

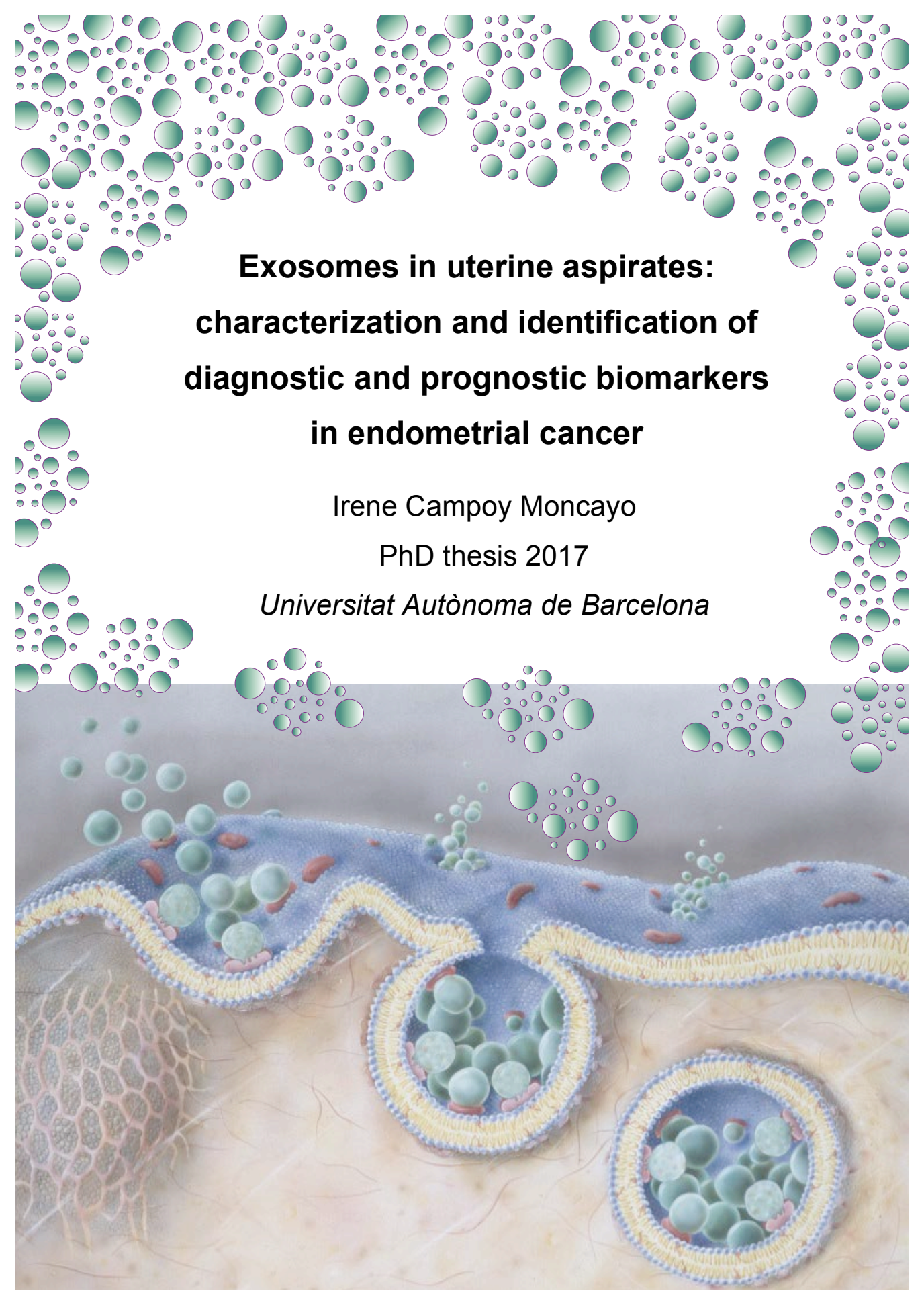


Universitat Autònoma de Barcelona

**ADVERTIMENT.** L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  [http://cat.creativecommons.org/?page\\_id=184](http://cat.creativecommons.org/?page_id=184)

**ADVERTENCIA.** El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <http://es.creativecommons.org/blog/licencias/>

**WARNING.** The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>

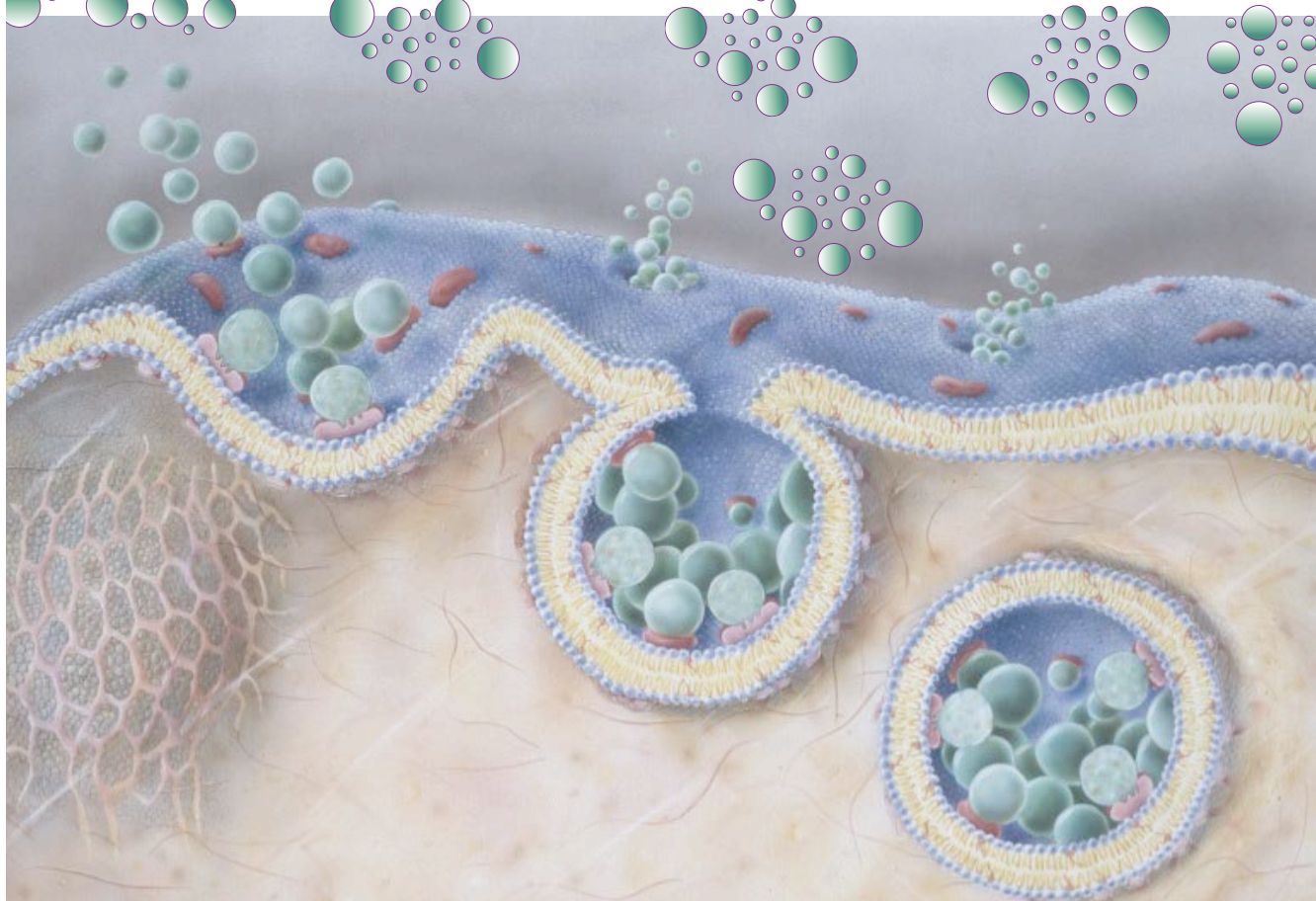


**Exosomes in uterine aspirates:  
characterization and identification of  
diagnostic and prognostic biomarkers  
in endometrial cancer**

Irene Campoy Moncayo

PhD thesis 2017

*Universitat Autònoma de Barcelona*







**EXOSOMES IN UTERINE ASPIRATES:  
CHARACTERIZATION AND IDENTIFICATION OF  
DIAGNOSTIC AND PROGNOSTIC BIOMARKERS IN  
ENDOMETRIAL CANCER**

Doctoral thesis presented by  
**Irene Campoy Moncayo**

to obtain the degree of  
**Doctor for the *Universitat Autònoma de Barcelona* (UAB)**

Doctoral thesis performed at the Vall d'Hebrón Research Institute, in the Group of Biomedical Research in Gynecology, under the supervision of  
**Dr. Jaume Reventós, Dr. Eva Colás and Dr. Antonio Gil.**

Doctoral study in Cellular Biology, Department of Cellular Biology, Physiology and Immunology, Faculty of Medicine at the UAB, under the supervision of  
**Dr. Imma Ponsa.**

**Universitat Autònoma de Barcelona, 20<sup>th</sup> June 2017**

Dr. Jaume Reventós  
(director)

Dra. Eva Colás  
(director)

Dr. Antonio Gil  
(director)

Dra. Imma Ponsa  
(tutor)

Irene Campoy  
(student)



Al meu pare  
A mi madre  
Al tete



**AGRAÏMENTS**  
**AGRADECIMIENTOS**  
**ACKNOWLEDGEMENTS**

---





Tot i que, de vegades, els camins no són fàcils, sempre ens hem d'aturar i pensar que som on som, i som qui som, gràcies a la gent que ens acompanya durant les diferents etapes de la vida.

En primer lloc vull mostrar la meva gratitud als meus directors de tesi, el Dr. Jaume Reventós, el Dr. Antonio Gil i la Dra. Eva Colás, per brindar-me l'oportunitat de realitzar aquest treball i per confiar en les meves possibilitats. Jaume, gràcies per obrir-me les portes del lab 209 en el seu dia i ensenyar-me, amb el teu exemple, que l'èxit s'aconsegueix amb perseverança. Antonio, gracias por tomar el relevo y seguir apoyando nuestros proyectos. Eva, gracias por creer en mi desde el principio y por todo el trabajo realizado hasta llegar a aquí.

Si intento recordar el meu primer dia al lab, em ve al cap la Laura. Lau, gracias por enseñarme todo lo que sabias durante los primeros meses, cuando estaba de prácticas de máster; en el fondo, hasta nos lo pasábamos bien bajando cada día a ratones! He aprendido muchísimo de ti y contigo, tanto en lo profesional como en lo personal. Y siguiendo con las endométricas... mi Jelen, mi siamesa, mi enganxineta. Parece que hayan pasado mil años desde aquellos viernes por la noche, en cultivos... cuantas risas, cuantos lloros, cuantas chorradas que sólo tu y yo entendíamos (Flipper, estsdsss, huevos fritos y sentir penita al ver musculazos que se comen a las epiteliales bonitas...). Pero luego empezamos con las estancias... "Tu a Luxemburgo y yo a Canadá" sería el título de nuestra película. A pesar de los manotazos que siempre te doy para que no te muerdas las uñas, sabes que te quiero y que nada de esto hubiese sido lo mismo sin ti. Lu, llucet del meu cor! IreLu ens va portar molts mals de caps, però també molts moments juntes que no tenen preu. Juntes hem plorat, hem rigut i hem cantat "la oreja de van Gogh" mentre fèiem figures amb el nostre estimat Illustrator. T'entenc i m'entens. Tati-xula i Blanki, també em veniu al cap quan recordo els primer dies al lab. Em sonàveu de la uni, però no tenia el plaer de conèixer-vos. Un dels primer records que tinc "extra-lab" amb vosaltres va ser la despedida i el viatge a Istanbul per la boda de la Tugçé. Allà ja vaig començar a veure que gaudiríem de molts moments juntes... i així ha sigut. Y no me olvido de ti Tamara! Taaaa, cuantas horas habremos compartido en el NanoSight del IQS? Os acordáis las 4 metidas en los baños turcos? Que gran viaje!

Me n'adono que més que companyes de feina hem anat fent una pinya. Juntes hem fet escapades de cap de setmana; ens hem disfressat, hem viatjat per diversos indrets del món (Turquía, València, Luxemburg, Washington, Grècia, Madrid...). Els anys han anat passant, i entre estada i estada, el grup anava incorporant a gent nova: Alfonso, Gabriel, Núria, Anna, Cristian, Berta...No em vull oblidar dels que ja fa temps que van marxar: Raúl, Marta

Llauradó, Marta García, Núria i Melania; ni de la gent de pràctiques que ha anat passant (en especial de la Rusiuu). Marina, un gràcies especial per estar sempre disposada a ajudar-me.

I com no, menció especial a les meves “Guais” i a la Mire. Mire, gràcies per totes les converses, l’ajuda i els ànims del dia a dia! Per deixar-me dormir al teu sofà de NYC i per aguantar-me al despatx (el despatx més molón i kuki). I a vosaltres Guais, què dir-vos que no ens haguem demostrat ja amb una abraçada a 3 en qualsevol W.C? Gràcies per tot, per ser-hi, per escoltar, per ajudar, per animar, per entendre’m, per riure, per plorar, pels detalls, per les notes de veu, pels esmorzars, pels dinars i per les birres! Shuli, Eli (tzabet), us estimo molt, de veritat.

Per acabar amb la família del VHIR, donar gràcies a la Rosa, l’Eulàlia, la Pilar, el Rai, el Tau i la Rachida, perquè tots vosaltres també heu aportat el vostre granet de sorra per arribar fins a aquí.

I would like to thank also to everyone that helped me somehow during my time in Montreal: Marta Sesé, Alba, Noelia, Debbie, Raik, Almer, Jacob, Francis, Dev, Dhanaraman, Nebiyu, Christel, Fred, Evgeny, Thi Thanh and Aditya. An especial thanks to Dr. Pierre Thibault for his help and kindness, and for letting me be part of his wonderful lab team for a while. Eric, thanks for making the hours at the platform easier (and for realizing when we got a new french manicure done!) Christina, Christine, Peter and Sibylle: thanks for sharing nice moments with me, for the weekend trips, the BBQs and for your help. Hope to see you again soon. Sibylle, mon amie, merci! Thanks for being the nicest person ever and for your unconditional help; For your post cards, your messages and everything you have done for me since our first coffee break. Love you tons.

També vull agrair la infinita paciència dels meus amics. A les meves amigues de tota la vida: a les “Rasus” (Rosi i Alba) per donar-me ànims i creure en mi; al meu putot (Anna), perquè amb una mirada ja ens entenem, perquè som la nit i el dia en moltes coses, però em complementes i no sabria viure sense tu; al meu cigronet (Pat) per ser-hi sempre i recolzar-me, pels teus petits detalls que signifiquen tant per mi. A les meves amigues de la uni: al meu pebrotus (Iris), per passar d’odiar-me a estimar-me i per posar-me maquinorra al cotxe; al conjunt de les Bionenes (Laura, Maria Sitges, Moni, Gemma, Anna Font, Marta, Coco, Neus, Gela i Maria) perquè sempre sabeu donar el toc màgic a qualsevol celebració i recolzar a qualsevol bionena necessitada; al meu xurru (Gela), per tot lo compartit i per ser com ets; a la Maria, per ser-hi sempre, per poder comptar amb tu en els bons i, sobretot, en els mals moments i perquè *juntes* hem pogut i podrem sempre amb tot. A les meves nenes del màster (Lulu, Marta and Kate, my BioLac girls!).

Finalment, donar les gràcies a tota la meva estimada família per tot el suport, ara i sempre. Als meus pares per tot el que sóc i el que tinc a la vida. Al meu germà per ser sempre un exemple a seguir i saber-me donar aquell punt de tranquil·litat i una visió pràctica dels problemes. A la Tita Mari, por todas las velitas a San Pancraccio. A Heidi por ser mi rubia, mi confidente y mi amiga, no sólo mi cuñada. Als Mu-iaios (Hugo & Sara) i al Jordi i la Lluïsa, per tots els dinars, sopars i pa de pessic compartits. A la Iria, per ser forta i haver tornat a Barcelona. I al Wiron, per estimar-me i donar-me forces cada dia, pels post-its als entrepans, per fer-me riure, per no deixar-me caure, per animar-me, per creure en mi, escoltar-me i aguantar-me.

A la abuela, al abuelo i al yayo, que sé que m'ajuden amagats darrere d'una estrella. I a la yaya, que espero que la última cosa que obliidi el seu cap és que l'estimo.

A l'Arlet i a la Jana, per omplir-me d'una indescriptible alegria amb la seva arribada.



# Table of contents

<b>INDEX OF FIGURES</b> .....	<b>15</b>
<b>INDEX OF TABLES</b> .....	<b>17</b>
<b>ABBREVIATIONS &amp; ACRONYMS</b> .....	<b>19</b>
<b>INTRODUCTION</b> .....	<b>25</b>
1. THE ENDOMETRIUM .....	27
1.1 The uterus .....	27
1.2 The endometrial cycle .....	29
1.2.1 Proliferative endometrium.....	29
1.2.2 Secretive endometrium.....	29
1.2.3 Menstrual endometrium.....	30
1.2.4 Endometrium at pregnancy .....	30
1.2.5 Atrophic endometrium .....	30
2. ENDOMETRIAL CANCER.....	31
2.1 Epidemiology.....	31
2.2 Risk factors .....	32
2.2.1 Long-term unopposed endogenous and exogenous estrogen exposure .....	33
2.2.2 Obesity .....	35
2.2.3 Diabetes and hypertension.....	35
2.2.4 Genetic factors .....	36
2.3 Protective factors .....	36
2.4 Diagnosis .....	37
2.4.1 Signs and symptomatology .....	37
2.4.2 Screening recommendations.....	37
2.4.3 Suspected diagnosis: clinical examination .....	38
2.4.4 Confirmatory diagnosis: pathological examination .....	39
2.5 Endometrial lesions.....	41
2.5.1 Endometrial polyps.....	41
2.5.2 Endometrial hyperplasia .....	41
2.6 Classification of endometrial cancer .....	43
2.6.1 Dualistic model – Clinico-pathological classification.....	43

2.6.2	Histological type .....	45
2.6.3	FIGO staging .....	46
2.6.4	Molecular classification: the TCGA model.....	48
2.7	Prognostic factors .....	49
2.8	Pre-operative risk assessment.....	51
2.8.1	Preoperative endometrial biopsies .....	51
2.8.2	Magnetic resonance imaging (MRI).....	52
2.9	Treatment.....	53
3.	BIOMARKERS.....	54
3.1	Biomarker definition and protein biomarkers.....	54
3.2	Challenges for the discovery of protein biomarkers and source sample selection .....	54
3.2	The biomarker pipeline.....	56
4.	EXTRACELLULAR VESICLES AND EXOSOMES .....	57
4.1	Brief history .....	57
4.2	Nomenclature of extracellular vesicles.....	59
4.3	Exosomes .....	61
4.3.1	Isolation methods .....	61
4.3.2	Physical properties .....	62
4.3.3	Biogenesis .....	63
4.3.4	Composition.....	67
4.4	Release and internalization.....	69
4.4.1	Mechanistic requirements for exosomes release .....	69
4.4.2	Interaction of exosomes with recipient cells .....	70
4.5	Functions of exosomes .....	71
4.6	Advantages of exosomes in protein biomarker research .....	72
	<b>OBJECTIVES .....</b>	<b>75</b>
	<b>MATERIALS AND METHODS .....</b>	<b>83</b>
1.	Uterine aspirates.....	85
1.1	Collection and processing of uterine aspirates .....	85
1.2	Selection of patients and inclusion/exclusion criteria .....	85
1.2.1	For the establishment of the ELVs isolation method .....	85
1.2.1	For the pilot studies to set up the proteomic approach for the discovery phase.....	87

1.2.2 For the discovery and the verification phase .....	87
1.2.2 For the validation in UAs .....	89
2. Cell culture .....	91
2.1 Cell lines .....	91
2.2 Culture conditions for ELVs isolation .....	91
2.2.1 ELVs-depleted media preparation .....	91
2.2.2 Cell Culture for ELVs enrichment .....	92
2.3 Culture conditions to prepare the Super-SILAC Mix .....	92
3. ELVs isolation .....	92
3.1 Establishment of the ELVs isolation method .....	92
3.1.1 Standard protocol .....	93
3.1.2 Filtration protocol .....	93
3.1.3 Sucrose cushion protocol .....	93
4. EVs visualization and measurement .....	95
4.1 Transmission electron microscopy .....	95
4.2 Nanoparticle tracking analysis .....	95
5. Protein techniques .....	96
5.1 Protein extraction .....	96
5.1.1 Protein extraction from MVs and ELVs .....	96
5.1.2 Protein extraction from soluble fractions .....	96
5.1.3 Protein extraction from UAs .....	97
5.2 Protein quantification .....	97
5.3 Western blotting .....	97
5.4 Protein digestion and sample preparation for Mass Spectrometry .....	98
5.4.1 Super-SILAC mix and the in-solution digestion protocol .....	98
5.4.2 Discovery Phase (In-Gel Digestion) .....	98
5.4.3 Verification Phase (In-Solution Digestion) .....	99
5.4.3 Validation in Uterine Aspirates .....	100
5.5 Mas Spectrometry analysis .....	101
5.5.1 Comparison of the cell lines-derived ELVs proteome .....	101
5.5.2 Discovery Phase .....	102
5.5.3 Verification Phase .....	104
5.5.4 Validation in UAs .....	106
6. Gene ontology analysis .....	106



7. Statistical analysis and Software .....	106
<b>RESULTS .....</b>	<b>109</b>
1. Establishment of a suitable protocol for the use of ELVs from UAs .....	111
1.1 Molecular characterization of ELVs from UAs.....	113
2. Identification of new protein biomarkers for EC in ELVs from UAs.....	118
2. 1 Proteomic analysis of ELVs derived from 5 EC cell lines for the creation of the Super-SILAC mix .....	118
2.2 Setting up the proteomic approach for the discovery phase .....	123
2.3 Protein identification and relative quantification with a SILAC-based LC-MS/MS discovery phase .....	125
2.4 GO analysis to understand the most relevant biological processes of ELVs in UAs .....	131
3. Verification of candidates by targeted proteomics in UAs ELVs.....	134
3.1 Verification of protein biomarkers in UAs ELVs .....	134
3.2 Development of predictors for EC diagnosis and prognosis in ELVs derived from UAs .....	137
4. Verification of candidates by targeted proteomics in the whole fluid of UAs ...	141
<b>DISCUSSION .....</b>	<b>145</b>
<b>CONCLUSIONS .....</b>	<b>163</b>
<b>JOURNAL PUBLICATIONS .....</b>	<b>167</b>
<b>BIBLIOGRAPHY .....</b>	<b>171</b>

## Index of Figures

<b>Figure 1.</b> Anatomy of the uterus.....	28
<b>Figure 2.</b> Leading sites of new cancer cases and deaths estimated for 2017 .....	31
<b>Figure 3.</b> Endometrial cancer stage and 5-year survival rate by race.....	32
<b>Figure 4.</b> Clinical examination.....	39
<b>Figure 5.</b> Endometrial biopsy .....	40
<b>Figure 6.</b> Progression from endometrial hyperplasia to EC. ....	42
<b>Figure 7.</b> Histology of 4 common types of EC.....	46
<b>Figure 8.</b> FIGO stages .....	47
<b>Figure 9.</b> The biomarker pipeline .....	57
<b>Figure 10.</b> Schematic representation of subtypes of extracellular vesicles (EVs) released by a cell.....	60
<b>Figure 11.</b> Ultrastructure of exosomes.....	63
<b>Figure 12.</b> Biogenesis, secretion and composition of exosomes. ....	65
<b>Figure 13.</b> Model for sorting of cargo into different MVB subpopulations .....	67
<b>Figure 14.</b> Composition of a typical exosome .....	69
<b>Figure 15.</b> Exosomes uptake .....	71
<b>Figure 16.</b> Publication statistics on extracellular vesicles using MS .....	74
<b>Figure 17.</b> Schematic representation of the three protocols tested to isolate ELVs from UAs.....	94
<b>Figure 18.</b> Electron microscopy image of ELVs and MVs isolated from UAs .....	112
<b>Figure 19.</b> Size distribution of ELVs and MVs by NTA.....	112
<b>Figure 20.</b> ELVs concentration measured by NTA.....	113
<b>Figure 21.</b> Immunoblot of MVs, ELVs and SF.....	114
<b>Figure 22.</b> Relative tetraspanins expression of ELVs .....	114
<b>Figure 23.</b> Particle size of ELVs from individual UAs.....	115
<b>Figure 24.</b> Correlation between total number of ELVs and UA's volume and protein concentration .....	115
<b>Figure 25.</b> Correlation between total number of MVs and UA's volume .....	116
<b>Figure 26.</b> Immunoblot of MVs, ELVs and SF of individual samples .....	116
<b>Figure 27.</b> Quantification of tetraspanin's expression of individual samples.....	117
<b>Figure 28.</b> Workflow of the experimental procedure from the discovery to the generation of predictive models.....	119

<b>Figure 29.</b> Workflow followed to compare the proteome of ELVs from 5 EC cell lines and from UAs.....	120
<b>Figure 30.</b> Immunoblot of EC cell line and their derived ELVs.....	120
<b>Figure 31.</b> EC cell line's derived ELVs size distribution by NTA .....	121
<b>Figure 32.</b> Number of proteins identified in each cell line and patient's derived ELVs .....	122
<b>Figure 33.</b> Comparisson of proteins identified from cell lines and UAs.....	122
<b>Figure 34.</b> The 3 selected cell lines represent the 47% of the UAs proteome ....	123
<b>Figure 35.</b> Coomassie staining of ELVs isolated from UAs with different amounts of blood contamination .....	124
<b>Figure 36.</b> Schematic representation of the discovery phase .....	125
<b>Figure 37.</b> Mode and size distribution of ELVs of samples of the discovery phase determined by NTA.....	126
<b>Figure 38.</b> ELVs and protein concentration of samples of the discovery phase ..	126
<b>Figure 39.</b> Immunoblot od ELVs of samples of the discovery .....	127
<b>Figure 40.</b> Relation of proteins identified and quantified by LC-MS/MS in each patient of the discovery.....	128
<b>Figure 41.</b> Match of proteins from the discovery with ExoCarta; Diagnostic/Prognostic potential of proteins differentially expressed and proteins up/down regulated in cancer .....	129
<b>Figure 42.</b> Workflow followed for the GO analysis .....	132
<b>Figure 43.</b> Diagnostic performance of biomarkers in ELVs from UAs.....	138
<b>Figure 44.</b> Prognostic performance of biomarkers in ELVs from UAs .....	138

## Index of Tables

<b>Table 1.</b> Clinico-pathological classification of endometrial cancer .....	44
<b>Table 2.</b> The 2009 FIGO staging .....	47
<b>Table 3.</b> The TCGA model .....	49
<b>Table 4.</b> Risk groups to guide adjuvant treatment.....	50
<b>Table 5.</b> Surgical treatment depending on staging.....	53
<b>Table 6.</b> Recommended adjuvant treatment .....	53
<b>Table 7.</b> Physicochemical characteristics of different types of secreted vesicles ..	59
<b>Table 8.</b> Tumor-derived exosomes involvement in the hallmarks of cancer .....	72
<b>Table 9.</b> Clinical and pathological features of patients used for Objective 1 .....	86
<b>Table 10.</b> Clinical and pathological features of patients used to set up the protocol for the proteomic approach.....	87
<b>Table 11.</b> Clinical and pathological features of patients used for Objective 2 .....	88
<b>Table 12.</b> Clinical and pathological features of patients used for Objective 3 .....	89
<b>Table 13.</b> Clinical and pathological features of patients used for Objective 3c .....	90
<b>Table 14.</b> . Features of the EC cell lines used in this work.....	91
<b>Table 15.</b> Proteins identified by LC-MS/MS after in-gel or in-solution digestion ..	125
<b>Table 16.</b> List of selected candidates for LC-SRM .....	130
<b>Table 17.</b> Reduced list of 22 proteins for GO analysis .....	133
<b>Table 18.</b> Statistical inference results of prognostic biomarkers in ELVs.....	135
<b>Table 19.</b> Statistical inference results of diagnostic biomarkers in ELVs .....	136
<b>Table 20.</b> Summary table of the results of the logistic regression model for the ROC analysis.....	139
<b>Table 21.</b> Summary of proteins monitored, detected and quantified in each cohort from the discovery of the candidates to the verification in ELVs and UAs.....	140
<b>Table 22.</b> Top-10 diagnostic candidates based on AUC values obtained in ELVs and UAs .....	142
<b>Table 23.</b> Top-10 prognostic candidates based on AUC values obtained in ELVs and UAs .....	142
<b>Table 24.</b> Quantitative proteomics studies performed on extracellular vesicles of different origin .....	156



# ABBREVIATIONS & ACRONYMS

---



<b>ABC</b>	Amonium bicarbonate
<b>ARID1</b>	Adenine, Thymine-rich interactive domain-containing protein 1A
<b>ASR(W)</b>	Age-Standardized Rate worldwide
<b>AUB</b>	Abnormal uterine bleeding
<b>AUC</b>	Area under the ROC curve
<b>BRCA1</b>	Breast cancer gene 1
<b>BSA</b>	Bovine serum albumin
<b>BT</b>	Brachytherapy
<b>CI</b>	Confidence interval
<b>cm</b>	Centimeters
<b>cryo-EM</b>	Cryo electron microscopy
<b>CT</b>	Computed Tomography
<b>CTNNB1</b>	Catenin (cadherin-associated protein), beta 1 ( $\beta$ -catenin)
<b>Da</b>	Daltons
<b>DMEM F-12</b>	Dulbecco's modified eagle medium: nutrient mixture F-12
<b>DNA</b>	Deoxyribonucleic acid
<b>DTT</b>	Dithiothreitol
<b>EC</b>	Endometrial Cancer
<b>ECL</b>	Enhanced chemiluminescence
<b>EDTA</b>	Ethylenediaminetetraacetic acid
<b>EEC</b>	Endometrioid Endometrial Cancer
<b>EIN</b>	Endometrial Intraepithelial Neoplasia
<b>ELV</b>	Exosome-like vesicle
<b>ESCRT</b>	Endosomal sorting complexes required for transport
<b>ESGO</b>	European Society of Gynaecologic Oncology
<b>ESMO</b>	European Society for Medical Oncology
<b>ESTRO</b>	European Society for Radiotherapy and Oncology
<b>EV</b>	Extracellular vesicle
<b>FA</b>	Formic acid
<b>FACS</b>	Fluorescence-activated cell sorting
<b>FASP</b>	Filter-Aided Sample Preparation
<b>FBS</b>	Fetal bovine serum
<b>FBXW7</b>	F-box/WD repeat-containing protein 7
<b>FDR</b>	False discovery rate
<b>FGFR2</b>	Fibroblast growth factor receptor
<b>FIGO</b>	International Federation of Gynecology and Obstetrics
<b>GO</b>	Gene ontology
<b>hCG</b>	Human chorionic gonadotropin
<b>HER2</b>	Human Epidermal Growth Factor Receptor 2



<b>HNPCC</b>	Hereditary non-polyposis colorectal cancer
<b>HRP</b>	Horseradish peroxidase
<b>HUAV</b>	Hospital Universitari Arnau de Vilanova
<b>HUVH</b>	Hospital Universitari Vall d'Hebron
<b>IAM</b>	Iodoacetamide
<b>Ig</b>	Immunoglobulin
<b>IGF1</b>	Insulin-like growth factor 1
<b>ILV</b>	Intraluminal vesicles
<b>ISEV</b>	International Society of Extracellular Vesicles
<b>ISGP</b>	International Society of Gynecological Pathologists
<b>iTRAQ</b>	Isobaric tags for relative and absolute quantitation
<b>KRAS</b>	Kirsten rat sarcoma viral oncogene homolog oxidase-like 2 LOXL2 Lysyl
<b>LBPA</b>	lysobisphosphatidic acid
<b>LC-MS</b>	Liquid chromatography coupled to mass spectrometry
<b>LC-MS/MS</b>	Liquid chromatography tandem mass spectrometry
<b>LTQ</b>	Linear trap quadrupole
<b>LVSI</b>	Lymphovascular space invasion
<b>Mb</b>	Megabase
<b>MI</b>	Myometrial invasion
<b>miRNA</b>	micro RNA
<b>ml</b>	Milliliter
<b>MLH1</b>	MutL homolog 1
<b>mm</b>	Millimeters
<b>mM</b>	Millimolar
<b>MRI</b>	Magnetic resonance imaging
<b>MRM</b>	Multiple reaction monitoring
<b>mRNA</b>	Messenger RNA
<b>MS</b>	Mass spectrometry
<b>MSI</b>	Microsatellite instability
<b>MV</b>	Microvesicle
<b>MVB</b>	Multivesicular body
<b>MVE</b>	Multivesicular endosome
<b>NEEC</b>	Non Endometrioid Endometria Cancer
<b>ng</b>	Nanograms
<b>nL</b>	Nanoliter
<b>nm</b>	Nanometers
<b>NPF</b>	Negative predictive value
<b>NTA</b>	Nanoparticle tracking analysis
<b>PBS</b>	Phosphate buffered saline

<b>PI3K</b>	Phosphoinositide-3-kinase
<b>PIK3CA</b>	Phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic
<b>PIK3R1</b>	Phosphatidylinositol 3-kinase regulatory subunit alpha
<b>POCS</b>	Polycyst ovary syndrome
<b>POLE</b>	Polimerase $\epsilon$
<b>ppm</b>	Parts per million
<b>PPP2R1A</b>	Serine/threonine-protein phosphatase 2A
<b>PTEN</b>	Phosphatase and tensin homolog
<b>PTM</b>	Post-translational modification
<b>PVDF</b>	Polyvinylidene difluoride
<b>RLP22</b>	Receptor-like protein 22
<b>RNA</b>	Ribonucleic acid
<b>ROC</b>	Receiver operating characteristic
<b>rpm</b>	Revolutions per minute
<b>RT</b>	Radiotherapy
<b>RT-qPCR</b>	Real time quantitative PCR
<b>SDS-PAGE</b>	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
<b>SEGO</b>	Sociedad Española de Ginecología y Obstetricia
<b>SID</b>	Stable isotope dilution
<b>SILAC</b>	Stable isotope labeling by amino acids in cell culture
<b>SN</b>	Supernatant
<b>SRM</b>	Selected reaction monitoring
<b>TBS</b>	Tris-buffered saline
<b>TCGA</b>	The Cancer Genome Atlas
<b>TDE</b>	Tumor-derived exosomes
<b>TEM</b>	Transmission electron microscopy
<b>TMT</b>	Tandem mass tags
<b>TP53</b>	Tumour protein p53
<b>TSPAN</b>	Tetraspanines
<b>TVS</b>	Transvaginal ultrasonography
<b>UA</b>	Uterine aspirate
<b>UMCF</b>	University Medical Center of Freiburg
<b>WHO</b>	World Health Organization
<b><math>\mu</math>g</b>	Micrograms
<b><math>\mu</math>l</b>	Microliters
<b><math>\mu</math>m</b>	Micrometers



# INTRODUCTION

---



# 1. THE ENDOMETRIUM

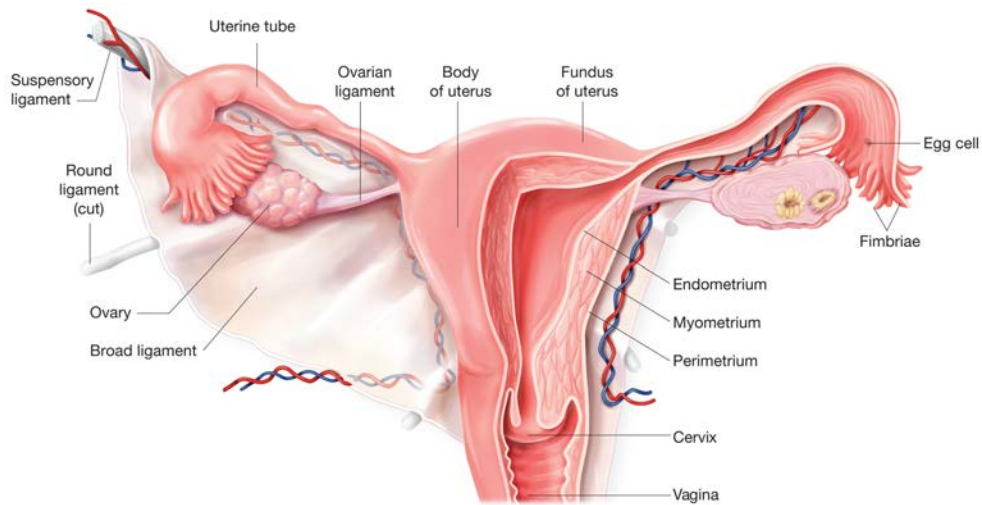
## 1.1 The uterus

The uterus or womb is the organ of the female reproductive system responsible for the development of the fetus during pregnancy. It is located within the pelvic cavity, along the body's midline, posterior to the urinary bladder and anterior to the rectum. It is a hollow, thick walled, muscular organ with the shape of an inverted pear that measures on average 7 cm long, 5 cm wide and 2.5 cm in diameter. The uterus is divided into four anatomic parts:

- The **fundus** is the domed upper portion of the uterus, opposite from the cervix between the points of insertion of the fallopian tubes.
- The **body** or corpus is the wider region of the uterus superior to the cervix.
- The **isthmus** is a slight constriction that separates the body of the uterus and the cervix.
- The **cervix** is the narrow inferior region that connects the uterus to the vagina below and it acts as a sphincter muscle to control the flow of material into and out of the uterus.

The walls of the body of the uterus are thicker than those of the cervix because they provide protection and support to the developing fetus and contain muscles that propel the fetus out of the mother's body during childbirth. Three different tissue layers compose the wall of the uterus: the perimetrium, the myometrium and the endometrium (Figure 1).

The **perimetrium** or *tunica serosa* is the outermost layer that forms the external skin of the uterus. It is a part of the visceral peritoneum that protects the uterus from friction by forming a smooth layer of simple squamous epithelium along its surface and by secreting watery serous fluid to lubricate its surface<sup>1,2</sup>.



**Figure 1.** Anatomy of the uterus. (Image from [www.anatomy-medicine.com](http://www.anatomy-medicine.com))

The **myometrium** or *tunica muscularis* is the intermediate layer that constitutes the thickness of the uterine wall. It consists of three strata of muscle fibers extending in all directions and provides to the uterus the necessary strength during labor contractions<sup>1,2</sup>.

The **endometrium** or *tunica mucosa* is a smooth layer highly vascularized that forms the inner mucosal lining of the uterus<sup>1,2</sup>. It is hormonally regulated and suffers periodic changes, which are the basis of the menstrual cycle. In addition to a population of immune cells, the endometrium is composed of:

- Epithelial compartment: composed of a monolayer of simple, columnar and polarized cells with superficial cilia, lining the uterine cavity together with a luminal and a glandular component.
- Stromal compartment: connective tissue mainly made of fibroblasts and extracellular matrix.
- Vascular compartment: a complex vascular network that starts in the myometrium.

Based on the involvement in the changes of the menstrual cycle, the endometrium can be divided into two layers: the *basalis* and the *functionalis*.

- The *basalis* is formed by the deeper glandular folds highly vascularized. Its thickness remains constant and serves as a regenerative area for the *functionalis*.
- The *functionalis* is composed of columnar epithelium and contains the secreting glands. It is the luminal part of the endometrium. This layer is shed during menstruation and is restored again under ovarian steroid hormones stimulation, with the aim of preparing the endometrium each cycle for the implantation and nutrition of a fertilized egg.

## 1.2 The endometrial cycle

In synchrony with the menstrual cycle, there are a series of changes in the histology of the endometrium that can be described in four phases or states of the endometrium<sup>3</sup>.

### 1.2.1 Proliferative endometrium

The proliferative endometrium is the endometrial tissue during the proliferative phase, right before the ovulation (follicular phase, days 5 to 14). The epithelial layer suffers a cellular proliferation due to an estrogenic stimulation and, consequently, there is an increase in the thickness. At this point, the concentration of estrogen receptors is the highest. The endometrium is composed of rectilinear endometrial glands delimited by pseudo-stratified nuclear cells and mitotic components. There is a dense stroma formed of cells with scant cytoplasm.

