVs transcriptomic library composition

The majority of transcripts contained in UF-EVs were from protein-coding genes (76.96%), while ribosomal mitochondrial RNAs (Mt-rRNAs) represented 15.36% of transcripts. A fraction of long non-coding RNAs (lncRNAs) was present (5.93%), whereas antisense and processed pseudogenes categories accounted for <1% (**Figure 10**).

Figure 10. UF-EVs transcriptomic library composition

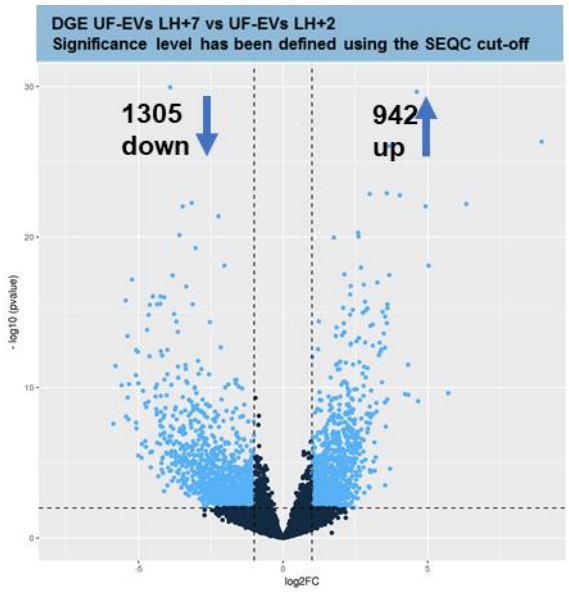


Pie charts showing the average fraction of each RNA biotype. Biotypes were defined according to Gencode annotations.

3.4 Transcriptomic profile of UF-EVs collected before and during the window of implantation

DGE analysis was performed on paired LH+7 vs LH+2 UF-EV samples from the fertile women included, without adjusting for any batch effect because paired samples from each woman were sequenced in the same batch. Of a total of 14228 'expressed' genes (i.e. genes with ≥ 1 CPM in at least n=11 saples), 2247 were differentially 'expressed' during the window of implantation: more specifically 942 gene transcripts were more abundant and 1305 less abundant during the window of implantation (day LH+7) compared to before (day LH+2) (**Figure 11**).

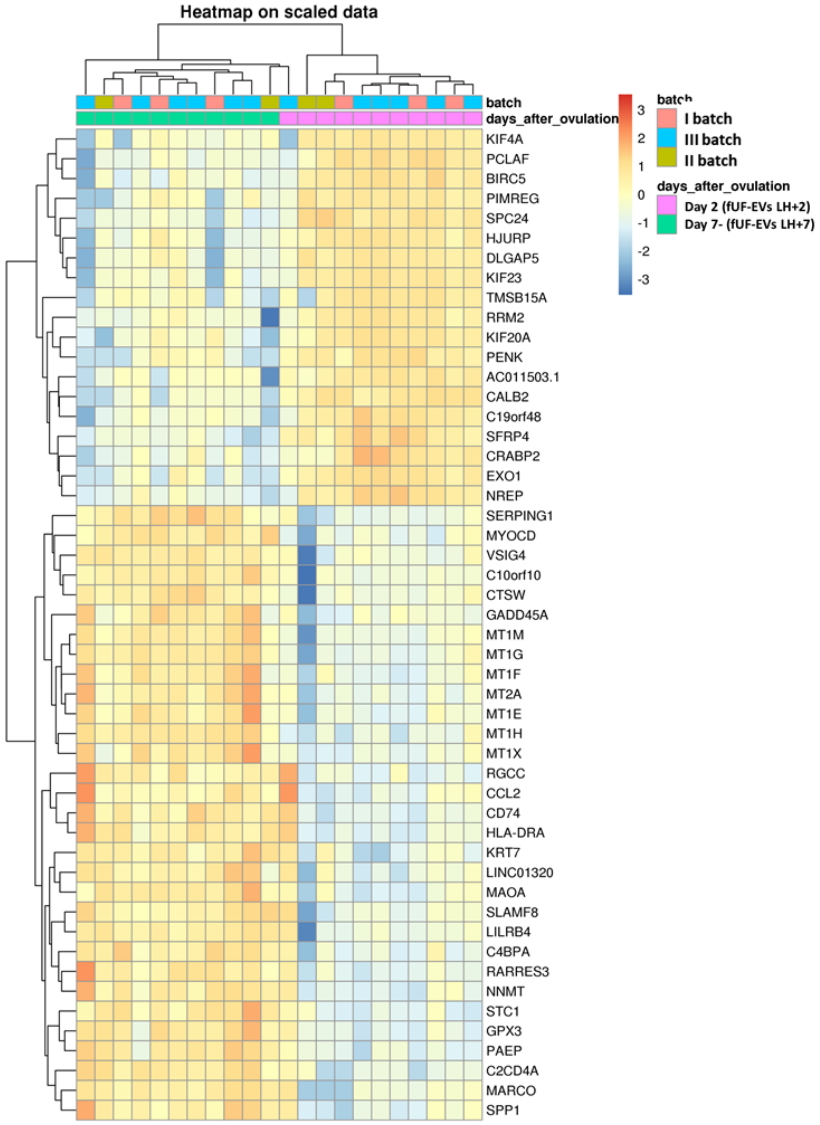
Figure 11. Volcano plot of significantly differentially detected RNAs in UF-EVs of the receptive (LH+7) versus the pre-receptive phase (LH+2)



DGE: Differential Gene Expression; FC: Fold Change; UF: Uterine Fluid; E: Endometrial; SEQC: Sequencing Quality Control Consortium.

The heatmap for the 'expression' of the top 50 most significant genes in the comparison shows a clear distinction between samples obtained before and during the window of implantation and is reported below (**Figure 12**), whereas the relative gene list is reported in **Table 3**.

Figure 12. Heatmap for the 'expression' of the top 50 most significant genes in the comparison of UF-EVs collected in the receptive (LH+7) versus the pre-receptive phase (LH+2)



f-: fertile volunteer. Expression values were scaled so that their sum is the same for each gene.

The top significant genes in the DGE analysis comparing the transcriptional profile of UF-EVs collected in the receptive (LH+7) versus the pre-receptive phase (LH+2) are shown in **Table 3**.

Table 3. DGE of the top statistically significant 50 genes with a difference between UF-EVs of fertile women in the receptive (LH+7) compared to the pre-receptive phase (LH+2)

GeneID	log2FoldChange	pvalue	padj	description
MT1G	6,76	2,61E-55	3,72E-51	metallothionein 1G [HGNC:7399]
PAEP	6,11	7,40E-48	5,27E-44	progestagen associated endometrial protein [HGNC:8573]
CALB2	-7,62	2,65E-37	1,26E-33	calbindin 2 [HGNC:1435]
MT1M	4,47	8,49E-35	3,02E-31	metallothionein 1M [HGNC:14296]
SPP1	3,91	2,96E-34	8,41E-31	secreted phosphoprotein 1 [HGNC:11255]
SFRP4	-3,90	1,13E-30	2,67E-27	secreted frizzled related protein 4 [HGNC:10778]
MYOCD	4,62	2,20E-30	4,47E-27	myocardin [HGNC:16067]
MT1H	8,94	4,65E-27	8,28E-24	metallothionein 1H [HGNC:7400]
NNMT	3,65	8,95E-27	1,41E-23	nicotinamide N-methyltransferase [HGNC:7861]
C10orf10	3,59	1,23E-23	1,75E-20	chromosome 10 open reading frame 10 [HGNC:23355]
MT1E	3,00	1,38E-23	1,78E-20	metallothionein 1E [HGNC:7397]
GPX3	4,04	1,68E-23	1,99E-20	glutathione peroxidase 3 [HGNC:4555]
PCLAF	-3,16	5,39E-23	5,90E-20	PCNA clamp associated factor [HGNC:28961]
C2CD4A	6,33	6,36E-23	6,47E-20	C2 calcium dependent domain containing 4A [HGNC:33627]
MARCO	4,93	9,15E-23	8,39E-20	macrophage receptor with collagenous structure [HGNC:6895]
BIRC5	-3,47	9,43E-23	8,39E-20	baculoviral IAP repeat containing 5 [HGNC:593]
CRABP2	-2,23	4,14E-22	3,46E-19	cellular retinoic acid binding protein 2 [HGNC:2339]
MT1X	2,59	5,09E-21	4,02E-18	metallothionein 1X [HGNC:7405]
SPC24	-3,59	7,41E-21	5,55E-18	SPC24, NDC80 kinetochore complex component [HGNC:26913]
MT2A	2,61	9,10E-21	6,47E-18	metallothionein 2A [HGNC:7406]
RARRES3	1,76	1,09E-20	7,40E-18	retinoic acid receptor responder 3 [HGNC:9869]

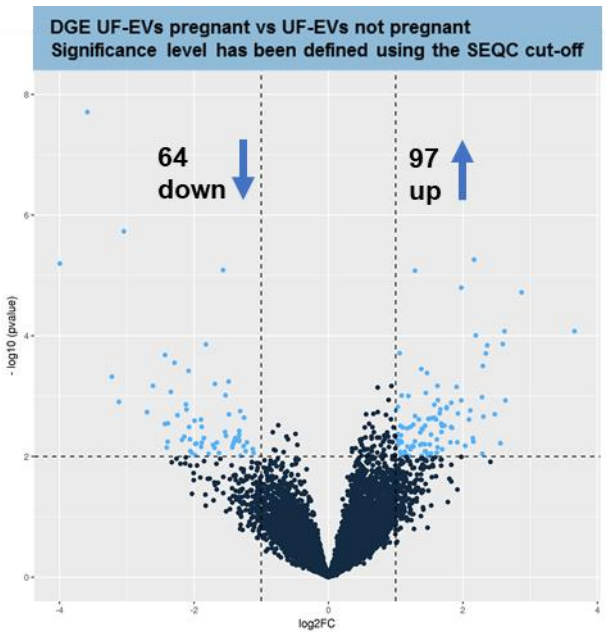
NREP	-3,03	5,38E-20	3,48E-17	neuronal regeneration related protein [HGNC:16834]
C19orf48	-2,03	7,86E-19	4,70E-16	chromosome 19 open reading frame 48 [HGNC:29667]
C4BPA	5,03	7,93E-19	4,70E-16	complement component 4 binding protein alpha [HGNC:1325]
SERPING1	2,69	1,07E-18	6,11E-16	serpin family G member 1 [HGNC:1228]
CD74	2,12	3,02E-18	1,65E-15	CD74 molecule [HGNC:1697]
LILRB4	3,67	3,35E-18	1,77E-15	leukocyte immunoglobulin like receptor B4 [HGNC:6608]
KIF4A	-3,82	3,55E-18	1,80E-15	kinesin family member 4A [HGNC:13339]
PENK	-5,22	6,81E-18	3,34E-15	proenkephalin [HGNC:8831]
MT1F	3,26	1,04E-17	4,92E-15	metallothionein 1F [HGNC:7398]
LINC01320	2,79	1,48E-17	6,79E-15	long intergenic non-protein coding RNA 1320 [HGNC:50526]
HLA-DRA	2,34	1,80E-17	8,02E-15	major histocompatibility complex, class II, DR alpha [HGNC:4947]
AC011503.1	-3,35	1,94E-17	8,35E-15	
RGCC	2,32	6,62E-17	2,77E-14	regulator of cell cycle [HGNC:20369]
KIF20A	-4,50	8,76E-17	3,51E-14	kinesin family member 20A [HGNC:9787]
KIF23	-4,24	8,89E-17	3,51E-14	kinesin family member 23 [HGNC:6392]
DLGAP5	-4,11	1,02E-16	3,94E-14	DLG associated protein 5 [HGNC:16864]
EXO1	-5,44	1,64E-16	6,13E-14	exonuclease 1 [HGNC:3511]
STC1	2,83	2,03E-16	7,39E-14	stanniocalcin 1 [HGNC:11373]
HJURP	-4,24	2,74E-16	9,74E-14	Holliday junction recognition protein [HGNC:25444]
RRM2	-3,12	2,92E-16	9,87E-14	ribonucleotide reductase regulatory subunit M2 [HGNC:10452]
PIMREG	-4,35	2,92E-16	9,87E-14	PICALM interacting mitotic regulator [HGNC:25483]
CTSW	3,61	2,98E-16	9,87E-14	cathepsin W [HGNC:2546]
TMSB15A	-4,61	3,22E-16	1,04E-13	thymosin beta 15a [HGNC:30744]
GADD45A	2,08	5,09E-16	1,61E-13	growth arrest and DNA damage inducible alpha [HGNC:4095]
SLAMF8	3,61	5,48E-16	1,68E-13	SLAM family member 8 [HGNC:21391]

MAOA	2,92	5,53E-16	1,68E-13	monoamine oxidase A [HGNC:6833]
CCL2	2,41	7,01E-16	2,08E-13	C-C motif chemokine ligand 2 [HGNC:10618]
KRT7	2,85	7,77E-16	2,26E-13	keratin 7 [HGNC:6445]
VSIG4	3,44	9,40E-16	2,67E-13	V-set and immunoglobulin domain containing 4 [HGNC:17032]

3.5 Transcriptomic profile of UF-EVs collected from patients with successful versus failed implantation

DGE was performed accounting for the sequencing batch as a covariate. Of a total of 14593 'expressed' genes (i.e. genes with ≥ 1 CPM in at least n=19 saples), 161 were differentially 'expressed' between the two groups. More specifically, 97 transcripts were more abundant whereas 64 transcripts were less abundant in patients with successful outcome (**Figure 13**).

Figure 13. Volcano plot of the significantly differentially detected RNAs in UF-EVs of ART patients with successful versus failed implantation



DGE: Differential Gene Expression; FC: Fold Change; UF: Uterine Fluid; E: Endometrial; SEQC: Sequencing Quality Control Consortium

The top significant genes in the DGE analysis comparing the transcriptional profile of UF-EVs derived from patients with successful versus failed implantation are shown in **Table 4**.

Table 4. DGE of the top statistically significant 50 genes with a difference between LH+7 UF-EVs derived from patients with successful versus failed implantation

GeneID	log2 FoldChange	pvalue	padj	description
MTRNR2L1	-3,59	1,97E-08	2,87E-04	MT-RNR2-like 1 [HGNC:37155]
DUSP1	-3,04	1,86E-06	1,36E-02	dual specificity phosphatase 1 [HGNC:3064]
CLU	2,17	5,48E-06	2,03E-02	clusterin [HGNC:2095]
PITX1	-4,00	6,35E-06	2,03E-02	paired like homeodomain 1 [HGNC:9004]
AL391834.2	-1,57	8,15E-06	2,03E-02	
IFI6	1,29	8,35E-06	2,03E-02	interferon alpha inducible protein 6 [HGNC:4054]
RPL10P9	1,98	1,60E-05	3,33E-02	ribosomal protein L10 pseudogene 9 [HGNC:35579]
KNL1	2,87	1,91E-05	3,48E-02	kinetochore scaffold 1 [HGNC:24054]
AOC1	3,66	8,38E-05	1,23E-01	amine oxidase, copper containing 1 [HGNC:80]
PPP1R26-AS1	2,62	8,43E-05	1,23E-01	PPP1R26 antisense RNA 1 [HGNC:48717]
AC006978.1	2,19	9,85E-05	1,31E-01	
ANKRD22	2,59	1,37E-04	1,50E-01	ankyrin repeat domain 22 [HGNC:28321]
ABHD3	-1,82	1,39E-04	1,50E-01	abhydrolase domain containing 3 [HGNC:18718]
LINC02486	2,36	1,44E-04	1,50E-01	long intergenic non-protein coding RNA 2486 [HGNC:25288]
OAS1	1,06	1,94E-04	1,79E-01	2'-5'-oligoadenylate synthetase 1 [HGNC:8086]
PIK3R6	2,34	1,96E-04	1,79E-01	phosphoinositide-3-kinase regulatory subunit 6 [HGNC:27101]

FAM83D	-2,43	2,09E-04	1,79E-01	family with sequence similarity 83 member D [HGNC:16122]
RGS2	-2,29	2,80E-04	2,27E-01	regulator of G protein signaling 2 [HGNC:9998]
PIK3C2B	2,30	3,17E-04	2,44E-01	phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 beta [HGNC:8972]
LNPK	1,38	3,56E-04	2,60E-01	lunapark, ER junction formation factor [HGNC:21610]
LRMP	-2,08	3,81E-04	2,65E-01	lymphoid restricted membrane protein [HGNC:6690]
CCL2	1,46	4,15E-04	2,76E-01	C-C motif chemokine ligand 2 [HGNC:10618]
KRT4	-3,22	4,75E-04	3,02E-01	keratin 4 [HGNC:6441]
ZNF841	-1,49	5,74E-04	3,49E-01	zinc finger protein 841 [HGNC:27611]
EVPL	-1,69	6,29E-04	3,50E-01	envoplakin [HGNC:3503]
FLG	-2,61	6,74E-04	3,50E-01	filaggrin [HGNC:3748]
LINC01637	1,62	6,75E-04	3,50E-01	long intergenic non-protein coding RNA 1637 [HGNC:52424]
CCNF	1,91	7,05E-04	3,50E-01	cyclin F [HGNC:1591]
KRT6A	-2,34	8,52E-04	4,01E-01	keratin 6A [HGNC:6443]
RTN2	1,45	8,86E-04	4,04E-01	reticulon 2 [HGNC:10468]
PPL	-1,53	9,68E-04	4,15E-01	periplakin [HGNC:9273]
LRRC45	1,09	9,79E-04	4,15E-01	leucine rich repeat containing 45 [HGNC:28302]
HSD17B1	1,19	9,96E-04	4,15E-01	hydroxysteroid 17-beta dehydrogenase 1 [HGNC:5210]
ZNF385C	2,28	1,04E-03	4,22E-01	zinc finger protein 385C [HGNC:33722]
AC011477.2	1,63	1,17E-03	4,31E-01	
LINC00621	2,63	1,18E-03	4,31E-01	long intergenic non-protein coding RNA 621 [HGNC:44227]
NUPR2	1,97	1,18E-03	4,31E-01	nuclear protein 2, transcriptional regulator [HGNC:44164]

SPINK7	-3,12	1,24E-03	4,43E-01	serine peptidase inhibitor, Kazal type 7 (putative) [HGNC:24643]
RAB38	1,83	1,29E-03	4,48E-01	RAB38, member RAS oncogene family [HGNC:9776]
AC078883.1	-2,12	1,37E-03	4,61E-01	
USP27X	1,57	1,39E-03	4,61E-01	ubiquitin specific peptidase 27, X-linked [HGNC:13486]
CILP	1,03	1,51E-03	4,84E-01	cartilage intermediate layer protein [HGNC:1980]
DNAJB7	1,76	1,53E-03	4,84E-01	DnaJ heat shock protein family (Hsp40) member B7 [HGNC:24986]
NCAM1	1,76	1,59E-03	4,84E-01	neural cell adhesion molecule 1 [HGNC:7656]
SPN	1,66	1,66E-03	4,84E-01	sialophorin [HGNC:11249]
CRYBG2	-2,11	1,67E-03	4,84E-01	crystallin beta-gamma domain containing 2 [HGNC:17295]
RASGRP1	2,11	1,72E-03	4,84E-01	RAS guanyl releasing protein 1 [HGNC:9878]
ATF5	1,00	1,74E-03	4,84E-01	activating transcription factor 5 [HGNC:790]
ZNF845	-1,31	1,77E-03	4,84E-01	zinc finger protein 845 [HGNC:25112]
CRNN	-2,70	1,84E-03	4,84E-01	cornulin [HGNC:1230]

We also searched for genes only expressed by either by women with a successful versus a negative outcome, by defining them as 'selectively expressed' in LH + 7 UF-EV samples from women who achieved pregnancy if they had ≥1 CPM in at least more than half of these samples (10/19) and ≥1 CPM in less than one-fifth of the LH + 7 UF-EV samples from women not achieving pregnancy (4/23 samples). This identified 14 transcripts selectively detected in UF-EVs of women with a successful implantation: *AC114491.1*, *AC008608.2*, *PMS2P5*, *C10orf99*, *NPTN-IT1*, *AC012358.3*, *ANKRD18A*, *GLIS2-AS1*, *AC011447.7*, *AL009174.1*, *C1QTNF2*, *TMED6*, *AC016355.1* and *AL021392.1* (Figure). Similarly, genes were defined as 'selectively expressed' in LH + 7 UF-EV

samples of women with a negative outcome if they had ≥1 CPM in at least n = 10 out of n = 23 of these samples and ≥1 CPM in less than one-fifth of the LH + 7 UF-EV samples from women who achieved pregnancy (3/19 samples). The five genes detected only in UF-EVs of patients who failed implantation were *CD200R1*, *FAM66B*, *AL391834.1*, *WNT9B* and *CEC*R7 (**Figure 14**). Figure lower panel reports levels of the genes with the most statistically significant 'selective expression' in UF-EV samples collected from one of the two groups (*PMS2P5*, *ANKRD18A*, *GLIS2-AS1*, *NPTN-IT1* and *CERC7*).