### 1.2.2 Secretive endometrium

The secretive endometrium corresponds to the endometrial tissue during the secretory phase, posterior to ovulation (luteal phase, days 14 to 28). The corpus luteum secretes progesterone, which increases the luminal secretions of the glands. The stromal cells become cubical and turn adenomatous. At this time, the endometrium can thicken up to 8 mm. It is during this phase when the endometrium reaches its highest receptivity for the blastocyst implantation, commonly known as the “window of implantation” (days 20 to 24).



### **1.2.3 Menstrual endometrium**

The menstrual endometrium is the endometrial tissue during the menstrual phase (days 1 to 5). If implantation does not occur, within approximately two weeks the corpus luteum will involute, causing a sharp drop in levels of both progesterone and estrogen. The hormone drop causes the uterus to shed its lining in a process termed menstruation. First, there is a structural detachment with disintegration of endometrial glands and stroma, along with leucocyte infiltration and red blood cells extravasation. Later, in a second stage and as a consequence of the hormonal stimulation, the endometrium starts its regeneration from the basal endometrium<sup>3</sup>.

### **1.2.4 Endometrium at pregnancy**

The gestational endometrium is the endometrial tissue when the blastocyst has been implanted. When this happens, human chorionic gonadotropin (hCG) signals the corpus luteum to continue progesterone secretion, thereby maintaining the thick lining of the uterus and providing an area rich in blood vessels in which the zygote can develop.

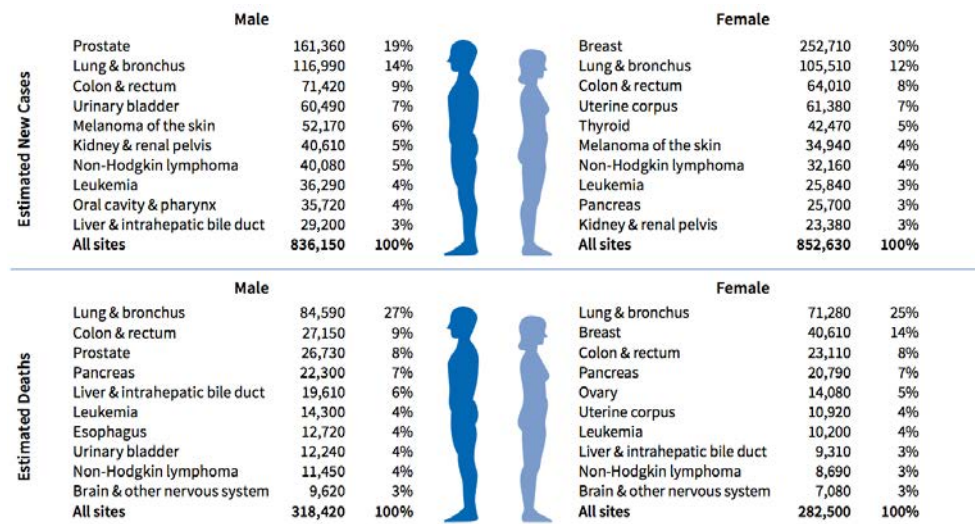
### **1.2.5 Atrophic endometrium**

The atrophic endometrium appears when the woman's reproductive life ends, at the onset of menopause. During this transitional phase, ovulation ceases to occur and, in consequence, there is a lack of progesterone stimuli that triggers important histological changes in the endometrium, which loses its capacity to proliferate and secrete. The epithelial tissue is reduced to a monolayer and the distinction between the basal and the functional layer disappears. The endometrial glands lose their morphology and turn spherical. The stromal component becomes more abundant, with a dense and fibrotic morphology. Estrogen stimuli, however, may continue, as the androgens secreted by the ovaries during menopause can be transformed into estrogens thanks to an aromatase enzyme present in the adipose tissue. High levels of estrogens, especially of estradiol, are usually associated with endometrial hyperplasia, because the hormone can bind its receptors, causing the development of architectural and cytological atypia<sup>1,3</sup>.

## 2. ENDOMETRIAL CANCER

### 2.1 Epidemiology

Endometrial cancer (EC) is the most common gynecologic malignancy of the female genital tract and the fourth most common cancer in women in the United States (US) with 61,380 estimated new cases diagnosed in 2017<sup>4</sup>. In terms of mortality, it occupies the sixth position, with 10,920 estimated deaths in the US (Figure 2). Worldwide, around 320,000 EC cases were estimated during 2012, with an estimated incidence ASR(W) of 8.2/100,000 and an accumulated risk of 0.97% (0-74 years)<sup>5</sup>. In Spain, the estimation was of 5,121 cases diagnosed in 2012, with an accumulated risk of 1.4%. EC is the third more prevalent cancer in Spain<sup>6</sup>. Based on the latest studies available, the incidence of EC in Catalonia is of around 20 cases in 100,000 women<sup>7,8</sup> with 730 new cases reported each year, corresponding to 6,1% of the population<sup>9</sup>.



Estimates are rounded to the nearest 10, and cases exclude basal cell and squamous cell skin cancers and in situ carcinoma except urinary bladder.

©2017, American Cancer Society, Inc., Surveillance Research

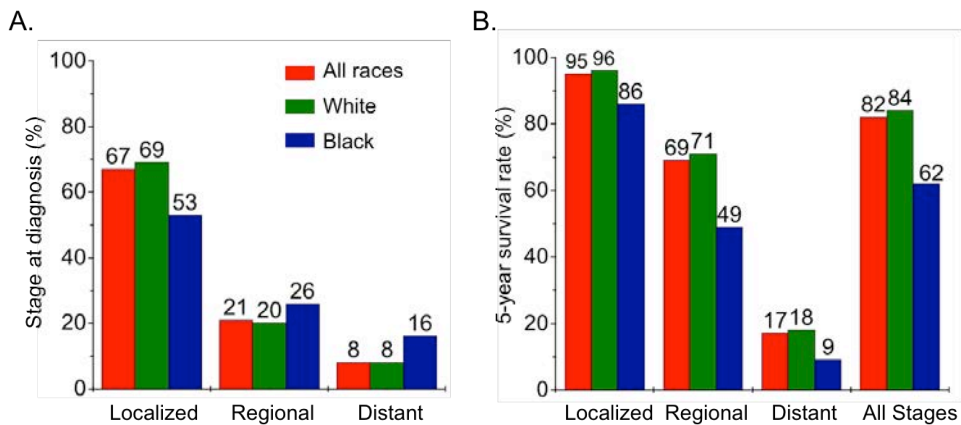
**Figure 2.** Leading sites of new cancer cases and deaths estimated for 2017. Image by Siegel et al.<sup>10</sup>.

The incidence of some types of tumor has grown progressively in the last years and it is expected to grow even more in the upcoming years<sup>6</sup>. In line with this, from 2004 to 2013, the incidence rate of EC increased by 1% per year among white

women and by 3% among black women<sup>10</sup>, probably due to factors like a raise in life expectancy of the population or an increased proportion of women with obesity (and its associated pathologies such as diabetes and hypertension)<sup>11</sup>.

As well, cancer death rate is increasing for EC. From 2005 to 2014, EC death rate increased by about 1% per year among white and 2% among black women<sup>10</sup>.

The elevated incidence rates of EC do not translate into high mortality rates. Thanks to the early presentation of disease-related symptoms (see section 2.4.1), EC is mostly detected at its initial stages, when the tumor is still confined to the uterus and presents a good prognosis and a 5-years survival rate of 95% (Figure 3A). But, unfortunately, there are still 30% of patients diagnosed at advanced stages of the disease, and present a bad prognosis and a drastic decrease in the 5-years survival rate, which goes down to a 68% in cases of regional metastasis, and even lower (17%) if the metastasis is distant (Figure 3B).



**Figure 3.** Endometrial cancer stage and 5-year survival rate by race. (A) Distribution of endometrial cancer by race and stage at diagnosis, United States, 2006 to 2012. (B) Five-year survival rates by stage at diagnosis and race, United States, 2006 to 2012. Image adapted from Siegel et al.<sup>4</sup>.

## 2.2 Risk factors

Although the etiology of EC remains unclear, there are several risk factors proven to be associated with its development. More than 80% of EC cases correspond to the Type I or endometrioid endometrial cancer (EEC) and these are mainly caused by estrogen exposure excess. In contrast, the expression of this steroid receptor decreases with higher tumor stage and grade and is usually absent in Type II

tumors<sup>12</sup> (see section 2.6 for EC classification). The most important risk factors are: (1) the association with long-term exposure to endogenous or exogenous estrogen, (2) obesity, (3) diabetes mellitus and hypertension, and (4) genetic factors.

### **2.2.1 Long-term unopposed endogenous and exogenous estrogen exposure**

Estrogens are steroid sexual hormones produced by the ovaries and by the placenta during pregnancy and, in a smaller proportion, by the adrenal glands. Prolonged exposure to estrogen (especially unopposed by progesterone) promotes uncontrolled cell proliferation of the endometrium leading to an increase in its thickness. Further, through a complex downstream cascade of transcriptional changes that may include the modulation of tumor suppressor functions, estrogen inhibits apoptosis and may also increase the rate of mutagenesis through free radical formation<sup>13</sup>. These alterations may end up causing endometrial hyperplasia, considered a precursor lesion of EC (see section 2.5.2).

#### *2.2.1.1 Endogenous estrogen exposure*

Common conditions related to excessive endogenous estrogen (or hyperestrogenism) and, consequently, to an increased risk of developing EC are: age, early menarche, late menopause, infertility, nulliparity and chronic anovulation.

- **Age**

EC occurs in women older than 50 years in more than 90% of the cases and is detected at a mean age between 62.6 and 68.7 years, with the highest number of cases detected at the age of 65<sup>14</sup>. Even though EC is traditionally thought to be a cancer of the postmenopausal period, 14% of cases are diagnosed in premenopausal women, 5% of whom are younger than 40 years<sup>15,16</sup>. Estrogens have a bigger impact after menopause, because the compensatory levels of progesterone produced by the ovaries before the menopause are no-longer secreted, leading to unopposed estrogen excess.

- **Menarche and menopause**

Early menarche is related to an earlier onset of ovulatory cycles and, thus, an earlier exposure to estrogens. Late age at menopause is also associated to a longer exposure to the hormone. Both situations lead to an increased number of menstrual cycles<sup>17</sup> and are risk factors for the development of EC<sup>17-19</sup>.

- **Infertility and nulliparity**

Both, infertility and nulliparity are risk factors for the development of EC, probably related to the high frequency of anovulatory cycles. In women older than 40 years, infertility is one of the main causes of EC<sup>20</sup>, because irregular menstrual cycles or infrequent ovulations together with chronic anovulations are associated with estrogen production and progesterone deficiency<sup>21</sup>. Furthermore, absence of pregnancy and lactation in nulliparity is translated into a higher number of ovulatory menstrual cycles, associated with a 2-to 3-fold increase in the risk of EC<sup>22</sup>.

- **Chronic anovulation**

The most frequent ovulatory disorder causing chronic anovulation and, hence, chronic infertility if it is not treated, is the polycystic ovary syndrome (PCOS). Women diagnosed with PCOS triplicate the risk of developing EC, as this syndrome is characterized by progesterone deficiency, which provokes irregular menstrual cycles or even anovulation<sup>23,24</sup>.

### *2.2.1.2 Exogenous estrogen exposure*

- **Estrogen therapy**

To palliate the menopause symptomatology, estrogen has been used as a therapy, although it has been demonstrated that its use (without progesterone) increases 5 times the risk of developing EC as well as increases the incidence of hyperplasia from 20 to 50 % after one year of treatment. This risk is dose and time dependent, and persists even after ceasing the treatment<sup>25-28</sup>.

- **Hormone replacement therapy**

When it was demonstrated that postmenopausal women under replacement treatment based exclusively on estrogens had an increased incidence of EC, a new therapy combining estrogen and progestin started to be used. Progestin, as an estrogen antagonist, acts diminishing the unopposed estrogen exposure<sup>25,29</sup>.

Despite most studies have also reported an increased risk of EC with this combined therapy there is still some controversial thoughts, as other studies have shown a decrease in the risk or even no association at all<sup>22,28,30-32</sup>.

#### ▪ Tamoxifen

Tamoxifen is the usual hormonal anti-estrogen therapy for hormone receptor-positive breast cancers<sup>33</sup>. Even though tamoxifen is an antagonist of the estrogen receptor in breast cancer, it behaves as an agonist in other tissues such as the endometrium. It seems that the endometrial activity under tamoxifen influence depends on the menopausal status<sup>34</sup>. While the increased risk of EC with tamoxifen in postmenopausal women is well described, there is no evidence of the same effect for premenopausal women.

### 2.2.2 Obesity

It has been demonstrated that obesity is one of the main factors responsible for increasing incidence of EC in developed countries<sup>35-37</sup>, as obese postmenopausal women tend to suffer from chronic estrogenic stimulation, due to a higher rate of conversion of adrenal precursors into estrogen and estradiol occurring in the adipose tissue. In addition, the adipose tissue is rich in aromatase enzymes, which convert androgens (produced by the adrenal cortex and postmenopausal ovaries) into estrogens<sup>38</sup>. Also, obesity is associated with higher levels of insulin and IGF1; both of them are ligands of the known EC activated signaling pathway PI3K<sup>36,39</sup>.

### 2.2.3 Diabetes and hypertension

Different studies have described the association of diabetes mellitus<sup>17,40,41</sup> and hypertension<sup>41,42</sup> to EC, though the exact causes are still not well understood. One possible explanation is the fact that both conditions are frequently associated with obesity and, thus, to an increased estrogen exposure<sup>43,44</sup>. The tumorigenic effect of insulin might be mediated by the insulin receptors in the target cells or by alterations in the endogenous hormonal metabolism as a result of the hyperinsulinemia, promoting the synthesis of IGF1, which regulates cell proliferation<sup>45</sup>.

## 2.2.4 Genetic factors

The hereditary component related to an increased risk of EC only accounts for roughly a 5 to 10% of all the reported cases. The main inherited factors associated with an increased risk of developing endometrial cancer are the lynch syndrome and the BRCA mutation.

### 2.2.4.1 Lynch syndrome

Lynch syndrome, also referred to as *hereditary non-polyposis colorectal cancer* (HNPCC), is an inherited disorder based on an autosomal dominant mutation that impairs DNA mismatch repair. The prevalence of this syndrome in the general population is approximately 1 in 500 to 1 in 10,000, and EC is the second most common cancer in patients carrying the mutation<sup>46,47</sup>.

### 2.2.4.2 BRCA

Women with mutated BRCA genes present higher risk of developing breast and ovarian cancer and, even though some studies suggest that BRCA1 mutations are associated with higher chances of developing serous EC, the relation is not clear enough yet<sup>48,49</sup>.

## 2.3 Protective factors

In contrast to what has been said for the risk factors, decreased estrogen exposure is associated with a lower incidence of EC. Most of the identified protective factors are related to increased levels of progesterone, resulting in the thinning and atrophy of the uterine glands, such as delayed menarche, women with several pregnancies and/or longer lactation periods and use of oral contraceptives<sup>25,50,51</sup>.

In spite of the numerous adverse effects on health, smoking has been associated to a decreased risk of EC, probably due to its anti-estrogenic effect<sup>18,52,53</sup>.

And finally, among numerous beneficial effects on health, physical activity also offers protection against EC<sup>54</sup>.

## 2.4 Diagnosis

### 2.4.1 Signs and symptomatology

The most common symptom of EC is abnormal uterine bleeding (AUB), which is present in around 90% of the patients. Pre- and perimenopausal women can experience irregular bleeding periods due to the hormonal changes<sup>55,56</sup> but when this happens in postmenopausal woman, is cause for alarm and should be evaluated by clinicians<sup>57</sup>. The probability of EC in postmenopausal women presenting AUB is about 5-10%, but this risk increases with age and when other risk factors are present<sup>58</sup>. Despite being the most frequent symptom of EC, there exists a spectrum of potential causes of AUB<sup>59</sup>.

Other recurrent symptoms of EC include: lower abdominal pain or pelvic cramping, thin white or clear vaginal discharge in postmenopausal women, alterations in bowel or bladder functions, anemia and shortness of breath. Nevertheless, these are less frequent and/or associated with more advanced stages of the disease.

### 2.4.2 Screening recommendations

Despite the absence of reliable screening tools, EC is usually diagnosed at early stages thanks to the appearance of AUB. Nevertheless, some screening recommendations exist and were revised and updated during the 2014 ESMO-ESGO-ESTRO Consensus Conference on Endometrial Cancer<sup>60</sup> and complement the ESMO clinical practice guidelines<sup>61</sup>.

There is **no indication that population-based screening has an impact** in the early detection of EC among women within the average EC risk and without symptoms. Also, there is no evidence that screening by ultrasonography reduces mortality in these cases. On the other hand, screening asymptomatic women with Lynch syndrome is recommended<sup>62,63</sup>.

**Women at increased risk for EC** due to a history of unopposed estrogen therapy, women with late menopause, tamoxifen therapy, nulliparity, infertility or failure to ovulate, obesity, diabetes or hypertension should be properly informed of the risks and symptoms of endometrial cancer and strongly encouraged to report any



unexpected bleeding or spotting to their gynecologist, though routine surveillance in asymptomatic women of these category is not necessary.

**Women with high risk for endometrial mutation** include those diagnosed with HNPCC, with a family history of the mutation and those without genetic testing results but from families with a suspected autosomal dominant predisposition to colon cancer. For this group of women, an annual screening beginning at age 35 is recommended. As the efficacy of the available screening tools is limited, if the maternal wish has been fulfilled, particularly by age 35-40 years, prophylactic surgery is an option that should carefully be considered<sup>64</sup>.

### **2.4.3 Suspected diagnosis: clinical examination**

Any women presenting with AUB and/or any other symptom related to EC is susceptible to suffer the disease and hence should be examined. The standard strategy consists in pelvic examination and transvaginal ultrasonography (Figure 4).

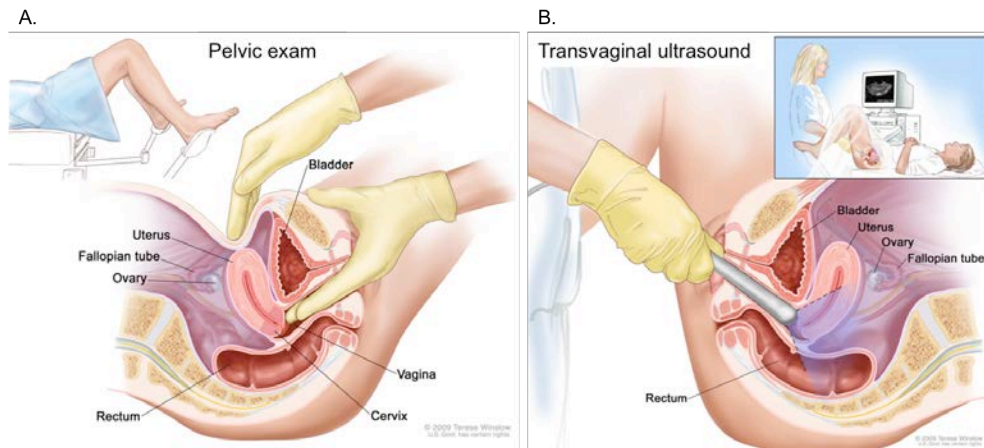
#### *2.4.3.1 Pelvic examination*

The pelvic examination is performed by a gynecologist who, first, inspects the vulva for irritations, lesions or abnormal vaginal discharge, and then also the internal organs to evaluate if they are enlarged or tender. A speculum is used to examine the cervix and the vaginal walls. Normally, no significant alterations are found with respect to the size, shape and consistency of the uterus in early stages of the disease but are typically encountered in advanced stages.

#### *2.4.3.2 Transvaginal ultrasonography*

Transvaginal ultrasonography (TVS) is the diagnostic imaging technique of choice to evaluate the endometrium of women with AUB<sup>11</sup>. This is a simple and non-invasive technique that allows discarding pathologies such as polyps or myomas as well as assessing the thickness of the endometrium. To determine when an increase of endometrial thickness is caused by EC, a cut-off of <5 mm has been set, with an associated sensitivity and specificity of 90% and 54%, respectively<sup>65,66</sup>. A recent meta-analysis showed that, lowering the cut-off reached a diagnostic accuracy of 95% sensitivity and 47% specificity if set to ≤4 mm, and of 98% sensitivity and 35% specificity in case of lowering the cut-off to ≤3 mm<sup>67</sup>. Despite

TVS presents a high sensitivity, the low specificity is a handicap, as other benign conditions increase the endometrial thickness and, hence, a definitive diagnosis usually requires endometrial sampling.



**Figure 4.** Clinical examination. (A) Pelvic exam. (B) Transvaginal ultrasonography. Image from [teresewinslow.com](http://teresewinslow.com).

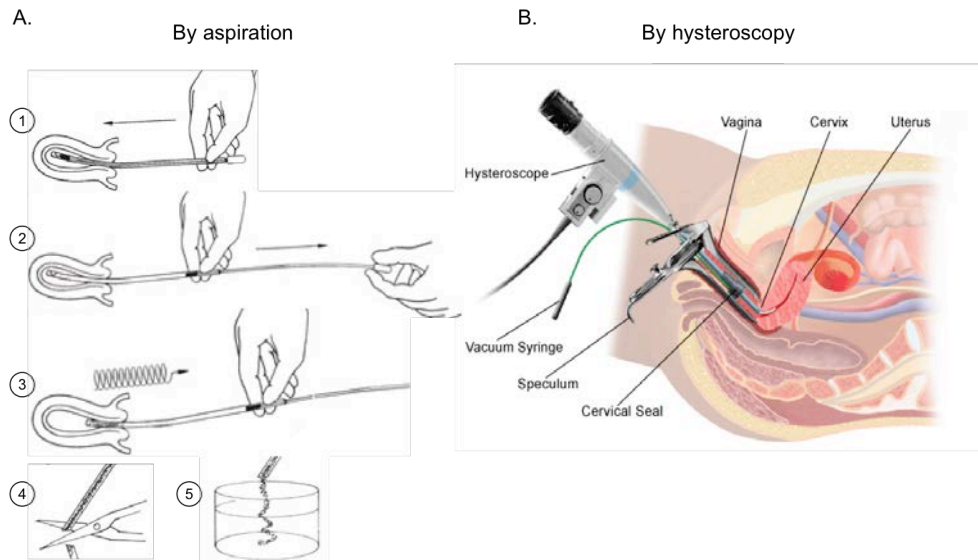
#### 2.4.4 Confirmatory diagnosis: pathological examination

When clinical suspicion of EC exists, a pathological examination of an endometrial biopsy must be performed to diagnose or exclude malignancy. There are two main procedures to obtain an endometrial biopsy (a small sample of the uterine lining) to be observed microscopically by the pathologists for abnormal cells: by aspiration or by hysteroscopy (Figure 5).

**Endometrial biopsy by aspiration** is performed blindly, by aspiration of the endometrial fluid from inside the uterine cavity using a soft straw-like device (pipelle). Even though several studies reported a sensitivity of more than 90% and specificity higher than 80% for the detection of EC with a biopsy obtained by aspiration<sup>65,68</sup>, unfortunately this process has a diagnostic failure and an inadequate sampling rate of 8% and 15%, respectively; which is increased in postmenopausal women up to 12% and 22%<sup>69</sup>. In those undiagnosed cases, a biopsy guided by hysteroscopy needs to be performed.

**Endometrial biopsy guided by hysteroscopy** is obtained by first placing the hysteroscope in the vagina to enable the visualization of the uterine cavity and

then, introducing the catheter (a thin tube with an internal piston) through the cervical opening into the uterine cavity to collect small pieces of selected endometrial tissue.



**Figure 5. Endometrial biopsy.** (A) Endometrial biopsy obtained by aspiration. (B) Endometrial biopsy obtained by hysteroscopy. Image adapted from Genesis Medical Ltd. and teresewinslow.com.

Comparing both methods, the aspiration is less expensive, less painful, does not require anesthesia, is faster and no dissemination of EC cells occurs during the collection of the sample, while the hysteroscopy is more expensive, requires previous blood tests as well as, sometimes, anesthesia. Also, it is considered an invasive procedure that presents more complications, including uterine perforation, hemorrhage and possible damage to other organs<sup>70</sup>. Moreover, it has been reported an increased risk of dissemination of EC cells in the peritoneal cavity in women who underwent this procedure<sup>71</sup>.

For all the reasons stated, the Spanish Society of Gynecology and Obstetrics (SEGO) recommends the endometrial biopsy by aspiration as the first method of choice for diagnosis. Nevertheless, when this is not suitable for the patient for any reason or when the results are not conclusive, a biopsy by hysteroscopy should be performed for a confirmatory diagnosis<sup>11,72</sup>.

As mentioned earlier, although AUB is the most frequent symptom of EC, there are other benign pathologies associated with this symptomatology. In fact, only 5-10% of women with AUB will be diagnosed with EC. Next, endometrial lesions presenting with AUB associated with progression to EC are reviewed.

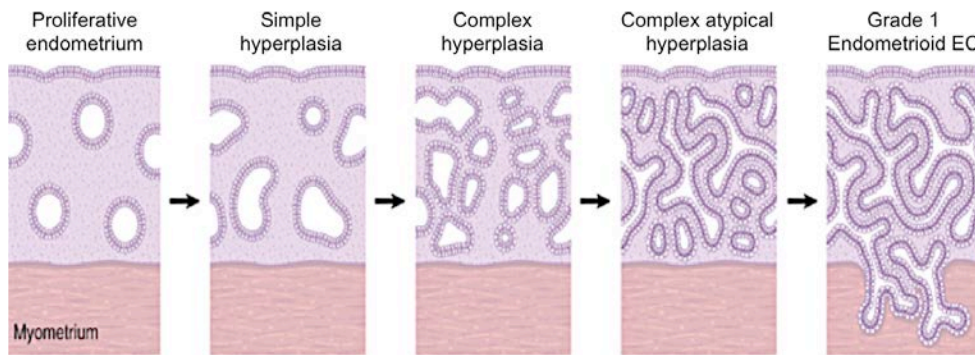
## 2.5 Endometrial lesions

### 2.5.1 Endometrial polyps

Endometrial polyps or uterine polyps are the most common cause of abnormal uterine bleeding in both premenopausal and postmenopausal women<sup>73,74</sup>. These lesions are hyperplastic overgrowths of endometrial glands and stroma that range in size from a few millimeters to several centimeters and can attach to the uterine wall by a large base or a thin stalk<sup>75</sup>. The stroma within the polyps may resemble that of a proliferative endometrium, but it is often fibrotic. There may appear as just one or many polyps and, in fully developed polyps, the presence of large, thick-walled blood vessels is usual. Endometrial polyps are usually benign, but they may contain premalignant or malignant tissue changes<sup>1</sup>.

### 2.5.2 Endometrial hyperplasia

Endometrial hyperplasia is a condition that involves the presence of an abnormal proliferation of the endometrium that results in increased volume of endometrial tissue with alterations of glandular architecture and an endometrial gland to stroma ratio greater than 1:1. Endometrial hyperplasia incidence is vaguely reported to be around 200,000 new cases per year in Western countries<sup>76</sup>. Although the pathogenesis of this lesion remains unclear, most of the cases are thought to result from excessive or unopposed estrogen stimulation, though in some cases is the result of an aberrant response of endometrial glands to normal levels of estrogen<sup>77</sup>. The World Health Organization (WHO) system is known as the classical classification system, and in 1994 proposed the four most widely used categories: **(i) simple hyperplasia, (ii) complex hyperplasia, (iii) simple atypical hyperplasia and (iv) complex atypical hyperplasia**, using glandular complexity and nuclear atypia as criteria<sup>78,79</sup>.



**Figure 6.** Progression from endometrial hyperplasia to EC. Image from Robins et al. 2015<sup>80</sup>.

The risk of progression to an endometrial neoplasm is different for each type of hyperplasia. Both simple and complex hyperplasias without atypia regress if estrogen excess is withdrawn. Only 1% of simple hyperplasias and 3% of complex hyperplasias progress to endometrial carcinoma. But the risk of progression to carcinoma of hyperplasias with atypia is higher, 8% for simple hyperplasia with atypia and up to 29% for complex hyperplasia with atypia<sup>81</sup>.

A major weakness of the WHO system is the high inter-observer variability among pathologist reviewing the same slide to evaluate nuclear atypia, which is the most important indicator of malignancy<sup>82,83</sup>. A new classification system that reduces this inter-observer variability is the one proposed by the International Society of Gynecological Pathologists (ISGP), who introduced the term **Endometrial Intraepithelial Neoplasia (EIN)**<sup>84</sup>.

Conceptually, the EIN system is the integration of all data available from the last fifty years into an evidence-based diagnostic schema that matches current concepts of pathogenesis. It is based on objective correlation of computer-assisted morphometric data with molecular and clinical progression annotations using hematoxylin-eosin-stained slides. This new system segregates biopsies in 3 categories: (i) benign (benign endometrial hyperplasia), (ii) premalignant (endometrial intraepithelial neoplasia), and (iii) malignant (well-differentiated endometrioid EC). There are three morphometric variables with predictive power to separate endometrial biopsies by clinical cancer outcome: the volume percentage of stroma (amount of stroma in a given area, or gland to stroma ratio), the glandular complexity, and the nuclear atypia (assessed by the standard deviation of the shortest nuclear axis, which is a measure of pleomorphism). The

combination of these morphometric variables into a mathematic formula results in the D-score, which correlates with progression of EIN to EC as follows:  $D < 1$ , high rate of progression to EC;  $D > 1$ , benign or really low probability to progress to EC, specifically the endometrioid or type I subtype (see section 2.6.1)<sup>85</sup>.

In the recent years, the EIN classification system has gained numerous adepts among the pathology community. Besides its improved reproducibility, the EIN system has been confirmed as prognostic in several retrospective studies<sup>86</sup> and establishes clear treatment categories that facilitates therapeutic decisions.

## 2.6 Classification of endometrial cancer

EC is compounded by a biologically and histologically diverse group of neoplasms with different pathogenesis. To reflect this heterogeneity on current classification systems parameters such as anatomical site of the tumor, histological type and grade, and the clinical and pathological extent of the disease are evaluated.

### 2.6.1 Dualistic model – Clinico-pathological classification

The dualistic model of classification of EC has been used since 1983, when it was first described by Bokhman et al.<sup>87</sup>, and it is based on clinical, pathological and molecular features. Following this classification model, two categories have been described: **type I or endometrioid endometrial cancer (EEC)**, and **type II or non-endometrioid endometrial cancer (NEEC)**. Table 1 summarizes the main features of EEC and NEEC tumors.

#### 2.6.1.1 Type I (EEC)

Type I is the most common subtype, representing around 80-90% of ECs<sup>60</sup>. These type of tumors usually develop in pre- or perimenopausal women and are usually associated with previous endometrial hyperplasia with or without atypia<sup>78</sup>. Type I tumors include the endometrioid and mucinous histology. They express estrogen and progesterone receptors. They are commonly diagnosed as well-differentiated and low-grade tumors, and patients with EEC present an overall good prognosis.

Regarding molecular alterations, type I ECs present a high mutational frequency and microsatellite instability<sup>88</sup>. The PI3K pathway is mutated in more than 90% of the cases, including simultaneous loss of PTEN and PIK3CA mutations<sup>89</sup>. Beta-

catenin is mutated in 20-25% of the cases<sup>90</sup>, KRAS in about 20% and FGFR2 in 12% of tumors<sup>91</sup>.

**Table 1.** Clinico-pathological classification of endometrial cancer. Adapted from Morice et al. 2016<sup>92</sup>.

Features	Type I (EEC)	Type II (NEEC)
<b>Age</b>	Pre- and perimenopausal	Postmenopausal
<b>Histology</b>	Endometrioid, Mucinous	Serous, Clear-cell
<b>Prevalence</b>	80-90%	10-20%
<b>Diagnosis</b>	Early stages	Advanced stages
<b>Grade</b>	Low-grade	High-grade
<b>Hormone dependence</b>	Hormone-dependent	Hormon-independent
<b>Growth</b>	Slow evolution	Aggressive evolution
<b>TP53 mutations</b>	Rare	90% Serous, 35% Clear-cell
<b>PI3K alterations</b>	PTEN mutation (75-85%) PIK3CA mutation (50-60%)	· Serous: PTEN mutation (11%), PIK3CA amplification (45%), PIK3CA mutation (35%) · Clear-cell: PTEN loss (80%), PIK3CA mutation (18%)
<b>KRAS mutation</b>	20-30%	3% Serous
<b>EGFR mutation</b>	12%	5% Serous
<b>ERBB alterations</b>	None	· Serous: ERBB amplification (21-47%) · Clear-cell: ERBB mutation (12%) and amplification (16%)
<b>Prognosis</b>	Good (overall survival 85% at 5 years)	Poor (overall survival 55% at 5 years)

### 2.6.1.2 Type II (NEEC)

Type II EC only represents around 10-20% of the cases and occur normally in older, postmenopausal woman. In contrast with type I, type II are non-estrogen associated tumors. The most common histology is the serous followed by the clear-cell. Type II tumors are normally diagnosed at advanced stages of the disease and about 20% of the patients present with myometrial invasion and/or lymph node affection. They are all high-grade tumors and significantly correlate with a worse prognosis.

Regarding molecular alterations, type II tumors are described to bear chromosomal instability and TP53 mutation<sup>93</sup>, but each histological subtype presents different characteristics. Serous tumors present TP53 mutations in up to 90% of the cases<sup>94</sup> and HER2/neu (ErbB2) overexpression in 14% to 80% of the cases, with HER2 amplification ranging from 21% to 47%<sup>95</sup>. Also, serous tumors present PIK3CA

amplification in 45% and mutation in 35% of the cases. Tumors with clear cell histology present loss of PTEN in about 80% of the cases and inactivation mutations in the ARID1 gene in 20-40% of the cases<sup>96</sup>.

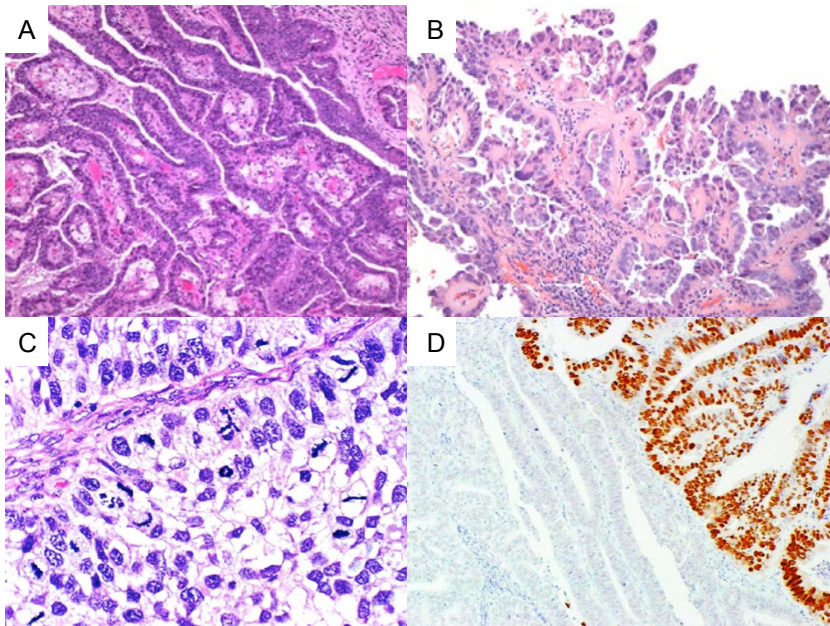
## 2.6.2 Histological type

Tumors of the endometrium comprise several distinct histological types that are associated to different overall prognosis in each case. Thus, EC should be typed according to the 2014 WHO classification<sup>97</sup> as follows.

The Type I or **endometrioid ECs** (EEC) are characterized by an endometrioid histology, which covers a wide range of different glandular growth patterns and differentiation grades<sup>98</sup>. The WHO has described several subtypes of EEC: adenocarcinoma with squamous differentiation (the most common variant), villoglandular carcinoma, secretory carcinoma and also mucinous adenocarcinomas.

Among the non-endometrioid ECs (NEEC), **serous** and **clear cell** are the most frequent subtypes<sup>98</sup>, and both are highly invasive and aggressive tumors associated with a poor prognosis. Other NEEC are the **neuroendocrine**, the **undifferentiated** and the **dedifferentiated** carcinomas. **Mixed adenocarcinomas** are also frequent, and are tumors composed of two or more different histological subtypes, in which the minor represented type must comprise a minimum of 5% of the total tumor volume, and at least one of the subtypes must be a type II; it is known that  $\geq 25\%$  of representation of type II is associated with a worse prognosis<sup>99</sup>. Finally, **carcinosarcomas**, or malignant mixed Müllerian tumors, are mixed epithelial and mesenchymal tumors that receive the same treatment as aggressive type II tumors<sup>97</sup>.





**Figure 7. Histology of 4 common types of EC. (A) Endometrioid adenocarcinoma, grade 1. (B) Serous adenocarcinoma. (C) Clear cell adenocarcinoma. (D) Mixed adenocarcinoma composed of endometrioid carcinoma (left) and serous (right). The brown staining is nuclear p53, only positive in the serous component**

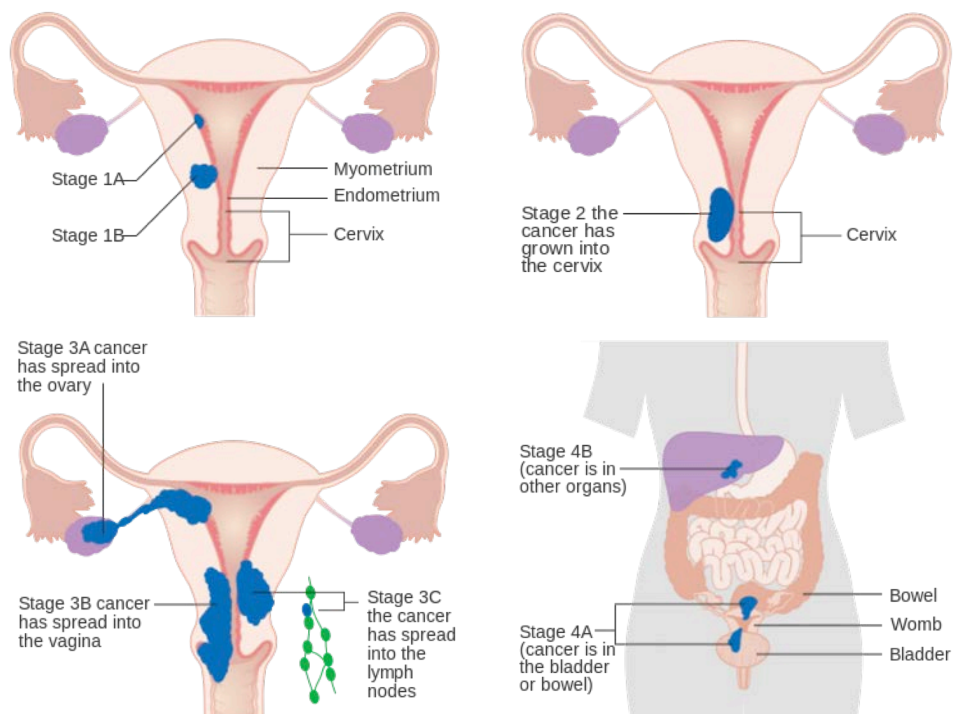
### 2.6.3 FIGO staging

The International Federation of Gynecology and Obstetrics (FIGO) created their EC classification and staging system in 1958. The system changed from a clinical to a surgical staging strategy in 1988 and updated in 2009, in order to solve problems of reproducibility, accuracy and predictive value, detected during the previous years.

The FIGO staging is a description of the extent to which the cancer has spread and it takes into account the following parameters: tumor size and location, myometrial invasion, cervical involvement, extension of tumor to the Fallopian tubes and the ovaries, lymphovascular space invasion, lymph nodes affection, tumor grade and histology. The different stages are numbered progressively from I to IV, and are summarized in Table 2 and depicted in Figure 8.

**Table 2.** 2009 FIGO staging, adapted from Plataiotis et al. <sup>100</sup>.

FIGO Stage	Description
<b>I</b>	Tumor confined to the corpus uteri
<b>IA</b>	No or <50% of myometrium invasion
<b>IB</b>	≥50% of myometrium invasion
<b>II</b>	Tumor invades cervical stroma but does not extend beyond the uterus
<b>III</b>	Local and/or regional spread of the tumor
<b>IIIA</b>	Tumor invades serosa of the corpus uteri and/or adnexae
<b>IIIB</b>	Vaginal and/or parametrial involvement
<b>IIIC1</b>	Positive pelvic lymph nodes
<b>IIIC2</b>	Positive paraortic lymph nodes with or without pelvic nodes
<b>IV</b>	Tumor invades bladder/bowel mucosa, and/or distant metastases
<b>IVA</b>	Tumor invasion of bladder and/or bowel mucosa
<b>IVB</b>	Distant metastases including intra-abdominal and/or inguinal lymph nodes

**Figure 8.** FIGO stages. Image from Cancer research UK (CRUK).

## 2.6.4 Molecular classification: the TCGA model

Although the dualistic model has been broadly used in clinical-decision making algorithms defining high-risk patients, its prognostic value remains limited because 20% of type I ECs relapse, whereas 50% of type II cancers do not. Moreover, 15-20% of endometrioid tumors are high-grade, and where they fit into the dualistic model is unclear. This exemplifies that some ECs present shared features of both classification groups.

Analyses of The Cancer Genome Atlas (TCGA) focusing on endometrioid and serous EC further emphasize the heterogeneity of the disease by identifying 4 different molecular subgroups and proposing a new classification model based on an integrated genomic characterization for ECs. The TCGA research network used array and sequencing based technologies to characterize the genomic, transcriptomic and proteomic profile of 373 endometrial carcinomas. The 4 new molecular subtypes of EC are described below and summarized in Table 3<sup>101</sup>:

**POLE ultramutated:** around 10% of endometrioid ECs fall in this category, which is the smallest group, and defines a unique subset that is characterized by mutations in the exonuclease domain of POLE, high mutation load and an excellent prognosis. About 60% of POLE ultramutated ECs are high-grade endometrioid lesions. It comprises few copy-number aberrations and increased frequency of C → A conversions.

**Microsatellite instability hypermutated:** this group is characterized by MLH1 promoter methylation and a high mutation rate. Involves endometrioid tumors and PTEN is the most commonly mutated gene.

**Copy-number-low:** this group comprises mainly endometrioid grade 1 or 2 tumors, with a low mutation rate, affecting mainly the PTEN gene.

**Copy-number-high:** this group is mostly composed of serous-like tumors. The most frequent mutation occurs in TP53.

**Table 3.** The four genomic classes identified by The Cancer Genome Atlas Network by combining information on mutations, copy-number aberrations, and microsatellite instability. General characteristics of the genomic classes are shown. Mb=megabase. MSI=microsatellite instability, adapted from Murali et al.<sup>102</sup>.

	<b>POLE (ultramutated)</b>	<b>MSI (hypermutated)</b>	<b>Copy-number low (endometrioid)</b>	<b>Copy-number high (serous-like)</b>
<b>Copy-number aberrations</b>	Low	Low	Low	High
<b>MSI/MLH1 methylation</b>	Mixed MSI high, low, stable	MSI high	MSI stable	MSI stable
<b>Mutation rate</b>	Very high	High	Low	Low
<b>Genes commonly mutated (prevalence)</b>	POLE (100%) PTEN (94%) PIK3CA (71%) PIK3R1 (65%) FBXW7 (82%) ARID1A (76%) KRAS (53%) ARID5B (47%)	PTEN (88%) RPL22 (37%) KRAS (35%) PIK3CA (54%) PIK3R1 (40%) ARID1A (37%)	PTEN (77%) CTNNB1 (52%) PIK3CA (53%) PIK3R1 (33%) ARID1A (42%)	TP53 (92%) PPP2R1A (22%) PIK3CA (47%)
<b>Histological type</b>	Endometrioid	Endometrioid	Endometrioid	Serous, endometrioid, mixed (serous + endometrioid)
<b>Tumor grade</b>	Mixed (Grades 1-3)	Mixed (Grades 1-3)	Grades 1 and 2	Grade 3
<b>Progression-free survival</b>	Good	Intermediate	Intermediate	Poor

## 2.7 Prognostic factors

Although EC is generally detected at early stages, when the tumor is still confined within the uterus, some patients still recur, most of them three years after receiving the first treatment<sup>103</sup>. Hence, it is important to establish accurate predictive and prognostic factors in order to identify subgroups of patients presenting poor prognosis, in order to select the most appropriate primary and adjuvant treatment.

The prognostic factors for EC have long been divided into uterine or extra-uterine factors. Uterine factors include: histological type, histological grade, depth of myometrial invasion, vascular invasion, presence of atypical endometrial hyperplasia, cervical involvement, DNA ploidy and S-phase fraction, and hormone receptor status. On the other hand, extra-uterine factors include: positive peritoneal cytology, adnexal involvement, pelvic and paraaortic lymph node metastasis, and peritoneal metastasis. In general, the most significant factors are the type and

histological grade of the tumor and the extent of myometrial and lymphovascular invasion<sup>104</sup>.

The FIGO stage is the most important individual prognostic factor and it is normally used as a reference, presenting a significant survival reduction with advanced FIGO stages<sup>4</sup> (Figure 3B). Even though regional dissemination to lymph nodes is also considered an important prognostic factor, the role of lymphadectomy in women with early-stage tumors remains controversial since neither global nor recurrence-free survival benefit has been demonstrated<sup>105-107</sup>.

The prognostic factors required to correctly stage a tumor and determine the risk of recurrence of each patient are evaluated before, during and after the surgical treatment. The extent of the surgery is decided based on the **pre-operative staging** (type and grade of the tumor, depth of myometrial invasion and extent of cervical involvement) and the medical condition of the patient. The final staging of the tumor is always determined after surgery and it is then called **clinical staging**.

Clinicians are able to classify patients at different risk of recurrence thanks to the stage and the histology of the tumors as shown in Table 4. This risk group classification system is also used to determine the most suitable adjuvant treatment.

**Table 4.** Risk groups to guide adjuvant treatment. FIGO 2009 is used; Molecular factors and tumor size were considered but not included; nodal status may be considered for treatment recommendations. EEC, endometrioid endometrial cancer; NEEC, non-endometrioid endometrial cancer; MI, myometrial invasion; LVSI, lymphovascular space involvement. Adapted from Colombo et al.<sup>60</sup>.

Risk group	Description
<b>Low</b>	Stage I EEC, grade 1-2, <50% MI, LVSI negative
<b>Intermediate</b>	Stage I EEC, grade 1-2, ≥50% MI, LVSI negative
<b>High-intermediate</b>	· Stage I EEC, grade 3, <50% MI, regardless of LVSI status · Stage I EEC, grade 1-2, regardless of MI, LVSI unequivocally positive
<b>High</b>	· Stage I EEC, grade 3, ≥50% MI, regardless of LVSI status · Stage II · Stage III EEC, no residual disease · NEEC (serous, clear cell or undifferentiated or carcinosarcoma)
<b>Advanced</b>	· Stage III residual disease · Stage IVA
<b>Metastatic</b>	· Stage IVB

While the intermediate and high-intermediate risk groups present a 5-years risk of recurrence of about 20-25%, the high and the advance risk groups present a 30-65% and, despite the fact that the non-endometrioid endometrial cancers represent only around a 10% of all the diagnosed endometrial cancers, this group alone accounts for more than 50% of total recurrence and deaths<sup>108-110</sup>.

## **2.8 Pre-operative risk assessment**

Treatment options for EC patients were revised during the 2014 ESMO-ESGO-ESTRO Consensus Conference on Endometrial Cancer<sup>60</sup>. Following their recommendations, an extensive evaluation is mandatory before surgery and must include: family history; general assessment and inventory of comorbidities; geriatric assessment, if appropriate; clinical examination, including pelvic examination; transvaginal or transrectal ultrasound; and complete pathology assessment (histological type and grade) of an endometrial biopsy or curettage specimen.

The main role of preoperative risk assessment is to correctly classify patients into those groups of risk for lymphatic dissemination and disease recurrence to define the most appropriate surgical treatment.

### **2.8.1 Preoperative endometrial biopsies**

Endometrial biopsies obtained by aspiration and/or guided by hysteroscopy not only serve as a confirmatory diagnostic sample, but also to assess the tumor grade and histological subtype of the EC cases. Nevertheless, several studies have reported discrepancies between pre- and post-operative biopsies, which could lead to a misclassification and the use of inappropriate therapeutic strategies<sup>111,112</sup>. It has been shown that between 22% and 40% of EEC classified as grade 1 on the preoperative biopsy were upgraded when analyzing the final surgical specimen; For those EEC preoperatively classified as grade 2, 26% to 56% of cases were upgraded; And, in addition, 8% to 60% of tumors initially classified as grade 3 were downgraded and/or classified as NEEC in the final surgical pathology<sup>113-115</sup>. On the other hand, it has been reported a high concordance between pre- and postoperative pathological determination of up to 93% for high-risk NEEC histologies<sup>115</sup>. Factors that may partially explain these reported discordances

include the interobserver variability and cases with insufficient amount of sample satisfactory for analysis<sup>116,117</sup>.

Importantly, intra-tumor genetic heterogeneity represents a challenge that hampers the correct characterization of tumor samples and, even though diagnosing endometrial cancer based exclusively on genetic alterations is currently unfeasible, it has been reported that the genetic analysis of uterine aspirates is useful to detect this heterogeneity; this study showed that mutations were mainly found in uterine aspirates and not in their paired tissue specimens<sup>118</sup>.

## **2.8.2 Magnetic resonance imaging (MRI)**

MRI is considered the preferred imaging technique for preoperative staging, specially to assess myometrial invasion, and has a high interobserver concordance<sup>119</sup>. However, a major restriction of this technique is the poor detection of lymph node metastases<sup>120</sup>. Transvaginal ultrasonography, if performed by an experienced radiologist, has been suggested to have a similar accuracy to that of MRI for assessment of myometrial and cervical invasion. This technique is less costly than MRI but cannot be used to determine lymph node metastases<sup>121,122</sup>. If MRI is not available, CT (Computed Tomography) can be used to determine extrauterine disease. In general, the major limitation of imaging techniques is poor detection of lymph node metastasis.

Some studies have emphasized the high accuracy of 18F-FDG PET/CT (18F-fluorodeoxyglucose Positron Emission Tomography/Computed Tomography) in detection of myometrial and cervical invasion and lymph node metastatic disease, but, although its prognostic value has been shown for advanced stage cases, its use in preoperative risk assessment in early stage cases remains questionable<sup>123</sup>.

Emerging molecular imaging techniques (i.e. hybrid PET/MRI) might improve diagnostic accuracy by better soft tissue contrast, multiplanar image acquisition and functional imaging<sup>124</sup>.

## 2.9 Treatment

Surgery is the primary treatment of endometrial cancer, and its extent should be adapted to the medical condition of the patient. Minimally invasive surgery is recommended in the surgical management of low-and intermediate-risk endometrial cancer. The surgical procedures recommended depending on the staging of the tumor are summarized on Table 5.

**Table 5. Surgical treatment depending on staging, adapted from Colombo et al.<sup>61</sup>.**

Preoperative staging	Recommended surgical procedure
<b>Stage I</b>	
IA; Grade 1-2	Hysterectomy with bilateral salpingo-oophorectomy
IA; Grade 3	Hysterectomy with bilateral salpingo-oophorectomy ± bilateral pelvic-paraortic lymphadenectomy
IB; Grade 1-2-3	Hysterectomy with bilateral salpingo-oophorectomy ± bilateral pelvic-paraortic lymphadenectomy
<b>Stage II</b>	Radical hysterectomy with bilateral salpingo-oophorectomy and bilateral pelvic-paraortic lymphadenectomy
<b>Stage III</b>	Maximal surgical cytoreduction with a good performance status
<b>Stage IV</b>	
IVA	Anterior and posterior pelvic exenteration Systemic
IVB	Systemic therapeutical approach with palliative surgery
<b>Serous &amp; Clear cell</b>	Hysterectomy with bilateral salpingo-oophorectomy, pelvic and paraaortic lymphadenectomy, omentectomy, appendectomy and peritoneal biopsies

Most of the patients diagnosed with endometrial cancer fall in the low-risk of recurrence group and, consequently, they are solely treated by surgery. Table 6 shows the recommended adjuvant treatment for each stage of the disease.

**Table 6. Recommended adjuvant treatment. NPF, negative predictive factor; BT, brachytherapy; RT, radiotherapy. Adapted from Colombo et al.<sup>61</sup>.**

Stage	Recommended adjuvant treatment
<b>Stage I</b>	
IA; Grade 1-2	Observation
IA; Grade 3	Observation or vaginal BT (if NPF: pelvic RT and/or adjunctive chemotherapy could be considered)
IB; Grade 1-2	Observation or vaginal BT (if NPF: pelvic RT and/or adjunctive chemotherapy could be considered)
IB; Grade 3	Pelvic RT (if NPF: combination of radiation and chemotherapy could be considered)
<b>Stage II</b>	Pelvic RT and vaginal BT · If grade 1–2 tumor, myometrial invasion <50%, negative LVS1 and complete surgical staging: BT alone · If NPF: chemotherapy ± RT
<b>Stage III-IV</b>	Chemotherapy · If positive nodes: sequential RT · If metastatic disease: chemotherapy – RT for palliative treatment



## 3. BIOMARKERS

### 3.1 Biomarker definition and protein biomarkers

A biomarker is a measurable indicator of a specific biological state, particularly a state related to the risk of contraction, the presence, the stage of a disease, or prediction of its development. Even though historically the term “biomarker” included physical traits or physiological metrics (i.e. body temperature is a well-known biomarker for fever, and blood pressure is used to determine the risk of stroke), the term has expanded to molecular biomarkers in the last decades.

The term “**molecular biomarker**” is a broad concept that encompasses a variety of components such as specific cells, proteins, hormones, enzymes, molecules, RNAs, miRNAs, genes and specific mutations, among others. In biomarker research, a broad spectrum of strategies has been adopted for their study, such as transcriptomics, metabolomics, proteomics and immunology. Among the different mentioned components with potential to be disease indicators, **proteins** present remarkable advantages. In the first place, the diversity of existing proteins is much higher than that of DNA or RNA, because alternative splicing and post-translational modifications create multiple protein forms from a single gene; the human genome is composed of about 20,300 genes that can potentially produce up to 1.8 million different proteins<sup>125,126</sup>. This tremendous diversity rises up the chances to identify a specific protein, or a combination of them, associated with a certain condition, such as cancer. Importantly, a great proportion of the proteome is detectable in a variety of body fluids. Proteomics holds special promise for biomarker discovery and, the development of novel protein biomarkers to improve diagnosis in EC is the general focus of this thesis.

### 3.2 Challenges for the discovery of protein biomarkers and source sample selection

The utility and importance of biomarkers has been recognized by substantial public and private funding, and biomarker discovery efforts are now commonplace in both

academic and industrial settings. Despite this, few novel biomarkers are used in clinical practice, and their rate of introduction is failing and reflects the long and difficult path from candidate discovery to clinical utility. Two of the most important challenges that must be faced to succeed in biomarker research include the selection of the most suitable sample as the source of biomarkers and the selection of the appropriate technology.

Traditionally, the ultimate goal of biomarker discovery has been the development of a blood test. Human plasma has been the most used matrix for discovery studies until now. Human plasma is described as the most complete human proteome, because it is a circulating representation of the whole body and of the physiological and pathological processes. The total protein content in plasma is very constant across the general population, facilitating the comparison of specific protein levels between patients. Blood plasma/serum is collected in a rapid, easy and minimally invasive way to be routinely analyzed in clinical laboratories. Despite its advantages, human plasma remains incompletely characterized because it contains tens of thousands of core proteins and span ten to eleven orders of magnitude in protein abundance, hindering the discovery of the low abundance proteins<sup>127</sup>. Thus, this is one of the most challenging samples to be analyzed by proteomic techniques.

Alternative body fluids, and especially proximal bodyfluids, such as urine, cerebrospinal fluid and, for our interest, uterine aspirates, can also be used for discovery studies. In contrast to blood, proximal fluids are in direct contact or close to a certain organ or part of the body; it has been shown that these proximal fluids are highly enriched in proteins derived from diseased tissue and reflect microenvironmental or systemic effects of the disease<sup>128,129</sup>, setting them as an attractive source of biomarkers. Nevertheless, some limitations of proximal fluids should be taken into consideration. Invasiveness to collect each specific fluid might vary, and minimal patient's discomfort should be pursued. Moreover, the proteome of body fluids is much less characterized than blood and is also less constant regarding total protein concentration. Small sample volume, low protein yield and/or frequent blood contamination are some other limitations that must be faced when working with proximal body fluids<sup>130,131</sup>.

In line with this, uterine aspirates might represent an ideal source of biomarkers for EC and other gynecological diseases. As explained earlier, uterine aspirates are the endometrial biopsy obtained by aspiration, and is formed by a mixture of fluid and cells that represent the uterine cavity. Until now, the use of uterine aspirates as a source of biomarkers has focused on the cellular fraction from the transcriptomic and genomic point of view. A work done previously in our group identified and clinically validated a RT-qPCR-based assay of 5 genes that improved EC diagnosis<sup>132</sup>. To date, very few proteomic studies have used uterine aspirates and none of them have used this sample as a source of protein biomarkers for EC<sup>133–135</sup>.

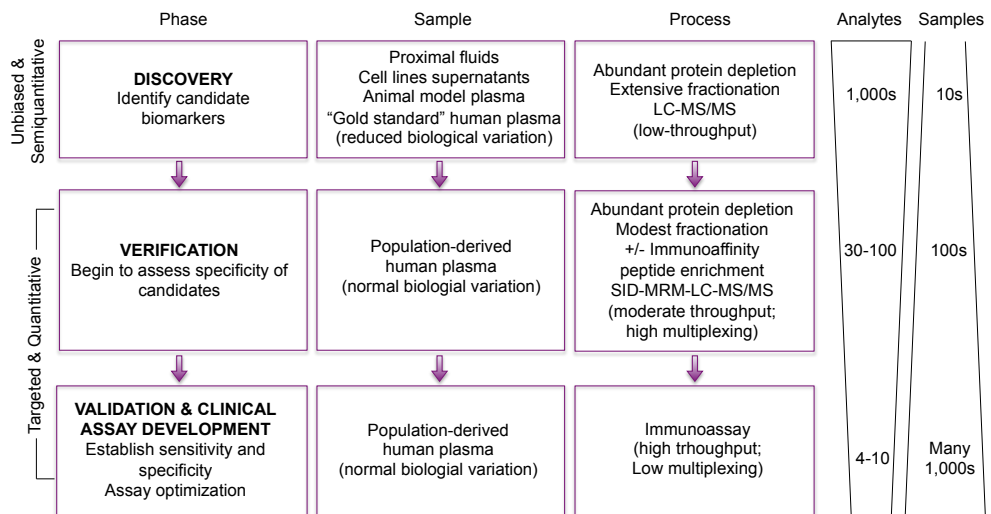
Although single markers may serve in some cases, there is growing consensus that multiple markers used either individually or as part of integrated panel result in a better performance. Comparative proteomics can be used to identify and verify such biomarkers and panels. Up to date, immunoaffinity capture is the most effective method to detect and quantify proteins present at or below nanogram/milliliter levels in blood, where many disease-specific biomarkers are thought to be. Unfortunately, this approach is not suitable for de novo studies, as it depends on available antibodies and, hence, the limitations of affinity approaches in general leave mass spectrometry (MS) as the principal enabling technology for unbiased candidate protein discovery.

## 3.2 The biomarker pipeline

The ideal process flow for the development of novel protein biomarkers is depicted in Figure 9.

The goal of the **discovery phase** is to define, for the first time, the differential expression of specific proteins between states, avoiding “contamination” by other diseases or confounding conditions (i.e. disease vs. normal tissue). In this phase, model systems such as mouse models and cell lines, or a variety of human samples, can be used. The results of a discovery phase are list of proteins found to be differentially expressed between the normal and the disease states based on semiquantitative assessments of relative peptide abundance in the MS data or, more recently, the use of exogenous isotopic labeling. These lists normally contain from tens to several hundreds of candidates and include a certain proportion of

false-positives. Discovery phase lists enter the subsequent phases of the biomarker development pipeline: **verification and validation phases**, which will now replace the unbiased experimental paradigm with targeted quantitative approaches. In the next steps, the analysis is extended to a larger number of samples and the candidate's list is reduced progressively. Validated biomarkers may hopefully be commercialized, if they show promising results.



**Figure 9.** The biomarker pipeline. Process flow for the development of novel protein biomarker candidates. LC-MS/MS, liquid chromatography tandem mass spectrometry; SID, stable isotope dilution; MRM, multiple reaction monitoring. Image adapted from Rifai et al.<sup>128</sup>.

## 4. EXTRACELLULAR VESICLES AND EXOSOMES

### 4.1 Brief history

In multicellular organisms, intercellular communication can be mediated through direct cell-cell contact or transfer of secreted molecules to short or long distances in the organism. In the last three decades, a new consensus that recognizes extracellular vesicles (EVs) as a mode of communication between and among cells and tissues has emerged, changing our understanding of human physiology and opening new possibilities in clinical practice<sup>136,137</sup>. The release of apoptotic bodies during apoptosis has been long studied, but the idea that perfectly healthy cells

can also shed vesicles from their plasma membranes is a concept that only recently has gained interest for the scientific community<sup>136</sup>. Alive cells release two main types of EVs, which differ in size, biogenesis, mechanism of release and molecular composition: the vesicles of endosomal origin, called exosomes, and the vesicles shed from the plasma membrane, named ectosomes, microparticles or microvesicles (MV)<sup>136,138</sup>. There is growing interest on the study of EVs in general but especially exosomes have been the subject of intensive research in recent years and are also the main interest of this work.

Exosomes were first described in the mid-eighties in two papers published virtually at the same time. In 1983, Harding et al. and Pan et al. described exosomes in reticulocyte maturation as small (50-150 nm) membranous vesicles involved in the externalization of the transferrin receptor, released during reticulocyte differentiation as a consequence of multivesicular endosome (MVE) fusion with the plasma membrane<sup>139,140</sup>. The two studies together set the basis for a novel vesicle secretion pathway model, but this new field of research grew slowly and very few publications appeared over the next decade. During that time, exosomes were only considered a means for cells to discard unwanted molecular components. However, in 1996 Garça Raposo published a groundbreaking paper that revived the interest on the field of exosomes biology, suggesting, for the first time, that exosomes could be important mediators of intercellular communication, rather than mere carriers of garbage molecules. They also stated that exosomes might play a role in antigen presentation<sup>141</sup>. Two years later, Zitvogel et al. dug deeper and demonstrated the release of exosomes by human dendritic cells and the ability of those exosomes to suppress the growth of established tumors *in vivo*<sup>142</sup>. These findings were the precursors of the explosion of numerous publications related to the immune function of exosomes *in vitro* and *in vivo*<sup>143,144</sup>. The interest on exosomes in general keeps growing fast and we can find around 5 thousand publications on exosomes to date. In addition to their importance to fundamental mechanisms of intercellular communication, signaling and regulation, exosomes and other extracellular vesicles may have important clinical applications in the future<sup>136</sup>. There is growing interest in the potential for diagnostics based on the analysis of exosomes, as they may bear the protein or RNA signatures of pathological or physiological states of their source cells. Furthermore, exosomes

have been purified from numerous body fluids, such as urine, seminal fluid, breast milk, amniotic fluid, bronchoalveolar fluid, nasal secretions, bile, cerebrospinal fluid, tumor effusions, saliva, and blood plasma<sup>145</sup>. The potential of exosomes as a diagnostic tool is undeniable.

## 4.2 Nomenclature of extracellular vesicles

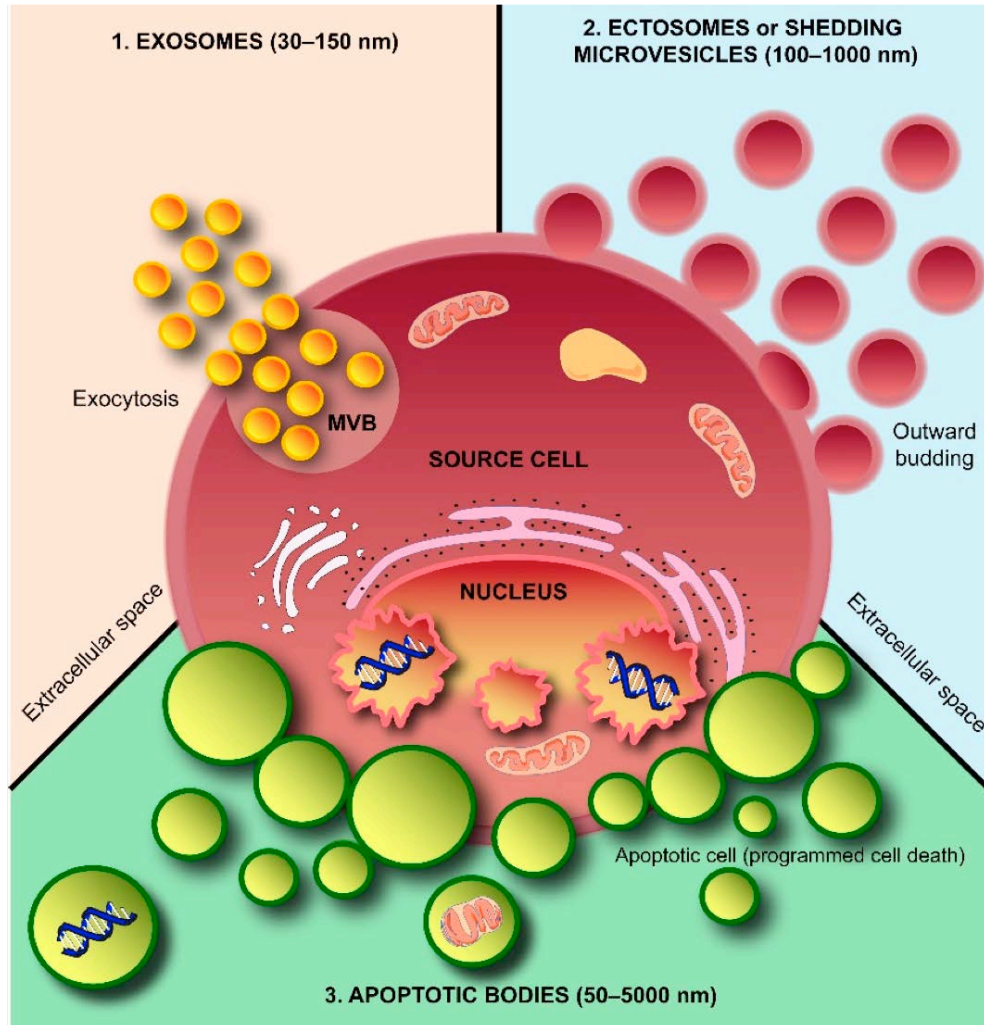
Although EVs constitute a promising platform for biomarker development, the terminology used to describe these vesicles has not been yet standardized and researchers have invented dozens of different names. Most of these names reflect specific functions (i.e. *calcifying matrix vesicles* that initiate bone formation), or refer to their cell of origin (i.e. *prostasomes* released by prostate epithelium). While this diversity might only be useful for specialized fields and generates confusion, more generic names such as “exosomes” and “microvesicle” have broader utility. Lamentably, these generic terms mean different things to different investigators so they lack from a uniform use and are often misleading. Hence, there is an urgent need to reach standardization for EVs nomenclature<sup>146,147</sup>.

Nowadays, it is considered that the most important criteria for classification of EVs are size, density, morphology, lipid composition and subcellular origin, summarized in Table 7 and Figure 10.

**Table 7.** Physicochemical characteristics of different types of secreted vesicles.

\* Appearance by electron microscopy is only an indication of vesicle type and should not be used to define vesicles, as their microscopic appearance can be influenced by the fixation and phase contrast techniques used. ND, not determined. Adapted from Ciardiello et al. 2016) and Théry et al. 2009<sup>138,143</sup>.

Feature	Exosomes	Microvesicles	Apoptotic bodies
Size	30-150 nm	50-1000 nm	50-5.000 nm
Density in sucrose	1,13-1,19 g/ml	ND	1,16-1,28 g/ml
Appearance by electron microscopy*	Cup shape	Irregular shape and electron dense	Heterogeneous
Sedimentation	100.000xg	10.000xg	1.200xg, 10.000xg or 100.000xg
Lipid composition	Enriched in cholesterol, sphingomyelin and ceramide; contain lipid rafts; expose phosphatidylserine	Expose phosphatidylserine	ND
Main protein markers	Tetraspanins (CD63, CD9, CD81), Alix and TSG101	Integrins, selectins and CD40 ligand, ARF6	Histones, Annexine V, Caspase 3
Intracellular origin	Endosomes	Plasma membrane	ND



**Figure 10.** Schematic representation of subtypes of extracellular vesicles (EVs) released by a cell. Three subtypes of EVs, namely exosomes, shedding microvesicles or ectosomes and apoptotic bodies, are known to be secreted by a cell into the extracellular space. Exosomes are released by exocytosis. Whereas shedding microvesicles or ectosomes are secreted by outward budding of the plasma membrane. Apoptotic bodies are released by dying cells during the later stages of apoptosis so that cell debris can easily be eliminated by neighboring and immune system cells. MVB, multivesicular body. Image from Kalra et al.<sup>148</sup>

In line with the trend of the latest publications and the recommendations of the International Society of Extracellular Vesicles (ISEV), in this work the term “exosomes” will be used to refer to vesicles that are released from cells as a result of MVBs fusing with the plasma membrane, and “microvesicles” (MV) when referring to the larger vesicles released by budding from the plasma membrane.

Since the physical differences between exosomes and MV are relatively small and a significant overlap occurs as far as their sizes are concerned, separation of these two classes of EVs is relatively difficult<sup>148</sup>. Since fractions obtained after centrifugation at 100.000xg, enriched in exosomes, may also contain other types of vesicles, we prefer to use the term “exosomes-like vesicles” (ELVs). Note that in the publication derived from the Objective 1 of this thesis (see section “Journal publications, [Paper A]), the abbreviation used for the term “exosomes-like vesicles” was EVs, and should not be confused with the general term “extracellular vesicles”.

## 4.3 Exosomes

### 4.3.1 Isolation methods

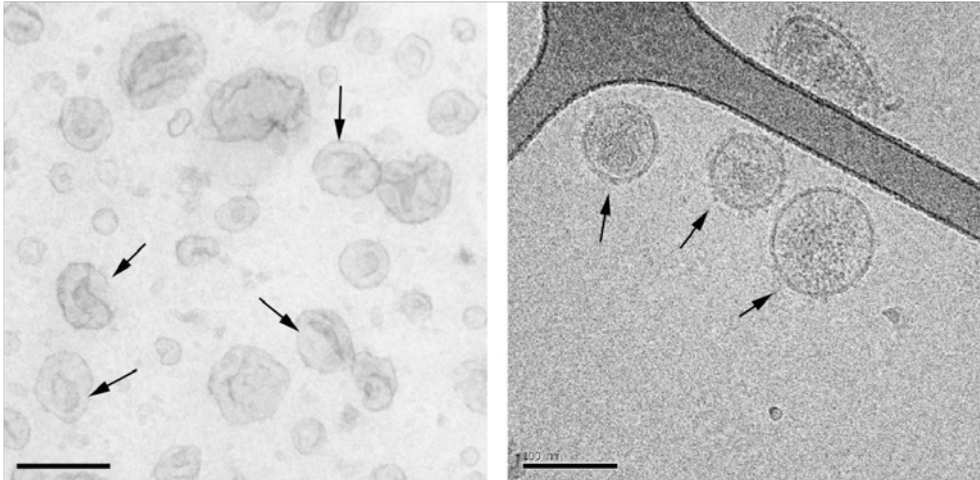
The most widely used methods for isolating exosomes are centrifugation-based procedures. Differential centrifugation involves a series of centrifugations, which successively increase in speed and time, and thus sequentially pellet smaller particles. At each centrifugation run, the pellet is discarded and the supernatant is used for the subsequent faster and longer centrifugation step until the last run (usually at 100.000-120.000xg), whose purpose is to pellet exosomes<sup>149–152</sup>. However, the major disadvantage of using a series of differential centrifugation steps coupled with ultracentrifugation is its inefficiency in separating EV subtypes<sup>146</sup>, though centrifugation at 10.000xg sediments large EVs such as MVs and apoptotic bodies. To avoid co-isolation of different EV subtypes, improvements proposed for this method include a filtration step (0,1 - 0,2  $\mu\text{m}$  pore size). To increase purity of the isolated exosomes, ultracentrifugation can be used in conjunction with density gradient centrifugation methods (i.e. sucrose, sucrose-D<sub>2</sub>O and iodixanol, commercially known as OptiPrep®)<sup>151</sup>, which separates exosomes according to their buoyant density. Whilst optimal exosome isolation can be achieved through density gradient centrifugation, the technique requires more sample, is tedious and time consuming. In an effort to make exosomes isolation feasible for a clinical application, several alternative methods have been suggested, such as immunoaffinity capture<sup>151</sup>, polymer-mediated precipitation<sup>153</sup>



and filtration-based protocols<sup>154</sup>. Despite being faster and easier than ultracentrifugation, these methods have some limitations related to the fact that exosomes are usually co-purified with high amounts of protein complexes, thus decreasing the purity of the sample obtained. Further, one of the inherent problems with immunocapture techniques is that the negative population (i.e. CD9 negative when CD9 immune beads are used) is often ignored<sup>146</sup>. As interest in the physiological and the pathological role of exosomes grows, many commercial kits that allow “easy and quick isolation procedures” are now routinely developed and are available for use. While most of these kits isolate/precipitate exosomes, the kits invariably suffer from co-isolation of other EVs and protein complexes. Hence, caution while interpreting data obtained from precipitation kits must be taken<sup>155</sup>. Currently, there is no gold standard method for isolating exosomes and, therefore, the method of choice should be determined based on the pathological or biological question of interest.

### **4.3.2 Physical properties**

Due to the influence of several variables such as the isolation method and the detection limit of the detection technique applied, a variety of discrepant size ranges have been reported to describe exosomes<sup>148</sup>. Most of the studies have worked on 50-100 nm vesicles but others have reported smaller (30 nm) or bigger (150 or 200 nm) exosomes-like vesicles<sup>156–158</sup>. Transmission electron microscopy (TEM) has been the most widely used technique to measure exosome preparations, but the fixation step needed to analyze the sample represents a drawback as vesicles tend to shrink during this process and are thus under-sized. In the recent years, new measuring technologies such as nanoparticle tracking analysis (NTA) have appeared, allowing the analysis of particles in suspension and revealed that the actual size of exosomes might be slightly larger than previously thought<sup>159</sup>.



**Figure 11.** Ultrastructure of exosomes. (Left) Exosomes isolated from melanoma cells were contrasted with uranyl-acetate and embedded as whole mount preparations in methylcellulose. Note their artificial cup-shape appearance (examples are indicated with arrows). (Right) Exosomes from prostate epithelial cells were directly frozen and observed by cryo-EM without chemical fixation or contrasting. Exosomes appear round and are visualized with improved resolution (arrows). The elongated structure (top right of the micrograph) is the Formvar film on the EM grid. Bars, 100 nm. Image from Raposo et al.<sup>136</sup>.

In line with the aforementioned technique drawback, the morphology of exosomes has traditionally been described as “cup-shaped” after fixation, adhesion, negative staining and visualization by TEM. However, advanced techniques like cryo-electron microscopy (cryo-EM) permitted the observation that the cup-shaped morphology was really an artifact related to fixation for TEM, and demonstrated that their actual shape is rounded<sup>136,160,161</sup> (Figure 11).

Exosomes float at a density between 1,13 and 1,19 g/mL, and can be pelleted by ultracentrifugation at 100.000xg or higher<sup>162</sup>.

It is worth noting that the comparison of these properties between the different types of vesicles described in Table 7 illustrates the difficulty to discriminate between them, as their sizes, densities and sedimentation speeds often overlap.

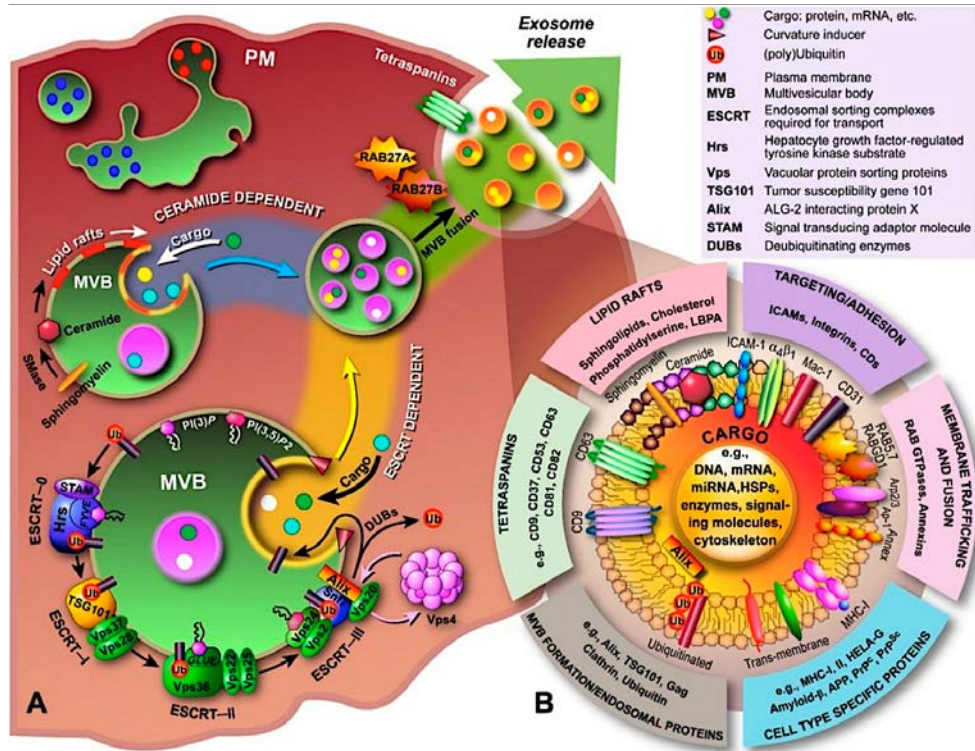
### 4.3.3 Biogenesis

Exosomes are formed within the endosomal network, a membranous compartment that sorts the different intraluminal vesicles and directs them to their appropriate destinations, including lysosomes and cell surface membranes<sup>163</sup>. When early endosomes matures to a late endosome by acidification, intraluminal vesicles

(ILVs) are formed by reverse budding from the cytoplasm into the lumen of the endosome. During this budding process, mRNA, miRNA, DNA, proteins and lipids are packed inside the ILVs. After budding of ILVs, the late endosome is called multivesicular body (MVB), which can either traffic to lysosomes for degradation (degradative MVBs) or, alternatively, to the plasma membrane (exocytic MVBs). When MVBs fuse with the plasma membrane, they release their content (the ILVs) into the extracellular space, and it is then when ILVs are referred to as “exosomes”<sup>148</sup>. ILVs (and thus exosomes) can be generated at the endosomal limiting membrane by at least two mechanisms: ESCRT-dependent and ESCRT-independent (Figure 12).

- *ESCRT-dependent mechanism*

The process of ILV formation starts when the endosomal membrane is reorganized into specialized tetraspanins-enriched microdomains, with the involvement of CD9 and CD63, that function to cluster the ILV formation machinery<sup>164</sup>. The ESCRT machinery consists of four cytosolic protein complexes (ESCRT 0, I, II and III) that are recruited to endosomes by membrane proteins that have been tagged, usually with ubiquitin, on their cytosolic domains. The ubiquitin tag is recognized by ESCRT-0 in the presence of abundant PI3P. Then, ESCRT-I (TSG101 and Vps28) is recruited to the endosomal membrane and forms an ESCRT-0/ESCRT-I complex. Next, segregation of ubiquitinated proteins into microdomains occurs and mobilization of ESCRT-II (Vps22) to the membrane. ESCRT-I and ESCRT-II then initiate reverse budding of nascent ILVs within MVBs and uptake of cytosolic cargo (i.e. RNAs and proteins). Recruitment of ESCRT-III subunits (Alix and Vps2) by ESCRT-II and oligomerization of ESCRT-III subunits inside the neck of the nascent ILVs results in closing of the cargo-containing vesicle and pinching off of the vesicles<sup>148</sup>. Overall, the components of ESCRT-0, I and II are responsible for sequestering ubiquitinated proteins at the endosomal membrane, whereas ESCRT-III contributes towards vesicle closure and detachment of ILVs from the membrane<sup>165–168</sup>.