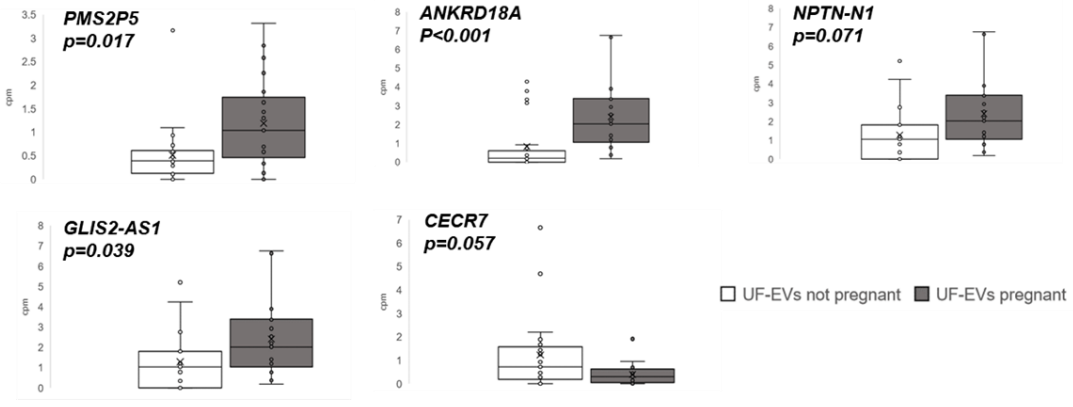
Figure 14. Transcripts 'selectively expressed' in UF-EVs from women with successful versus failed implantation

A

	GENE ID	Description	p value *	n° UF-EVs not pregnant	n° UF-EVs pregnant
UF-EVs pregnant	*AC114491.1*	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 4, 9kDa (NDUFA4) pseudogene	0,197	3	10
	AC008608.2	novel transcript (Lnc-RNA)	0,027	3	10
	PMS2P5	PMS1 Homolog 2, Mismatch Repair System Component	0,017	2	10
	C10orf99	Chromosome 10 Open Reading Frame 99	0,012	4	10
	NPTN-IT1	NPTN Intronic Transcript 1 (Lnc-RNA)	0,071	4	10
	AC012358.3	Novel Transcript, Antisense To CCDC88A (Lnc-RNA)	0,01	4	10
	ANKRD18A	Ankyrin Repeat Domain 18A	< .001	4	14
	GLIS2-AS1	GLIS2 antisense RNA 1 (Lnc-RNA)	0,039	1	10
	AC011447.7	Lnc-RNA (TEC)	0,122	4	10
	AL009174.1	Ribosomal Protein S24 (RPS24) Pseudogene	0,036	4	11
	C1QTNF2	Complement C1q Tumor Necrosis Factor-Related Protein 2	0,08	4	10
	TMED6	Transmembrane P24 Trafficking Protein 6	0,234	4	10
	AC016355.1	Lnc-RNA	0,002	1	10
	AL021392.1	Novel Transcript (Lnc-RNA)	0,102	3	10
UF-EVs not pregnant	*CD200R1*	Cell Surface Glycoprotein CD200 Receptor 1	0,113	12	2
	FAM66B	Family With Sequence Similarity 66 Member B (Lnc-RNA)	0,275	10	2
	AL391834.1	novel transcript, antisense to DENND4C (Lnc-RNA)	0,012	11	2
	WNT9B	Wnt Family Member 9B	0,25	10	2
	CECR7	Cat Eye Syndrome Chromosome Region, Candidate 7 ((Lnc-RNA)	0,057	10	1

Note: The frequency of dection of these transcripts is statistically significant by Chi-square test (p<0,05). *Mann-Whitney U test.

B

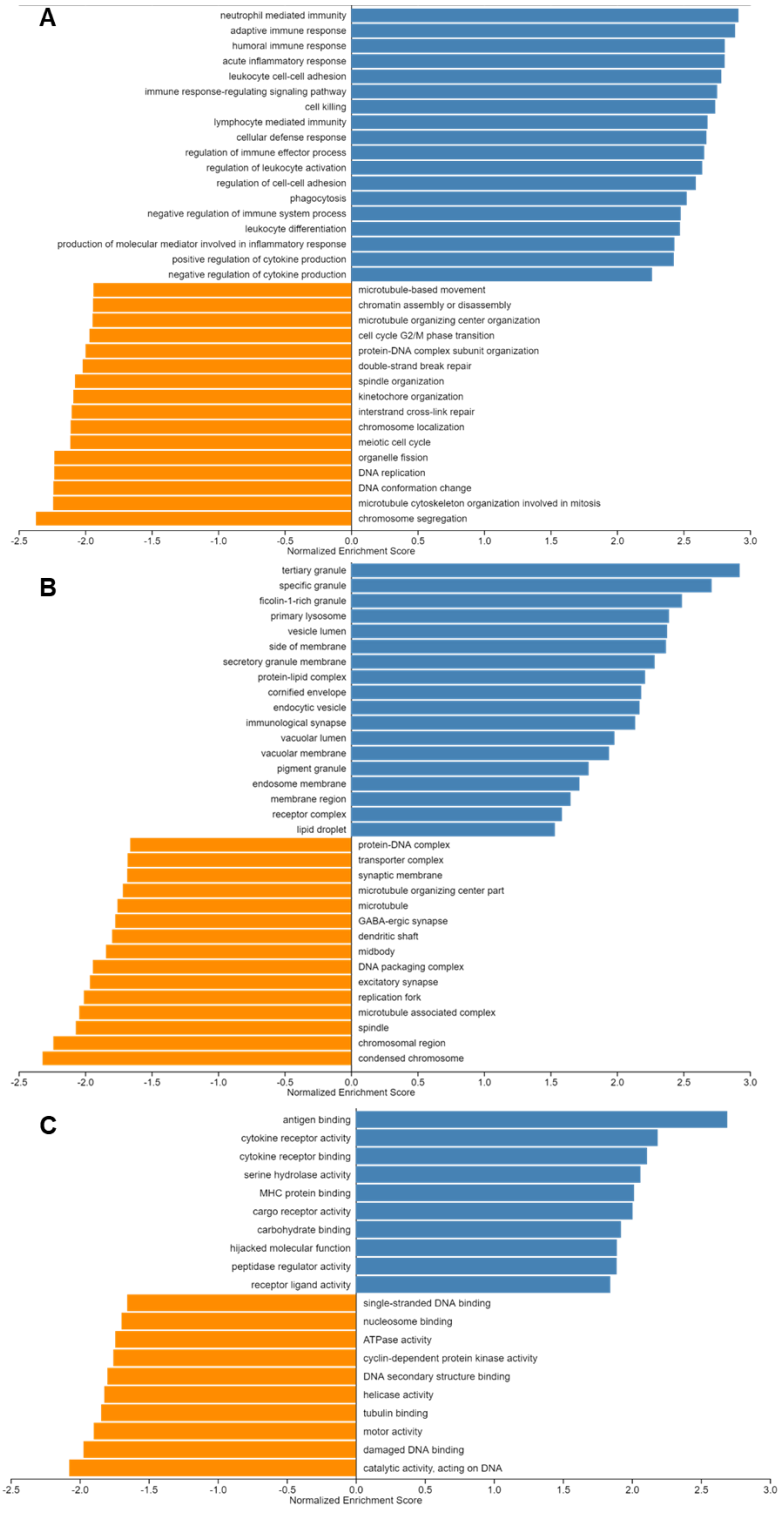


A) Lists of all the 'selectively expressed' genes; B) Box-plot of 'expression' levels of the five genes with the most statistically significant 'selective expression 'in one of the two groups.

3.6 Gene set enrichment analysis during the window of implantation

The pre-ranked gene list of detected transcripts based on their log2 fold change (log2FC) value in the LH+7 versus LH+2 comparison was used for Gene Set Enrichment Analysis (GSEA), with the WEB-based GEne SeT AnaLysis Toolkit (WebGestalt, http://www.webgestalt.org/) and setting a threshold of a FDR < 0.05 with a minimum of 20 genes per category. Analysis of gene ontology revealed that the window of implantation was highly enriched with genes associated with biological processes within the immunological domain including *neutrophil mediated immunity* [Normalized Enrichment Score (NES) = 2.91], *adaptive immune response* (NES = 2.88), *humoral immune response* (NES = 2.80) and *regulation cell-to-cell adhesion* (NES = 2.5), while a significant depletion of processes involved in cellular division, such as *DNA replication* (NES = −2.23) and *cell cycle G2/M transition* (NES = −1.96) was observed (**Figure 15A**). Cellular component analysis (**Figure 15B**) revealed an enrichment in vesicles and immune system components: *tertiary granule* (NES = 2.92), *vesicles lumen* (NES = 37) and *endocytic vesicles* (NES = 2.17) and a depletion of transcripts associated with DNA replication such as *replication fork* (NES = −2.01) and *DNA packaging complex* (NES = −1.94). In terms of molecular functions (**Figure 15C**), we found *antigen binding* (NES = 2.69) and *MHC protein binding* (NES = 2.01) to be enriched and *helicase activity* (NES = −1.82) and *cyclin dependent protein kinase activity* (NES = −1.75) to be depleted.

Figure 15. GSEA Gene Ontology analysis of UF-EVs transcripts enriched or depleted during the receptive phase

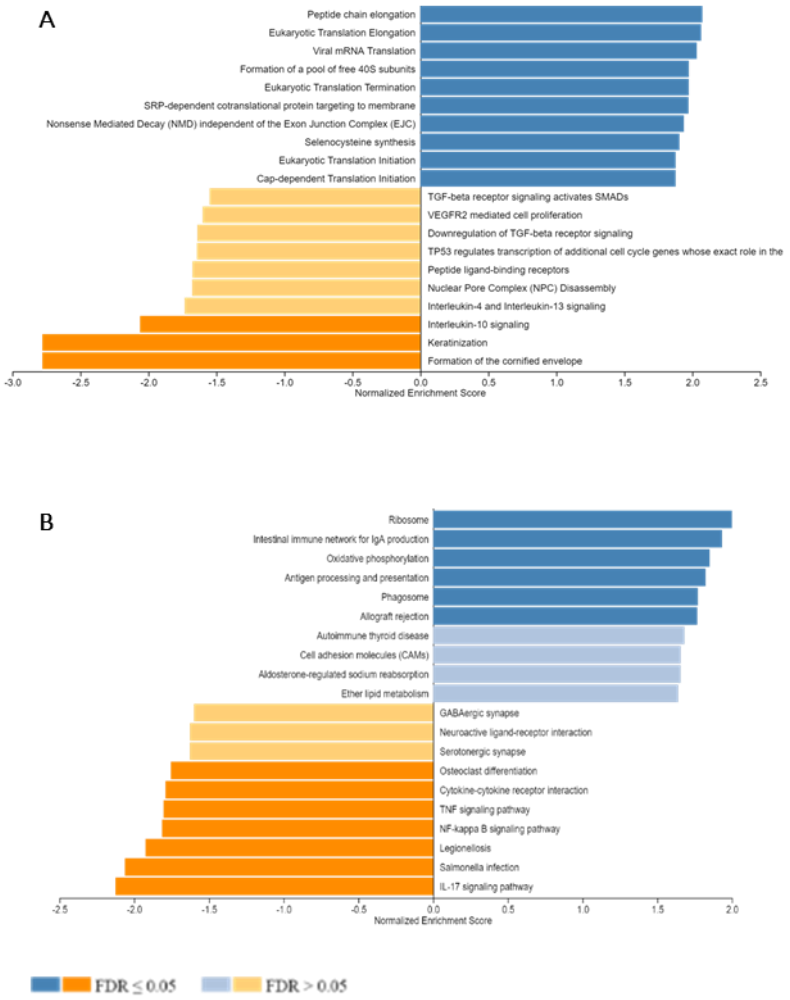


Bar graphs are categorized by gene ontology annotations. A) Biological Process; B) Cellular Component; C) Molecular Function. The NES of most enriched (blue bars) and most depleted (orange bars) pathways on day LH+7 compared to day LH+2 are plotted.

3.7 Gene set enrichment analysis in relation to successful versus failed implantation

The pre-ranked gene list of detected transcripts based on their log2 fold change (log2FC) value in the successuful versus failed implantation comparison was used for GSEA, with the WEB-based GEne SeT AnaLysis Toolkit (WebGestalt, http://www.webgestalt.org/) and setting a threshold of a FDR < 0.01 with a minimum of 20 genes per category. Analysis of gene ontology revealed that successful implantation was highly enriched with genes associated with biological processes within the immunological domain including including *tumor necrosis factor superfamily cytokine production* (NES = 2.02), *natural killer cell activation* (NES = 1.96), *cell killing* (NES = 1.83) and *response to type I interferon* (NES = 1.80) while a significant depletion of processes involved in *epidermis development* (NES = −1.88) was observed. In terms of pathways enriched in patients who achieved implantation (**Figure 16**), among the most significantly enriched were *antigen processing and presentation* (NES = 1.82), *allograft rejection* (NES = 1.76), *MHC protein binding* (NES = 2.01), *cell adhesion molecules (CAMs)* (NES = 1.66), whereas *NF-kappa B signaling pathway* (NES = 1.81), *Interleukin-10 signalling* (NES = −2.06) and *IL-17 signalling pathway* (NES = −2.12) were depleted.

Figure 16. GSEA pathway analysis of UF-EVs transcripts enriched or depleted in women with successful versus failed implantation

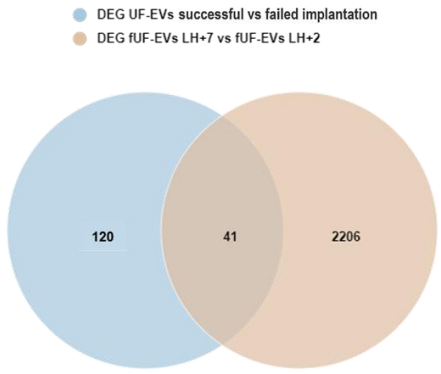


Bar graphs of pathway analysis using A) Reactome or B) KEGG databases are reported. The NES of most enriched (blue bars) and most depleted (orange bars) pathways in women with successful versus failed implantation are plotted. Terms with a FDR < 0.05 are represented with dark blue or orange bars; terms with a FDR≥ 0.05 are represented with light blue or orange bars

3.8 Intersection genes differentially 'expressed' both in relation to the window of implantation and to successful implantation

We verified which genes were differentially 'expressed' in both the window of implantation of fertile women and in the UF-EVs of women with successful versus failed implantation, and identified 41 genes with a significant difference in both comparisons (**Figure 17** and **Table 5**).

Figure 17. Venn graph of the significant genes identified by both DGE analyses (LH+7 vs LH+2 and successful vs failed implantation)



Out of these 41 genes, 24 showed the same positive or negative 'expression' trend whereas 17 had opposite 'expression' trend between the two comparisons.