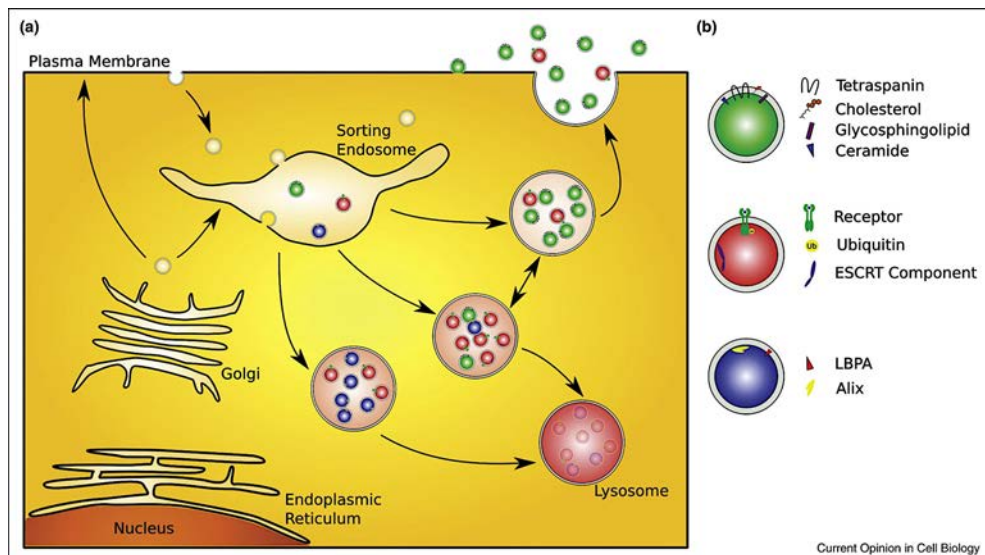
**Figure 12.** Biogenesis, secretion and composition of exosomes. (A) The biogenesis and secretion of exosomes is believed to be mediated via ceramide and/or ESCRT-dependent pathway. The ceramide-dependent pathway is based on the formation of lipid rafts in which sphingomyelin is converted to ceramide by sphingomyelinases. These ceramide-enriched domains have structural imbalances between monoleaflets causing the membrane to bend inward. In the ESCRT-dependent pathway, components of the ESCRT machinery are sequentially recruited to the endosomal membrane, which starts with Hrs, and bind to phosphatidylinositol-3-phosphate (PI(3)P) and the 3,5-bisphosphate (PI(3,5)P<sub>2</sub>) through lipid binding domains (i.e. FYVE, GLUE), and to the ubiquitinated protein (ESCRT-0). ESCRT-I and -II drive budding of ILVs, during which cargo is transported into the lumen, and ESCRT-III is recruited by Alix to complete budding and drive vesicle scission (spiral formation and pulling). DUB deubiquitinates the protein and Vps4 recycles the ESCRT machinery. The now formed MVB is transported to the plasma membrane and through fusion, the ILVs are released into the extracellular environment and are now called "exosomes". (B) Exosomal luminal cargo predominantly consists of mRNA, miRNA and gDNA fragments, and a myriad of different proteins depending on the cell of origin. Generally, proteins involved in MVB formation, tetraspanins, membrane transport and fusion, transmembrane proteins, cytoskeletal components and proteins of cytosolic origin are part of exosomes. Image from Kalra et al.<sup>148</sup>

#### ▪ ESCRT-independent mechanisms

Interestingly, ESCRT-independent ILV formation has also been described. One study performed simultaneous depletion of key-units of all four ESCRT-complexes (Hrs, Tsg101, Vps22 and Vps24) in mammalian cells and found that did not impair

ILVs formation in MVBs, although these ILVs appeared more heterogeneous<sup>169</sup>. Another study conducted in melanocytes proposed the requirement of CD63 in the ESCRT-independent mechanisms and showed that the involvement of the different sorting complexes has important implications for the distinct fates of ILVs (i.e. secretion or degradation)<sup>170</sup>. Further, a more recent study showed that EGF stimulation promotes the formation of large ESCRT-dependent ILVs, whereas depletion of the ESCRT-0 component, Hrs, promotes the formation of a uniformly sized population of small ILVs, the formation of which requires CD63; Up-regulation of CD63-dependent ILV formation by Hrs depletion indicates that Hrs and CD63 regulate competing machineries required for the generation of distinct ILV subpopulations.

The ESCRT-independent pathways of exosome's biogenesis seem to be driven by the presence of certain lipids. One of these lipid-based mechanisms is the LBPA, which is enriched in ILVs and has the capacity to drive budding of membranes into acidic liposomes just by the pH gradient across the membrane<sup>171-173</sup>. A different lipid-based mechanism implicates the sphingolipid ceramide, which in high concentrations appear to help MVB contents escape lysosomal digestion in favor of release as exosomes<sup>174</sup>. Cholesterol is another lipid implicated in MVB/ILV biogenesis. In one study, large and cholesterol-rich endosomes accumulated as a result of blocking the function of the ESCRT-III interacting protein Vsp4<sup>175,176</sup>. In other studies, knocking down the ESCRT-0 complex member, Hrs/Vps27, produced endosomal accumulation of LDL-derived cholesterol, while knock-down of other ESCRT-0, -I, -II and -III members did not produce any effect on intracellular cholesterol distribution<sup>177</sup>. Also, exosome's surface is enriched in lipid-raft microdomains, specially sphingolipids, that might be involved in the initiation of vesicles formation, as lipid-rafts are weak points, prone to outward bending<sup>174,178</sup>. Taken all together, ILVs (exosomes) size is influenced by their cargo and mechanism of formation and suggest a competitive relationship between ESCRT-dependent and -independent mechanisms of ILV formation within single MVBs<sup>179</sup>. Lipid-driven mechanisms are shared between both mechanisms. The fact that ESCRT-dependent and independent pathways exists should not be interpreted as contradictory but as a possible explanation to the presence of heterogeneous populations of MVBs and exosomes (Figure 13)<sup>137</sup>.



**Figure 13.** Model for sorting of cargo into different MVB subpopulations. (A) Different hypothetical MVB subclasses with distinct populations of ILVs (red, green and blue) are shown. The putative composition of these ILVs is shown in the right panel. Whether the MVBs contain a mixture of different ILVs as depicted in the figure is unknown. (B) At least three different subtypes of ILVs may coexist. The molecules shown represent a selection of proteins and lipids that define different subclasses of ILVs. Image from Simons et al.<sup>137</sup>

#### 4.3.4 Composition

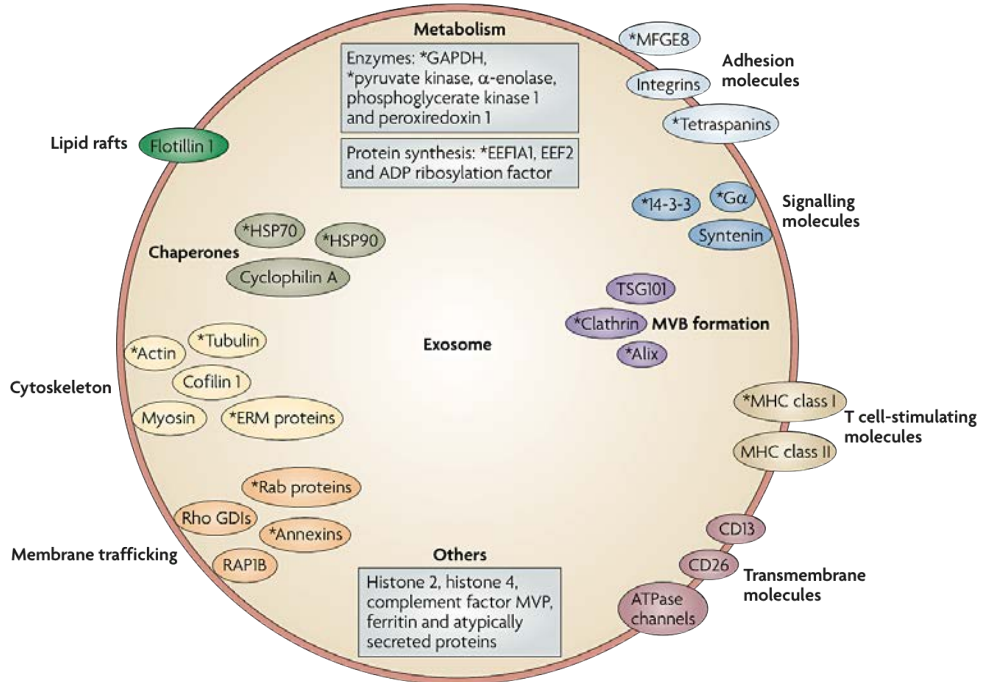
Exosomes are phospholipid bilayer-enclosed vesicles that contain proteins, mRNAs and miRNAs. Some of the contents are related to the exosomes biogenesis and therefore common to all exosomes populations, while other are specific of the originating cell. To date, 286 studies have assembled their results in a public database called “ExoCarta”, and hosts 41.860 protein, more than 7540 RNA and 116 lipid molecules (<http://www.exocarta.org>)<sup>180</sup>. For the purpose of this thesis, we will focus on the protein content.

Since their discovery, a better description of the protein composition of exosomes has been one of the main goals of researchers in the field and a wide range of techniques has been used: immunoblotting, FACS, immuno-EM, and more recently, the use of high-throughput proteomic approaches. A large number of studies led to the conclusion that exosomes contain a specific set of proteins, some of which depend on the cell of origin and some others are present regardless of the cell type<sup>143</sup> (Figure 14), including proteins involved in membrane trafficking (Rab

GTPases, annexins), MVE biogenesis (clathrin, Alix, TSG101), chaperones (HSC70, HSP90), adhesion molecules (Mfge8, integrins and tetraspanins such as CD63, CD81, CD9, CD82), cytoskeleton molecules (actin, myosin, tubulin), and lipid-rafts associated proteins (FLOT1), among others. But, besides the assortment of common markers, exosomes also bear a spectrum of proteins specific of the cell type that secretes them<sup>181-183</sup>. Worth mentioning, vesicles secreted by all cell types carry proteins from endosomes, the plasma membrane and the cytosol, but lack proteins from the mitochondria, the nucleus, the endoplasmic reticulum and the Golgi, highlighting that the formation of exosomes is a specific process, rather than a randomly packaging process of cell fragments.

The relevance of protein sorting into exosomes is still poorly understood. Most of the studies are interested in the role of exosomes once they are secreted and describe how they affect their target cells. Nevertheless, it is reasonable to think that protein sorting into exosomes is equally relevant for both the secreting and target cell. Presently, mainly two exosomal-sorting mechanisms have been suggested: post-translational modifications of proteins and trafficking via microdomains. Ubiquitination is the most widely known post-translational modification involved in exosomal sorting, where mono- or short ubiquitin chains can be recognized by the ESCRT machinery, driving incorporation into ILVs<sup>184</sup>. But, interestingly, most of the ubiquitinated proteins found in exosomes are not membrane-integrated<sup>185</sup> and, thus, membrane proteins require a different mechanism for ILV incorporation. Lipid and protein-based microdomains have been suggested to target proteins into exosomes, specially the TSPAN protein microdomains<sup>186,187</sup>.





**Figure 14.** This schematic showing the composition of a typical exosome is based on data from 15 proteomic analyses carried out on exosomes purified from cultured cells and from biological fluids. Proteins found in at least 30% of different exosomes are listed and proteins present in at least 50% of exosomes are indicated by an asterisk. EEF, eukaryotic translation elongation factor; ERM, ezrin, radixin and moesin; GAPDH, glyceraldehyde 3-phosphate dehydrogenase-activating protein; HSP, heat shock protein; MFGE8, milk fat globule EGF factor 8 protein; MVB, multivesicular body; MVP, major vault protein; RAP1B, RAS related protein 1B; Rho GDI, Rho GDP dissociation inhibitor; TSG101, tumour susceptibility gene 101. Image from Théry et al.<sup>143</sup>

## 4.4 Release and internalization

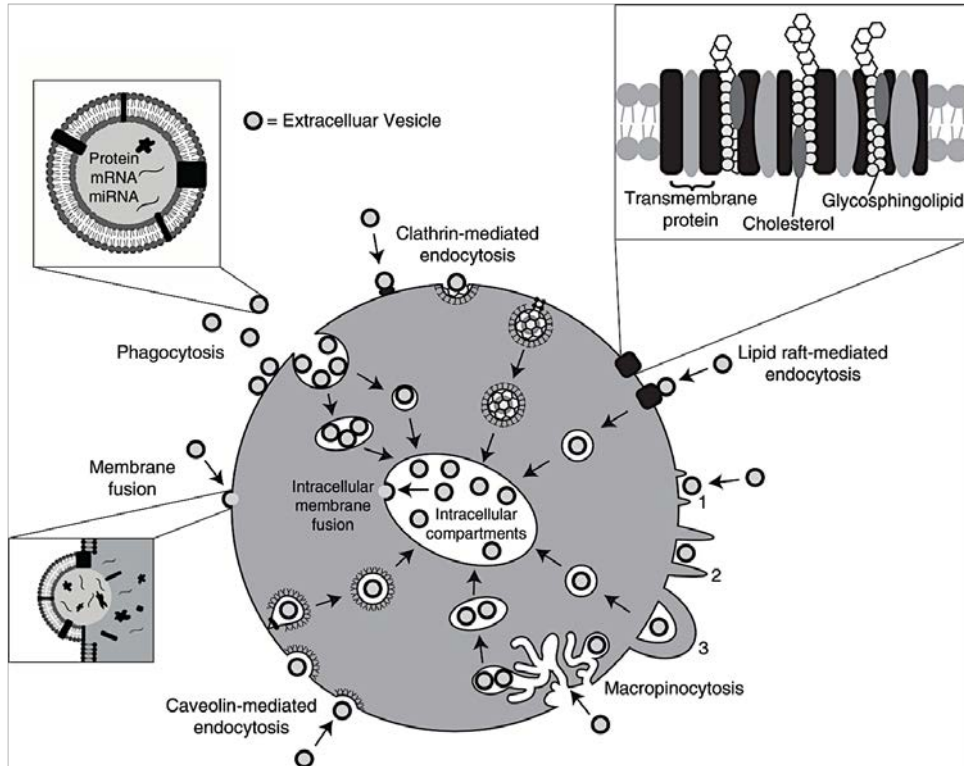
### 4.4.1 Mechanistic requirements for exosomes release

Formation of MVB and protein sorting into ILVs, transport of MVB to the plasma membrane and, finally, fusion of the luminal and the plasma membranes are required events prior to exosomes release. Similar to the process of transporting endosomes from the plasma membrane to the nucleus back and forth along microtubules, several Rab-GTPases (Rab11, Rab27a/b, Rab35) have been implicated in MVB movement<sup>188–190</sup>. SNARE molecules are key components of the protein machinery that mediates membrane fusion and, up to date, only two SNAREs have been described in exosomes release (SNAP23 and Syntaxin1a)<sup>191,192</sup>.



#### 4.4.2 Interaction of exosomes with recipient cells

Exosomes secreted by one cell must interact with a recipient cell to induce changes in its physiology but how exosomes penetrate into cells is still under debate. As depicted in Figure 15, the main routes of exosomes uptake are: clathrin-mediated endocytosis, lipid raft-mediated endocytosis, macro-pinocytosis, caveolin-dependent endocytosis, membrane fusion and phagocytosis<sup>193</sup>. From the mechanistic point of view, these can be grouped in 3 categories: (i) direct contact between surface molecules of exosomes and cells, (ii) endocytosis of exosomes, and (iii) fusion between the cell and the exosomal membranes<sup>143</sup>. In the case of direct contact, integrins and other adhesion molecules on the surface of exosomes and on the plasma membrane, or specific ligand and receptors are involved<sup>194</sup>. In some cases, this first step might be enough to induce changes in the physiological state of the recipient cell but, in other cases, exosomes need to be internalized to accomplish their function. This internalization can be achieved by receptor-mediated endocytosis of individual exosomes, or by phagocytosis of exosomes aggregates. Nevertheless, the precise mechanism of binding and internalization of exosomes remains unclear and is a currently area of extensive study<sup>193</sup>.



**Figure 15.** Exosomes uptake. Pathways shown to participate in exosomes uptake by target cells. Image from Mulcahy et al<sup>193</sup>.

## 4.5 Functions of exosomes

Due to the fact that excretion via EVs probably requires a significant amount of energy, it has been proposed that EVs are preserved through evolution.

Nowadays, exosomes are widely known to be mediators of intercellular communications and part of the intercellular milieu. Under healthy conditions, exosomes contribute to the regulation and maintenance of physiological conditions, acting as immune-modulators or participating in processes such as programmed cell death, cell-cell signaling, protein trafficking, angiogenesis, inflammation and coagulation<sup>195,196</sup>. Moreover, in developmental biology, exosomes have been described as morphogen transporters, playing an essential role in tissue patterning<sup>145,197</sup>.

Researches on exosomes have considerably increased over the past decade, not only focus on the understanding of the exosome's function in physiological conditions, but also in pathogenic states, such as cancer. Exosomes have been reported to be involved in all stages in cancer development: (i) tumorigenic transformation, (ii) tumor growth, (iii) angiogenesis, (iv) modulation of immune responses, and (v) induction of mechanisms to acquire therapy resistance<sup>198–201</sup>. As summarized in Table 8, it has been recently reviewed the involvement of tumor-derived exosomes (TDE) in tumorigenesis according to the hallmarks of cancer, which were compiled by Hanahan and Weinberg over a decade ago<sup>202</sup>.

**Table 8.** Summary of tumor-derived exosomes involvement in the hallmarks of cancer. Adapted from Meehan 2015<sup>301</sup>.

Hallmark	Role of TDE	Relevant TDE cargo
Sustaining proliferative signaling	Capable of transferring the proliferative phenotype of donor cells to recipient cells	EGFRvIII, KRAS, phosphorylated proteins, DeltaNP73 mRNA, Let-7 miRNA and other miRNA
Evading growth suppressors	Poorly defined	H-ras and N-ras transcripts or Rab proteins, miRNA and PTEN
Resisting cell death	Direct and indirect roles in resistance to cell death	Bcl2, Bax, survivin, miR-21 and mutant TP53
Enabling replicative immortality	Not established	-
Inducing angiogenesis	Confer proangiogenic properties via cell-cell communication and also by mediating cellular contents	VEGF, FGF, angiopoietin 1, ephrin A3, MMP 2, MMP9, miR-23b, D114 and EGFR
Activating invasion and metastasis	Mediate the metastatic process by promoting invasion and supporting tumor growth	Amphiregulin, EGFR, MET, miR-200, miR122, vimentin, MMPs, integrins and LMP-1
Reprogramming of energy metabolism	Poorly defined	Glycolysis enzymes and lactate dehydrogenase
Evading immune destruction	Immune system modulation	TNF, Fas ligand, TRAIL, MHC class I and II, galactin-9, CD63, CD73, CD39, TGFB1, NKG2D, hsp72, PGE2, TGF- $\beta$ and MyD88
Genomic instability and mutations	No evidence to suggest that TDE are able to induce sustained genomic instability or mutations	-
Tumor promoting inflammation	May enable immune cells to generate an inflammatory response	KIT, miR-21 and miR-29a

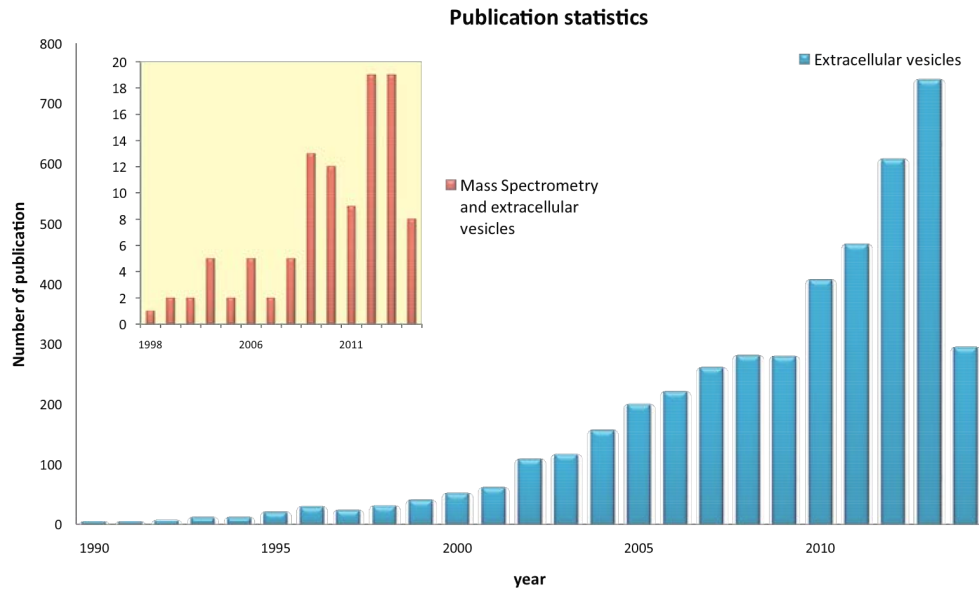
## 4.6 Advantages of exosomes in protein biomarker research

Tumor cells have been reported to secrete increased amounts of exosomes<sup>203</sup>. Since these tumor-derived exosomes carry the tumor-specific genomic and proteomic signatures, they are currently considered ideal and unique targets for cancer detection<sup>148</sup>. It is well established that the earlier the cancer is diagnosed, the better the survival rate but, to date, most detection methods are image-based and have the limitation of detecting small early-stage tumors. Furthermore, other

diagnostic procedures, such as biopsies, are invasive and require an already visible tumor for sampling. Also, tumor biopsies do not reflect tumor heterogeneity. Hence, a main goal in oncological research is the establishment of reliable and non-invasive diagnostic methods. In this regard, the presence of exosomes in body fluids will help moving towards working with liquid biopsies<sup>204,205</sup>. Few high-abundance proteins make up to 97% of body fluids, hindering the detection of the low-abundance proteins, which are generally the most promising candidates for biomarker discovery. Exosomes analysis reduces the complexity of bodily fluids thereby aiding in the detection of low abundance proteins and solving the dynamic-range problem. Also, exosomes contain RNA, lipids, proteins and metabolites that are reflective of the cell type of origin and may contain disease causing proteins (i.e. mutant proteins). The exosomal content represents the current disease state, because they are released by living cells, differently from cell-free DNA, that can be released by apoptotic cells. Moreover, the membranous structure of exosomes grants that their luminal contents are protected from degradation by extracellular proteases. Importantly, exosomes are highly stable in storage conditions for extended periods of time, maintaining the quality of their bio-active material<sup>206</sup>.

Proteins are the best-characterized compounds of exosomes so far. In the recent years, the number of papers that use mass spectrometry-based analysis has grown fast, mainly thanks to the rapid methodology advances in the field (Figure 16). Most of the current proteomic research is conducted in the “bottom-up” fashion (analysis of enzymatically digested proteomes), and thousands of proteins can be routinely identified in a matter of hours using the most advanced mass spectrometry and liquid chromatography methods. Contemporary high-resolution and high-sensitivity mass spectrometry, high-resolution liquid chromatography, and data searching algorithms allow for the in-depth proteomic profiling and identification of post-translational modifications (PTM). It is also possible to quantify changes in protein expressions within exosomes using stable isotopic labeling (SILAC, TMT, and iTRAQ) or in a label-free manner (peptide spectral match counts, intensity, and peak area integration). The major obstacles hindering the detailed proteomic analysis of exosomes are the difficulties in purifying true exosomal proteins from high dynamic-range source fluids in sufficient quantities for LC-MS analysis. The application of cutting-edge LC-MS techniques to the

proteomic analysis of exosomes could eventually allow for the detailed profiling of EV proteomes. However, the clinical application of such biomarkers would require a robust and simple EV isolation<sup>207–209</sup>.



**Figure 16.** Publication statistics on extracellular vesicles showing the number of publications which used mass spectrometry methods. (topic search was performed using terms: “exosome” or “microvesicle” or “platelet” or “apoptotic body” or “extracellular vesicle” in Web of Science database (execution date October 2014). References found in the search were exported into EndNote and re-searched with the term “Mass Spectrometry” (see insert). Image from Pocsfalvi et al.<sup>207</sup>.

# OBJECTIVES

---



## Background

Endometrial cancer (EC) is the most common cancer of the female genital tract and the fourth most frequent cancer in women in the United States. Discouraging, EC deaths are increasing rapidly and rates grew about a 2% per year from 2010 to 2014. Even though early-stage diagnosed EC restricted to the uterus represent 67% of the cases with a 5-year survival rate of 95%, it still remains around a 30% of patients diagnosed at advanced stages of the disease, when the 5-year survival rate decreases dramatically to 69% and 17% in case of regional and distant spreading, respectively. These statistics highlight the urgent need to develop improved methods for early EC detection.

The current diagnostic process is laborious and includes a pelvic examination and a transvaginal ultrasonography followed by a confirmatory histopathological evaluation of an endometrial biopsy, which can either be obtained by aspiration or by hysteroscopy. Although the aspiration method is the less invasive option, it lamentably presents an insufficient sampling rate of up to 22% and fails to provide a diagnosis in 12% of cases. Moreover, the accuracy to determine the histological subtype of the tumor is limited, with up to 20% of discrepancy. In those cases, a biopsy guided by hysteroscopy is required, although this technique increases the risk of uterine perforation and hemorrhage, among other complications, as well as derived sanitary costs.

To improve EC diagnosis, the use of uterine aspirates meet some features that makes this biofluid a promising source of biomarkers for screening, diagnosis and monitoring of pathologies related to the female genital tract: (i) sampling of uterine aspirates is minimally invasive and it is feasible to retrieve between 50 and 1000  $\mu$ l of sample from a patient, which is sufficient to extract material for molecular studies; (ii) uterine aspirates are in direct contact with the uterine cavity, representing faithfully its complex environment. Uterine aspirate's composition derives from secretions and cells flaking from a variety of surrounding tissues, from the luminal epithelium and glands, from proteins selectively transudated from blood, and likely contributors from tubal fluid; and (iii) uterine aspirates exhibit



molecular alterations present in EC and are useful to study the EC tumor heterogeneity.

Nowadays, molecular screening technologies allow the consecution of high-throughput discovery and validation studies on uterine aspirates for the development of non-invasive tests. Although research in this field appears to be promising for transcriptomic and genomic approaches, untargeted proteomic-based studies present serious limitations. Uterine aspirates highly express abundant proteins coming from plasma, such as albumin and gamma globulins, along with hemoglobin; and those mask the detection and analysis of less abundant proteins, which are generally the most promising candidates for biomarkers discovery. To solve this dynamic-range problem, but also to expand research in the field of biomarkers discovery for gynecological pathologies, exosome-like vesicles (ELVs) arise as a promising source of EC biomarkers.

## Working hypothesis

Under the assumption that exosome-like vesicles exist in uterine aspirates and that, in the context of EC, a high proportion of these vesicles might be released by tumor cells, we hypothesize that the proteomic study on exosome-like vesicles from uterine aspirates of women under the suspicion of suffering from endometrial cancer will provide a set of biomarkers useful to diagnose and differentiate between histological subtypes of endometrial cancer. We surmise that improved mass spectrometric technics applied to the study of the exosomal fraction will present advantages as compared to the study of the whole uterine aspirate sample.

## General objective

The main objective of this work is to improve minimally-invasive diagnosis of EC by identifying a proteomic signature in exosomes derived from uterine aspirates. This might help to improve the detection and management of EC in the near future, reducing the number of invasive biopsies (i.e. performed by hysteroscopy) and its associated complications and health care costs.

To do that, we have defined the tree following 3 specific subobjectives:

## **Specific objectives**

### **1. Establishment of a suitable protocol for the use of ELVs from uterine aspirates**

#### **1a. Determining whether ELVs exist in the fluid fraction of uterine aspirates and comparing three ultracentrifugation-based protocols for their isolation**

Despite new techniques are emerging that might eventually take the relay, the most widely used methods to isolate exosomes from biological fluids are still centrifugation-based but, to date, no references are found in the literature to standardize methods to isolate them from uterine aspirates. Thus, the first objective of this project is to compare three different ultracentrifugation-based protocols (“standard”, “filtration” and “sucrose”) to isolate ELVs from UAs. Parameters such as vesicles purity, reproducibility of the method and time/cost effectiveness will be evaluated.

#### **1b. Characterizing ELVs from uterine aspirates from a structural perspective**

The isolated ELVs will be visualized, measured and counted to verify enrichment in exosomes by using electron microscopy and Nanoparticle Tracking Analysis.

#### **1c. Characterizing ELVs from uterine aspirates at molecular level**

Enrichment in exosomes will also be assessed by determining the presence of typical exosome markers by immunoblotting.

### **2. Identification of new protein biomarkers for EC in ELVs from uterine aspirates.**

#### **2a. Proteomic analysis of ELVs derived from 5 EC cell lines for the creation of a Super-SILAC mix**

Clinical evaluation of protein biomarkers requires an accurate quantification. In this regard, the combination of uterine aspirates ELV's proteome with known amounts of heavy standards (super-SILAC mix) will enable robust relative quantification of

putative candidates. For this purpose, we will study the proteome of ELVs derived from 5 EC cell lines to select the combination of those that better represents the proteome of uterine aspirates ELVs and that will compose the Super-SILAC mix (internal standard).

### **2b. Optimizing methods for protein extraction and digestion**

ELVs from uterine aspirates yield an extremely low amount of protein so it is necessary to carefully set up the conditions for the proteomic analysis. In here, in-gel and in-solution digestion methods will be compared.

### **2c. Identifying diagnostic biomarkers in ELVs derived from uterine aspirates by a SILAC-based LC-MS/MS approach (Discovery Phase)**

In this phase, we will compare the proteomic profile of ELVs isolated from uterine aspirates of EC patients to those isolated from control patients in order to identify EC diagnostic biomarkers; and the proteomic profile of ELVs isolated from EEC patients compared to the ones isolated of NEEC patients to identify biomarkers to differentially diagnose those histological subtypes. GO Analysis of the label-free data will be conducted to reveal overrepresented biological processes.

## **3. Verification of candidates by targeted proteomics**

### **3a. Verification of protein abundance changes in UAs ELVs by SRM**

The most promising candidate biomarkers found in Objective 2 will be verified by a targeted proteomic approach, the LC-MS/MS operated in Selected Reaction Monitoring (SRM), in a bigger and independent cohort of patients.

### **3b. Building and evaluating predictive models**

The classification ability of each candidate will be evaluated. Sensitivity and specificity will be determined in a panel of protein biomarkers to assess its performance in clinical practice.

### **3c. Validation of candidates in UAs**

The most promising candidate biomarkers found in Objective 2 will also be evaluated by targeted proteomics (SRM) in the whole fluid fraction of uterine aspirates of an independent cohort of patients in order to assess the potential of

ELVs biomarkers in the whole fluid and consequently, evaluate the need of ELVs isolation in a future clinical application.



# **MATERIALS AND METHODS**

---



# 1. Uterine aspirates

## 1.1 Collection and processing of uterine aspirates

Patients were recruited at three different institutions: HUVH (Hospital Universitari Vall d'Hebron, Barcelona, Spain), HUAV (Hospital Universitari Arnau de Vilanova, Lleida, Spain) and UMCF (University Medical Center of Freiburg, Freiburg, Germany). Each participating institution obtained ethical approval and samples were obtained after the participants signed the informed consent.

UAs were obtained by aspiration with a Cornier Pipelle (Gynetics Medical Products). Samples were placed in 1.5 mL tubes and kept on ice through all the processing which included addition of phosphate buffered saline (PBS) in a 1:1 ratio (v/v), gently pipetting of the sample, and centrifugation at 2500g (4 °C) in a F45-30-11 rotor (Eppendorf Microcentrifuge 5417R) for 20 min to remove the cellular fraction. The remaining supernatant (SN) fraction, was then aliquoted and frozen at -80 °C until needed.

## 1.2 Selection of patients and inclusion/exclusion criteria

### 1.2.1 For the establishment of the ELVs isolation method

A total of 33 pre-menopausal patients with benign gynecological diseases or healthy donors were recruited at the Department of Gynecological Oncology of the HUVH. A description of the clinic-pathological features of all participating patients for this objective is detailed in Table 9. An inclusion criterion was pre-menopause. Women who had been treated previously for gynecological pelvic cancer, as well as patients positive for the human immunodeficiency virus and/or the hepatitis virus were excluded.

To compare isolation protocols in Objective 1, 27 UAs SNs (samples 1-27, Table 9) were pooled together and divided into 20 aliquots containing 445 µL (Figure 17).



**Table 9.** Clinical and pathological features of patients used for Objective 1. Age, diagnosis and starting volume of uterine aspirates' fluid fraction are detailed. Samples 1-27 were pooled to compare ELVs isolation protocols; samples 28-33 were individually used for ELVs characterization by NTA and immunoblot.

Patient	Age	Diagnosis	UAs SN volume ( $\mu\text{L}$ )	Experiment
1	22	Mucinous cystadenoma	600	Pool
2	30	Endometrioma	760	Pool
3	24	Follicular cyst	550	Pool
4	36	Papillary-serous cystadenoma	1000	Pool
5	39	Endometrioma	1000	Pool
6	40	Endometrioma	650	Pool
7	23	Endometrioma	300	Pool
8	38	Mucinous cystadenoma	500	Pool
9	29	Ovarian fibroma	600	Pool
10	37	Endometrioma	450	Pool
11	37	Endometrioma + uterine leiomyoma	250	Pool
12	33	Tubo-ovarian abscess + Endometrioma	120	Pool
13	44	Serous cystadenoma	600	Pool
14	40	Endometrioma	400	Pool
15	21	Serous cystadenoma	400	Pool
16	38	Mucinous cystadenoma	600	Pool
17	31	Mucinous cystadenoma	100	Pool
18	51	Proliferative endometrium	500	Pool
19	28	Mid-secretory endometrium	500	Pool
20	47	Endometrial polyp	1000	Pool
21	45	Uterine leiomyoma	500	Pool
22	38	Uterine leiomyoma	300	Pool
23	46	Normal endocervical tissue	600	Pool
24	50	Atrophic endometrium	350	Pool
25	42	Endometrial polyp	160	Pool
26	36	Endometrial polyp	400	Pool
27	45	Endometrioma	500	Pool
28	30	Endometrioma	904	Individual
29	36	Papillary serous cystadenoma	1240	Individual
30	53	Uterine submucosal leiomyoma	434	Individual
31	28	Normal endocervical tissue	330	Individual
32	47	Endometrial polyp	678	Individual
33	51	Secretory endometrium	1000	Individual

### 1.2.1 For the pilot studies to set up the proteomic approach for the discovery phase

A total of 9 patients were recruited at the Department of Gynecological Oncology of the HUVH to set up the proteomic approach for the Discovery phase. Inclusion criteria were a minimum age of 40 years, post menopause, and AUB. Patients positive for the human immunodeficiency virus and/or the hepatitis virus were excluded. From the 9 women, 3 were diagnosed with endometrioid EC or Type I (EC1), 3 with non-endometrioid EC or Type II (EC2), and 3 were non-EC woman (CTRL). Clinico-pathological information is presented in Table 10. These samples were used to perform two pilot studies intended to define the best sample preparation and proteomic approach before moving to the discovery phase study: (i) we use the 9 samples to understand the variability and blood content of the samples, and (ii) we use 2 samples to compare the in-gel and the in-solution digestion protocols to ensure the maximum number of identifications; samples selected for this pilot study are in bold in table 10.

**Table 10.** Clinical and pathological features of patients used to set up the protocol for the proteomic approach. The bloodiness of each UA is presented based on the color.

Patient	Age	Diagnosis	Grade	FIGO	Group	UAs Appearance	UAs SN volume (μL)
1	72	Endometrioid EC	G3	IA	EC1	Clear	480
<b>2</b>	66	Endometrioid EC	G2	IB	EC1	Pinkish	630
3	66	Endometrioid EC	G2	II	EC1	Reddish	750
4	73	Carcinosarcoma	G3	IB	EC2	Clear	765
5	68	Serous EC	G3	IA	EC2	Pinkish	2600
6	89	Clear cell EC	G3	IA	EC2	Very dark	1045
7	78	Hyperplasia	-	-	CTRL	Pinkish	190
<b>8</b>	59	Adenomyosis	-	-	CTRL	Clear red	670
9	53	Myoma	-	-	CTRL	Reddish	650

### 1.2.2 For the discovery and the verification phase

A total of 30 women for the discovery phase of Objective 2 (10 EC1, 10 EC2 and 10 CTRL) and 107 for the verification phase of Objective 3 (25 EC1, 21 EC2 and 41 CTRL) were recruited. Inclusion criteria were attendance to the Department of

Gynecological Oncology for suspicion of EC at any of the participating centers, a minimum age of 40 years, post menopause, and AUB. Exclusion criteria were previous gynecological pelvic cancer, and infection by human immunodeficiency virus and/or the hepatitis virus. The control groups are formed by women presenting abnormal vaginal bleeding but diagnosed of a benign pathology or with a normal endometrium, and the EC2 groups only include non-endometrioid serous EC. A description of the clinical and pathological features of these patients is detailed in Table 11 for the discovery cohort and Table 12 for the validation cohort.

**Table 11.** Clinical and pathological features of patients used for Objective 2, the discovery phase. EC1, endometrioid endometrial cancer type I; EC2, serous endometrial cancer type II; CTRL, control. \* One case with undetermined myometrial infiltration. \*\* One case with undetermined lymphovascular invasion.

Discovery cohort (A)	EC1 (n=10)	EC2 (n=10)	CTRL (n=10)
<b>Age (years)</b>			
<b>Median</b>	69.5	68.5	65.5
<b>Minimum</b>	52	40	51
<b>Maximum</b>	86	85	78
<b>Collection center</b>			
<b>HUVH</b>	10	8	10
<b>HUAV</b>	-	2	-
<b>UMCF</b>	-	-	-
<b>Uterine condition</b>			
<b>Premenopausal</b>	-	-	-
<b>Postmenopausal</b>	10	10	10
<b>Histologic grade</b>			
<b>Grade 1</b>	1	-	
<b>Grade 2</b>	5	-	
<b>Grade 3</b>	4	10	
<b>FIGO stage</b>			
<b>IA</b>	3	4	
<b>IB</b>	4	1	
<b>II</b>	2	1	
<b>IIIA</b>	-	1	
<b>IIIB</b>	-	-	
<b>IIIC1</b>	-	3	
<b>IIIC2</b>	-	-	
<b>IVA</b>	1	-	
<b>IVB</b>	-	-	
<b>Myometrial invasion</b>			
<b>&lt;50%</b>	6	5	
<b>&gt;50%</b>	3*	5	
<b>Lymphovascular invasion</b>			
<b>Yes</b>	3	3	
<b>No</b>	6**	7	

**Table 12.** Clinical and pathological features of patients used for Objective 3, the verification phase. EC1, endometrioid endometrial cancer type I; EC2, serous endometrial cancer type II; CTRL, control.

Verification cohort (B)	EC1 (n=45)	EC2 (n=21)	CTRL (n=41)
<b>Age (years)</b>			
<b>Median</b>	67	74	57
<b>Minimum</b>	50	56	45
<b>Maximum</b>	88	93	80
<b>Collection center</b>			
<b>HUVH</b>	35	14	33
<b>HUAV</b>	10	7	-
<b>UMCF</b>	-	-	8
<b>Uterine condition</b>			
<b>Premenopausal</b>	-	-	-
<b>Postmenopausal</b>	45	21	41
<b>Histologic grade</b>			
<b>Grade 1</b>	11	-	
<b>Grade 2</b>	27	-	
<b>Grade 3</b>	7	21	
<b>FIGO stage</b>			
<b>IA</b>	24	7	
<b>IB</b>	21	1	
<b>II</b>	-	3	
<b>IIIA</b>	-	2	
<b>IIIB</b>	-	1	
<b>IIIC1</b>	-	-	
<b>IIIC2</b>	-	3	
<b>IVA</b>	-	1	
<b>IVB</b>	-	1	

### 1.2.2 For the validation in UAs

A total of 67 women were recruited at the Department of Gynecological Oncology of the participating centers (22 EC1, 20 EC2 and 25 CTRL). In this cohort, both pre- and post- menopausal women were included, with no age-restriction. All women were patients under the suspicion of EC by presenting AUB and/or a thickness of the endometrium higher than 4mm for postmenopausal women and 8mm for premenopausal women, based on the results of a transvaginal ultrasonography. Patients previously diagnosed with a gynecological pelvic cancer or infected with the human immunodeficiency virus and/or the hepatitis virus were discarded. Clinico-pathological features are detailed in Table 13.

**Table 13.** Clinical and pathological features of patients used for Objective 3c, the validation of candidates in UAs.

Validation in UAs cohort (C)		EC1 (n=22)	EC2 (n=20)	CTRL (n=25)
<b>Age (years)</b>				
	Median	69.5	73.5	54
	Minimum	38	51	32
	Maximum	87	93	80
<b>Collection center</b>				
	HUVH	22	12	22
	HUAV	-	8	-
	UMCF	-	-	3
<b>Uterine condition</b>				
	Premenopausal	3	1	5
	Postmenopausal	19	19	20
<b>Histologic grade</b>				
	Grade 1	1	-	
	Grade 2	16	-	
	Grade 3	5	20	
<b>FIGO stage</b>				
	IA	12	5	
	IB	6	-	
	II	3	3	
	IIIA	-	2	
	IIIB	-	1	
	IIIC1	-	-	
	IIIC2	1	6	
	IVA	-	2	
	IVB	-	1	
<b>Myometrial invasion</b>				
	<50%	14	12	
	>50%	8	8	
<b>Lymphovascular invasion</b>				
	Yes	4	11	
	No	18	9	

## 2. Cell culture

### 2.1 Cell lines

We used the commercially available EC cell lines AN3CA, KLE, RL95.2, Hec1a and Ishikawa. Each cell line was grown at 37°C in a 5% CO<sub>2</sub> humidified atmosphere in its preferred cell culture media supplemented with 10% FBS and 1% penicillin-streptomycin (Gibco, LifeTechnologies, USA). A complete description of each cell line origin and culture media is described in Table 14.

**Table 14.** Features of the EC cell lines used in this work.

Cell line	Culture media	Origin
<b>AN3CA</b>	DMEM/F-12	Poorly differentiated endometrial adenocarcinoma from a 55-year-old Caucasian woman. Derived from a metastatic lesion in the lymph node and related to the malignant disorder acanthosis nigricans <sup>1</sup> .
<b>KLE</b>	DMEM/F-12	Poorly differentiated endometrial carcinoma from a 64-year old Caucasian woman <sup>2</sup> .
<b>RL95.2</b>	DMEM/F-12	Grade 2 moderately differentiated adenosquamous carcinoma of the endometrium from a 65-year-old Caucasian woman <sup>3</sup> .
<b>Hec1a</b>	McCoy's 5A	Grade 2 endometrial adenocarcinoma from a 71-year-old woman (stage IA) <sup>4</sup> .
<b>Ishikawa</b>	DMEM/F-12	Grade 1 endometrial adenocarcinoma from a 39-year-old woman. Well-differentiated adenocarcinoma <sup>5</sup> .

### 2.2 Culture conditions for ELVs isolation

#### 2.2.1 ELVs-depleted media preparation

The previous media's ELVs depletion is crucial to ensure that there is no contamination of ELVs coming from fetal bovine serum (FBS) in the media that will be used for culturing cells. To eliminate ELVs from the FBS, 20% FBS DMEM/F-12, McCoy's 5A, or SILAC-DMEM/F-12 were prepared and ultracentrifuged for 16 hours at 100,000xg at 4°C. The pellet containing serum-derived ELVs was discarded and the SN was diluted with FBS-free fresh media to a final concentration of 10% ELVs-depleted FBS media. This media was supplemented with 1% penicillin-streptomycin and passed through a 0,22 µm sterile filter (Corning).

## 2.2.2 Cell Culture for ELVs enrichment

AN3CA, KLE, RL95.2 and Ishikawa cells were cultured in ELVs-depleted DMEM/F-12, and Hec-1a cells were cultured in ELVs-depleted McCoy's 5A. Cells were seeded based on growth rate (doubling time) to achieve 80-90% confluence after 48-72 hours. Following the two to three days' culture, the conditioned media was harvested and frozen at -80 °C for posterior ELVs isolation.

## 2.3 Culture conditions to prepare the Super-SILAC Mix

SILAC labeling was performed by culturing AN3CA, RL-95 and KLE cells in ELVs-depleted SILAC-DMEM/F-12, namely DEMEM deprived of its natural lysine and arginine, and supplemented with  $^{13}\text{C}^{15}\text{N}$ -Lysine,  $^{13}\text{C}^{15}\text{N}$ -Arginine, Proline, dialyzed ELVs-free FBS (ThermoFisher Scientific) and 1% penicillin-streptomycin (Gibco, ThermoFisher Scientific). Labeled aminoacids and proline were purchased from Silantes GmbH, Germany. Cells were cultured for more than 10 doublings in the ELVs-depleted SILAC medium to reach complete protein labeling. Heavy aminoacids incorporation was assessed by LC-MS/MS analysis (data not shown). Once the labeling was ensured, cells were cultured for 48 - 72 hours in ELVs-depleted SILAC media, and the conditioned media was harvested and frozen at -80 °C for posterior ELVs isolation. Each cell line's ELVs protein content was extracted and quantified as explained later, and mixed in a 1:1:1 ratio, to create the Super-SILAC mix, also referred to as Internal Standard (IS).

# 3. ELVs isolation

## 3.1 Establishment of the ELVs isolation method

A pool of 27 UAs was used to compare three ELVs isolation protocols based on differential centrifugation: "Standard", "Filtration" and "Sucrose". In addition to ELVs, we collected fractions corresponding to MVs and proteins from the soluble fraction to monitor differences in the enrichment in ELVs. A schematic representation of the experimental work is depicted in Figure 17.

Each protocol was tested in quadruplicates in order to have three technical replicates for subsequent NTA analysis, protein quantification and analysis of

exosomal markers by western blot; and one replicate to observe the isolated ELVs by TEM.

### **3.1.1 Standard protocol**

ELVs were obtained from the SNs of UAs by differential centrifugation, following a modification of a previously described ELVs isolation protocol by They et al.<sup>149</sup>. Briefly, SNs were thawed and diluted in PBS to a final volume of 25 mL. A centrifugation step at 10,000 $\times$ g (4 °C) for 30 min was performed on a Thermo Scientific Heraeus MultifugeX3R Centrifuge (FiberLite rotor F15-8x-50c) to remove cell debris, macroparticles and apoptotic bodies. The resulting pellet enriched in MVs was resuspended in 50  $\mu$ L of PBS and frozen at -80 °C. Then, the supernatant was transferred to ultracentrifuge tubes (Beckman Coulter) and filled with PBS to perform a first ultracentrifugation step at 100,000 $\times$ g (4 °C) for 2 h on a Thermo Scientific Sorvall WX UltraSeries Centrifuge with an AH-629 rotor. The supernatant of this second centrifugation was the soluble fraction and was frozen at -80 °C. This first pellet was resuspended in PBS and again centrifuged at 100,000 $\times$ g (4 °C) for 1 h. The final pellet enriched in ELVs (possibly along with MVs and some remaining apoptotic bodies) was resuspended in 50  $\mu$ L of PBS. Five microliters from MVs and ELVs pellets were reserved at -80 °C for particle size distribution and quantification by NTA while the rest of the sample was frozen at -80 °C for protein extraction.

### **3.1.2 Filtration protocol**

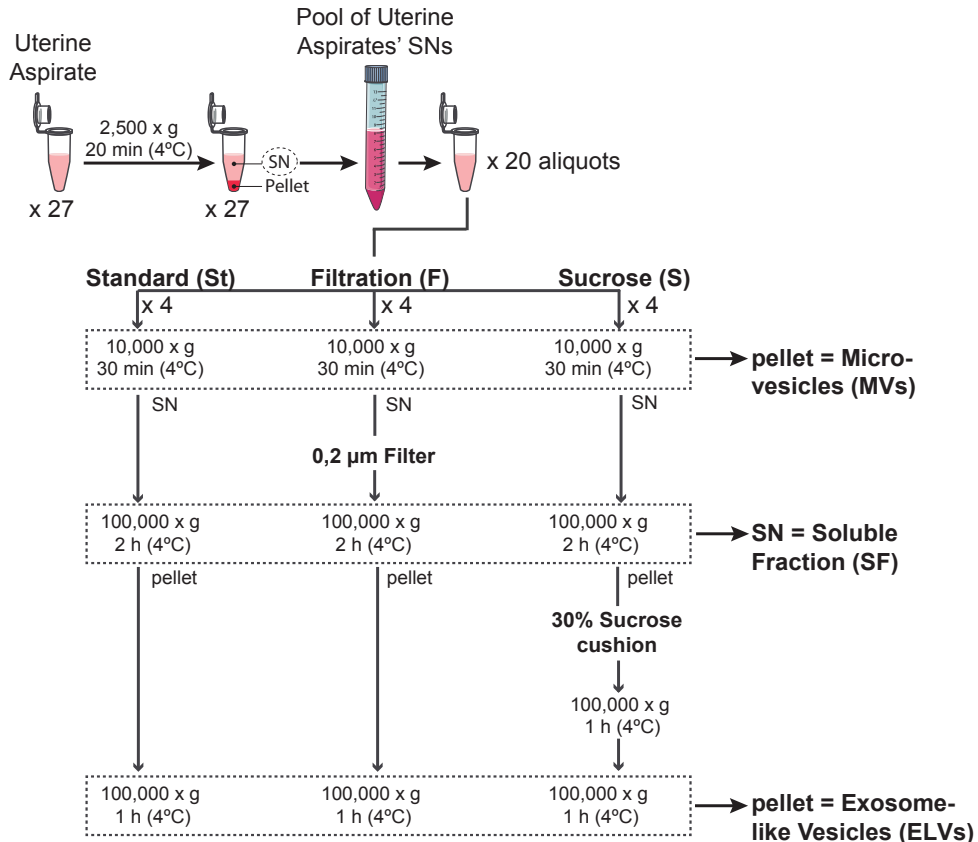
The Filtration protocol consisted in depleting the supernatant obtained after the 10,000 $\times$ g centrifugation of structures bigger than 200 nm using a sterile filter (Corning). The rest of the protocol remains the same as for the Standard.

### **3.1.3 Sucrose cushion protocol**

A 30 % sucrose cushion composed of 20 mM Tris-HCl, pH 7.4 in D<sub>2</sub>O (density from 1.13 to 1.19 g/mL) was added to the Standard protocol following the first ultracentrifugation at 100,000 $\times$ g. The sucrose cushion containing the 2 h pellet was then centrifuged for 1 h at 100,000 $\times$ g (4 °C). ELVs were recovered by poling the tube with a needle and were then washed with PBS for 1 h at 100,000 $\times$ g (4 °C).



The final pellet was resuspended in 50  $\mu\text{L}$  of PBS. Again, 5  $\mu\text{L}$  from MVs and ELVs pellets were reserved at  $-80\text{ }^{\circ}\text{C}$  for particle size distribution and quantification by NTA while the rest of the sample was frozen at  $-80\text{ }^{\circ}\text{C}$  for protein extraction.



**Figure 17.** Schematic representation of the three protocols tested to isolate ELVs from UAs.

### 3.2 Isolation of ELVs for the discovery and the verification phase

As concluded from Objective 1, we used the Standard protocol to isolate ELVs from the UAs of the patients recruited for the discovery and the verification phase.

The same procedure was applied to purify ELVs from conditioned media, with an additional initial centrifugation step at 500 x g for 5 minutes to eliminate dead cells followed by 20 minutes at 2.500 x g at 4°C to pellet any remaining big cell debris.

## 4. EVs visualization and measurement

### 4.1 Transmission electron microscopy

TEM imaging of MVs and ELVs was conducted at the Electron Microscopy Unit in the Centre Científic i Tecnològic of the University of Barcelona, settled at the Hospital Clínic.

Isolated MVs and EVs were negatively stained and analyzed per duplicate by TEM. Vesicles were fixed in 50  $\mu$ L of 4 % paraformaldehyde. Gold grids were incubated with samples for 1 min. After removing sample excess, negative staining was performed by incubation with uranyl acetate for 1 min. After washing, grids were dried overnight at room temperature. Samples were observed with a transmission electron microscope JEOL 1010 coupled to an Orius CCD camera (Gatan, Inc.), working at 80 kV with a tungsten filament.

### 4.2 Nanoparticle tracking analysis

Size distribution and quantification of isolated ELVs and MVs was determined by Nanoparticle Tracking Analysis (NTA) in collaboration with two different groups. For samples of Objective 1, NTA was conducted in collaboration with the Grup d'Enginyeria de Materials (GEMAT) of the Universitat Ramon Llull at the IQS (Institut Químic de Sarrià, Barcelona, Spain), lead by Dr. Salvador Borrós. For samples of Objectives 2 and 3, NTA was conducted in collaboration with the group of Dr. Hernando del Portillo, at the IGTP (Institut de Recerca Germans Trias i Pujol, Badalona, Spain).

Isolated EVs were analyzed by NTA using a NanoSight LM10 system (Malvern Instruments) equipped with a 405 nm laser and a Hamamatsu C11440 ORCA-Flash2.8 camera (Hamamatsu). Each sample was diluted (dilutions ranged from 1:5000 to 1:10000) with mili-Q water (Milli-Q Synthesis, Merck Millipore) prior to injection in the Nanosight instrument. Data was analyzed with the NTA software 2.3. Size and concentration of particles were determined by the following settings: camera level and detection threshold were set to maximum (15 or 16) and minimum (3–5), respectively; camera gain was set to 512; blur, minimum track length, and minimum expected size were set to “auto”. Readings for Objective 1

were taken in triplicates during 60 s at 18.87 frames/s, at room temperature ranging from 23-25°C; average measurements and standard deviation of the 3 recorded videos were calculated. Readings for Objective 2 and 3 were taken in single capture with automatic temperature monitoring.

## 5. Protein techniques

### 5.1 Protein extraction

#### 5.1.1 Protein extraction from MVs and ELVs

For Objectives 1 and 2, pellets of MVs and/or ELVs were resuspended in a lysis buffer composed of 40 mM Tris pH 8, 300 mM NaCl, 10 mM EDTA, 2 % Triton X-100 and 1:100 protease inhibitors (Sigma-Aldrich) in a 1:1 ratio (v/v). Then, samples were frozen at -20°C for at least 8 hours and then thawed on ice and sonicated five cycles of 5 seconds at amplitude 100% (Labsonic M, Sartorius Stedim Biotech) to ensure membrane disruption. The extracted proteins were stored at -20°C until needed.

To extract proteins from ELVs for the validation phase (Objective 3), the detergent contained in the lysis buffer was changed for <1% NP-40, to make the protein extraction already suitable for direct in-solution digestion and LC-MS/MS. The rest of the procedure remained the same.

#### 5.1.2 Protein extraction from soluble fractions

For objective 1, protein extraction of the soluble fraction was performed after protein precipitation. A 100% stock solution of acetonitrile was added to the soluble fraction in a 1:5 (v/v) ratio and was incubated at -20 °C overnight. Then, sequential centrifugations were performed at 14,000xg 4 °C for 30 and 15 min, respectively. Finally, the pellet was dried and resuspended in 500 µL of the lysis buffer containing Triton X-100.

### 5.1.3 Protein extraction from UAs

The fluid fractions of UAs were sonicated 5 cycles of 5 seconds each at 100% amplitude (Labsonic M, Sartorius Stedim Biotech) to disrupt microvesicles, protein aggregates, and/or mucus present in the sample. Then, the Albumin & IgG depletion spin trap Kit (GE Healthcare) was used following the manufacturer's instructions to remove albumin and immunoglobulin G from the samples. The extracted proteins were stored at -20°C until needed.

### 5.2 Protein quantification

Protein concentration was determined using the Bio-Rad DC Protein Assay (Bio-Rad Laboratories) following the manufacturer's recommendations. Samples, usually diluted 1 in 10, were compared in triplicates against serially diluted BSA as standard. Total protein concentration of UAs was determined in triplicates by the Bradford assay following the manufacturer's indications.

### 5.3 Western blotting

Proteins were separated by 10 % SDS-PAGE under reducing or non-reducing conditions and transferred to PVDF membranes. For blocking, membranes were soaked in 5 % non-fat dried milk in TBS-Tween20 (0.01 %). Proteins were immunodetected using primary antibodies incubated at 4°C overnight. Then the membranes were washed and incubated with a secondary HRP-coupled antibody for 1 hour at room temperature. Finally, HRP signal was revealed using the Immobilon Western Chemiluminescent HRP Substrate (Merck Millipore), or ECL Western Blotting System (GE Healthcare). When required, the intensity of the bands was quantified using the Image J software (v. 1.45s)<sup>210</sup>.

Primary antibodies: **mouse anti-CD9** (1:250; ref. 555370, BD Biosciences), **mouse anti-CD63** (1:1000; ref. OP171, Calbiochem), mouse anti-CD81 (1:1000; ref. sc-166028, Santa Cruz), mouse anti-TSG101 (1:500; ref. Ab83, Abcam), mouse anti-Flotillin-1 (1:250; ref. 610821, BD Biosciences), rabbit anti-Annexin V (1:1000; ref. ab108321, Abcam), mouse anti-Haptoglobin (1:1000; ref. ab13429, Abcam), and rabbit anti-GRP78 (1:1000; ref. ab21685). The primary antibodies in bold were separated under non-reducing conditions.

Secondary antibodies: rabbit anti- mouse Immunoglobulins/HRP, 1:2000, ref. P0260, Dako; and goat anti-rabbit Immunoglobulins/HRP, 1:2000, ref. P0448, Dako.

## **5.4 Protein digestion and sample preparation for Mass Spectrometry**

### **5.4.1 Super-SILAC mix and the in-solution digestion protocol**

To digest the proteins extracted from UAs and EC cell lines-derived ELVs, Filter-Aided Sample Preparation (FASP) was performed using a 10kDa molecular weight cutoff filter (Millipore), following the protocol described by Manza et al.<sup>211</sup>. Briefly, 20 µg of protein from each sample were loaded in the filter unit and washed twice with 8M urea, by centrifuging at 14000xg (4°C) for 15 minutes. Proteins were then reduced with 10 mM DTT for 1 hour at room temperature, and alkylated with 30 mM IAM (iodoacetamide) for 30 minutes in the dark. The reaction was stopped incubating with 37.5 mM ACN (acetonitrile) for 15 minutes. Next, 15 minutes centrifugation at 14000xg permitted the removal of all the solutions and, after that, the samples were washed once with 1M urea. The resulting concentrate was diluted with 40 µL of 1M urea containing 20 µg of trypsin and incubated at 37°C overnight. Finally, the samples were centrifuged 10 minutes at 14000xg and the eluted tryptic peptides were collected in a new tube, where they were acidified with 0,5 µl of formic acid. Samples were stored at -20°C until further LC-MS/MS analysis to compare the proteome of the 5 EC cell lines and to compare the in-gel and the FASP (in-solution) digestion protocols.

### **5.4.2 Discovery Phase (In-Gel Digestion)**

A total of 21 µg of IS (heavy) was added to 10.5 µg of each patient's sample (light). Every patient's sample with its IS spiked-in was boiled for 5 minutes with Laemmli buffer for gel separation. Ten-wells 1D-SDS PAGE gels (1.0 mm X 10 well NuPAGE 4-12% Bis-Tris Gel, invitrogen Thermo Fisher Scientific) were used for the electrophoretic separation. Samples were loaded randomly in 4 gels but ensuring the presence of samples from all 3 groups of study in each gel. After running the gels for 30 minutes at 150 volts, they were stained in a Coomassie

solution for 1 hour (0.1% Coomassie in 40% methanol/10% acetic acid solution, all purchased from Sigma). Then, each gel lane was excised in 10 bands and each band was chopped and frozen at -20°C in a 1.5 mL tube for subsequent digestion. Before the trypsin digestion protocol, all gel slices were destained with Coomassie destaining solution (50/50 acetonitrile/ddH<sub>2</sub>O). Once destained, gel pieces were covered with 10mM DTT in 50 mM NH<sub>4</sub>HCO<sub>3</sub> (ammonium bicarbonate or ABC) and incubated at 56°C for 30 minutes. Next, the DTT solution was removed and 100 µl of 55 mM IAM in ABC were added and incubated in dark for 30 minutes at room temperature. Then, the gel slices were washed with ABC and dehydrated with ACN before adding ABC containing 200 ng of trypsin. Samples were digested with trypsin for 8 hours at 37°C and then centrifuged at maximum speed for a few seconds to collect all the digestion solution supernatant and transfer it to a clean tube. To optimize the peptide extraction, 100 µL of extraction buffer (90% ACN/10% milliQ water) was added to the gel pieces and incubated for 15 minutes to centrifuge and collect the supernatant again. Finally, samples were evaporated to dryness in a Savant SPD131DDA SpeedVac Concentrator (Thermo Fisher Scientific) and suspended in 0.1% formic acid (FA) before LC-MS/MS analysis. This procedure applies also for the in-gel digestion performed on the two samples used for the comparison between in-gel and in-solution digestion.

#### **5.4.3 Verification Phase (In-Solution Digestion)**

To precipitate proteins, six volumes of cold acetone (Sigma) was added to each cold sample tube and incubated at -20°C overnight. Then, tubes were centrifuged for 10 minutes at 16000xg at 4°C and acetone was removed without disturbing the pellet. When samples were dry, they were dissolved in the corresponding volume of 6M urea to have each sample at a protein concentration of 1 µg/µl. From most of the 107 samples of the validation phase an aliquot of 20 µg of precipitated protein was used for in-solution digestion and, for those for which the amount of material was not sufficient (but higher than 10 µg) the whole sample was used. The protein samples were reduced for 1 hour at 37°C with 10 mM DTT and then alkylated for 30 minutes in dark at room temperature with 20 mM IAM. Samples were diluted with 200 mM ABC to have samples at 2M urea before the addition of Lysine C at a 1:10 ratio enzyme:protein (w:w). Samples were digested overnight with LysC at

37°C. Then, samples were diluted to less than 1 M urea before adding a 1:10 ratio enzyme:protein (w:w) of sequence-grade trypsin. Samples were incubated for 8 hours at 37°C and, after that, the reaction was stopped by adding neat FA (10% of final total volume). A clean-up step was performed following an off-line reverse phase purification protocol using UltraMicroSpin C18 300A silica columns to desalt the samples. Briefly, columns were conditioned with 100% methanol and equilibrated with 5% FA in water before loading twice the acidified samples into the columns. Then, the columns were washed with 5% FA in water and eluted with 5% FA in a 1:1 solution of ACN:water (v/v). Finally, the solvent of the eluted samples was evaporated to dryness using a speed-vac system and resuspended in 0.1% FA before the SRM analysis. All incubation steps were performed at 650 rpm shaking when possible.

### **5.4.3 Validation in Uterine Aspirates**

Samples containing 25 µg of protein were denatured by incubation for 20 min at 22°C under agitation with a solution of 10 M urea in 50 mM ABC. Then, the mixture was incubated 10 min in an ultrasonic bath (Branson 5510, Branson Ultrasonics). Next, samples were reduced for 60 min at 37°C with 200 mM DTT and alkylated for 30 min in the dark at 22°C with 200 mM IAM. The samples were then digested for 4 hours at 37°C with LysC at a 1:150 ratio enzyme:protein (w:w). Next, the samples were diluted with 50 mM ABC to a final urea concentration of 1M and then they were incubated overnight at 37°C with trypsin at a 1:50 ratio enzyme:protein (w:w). To quench the activity of trypsin, 1 µl of neat FA per 100 µl of solution was added. At this point, the heavy synthetic peptides were spiked in the samples. Next, samples were desalted by solid phase extraction (Sep Pak tC18, 50 mg, Waters). Finally, the eluates were evaporated to dryness in a vacuum centrifuge and suspended in 0.1% FA before LC- SRM analysis.

## 5.5 Mas Spectrometry analysis

### 5.5.1 Comparison of the cell lines-derived ELVs proteome

This part of the proteomic analysis was conducted in collaboration with the Proteomics Laboratory of the Vall Hebron Institute of Oncology (VHIO) in Barcelona, Spain.

#### 5.5.1.1 LC-MS configuration

Tryptic digests were analyzed using a linear ion trap Velos-Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). Instrument control was performed using Xcalibur software package, version 2.1.0 (Thermo Fisher Scientific, Bremen, Germany). Peptide mixtures were fractionated by on-line nanoflow liquid chromatography using an EASY-nLC system (Proxeon Biosystems, Thermo Fisher Scientific) with a two-linear-column system. Digests were loaded onto a trapping guard column (EASY-column, 2 cm long, ID 100  $\mu$ m and packed with Reprosil C18, 5  $\mu$ m particle size from Proxeon, Thermo Fisher Scientific) at a maximum pressure of 160 Bar. Then, samples were eluted from the analytical column (EASY-column, 10 cm long, ID 75  $\mu$ m and packed with Reprosil, 3  $\mu$ m particle size from Proxeon, Thermo Fisher Scientific). Separation was achieved by using a mobile phase from 0.1% FA (Buffer A) and 100% acetonitrile with 0.1% FA (Buffer B) and applying a linear gradient from 5 to 35% of buffer B for 60 min at a flow rate of 300 nL/min. Ions were generated applying a voltage of 1.9 kV to a stainless steel nano-bore emitter (Proxeon, Thermo Fisher Scientific), connected to the end of the analytical column. The LTQ Orbitrap Velos mass spectrometer was operated in data-dependent mode. A scan cycle was initiated with a full-scan MS spectrum (from m/z 300 to 1600) acquired in the Orbitrap with a resolution of 30,000. The 20 most abundant ions were selected for collision-induced dissociation fragmentation in the linear ion trap when their intensity exceeded a minimum threshold of 1000 counts, excluding singly charged ions. Accumulation of ions for both MS and MS/MS scans was performed in the linear ion trap, and the AGC target values were set to  $1 \times 10^6$  ions for survey MS and 5000 ions for MS/MS experiments. The maximum ion accumulation time was 500 and 200 ms in the MS and MS/MS modes, respectively. The normalized collision energy was set to 35%, and one microscan was acquired per spectrum. Ions subjected to MS/MS with a



relative mass window of 10 ppm were excluded from further sequencing for 20 s. For all precursor masses a window of 20 ppm and isolation width of 2 Da was defined. Orbitrap measurements were performed enabling the lock mass option ( $m/z$  445.120024) for survey scans to improve mass accuracy.

#### *5.5.1.2 Protein Identification*

LC-MSMS data was analyzed using the Proteome Discoverer software (Thermo Fisher Scientific) to generate mgf files. Processed runs were loaded to ProteinScape software (Bruker Daltonics, Bremen, Germany) and proteins were identified using Mascot (Matrix Science, London UK) to search against the SwissProt database. MS/MS spectra were searched with a precursor mass tolerance of 10 ppm, fragment tolerance of 0.05 Da, trypsin specificity with a maximum of 2 missed cleavages, cysteine carbamidomethylation set as fixed modification and methionine oxidation as variable modification. Significance threshold for the identifications was set to  $p < 0.05$ , minimum ions score of 20.

### **5.5.2 Discovery Phase**

This part of the proteomic analysis, as well as the previous proteomic setting up experiments, were conducted during a 10-months internship at Dr. Pierre Thibault's laboratory, in the Proteomics and Bioanalytical Mass Spectrometry Research Unit at the Institute of Research in Immunology and Cancer (IRIC) in Montreal, Canada.

#### *5.5.2.1 LC-MS configuration*

Peptides dissolved in 0.1% formic acid were first loaded on a 150  $\mu$ m ID x 20 cm nano-LC in a house packed column (Jupiter C18, 3  $\mu$ m, 300 Å, Phenomenex, Torrance, CA) and separated with an EASY nanoLC system (Thermo Scientific). For the elution of the peptides, a 1 hour linear gradient of 5–40% ACN (0.2% FA) at a constant flow rate of 600 nL/min was used. The LC system was coupled to a QExactive plus Orbitrap tandem mass spectrometer (Thermo Fisher Scientific) and RAW files were acquired with XCalibur software (Thermo Fisher Scientific). Tandem mass spectra were performed with a Top-12 method with precursor isolation window of  $m/z$  2.0. The resolution was 70,000 at  $m/z$  400 for the survey scan (with AGC  $1e^6$ , maximal injection time 200 ms and a scan range of  $m/z$  300–

2000) and 17.500 for MS/MS spectra (AGC at  $5e^5$ , maximal injection time 50 ms and scan range m/z 200-2000). Normalized collision energy (NCE) was set at 25 and exclusion time was set to 10 s.

#### *5.5.2.2 Protein identification*

For SILAC quantification, RAW files were analyzed using MaxQuant (version 1.3.0.3). Default parameters were taken: MS/MS fragment error tolerance of 20 ppm, Carbamidomethylatin (C) fixed and Oxidation (M) as well as Acetyl (Protein N-term) as variable modification. We selected arginine (Arg10) and lysine (Lys8) for the heavy label and applied the “match between run” option. The RAW files were searched against the HUMAN SwissProt database.

For the Presence/Absence analysis, a label free analysis was done with PEAKS 7.0 (<http://www.bioinform.com/>). For the database search, RAW files were searched against the Human Uniprot database containing 37254 entries. For the search parameters, precursor error tolerance was set at 10 ppm and for the MS2 fragments the tolerance was at 0.01Da. A maximum of 2 misscleavages for Trypsin were allowed. Carbamidometylation (C), Deamidation (NQ) and Oxidations (M) were variable modifications. Only peptides with false discovery rates (FDR) lower 1% were considered. For validation of potential biomarkers, PEAKS results were uploaded in Scaffold proteome software (<http://www.proteomesoftware.com/>).

#### *5.5.2.2 Statistical analysis*

The statistical analysis was performed to obtain two different outputs: (1) A qualitative data consisting of proteins that were present or absent in the different subgroups of patients; and (2) a quantitative data consisting of the expression measures obtained from SILAC ratios representing relative abundance of each protein vs. internal standards.

For the qualitative data, a Fisher exact test was applied to each protein to compare presence/absence between groups. For the quantitative data, a Student t-test was performed between each pair of groups in order to select differentially expressed proteins. The test was performed only for those proteins that were present in more than 4 individuals per group. In order to address the problem of multiple comparisons derived from performing many tests (one per protein), the p-values were adjusted to obtain a strong control over FDR using the Benjamini and Hockberg methods.

All analyses were performed using the statistical program "R" (R version 3.2, Copyright (C) 2015 The R Foundation for Statistical Computing).

### 5.5.3 Verification Phase

Selected reaction monitoring (SRM) analysis was performed in collaboration with the Proteomics Unit of the Centre for Genomic Regulation (CRG) and University Pompeu Fabra (UPF) in Barcelona.

#### 5.5.3.1 Monitored peptides selection

A total of 54 proteins were selected for the targeted experiment by SRM based on the results from the discovery phase (49), together with 5 candidates selected from previous results of our group (Table 16). Two unique peptides per protein were selected to be monitored by SRM based on their detectability in previous mass spectrometric experiments. For each selected peptide, an isotopically-labeled version ( $^{15}\text{N}_2$ ,  $^{13}\text{C}_6$ -Lysine,  $^{15}\text{N}_4$ ,  $^{13}\text{C}_6$ -Arginine) was bought and spiked in each sample to be used as internal standard. Internal standards were spiked in a concentration within the linear response range, which was established for each individual peptide based on experimental dilution curves (data not shown). In parallel, isotopically-labeled peptides were mixed and used to generate MS2 spectral library and retention time knowledge database.

#### 5.5.3.2 SRM configuration

A total of 85 endogenous peptides and their corresponding internal standards were measured by SRM in Lys-C and trypsin-digested samples from an independent cohort of 107 patients. The five most intense transitions per peptide were monitored on a triple quadrupole mass spectrometer (5500 Q-Trap, AB Sciex Instruments, Foster, CA, USA) equipped a reversed-phase chromatography 25-cm column with an inner diameter of 75  $\mu\text{M}$ , packed with 1.9  $\mu\text{M}$  C18 particles (NikkyoTechnos Co., Ltd. Japan) and a 2-cm pre-column (Acclaim PepMap 100, C18, 15  $\mu\text{M}$ , 100A). Loading buffer: H<sub>2</sub>O with 0.1% formic acid; eluting buffer: ACN, 0.1% formic acid. The flow rate used was 250 nL/min and a chromatographic gradient ranging from 5 to 40% eluting buffer in 40 min was used. Blank runs were performed between the SRM measurements of biological samples to avoid sample

carryover. Measurements were done in scheduled SRM mode, using a window of 300 seconds and a target cycle time of 1.5 seconds.

### *5.5.3.2 Data analysis*

Transition groups corresponding to the targeted peptides were evaluated with Skyline v2.5 based on the co-elution of the transition traces associated with a targeted peptide, both in its light and heavy form; the correlation between the light SRM relative intensities and the heavy counterpart. All peptide peaks were visually inspected. Areas of all transitions were used as input for the linear mixed-effects model implemented in the MSstats Skyline plug-in (v3.3) to calculate protein fold-changes and adjusted p-values among the different sample groups.

### *5.5.3.3 Development of predictors*

MSstats was used to estimate the quantity of proteins present in all samples based on log<sub>2</sub>-transformed transition areas, which were then used as input variables to a logistic regression model between defined groups. The classification ability of each protein was evaluated within 2/3 of the dataset using the area under the curve (AUC) as performance indicator. The most discriminative protein was selected as the first classifier. Most discriminative proteins were repeatedly added while increasing AUC values ( $\Delta\text{AUC} > 0.02$ ). The procedure of classification evaluation was repeated 500 times using a different subset of patients in each iteration. Sample subsets were generated by randomly selecting patients from the original cohort without replacement and balanced for each sample subgroup. A final consensus model was comprised of the combination of proteins, which were selected most in 500 repeats. The final model was fitted to the full dataset and the predictive accuracy was quantified using the area under the ROC curve, sensitivity, specificity, and accuracy. The pROC package for the statistical program "R" was used to draw ROCs, to calculate AUCs and other performance values (i.e. sensitivity, specificity, and accuracy); these were obtained considering the "optimal cutoff" as the point where the sum of sensitivities and specificities reached the maximum value.

### 5.5.4 Validation in UAs

To further validate protein candidates in the whole fluid fraction of UAs of a new cohort of patients (cohort C, n=67), we followed the same steps and criteria than in the verification in ELVs from UAs (Section 5.5.3). In this case, a total of 51 protein candidate biomarkers were selected, and 75 endogenous peptides and their corresponding internal standards were measured by SRM. For this experiment, those proteins that were not detected in any patient of the verification phase in ELVs were not included for monitoring in UAs, and 3 additional proteins of interest for previous results obtained in the laboratory were added (KP YM, PIG3 and CLIC1).

“SRM configuration”, “Data analysis” and “Development of predictors” sections apply the same way for this new experiment.

## 6. Gene ontology analysis

For the gene ontology analysis, we used the data generated from the label free analysis of the discovery phase. To ensure robustness of the protein identifications, we filtered-out low frequent proteins and we kept only proteins identified by at least five different peptides, being a minimum of one of those a unique peptide. We removed redundancy data manually.

GO functional classification and overrepresentation tests for biological process, molecular functions, pathways and protein class were performed using the free on-line available *Panther* software (Protein Analysis Through Evolutionary Relationships)<sup>212</sup>. The statistical method used to conduct the overrepresentation test is the binomial test. *Funrich* software (Functional Enrichment analysis tool)<sup>213</sup> was also used for analysis and graphical representation of the proteomic data.

## 7. Statistical analysis and Software

For Objective 1, statistical analyses were performed using GraphPad Prism 5 software. The Student's t test was applied to compare means of ELVs concentration, particle size distribution, and expression of tetraspanins. The

Pearsons' Rho test was used to analyze correlation. The probability of  $p < 0.05$  was considered statistically significant.

For Objectives 2 and 3, statistical analyses have been detailed together with the mass spectrometry analysis for better comprehensiveness.

Adobe Illustrator CS5, Adobe Photoshop CS5, GraphPad Prism 6 and Microsoft Power Point have been used to generate the graphic material presented in this work.



# RESULTS

---



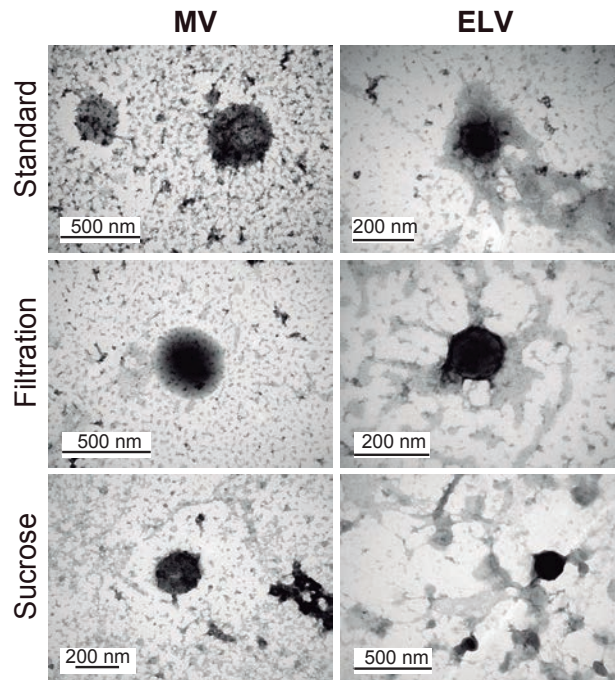


## RESULTS OF OBJECTIVE 1

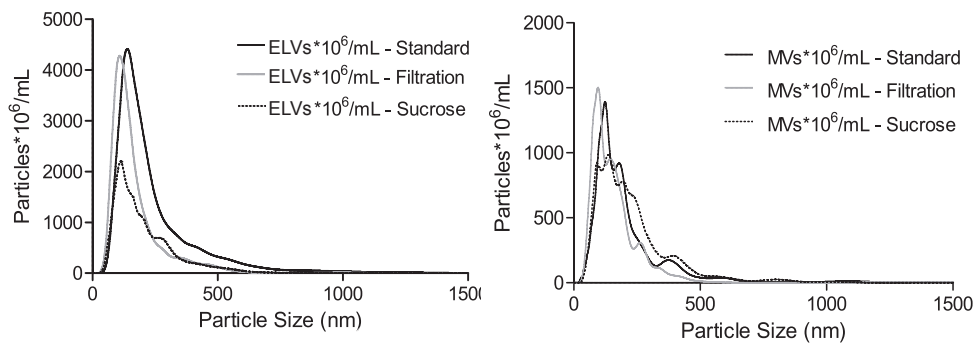
*The results obtained for Objective 1 partially overlap with those presented in the doctoral thesis of Lucia Lanau. This is part of a work done together in collaboration that resulted in the publication of a paper of which we both are first co-authors<sup>214</sup>.*

### **1. Establishment of a suitable protocol for the use of ELVs from UAs**

UAs are used in the diagnostic process of EC, yet further applications could emerge if its complex milieu was simplified. ELVs isolated from UAs could become an attractive source of biomarkers, but there is a need to standardize isolation protocols. The first objective of this thesis was to determine whether ELVs exist in the fluid fraction of UAs and to compare protocols for their isolation, characterization, and analysis. For that, we collected UAs from 33 pre-menopausal women suffering from benign gynecological diseases. The fluid fraction of 27 of those aspirates were pooled and split into equal volumes to evaluate three differential centrifugation-based procedures: (1) a standard protocol, (2) a filtration protocol, and (3) a sucrose cushion protocol. Characterization of isolated vesicles was assessed by electron microscopy, nanoparticle tracking analysis, and immunoblotting. Then, the selected protocol was tested in 6 non-pooled samples. All three protocols permitted the isolation of ELVs of the expected round cup shape, as observed by TEM (Figure 18). Both ELVs and MVs appeared as well-defined bilayer vesicles but, notably, the size of all ELVs was smaller than that for MVs, especially in the case of Standard protocol. To further investigate ELVs concentration and size distribution, samples were analyzed by NTA (Figure 19 right). The population of ELVs isolated by the Standard and Filtration protocols followed a uniform size distribution with a unique peak corresponding to a mode of  $135 \pm 5$  and  $115 \pm 3$  nm, respectively. For the Sucrose protocol, the distribution was not uniform; the mode was  $135 \pm 42$  nm but presented an additional peak around 300 nm, and a high standard deviation was observed indicating less reproducibility of this isolation protocol.