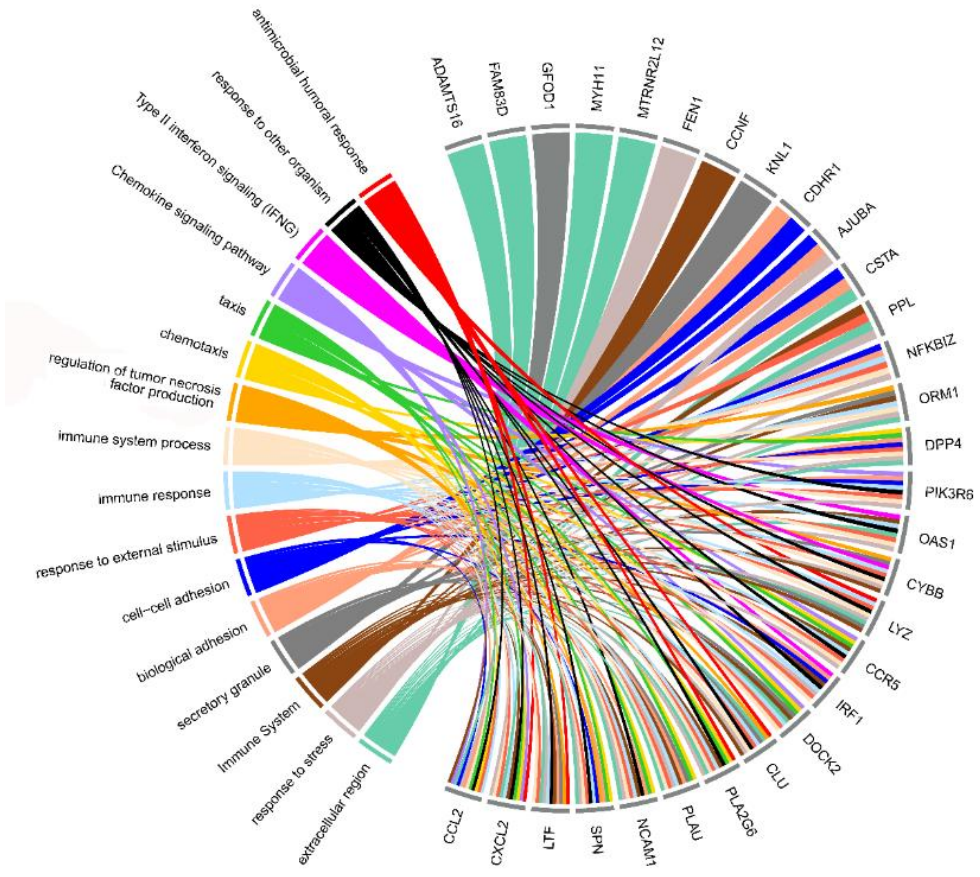
Table 5. List of 41 genes differentially 'expressed' in both DGE analyses between LH+7 versus LH+2 UF-EVs and between UF-EVs of women with successful versus failed implantation

Gene Symbol	Gene Name	log2FC successful vs failed implantation	log2FC LH+7 vs LH+2
CCL2	C-C motif chemokine ligand 2	1,46	2,41
CYBB	cytochrome b-245 beta chain	1,08	2,42
CLU	clusterin	2,17	2,15
ADAMTS16	ADAM metallopeptidase with thrombospondin type 1 motif 16	-1,85	-5,58
KNL1	kinetochore scaffold 1	2,87	-3,81
LZTS1	leucine zipper tumor suppressor 1	-1,63	-3,88
LYZ	lysozyme	-1,39	2,39
DPP4	dipeptidyl peptidase 4	1,53	2,48
SPN	sialophorin	1,66	2,56
HOXB-AS3	HOXB cluster antisense RNA 3	1,03	1,66
AURKB	aurora kinase B	1,82	-2,15
PIK3R6	phosphoinositide-3-kinase regulatory subunit 6	2,34	2,36

OAS1	2'-5'-oligoadenylate synthetase 1	1,06	1,04
CCR5	C-C motif chemokine receptor 5 (gene/pseudogene)	1,68	2,32
DOCK2	dedicator of cytokinesis 2	1,20	2,28
ZFP69B	ZFP69 zinc finger protein B	2,15	-2,87
SNAI3-AS1	SNAI3 antisense RNA 1	-1,33	-2,35
TUNAR	TCL1 upstream neural differentiation-associated RNA	-1,52	-1,41
IRF1	interferon regulatory factor 1	1,03	1,32
CDHR1	cadherin related family member 1	1,69	1,87
GFOD1	glucose-fructose oxidoreductase domain containing 1	-1,33	2,13
MYH11	myosin heavy chain 11	1,73	2,01
LTF	lactotransferrin	1,69	1,60
MTRNR2L12	MT-RNR2 like 12	1,26	1,46
PLAU	plasminogen activator, urokinase	-1,21	1,48
FEN1	flap structure-specific endonuclease 1	1,04	-1,26
NCAM1	neural cell adhesion molecule 1	1,76	1,82
NFKBIZ	NFKB inhibitor zeta	-1,38	1,23
AJUBA	ajuba LIM protein	1,14	-1,84
CXCL2	C-X-C motif chemokine ligand 2	-2,43	1,51
ZNF70	zinc finger protein 70	1,68	-1,58
FAM83D	family with sequence similarity 83 member D	-2,43	-2,67
PLA2G6	phospholipase A2 group VI	1,58	-1,69
PPL	periplakin	-1,53	1,15
CCNF	cyclin F	1,91	-1,30
CSTA	cystatin A	-2,03	1,32
ORM1	orosomucoid 1	2,56	2,41
GAS6-AS2	GAS6 divergent transcript	1,81	-2,49
AC098613.1		2,29	1,89
AC016582.3	novel zinc finger protein pseudogene	-1,31	-2,22
LINC02486	long intergenic non-protein coding RNA 2486	2,36	-1,83

Enrichment analysis for identification of biological processes and pathways connected to these genes (performed with g:Profiler software) revealed them to be mostly involved in biological processes such as *cell-cell adhesion*, *immune system process, chemotaxis, response to external stimuli*. Three significantly enriched pathways were identified: KEGG pathway of *chemokine signaling*, Reactome pathway of *immune System* and a WikiPathway term of *type II interferon signaling*. A significant number of the genes were also connected with the *extracellular region* and *secretory* cell compartments (**Figure 18**).

Figure 5. N=41 intersection genes 'expressed' in both the window of implantation of fertile women and in the UF-EVs of women with successful versus failed implantation.

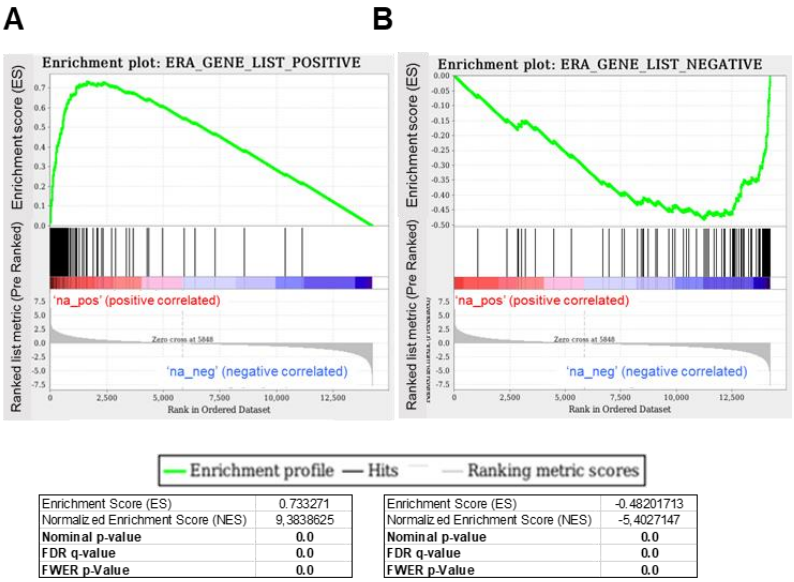


Genes are shown on the right side; their single or multiple GeneOntology biological processes on the left side.

3.9 Exploratory analyses of genes in common with current diagnostic tests and literature

To evaluate whether the 238 genes comprised in the biopsy-based ERA® test (of which n = 143 are described to be up-regulated and n = 95 to be down-regulated during the implantation window, Díaz-Gimeno et al., 2011) are also enriched in UF-EVs, preranked-GSEA (Mootha *et al.* 2003; Subramanian *et al.*, 2005) was performed on the log2FC data resulting from our LH+7 versus LH+2 comparison. GSEA indeed highlighted that UF-EVs have a similar transcriptional profile to that reported in the receptive endometrial tissue as defined by ERA®: NES of 9.38 ($P < 0.001$) and of -5.40 ($P < 0.001$) was found for up-regulated and down-regulated transcripts respectively (**Figure 19**). The GSEA leading edge, i.e. the subset of our gene set that contributed most to the enrichment score, counted 106 enriched and 64 depleted transcripts.

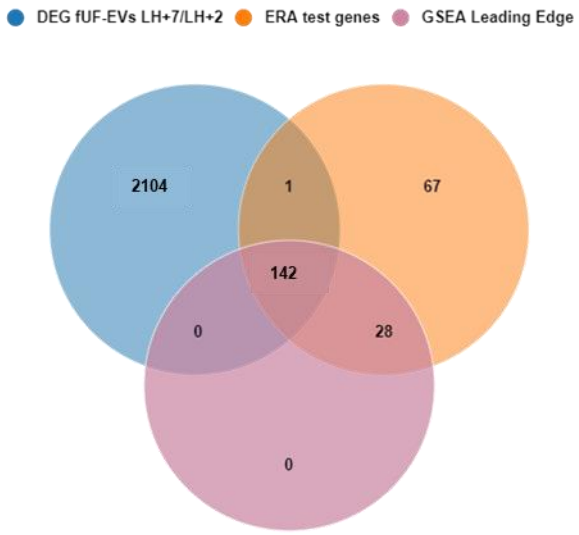
Figure 6. GSEA curves obtained for the n = 143 and n = 95 ERA® gene-sets expected to be respectively up-regulated (A) and down-regulated (B) during the implantation window



One hundred and twenty-six out of 143 ERA genes expected to be up-regulated during the implantation window, were enriched in fUF-EVs LH+7 (NES=9.38; p<0.001) and 93 genes of 95 ERA genes expected to be down-regulated during the implantation window, were depleted in fUF-EVs LH+7 (NES=-5.40; p<0.001).

Comparing the lists of leading-edge genes with ERA® genes and with differentially 'expressed' genes (DEGs of UF-EVs LH+7 versus LH+2), we found 142 genes in common among these three lists (**Figure 20**).

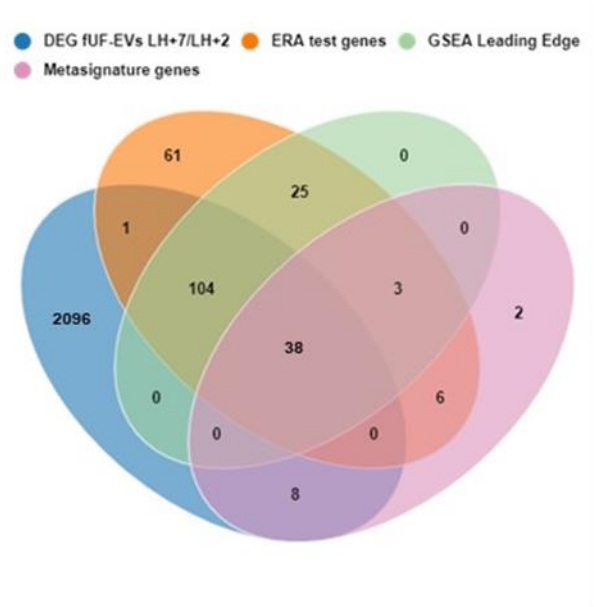
Figure 20. Venn graph showing the overlap between genes in the ERA® test, differentially 'expressed' genes in UF-EVs during the receptive phase and genes in the GSEA Leading Edge



As shown in the Venn diagram (**Figure 20**), only one gene (which is *BIRC3*) belongs to both the list of ERA® genes and of DEGs but not to leading-edge genes, as it shows more abundance during the implantation window in UF-EVs but down-regulation in tissues according to the ERA® results.

A further intersection was performed with the list of n=57 genes composing the 'metasignature' of endometrial receptivity-associated transcripts based on the latest meta-analysis of current literature (Altmäe et al., 2017). This intersection identified 38 genes reported in Table that thus: i) show differential 'expression' during the window of implantation according to both ERA® test and our LH+7 vs LH +2 UF-EVs comparison; ii) belong to the GSEA leading edge of genes with the most significant enrichment score; and iii) are identified as receptivity-associated genes by the current meta-analysis on endometrial transcriptome (**Figure 21**).

Figure 21. Venn graph showing the overlap between genes in the ERA® test, differentially 'expressed' genes in UF-EVs during the receptive phase, genes in the GSEA Leading Edge and genes belonging to the 'metasignature' of endometrial receptivity



The 38 significant genes belonging to the LH+7/LH+2 DGE analysis, to theERA test, to the leasing edge of the relative GSEA and to the endometrial receptivity 'metasignature' are shown in **Table 6**.

Table 6. List of 38 transcripts in common among the genes listed in the ERA test, in the Endometrial Metasignature, in the GSEA Leading Edge and in the DGE analysis between UF-EVs LH+7 versus LH+2

GeneID	log2FC	pvalue	padj	description
MT1G	6,76	2,6E-55	3,7E-51	metallothionein 1G [HGNC:7399]
PAEP	6,11	7,4E-48	5,3E-44	progestagen associated endometrial protein [HGNC:8573]
SPP1	3,91	3,0E-34	8,4E-31	secreted phosphoprotein 1 [HGNC:11255]
SFRP4	-3,90	1,1E-30	2,7E-27	secreted frizzled related protein 4 [HGNC:10778]
MT1H	8,94	4,7E-27	8,3E-24	metallothionein 1H [HGNC:7400]

NNMT	3,65	8,9E-27	1,4E-23	nicotinamide N-methyltransferase [HGNC:7861]
C10orf10	3,59	1,2E-23	1,7E-20	chromosome 10 open reading frame 10 [HGNC:23355]
GPX3	4,04	1,7E-23	2,0E-20	glutathione peroxidase 3 [HGNC:4555]
CRABP2	-2,23	4,1E-22	3,5E-19	cellular retinoic acid binding protein 2 [HGNC:2339]
C4BPA	5,03	7,9E-19	4,7E-16	complement component 4 binding protein alpha [HGNC:1325]
SERPING1	2,69	1,1E-18	6,1E-16	serpin family G member 1 [HGNC:1228]
GADD45A	2,08	5,1E-16	1,6E-13	growth arrest and DNA damage inducible alpha [HGNC:4095]
MAOA	2,92	5,5E-16	1,7E-13	monoamine oxidase A [HGNC:6833]
CFD	3,53	1,9E-15	5,0E-13	complement factor D [HGNC:2771]
IDO1	2,63	5,2E-14	1,0E-11	indoleamine 2,3-dioxygenase 1 [HGNC:6059]
DEFB1	2,84	1,0E-13	2,0E-11	defensin beta 1 [HGNC:2766]
OLFM1	-4,00	3,1E-13	5,4E-11	olfactomedin 1 [HGNC:17187]
CP	3,37	4,2E-13	6,9E-11	ceruloplasmin [HGNC:2295]
GNLY	3,32	6,9E-13	1,1E-10	granulysin [HGNC:4414]
EDN3	-4,61	8,5E-13	1,3E-10	endothelin 3 [HGNC:3178]
CLDN4	2,08	4,1E-11	4,8E-09	claudin 4 [HGNC:2046]
ARG2	2,36	5,2E-10	4,5E-08	arginase 2 [HGNC:664]
GBP2	2,27	5,7E-09	3,5E-07	guanylate binding protein 2 [HGNC:4183]
DKK1	2,71	1,1E-08	6,3E-07	dickkopf WNT signaling pathway inhibitor 1 [HGNC:2891]

ABCC3	3,55	3,9E-08	1,9E-06	ATP binding cassette subfamily C member 3 [HGNC:54]
ANXA4	1,46	4,7E-08	2,2E-06	annexin A4 [HGNC:542]
BCL6	1,51	1,4E-07	5,5E-06	B-cell CLL/lymphoma 6 [HGNC:1001]
DPP4	2,48	4,8E-07	1,6E-05	dipeptidyl peptidase 4 [HGNC:3009]
TSPAN8	2,51	5,9E-07	1,9E-05	tetraspanin 8 [HGNC:11855]
S100P	2,35	5,8E-06	1,3E-04	S100 calcium binding protein P [HGNC:10504]
LAMB3	2,04	6,0E-06	1,3E-04	laminin subunit beta 3 [HGNC:6490]
ARID5B	1,03	2,1E-05	3,8E-04	AT-rich interaction domain 5B [HGNC:17362]
IL15	1,33	4,3E-05	6,8E-04	interleukin 15 [HGNC:5977]
ID4	1,55	4,3E-05	6,8E-04	inhibitor of DNA binding 4, HLH protein [HGNC:5363]
ENPEP	1,50	5,1E-05	7,8E-04	glutamyl aminopeptidase [HGNC:3355]
EFNA1	1,69	6,6E-05	9,7E-04	ephrin A1 [HGNC:3221]
TCN1	2,36	2,9E-04	3,4E-03	transcobalamin 1 [HGNC:11652]
SLC1A1	2,47	3,4E-04	3,8E-03	solute carrier family 1 member 1 [HGNC:10939]

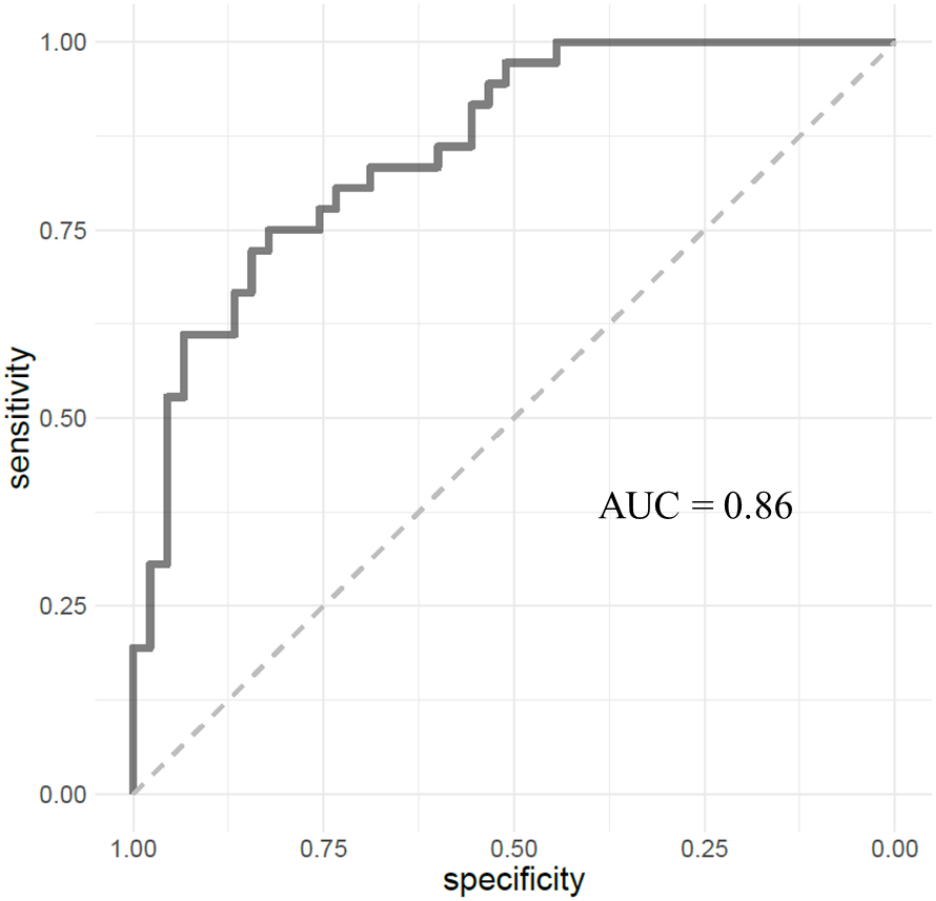
3.10 Validation of the UF-EVs transcriptome as a test for a receptive endometrium in the global cohort of PGT-A patients

After including a validation cohort (n=52) and thus increasing the total number of UF-derived EVs samples (n=85; excluding low quality samples), n=19 genes were confirmed as being differentially 'expressed' between women who achieved pregnancy and women who failed to conceive and more specifically *RPL10P9, BMP4, LINC00621, PTGES, USP27X, AURKB, BAALC, AC245060.5, MALAT1, CLU, BBS5, RAB38, RTN2 ,ZNF571, DUSP1, PLA2G6, ENPP5, LNPK,ASPA*. Among them we can

find the most significant differentially expressed transcripts already identified in Aim 2 (such as *CLU, DUSP1*), validating the study results. Furthermore, the NTA on all samples collected from the validation cohort also confirms the same trend as the samples analyzed in the study: comparison of size distributions shows that EVs isolated from UF of women who failed to conceive are larger than EVs collected from women who achieved pregnancy. In particular, the 90 percentiles of the size distributions are significantly different between the two sample types ($p = 0.02$).

The n=19 genes identified were included into a logistic regression model with forward stepwise entry selection, in order to identify a smaller set of genes capable of discriminating between pregnancy and non-pregnancy. *RPL10P9, BMP4, CLU, DUSP1, PTGES* genes emerged as the ones selected for the prediction model, to which the 90th percentile value of EV size distribution was also added. Finally, propensity score was calculated for each sample based on the RNA expression of chosen genes and vesicles dimensions. A ROC curve was then calculated (**Figure 22**) describing the predictive characteristics of the model (Area Under the Curve, AUC = 0.86, 95%CI = 0.78-0.94, $p=2.8 \times 10^{-8}$) with respect to successful versus failed implantation.

Figure 22. Evaluation of EV-RNA profile and diameter as predictive biomarkers for clinical pregnancy in ART by ROC curve



AUC: Area Under the Curve. ROC curve analysis for clinical pregnancy prediction by RPL10P9, BMP4, CLU, DUSP1, PTGES transcript level in extracellular vesicles and 90th percentile value of extracellular vesicles' size distribution is shown (AUC = 0.86, 95%CI = 0.78-0.94, p=2.8 x 10^{-8}).

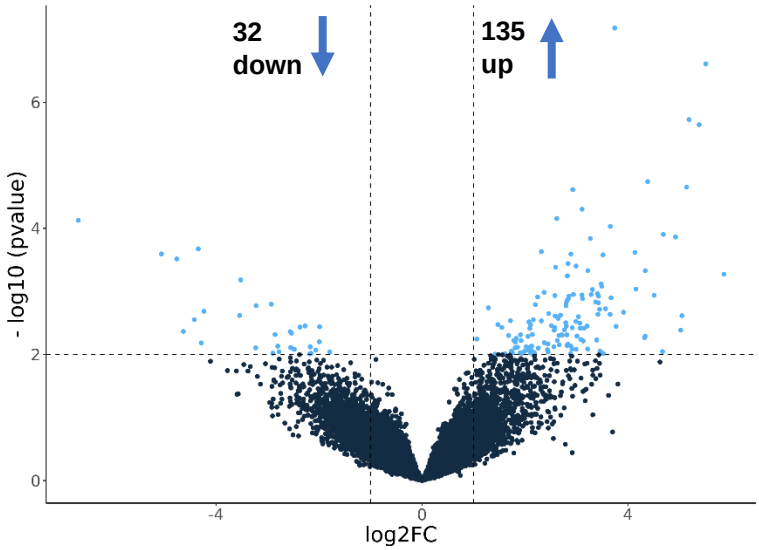
3.11 Subgroup transcriptomic analyses of patients with Recurrent Implantation Failure

The total cohort of ART patients included comprised n=16 women who underwent PGT-A due to a history of RIF, with a mean (± SD) age of 36.6 (± 3.4) and 5.1 ± 2.1 (mean ± SD) previous failed embryo transfers. After their first euploid embryo transfer, n=6 (37.5%) achieved successful pregnancy and n=10 (62.5%) failed implantation. The UF-EVs content of RIF women with failed implantation also after euploid blastocyst transfer was thus compared to the control group of fertile women who underwent ART and PGT-A due to a genetic disease and achieved successful

pregnancy (n=28), as well as to patients with a diagnosis of RIF who were instead successfully treated with PGT-A.

Of the 17191 'expressed' genes found, 135 were significantly upregulated and 32 were significantly downregulated in the RIF group with failed implantation versus the fertile group with implantation (**Figure 23**). Results were obtained using DESeq2 and significance was defined according to SEQC cutoff, that selects genes showing at the same time a raw P-value < 0.01 and a |log2FC| >1.

Figure 23. Volcano plot of the significantly differentially detected RNAs in UF-EVs from RIF patients without implantation and fertile women with implantation



Results of DGE analysis comparing the transcriptional profile of UF-EVs derived from RIF patients with failed implantation versus fertile controls with successful implantation are shown in **Table 7**.

Table 7. DGE of the top statistically significant 50 genes with a difference between RIF patients with failed implantation and fertile women undergoing PGT-A for genetic diseases, with successful implantation.

GeneID	log2FoldChange	pvalue	padj	Significance (FDR)	Significance (SEQC)	description
DUSP1	3.74	6.62E-08	0.001137573	1	1	dual specificity phosphatase 1 [HGNC:3064]
AC127526.5	5.51	2.45E-07	0.002108968	1	1	novel transcript
FMO2	5.18	1.88E-06	0.0097215	1	1	flavin containing monooxygenase 2 [HGNC:3770]
MYO7B	5.38	2.26E-06	0.0097215	1	1	myosin VIIB [HGNC:7607]
AREG	4.38	1.81E-05	0.059555565	0	1	amphiregulin [HGNC:651]
LY6G6C	5.14	2.22E-05	0.059555565	0	1	lymphocyte antigen 6 family member G6C [HGNC:13936]
IPMK	2.93	2.43E-05	0.059555565	0	1	inositol polyphosphate multikinase [HGNC:20739]
RGS2	3.10	4.96E-05	0.106669075	0	1	regulator of G protein signaling 2 [HGNC:9998]
AP001062.1	2.62	7.00E-05	0.12806334	0	1	novel transcript, antisense to C21orf2

AC118754.1	-6.67	7.45E-05	0.12806334	0	-1	sphingolipid transporter 2 [NCBI:124976]
STX17-AS1	3.65	9.32E-05	0.14557703	0	1	STX17 antisense RNA 1 [HGNC:51174]
PITX1	4.68	0.00012459	0.17788217	0	1	paired like homeodomain 1 [HGNC:9004]
SERPINB2	4.92	0.00013695	0.17788217	0	1	serpin family B member 2 [HGNC:8584]
PTGS2	3.26	0.00014486	0.17788217	0	1	prostaglandin-endoperoxide synthase 2 [HGNC:9605]
IGHA2	-4.34	0.00021071	0.22566922	0	-1	immunoglobulin heavy constant alpha 2 (A2m marker) [HGNC:5479]
AGO4	2.31	0.00023409	0.22566922	0	1	argonaute RISC component 4 [HGNC:18424]
FAM25A	4.13	0.00024073	0.22566922	0	1	family with sequence similarity 25 member A [HGNC:23436]

AC011298.1	-5.06	0.00025379	0.22566922	0	-1	novel transcript
LRMP	2.89	0.00025521	0.22566922	0	1	lymphoid restricted membrane protein [HGNC:6690]
CRCT1	3.51	0.00026254	0.22566922	0	1	cysteine rich C-terminal 1 [HGNC:29875]
AOC1	-4.7637044	0.00030508	0.24974252	0	-1	amine oxidase copper containing 1 [HGNC:80]
FOS	2.83	0.00036122	0.28225821	0	1	Fos proto-oncogene, AP-1 transcription factor subunit [HGNC:3796]
SPINK5	2.99	0.00039558	0.29539927	0	1	serine peptidase inhibitor, Kazal type 5 [HGNC:15464]
DENND3	2.59	0.0004124	0.29539927	0	1	DENN domain containing 3 [HGNC:29134]
RAB26	3.22	0.0004681	0.30990287	0	1	RAB26, member RAS oncogene family [HGNC:14259]

DUOXA2	4.33	0.0004687	0.30990287	0	1	dual oxidase maturation factor 2 [HGNC:32698]
LOR	5.86	0.00053376	0.3398485	0	1	loricrin [HGNC:6663]
ZNF665	2.82	0.00056572	0.34733332	0	1	zinc finger protein 665 [HGNC:25885]
KNL1	-3.52	0.00065666	0.38926483	0	-1	kinetochore scaffold 1 [HGNC:24054]
AL035587.1	3.48	0.00075932	0.43511642	0	1	novel transcript
CRNN	3.49	0.00082941	0.45994909	0	1	cornulin [HGNC:1230]
SPINK7	4.15	0.0009161	0.48632111	0	1	serine peptidase inhibitor, Kazal type 7 (putative) [HGNC:24643]
OLR1	3.31	0.00093355	0.48632111	0	1	oxidized low density lipoprotein receptor 1 [HGNC:8133]
EVPL	2.37	0.00103353	0.4864929	0	1	envoplakin [HGNC:3503]
AC092422.1	3.40	0.00108164	0.4864929	0	1	novel transcript

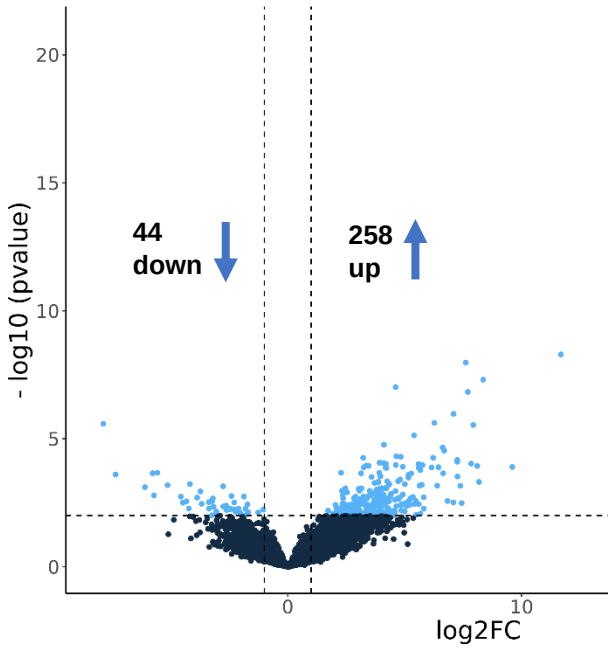
KRT13	2.94	0.0011 1732	0.4864 929	0	1	keratin 13 [HGNC:6415]
HIST1H 3H	3.28	0.0011 1961	0.4864 929	0	1	histone cluster 1 H3 family member h [HGNC:4775]
TREX2	3.10	0.0011 3714	0.4864 929	0	1	three prime repair exonuclease 2 [HGNC:12270]
TIGD3	4.51	0.0011 5359	0.4864 929	0	1	tigger transposable element derived 3 [HGNC:18334]
ABHD3	2.58	0.0011 6467	0.4864 929	0	1	abhydrolase domain containing 3 [HGNC:18718]
SPRR3	2.85	0.0011 8409	0.4864 929	0	1	small proline rich protein 3 [HGNC:11268]
CXCL2	3.42	0.0011 8857	0.4864 929	0	1	C-X-C motif chemokine ligand 2 [HGNC:4603]
NAMP TP1	2.24	0.0012 2143	0.4883 1816	0	1	nicotinamide phosphoribosyltransferase pseudogene 1 [HGNC:17633]

RNASE 7	3.67	0.0012 6549	0.4934 47	0	1	ribonuclease A family member 7 [HGNC:19278]
MMP2 5	3.11	0.0012 9696	0.4934 47	0	1	matrix metallopeptidase 25 [HGNC:14246]
C5AR1	2.84	0.0013 2037	0.4934 47	0	1	complement C5a receptor 1 [HGNC:1338]
MYOM 1	2.81323 709	0.0014 0537	0.5140 3657	0	1	myomesin 1 [HGNC:7613]
PLP1	3.37	0.0014 7108	0.5268 6214	0	1	proteolipid protein 1 [HGNC:9086]
LYPD2	3.42	0.0015 1278	0.5307 3872	0	1	LY6/PLAUR domain containing 2 [HGNC:25215]
CSTA	2.79	0.0015 4781	0.5321 6943	0	1	cystatin A [HGNC:2481]

The UF-EVs content of RIF women with failed implantation after euploid blastocyst transfer (n=10) was then compared to the group of RIF women who achieved successful pregnancy after euploid embryo transfer (n=6). Of the 17191 'expressed' genes found, 258 were significantly upregulated and 44 were significantly downregulated in the RIF group with failed implantation versus the RIF group with implantation (**Figure 24**). Results were obtained using DESeq2 and significance was

defined according to SEQC cutoff, that selects genes showing at the same time a raw P-value < 0.01 and a |log2FC| >1.

Figure 24. Volcano plot of the significantly differentially detected RNAs in UF-EVs from RIF patients without implantation and RIF patients with implantation.



Results of DGE analysis comparing the transcriptional profile of UF-EVs derived from RIF patients with failed implantation versus RIF patients with successful implantation are shown in **Table 8**.

Table 8. DGE of the top statistically significant 50 genes with a difference between RIF patients with failed implantation and RIF patients with successful implantation

GeneID	log2FoldChange	pvalue	padj	FDR significance	SEQC significance	description
DEFA4	24.08	1.42E-21	2.45E-17	1	1	defensin alpha 4 [HGNC:2763]
CCL7	24.21	1.24E-19	1.07E-15	1	1	C-C motif chemokine

						ligand 7 [HGNC:10634]
NLRP12	11.68	5.05E-09	2.89E-05	1	1	NLR family pyrin domain containing 12 [HGNC:22938]
ASPG	7.61	1.05E-08	4.52E-05	1	1	asparaginase [HGNC:20123]
MYO7B	8.36	4.91E-08	0.000168682	1	1	myosin VIIB [HGNC:7607]
DUSP1	4.61	9.63E-08	0.0002758	1	1	dual specificity phosphatase 1 [HGNC:3064]
DEFB4A	7.70	1.47E-07	0.000360801	1	1	defensin beta 4A [HGNC:2767]
LILRA1	7.09	1.06E-06	0.002277523	1	1	leukocyte immunoglobulin like receptor A1 [HGNC:6602]
OSM	6.27	2.38E-06	0.004440838	1	1	oncostatin M [HGNC:8506]
AC011298.1	-7.89	2.58E-06	0.004440838	-1	-1	novel transcript
PADI4	7.93	2.88E-06	0.004506972	1	1	peptidyl arginine deiminase 4 [HGNC:18368]
IL1B	5.40	7.37E-06	0.01056168	1	1	interleukin 1 beta [HGNC:5992]
RGS2	4.11	1.70E-05	0.022507304	1	1	regulator of G protein

						signaling 2 [HGNC:9998]
P2RX1	6.62	2.19E-05	0.026944062	1	1	purinergic receptor P2X 1 [HGNC:8533]
DUOXA2	6.69	2.92E-05	0.033496591	1	1	dual oxidase maturation factor 2 [HGNC:32698]
G0S2	4.64	4.83E-05	0.050099036	0	1	G0/G1 switch 2 [HGNC:30229]
NLRP3	4.84	5.04E-05	0.050099036	0	1	NLR family pyrin domain containing 3 [HGNC:16400]
CASS4	6.22	5.49E-05	0.050099036	0	1	Cas scaffold protein family member 4 [HGNC:15878]
TNFSF13B	3.22	5.54E-05	0.050099036	0	1	TNF superfamily member 13b [HGNC:11929]
TIGD3	7.25	6.68E-05	0.057385962	0	1	tigger transposable element derived 3 [HGNC:18334]
DEFA3	7.25	7.91E-05	0.060820827	0	1	defensin alpha 3 [HGNC:2762]
LRMP	3.93	8.60E-05	0.060820827	0	1	lymphoid restricted

						membrane protein [HGNC:6690]
ADGRE2	4.08	8.79E-05	0.060820827	0	1	adhesion G protein-coupled receptor E2 [HGNC:3337]
TREML2	7.84	9.34E-05	0.060820827	0	1	triggering receptor expressed on myeloid cells like 2 [HGNC:21092]
PTGS2	4.19	9.49E-05	0.060820827	0	1	prostaglandin-endoperoxide synthase 2 [HGNC:9605]
MGAM	5.59	9.61E-05	0.060820827	0	1	maltase-glucoamylase [HGNC:7043]
CXCL8	4.62	0.000105614	0.060820827	0	1	C-X-C motif chemokine ligand 8 [HGNC:6025]
PROK2	5.76	0.000106314	0.060820827	0	1	prokineticin 2 [HGNC:18455]
CD37	3.36	0.000111481	0.060820827	0	1	CD37 molecule [HGNC:1666]
SRGN	3.46	0.000114287	0.060820827	0	1	serglycin [HGNC:9361]
FUT7	8.10	0.0001151	0.060820827	0	1	fucosyltransferase 7 [HGNC:4018]
MMP25	4.92	0.000124454	0.060820827	0	1	matrix metallopeptidase 25

						[HGNC:14246]
RPS4Y1	9.60	0.000126365	0.060820827	0	1	ribosomal protein S4 Y-linked 1 [HGNC:10425]
SERPINB2	6.44	0.000128518	0.060820827	0	1	serpin family B member 2 [HGNC:8584]
RNF19B	3.24	0.000130332	0.060820827	0	1	ring finger protein 19B [HGNC:26886]
KCNH7	6.15	0.000131703	0.060820827	0	1	potassium voltage-gated channel subfamily H member 7 [HGNC:18863]
RGS1	4.18	0.000133406	0.060820827	0	1	regulator of G protein signaling 1 [HGNC:9991]
SULT1B1	5.60	0.000135917	0.060820827	0	1	sulfotransferase family 1B member 1 [HGNC:17845]
TAGAP	4.21	0.00013798	0.060820827	0	1	T cell activation RhoGTPase activating protein [HGNC:15669]
ZNF812P	5.67	0.000167451	0.071966361	0	1	zinc finger protein 812, pseudogene

						[HGNC:33242]
ARHGDIB	2.28	0.000212071	0.08290495	0	1	Rho GDP dissociation inhibitor beta [HGNC:679]
LINC02126	-5.58	0.000213463	0.08290495	0	-1	long intergenic non-protein coding RNA 2126 [HGNC:52983]
CHI3L1	3.91	0.000215836	0.08290495	0	1	chitinase 3 like 1 [HGNC:1932]
RHBDL1	-5.78	0.000222048	0.08290495	0	-1	rhomboid like 1 [HGNC:10007]
AC034199.1	6.64	0.000222394	0.08290495	0	1	novel transcript
SOCS3	3.13	0.000223198	0.08290495	0	1	suppressor of cytokine signaling 3 [HGNC:19391]
SPX	5.28	0.000226661	0.08290495	0	1	spexin hormone [HGNC:28139]
NLRP5	-7.37	0.000246735	0.088367071	0	-1	NLR family pyrin domain containing 5 [HGNC:21269]
PLP1	5.16	0.0002526	0.088621518	0	1	proteolipid protein 1 [HGNC:9086]

RHOH	3.96	0.000292023	0.100182302	0	1	ras homolog family member H [HGNC:686]

Of the genes with a significantly different 'expression', 63 were consistently significantly enriched (n=61) or depleted (n=2) in RIF patients who failed implantation compared to both fertile controls with successful implantation after PGT-A and to RIF patients with successful implantation. The two genes belonging to this intersection and having the highest statistically significant difference are MYO7B and DUSP1 (**Table 9**). The two depleted genes in this intersection are IGHA2, encoding for the immunoglobulin heavy constant alpha 2 protein (part of monomeric IgA immunoglobulin complex and secretory dimeric IgA immunoglobulin complex) and AC011298.1, a novel transcript.