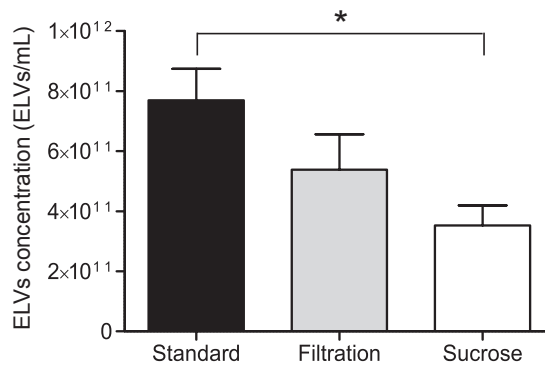
**Figure 18.** . Electron microscopy image of negatively stained ELVs and MVs isolated from UAs by the three different centrifugation-based protocols tested.



**Figure 19.** Size distribution of isolated ELVs (left) and MVs (right) by NTA.

Differently, all MVs preparations presented a heterogeneous distribution, and a lower concentration than that for ELVs (Figure 19 left).

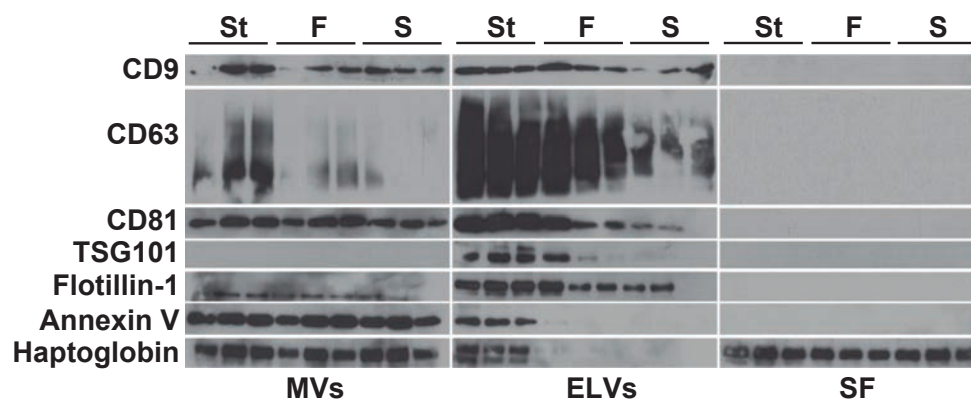
A reduction in ELVs concentration was seen as more steps were added to the isolation protocol; significant differences were observed between Standard and Sucrose protocols ( $p = 0.029$ ), and the same tendency was observed when comparing the Standard and Filtration protocols (Figure 20).



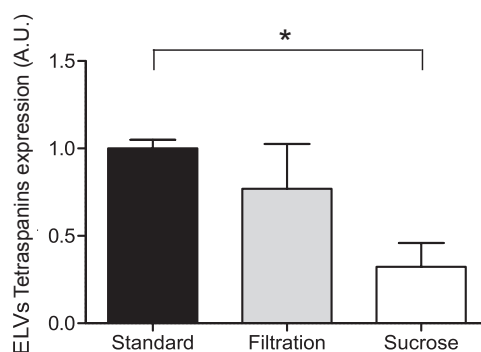
**Figure 20.** ELVs concentration measured by NTA.

## 1.1 Molecular characterization of ELVs from UAs

To evaluate the purity of ELVs obtained from each isolation protocol, we performed an immunoblot loading equal amounts of protein, and demonstrated that the expression of the tetraspanins CD63, CD9, and CD81 (all considered late endosomal markers enriched in ELVs) was significantly higher in the Standard compared to the Sucrose protocol ( $p = 0.001$ ) (Figures 21 and 22). The same tendency was observed for TSG101, a known endosomal origin marker, and Flotillin-1, an element of the membrane lipid rafts. These two markers were practically undetectable in MVs preparations, indicating a different biogenesis of these vesicles. As expected, Annexin V, a marker of MVs, was highly expressed in all MVs preparations; however, it was also detected in ELVs derived from the Standard protocol suggesting that the smallest MVs populations might have precipitated at  $100,000 \times g$  or that specificity of this marker is arguable. None of the MVs or ELVs markers were detected in the soluble fraction, but haptoglobin (an abundant protein found in blood) was highly expressed (Figure 21).



**Figure 21.** Immunoblot of isolated MVs, ELVs and SF (done in triplicates) against ELVs/MVs related markers and Haptoglobin.

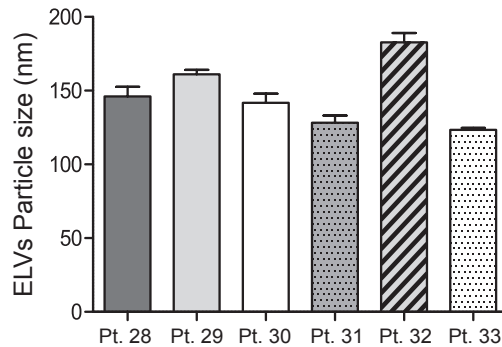


**Figure 22.** Relative tetraspanins expression of ELVs. The average of tetraspanins (CD9, CD63 and CD81) expression of each protocol was normalized to the Standard in order to have a relative measurement of ELVs purity.

Altogether, we demonstrated that all protocols were able to enrich the sample in ELVs. Nevertheless, we selected the Standard protocol for further applications since it yielded a higher ELVs concentration, while maintained higher levels of ELVs-related markers and better reproducibility than the other tested protocols.

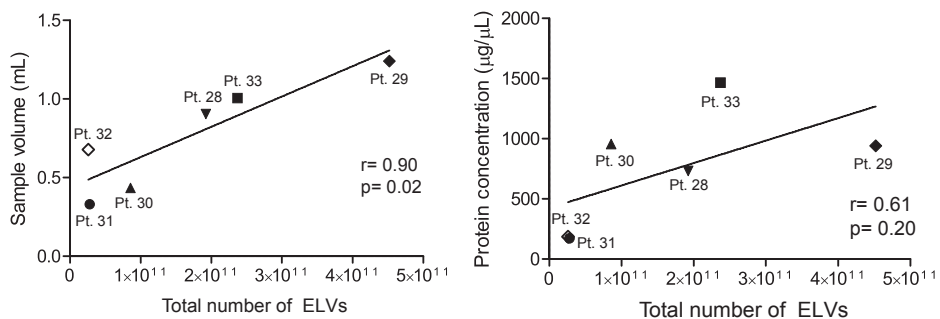
To confirm that the enrichment in ELVs following the Standard protocol holds when analyzing individual samples, we recovered ELVs, MVs and soluble proteins from the fluid fraction of 6 non-pooled UAs (samples 28–33, Table 9).

Concomitant to our previous observations in the pooled samples analysis, we observed that all ELVs preparations from individual UAs had a similar size distribution, presenting a mode of 120–160 nm (Figure 23).



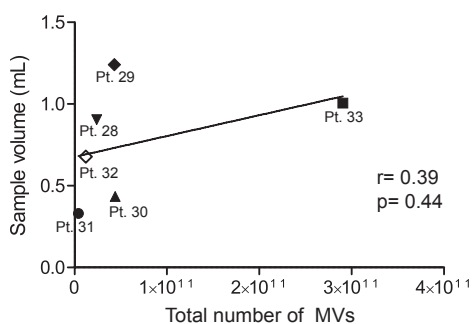
**Figure 23.** Particle size of ELVs isolated from individual UAs.

The particles concentration differed clearly between patients but the total number of isolated ELVs significantly correlated with the initial volume of UAs fluid fraction ( $r = 0.90$ ,  $p = 0.02$ ) (Figure 24 left), but not with protein concentration (Figure 24 right).



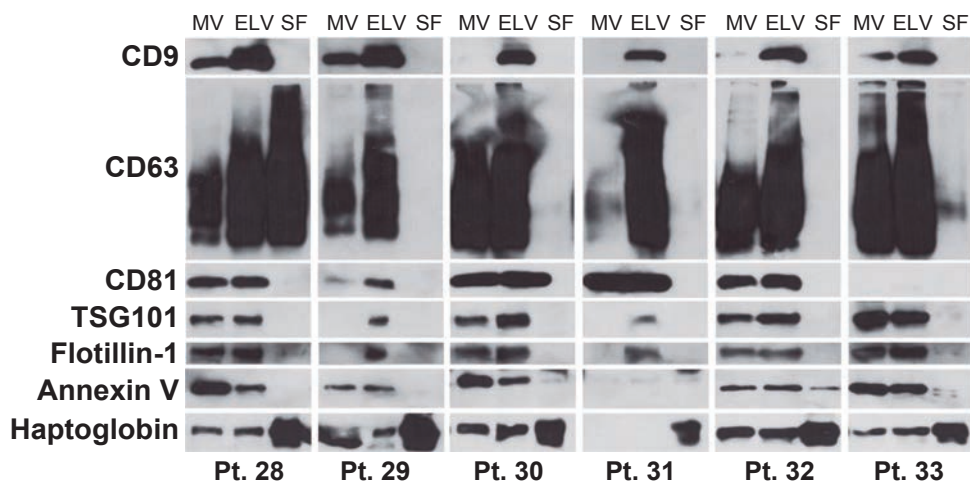
**Figure 24.** (A) Correlation between total number of isolated ELVs and starting volume of UA's fluid fraction. (B) Correlation between total number of isolated ELVs and ELV's protein concentration.

On the other hand, no correlation was observed between the number of MVs and sample volume (Figure 25).



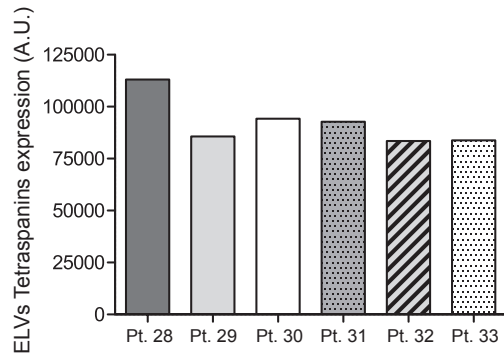
**Figure 25.** Correlation between total number of isolated MVs and starting volume of UA's fluid fraction.

ELVs markers were expressed in both ELVs and MVs preparations from all patients (Figure 26). As seen previously, tetraspanins expression was higher in ELVs than in MVs, indicating that we isolated a population of vesicles enriched in ELVs.



**Figure 26.** Characterization of MVs, ELVs and soluble fraction proteins by immunoblot against ELVs/MVs related markers and Haptoglobin.

Altogether, these results indicate that the Standard protocol is suitable to obtain ELVs from individual UAs.



**Figure 27.** Quantification of tetraspanins (CD9, CD63 and CD81) expression of ELVs for each individual sample.



## RESULTS OF OBJECTIVE 2

## 2. Identification of new protein biomarkers for EC in ELVs from UAs

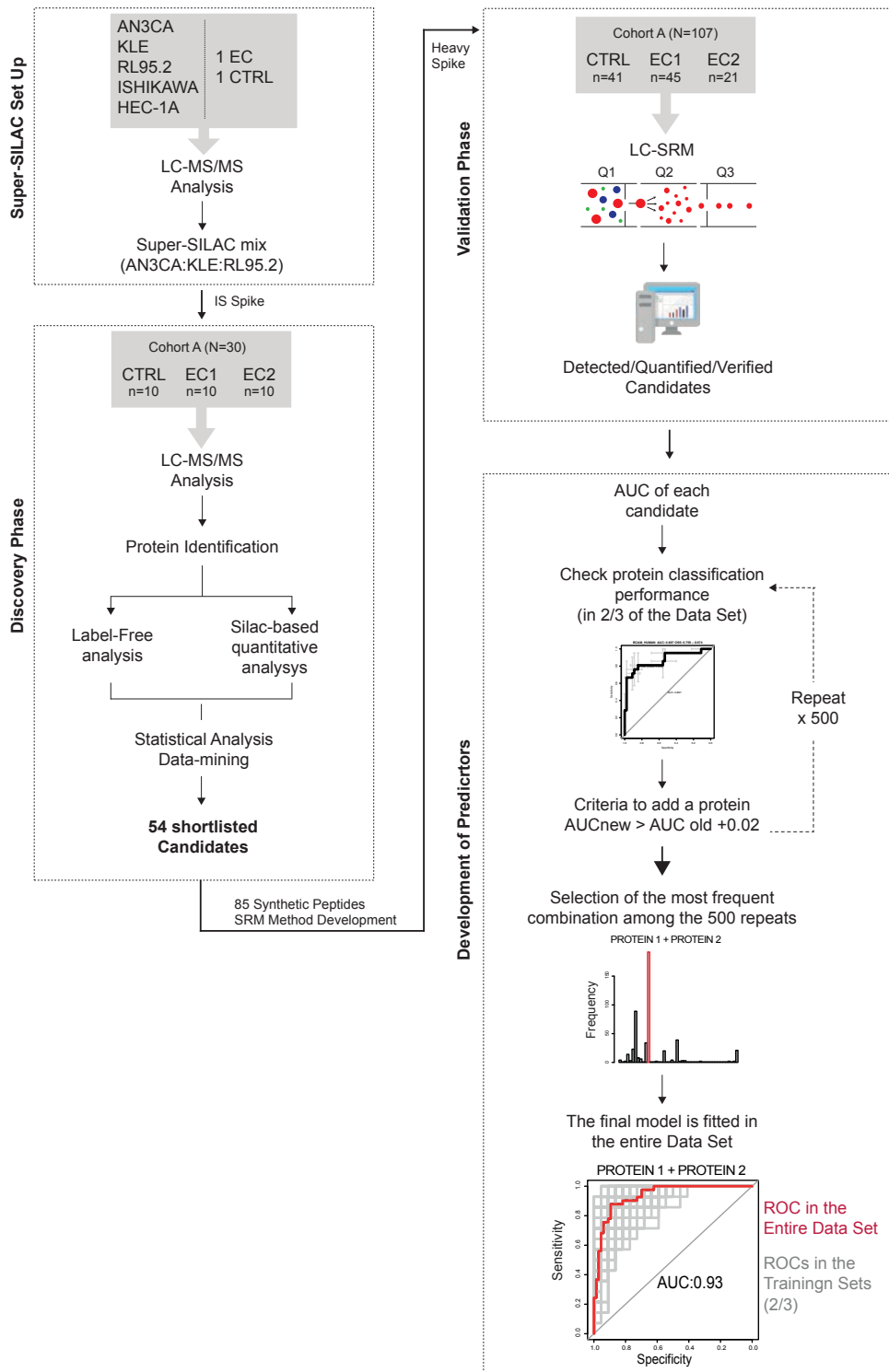
A general outlook of the workflow followed from the discovery to the verification of EC biomarkers is presented in Figure 28. As stated in objective 2, the discovery phase aimed to establish a Super SILAC-based mass spectrometry approach for the discovery of diagnostic biomarkers in ELVs obtained from uterine aspirates. Prior to this, we conducted a set of experiments to develop the Super-SILAC mix and to evaluate the optimum sample preparation and analytical methodology.

### 2. 1 Proteomic analysis of ELVs derived from 5 EC cell lines for the creation of the Super-SILAC mix

Clinical evaluation of protein biomarkers requires an accurate quantification. In this regard, the combination of ELV's proteome with known amounts of heavy standards enables robust relative quantification of putative candidates.

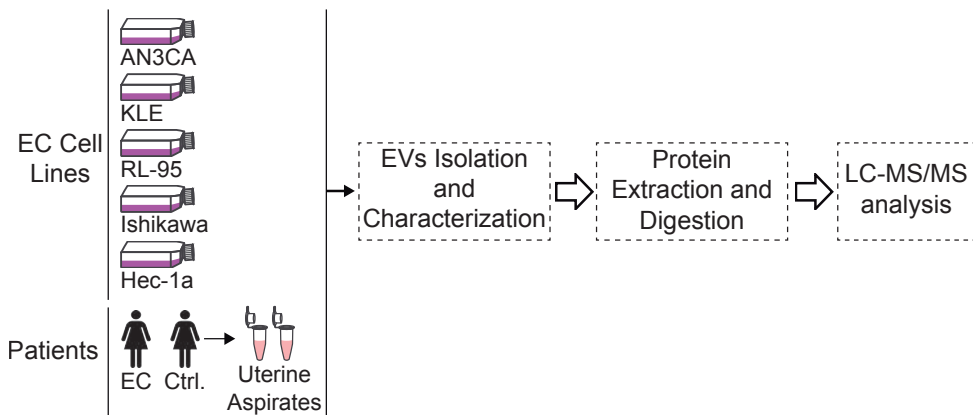
The super-SILAC approach is a variant of the conventional SILAC strategy applicable to tissue samples or biological fluids<sup>215</sup>. It consists in combining different cell lines (called super-SILAC mix), to be used as a spike-in standard. This approach enables the relative quantification of the proteome under investigation and allows the comparison of multiple samples.

The first step to develop an optimal Super-SILAC mix is to compare the proteome of different cell lines with that of our sample of interest in order to fulfill two criteria: the proteome of ELVs derived from cell lines should differ as much as possible from one another but still be similar to that of ELVs derived from UAs, in order to cover the maximum proteome of the sample of study, but not to include redundant information. To investigate that, we followed the workflow depicted in Figure 29.



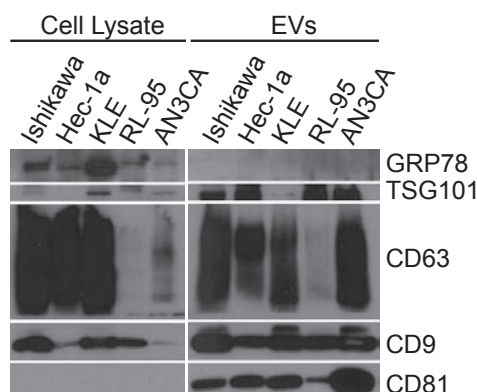
**Figure 28.** Workflow of the experimental procedure followed from the discovery phase to the generation of predictive models.

To characterize and compare the proteome of extracellular vesicles secreted by AN3CA, KLE, RL95.2, Ishikawa and Hec-1a cells, ELVs were isolated from their conditioned media. Since it has been suggested that ELV abundance in cell culture medium varies with culture time and as a factor of confluence, we tried to harvest the conditioned media after the same time and at the same confluence for all cell lines. After ultracentrifugation, pellets enriched in ELVs were characterized. Likewise, ELVs from two UAs (samples T1\_2 and CTRL\_1, from the discovery cohort) were isolated.



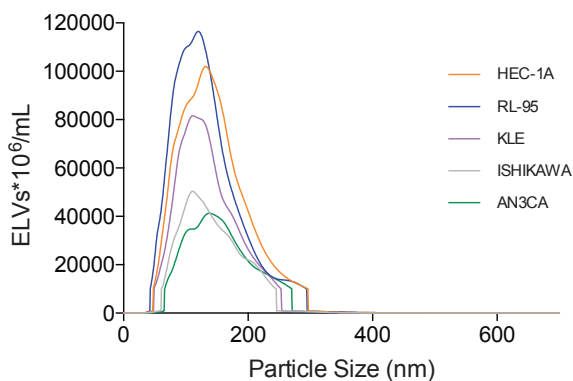
**Figure 29.** Schematic representation of the workflow followed to compare the proteome of ELVs from 5 EC cell lines and from UAs.

First, ELVs protein content was extracted and quantified to know the yield, which ranged from 67  $\mu$ g to 116  $\mu$ g. As shown in Figure 30, Western Blot analysis confirmed the presence of an enriched population of ELVs in our sample, as seen by the expression of several ELVs markers such as TSG101, CD63, CD9 and CD81; and the absence of GRP78.



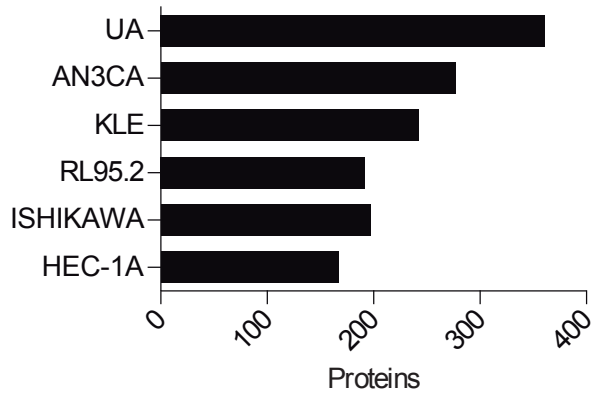
**Figure 30.** Characterization of EC cells and its derived ELVs by immunoblot against ELVs related markers and the ER marker GRP78.

Nanoparticle tracking analysis of EC cell line's ELVs reported sizes consistent with those previously reported for UAs (Figure 31).

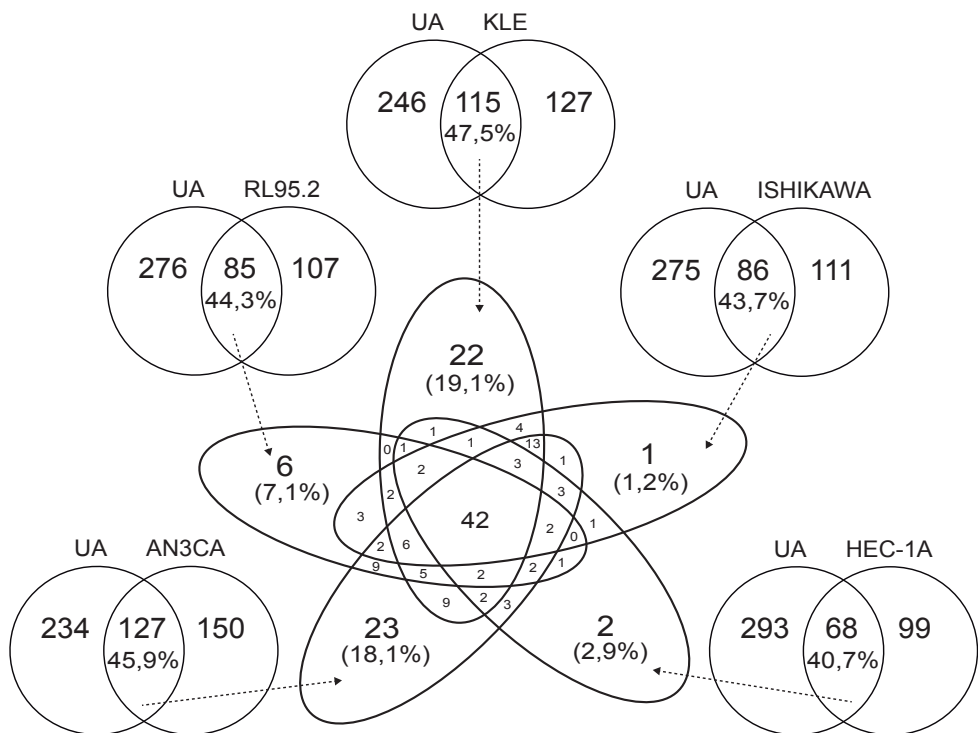


**Figure 31.** EC cell line's derived ELVs size distribution determined by NTA.

All samples, ELVs derived from cell lines and the two patients, were digested using the filter-aided sample preparation (FASP) method and injected on a linear ion trap Velos-Orbitrap mass spectrometer to generate the list of identified proteins. A similar number of proteins were detected in all cell line's ELVs, being AN3CA and HEC-1a the ones with the highest and the lowest number, respectively (Figure 32). In order to compare the cell line's ELVs to the patient's UA ELVs, we treated the information of both patients as a single dataset. Overall, around a 45% of the cell line's ELVs proteome overlapped with that of the UAs ELVs. Noteworthy, using two similar cell lines in the mixture does not contribute to the diversity of the represented UA and hence this effect should be avoided. Thus, we looked for proteins in common with UAs and unique for each cell line (Figure 33), finding that AN3CA, KLE and RL95.2 contributed with a higher number of unique proteins to the representation of the patient's UA ELVs. Ishikawa and Hec-1a cell lines were discarded because their proteome just represented UAs ELVs with 1 and 2 unique proteins, respectively.

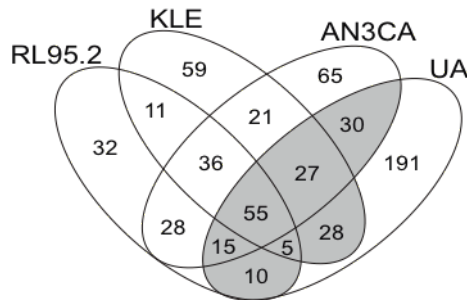


**Figure 32.** Number of proteins identified in each cell line and patient's derived ELVs



**Figure 33.** Comparison of proteins identified from cell lines and UA. Proteins that overlap are then compared between them to find which cell lines present the highest proportion of unique proteins in common with UAs.

Thanks to this study, we could develop the Super-SILAC mix. Theoretically, with the combination of AN3CA, KLE, and RL95.2, it was possible to represent about a 47% of the UAs EVs proteome (Figure 34).

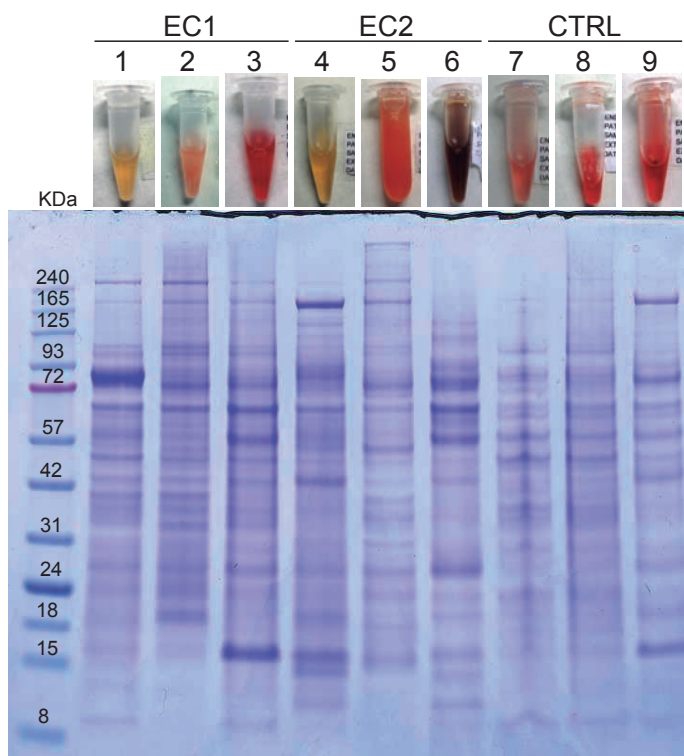


**Figure 32.** The 3 selected cell lines represent the 47% of the UAs ELVs proteome (gray shadow).

## 2.2 Setting up the proteomic approach for the discovery phase

As mentioned earlier, we had no references of previous proteomic studies on ELVs isolated from UAs, so we needed to perform some proof-of-concept experiments in order to define the best sample preparation and proteomic approach before moving to the discovery phase study.

First of all, we wondered if the different blood amount of the samples would be translated into different protein content in the exosomes. To check the variability among the samples regarding bloodiness, we took the isolated ELVs from 9 UAs samples (UAs' features are detailed in Table 10) covering from *clear* to *very bloody* samples within the 3 samples selected in each group (3 EC1, 3 EC2 and 3 CTRL). We separated the proteins in a gel by electrophoresis and stained them with Coomassie. As observable in Figure 35, there was a lot of variability among patients, including those classified within the same group. However, we did not observe a common pattern of bands for the bloodiest samples. We concluded that the major factor for heterogeneity between patients is not related to the blood content, but that this variability should be controlled in our discovery study.



**Figure 33.** Coomassie staining of a gel showing the pattern of bands corresponding to the proteome of ELVs isolated from UAs containing different amounts of blood contamination.

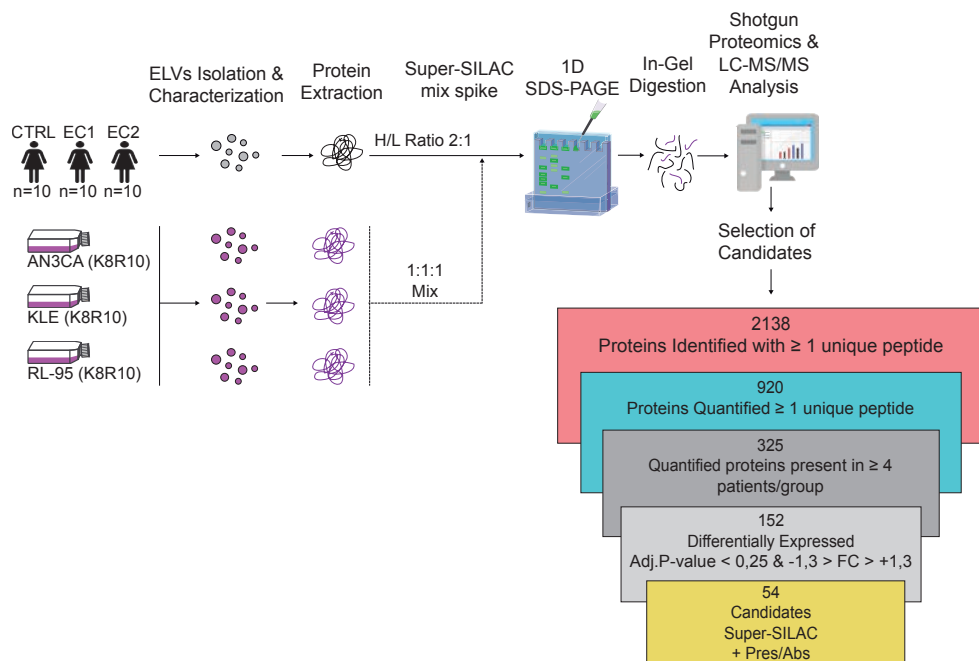
Next, we wanted to compare the in-gel and the in-solution digestion protocols in order to select the methodology that maximized the number of identifications. For that, we used samples 2 and 8 from the Pilot Studies cohort, we performed both digestion protocols, and compared the number of proteins identified by LC-MS/MS. Interestingly, the in-gel procedure not only permitted the identification of a higher number of proteins in both samples, but also enabled the identification of more known ELVs markers (Table 15). Of note, the number of proteins identified in sample 2 (EC1) is higher than in sample 8 (CTRL) using both digestion protocols, suggesting that ELVs from EC patients might have a higher proportion of proteins in their cargo as compared with ELVs from healthy donors.

**Table 15.** Comparison of the number of proteins identified by LC-MS/MS in samples 2 and 8 after in-gel or in-solution digestion. The presence of some ELVs markers has been evaluated.

Sample	Proteins identified	CD81	TSG101	CD63	CD9
2 in-gel	904	yes	no	yes	yes
2 in-solution	416	yes	no	no	yes
8 in-gel	435	yes	yes	yes	yes
8 in-solution	232	yes	no	no	yes

Thus, even though the in-gel procedure is time-consuming and more expensive, we selected this method as it enables the identification of a higher number of proteins, which is important for discovery studies.

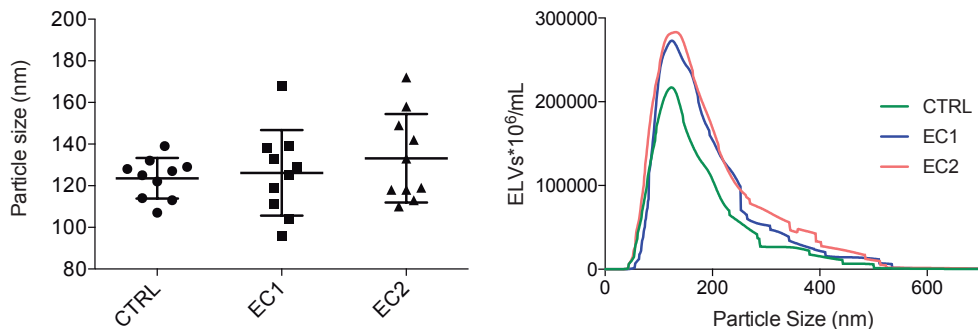
### 2.3 Protein identification and relative quantification with a SILAC-based LC-MS/MS discovery phase



**Figure 34.** Schematic representation of the discovery phase.

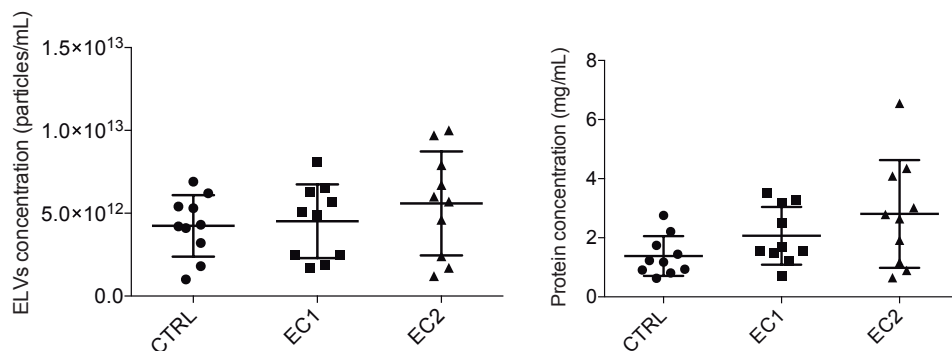


For the discovery phase of this project (Figure 36), uterine aspirates from 10 EC1 patients, 10 from EC2 and 10 Control patients (N=30) were used to isolate ELVs. To confirm the correct enrichment in ELVs, samples were subjected to several control analyses. NTA data revealed that all patient's UAs contained a homogeneous mixture of vesicles with mode sizes varying between 96 and 172 nm (Figure 37).



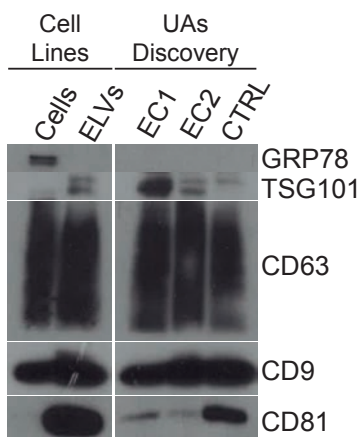
**Figure 35.** Mode (left) and size distribution (right) of ELVs isolated from the 30 samples included in the discovery phase determined by NTA. The size distribution is represented by the mean of the 10 samples of each group.

We observed a tendency towards higher ELVs concentration and protein content in ELVs from EC patients compared to controls although it was not significant (Figure 38).



**Figure 36.** (Left) ELVs and (right) protein concentration of the 30 samples used for the discovery phase.

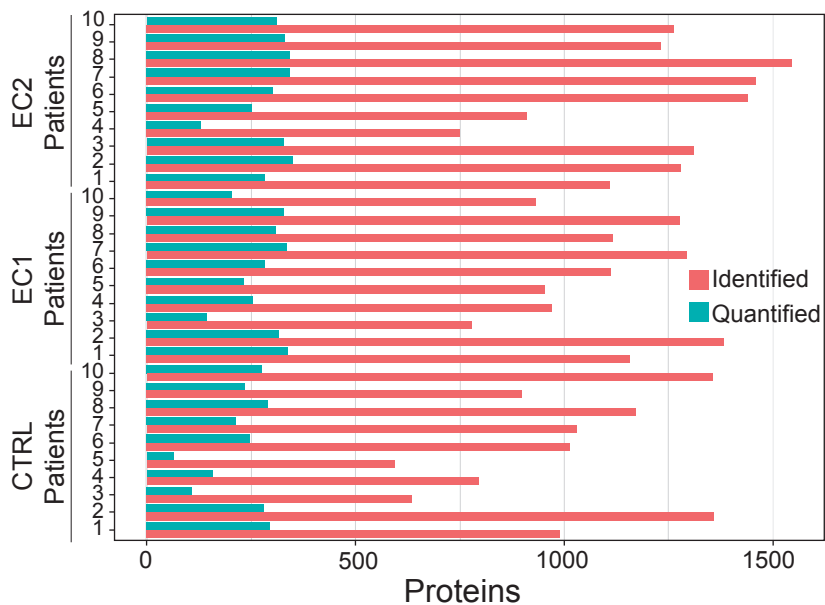
The ELV markers TSG101, CD63, CD9 and CD81 were detected in the pooled samples of each study group. The absence of GRP78 discarded a possible ER-contamination in our preparations (Figure 39).



**Figure 37.** Characterization by immunoblot against ELVs related markers of ELVs derived from the 30 patients of the discovery. Samples from each study group have been pooled. Cells lysates serve as a positive control for GRP78.

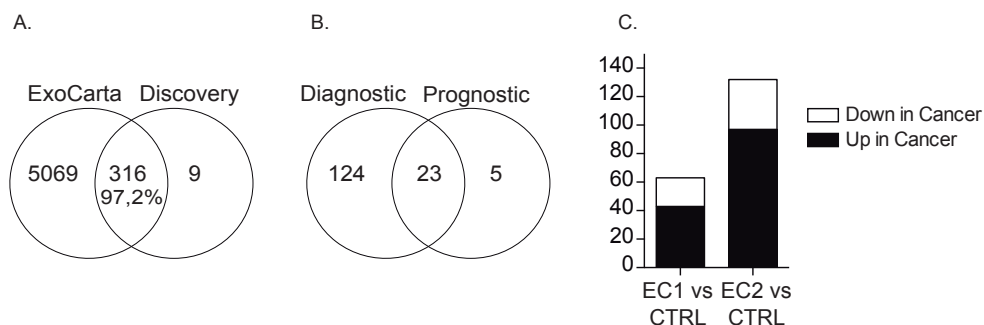
Once ensured the purity of the vesicles isolated, the same amount of Super-SILAC mix (internal standard) was spiked in each of the 30 samples (EC1 n=10, EC2 n=10 and CTRL n=10) in a 2:1 *Heavy/Light* ratio, right before the separation by 1D-SDS-PAGE. Each gel lane was sliced into 10 bands, and each band was subjected to trypsin digestion to extract peptides, resulting in 300 individual LC-MS/MS runs. Protein database searching of MS/MS data resulted in the overall identification of 2138 proteins, considering only those proteins identified with at least 1 unique peptide, and those peptides with a false discovery rate (FDR) lower than 1%. The MS/MS spectra were further processed with the software MaxQuant leading to the quantification of 920 proteins, that is an overall quantification rate of a 43% of the identified proteome. In Figure 40, results for the identification and quantification of proteins are shown for each individual patient. As seen in the figure, the percentage of quantified proteins for each individual patient is lower, about a 25%. Due to the high heterogeneity among patients seen in our pilot study, we only rely on quantified proteins that were present in more than 4 patients per group to perform the statistical analysis. Among the 325 proteins that fulfill this criteria, a total of 152 proteins were differentially expressed with adjusted p-value < 0,25 and

fold-change lower than -1,3 or higher than 1,3. From those, 147 proteins are potential diagnostic biomarkers (differential from the comparison CTRL vs. EC), and 28 proteins are potential prognostic biomarkers (differential from the comparison EC1 vs. EC) (Figure 41B). As seen in Figure 41C, most of the differentially expressed proteins are up-regulated in cancer.



**Figure 38.** Relation of proteins identified and quantified by LC-MS/MS in each patient.

We compare the list of 325 proteins against the ExoCarta database of exosomal proteins, the results showed that only 9 proteins have not been previously reported in ELVs but the vast majority of them (97,2%) are proteins known to be contained in exosomes, reinforcing the assumption that we are working with correctly isolated ELVs (Figure 41A). Also, the published top 25 proteins that are often identified in exosomal studies as exosomal markers were all found in our list except for ACTB and ALB ([http://exocarta.org/exosome\\_markers](http://exocarta.org/exosome_markers)).



**Figure 39.** (A) Match of the 325 proteins considered for the differential expression analysis with the ExoCarta database showing that 97,2% of the proteins in our list are described in exosomes. (B) Diagnostic and/or prognostic potential of the 152 differentially expressed proteins. (C) Bar graphs showing that most of the differentially expressed proteins are up-regulated in cancer.

In parallel to the quantitative analysis, a *presence/absence* study was conducted in order to take into consideration proteins that failed to be quantified because they lacked the heavy counterpart but were still relevant for being present in a certain group. For this, RAW files were reanalyzed with the PEAKS software and a Fisher Test was performed to select those proteins significantly present in a group ( $p$ -value < 0.001). A total of 30 candidates were included following this analysis, corresponding to 29 potential diagnostic biomarkers and 9 potential prognostic biomarkers (see table 16).