Table 9. Genes with the highest statistically significant difference between RIF patients failing implantation and both fertile women with implantation and RIF patients achieving implantation

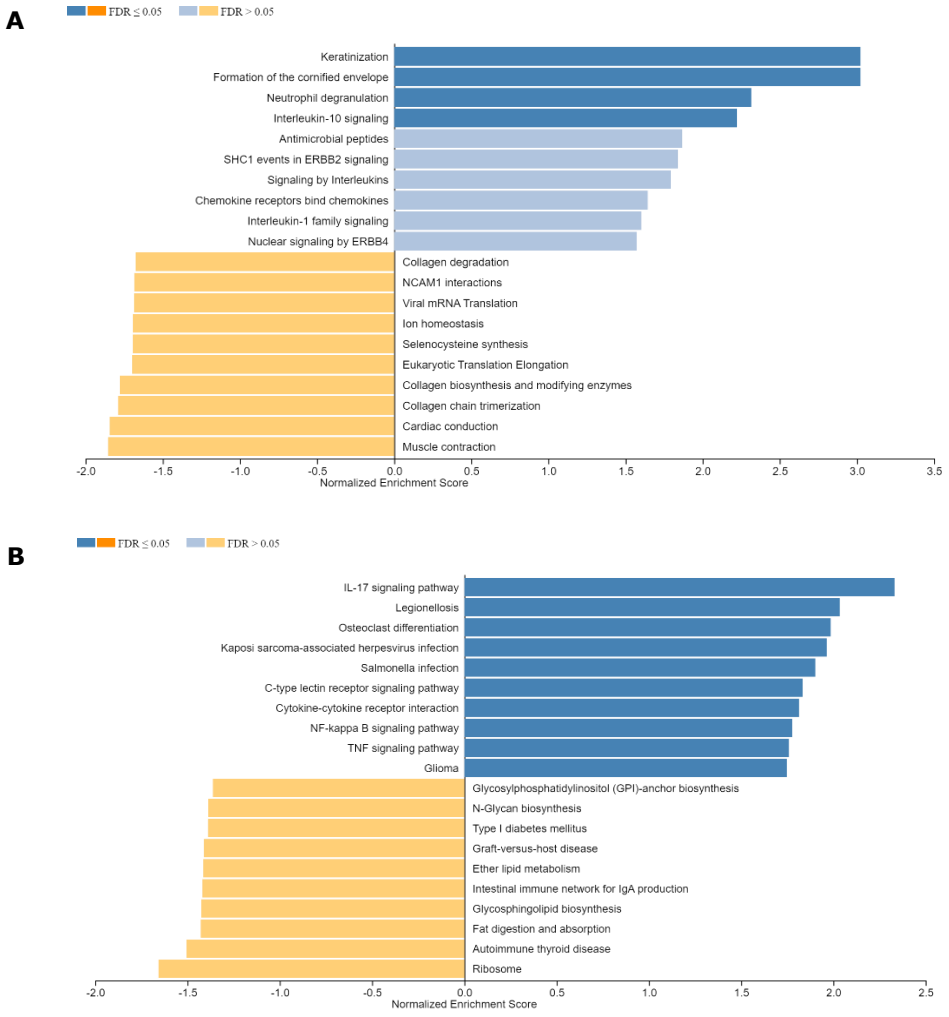
Gene Symbol	Gene Name	log2FC RIF failing implantation vs fertile controls achieving implantation	log2FC RIF failing implantation vs RIF achieving implantation
MYO7B	myosin VIIB *[HGNC:7607]*	5.38	8.36
DUSP1	dual specificity phosphatase 1 *[HGNC:3064]*	3.74	4.61

3.12 Pathway enrichment in patients with persistent implantation failure versus fertile patients with implantation

All detected gene transcripts (17191) were there sorted based on their log2FC value between UF-EVs from RIF patients without implantation and fertile controls undergoing PGT-A for genetic diseases with implantation, resulting in a pre-ranked gene list that was further used in a GSEA (WebGestalt). Among different molecular

Reactome and KEGG pathways enriched in EVs from women with persistent implantation failure, the most significantly enriched were *keratinization* (NES 3.023), *formation of the cornified envelope* (NES 3.023), *neutrophil degranulation* (NES 2.315), *IL-10 signalling* (NES 2.222), IL-*17 signalling pathway* (NES 2.331). In contrast, *muscle contraction* (NES -1.853) and *ribosome* (NES -1.656) were the most significantly depleted (**Figure 25**).

Figure 25. GSEA pathway analysis of UF-EVs transcripts enriched or depleted in RIF patients with failed implantation and fertile women with implantation.



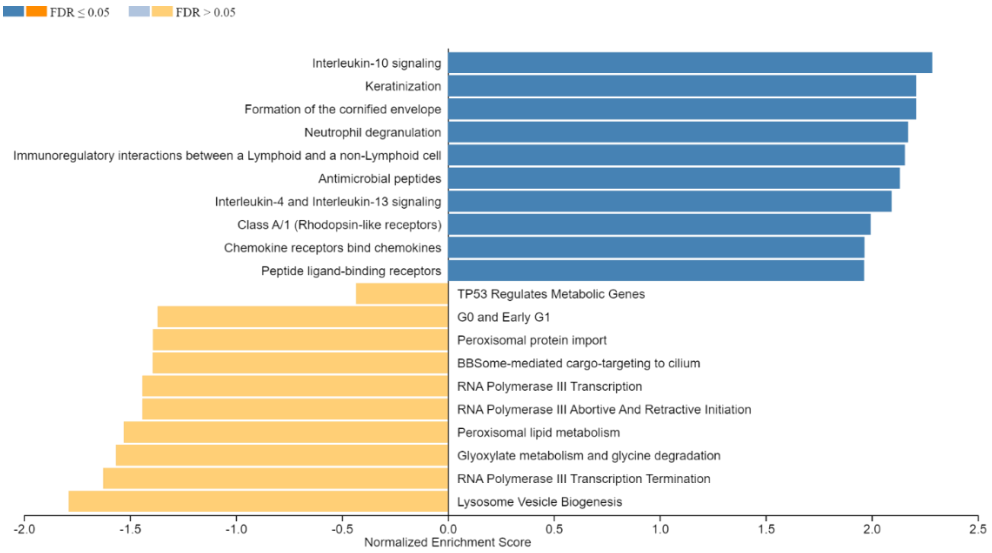
Bar graphs of pathway analysis using A) Reactome or B) KEGG databases are reported. The NES of most enriched (blue bars) and most depleted (orange bars) are plotted.

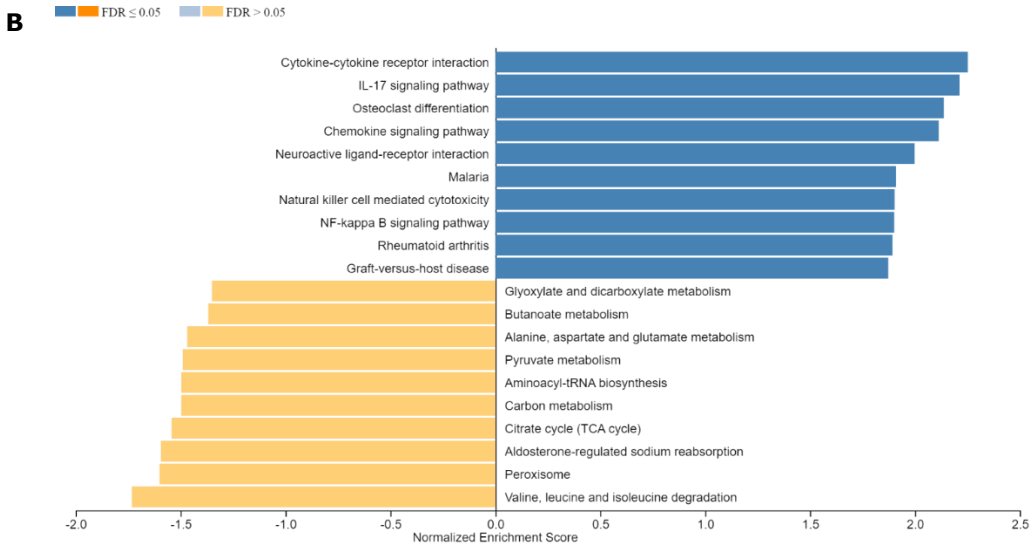
3.13 Pathway enrichment in patients with persistent implantation failure versus patients successfully treated by euploid blastocyst transfer

All detected gene transcripts (17191) were sorted based on their log2FC value between UF-EVs from RIF patients without implantation and RIF patients with implantation, resulting in a pre-ranked gene list that was further used in a GSEA (WebGestalt). Among different molecular Reactome and KEGG pathways enriched in EVs from women with persistent implantation failure, the most significantly enriched were *IL-10 signalling* (NES 2.286), *keratinization* (NES 2.210), *formation of the cornified envelope* (NES 2.210), *neutrophil degranulation* (NES 2.171), *Cytokine-cytokine receptor interaction* (NES 2.251), *IL-17 signaling pathway* (NES 2.213). In contrast, *lysosome vescicle biogenesis* (NES -1.789) and *valine, leucine and isoleucine degradation* (NES -1.732) were the most significantly depleted (**Figure 26**).

Figure 26. GSEA pathway analysis of UF-EVs transcripts enriched or depleted in RIF patients with failed implantation versus RIF patients with successful implantation.

A





Bar graphs of pathway analysis using A) Reactome or B) KEGG databases are reported. The NES of most enriched (blue bars) and most depleted (orange bars) are plotted.

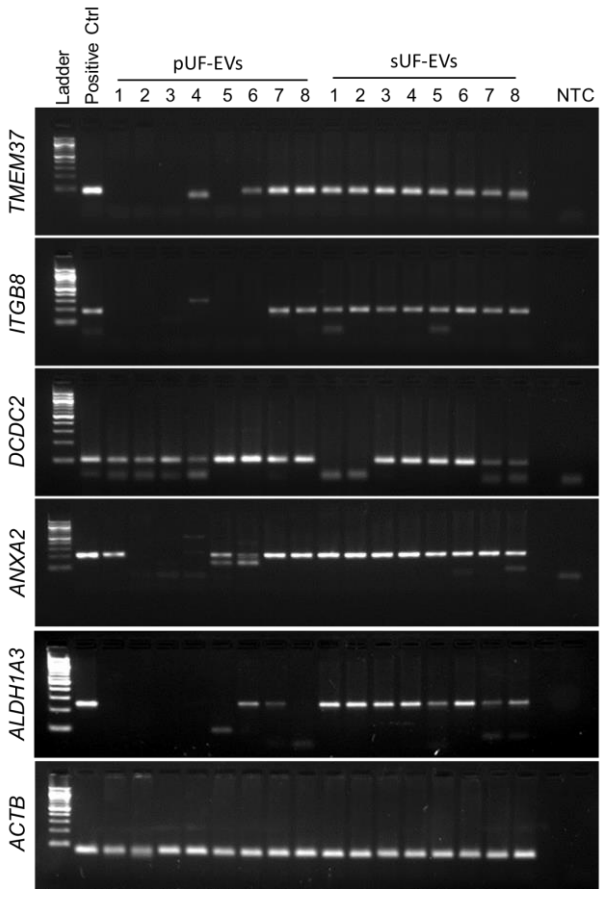
3.14 Validation of the UF-EVs transcriptome as a test for a receptive endometrium in RIF patients

To test whether the UF-EVs transcriptome might be used as a test for a receptive endometrium in RIF patients, we selected genes whose RPKM showed statistically different counts and a *p* value < 0.0001 between RIF patients failing to achieve implantation after euploid ET and fertile patients undergoing PGT-A for genetic diseases with successful pregnancy. This approach identified n=24 genes (*AC021188.1, AC068057.1, AC068137.4, TIMM8AP1, AC083900.1, HEMK1, MIR571, RIOK2, BLVRA, SNORA84, MS4A13, TMEM223, RAB1B, RPL13P5, SPCS2P1, AC004223.2, CHMP6, MRPL12, AC011444.2, MIR27A, AC008750.3, CR392039.4, AP1B1, PQBP1*), based on which we performed two-step cluster analysis between the two groups. The cluster membership analysis was then also applied in the third cohort of RIF patients who achieved implantation after euploid transfer (n=6). Overall, this analysis correctly classified n=33 out of n=35 patients as belonging to the successful implantation group and n=8 out of n=10 patients as belonging to the failed implantation group (i.e. a test sensitivity of 94.3% and a specificity of 80.0%).

3.15 Validation of RNASeq results

Two sets of genes were selected for a validation of the RNA-Seq analyses: i) one set derived from the analysis on the proliferative versus secretory phase (preliminary data), chosen taking into account their biological significance in endometrial receptivity (TMEM37, ITGB8, DCDC2, ANXA2, ALDH1A3); ii) one set derived from the successful versus failed implantation analysis (NPTN-IT1, CECR7, MTRNR2L1, CLU, DUSP1). For the former, the PCR followed by electrophoresis showed a stable presence of *TMEM37, ANXA2, ALDH1A3* and *ITGB8* transcripts in EVs derived from women in secretory phase, while *DCDC2* transcript was detected mostly in proliferative phase (**Figure 27**), confirming RNA-seq results.

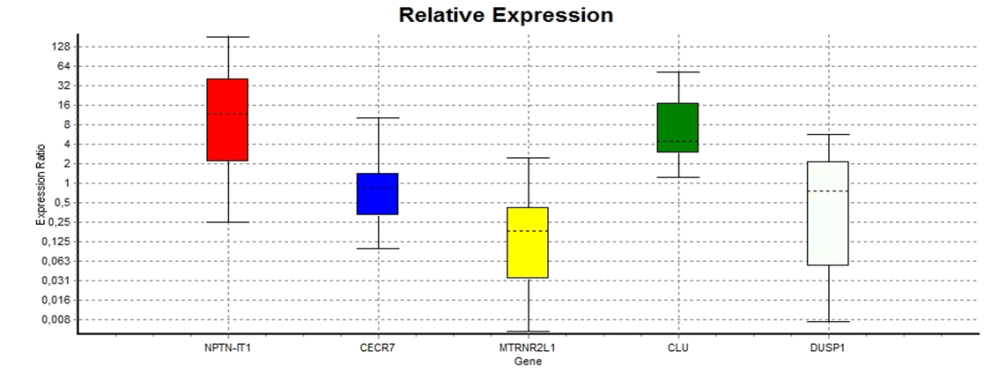
Figure 27. Validation of RNASeq results by qPCR (first set of genes).



NTC: No template controls; p: proliferative phase; s: secretory phase. Gel electrophoresis of PCR products is shown.

Figure 28 shows the relative PCR expression levels for the second set of genes. With both normalization methods used, the expression profile detected by RNA-seq was confirmed by qPCR reflecting the trend described by the RNA-seq, but without statistical significance.

Figure 28. Validation of RNASeq results by qPCR (second set of genes)



A

Gene	Type	Expression	Std. Error	95%C.I.	P(H1)
NPTN-IT1	TRG	8,159	0,955 - 48,333	0,302 - 157,824	0,232
ACTB	REF	0,612			
GAPDH	REF	1,633			
CECR7	TRG	0,831	0,213 - 3,264	0,115 - 9,136	0,890
MTRNR2L1	TRG	0,133	0,013 - 0,893	0,005 - 2,220	0,257
CLU	TRG	6,623	2,172 - 24,944	1,362 - 47,907	0,067
DUSP1	TRG	0,329	0,030 - 2,221	0,010 - 4,951	0,583

B

Gene	Type	Expression	Std. Error	95%C.I.	P(H1)
NPTN-IT1	TRG	4,331	0,772 - 27,613	0,344 - 85,119	0,347
CECR7	TRG	0,441	0,188 - 1,539	0,131 - 2,039	0,244
MTRNR2L1	TRG	0,071	0,008 - 0,367	0,004 - 0,696	0,074
CLU	TRG	3,516	1,137 - 18,746	0,875 - 26,593	0,131
DUSP1	TRG	0,175	0,029 - 0,917	0,011 - 1,133	0,142

Genes were chosen among the significantly different one in the successful versus failed implantation DGE analysis. Upper panel: results REST software analysis for relative expression are shown. Lower panel: expression levels using two normalization methods are shown. A) normalization using ACTB and GAPDH, standard method; B) normalization considering the same total amount of cDNA used for PCR (total ng).

4. DISCUSSION

4.1 Interpretation of the core results

The preliminary phase of the project confirmed that UF-EVs can be efficiently isolated and that their transcriptome is highly correlated to that of concomitantly sampled endometrial biopsies. The project then proceeded and showed that:

i) The transcriptomic content of UF-EVs markedly changes between the nonreceptive (LH+2) and putative receptive (LH+7) phase of the menstrual cycle, with n=2247 differentially 'expressed' genes. This result

supports the rationale of using UF-EVs to identify a transcriptomic signature of endometrial receptivity;

ii) The LH+7 transcriptomic content of ART patients who subsequently achieve implantation partly differs from that of patients who subsequently fail implantation, with n=161 differentially 'expressed' genes, of which n=41 are also significantly regulated in the LH+2 versus LH+7 comparison. The UF-EVs of ART patients also show a difference in size with regards to the outcome, with a higher mean size of EVs being found in patients with failed versus successful implantation. After including the validation cohort and using a propensity score – adjusted multivariate logistic model encompassing the n=5 most significant different genes as well as EVs size, a significant ROC curve for prediction of successful implantation based on UF-EVs was calculated, with an AUC=0.86 (95% CI 0.78-0.94, $p=2.8x10^{-8}$). This result is a promising proof of concept for a novel test for endometrial receptivity based on few selected information contained in UF-EVs;

iii) A subgroup of n=38 genes that are significantly regulated in LH+7 versus LH+2 UF-EVs also belong to the gene set of the ERA® test and to the endometrial tissue 'metasignature' of receptivity based on current literature, providing a potential rationale for their use in a 'liquid biopsy' adaptation of the currently available ERA® test;

iv) When focusing on the subgroup of n=16 patients with a diagnosis of RIF, we found n=167 genes differentially 'expressed' in patients who show persistent implantation failure compared to fertile women undergoing PGT-A for genetic diseases. After including the validation cohort and using a two-step cluster analysis encompassing the n=24 most significant genes in the comparison, we tested whether this model could correctly classify RIF patients as successfully or unsuccessfllly treated by PGT-A and found a test sensitivity of 94.3% and a specificity of 80.0%. Despite the small group of samples, this preliminary result reassures about the potential use of UF-EVs as a diagnostic tool for endometrial receptivity also in the subgroup of RIF patients.

4.2 Feasibility of the approach: technical considerations

Our study represents the largest transcriptomic study on human EVs so far. From a technical perspective, we can observe that:

-Of a total of n=129 UF-EVs samples, n=4 in the fertile volunteers group, n=7 in the PGT-A cohort of Aim 2 and n=9 in the PGT-A validation cohort of Aim 3 showed unsuccessful quality for RNAseq library preparation (n=20/129, 15.5%). This overall good output allowed as to avoid any pooling of the samples as opposed to other studies on EVs (Campoy et al., 2016), and is thus a strength of the study. Nonetheless, envisioning a patient-based personalized clinical use of UF-EVs transcriptome as a diagnostic tool, further optimization of the technique represents one of current goals. In this context, ultracentrifugation has recently been confirmed to be the most efficient method for UF-EVs isolation for the purpose of RNA-seq compared to both ultrafiltration and polymer-based precipitation (Li et al., 2021). Unfortunately, however, neither this publication nor a recent study recurring to transcriptomic analysis of UF-EVs for the early diagnosis of ovarian cancer provide any information about the proportion of low quality samples (Skryabin et al., 2022);

- in UF-EV samples, the great majority of transcriptomic reads seem to be distributed on the same consistent genes, as demonstrated by the limited changes in the number of 'expressed' genes when different cut-offs in CPM to define 'expression' were applied (14228 in the fertile volunteers' group, 14593 in the ART PGT-A population and 17191 in the RIF subpopulation).

4.3 Origin of uterine fluid-derived extracellular vesicles

The very significant correlation that we observed between the transcriptome of paired endometrial biopsies and of UF-EVs might to some extent suggest that these are more likely to be microvesicles or apoptotic bodies, both including portions of the cytoplasm of the origin cell. Similarly, the mean bigger size of UF-EVs retrieved from patients who subsequently fail implantation compared to patients with successful pregnancy could be speculated as a reflection of a higher proportion of apoptotic bodies in this subgroup of patients as these are known to be the largest type of EVs. However, assigning an extracellular vesicle to a particular biogenesis pathway remains speculative and current guidelines of the International Society for Extracellular Vesicles suggest not to operate such a distinction 'unless, for example, the EV is caught in the act of release by live imaging techniques or labelled with

specific markers' (Théry et al., 2018). We thus cannot distinguish the relative contribution of exosomes, microvesicles and apoptotic bodies to our findings, and neither their origin. Further insights regarding the origin of UF-EVs in terms of parental cells could however be drawn from a systematic comparison with the results from single-cell RNAseq studies of the endometrium, which is therefore reported below.

4.4 Molecular mechanisms of implantation reflected by the UF-EVs transcriptome

As the top upregulated genes in UF-EVs during the window of implantation we found MT1G - also known to be upregulated in the secretory endometrium (Ace et al., 2004), PAEP or glycodelin – involved in implantation and Natural Killer cells regulation (Lee et al., 2019) and CALB – implicated in endometrial gland epithelial function (Russell et al., 2014). Among the top significant differentially expressed genes in the successful versus failed implantation groups are: *DUSP1*, a protein-coding gene for a phosphatase that can dephosphorylate MAP kinase MAPK1/ERK2 and contributes to cellular response to environmental stress as well as to the inhibition of cellular proliferation. Recently, in vitro studies on human primary endometrial cells have implicated DUSP1 levels to be increased in a time-dependent manner during induction of decidualization (Xu et al., 2021); *CLU*, encoding for clusterin, a secreted glycoprotein that has also been previously suggested to play an immunosuppressive role during the receptive period, by inhibiting membrane attack of activated complement proteins and interacting with IgG at endometrial epithelial level (Tapia et al., 2008); DPP4, a known regulator of T cell function involved in implantation: expressed by the endometrial epithelium, it facilitates both blastocyst adhesion via fibronectin binding and trophoblast invasion through pericellular proteolysis of the extracellular matrix (Shimomura et al., 2006; Dolanbay et al., 2016). Of note, both *CLU* gene transcript and DPP4 were positively enriched both in the LH+7 versus LH+2 and in the successful versus failed implantation comparisons. In contrast, among transcripts 'selectively expressed' only in one of the two groups of patients based on successful versus failed implantation, 6/19 are novel lncRNAs, suggesting an additional regulatory function of RNA transported by EVs. For example, *CERC7*, detected only in UF-EVs of women failing to achieve pregnancy, regulates immune cell differentiation through the modulation of CTLA4 expression by targeting miR-429 (Yao et al., 2017). CTLA4, expressed by activated Foxp3þ Treg cells, participates in

the balance between effectors (i.e. innate immunity and NK cells, B cells, T helper [Th] 1 and Th17 immunity) and regulators (Th2 cells, regulatory T cells) (Vignali et al., 2008), which is essential for establishment of a pregnancy.

When focusing on the intersection genes that are significantly regulated in LH+7 versus LH+2 UF-EVs and also belong to the gene set of the ERA® test and to the endometrial tissue 'metasignature' of receptivity based on current literature, we confirmed *MT1G* and *PAEP* or glycodelin – whose functions in implantation are described above. Of note, belonging to this intersection of relevant genes is also CRABP2 (encoding for cellular retinoic acid-binding protein 2), that was recently described among the n=33 genes identified as prioritized in the regulation of the window of implantation based on WGCNA (Sebastian-Leon et al., 2021).

4.5 Molecular mechanisms of RIF reflected by the UF-EVs transcriptome

In the subgroup analysis on RIF patients, among the top significant differentially expressed genes between patients with persistent implantation failure and successful outcomes are: DUSP1, already described above, found to be increased in a time-dependent manner during induction of decidualization (Xu et al., 2021). Thus, it may be hypothesized that increased 'expression' of *DUSP1* in UF-EVs of women with a negative outcome might indicate a dyssynchronous endometrium; *FMO2*, which encodes a member of the flavin containing dimethylaniline monoxygenase family - that catalyzes the oxygenation of a wide range of foreign chemicals. However, most humans have a truncated form of the FMO2 protein without any functional activity (Phillips and Shephard, 2008); *MYO7B*, which encodes myosin VIIB, a protein mainly expressed in brush border microvilli of epithelial cells in the intestines and kidneys. Myosin VIIB is involved in anchoring membrane surface receptors to the actin cytoskeleton (Henn and De La Cruz, 2005) and in proper function of transporting epithelia (Chen et al., 2001). Increased levels of myosin VIIB have been found in nephrotic rats and related to increased podocyte vesicle transport (Tojo et al., 2017). To date, the potential role of Myosin VIIB in the endometrium remains however unknown. Of note, *MYO7B* and *DUSP1* are the two genes with the highest statistically significant difference also between RIF patients with persistent implantation failure versus RIF patients who are successfully managed by PGT-A and euploid blastocyst transfer. Among genes enriched in both DGE comparisons focusing on RIF patients

are also C-X-C motif chemokine ligand (CXCL)1, CXCL2, CXCL8, and prostaglandin-endoperoxide synthase 2 (PTGS2), that belong to both the IL-10 and IL-17 pathway.

Thus, our study targets the presence in UF-EVs of possible RIF predictors that belong to the immunological panel, a relevant domain where an intimate relationship between pro-inflammatory and anti-inflammatory components such as the IL-10 and IL-17 pathways has been described for successful blastocyst implantation (Zhao et al., 2021).

4.6 Relationships with results from single-cell RNAseq studies

In the last two years, the high resolution of single-cell RNA-sequencing has been applied to the study of the human endometrium identifying unciliated epithelium, stromal fibroblasts (the two major contributing cell types), ciliated epithelium, endothelium, macrophages and lymphocytes as the main cellular components, with the presence of an additional population of smooth muscle cells (Wang et al., 2020; Garcia-Alonso et al., 2021; Zhang et al., 2022). We have thus systematically searched the available literature for data that could integrate our results in terms of UF-EVs origin based on the available single-cell transcriptomic atlas of the human endometrium. Based on this, during the transition to the window of implantation *PAEP* (progestagen associated endometrial protein), *GPX3* (glutathione peroxidase 3) and *MAOA* –transcripts that are among the n=38 intersection genes comprising our study, the ERA test and the metasignature by Altmäe et al. (2017) - are upregulated in the unciliated epithelium. More specifically, their upreglation occurs first in the glandular and then in the luminal epithelium. Also *DPP4,* again belonging to our n=38 intersection genes, is shown to be activated in both glandular and luminal endometrial epithelium (Wang et al., 2020). *FOS* - a transcript we found to be upregulated both in the LH+7 vs LH+2 DGE analysis and in the DGE analysis comparing RIF patients with persistent negative outcome vs genetic controls achieving pregnancy- is also enriched in the unciliated epithelium. Similarly, metallothionein I genes (*MT1E, MT1F, MT1G, MT1X*), also belonging to the n=38 intersection genes in our study, are described to be a key regulatory module associated with the window of implantation in endometrial unciliated epithelium (Wang et al., 2020). Among genes belonging to our n=38 intersection genes, *DKK1* and *IL15* are instead upregulated in stromal fibroblasts and *GNLY* was described to be expressed by endometrial lymphocytes (Wang et al., 2020), similar to *IL12R* which

we found to be enriched in our DGE analysis between patients with successful vs failed implantation. Also CYBB – enriched in our LH+7 versus LH+2 and in our successful versus failed implantation DGE analyses – has been reported as a discriminatory gene that is over-expressed by endometrial macrophages (Wang et al., 2020). Taken together these findings suggest that uterine fluid-derived extracellular vesicles encompass the complexity of the endometrial tissue, also reflecting the function of immune cells.

4.7 Relationships with results from UF-EVs proteomic studies

Due to considerable technical difficulties, a proteomic study of UF-derived exosomes based on high-resolution quantitative mass spectrometry from fertile women has only recently been reported, and has identified a total of n=200 and n=206 proteins in proliferative and secretory phase UF-derived exosomes respectively, of which 64 proteins (30 upregulated, fold change >1.5, and 34 uniquely identified) were found in higher abundance in the secretory phase (Rai et al., 2021). Secretory phase UF-exosomes were enriched in invasion-related proteins and among them is WAP four-disulfide core domain protein 2, which is enriched in UF-EVs both at a proteomic level and at *WFDC2* transcript level in our successful vs failed implantation DGE analysis. In the proteomic study by Rai et al., (2021) UF-derived exosomes collected from infertile women were also analyzed and showed 79 down-regulated proteins compared to fertile women. In this context, Cystatin A and small proline-rich protein 3 showed a significant depletion at protein level and at transcript (*CSTA* and *SPRR3* respectively) levels in our our failed vs successful implantation DGE analyses (and more specifically both in the one including the total population and in the one including on RIF patients with negative outcome compared to fertile genetic controls with positive outcome). However, due to the very preliminary nature of this first proteomic study on UF-EVs, significant conclusions inherent to the relationship between UF-EVs transcriptional/proteic content seem currently unwarranted.

4.8 Limitations and current advancements of the project

At the present stage, the main limitation of this study is the timeshift between the uterine fluid collection and the embryo transfer cycle. We chose this approach as the most appropriate in relation to the current scarcity of data regarding the safety of UF retrieval in the same cycle of an embryo transfer, currently comprising three small

case-control studies all using direct UF aspiration rather than flushing (Boomsma et al., 2009; van der Gaast et al., 2003; Hou et al., 2022). Our perspectives are thus to validate our results and proceed to increasing data on the safety of performing fluid sampling procedure on the same cycle of the embryo transfer if necessary. For example, same-cycle 'liquid biopsy' might become advisable in patients with an "unexpected" negative outcome despite the diagnosis of a receptive endometrium based on UF-EVs of the preceeding cycle. As a further limitation, membrane-bound RNAs—rather than the internal RNA content of EVs—might have to some extent contributed to our RNA-seq results. This is due to the fact that - to enhance the consistency and clinical feasibility of our methods (O'Brien et al., 2020) - we did not treat pelleted EVs with RNase, potentially resulting in incomplete separation of RNAs associated with EV membranes. Nonetheless, we recognize this as a minor limitation given that the core of this project is the quest for biomarkers – rather than for physio-pathological pathways.

In line with these perspectives, we are currently extending the project on the use of UF as a 'liquid biopsy' for endometrial receptivity with a further observational phase (supported by a grant from the Italian Ministry of Health, *Bando Ricerca Finalizzata 2019*) in which we are recruiting n=309 infertile patients undergoing their first frozen-thawed embryo transfer attempt and obtaining their LH+7 UF for concomitant:

a) UF-EVs transcriptomic analysis and validation of our current results
b) UF microbiome analysis

Ultrasound features of the uterine walls and endometrium are also being recorded, for their integration in predictive models related to embryo transfer outcomes and based on multi-modal high throughput techniques.

4.9 Therapeutic potentials

As an important tool for intercellular communication, exosomes have been proposed as ideal candidates for therapeutic applications (Yanez-Mo et al., 2015). As a matter of fact, while cell therapies present some limitations in terms of safety, extracellular vesicles are less immunogenic and more biologically stable compared to parent cells (Liao et al., 2021). In a recent *in vitro* study, UF-derived exosomes have been used as a delivery platform for hCG, suggesting a beneficial effect for treatment with hCG-

loaded exosomes compared to hCG alone in a model mimicking endometrial receptivity (Hajipour et al., 2021).

In this context, the recent collection of uterine fluid from endometrial organoids might represent an advancement in the field of endometrial therapeutics, as it represents a novel source of uterine fluid (Simintiras et al., 2021) and potentially of uterine-fluid derived extracellular vesicles. As a matter of fact, for clinical applications, it is methods to increase and standardize uterine fluid and EVs production would be of utmost relevance (Liao et al., 2021). However, whether organoids will indeed become a useful tool in the advancement of uterine fluid studies remains to be proved.

5.0 MATERIALS AND METHODS

Declaration of the author: Please note that Materials and Methods of the project have already been published in a recent article of which I am a co-author (Giacomini et al., 2021), with the exception of Statistical Analyses used for validation of results.

5.1 Subjects

The participants for Aim 1 were cycling women with proven natural fertility (i.e. one or more previous live births resulting from natural conception) attending the Gynecology & Obstetrics Unit of IRCCS Ospedale San Raffaele for office consultation, with a body mass index (BMI) <30, aged 20 to 40 years, nonsmokers and taking no medication. Uterine fluid samples were collected at the mid-secretory, receptive phase 7 days after the LH surge (LH+7) and compared to nonreceptive LH+2 samples.

The participants for Aim 2 were women undergoing ovarian stimulation and their first euploid embryo transfer in an ART cycle with PGT-A. Inclusion criteria comprised age ≤38 years at time of liquid biopsy, having regular menstrual cycles of 25–35 days. ART treatments were started after at least one year of trying to conceive spontaneously; all included women underwent routine fertility investigations and had a clinical indication for ART and PGT-A treatments.

Full description of the two populations included is reported below (**Table 10**), while **Table 11** shows the comparison between patients with failed or successful implantation within the ART group.

Table 10. Descriptive characteristics of subjects who partecipated in the experimental study (Aim 1, Aim 2)

(A) Descriptive characteristics of fertile controls (n=14)

Characteristics	Mean ± SD or *n*	Range or %
Age (years)	36.2 ± 5.5	23 - 42
BMI (Kg/m^2)	21.9 ± 1.6	19.8 - 25.7
Parity		
Primiparous	9	64.3
Multiparous	5	35.7
Previous spontaneous abortions		
None	14	100%

(B) Descriptive characteristics of patients undergoing PGT-A (n=49)

Characteristics	Mean ± SD or n	Range or %
Age (years)	35.6 ± 3.6	25 - 43
BMI (Kg/m^2)	21.6 ± 2.5	16.9 - 28
Basal FSH (IU/L)	7.0 ± 2.3	1.3 - 14.0
Basal AMH (ng/mL)	2.9 ± 1.9	0.7 - 8.6
AFC (n)	13.1 ± 5.2	5.0 - 35.0
Previous failed embryo transfers (n)	1.9 ± 2.8	0 - 12
Parity		
Nulliparous	44	89.8
Primiparous	4	8.2
Multiparous	1	2.0
Previous spontaneous abortions		
None	27	55.1
One	10	20.4
Two or more	12	24.5
Indication to PGT-A		
Monogenic disease	15	30.6
Recurrent Implantation Failure	14	28.6
Chromosomal abnormality	10	20.4
Unexplained Recurrent Abortion	5	10.2
Other	5	10.2
Not Pregnant after ART cycle	27	55.1
Pregnant after ART cycle	22	44.9

Table 11. Characteristics of PGT-A patients (Aim 2) grouped by implantation outcome

Characteristics	Successful implantation (n=19)	Failed implantation (n=23)	*p*
Age (years)	35.6 ± 3.2	35.6 ± 4.2	1
BMI (Kg/m^2)	21.7 ± 2.4	21.6 ± 2.8	0.98
Basal FSH (IU/L)	7.4 ± 2.0	6.2 ± 2.2	0.09
Basal AMH (ng/mL)	2.8 ± 1.7	3.1 ± 2.2	0.99
AFC (n)	12.3 ± 3.8	13.8 ± 6.4	0.48
Previous failed embryo transfers (n)	1.4 ± 2.0	2.3 ± 2.8	0.35
Stimulation protocol data			
Gn-RH antagonist protocol	19 (100)	23 (100)	1.00§
Total dose of rFSH/hMG	2386 ± 1163	2809 ± 1235	0.48
Number of blastocysts transferred (n)	1 (100)	1 (100)	1.00§
Indication to PGT			
Monogenic disease	7	5	
Recurrent Implantation Failure	5	8	
Chromosomal abnormality	4	6	0.66§
Other	2	1	
Unexplained Recurrent Abortion	1	3	

Data are expressed as mean ± SD or number (percentage), p values are reported for Wilcoxon test or §Pearson Chi Square

Subjects included for Aim 3 (n=52; validation cohort) were a further group of patients undergoing ART and PGT-A and fulfilling the same inclusion criteria as those included in Aim 2 (**Table 11**).