Altogether, a final list of 54 candidates was generated combining (i) the relative quantification analysis, (ii) the presence/absence study, and (iii) biological criteria such as the exclusion of proteins whose family was overrepresented in the candidate list (i.e. ribosomal proteins and histones), or proteins down-regulated in cancer; always respecting the statistical restrictions stated above. Additionally, five proteins of the lab's interest were added at this point for further verification by SRM (Table 16). From those 54 candidate biomarkers, 50 corresponded to diagnostic biomarkers and 15 to prognostic biomarkers, from which 37 had only diagnostic potential, 2 had only prognostic potential and 13 had both, diagnostic and prognostic potential.

**Table 16.** List of selected candidates for LC-SRM. RQ, Relative quantification analysis; P/A, Presence/Absence analysis; Lab, proteins known to be relevant for previous results of our group.

Entry Name	Gene Name	UniProt Number	Protein Name	Adj.p.val	FC	Significant Comparissons	Analysis
ANXA4_HUMAN	ANXA4	P09525	Annexin A4	0,07	2,06	T2 vs CTRL	RQ
BCAM_HUMAN	BCAM	P50895	Basal cell adhesion molecule	0,19	1,33	T2 vs CTRL	RQ
CD59_HUMAN	CD59	P13987	CD59 glycoprotein	0,02	3,45	T1 vs CTRL	RQ
				0,21	2,55	T1 vs T2	
CD81_HUMAN	CD81	P60033	CD81 antigen	0,02	3,40	T1 vs CTRL	RQ
				0,19	2,54	T1 vs T2	
CLH1_HUMAN	CLTC	Q00610	Clathrin heavy chain 1	0,07	1,43	T2 vs CTRL	RQ
				0,16	2,73	T1 vs CTRL	RQ
H10_HUMAN	H1F0	P07305	Histone H1.0	0,10	2,02	T2 vs CTRL	RQ
				0,19	-1,64	T1 vs CTRL	RQ
ILF2_HUMAN	ILF2	Q12905	Interleukin enhancer-binding factor 2	0,21	-2,22	T1 vs T2	
MVP_HUMAN	MVP	Q14764	Major vault protein	0,00	2,44	T2 vs CTRL	RQ
PODXL_HUMAN	PODXL	O00592	Podocalyxin	0,21	1,37	T1 vs T2	RQ
PPIA_HUMAN	PPIA	P62937	Peptidyl-prolyl cis-trans isomerase A	0,02	3,21	T1 vs CTRL	RQ
				0,19	2,34	T1 vs T2	
PSB3_HUMAN	PSMB3	P49720	Proteasome subunit beta type-3	0,10	2,16	T2 vs CTRL	RQ
RL11_HUMAN	RPL11	P62913	60S ribosomal protein L11	0,08	2,27	T2 vs CTRL	RQ
				0,09	2,65	T1 vs CTRL	RQ
RL12_HUMAN	RPL12	P30050	60S ribosomal protein L12	0,07	2,00	T2 vs CTRL	RQ
				0,21	2,32	T1 vs CTRL	
RL13A_HUMAN	RPL13A	P40429	60S ribosomal protein L13a	0,21	2,08	T2 vs CTRL	RQ
				0,01	4,37	T1 vs CTRL	
RL29_HUMAN	RPL29	P47914	60S ribosomal protein L29	0,07	2,68	T2 vs CTRL	RQ
				0,10	3,71	T1 vs CTRL	RQ
RS16_HUMAN	RPS16	P62249	40S ribosomal protein S16	0,10	3,71	T1 vs CTRL	RQ
				0,07	2,64	T2 vs CTRL	RQ
RSSA_HUMAN	RPSA	P08865	40S ribosomal protein SA	0,19	-2,40	T1 vs T2	
RUVB1_HUMAN	RUVBL1	Q9Y265	RuvB-like 1	0,01	1,34	T2 vs CTRL	RQ
TERRA_HUMAN	VCP	P55072	Transitional endoplasmic reticulum ATPase	0,02	1,47	T1 vs CTRL	RQ
Entry Name	Gene Name	UniProt Number	Protein Name	Adj.p.val	Significant presence in	Significant Comparissons	Analysis
ADA10_HUMAN	ADAM10	O14672	Disintegrin and metalloproteinase domain-containing protein 10	5,05E-05	T2	T2 vs CTRL	P/A
				8,91E-04	T1	T1 vs CTRL	
AGR2_HUMAN	AGR2	O95994	Anterior gradient protein 2 homolog	8,91E-04	T2	T2 vs CTRL	P/A
				8,91E-04	T1	T1 vs CTRL	
AGRIN_HUMAN	AGRN	O00468	Agrin	8,91E-04	T2	T2 vs CTRL	P/A
				5,90E-04	T1	T1 vs CTRL	
AR6P1_HUMAN	ARL6IP1	Q15041	ADP-ribosylation factor-like protein 6-interacting protein 1	5,90E-04	T2	T2 vs CTRL	P/A
				5,90E-04	T1	T1 vs CTRL	
CD14_HUMAN	CD14	P08571	Monocyte differentiation antigen CD14	1,77E-04	T2	T2 vs CTRL	P/A
				1,77E-04	T2	T1 vs T2	P/A
CLD6_HUMAN	CLDN6	P56747	Claudin-6	1,77E-04	T2	T1 vs CTRL	P/A
				1,00E-04	T1	T1 vs CTRL	P/A
FAS_HUMAN	FASN	P49327	Fatty acid synthase	1,00E-04	T2	T2 vs CTRL	P/A
				3,12E-06	T1	T1 vs CTRL	
PGBM_HUMAN	HSPG2	P98160	Basement membrane-specific heparan sulfate proteoglycan core protein	3,12E-06	T2	T2 vs CTRL	P/A
				6,59E-06	T1	T1 vs CTRL	
SY1C_HUMAN	IARS	P41252	Isoleucine tRNA ligase, cytoplasmic	6,59E-06	T2	T2 vs CTRL	P/A
				1,77E-04	T2	T2 vs CTRL	P/A
				1,77E-04	T2	T1 vs T2	
IF2B3_HUMAN	IGF2BP3	O00425	insulin-like growth factor 2 mRNA binding protein 3	6,01E-06	T2	T2 vs CTRL	P/A
ITA3_HUMAN	ITGA3	P26006	Integrin alpha-3	6,66E-06	T2	T2 vs CTRL	P/A
ITB3_HUMAN	ITGB3	P05106	Integrin beta-3	6,59E-06	T1	T1 vs CTRL	P/A
				6,59E-06	T2	T2 vs CTRL	P/A
IMB1_HUMAN	KPNB1	Q14974	Importin subunit beta-1	6,59E-06	T2	T2 vs CTRL	P/A
				9,99E-08	T2	T2 vs CTRL	P/A
LAMP2_HUMAN	LAMP2	P13473	Lysosome-associated membrane glycoprotein 2	2,86E-05	T1	T1 vs CTRL	P/A
MLEC_HUMAN	MLEC	Q14165	Malectin	2,86E-05	T2	T2 vs CTRL	P/A
				1,77E-04	T2	T2 vs CTRL	P/A
				1,77E-04	T2	T1 vs T2	
MX1_HUMAN	MX1	P20591	Interferon-induced GTP-binding protein Mx1	1,77E-04	T2	T2 vs CTRL	P/A
				1,77E-04	T2	T1 vs CTRL	P/A
				1,77E-04	T2	T1 vs T2	
PLD3_HUMAN	PLD3	Q8IV08	Phospholipase D3	2,86E-05	T1	T1 vs CTRL	P/A
				2,86E-05	T2	T2 vs CTRL	P/A
PSMD2_HUMAN	PSMD2	Q13200	26S proteasome non-ATPase regulatory subunit 2	1,77E-04	T2	T2 vs CTRL	P/A
				1,77E-04	T2	T1 vs T2	
RAB8A_HUMAN	RAB8A	P61006	Ras-related protein Rab-8A	8,91E-04	T1	T1 vs CTRL	P/A
				8,91E-04	T2	T2 vs CTRL	P/A
S10AC_HUMAN	S100A12	P80511	Protein S100-A12	1,77E-04	T2	T2 vs CTRL	P/A
				1,77E-04	T2	T1 vs T2	
SH3L3_HUMAN	SH3BGR13	Q9H299	SH3 domain-binding glutamic acid-rich-like protein 3	1,10E-06	T1	T1 vs CTRL	P/A
				1,10E-06	T2	T2 vs CTRL	P/A
LAT1_HUMAN	SLC7A5	Q01650	Large neutral amino acids transporter small subunit 1	2,31E-05	T2	T2 vs CTRL	P/A
				2,31E-05	T2	T1 vs T2	P/A
SMD3_HUMAN	SNRPD3	P62318	Small nuclear ribonucleoprotein Sm D3	9,99E-08	T1	T1 vs CTRL	P/A
RUXE_HUMAN	SNRPE	P62304	Small nuclear ribonucleoprotein E	5,90E-04	T1	T1 vs CTRL	P/A
SORT_HUMAN	SORT1	Q99523	Sortilin	5,90E-04	T2	T2 vs CTRL	P/A
				2,31E-05	T2	T2 vs CTRL	P/A
				2,31E-05	T2	T1 vs T2	
SSRA_HUMAN	SSR1	P43307	Translocon-associated protein subunit alpha	5,82E-05	T1	T1 vs CTRL	P/A
				5,82E-05	T2	T2 vs CTRL	P/A
TACD2_HUMAN	TACSTD2	P09758	Tumor-associated calcium signal transducer 2	1,77E-04	T2	T2 vs CTRL	P/A
				1,77E-04	T2	T1 vs T2	
VAC14_HUMAN	VAC14	Q08AM6	Protein VAC14 homolog	1,77E-04	T2	T2 vs CTRL	P/A
				7,02E-04	T2	T2 vs CTRL	P/A
VAMP8_HUMAN	VAMP8	Q9BV40	Vesicle-associated membrane protein 8	8,91E-04	T1	T1 vs CTRL	P/A
VPS35_HUMAN	VPS35	Q96QK1	Vacuolar protein sorting-associated protein 35	8,91E-04	T2	T2 vs CTRL	P/A
ANXA2_HUMAN	ANXA2	P07355	Annexin A2	n.a	n.a	n.a	Lab.
CD166_HUMAN	ALCAM	Q13740	CD166 antigen	n.a	n.a	n.a	Lab.
MMP9_HUMAN	MMP9	P14780	Matrix metalloproteinase-9	n.a	n.a	n.a	Lab.
PDIA1_HUMAN	PDIA1	P07237	Protein disulfide-isomerase	n.a	n.a	n.a	Lab.
TNR6_HUMAN	FASN	P25445	Tumor necrosis factor receptor superfamily member 6	n.a	n.a	n.a	Lab.

## 2.4 GO analysis to understand the most relevant biological processes of ELVs in UAs

For this part of the study, the label-free results obtained from the PEAKS analysis were used. A general overview of the steps followed for the analysis is depicted in Figure 42. Our results indicated that from the 1457 proteins identified in Control patients, only 624 proteins were present in four or more patients. Furthermore, in EC patients, from 1572 identified proteins in EC1 and 2030 proteins in EC2, only 715 and 932 were present in 4 or more patients, respectively. We used these lists of proteins as the more representative proteins for each condition (CTRL, EC1 and EC2 patients) to perform a GO biological process overrepresentation test. We found biological adhesion (30 proteins EC1 + 45 proteins EC2); mitosis (28 proteins EC1 + 39 proteins EC2); and exocytosis (21 proteins EC1 + 25 proteins EC2) as the most overrepresented processes in cancer. These processes included a total of 108 proteins (Figure 42, List 1).

Additionally, we found that from 715 and 932 proteins identified in EC1 and EC2 respectively, a total of 153 were exclusively present in both types of EC and were not present in any control patient (Figure 42, List 2). An overrepresentation test of this list of proteins indicated that metabolic processes such as translation (GO:0006412) and tRNA aminoacylation for protein translation (GO:0006418) were overrepresented with fold changes of 5.24 and 16.63 respectively. Other biological processes overrepresented were chromatin organization (GO:0006325, FC=6.44), cellular component organization (GO:0016043, FC=3.12), vesicle-mediated transport (GO:0016192, FC=3.23), and intracellular protein transport (GO:0006886 FC=3.05).

Finally, by comparing the list of proteins produced by both analysis, we obtained a reduced list of 22 proteins highly enriched in GO biological processes terms of biological adhesion and cellular process (cell cycle, cell proliferation, cell communication, cytokinesis, chromosome segregation and cellular component movement) (Figure 42 and Table 17).

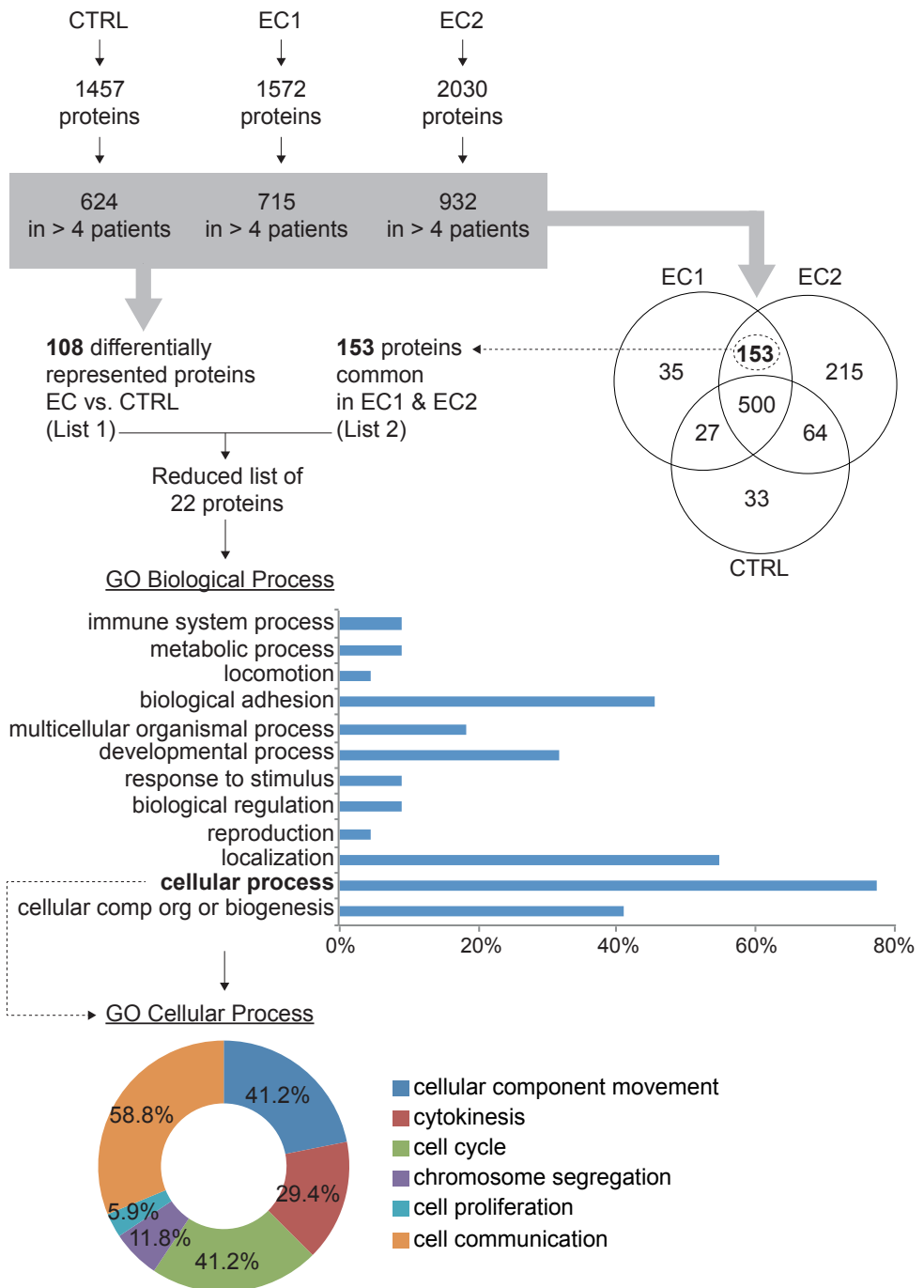


Figure 40. Workflow followed for the GO analysis.

**Table 17.** Reduced list of 22 proteins obtained by comparing the most representative proteins of each condition group and the 153 proteins found exclusively in cancer patients.

<b>Entry Name</b>	<b>UniProt Number</b>	<b>Protein Name</b>
AGRIN_HUMAN	O00468	Agrin
ARP3_HUMAN	P61158	Actin-related protein 3
COPB_HUMAN	P53618	Coatomer subunit beta
DYH9_HUMAN	Q9NYC9	Dynein heavy chain 9, axonemal
FBLN1_HUMAN	P23142	Fibulin-1
FCN1_HUMAN	O00602	Ficolin-1
ICAM1_HUMAN	P05362	Intercellular adhesion molecule 1
ITA3_HUMAN	P26006	Integrin alpha-3
ITAM_HUMAN	P11215	Integrin alpha-M
ITB2_HUMAN	P05107	Integrin beta-2
LAMC1_HUMAN	P11047	Laminin subunit gamma-1
LYN_HUMAN	P07948	Tyrosine-protein kinase Lyn
MYH14_HUMAN	Q7Z406	Myosin-14
MYO6_HUMAN	Q9UM54	Unconventional myosin-VI
MYOME_HUMAN	Q5VU43	Myomegalin
PLXB2_HUMAN	O15031	Plexin-B2
SC23A_HUMAN	Q15436	Protein transport protein Sec23A
SCAM2_HUMAN	O15127	Secretory carrier-associated membrane protein 2
SCAM3_HUMAN	O14828	Secretory carrier-associated membrane protein 3
SEPT9_HUMAN	Q9UHD8	Septin-9
TBB2A_HUMAN	Q13885	Tubulin beta-2A chain
VAMP8_HUMAN	Q9BV40	Vesicle-associated membrane protein 8



## RESULTS OF OBJECTIVE 3

### 3. Verification of candidates by targeted proteomics in UAs ELVs

Once we obtained the list of 54 candidates described in the previous section, we aimed to verify the potential of those candidates by using LC-SRM in a new and bigger cohort of patients. As in the discovery phase study, the ELVs fraction of UAs was the selected sample of analysis. In addition to evaluate the individual potential of each marker to diagnose EC and differentiate between histological subtypes, in here we also sought to generate diagnostic and prognostic models by combining different candidates. Moreover, we assessed the classification power of each candidate in the whole fluid fraction of UAs in an independent cohort. The summary of proteins monitored, detected and quantified in each cohort is presented in Table 21.

#### 3.1 Verification of protein biomarkers in UAs ELVs

A total of 107 patients were recruited for the verification phase (cohort B), divided in three groups: EC1 (n=45), EC2 (n=21) and CTRL (n=41). From those, ELVs from UAs were isolated and characterized as seen in previous sections (data not shown). The list of 54 protein candidate biomarkers generated in the discovery phase was verified by LC-SRM. To do so, two unique tryptic peptides were selected per protein, from which a total of 85 endogenous peptides were finally monitored by scheduled-SRM. However, we were able to only detect 69 of the 85 peptides corresponding to 51 proteins; three of the candidates (TNR6, CLH1 and PSB3) were not detected in any of the samples.

As a result of this SRM experiment, we observed that 43 out of the 48 (89.6%) potential diagnostic biomarkers (i.e. significant differences in abundance were observed between CTRL and EC) were also significant in this verification phase (adj.pvalue < 0.01), thus confirming their potential as individual diagnostic biomarkers (Table 19). A total of 29 out of the 45 quantified proteins presented high accuracy to individually discriminate between EC and CTRL cases (AUC

values higher than 0.75, highlighted in bold in Table 19). The 5 most significant individual biomarkers were AGRIN (AUC= 0.90, CI95: 0.85-0.96), TACD2 (AUC= 0.87, CI95: 0.81-0.94), SORT (AUC=0.86, CI95: 0.79-0.93), MVP (AUC= 0.86, CI95: 0.78-0.93) and FAS (AUC= 0.85, CI95: 0.78-0.92). Interestingly, two proteins that were not originally significant in this comparison, were verified as potential diagnostic biomarkers; those are PODXL and CLD6. While CLD6 resulted significant in CTRL vs. EC and in EC1 vs. EC2, PODXL was just verified as significant between CTRL vs. EC.

Regarding the potential prognostic biomarkers (i.e. comparison of EC1 vs. EC2) only 4 proteins out of the 15 (26,6%) candidates were found to be differentially expressed in the verification phase (adj. p-value < 0.01). The verified biomarkers were CLD6, IF2B3, PLD3 and MX1 (Table 18). Curiously, the protein BCAM, which was originally monitored for being significant in the comparison between EC and CTRL, was not verified as diagnostic biomarker but appeared significant in the comparison between EC1 and EC2. A total of 4 out of the 45 quantified proteins presented high accuracy to individually discriminate between EC1 and EC2 cases with AUC values higher than 0.75: CLD6 (AUC= 0.88, CI95: 0.76-1.00), BCAM (AUC= 0.87, CI95: 0.76-0.97), IF2B3 (AUC=0.80, IC95: 0.68-0.93) and PLD3 (AUC= 0.79, IC95: 0.66-0.93).

**Table 18.** Statistical inference results of prognostic biomarkers in ELVs. Only the 15 proteins considered for this comparison are shown. (\*) BCAM was not originally monitored for being significant in this comparison but was validated as prognostic marker. Proteins in red were not validated. "n.q", proteins not quantified in the verification phase. AUC values higher than 0.75 are highlighted in bold.

Entry Name	UniProt Number	FC	log2FC	adj.pvalue	AUC	IC95%
CLD6_HUMAN	P56747	4,06	2,02	9,63E-13	<b>0,88</b>	0,76 - 1,00
BCAM_HUMAN (*)	P50895	2,12	1,08	4,60E-08	<b>0,87</b>	0,76 - 0,97
IF2B3_HUMAN	O00425	3,27	1,71	5,28E-06	<b>0,80</b>	0,68 - 0,93
PLD3_HUMAN	Q8IV08	2,01	1,01	6,72E-05	<b>0,79</b>	0,66 - 0,93
MX1_HUMAN	P20591	3,49	1,80	7,91E-04	0,74	0,61 - 0,87
PODXL_HUMAN	O00592	0,40	-1,33	6,87E-02	0,29	0,16 - 0,42
SSRA_HUMAN	P43307	1,65	0,72	4,83E-01	0,62	0,47 - 0,77
ILF2_HUMAN	Q12905	1,54	0,62	4,83E-01	0,59	0,44 - 0,74
CD81_HUMAN	P60033	1,49	0,58	5,21E-01	0,61	0,46 - 0,76
VAC14_HUMAN	Q08AM6	1,34	0,42	5,33E-01	n.q	n.q
SMD3_HUMAN	P62318	1,35	0,43	5,52E-01	0,56	0,41 - 0,72
CD59_HUMAN	P13987	0,82	-0,28	7,40E-01	0,41	0,26 - 0,56
SH3L3_HUMAN	Q9H299	1,11	0,15	7,70E-01	0,50	0,34 - 0,67
RSSA_HUMAN	P08865	1,11	0,15	7,82E-01	0,51	0,36 - 0,67
RAB8A_HUMAN	P61006	0,96	-0,05	9,31E-01	0,49	0,32 - 0,66
PPIA_HUMAN	P62937	0,98	-0,03	9,62E-01	0,48	0,32 - 0,64

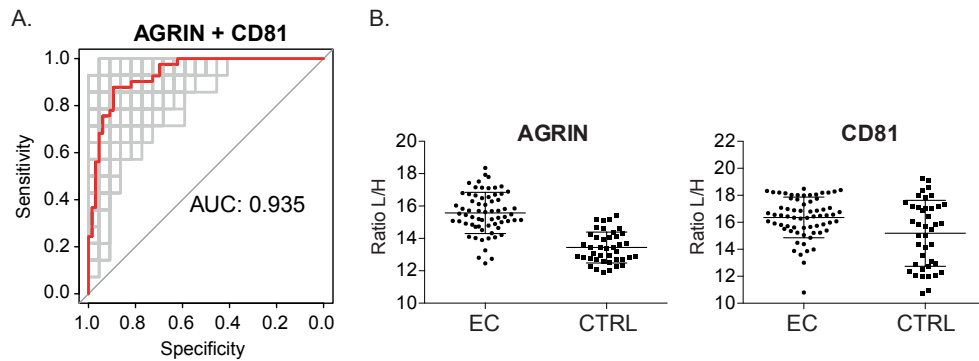
**Table 19.** Statistical inference results of diagnostic biomarkers in ELVs. *PODXL*, *CLD6* and *CD166* were not considered for this comparison. Proteins in red were not validated (adj.pvalue > 0.01). "n.q", proteins not quantified in the verification phase. AUC values higher than 0.75 are highlighted in bold.

Entry Name	UniProt Number	FC	log2FC	adj.pvalue	AUC	IC95%
AGRIN_HUMAN	O00468	4,46	2,16	1,25E-13	<b>0,90</b>	0,85 - 0,96
MVP_HUMAN	Q14764	7,50	2,91	3,01E-10	<b>0,86</b>	0,78 - 0,93
TACD2_HUMAN	P09758	4,76	2,25	3,01E-10	<b>0,87</b>	0,81 - 0,94
FAS_HUMAN	P49327	4,90	2,29	4,30E-10	<b>0,85</b>	0,78 - 0,92
SYIC_HUMAN	P41252	9,74	3,28	1,20E-09	<b>0,84</b>	0,77 - 0,91
VAMP8_HUMAN	Q9BV40	4,17	2,06	1,57E-09	<b>0,84</b>	0,76 - 0,92
SORT_HUMAN	Q99523	3,10	1,63	1,57E-09	<b>0,86</b>	0,79 - 0,93
LAT1_HUMAN	Q01650	5,65	2,50	4,51E-09	<b>0,84</b>	0,76 - 0,91
TERA_HUMAN	P55072	4,93	2,30	9,49E-09	<b>0,82</b>	0,74 - 0,90
RUVB1_HUMAN	Q9Y265	4,76	2,25	9,49E-09	<b>0,82</b>	0,74 - 0,90
RSSA_HUMAN	P08865	3,63	1,86	9,49E-09	<b>0,82</b>	0,74 - 0,90
SMD3_HUMAN	P62318	4,38	2,13	1,09E-08	<b>0,82</b>	0,75 - 0,90
ADA10_HUMAN	O14672	2,44	1,29	2,21E-08	<b>0,80</b>	0,71 - 0,89
RPL13A_HUMAN	P40429	10,98	3,46	3,01E-08	<b>0,80</b>	0,72 - 0,88
PGBM_HUMAN	P98160	3,95	1,98	3,05E-08	n.q	n.q
RL11_HUMAN	P62913	8,32	3,06	3,53E-08	<b>0,80</b>	0,71 - 0,88
IMB1_HUMAN	Q14974	3,27	1,71	3,53E-08	<b>0,80</b>	0,71 - 0,89
AGR2_HUMAN	O95994	5,84	2,55	5,10E-08	<b>0,79</b>	0,69 - 0,88
ITA3_HUMAN	P26006	2,26	1,18	5,10E-08	<b>0,81</b>	0,73 - 0,89
RUXE_HUMAN	P62304	3,03	1,60	7,34E-08	<b>0,78</b>	0,70 - 0,87
RL12_HUMAN	P30050	9,82	3,30	7,80E-08	<b>0,79</b>	0,71 - 0,88
RS16_HUMAN	P62249	9,04	3,18	1,36E-07	<b>0,79</b>	0,71 - 0,87
PSMD2_HUMAN	Q13200	5,09	2,35	3,02E-07	<b>0,79</b>	0,70 - 0,87
MX1_HUMAN	P20591	3,94	1,98	3,77E-07	<b>0,81</b>	0,73 - 0,89
VPS35_HUMAN	Q96QK1	2,00	1,00	7,86E-07	<b>0,78</b>	0,69 - 0,87
ILF2_HUMAN	Q12905	3,20	1,68	8,51E-07	<b>0,79</b>	0,70 - 0,88
PDIA1_HUMAN	P07237	2,65	1,41	9,50E-07	<b>0,77</b>	0,68 - 0,86
MMP9_HUMAN	P14780	1,97	0,98	7,67E-06	<b>0,79</b>	0,70 - 0,88
ANXA4_HUMAN	P09525	3,58	1,84	1,20E-05	0,72	0,61 - 0,82
RAB8A_HUMAN	P61006	3,18	1,67	1,20E-05	0,73	0,62 - 0,84
SH3L3_HUMAN	Q9H299	2,32	1,22	1,28E-05	0,72	0,62 - 0,83
RL29_HUMAN	P47914	7,00	2,81	1,55E-05	0,74	0,64 - 0,83
PLD3_HUMAN	Q8IV08	1,74	0,80	1,82E-05	<b>0,75</b>	0,65 - 0,85
PPIA_HUMAN	P62937	3,86	1,95	2,86E-05	0,71	0,60 - 0,82
ANXA2_HUMAN	P07355	3,49	1,80	3,66E-05	0,70	0,59 - 0,81
S10AC_HUMAN	P80511	7,75	2,95	5,40E-05	n.q	n.q
CD14_HUMAN	P08571	3,89	1,96	6,48E-05	n.q	n.q
SSRA_HUMAN	P43307	3,05	1,61	6,67E-05	0,72	0,61 - 0,82
LAMP2_HUMAN	P13473	2,16	1,11	7,78E-05	0,72	0,62 - 0,83
PODXL_HUMAN	O00592	2,86	1,52	2,61E-04	0,71	0,61 - 0,81
CLD6_HUMAN	P56747	1,83	0,87	2,61E-04	0,67	0,57 - 0,77
IF2B3_HUMAN	O00425	1,97	0,98	4,27E-04	<b>0,77</b>	0,68 - 0,86
CD59_HUMAN	P13987	2,80	1,49	6,17E-04	0,64	0,52 - 0,76
MLEC_HUMAN	Q14165	2,05	1,04	2,55E-03	0,68	0,57 - 0,78
H10_HUMAN	P07305	1,65	0,72	6,93E-03	n.q	n.q
CD166_HUMAN	Q13740	1,59	0,67	9,42E-03	0,66	0,55 - 0,77
<b>CD81_HUMAN</b>	P60033	2,01	1,00	1,15E-02	0,63	0,51 - 0,75
<b>AR6P1_HUMAN</b>	Q15041	2,47	1,31	3,43E-02	n.q	n.q
<b>BCAM_HUMAN</b>	P50895	1,24	0,31	3,78E-02	0,63	0,52 - 0,74
<b>VAC14_HUMAN</b>	Q08AM6	1,61	0,68	5,29E-02	n.q	n.q
<b>ITB3_HUMAN</b>	P05106	1,33	0,41	5,61E-02	0,61	0,50 - 0,73

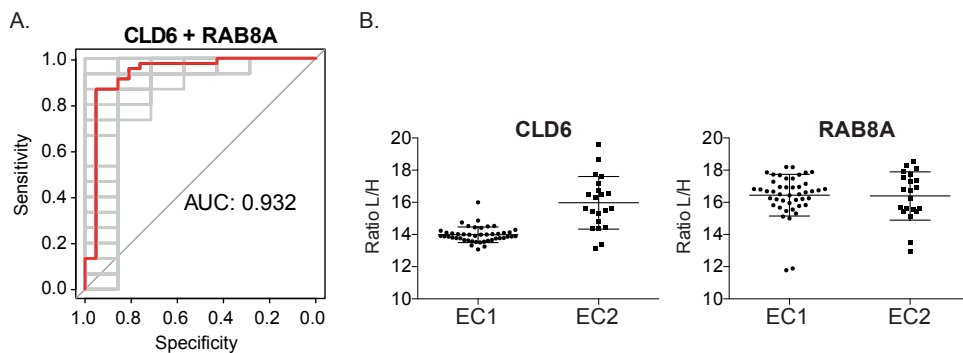
## 3.2 Development of predictors for EC diagnosis and prognosis in ELVs derived from UAs

Previous studies described that combining multiple proteins might improve the diagnostic performance over the use of single biomarkers, as single markers may not necessarily reflect the multifactorial nature of cancer (i.e. EC), neither have the necessary prediction power for patient classification<sup>216</sup>. Therefore, in order to compute a protein combination with a good diagnostic power, we fitted estimated abundances of all complete quantified proteins (45 candidates) into a logistic regression model evaluated in 2/3 of the dataset and checked the AUC value for each protein combination. As explained in the materials and methods section, this process was repeated 500 times and the most discriminative panel was selected based on the most repeated outcome. In order to minimize the number of proteins included in the predictors, whilst maximizing the sensitivity and specificity of combining different proteins, the model only included an additional protein to the predictor while increasing AUC values ( $\Delta\text{AUC} > 0.02$ ).

The diagnostic panel composed by AGRIN and CD81 was selected 192 times over the 500 repetitions, providing an AUC of 0.935 (CI95: 0.89–0.98). Curiously, the panel was composed by the best individual diagnostic biomarker, AGRIN (AUC= 0.90, CI95: 0.85–0.96), but also by CD81, which as individual biomarker presented a limited diagnostic potential (AUC= 0.63, CI95: 0.51–0.75). This result evidenced that the combination of two different proteins improved the ability to distinguish benign from EC cases (Figure 43), and that this combination is independent of the individual potential of each biomarker as presumably, the most significant biomarkers might provide redundant information.



**Figure 41.** Diagnostic performance of biomarkers in discriminating EC from CTRL patients in ELVs isolated from UAs. (A) ROC curve of EC versus CTRL patients of the 2-protein panel. (B) Box-plots showing the distribution of the light/heavy (L/H) ratios obtained by LC-SRM across the 66 EC and the 41 CTRL patients of the two proteins that compose the diagnostic panel.



**Figure 42.** Prognostic performance of biomarkers in classifying EC patients in EC1 or EC2, in ELVs isolated from UAs. (A) ROC curve of EC1 versus EC2 patients of the 2-protein panel. (B) Box-plots showing the distribution of the light/heavy (L/H) ratios obtained by LC-SRM across the 41 EC1 and the 21 EC2 patients of the two proteins that compose the prognostic panel.

Next, we also evaluated protein combinations that could correctly classify EC1 and EC2 tumors. Again, we found that the combination of two proteins (CLD6 and RAB8A) had a better performance (AUC= 0.93, CI95: 0.85-1.00) (Figure 44) than any of the three best individual prognosis biomarkers (CLD6, BCAM and IF2B3). The prognostic combination included RAB8A, which exhibited an AUC value of 0.49 (IC95: 0.32-0.66) as individual biomarker. Sensitivity, specificity and cut points of these proteins are detailed in Table 20.

ROC EC vs. CTRL (ELVs)					
Protein/s	Specificity	Sensitivity	AUC	Cut point	CI95%
<b>AGRIN + CD81</b>	0.88	0.89	0.94	0.49	0.89-0.98
<b>AGRIN</b>	0.90	0.80	0.90	0.70	0.85-0.96
<b>CD81</b>	0.46	0.86	0.63	0.56	0.51-0.75
<b>TACD2</b>	0.90	0.82	0.87	0.68	0.81-0.94
<b>MVP</b>	0.93	0.70	0.86	0.77	0.78-0.93
ROC EC1 vs. EC2 (ELVs)					
Protein/s	Specificity	Sensitivity	AUC	cut point	CI95%
<b>CLD6 + RAB8A</b>	0.87	0.95	0.93	0.19	0.85-1.00
<b>CLD6</b>	0.82	0.91	0.88	0.21	0.76-1.00
<b>RAB8A</b>	0.84	0.43	0.49	0.32	0.32-0.66
<b>BCAM</b>	0.84	0.81	0.87	0.24	0.76-0.97
<b>IF2B3</b>	0.84	0.81	0.80	0.24	0.68-0.93

**Table 20.** Summary table of the results of the logistic regression model for the ROC analysis. Sensitivity, specificity, AUC, cut point and the upper and lower limits of the confidence interval (CI95%) of the most significant individual candidates and the best combination panel are shown.

**Table 21.** Summary of proteins monitored, detected and quantified in each cohort from the discovery of the candidates to the verification in ELVs and UAs

Entry Name	Gene Name	UniProt Number	ELVs of cohort A (n=30)			ELVs of cohort B (n=107)			UAs of cohort C (n=67)		
			Significant in Comparison of the Discovery	Monitored by LC-SRM	Detected by LC-SRM	Quantified for AUC estimation	Monitored by LC-SRM	Detected by LC-SRM	Quantified for AUC estimation	Monitored by LC-SRM	Detected by LC-SRM
ADA10_HUMAN	ADAM10	O14672	EC1_CTRL & EC2_CTRL	✓	✓	✓	✓	✓	✗	✗	
AGR2_HUMAN	AGR2	O95994	EC1_CTRL & EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
AGRIN_HUMAN	AGRN	O00468	EC1_CTRL & EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
ANXA4_HUMAN	ANXA4	P09525	EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
AR6P1_HUMAN	ARL6IP1	Q15041	EC1_CTRL & EC2_CTRL	✓	✓	✗	✓	✓	✗	✗	
BCAM_HUMAN	BCAM	P50895	EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
CD14_HUMAN	CD14	P08571	EC1_CTRL & EC2_CTRL	✓	✓	✗	✗	-	-	-	
CD59_HUMAN	CD59	P13987	EC1_CTRL & EC1_EC2	✓	✓	✓	✓	✓	✓	✓	
CD81_HUMAN	CD81	P60033	EC1_CTRL & EC1_EC2	✓	✓	✓	✓	✓	✗	✗	
CLD6_HUMAN	CLDN6	P56747	EC1_EC2	✓	✓	✓	✓	✓	✗	✗	
CLH1_HUMAN	CLTC	Q00610	EC2_CTRL	✓	✗	✗	✗	-	-	-	
FAS_HUMAN	FASN	P49327	EC1_CTRL & EC2_CTRL	✓	✓	✓	✓	✓	✗	✗	
H10_HUMAN	H1FO	P07305	EC1_CTRL & EC2_CTRL	✓	✓	✗	✓	✓	✓	✓	
IF2B3_HUMAN	IGF2BP3	O00425	EC2_CTR & EC1_EC2	✓	✓	✓	✓	✓	✗	✗	
ILF2_HUMAN	ILF2	Q12905	EC1_CTRL & EC1_EC2	✓	✓	✓	✓	✓	✓	✓	
IMB1_HUMAN	KPNB1	Q14974	EC1_CTRL & EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
ITA3_HUMAN	ITGA3	P26006	EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
ITB3_HUMAN	ITGB3	P05106	EC2_CTRL	✓	✓	✓	✓	✓	✗	✗	
LAMP2_HUMAN	LAMP2	P13473	EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
LAT1_HUMAN	SLC7A5	Q01650	EC1_CTRL & EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
MLEC_HUMAN	MLEC	Q14165	EC1_CTRL & EC2_CTRL	✓	✓	✓	✓	✓	✗	✗	
MVP_HUMAN	MVP	Q14764	EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
MX1_HUMAN	MX1	P20591	EC2_CTR & EC1_EC2	✓	✓	✓	✓	✓	✓	✓	
PGBM_HUMAN	HSPG2	P98160	EC1_CTRL & EC2_CTRL	✓	✓	✗	✗	-	-	-	
PLD3_HUMAN	PLD3	Q8IV08	EC2_CTR & EC1_EC2	✓	✓	✓	✓	✓	✓	✓	
PODXL_HUMAN	PODXL	O00592	EC1_EC2	✓	✓	✓	✓	✓	✓	✓	
PPIA_HUMAN	PPIA	P62937	EC1_CTRL & EC1_EC2	✓	✓	✓	✓	✓	✓	✓	
PSB3_HUMAN	PSMB3	P49720	EC2_CTRL	✓	✗	✗	✗	-	-	-	
PSMD2_HUMAN	PSMD2	Q13200	EC1_CTRL & EC2_CTRL	✓	✓	✓	✓	✓	✗	✗	
RAB8A_HUMAN	RAB8A	P61006	EC2_CTR & EC1_EC2	✓	✓	✓	✓	✓	✓	✓	
RL11_HUMAN	RPL11	P62913	EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
RL12_HUMAN	RPL12	P30050	EC1_CTRL & EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
RL13A_HUMAN	RPL13A	P40429	EC1_CTRL & EC1_EC2	✓	✓	✓	✓	✓	✓	✓	
RL29_HUMAN	RPL29	P47914	EC1_CTRL & EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
RS16_HUMAN	RPS16	P62249	EC1_CTRL	✓	✓	✓	✓	✓	✓	✓	
RSSA_HUMAN	RPSA	P08865	EC2_CTR & EC1_EC2	✓	✓	✓	✓	✓	✓	✓	
RUVB1_HUMAN	RUVBL1	Q9Y265	EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
RUXE_HUMAN	SNRPE	P62304	EC1_CTRL	✓	✓	✓	✓	✓	✓	✓	
S10AC_HUMAN	S100A12	P80511	EC1_CTRL & EC2_CTRL	✓	✓	✗	✗	-	-	-	
SH3L3_HUMAN	SH3BGR13	Q9H299	EC2_CTR & EC1_EC2	✓	✓	✓	✓	✓	✓	✓	
SMD3_HUMAN	SNRPD3	P62318	EC2_CTR & EC1_EC2	✓	✓	✓	✓	✓	✓	✓	
SORT_HUMAN	SORT1	Q99523	EC1_CTRL & EC2_CTRL	✓	✓	✓	✓	✓	✗	✗	
SSRA_HUMAN	SSR1	P43307	EC2_CTR & EC1_EC2	✓	✓	✓	✓	✓	✗	✗	
SYIC_HUMAN	IARS	P41252	EC1_CTRL & EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
TACD2_HUMAN	TACSTD2	P09758	EC1_CTRL & EC2_CTRL	✓	✓	✓	✓	✓	✗	✗	
TERA_HUMAN	VCP	P55072	EC1_CTRL	✓	✓	✓	✓	✓	✓	✓	
VAC14_HUMAN	VAC14	Q08AM6	EC2_CTR & EC1_EC2	✓	✓	✗	✓	✓	✗	✗	
VAMP8_HUMAN	VAMP8	Q9BV40	EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
VPS35_HUMAN	VPS35	Q96QK1	EC1_CTRL & EC2_CTRL	✓	✓	✓	✓	✓	✓	✓	
ANXA2_HUMAN*	ANXA2	P07355	n.a	✓	✓	✓	✓	✓	✓	✓	
CD166_HUMAN*	ALCAM	Q13740	n.a	✓	✓	✓	✓	✓	✗	✗	
CLIC1_HUMAN*	CLIC1	O00299	n.a	✗	-	-	✓	✓	✓	✓	
KPYM_HUMAN*	PKM	P14618	n.a	✗	-	-	✓	✓	✓	✓	
MMP9_HUMAN*	MMP9	P14780	n.a	✓	✓	-	✓	✓	✓	✓	
PDIA1_HUMAN*	PDIA1	P07237	n.a	✓	✓	✓	✓	✓	✓	✓	
PIGR_HUMAN*	PIGR	P01833	n.a	✗	-	-	✓	✓	✓	✓	
TNR6_HUMAN*	FASN	P25445	n.a	✓	✗	✗	✓	✓	-	-	
				54	51	45	51	37	37		

## 4. Verification of candidates by targeted proteomics in the whole fluid of UAs

In order to understand whether it would be feasible to transfer the use of these biomarkers to a sample that is easier to access and does not require the isolation of the extracellular vesicles, we aimed to study the biomarkers identified in exosomes in the whole fluid fraction of UAs. For that, a total of 67 patients (cohort C) were selected and divided into three groups: EC1 (n=22), EC2 (n=20) and CTRL (n=25).