Table 12. Descriptive characteristics of patients undergoing PGT-A (n=52)who participated in the validation cohort (Aim 3)

Characteristics	Mean ± SD or n	Range or %
Age (years)	37.0 ± 3.3	25 - 40
BMI (Kg/m^2)	19.9 ± 2.0	18 - 29
Basal FSH (IU/L)	6.1 ± 2.2	1.3 – 14.0
Basal AMH (ng/mL)	3.6 ± 2.2	0.7 – 10.5
AFC (n)	13.2 ± 4.8	4.0 – 35.0
Previous failed embryo transfers (n)	0.61 ± 1.8	0 - 9
Parity		
Nulliparous	43	82.7
Primiparous	9	17.3
Multiparous	0	0.0
Previous spontaneous abortions		
None	30	57.7
One	5	9.6
Two or more	17	32.7
Indication to PGT-A		
Monogenic disease	13	25
Recurrent Implantation Failure	5	9.6
Chromosomal abnormality	14	26.9
Unexplained Recurrent Abortion	11	21.1
Other	9	17.3
Not Pregnant after ART cycle	29	55.7
Pregnant after ART cycle	23	44.2

5.2 ART procedures

ART procedures were performed according to standard protocols. Women underwent controlled ovarian hyperstimulation and Intra Cytoplasmic Sperm Injection (ICSI) was performed 2–4 hours after follicle aspiration. Blastocysts obtained were biopsied on day 5-7 of development and frozen. Euploid frozen-thawed embryos were transferred on day LH+7 of ultrasound and LH surge monitored natural cycles. In these women, uterine fluid samples were collected on day 7 after the LH surge of the

natural cycle (LH+7) immediately preceeding that of the embryo transfer. Clinical outcomes considered was the clinical pregnancy rate, defined as the number of gestational sacs confirmed by ultrasound performed at week 6 of gestation.

5.3 Uterine fluid collection

Uterine fluid samples were collected by lavage of the uterine cavity with 2,5ml of sterile saline through a Cook® Silicone Balloon HSG (histerosonography) catheter, after removal of cervical/vaginal discharge or mucus from the external *os* and with the silicon balloon inflated at the internal cervical *os* in order to avoid contamination by the cervical and vaginal environment. To the same purpose, upon withdrawal of the catheter and before uterine fluid collection in a 15 ml Falcon® tube, the outside was wiped off with a sterile gauze.

5.4 Isolation of UF-EVs

Uterine fluid was freshly processed and subjected to centrifugation 1200 x g for 10 min at 4°C according to the standard procedure to eliminate cell debris. The supernatant was separated and stored at 4°C while mucus within the sample was retrieved, suspended in PBS (pH 7.6, 1:1) (Euroclone S.p.a, Pero, MI, Italy), subjected to vortexing for 3-5 min and then to centrifugation (800 x g for 5 min) for EVs release. The procedure was repeated several times until the mucus was fully dissociated (i.e. no longer evident upon observation) (Ng et al., 2013). The supernatant from mucus dissociation was then added to the liquid fraction of UF and subjected to sequential centrifugations to isolate EVs, following the protocol described by Théry et al. (2018) and modified as reported below. Supernatants were centrifuged (300 x g for 10 min) to remove any remaining cells, and then to a second centrifugation (2000 x g for 20 min) to remove dead cells and debris. The supernatants were diluted in 1:1 or more PBS and transferred to sterilized 1.5-ml polypropylene tubes (Beckman Coulter Inc., Brea, CA, USA, item no. 357448) and then ultra-centrifuged at 110 000 x g for 2 h at 4°C in a TLA-55 rotor (Beckman Coulter Inc.) using an Optima TLX centrifuge (Beckman Coulter Inc.) to pellet EVs. The pellets were resuspended in PBS, again centrifuged at 110 000 x g for 90 min and finally resuspended in a total volume of 110 ml of PBS (filtered three times with 0.1 mm filter). The final UF-EV resuspension was split in two aliquots, 100 µl for RNA

Next Generation Sequencing (NGS) and 10 µl for nanoparticle tracking analysis (NTA). Aliquots of UF-EVs were preserved at -80°C until processing.

5.5 Transmission electron microscopy

Preparation for TEM analysis was done using the method described by Théry et al. Briefly, freshly purified UF- EVs were absorbed on glow discharged carbon coated formvar copper grids, washed with water, contrasted with 2% uranyl acetate and air-dried. Grids were observed with a Zeiss LEO 512 transmission electron microscope. Images were acquired by a 2k x 2k bottom-mounted slow-scan Proscan camera controlled by EsivisionPro 3.2 software.

5.6 Western blotting

For western blotting analysis, endometrial tissues were lysed in lysis buffer (50 mM Hepes, pH 7.5, 150 mM NaCl, 1 mM EDTA, 2.5 mM EGTA, 10% glycerol, 0.1% Tween20, 10 mM β-glycerophosphate) with 1% Protease inhibitor cocktail (Merck KGaA, Darmstadt, Germany) while isolated EVs (20 µg as evaluated by Bradford assay and by NanoDrop8000 measurement) were lysed directly in reducing Laemmli buffer [0.25 M Tris–HCl (pH 6.8), 40% glycerol, 8% SDS, 5% 2-mercaptoethanol and 0.04% bromophenol blue] and boiled for 5 minutes at 95 °C. For tetraspanins detection, 10 µg of isolated EVs were lysed in non-reducing sample buffer (without 2-mercaptoethanol). For the detection of EV marker, total proteins were resolved by SDS-PAGE (SDS-polyacrylamide gel electrophoresis), electrophoretically transferred to polyvinylidene fluoride membranes, blocked in 5% non-fat powdered milk in TBS-T (0.5% Tween-20) and the membranes were incubated with the following antibodies: anti-CD63 (1:1000; BD Pharmingen, #556019, San Jose, CA, USA), anti-CD9 (1:1000, BD Pharmingen, #555370, San Jose, CA, USA), anti-TSG101 (1:500, Novus Bio, #NB200-112, Littleton, CO, USA), anti-calnexin (1:1000, Sigma, #C7617, Sigma/Aldrich, St. Louis, MO, USA), anti-TOM20 (1:1000; Abcam, #ab78547;Cambridge, MA) and anti-TIM44 (1:1000; BD Transduction Laboratories, Franklin Lakes, New Jersey, USA). Protein bands were detected using X-ray film and enhanced chemiluminescence reagent (ECL, Amersham, Buckinghamshire, UK).

5.7 Nanoparticle tracking analysis

Extracellular vesicles amount and size distribution were analyzed by NTA by laser light scattering performed using a NanoSight LM10-HS microscope (NanoSight Ltd., Amesbury, UK), as previously described (Zarovni et al. 2015). The NanoSight system was calibrated with polystyrene latex microbeads of 50, 100, and 200 nm prior to analysis and auto settings were used for blur, minimum track length and minimum expected particle size. UF-EVs (3 µl) were diluted 1–300 with PBS, filtered three times with a 0.1 mm filter and concentrations adjusted, if necessary, in order to specifically fit the optimal working range (20–120 particles\per frame) of the instrument. Five 30-s videos were recorded of each sample with camera level set at 13 and detection threshold set at 5. Videos were analyzed with NTA software (NTA 3.2 Dev Build 3.2.16, NanoSight Ltd., Amesbury, UK) to determine the concentration and size of measured particles with corresponding standard error.

5.8 RNA isolation, sequencing library construction and batch effect management

Total RNA content of the extracellular vesicles was isolated with RNeasy micro kit (Qiagen, Hilden, Germania) according to the manufacturer's protocol. Total RNA sequencing (whole-transcriptome analysis) was carried out with TruSeq Stranded total RNA kit (Illumina, San Diego, California, USA) in order to analyze the most complete picture of the transcriptome (including coding and multiple forms of noncoding RNA). cDNA libraries were prepared using the SMART-SeqVR v4 UltraVR Low Input RNA Kit (Takara Bio Inc., Otsu, Shiga, Japan) according to the manufacturer’s instructions. Sequencing was performed on both Illumina NextSeq 500 and Illumina NovaSeq 6000 to obtain 15 million single end reads per sample. RNAseq analysis was performed in three batches, each containing samples derived both from fertile and from both groups of ART patients divided based on the outcome (i.e. successful and failed implantation). In addition, since the effect of different library preparation batches was quite evident, the expression values were adjusted for the batch effect in a regression model.

5.9 RNA Sequencing

The raw reads produced from sequencing were trimmed using Trimmomatic, version 0.32, to remove adapters and to exclude low-quality reads from the analysis. The remaining reads were then aligned to the human genome GRCh38, annotated according to Gencode basic annotations version 27, using STAR, version 2.5.3a. Reads were assigned to the corresponding genomic features using featureCounts via Rsubread package. Quality of sequencing and alignment was assessed using FastQC and MultiQC tools. Differences in quantity and quality of the genomic starting material resulted in different depths of sequencing, thus producing RNA libraries of different sizes. For this reason, RNAseq data were normalized for different library sizes by calculating the Counts Per Million (CPM) values, and also for different gene lengths by calculating the Read Per Kilobase Million (RPKM) values. CPM was used to define the 'expressed' genes (Chen *et al.*, 2016), while RPKM were considered for exploratory two-step cluster analysis. 'Expressed' genes were defined as those genes showing at least 1 CPM read in a selected number of samples based on the size of the smaller group in a comparison (Chen *et al.*, 2016). CPM values were also used for comparisons and correlation analysis among groups of samples. Seven of the n=49 samples obtained from in the study cohort (Aim 2) were not included in the analysis due to a low number of total read counts (<2 million reads per sample) and the remaining n=42 were grouped by implantation outcome (with n=19 patients defined by successful implantation and n=23 defined by failed implantation); similarly, n=9 of the n=52 samples from the validation cohort (Aim 3) were also excluded due to low quality (<2 million reads per sample). Analysis of biotype composition of each sample was performed according to Gencode biotype definition (https://www.gencodegenes.org/pages/biotypes.html).

5.10 Validation of RNAseq results by PCR

For validation of RNA-seq analyses, two sets of genes (ANXA2, ITGB8, ALDH1A3, DCDC2, TMEM37 and NPTN-IT1, CECR7, MTRNR2L1, CLU, DUSP1) were validated in UF-EVs from two subsequent cohorts of women: n=16 volunteers (n=8 in the proliferative and n=8 in the secretory phase of the menstrual cycle) and n=10 ART patients (who achieved implantation, n=5; and who failed to achieve implantation, n=5). Since the RNA content of EVs was expected to be limited, UF-EVs were subjected to RNA extraction, reverse transcription and whole transcriptome

amplification using REPLI-g Cell whole genome amplification (WGA) & whole transcriptome amplification (WTA) Kit (Qiagen) according to the manufacturer's instructions. Briefly, lysis buffer was added to 13 µl of UF-EVs resuspended in PBS and heated at 95 °C for 5 min to lyse and release the UF-EV contents. Lysed UF-EVs were used for WTA of total RNA. Genomic DNA was removed, cDNA was synthesized and subjected to ligation and amplification steps. The Real Time (RT)-reaction was included in the kit which used T-Script reverse transcriptase combined with random and oligo-dT primers. REPLI-g SensiPhi DNA Polymerase with high proofreading was used for isothermal amplification of cDNA. The DNA derived from amplification of cDNA was used to perform Polymerase Chain Reaction (PCR) protocol. In order to quantify and perform the PCR amplification using the same amount of cDNAs from each sample, purification of amplified DNA by LiCl/EtOH precipitation was performed, according to supplementary protocol instructions of REPLI-g kit. In order to ensure optimal normalization of results on total amplified cDNA quantity, 10 ng of the amplified cDNA of UF-EVs were subjected to PCR using SYBR Green Master Mix (Applied Biosystems, Foster City, CA, USA). The PCR amplification was carried out for 40 cycle with annealing temperature of 60°C, using: *ANXA2* (Fwd 5'-GGACGCGAGATAAGGTCCTG-3', Rev 5'-ACCATTTCTGGACGCTCAGG-3'); *ITGB8* (Fwd 5'-TGGCACTTCAGATGCAGTGT-3', Rev 5'- GCTCCATGTTGAGTTGTGCG-3'); *ALDH1A3* (Fwd 5'-CCTGAGTATTTCACTGGCAGGT-3', Rev 5'-TCTGTTACGGGCCCTCATTT-3') *DCDC2* (Fwd 5'-AAGACAGAGCAAGGCGTTCA-3', Rev 5'-AAGCATGCAAGCCTGAAAGC-3'); *ACTB* (Fwd 5'-GGCACCCAGCACAATGAAG, Rev 5'-CCGATCCACACGGAGTACTTG-3'), TMEM37 (Hs_TMEM37_2_SG QuantiTect Primer Assay #QT01530361) and *CECR7* (Fwd 5'-GGTCTGTGCTGAGGTGGATT-3', Rev 5' ACCATAGGGCGGTTTTTCTC-3');*DUSP1* (Fwd 5'-ACCACCACCGTGTTCAACTTC-3', Rev 5'-*TGGGAGAGGTCGTAATGGGG-3'); GAPDH (Fwd 5'-CTGGGCTACACTGAGCACC-3', Rev 5' AAGTGGTCGTTGAGGGCAATG-3'); MTRNR2L1 (Fwd* 5'-CTCGATGTTGGATCAGGACA-3', Rev 5'-CCTGGATTACTCCGGTCTGA-3'); *CLU* (Fwd 5'-ACATTTGGTGCCCAGAAGTC-3', Rev 5'-TGGGATGCTGAGCTCTTACA-3') primers and NPTN-1 TaqMan Assay (Hs00288947_s1, Applied Byosystems). Ten microliters of PCR products were loaded on a 2% agarose gel and stained with EtBr, using cDNA from endometrial cells as positive control. No template controls were also added. For the second set of genes, belonging to the ones with a significant differential 'expression' in the DGE analysis of women with successful versus failed implantation, results were analyzed using REST software which also considers the efficiency of the amplification reaction, for comparison of the expression of the UF-EV transcripts obtained from the two groups of patients. In this case, we used ACTB and GAPDH

as reference genes because they have similar CVs (coefficents of variation) between the different groups and the validation was performed using two normalization methods: a) normalizing using ACTB and GAPDH (standard method); b) considering the same total amount of cDNA used for PCR (total ng).

5.11 Differential Gene Expression analysis

Gene expression read counts were exported and analyzed in *R* to identify differential expressed genes (DEGs), using the DESeq Bioconductor library version 1.22.1, with no assumption about prior distribution (option betaPrior=FALSE) and applying independent filtering to the result. Several comparisons were performed. In each comparison, only genes identified as 'expressed' in the corresponding group of samples were considered. In all DGE comparisons, significance was defined using the cut-off suggested by SEQC consortium, that selects genes showing at the same time a raw P-value < 0.01 and a logFC greater than 1 or lower than -1 (SEQC-MAQC-III Consortium, https://doi.org/10.1038/nbt.2957). DGE analysis was performed starting from raw counts, since DESeq2, edgeR and Limma models apply an embedded normalization step.

5.12 Gene set enrichment analysis and over-representation analysis

Over-representation analyses for Gene Ontology (GO) terms and biological pathways (KEGG, Reactome and WikiPathway) were performed with the g:Profiler web tool (https://biit.cs.ut.ee/gprofiler). The obtained results were corrected for multiple testing by using the g:Profiler tailor-made algorithm g:SCS (Raudvere et al., 2019). The Webgestalt (http://www.webgestalt.org) platform was used for GSEA. To compare RNA-seq results with the transcriptional profile of the gene-set evaluated in the ERA® test (Díaz-Gimeno et al., 2011), results from the fUF-EVs LH + 7 versus fUF-EVs LH + 2 comparison and from the UF-EVs 'pregnant' versus UF-EVs 'not pregnant' comparison were subjected to a Preranked GSEA using GSEA software. NES was used to indicate the strength of the enrichment in GSEA. With NES being normalized for the number of genes in the considered pathway, the greater the NES absolute value, the stronger the enrichment. The GSEA leading edge, defined as the

subset of genes that contributed most to the enrichment score, was also presented (Mootha, et al. 2003; Subramanian et al. 2005).

5.13 Statistical analyses

Normality was assessed with the Shapiro–Wilk test. Paired/unpaired Student's t-test, non-parametric Wilcoxon and Kruskal–Wallis tests followed by the Dunn's Multiple comparison's tests were used as appropriate. The Chi-square test was used for the evaluation of gene transcripts detected only in UF-EVs of women who achieved a successful versus a failed pregnancy outcome. Logistic regression analysis and principal component analysis were used to assess the relationship between clinical outcomes and EV- RNA content, number or diameter. Various models were estimated, including EVs RNA signature and EVs size and number in relation to pregnancy outcomes. Significance of parameters was evaluated using the likelihood-ratio test and alternative ways of combining the significant variables were evaluated in order to obtain an optimal profile. Propensity score was also calculated and assigned to each sample based on the RNA expression of chosen genes and EV dimensions. ROC curve analysis and two-step cluster analysis were used to test the ability of the UF-EVs transcriptome to correctly classify samples as belonging to successful versus failed implantation groups. All results were expressed as mean ± SD or as median (interquartile range, IQR). Analyses and relative graphs were made in Prism 5.0 (GraphPad Software Inc., La Jolla, CA, USA) or JASP 0.11.1.0 (JASP Team, T. 2017 JASP, https://jasp-stats.org/) or with SPSS Statistics for Windows, Version 21 (IBM Corp., Armonk, NY, USA).

6. REFERENCES

Ace CI, Okulicz WC (2004) Microarray profiling of progesterone-regulated endometrial genes during the rhesus monkey secretory phase. *Reprod Biol Endocrinol* 2:54

Altmäe S, Koel M, Võsa U, Adler P, Suhorutšenko M, Laisk-Podar T, Kukushkina V, Saare M, Velthut-Meikas A, Krjutškov K, et al (2017) Meta-signature of human endometrial receptivity: a meta-analysis and validation study of transcriptomic biomarkers. *Sci Rep* 7(1):10077

Ata B, Kalafat E, Somigliana E (2021) A new definition of recurrent implantation failure on the basis of anticipated blastocyst aneuploidy rates across female age. *Fertil Steril* 116(5):1320-1327

Bentin-Ley U, Lindhard A, Ravn V, Islin H, Sørensen S (2011) Glycodelin in endometrial flushing fluid and endometrial biopsies from infertile and fertile women. *Eur J Obstet Gynecol Reprod Biol* 156(1):60-6

Blockeel C, Schutyser V, De Vos A, Verpoest W, De Vos M, Staessen C, Haentjens P, Van der Elst J, Devroey P (2008) Prospectively randomized controlled trial of PGS in IVF/ICSI patients with poor implantation. *Reprod Biomed Online* 17(6):848-54

Boomsma CM, Kavelaars A, Eijkemans MJ, Amarouchi K, Teklenburg G, Gutknecht D, Fauser BJ, Heijnen CJ, Macklon NS (2009) Cytokine profiling in endometrial secretions: a non-invasive window on endometrial receptivity. *Reprod Biomed Online* 18(1):85-94

Braga DPAF, Borges E Jr, Godoy AT, Montani DA, Setti AS, Zanetti BF, Figueira RCS, Eberlin MN, Lo Turco EG (2019) Lipidomic profile as a noninvasive tool to predict endometrial receptivity. *Mol Reprod Dev* 86(2):145-155

Busnelli A, Somigliana E, Cirillo F, Baggiani A, Levi-Setti PE (2021) Efficacy of therapies and interventions for repeated embryo implantation failure: a systematic review and meta-analysis. *Sci Rep* 11(1):1747

Campoy I, Lanau L, Altadill T, Sequeiros T, Cabrera S, Cubo-Abert M, Pérez-Benavente A, Garcia A, Borrós S, Santamaria A, et al (2016) Exosome-like vesicles in uterine aspirates: a comparison of ultracentrifugation-based isolation protocols. *J Transl Med* 14(1):180

Casado-Vela J, Rodriguez-Suarez E, Iloro I, Ametzazurra A, Alkorta N, García-Velasco JA, Matorras R, Prieto B, González S, Nagore D, et al (2009) Comprehensive proteomic analysis of human endometrial fluid aspirate*. J Proteome Res* 8(10):4622-32

Chen ZY, Hasson T, Zhang DS, Schwender BJ, Derfler BH, Mooseker MS, Corey DP (2001) Myosin-VIIb, a novel unconventional myosin, is a constituent of microvilli in transporting epithelia. *Genomics* 72(3):285-96.

Chen Y, Lun AT, Smyth GK (2016) From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Res* 5:1438

Cimadomo D, Capalbo A, Dovere L, Tacconi L, Soscia D, Giancani A, Scepi E, Maggiulli R, Vaiarelli A, Rienzi L, Ubaldi FM (2021) Leave the past behind: women's reproductive history shows no association with blastocysts' euploidy and limited association with live birth rates after euploid embryo transfers. *Hum Reprod* 36(4):929-940

Cozzolino M, Diaz-Gimeno P, Pellicer A, Garrido N (2020) Evaluation of the endometrial receptivity assay and the preimplantation genetic test for aneuploidy in overcoming recurrent implantation failure. *J Assist Reprod Genet* 37(12):2989-2997

Craciunas L, Gallos I, Chu J, Bourne T, Quenby S, Brosens JJ, Coomarasamy A (2019) Conventional and modern markers of endometrial receptivity: a systematic review and meta-analysis. *Hum Reprod Update* 25(2):202-223

Díaz-Gimeno P, Horcajadas JA, Martínez-Conejero JA, Esteban FJ, Alamá P, Pellicer A, Simón C (2011) A genomic diagnostic tool for human endometrial receptivity based on the transcriptomic signature. *Fertil Steril* 95(1):50-60

Dolanbay EG, Yardimoglu M, Yalcinkaya E, Yazir Y, Aksoy A, Karaoz E, Caliskan E (2016) Expression of trophinin and dipeptidyl peptidase IV in endometrial co-culture in the presence of an embryo: A comparative immunocytochemical study. *Mol Med Rep* 13(5):3961-8

European IVF-Monitoring Consortium (EIM) for the European Society of Human Reproduction and Embryology (ESHRE), Wyns C, De Geyter C, Calhaz-Jorge C, Kupka MS, Motrenko T, Smeenk J, Bergh C, Tandler-Schneider A, Rugescu IA et al (2021) ART in Europe, 2017: results generated from European registries by ESHRE. *Hum Reprod Open* 2021(3):hoab026

Evans J, Rai A, Nguyen HPT, Poh QH, Elglass K, Simpson RJ, Salamonsen LA, Greening DW (2019) Human Endometrial Extracellular Vesicles Functionally Prepare Human Trophectoderm Model for Implantation: Understanding Bidirectional Maternal-Embryo Communication. *Proteomics* 19(23):e1800423

Florio P, Bruni L, Galleri L, Reis FM, Borges LE, Bocchi C, Litta P, De Leo V, Petraglia F (2010) Evaluation of endometrial activin A secretion for prediction of pregnancy after intrauterine insemination. *Fertil Steril* 93(7):2316-20

Garcia-Alonso L, Handfield LF, Roberts K, Nikolakopoulou K, Fernando RC, Gardner L, Woodhams B, Arutyunyan A, Polanski K, Hoo R et al (2021) Mapping the temporal and spatial dynamics of the human endometrium in vivo and in vitro. *Nat Genet* 53(12):1698-1711

Gardner DK, Lane M, Calderon I, Leeton J (1996) Environment of the preimplantation human embryo in vivo: metabolite analysis of oviduct and uterine fluids and metabolism of cumulus cells. *Fertil Steril* 65(2):349-53

Giacomini E, Scotti GM, Vanni VS, Lazarevic D, Makieva S, Privitera L, Signorelli S, Cantone L, Bollati V, Murdica V, et al (2021) Global transcriptomic changes occur in uterine fluid-derived extracellular vesicles during the endometrial window for embryo implantation. *Hum Reprod* 36(8):2249-2274

Greco E, Bono S, Ruberti A, Lobascio AM, Greco P, Biricik A, Spizzichino L, Greco A, Tesarik J, Minasi MG, Fiorentino F (2014) Comparative genomic hybridization selection of blastocysts for repeated implantation failure treatment: a pilot study. *Biomed Res Int* 2014:457913

Greening DW, Nguyen HP, Elgass K, Simpson RJ, Salamonsen LA (2016) Human Endometrial Exosomes Contain Hormone-Specific Cargo Modulating Trophoblast Adhesive Capacity: Insights into Endometrial-Embryo Interactions. *Biol Reprod* 94(2):38

Hajipour H, Farzadi L, Roshangar L, Latifi Z, Kahroba H, Shahnazi V, Hamdi K, Ghasemzadeh A, Fattahi A, Nouri M (2021) A human chorionic gonadotropin (hCG) delivery platform using engineered uterine exosomes to improve endometrial receptivity. *Life Sci* 275:119351

Halperin R, Ron-El R, Golan A, Hadas E, Schneider D, Bukovsky I, Herman A (1995) Uterine fluid human decidua-associated protein 200 and implantation after embryo transfer. *Hum Reprod* 10(4):907-10

Henn A, De La Cruz EM (2005) Vertebrate myosin VIIb is a high duty ratio motor adapted for generating and maintaining tension. *J Biol Chem* 280(47):39665-76

Hou Z, He A, Zhang Q, Liu N, Liu D, Li Y, Xu B, Wang Y, Li S, Tian F, Liao T, Zhang Y, Cao J, Cao E, Li Y (2022) Endometrial fluid aspiration immediately prior to embryo transfer does not affect IVF/vitrified-warmed embryo transfer outcomes - a prospective matched cohort study. *Reprod Biomed Online* 44(3):486-493

Jain M, Samokhodskaya L, Mladova E, Panina O (2022) Mucosal biomarkers for endometrial receptivity: A promising yet underexplored aspect of reproductive medicine. *Syst Biol Reprod Med* 68(1):13-24

Khadem N, Mansoori M, Attaran M, Attaranzadeh A, Zohdi E (2019) Association of IL-1 and TNF-α Levels in Endometrial Secretion and Success of Embryo Transfer in IVF/ICSI Cycles. *Int J Fertil Steril* 13(3):236-239

Lédée-Bataille N, Olivennes F, Kadoch J, Dubanchet S, Frydman N, Chaouat G, Frydman R (2004) Detectable levels of interleukin-18 in uterine luminal secretions at oocyte retrieval predict failure of the embryo transfer. *Hum Reprod* 19(9):1968-73

Lee CL, Vijayan M, Wang X, Lam KKW, Koistinen H, Seppala M, Li RHW, Ng EHY, Yeung WSB, Chiu PCN (2019) Glycodelin-A stimulates the conversion of human peripheral blood CD16-CD56bright NK cell to a decidual NK cell-like phenotype. *Hum Reprod* 34(4):689-701

Leese HJ, Hugentobler SA, Gray SM, Morris DG, Sturmey RG, Whitear SL, Sreenan JM (2008) Female reproductive tract fluids: composition, mechanism of formation and potential role in the developmental origins of health and disease. *Reprod Fertil Dev* 20(1):1-8

Li T, Greenblatt EM, Shin ME, Brown TJ, Chan C (2021) Cargo small non-coding RNAs of extracellular vesicles isolated from uterine fluid associate with endometrial receptivity and implantation success. *Fertil Steril* 115(5):1327-1336

Liao Z, Liu C, Wang L, Sui C, Zhang H (2021) Therapeutic Role of Mesenchymal Stem Cell-Derived Extracellular Vesicles in Female Reproductive Diseases. *Front Endocrinol (Lausanne)* 12:665645

Liu C, Yao W, Yao J, Li L, Yang L, Zhang H, Sui C (2020) Endometrial extracellular vesicles from women with recurrent implantation failure attenuate the growth and invasion of embryos. *Fertil Steril* 114(2):416-425

Luddi A, Zarovni N, Maltinti E, Governini L, Leo V, Cappelli V, Quintero L, Paccagnini E, Loria F, Piomboni P (2019) Clues to Non-Invasive Implantation Window Monitoring: Isolation and Characterisation of Endometrial Exosomes. *Cells* 8(8):811

Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstråle M, Laurila E, et al. (2003) PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 34(3):267-73

Moreno I, Codoñer FM, Vilella F, Valbuena D, Martinez-Blanch JF, Jimenez-Almazán J, Alonso R, Alamá P, Remohí J, Pellicer A, et al. (2016) Evidence that the endometrial microbiota has an effect on implantation success or failure. *Am J Obstet Gynecol* 215(6):684-703

Murphy CR (2004) Uterine receptivity and the plasma membrane transformation. *Cell Res* 14(4):259-67

Ng YH, Rome S, Jalabert A, Forterre A, Singh H, Hincks CL, Salamonsen LA (2013) Endometrial exosomes/microvesicles in the uterine microenvironment: a new paradigm for embryo-endometrial cross talk at implantation. *PLoS One* 8(3):e58502

O'Brien K, Breyne K, Ughetto S, Laurent LC, Breakefield XO (2020) RNA delivery by extracellular vesicles in mammalian cells and its applications. *Nat Rev Mol Cell Biol* 21(10):585-606

Phillips IR, Shephard EA (2008) Flavin-containing monooxygenases: mutations, disease and drug response. *Trends Pharmacol Sci* 29(6):294-301

Rahiminejad ME, Moaddab A, Ganji M, Eskandari N, Yepez M, Rabiee S, Wise M, Ruano R, Ranjbar A (2016) Oxidative stress biomarkers in endometrial secretions: A comparison between successful and unsuccessful in vitro fertilization cycles. *J Reprod Immunol* 116:70-5

Rai A, Poh QH, Fatmous M, Fang H, Gurung S, Vollenhoven B, Salamonsen LA, Greening DW (2021) Proteomic profiling of human uterine extracellular vesicles reveal dynamic regulation of key players of embryo implantation and fertility during menstrual cycle. *Proteomics* 21(13-14):e2000211

Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J (2019) g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res* 2;47(W1):W191-W198

Rubio C, Bellver J, Rodrigo L, Bosch E, Mercader A, Vidal C, De los Santos MJ, Giles J, Labarta E, Domingo J, et al. (2013) Preimplantation genetic screening using fluorescence in situ hybridization in patients with repetitive implantation failure and advanced maternal age: two randomized trials. *Fertil Steril* 99(5):1400-7

Russell P, Hey-Cunningham A, Berbic M, Tremellen K, Sacks G, Gee A, Cheerala B (2014) Asynchronous glands in the endometrium of women with recurrent reproductive failure. *Pathology* 46(4):325-32

Sato T, Sugiura-Ogasawara M, Ozawa F, Yamamoto T, Kato T, Kurahashi H, Kuroda T, Aoyama N, Kato K, Kobayashi R, et al. (2019) Preimplantation genetic testing for aneuploidy: a comparison of live birth rates in patients with recurrent pregnancy loss due to embryonic aneuploidy or recurrent implantation failure. *Hum Reprod* 1;34(12):2340-2348

Sebastian-Leon P, Devesa-Peiro A, Aleman A, Parraga-Leo A, Arnau V, Pellicer A, Diaz-Gimeno P (2021) Transcriptional changes through menstrual cycle reveal a global transcriptional derepression underlying the molecular mechanism involved in the window of implantation. *Mol Hum Reprod* 27(5):gaab027

Shimomura Y, Ando H, Furugori K, Kajiyama H, Suzuki M, Iwase A, Mizutani S, Kikkawa F (2006) Possible involvement of crosstalk cell-adhesion mechanism by endometrial CD26/dipeptidyl peptidase IV and embryonal fibronectin in human blastocyst implantation. *Mol Hum Reprod* 12(8):491-5

Simintiras CA, Dhakal P, Ranjit C, Fitzgerald HC, Balboula AZ, Spencer TE (2021) Capture and metabolomic analysis of the human endometrial epithelial organoid secretome. *Proc Natl Acad Sci U S A* 118(15):e2026804118

Simón C, Gómez C, Cabanillas S, Vladimirov I, Castillón G, Giles J, Boynukalin K, Findikli N, Bahçeci M, Ortega I, et al. (2020) ERA-RCT Study Consortium Group. A 5-year multicentre randomized controlled trial comparing personalized, frozen and fresh blastocyst transfer in IVF. *Reprod Biomed Online* 41(3):402-415

Skryabin GO, Komelkov AV, Zhordania KI, Bagrov DV, Vinokurova SV, Galetsky SA, Elkina NV, Denisova DA, Enikeev AD, Tchevkina EM (2022) Extracellular Vesicles from Uterine Aspirates Represent a Promising Source for Screening Markers of Gynecologic Cancers. *Cells* 11(7):1064

Somigliana E, Busnelli A, Kalafat E, Viganò P, Ata B (2022) Recurrent implantation failure: a plea for a widely adopted rational definition. *Reprod Biomed Online* S1472-6483(22)00071-2

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102(43):15545-50

Tan Q, Shi S, Liang J, Zhang X, Cao D, Wang Z (2020) MicroRNAs in Small Extracellular Vesicles Indicate Successful Embryo Implantation during Early Pregnancy. *Cells* 9(3):645

Tapia A, Gangi LM, Zegers-Hochschild F, Balmaceda J, Pommer R, Trejo L, Pacheco IM, Salvatierra AM, Henríquez S, Quezada M, et al. (2008) Differences in the endometrial transcript profile during the receptive period between women who were refractory to implantation and those who achieved pregnancy. *Hum Reprod* 23(2):340-51

Théry C, Witwer KW, Aikawa E, Alcaraz MJ, Anderson JD, Andriantsitohaina R, Antoniou A, Arab T, Archer F, Atkin-Smith GK, et al. (2018) Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. *J Extracell Vesicles* 7(1):1535750

Tojo A, Hatakeyama S, Kinugasa S, Fukuda S, Sakai T (2017) Enhanced podocyte vesicle transport in the nephrotic rat. *Med Mol Morphol* 50(2):86-93

van der Gaast MH, Beier-Hellwig K, Fauser BC, Beier HM, Macklon NS (2003) Endometrial secretion aspiration prior to embryo transfer does not reduce implantation rates. *Reprod Biomed Online* 7(1):105-9

Vignali DA, Collison LW, Workman CJ (2008) How regulatory T cells work. *Nat Rev Immunol* 8(7):523-32

Vilella F, Moreno-Moya JM, Balaguer N, Grasso A, Herrero M, Martínez S, Marcilla A, Simón C (2015) Hsa-miR-30d, secreted by the human endometrium, is taken up by the pre-implantation embryo and might modify its transcriptome. *Development* 15;142(18):3210-21

von Grothusen C, Frisendahl C, Modhukur V, Lalitkumar PG, Peters M, Faridani OR, Salumets A, Boggavarapu NR, Gemzell-Danielsson K (2022) Uterine fluid microRNAs

are dysregulated in women with recurrent implantation failure. *Hum Reprod* 37(4):734-746

Xu C, Zhao W, Huang X, Jiang Z, Liu L, Cui L, Li X, Li D, Du M (2021) TORC2/3-mediated DUSP1 upregulation is essential for human decidualization. *Reproduction* 161(5):573-580

Wang W, Vilella F, Alama P, Moreno I, Mignardi M, Isakova A, Pan W, Simon C, Quake SR (2020) Single-cell transcriptomic atlas of the human endometrium during the menstrual cycle. *Nat Med* 26(10):1644-1653

Yakin K, Ata B, Ercelen N, Balaban B, Urman B (2008) The effect of preimplantation genetic screening on the probability of live birth in young women with recurrent implantation failure; a nonrandomized parallel group trial. *Eur J Obstet Gynecol Reprod Biol* 140(2):224-9

Yáñez-Mó M, Siljander PR, Andreu Z, Zavec AB, Borràs FE, Buzas EI, Buzas K, Casal E, Cappello F, Carvalho J, et al (2015) Biological properties of extracellular vesicles and their physiological functions. *J Extracell Vesicles* 4:27066

Yao K, Wang Q, Jia J, Zhao H (2017) A competing endogenous RNA network identifies novel mRNA, miRNA and lncRNA markers for the prognosis of diabetic pancreatic cancer. *Tumour Biol* 39(6):1010428317707882

Zarovni N, Corrado A, Guazzi P, Zocco D, Lari E, Radano G, Muhhina J, Fondelli C, Gavrilova J, Chiesi A (2015) Integrated isolation and quantitative analysis of exosome shuttled proteins and nucleic acids using immunocapture approaches. *Methods* Oct 1;87:46-58

Zhang X, Li Y, Chen X, Jin B, Shu C, Ni W, Jiang Y, Zhang J, Ma L, Shu J (2022) Single-cell transcriptome analysis uncovers the molecular and cellular characteristics of thin endometrium. *FASEB J* 36(3):e22193

Zhao Y, Zhang T, Guo X, Wong CK, Chen X, Chan YL, Wang CC, Laird S, Li TC (2021) Successful implantation is associated with a transient increase in serum pro-inflammatory cytokine profile followed by a switch to anti-inflammatory cytokine profile prior to confirmation of pregnancy. *Fertil Steril* 115(4):1044-1053