For this part of the study, a total of 51 proteins were analyzed (Table 21). Among those, 48 proteins were also monitored in the previous verification phase, and 3 proteins (KPYM, PIG3 and CLIC1) were newly incorporated as they have been recently described as biomarkers for EC in uterine aspirates<sup>217</sup>. Two unique tryptic peptides were selected per protein, from which a total of 75 endogenous peptides were finally monitored by scheduled-SRM. A total of 37 proteins were only detected and subsequently quantified for AUC estimation.

We evaluated the results obtained in this verification study in comparison to the one previously performed, in order to understand the feasibility to translate ELV's based biomarkers to the whole fluid of UAs. First, we observed that the detectability of our ELVs biomarkers in the whole fluid of UAs was very limited compared to the detectability when ELVs were analyzed; out of the 51 biomarkers detected in ELVs, only 34 (66,7%) were detected in UAs. A total of 13 out of the 44 (29.5%) proteins that were discovered in ELVs and monitored in UAs, were not detected when analyzed in the whole fluid of UAs.

Then, we compared the 10 best individual candidates based on AUC values for the comparison between EC and CTRL in both matrices, ELVs and UAs (Table 22). The 5 proteins with the highest AUC values in ELVs appeared also among the top-10 in UAs (AGRIN, MVP, VAMP8, SYIC and LAT1). The AUC values obtained in the different matrices are very similar, and the biggest difference is of 0.04 points (AGRIN  $AUC_{ELVs} = 0.90$  and  $AUC_{UAs} = 0.86$ ). However, the other 50% of proteins did not match between the UAs and ELVs lists. The best individual biomarker in



UAs was MMP9 (AUC= 0.91, CI95: 0.84-0.99). MMP9 does not appear in the ELVs top-10 list, which is reasonable because this protein was not identified in ELVs. Actually, this biomarker was included in the candidate's list based on previous results of our lab.

**Table 22.** Top-10 candidates based on AUC values obtained in ELVs and UAs for the comparison between EC and CTRL samples. The common proteins in both matrices are highlighted in blue.

AUCs EC vs. CTRL (ELVs)				AUCs EC vs. CTRL (UAs)			
Protein	AUC	95% CI		Protein	AUC	95% CI	
		Lower limit	Upper limit			Lower limit	Upper limit
AGRIN_HUMAN	0,90	0,85	0,96	MMP9_HUMAN	0,91	0,84	0,99
MVP_HUMAN	0,86	0,78	0,93	MVP_HUMAN	0,88	0,81	0,96
VAMP8_HUMAN	0,84	0,76	0,92	AGRIN_HUMAN	0,86	0,77	0,96
SYIC_HUMAN	0,84	0,77	0,91	MX1_HUMAN	0,86	0,77	0,95
LAT1_HUMAN	0,84	0,76	0,91	VAMP8_HUMAN	0,86	0,77	0,94
SMD3_HUMAN	0,82	0,75	0,90	SYIC_HUMAN	0,85	0,76	0,95
RSSA_HUMAN	0,82	0,74	0,90	IMB1_HUMAN	0,85	0,76	0,94
TERA_HUMAN	0,82	0,74	0,90	ILF2_HUMAN	0,84	0,74	0,94
RUVB1_HUMAN	0,82	0,74	0,90	RAB8A_HUMAN	0,84	0,74	0,94
ITA3_HUMAN	0,81	0,73	0,89	LAT1_HUMAN	0,83	0,72	0,94

Interestingly, the AUC values of the 10 best individual candidates for the comparison between EC1 and EC2 were higher in ELVs than in UAs. There are 5 proteins that appeared in the top-10 list proteins of both matrices (BCAM, PLD3, MX1, ITA3 and LAMP2). Among them, the highest differences are of 0.14 points (ITA3  $AUC_{ELVs} = 0.67$ ,  $AUC_{UAs} = 0.53$  and LAMP2  $AUC_{ELVs} = 0.65$ ,  $AUC_{UAs} = 0.51$ ).

**Table 23.** Top-10 candidates based on AUC values obtained in ELVs and UAs for the comparison between EC1 and EC2 samples. The common proteins in both matrices are highlighted in blue.

AUCs EC1 vs. EC2 (ELVs)				AUCs EC1 vs. EC2 (UAs)			
Protein	AUC	95% CI		Protein	AUC	95% CI	
		Lower limit	Upper limit			Lower limit	Upper limit
CLD6_HUMAN	0,88	0,76	1,00	BCAM_HUMAN	0,83	0,71	0,96
BCAM_HUMAN	0,87	0,76	0,97	PLD3_HUMAN	0,72	0,56	0,87
IF2B3_HUMAN	0,80	0,68	0,93	RUXE_HUMAN	0,66	0,49	0,83
PLD3_HUMAN	0,79	0,66	0,93	MX1_HUMAN	0,65	0,48	0,82
TACD2_HUMAN	0,75	0,61	0,88	ITA3_HUMAN	0,53	0,35	0,70
MX1_HUMAN	0,74	0,61	0,87	ILF2_HUMAN	0,53	0,34	0,71
ITA3_HUMAN	0,67	0,52	0,82	SMD3_HUMAN	0,52	0,34	0,70
ITB3_HUMAN	0,66	0,52	0,81	LAMP2_HUMAN	0,51	0,33	0,69
LAMP2_HUMAN	0,65	0,50	0,79	ANXA4_HUMAN	0,50	0,32	0,68
VAMP8_HUMAN	0,63	0,48	0,79	RSSA_HUMAN	0,49	0,30	0,67

Although some biomarkers described for ELVs have potential to be used in UAs samples, these should be further analyzed in future studies, as the results are not 100% transferable between both matrices. Likewise, it should be assessed to what extent the biomarker panels described for ELVs maintain a good performance when analyzed in UA samples.



# DISCUSSION

---



Endometrial cancer is the fourth most common cancer in women in developed countries with 61,380 new cases and 10,920 deaths estimated for 2017. Since 2004, its incidence and mortality rates are increasing annually worldwide. There are about 30% of EC patients diagnosed at advanced stages of the disease, presenting a bad prognosis and a drastic decrease in the 5-years survival rate<sup>4</sup>. Hence, there is a clear clinical need to develop improved methods for early EC detection.

Despite the fact that early detection of EC is associated to the presence of AUB in about a 93% of the cases, this symptom is not exclusive for EC and other benign disorders (i.e. myomas, polyps and hyperplasias) generate a similar symptomatology<sup>14</sup>. The current diagnostic process to discriminate between benign and EC patients is laborious and consists of a first pelvic examination and a transvaginal ultrasonography. Results from pelvic examination are frequently normal, especially in the early stages of the disease, thus limiting the possibility of an early EC diagnose. Changes in size, shape or consistency of the uterus and/or its surrounding supporting structures may exist when the disease is more advanced. Transvaginal ultrasonography, which is based on the measurement of the thickness of the endometrium, is highly sensitive but lacks specificity to distinguish whether a thicker endometrium is caused by a benign or a malignant process. Therefore, a confirmatory histopathological examination of an endometrial biopsy is always needed. Uterine aspirates (UAs) are obtained by a minimally invasive procedure by which the endometrial fluid and cells are aspirated from the uterine cavity using a Cornier pipelle. Diagnosis is obtained by the observation of abnormal cells present in the sample. Even though this is the preferred way to obtain an endometrial biopsy, it has been reported an inadequate sampling rate (sampling not possible for technical reasons) of 12% in post-menopausal and 15% in pre-menopausal woman, and a diagnostic failure rate (not enough tissue obtained for a pathologic diagnosis) of 22% in post- and 8% in pre-menopausal woman<sup>69,218</sup>. Moreover, the accuracy to determine the histological subtype of the tumor using an uterine aspirate is limited; many authors have described up to 20 % of discrepancy in comparison with the clinical staging of the tumor<sup>219</sup>. To overcome these limitations, a biopsy guided by hysteroscopy is required, although this technique is more invasive and increases the risk of uterine perforation and hemorrhage among other complications that, altogether result in elevated derived

sanitary costs<sup>70</sup>. For these reasons, **the present study focused on the improvement of the use of uterine aspirates for the early detection of EC by the identification and verification of new protein biomarkers.**

To improve EC diagnosis, the use of UAs meet some features that enhance the value of this biofluid as a promising source of biomarkers for screening, diagnosis and monitoring of pathologies related to the female genital tract: (i) sampling of UAs is minimally invasive and it is feasible to retrieve between 50 and 1000  $\mu$ L of sample from a patient, which is sufficient to extract material for molecular studies<sup>133,135</sup>; (ii) UAs are in direct contact with the uterine cavity, representing faithfully its complex environment. UA's composition derives from secretions and cells flaking from a variety of surrounding tissues, from the luminal epithelium and glands, from proteins selectively transudated from blood, and likely contributors from tubal fluid<sup>135</sup>. When an EC is present, UAs also contained EC cells, and thus, (iii) UAs exhibit molecular alterations present in EC and are useful to study the EC tumor heterogeneity<sup>118</sup>.

Nowadays, molecular screening technologies allow the consecution of high-throughput discovery and validation studies on complex samples, such as the UAs, for the development of non-invasive tests. UAs have been used for the search of potential biomarkers to easily diagnose benign gynecological diseases such as endometriosis<sup>134</sup>. Furthermore, malignant diseases, such as gynecological cancers, have also benefit from the identification of markers in uterine aspirations<sup>220,221</sup>. Following a transcriptomic approach, others and our lab previously developed<sup>222</sup> and clinically validated<sup>132</sup> a molecular test, which increased the efficacy, sensitivity and negative predictive value of UAs-based diagnosis for EC. Moreover, genomic approaches have been used in UAs to correctly classify EC patients based on the detection of clonality and genetic alterations<sup>223</sup>. Although research in this field appears to be promising for transcriptomic and genomic approaches, untargeted proteomic-based studies on UAs present serious limitations. UAs highly express abundant proteins coming from plasma, such as albumin and gamma globulins, along with hemoglobin; and those mask the detection and analysis of less abundant proteins, which are generally the most promising candidates for biomarkers discovery<sup>133,224</sup>. To solve this dynamic-range problem, but also to expand research in the field of biomarkers

discovery for gynecological pathologies, exosome-like vesicles (ELV) arise as a promising source of biomarkers.

ELVs are small membrane vesicles, ranging from 20 to 200 nm, which are released by multivesicular bodies (MVBs) fusing with the cell membrane<sup>140,225</sup>. ELVs are capable of transferring information from the cell of origin (mostly proteins, lipids, nucleic acids and sugars) to other cells, thereby influencing the recipient cell function<sup>137,226–229</sup>. To do that, cells release ELVs to circulation or proximal body fluids, and consequently, ELVs have been described in blood<sup>230</sup>, urine<sup>231</sup>, saliva, and breast milk<sup>232</sup> among other body fluids. This feature has fostered the field of biomarker discovery for many diseases.

**Taking all this into account, we first aimed to confirm the existence of ELVs in UAs. Next, we deeply profiled the proteome of UA's ELVs by a SILAC-based shotgun proteomic approach that was useful to identify potential candidate biomarkers of EC. Those biomarkers were finally verified in a larger and independent cohort of 107 patients by targeted proteomics (LC-SRM). This study permitted to develop predictors that showed great sensitivity and specificity for improving the diagnosis and prognosis of EC patients.**

In the first place, we demonstrated here that ELVs exist in the fluid fraction of UAs by comparing three protocols of isolation, all of them based on ultracentrifugation, as this has been the method of choice for concentrating and isolating ELVs in several body fluids<sup>233</sup>. Moreover, we carried out an extensive characterization describing their morphology, size and enrichment in well-known ELV markers. When comparing the Standard, Filtration, and Sucrose protocols, we observed that all of them were capable of isolating ELVs; but in particular, the Standard protocol permitted not only a higher recovery of ELVs, but also a higher enrichment in tetraspanins, thus confirming the purity of the isolated vesicles. Furthermore, this protocol was the simplest, most reproducible and less costly protocol investigated here. In addition, sonication was applied to successfully disrupt ELVs membranes to improve protein yield for subsequent proteomic analysis, as it is known that, besides tetraspanin-enriched membrane domains, ELVs are also enriched in proteins that associate with detergent-resistant lipid-protein rafts<sup>136</sup>. Regarding this,



a treatment to clean up ELVs membranes from extraneous adhered proteins could have been tested. Trypsin is often used to break protein interactions; this property could be applied to analyze those proteins specifically contained in ELVs. However, considering that the main ELVs markers and possibly other proteins of interest are transmembrane structures, this digestion could affect the extracellular domains compromising protein structure, function and interaction with other proteins.

At the time of presenting this thesis, a wide range of different uterine vesicles collected by a variety of procedures is described in the literature<sup>134,234–237</sup>. Vilella et al. proved that ELVs isolated from endometrial fluids are certainly secreted by the endometrial epithelium cells, and consequently, their content may reflect the physiologic state of the uterine cavity. In agreement with our findings, these results promote the use of ELVs in UAs to search for those alterations that may originate from anomalous cells in the female genital tract, as the same rationale has been performed in other body fluids, such as bronchoalveolar lavage fluid of asthmatic patients<sup>238</sup> and urine of prostate cancer patients<sup>239</sup>.

In conclusion for Objective 1, we confirmed the existence of exosome-like vesicles in the fluid fraction of uterine aspirates. They were successfully isolated by differential centrifugation giving sufficient proteomic material for further analyses. The Standard protocol was the best performing procedure since the other two tested protocols did not ameliorate neither yield nor purity of exosome-like vesicles. Certainly, this study contributes to standardize protocols and opens the door to conduct reliable and reproducible comparative studies using ELVs isolated from UAs to foster the field of biomarker research in gynecology.

As reviewed by Mittal et al.<sup>240</sup>, many studies have reported to find proteins with diagnostic potential<sup>241</sup> (i.e. CA-125<sup>242</sup>, CA72-4<sup>243</sup>, sFas<sup>244</sup>, and HE4<sup>245</sup>), prognostic potential<sup>246</sup> (i.e. L1CAM<sup>247</sup>, COX-2, Survivin, and c-erbB2<sup>248</sup>) and/or predictive value for EC<sup>249</sup>, however, none of them have had an impact in the daily clinical practice yet. Some authors attribute this to the use of a single proteomics approach, which is not enough to achieve a deep understanding of the protein function and is not able to discard false-negative and false-positive results<sup>250</sup>. Probably, the lack of prospective multi-centric validations or the unfeasibility to

reach optimal values of sensitivity and/or specificity contributes to this poor translation into clinical use. Also, this might be explained by the fact that these studies focus mostly on tissue, few on serum or plasma samples<sup>251-254</sup>, and even less on proximal body fluids. In serum, cancer-derived proteins are estimated to be 10 million times less abundant than common high abundant proteins secreted by normal cells<sup>240</sup>. Consequently, it is challenging to correctly identify and quantify tumor-associated proteins from the whole serum proteome. As stated earlier, the use of UAs and exosomes derived from uterine fluids might overcome these limitations.

So far, proteomic research performed in UAs has focused on embryo implantation, benign gynecological conditions or fertility, but not on endometrial cancer<sup>133,135,234</sup>. On the other side, Nguyen et al. have recently reviewed the works done on extracellular vesicles in the intrauterine environment<sup>255</sup>. Of note, none of the publications cited in that review explore the content of ELVs isolated from UAs to search EC diagnostic and/or prognostic biomarkers. They summarize that EVs have been identified in the uterine fluid secreted during the estrous/menstrual cycles of different species (i.e. sheep<sup>256,257</sup>, mice<sup>258</sup> and humans<sup>259</sup>). Other studies have proven that EVs are released from endometrial epithelial cells, suggesting a luminal/epithelial origin for EVs isolated from uterine fluids for the study of endometriosis<sup>260,261</sup>. The population of extracellular vesicles derived from the endometrium is varied and includes a mixture of small EVs, MVs and exosomes enclosing different proteins. Since the population is heterogeneous, it is probable that some exosome preparations obtained by ultracentrifugation from uterine fluids contain a small proportion of other types of EVs that could have co-precipitated. Nevertheless, and in line with our results, most of the studies on EVs isolated from uterine samples showed vesicles in the size range of exosomes and expressed the common markers (i.e. CD9, CD63, CD81, Alix, TSG101 and HSP70). A work done by Greening et al.<sup>262</sup> assert that the protein cargo of endometrial exosomes is regulated in a cyclical manner by estrogen and progesterone. This affirmation is the result of *in vitro* studies only, but it should be considered in order to study the behavior of the EC diagnostic and prognostic biomarkers described in this thesis in a premenopausal population, as their expression might vary according to the menstrual cycle. The work of this thesis has mainly focused on postmenopausal

women because most of the EC cases occur in this period, and only 14% of the cases affect premenopausal women. Only in the verification phase on UAs, we included a 13,4% of premenopausal women.

Working with UAs is a double-edged sword; on the one hand, it is more representative of the tumor heterogeneity than a tissue biopsy obtained from a single tumor area but, on the other hand, UA is itself a very heterogeneous sample regarding the volume obtained from each patient and the proportion of blood contamination. As a direct consequence of this, proteomic samples from ELVs isolated from UAs lack homogeneity among study groups and finding consistent biomarkers is challenging. For these reasons, we proposed an approach tackling the identification of a broad number of proteins (in-gel) in a semi-quantitative manner (super-SILAC).

The use of the Super-SILAC approach offers advantages over the classical SILAC or other chemical labeling techniques<sup>263</sup>. The use of the SILAC-labeled sample as a reference increases the applicability to samples that cannot be metabolically labeled, such as tissue samples, primary cells and body fluids. Moreover, the super-SILAC mix can be prepared in advance, stored for long periods and combined in appropriate ratios with the sample just before the preparation for MS analysis. Importantly, this approach allows for the comparison of multiple samples and the determination of a robust relative quantification.

For the Discovery phase of the present study, we took advantage of the benefits of using a super-SILAC mix composed of ELVs derived from EC cell lines as an internal standard (IS) for quantitative shotgun proteomics. The AN3CA, KLE and RL95.2 cell lines were selected after their proteomic profiling to be sure that we chose those cell lines that represented more accurately the proteome of UAs ELVs, without introducing redundant information. The IS developed covered 47% of “the patient proteome” (which included the UA ELV’s proteome of a control and an EC patient) according to the pilot study, thus enabling a fair use of this semi-quantitative approach to unveil biomarkers among EC and non-EC patients. This percentage was sustained in our discovery phase, as we were able to quantify 920 of the 2138 identified proteins. Moreover, the fact that our IS is composed of proteins contained in ELVs isolated from EC cancer cell lines increases the

chances of the quantifiable proteins to be related to cancer processes and to truly be ELVs-derived proteins. Also, it prevents from misleading results due to the quantification of possible contaminant proteins, such as those coming from blood.

Besides the use of the super-SILAC approach, another strength of our study is the sample fractionation of the discovery phase in order to increase the number of identified proteins. Identification of proteins from polyacrylamide gels offers a number of important advantages compared to gel-free approaches. Working with definite, molecular weight-separated proteins bands increases the dynamic range of analysis of protein mixtures because peptides produced by in-gel tryptic cleavage of each band are analyzed in separate experiments. Thus, proteins that would normally be undetectable in-solution due to high-abundance proteins of a specific molecular weight are accurately detected and analyzed. As stated in the technical paper of Matthias Mann group<sup>264</sup>, for complex mixtures analysis, spreading out the proteome over 10-20 gel slices dramatically increases the depth of analysis, and hence the number of identified proteins. At the same time, gel electrophoresis removes low molecular weight impurities, including detergents and buffer components, which are often detrimental for mass spectrometry analysis. According to this, we compared the in-gel and the in-solution digestion protocols in a pilot study and identified around double number of proteins with the in-gel procedure. Nevertheless, this approach also presents some drawbacks. Much higher concentration of the enzyme is required compared to in-solution digestion, which usually results in a significant background of autolysis products. Also, handling the excised bands increase the risk of contaminating the sample with keratins and might enhance chemical noise in analyzed samples. In our particular case, cutting each of the 30 lanes (each lane corresponding to a patient's sample spiked with the IS) in 10 bands resulted in 300 in-gel digestions and the corresponding 300 injections and runs in the mass spectrometer, which was very laborious, but also worthy as it allowed the identification and quantification of 2,138 and 920 proteins, respectively. It is worth highlighting that the number of proteins identified in our study is, as compared to the numbers reported in a review on mass spectrometry of extracellular vesicles<sup>207</sup>, in line with the highest reported to date (Table 24). From the 920 quantified proteins, 325 were considered for the analysis of differential expression for being present in at least 4 patients per group;

around 97% of these proteins were found on the ExoCarta database, confirming that we successfully enriched our samples in ELVs. Moreover, from the list of 152 differentially expressed proteins, 42 proteins (27,6%) were found among the ExoCarta's list of the top 100 proteins that are often identified in exosomes. These commonly identified proteins include some of those involved in membrane transport and fusion (i.e. small GTPase Rab proteins, annexins) and MVB biogenesis (i.e. clathrins, Alix/PDCD6IP, TSG101), as well as chaperones (i.e. HSPA8) and adhesion molecules (i.e. MFGE8 and tetraspanins such as CD63, CD81 and CD9). It is worth mentioning that while immunoblotting failed to detect significant amounts of GRP78 (HSPA5), which is a protein from the endoplasmic reticulum and is commonly used as a marker of cell contamination, MS identified it in all the patients of the discovery phase. Consequently, although our samples are enriched in ELVs, we assume there is a certain percentage of cellular contamination.

Interestingly, a high representation of ribosome proteins (RPLs and RPSs), translation elongation factors (EEFs), and histones were found among the 152 differentially expressed proteins and are also described in the ExoCarta database. The high representation of RPLs and RPSs could be explained by several reasons. The more plausible explanation is related to the fact that cancer cells display a number of abnormal properties to maintain their unrepressed growth and proliferation<sup>265</sup>. In this context, ribosome biogenesis and protein synthesis are critical cellular processes required for sustained cancer cell growth. Mutations in ribosomal protein genes have been found in different types of cancer including endometrial cancer<sup>266</sup>. In addition, RPs hold extra-ribosomal functions and some of them are involved in processes such as DNA repair, transcription, RNA processing and apoptosis<sup>267-269</sup>. Interestingly, one of our validated RPs (RP11) that was found differentially expressed between EC2 and CTRL patients is involved in a mechanism that leads to p53 activation<sup>267</sup>. Another possible explanation is the presumably presence of macrophage-derived ELVs in our preparations. Macrophages are a population of immune cells with a high degree of phenotypic and functional heterogeneity in the tumor microenvironment<sup>270</sup>. Under stimulation (with M-CFS, CCL2, IL-4, IL-10 and TGF- $\beta$ ), macrophages are recruited into the tumor local tissue and transformed into a population with tumor-supportive

properties named tumor-associated macrophages (TAMs)<sup>271,272</sup>. It has been reported that macrophages can regulate recipient cells by releasing exosomes<sup>273</sup> and that their protein cargo is mainly involved in RNA processing and proteolysis functions along with ribosome biogenesis. Taking all this into account, we could hypothesize that UAs contain a heterogeneous ELVs population, released by a variety of cell types and, among them, macrophages.

The presence of histones in exosomes was historically considered a contamination of apoptotic bodies<sup>274</sup>. However, histones are present in both cytoplasmic and nuclear pools<sup>275</sup> and are frequently described in exosome preparations<sup>274,276,277</sup>. In fact, a direct association of miRNA and histones has been reported for breast cancer cell line derived exosomes<sup>278</sup> and histone H3 modification was suggested to be essential for exosome release<sup>279</sup>.

Notwithstanding the foregoing mentioned advantages of our super-SILAC approach, we found interesting to explore the valuable information produced among the non-quantified proteins. Thus, to complement and ameliorate our study, we also analyzed the label-free data. The proteomic profiling of the ELVs isolated for the discovery phase revealed and confirmed that these vesicles are enriched in proteins related to important processes for cancer development. Biological processes such as adhesion, mitosis and exocytosis were overrepresented in tumor as compared to control samples. The overrepresentation test of proteins found exclusively on tumor samples unveiled that our ELVs were enriched in proteins participating in metabolic processes such as translation and tRNA aminoacylation for protein translation. These are essential processes for protein synthesis and are related to the ability of cancer cells to sustain chronic proliferation. Cellular processes enhanced in our ELVs included cell cycle, cell proliferation, cell communication, cytokinesis, chromosome segregation and cellular component movement, all of them related to the hallmarks of cancer. Adhesion, a biological process directly associated to EMT processes in EC progression and tumor dissemination was also overrepresented. Importantly, vesicle-mediated transport was also one of the overrepresented processes.

**Table 24.** Quantitative proteomics studies performed on extracellular vesicles of different origin. Adapted from Pocsfalvi et al.<sup>207</sup>.

	Biological Source of EVs	Stimuli/disease	Type of EVs isolated	Method used for protein quantitation	Number of identified proteins	Number of differentially expr. proteins	Ref.
Cell or microorganism culture	Human bladder carcinoma cell lines (T24, FL3 and SLT4)	metastatic (FL3 and SLT4) versus non-metastatic (T24) cells	EVs lumen and membrane fractions	4plex-iTRAQ	1040, 1333 (membrane replicate) 547, 523 (lumen replicate)	58 (FL3 and SLT4 versus T24)	(Jeppesen, et al., 2014)
	Human primary (SW480) and metastatic colorectal cancer cells (SW620)	Primary and metastatic cancer	EVs	Label-free SC	803 (SW480) 787 (SW620)	368 (SW480) 359 (SW620)	(Choi, et al., 2012)
	Human prostate cancer cells (PC3)	expression of polymerase I and transcript release factor in the caveolin-1 positive cell line PC-3	EVs (prostasome)	SILAC	369 (Prostasome) 372 (Secretome)	123 (Prostasome) 125 (Secretome)	(Inder, et al., 2014; Inder, et al., 2012)
	Brain tumor cells (murine) and Sera (human)	Human versus murine brain tumor exosomes	EVs (density gradient)	2D-DIGE	45	36	(Graner, et al., 2009)
	Human microvascular endothelial cells (HMEC-1)	hypoxia, TNF- $\alpha$ -induced activation, high glucose and mannose concentrations	Exosomes (gradient centrifugation)	8plex-iTRAQ	1354	19	(de Jong, et al., 2012)
	Mouse myoblasts (C2C12)	Myogenesis (myoblast transformation to myotubes)	Exosomes	Label-free SC	455	31 (myoblast) 78 (myotubes)	(Forterre, et al., 2014)
	B-cells	Infection with EBV, Kaposi sarcoma-associated (KSHV) or with Human gamma herpes virus	Exosomes (sucrose cushion)	2D-DIGE and Label free SC	871 (B-cells) 360 (viral)	187	(Meckes, et al., 2013)
	$\beta$ -cells	cytokine induced apoptosis	Microvesicles and Exosomes	SILAC	401 (Microvesicles) 191 (Exosomes)	366 (Microvesicles) 164 (Exosomes)	(Palmisano, et al., 2012)
	Human mesenchymal stem cells	Activin-A stimulus during <i>in vitro</i> osteogenic differentiation of human mesenchymal stem cells	Extracellular matrix and EVs	SILAC	293 (Extracellular matrix) 27 (EVs)	104 (Extracellular matrix) 12 (EVs)	(Alves, et al., 2013)
	Primary lymphoblast cells	WT and CD81-deficient animals	EVs	18O labeling and LC-MS/MS	692	25	(Perez-Hernandez, et al., 2013)
	Lymphocytic cells (H9)	HIV-1 infection	EVs	SILAC	770	14	(Li, et al., 2012)
	Trypanosoma cruzi culture	Secretion mechanism	EVs	Label-free SC	367	149	(Bayer-Santos, et al., 2012)
	Human monocyte-derived macrophages	fungal infections: $\beta$ -glucan-stimulation	Microvesicles	4plex-iTRAQ	540	145	(Cypryk, et al., 2014)
	Human platelet	Thrombin, Shear	Microvesicles	2-DE-PAGE MS	2307 2271	26	(Shai, et al., 2012)
	Primary hepatocytes and sera (rat)	liver toxins (galactosamine and Escherichia coli-derived lipopolysaccharide)	EVs (differential centrifugation)	Label-free 1D-LC-DIA-MS and 2D-LC-DIA-MS	412 557	88	(Rodríguez-Suárez, et al., 2014)
Milk (bovine)	Staphylococcus aureus infected cows	Exosomes	iTRAQ	300	94	(Reinhardt, et al., 2013)	
Biofluids	Urine (human)	Hernia and bladder cancer	EVs (differential centrifugation)	Isotopic dimethylation labeling	2964	107	(Chen, et al., 2012)
	Urine (human)	Pooled sample compared to single sample, two different age groups	EVs (differential centrifugation and double-cushion)	4plex-iTRAQ	378	21 6	(Raj, et al., 2012a)
	Urine (human)	Diabetic nephropathy	Mixed EVs population	Label-free SC	562	25	(Zubiri, et al., 2014)
	Urine (rat and human)	Acute kidney injury, cisplatin treatment	EVs (differential centrifugation)	2D-DIGE-MALDI-TOF/TOF and LC-ESI-MS/MS	40	27	(Zhou, et al., 2006)

Summarizing, we can robustly affirm that we described a population of vesicles that act as conveyors of important proteins to draw the big picture of cancer.

In conclusion for Objective 2, we developed a novel and reliable proteomic approach for the study of the protein content of ELVs isolated from UAs. We combined the advantages of working with exosomes with an in-gel procedure to overcome the dynamic-range issue inherent to the nature of UA samples. Moreover, the use of our house-made IS guarantees a robust selection of candidates for subsequent targeted MS experiments. To our knowledge, we are the first to describe the proteome of ELVs isolated from UAs in the context of EC. Our study led us to generate a list of 54 robust proteins that might allow diagnosis and prognosis of EC; this list was analyzed in a verification phase.

The verification phase is intended to determine whether there is enough evidence for potential clinical utility of a given candidate biomarker to warrant further investment in it for clinical validation studies, which are highly costly in terms of time, money and using of clinical samples<sup>280</sup>. The desired scenario after a discovery phase, in which thousands of proteins have been identified, is the use of technologies that allow an accurate protein quantification of several candidates simultaneously in a large cohort of patients. Targeted MS approaches, such as LC-SRM, enable the accurate detection and quantification of multiple low-abundance proteins in highly complex mixtures in a single run. In addition, it allows determining all possible combinations among the monitored candidates to define biomarker signatures that meet the standards of sensitivity and specificity required for clinical application. Moreover, LC-SRM does not rely on the existence and availability of antibodies, as it is the case for ELISA tests<sup>280,281</sup>.

Our verification study was performed by LC-SRM in a new, bigger and independent cohort of 107 patients meeting the same inclusion criteria as for the discovery phase. This study revealed that 89,6% of the candidates selected to discriminate between CTRL and EC patients were validated even though no sample fractionation was performed this time. These results confirmed that the strategy followed in the discovery phase permitted a robust prioritization of candidates and ratify the high sensitivity of the LC-SRM technique for the



verification phase. We further investigated the expression of candidates in the whole fluid fraction of UAs and, interestingly, we found a number of proteins that were exclusively belonging to ELVs, as they were undetectable in UAs. Importantly, this confirms a good enrichment in ELVs from the samples of the discovery phase. Although it is not on the scope of this thesis, the study of their functions will bring insights in understanding to what extent ELVs play a key role in EC associated processes.

Here, we defined two protein biomarker signatures: one to detect EC and the other to discriminate between the two most common histologies, namely EEC and SEC.

Although the most relevant individual biomarkers were AGRIN (AUC= 0.90), TACD2 (AUC= 0.87), SORT (AUC=0.86), MVP (AUC= 0.86) and FAS (AUC= 0.85), the diagnostic panel was composed of two proteins: AGRIN and CD81, whose individual potential was  $AUC_{\text{AGRIN}} = 0.90$  and  $AUC_{\text{CD81}} = 0.63$ . Agrin (AGRIN) is a large, multidomain heparan sulfate (HS) proteoglycan that cross-links the cytoskeleton of cells with the extracellular matrix, where it interacts with laminins and cell-surface receptors. It is localized to basement membranes and expressed in several tissues<sup>282</sup>. Agrin binds to and regulates growth factors such as TGF- $\beta$ <sup>283</sup>, known to be a key factor in the initiation of tumor invasion<sup>284</sup>. Dramatically elevated levels of Agrin have been reported in the tumor microvasculature of hepatocellular carcinoma (HCC)<sup>285</sup>. A recent study reported that Agrin promoted proliferation, migration, invasion, and EMT in HCC cell lines; similar pro-tumorigenic effects of Agrin have also been observed in oral squamous cell carcinoma<sup>286,287</sup>. Agrin appears in the ExoCarta database but, to the best of our knowledge, this is the first work that shows a relation between Agrin and EC. CD81 is member of the transmembrane 4 superfamily, also known as the tetraspanin family. CD81 participates in a variety of important cellular processes such as membrane organization, protein trafficking, cellular fusion and cell-cell interactions<sup>288</sup>. This glycoprotein is known to complex with integrins, other tetraspanins and with cell-specific partner proteins, providing cells with a signaling platform. CD81 is one of the most widely-used exosomal markers<sup>136</sup> and has also been shown to regulate cell migration and invasion, and has therefore been implicated in cancer progression. It has been recently shown that CD81 contributes

to tumor growth and metastasis<sup>288,289</sup>. CD81 is expressed in most types of cancer, including breast, lung, prostate, melanoma, brain cancer and lymphoma, and the overexpression or down-regulation of this molecule has been correlated with either good or bad prognosis<sup>289</sup>. In combination, the two proteins can discriminate EC with 89% sensitivity and 88% specificity, when are analyzed in ELVs isolated from the fluid fraction of UAs.

Additionally, we found that the combination of two other proteins (CLD6 and RAB8A) allowed discriminating between the EC histological subtypes EEC and SEC with a sensitivity of 95% and a specificity of 87%. The most relevant individual prognostic biomarkers were CLD6 (AUC= 0.88), BCAM (AUC= 0.87), IF2B3 (AUC=0.80) and PLD3 (AUC= 0.79). The importance of accurately discriminate between EEC and SEC cases lies on the definition of an optimal treatment for EC patients. Most of the EEC are low-grade tumors associated to a good prognosis. The standard surgical treatment for these group of patients consists of total hysterectomy and bilateral salpingo-oophorectomy; for the high-grade EEC cases, the surgical treatment is complemented with lymphadectomy<sup>92</sup>. On the other hand, SEC tumors are more aggressive and present a poor prognosis and, in consequence, these patients receive a more extensive surgery consisting of hysterectomy, bilateral salpingo-oophorectomy, pelvic and para-aortic lymphadectomy, omentectomy, appendectomy and peritoneal biopsies.

CLD6, the fetal tight junction molecule claudin 6, is virtually absent from any normal tissue, but it is aberrantly and frequently expressed in various cancer types such as ovarian, lung, gastric, breast and pediatric cancers<sup>290-295</sup>. The genes of other members of the family (claudin 3 and 4) have been described in uterine serous papillary carcinoma<sup>296</sup> and carcinosarcoma<sup>297</sup>. However, we have not found any references relating CLD6 to EC in the literature. RAB8A is a small GTPase belonging to the rab family of ras-GTPases involved in intercellular trafficking, and is also a marker of recycling endosomes. Alterations in Rab-GTPase expression or activity can cause defects in cell adhesion, motility and invasion, leading to neurologic diseases, lipid storage disorders or cancer<sup>298</sup>. A significant increase of RAB8A in EC versus healthy controls tissues has been previously described<sup>299</sup>. Sun et al. reported a possible regulation of muscle GLUT 4 traffic in response to

insulin by RAB8A (and also RAB13), proposing an explanation to the cause of EC in patients with obesity and diabetes<sup>300</sup>. Both CLD6 and RAB8A have been previously found in exosome preparations, as reported in the ExoCarta database.

In conclusion for Objective 3, we used LC-SRM to successfully validate a high proportion of the candidates selected from the discovery phase. LC-SRM together with bioinformatics allowed the generation of a diagnostic and a prognostic panel, capable of improving the current diagnostic process.

As general remarks, our study presents some strength and limitations that are worth taking into consideration.

From the population point of view, our study focused on the two most common histological subtypes of EC (EEC and SEC) and embraced the vast diversity of women presenting with AUB and/or thickening of the endometrium entering the diagnostic process for suspicion of EC (i.e. women with polyps, myomas, endometrial hyperplasia or normal endometrium). With respect to age and endometrial status of women enrolling our studies, we centered on post-menopausal women older than 40 years for the discovery and verification phases, and included a proportion of pre-menopausal women for the verification of candidates in the whole fluid fraction of UAs. In this regard, it could have been interesting to explore a bigger cohort of pre-menopausal women from the discovery phase and to include a study group of the different types of hyperplasias, as it is a precursor lesion of EC. Also, increasing the representation of different tumor grades among the EC cases would be of great interest to evaluate the potential of the biomarkers to discriminate among tumor grades. The accurate definition of tumor grade, as well as histological subtype, help clinicians to determine the best surgical treatment at the moment of diagnosis.

From a technical point of view, we were able to standardize protocols to isolate exosomes from UAs providing highly ELVs-enriched samples for subsequent proteomic analysis. For the discovery phase, we designed a novel super-SILAC approach combined with in-gel sample fractionation and digestion that permitted the identification of an elevated number of low abundance proteins contained in

ELVs that, otherwise, would have been impossible to detect with shotgun proteomics. The advantage of using multiplex-targeted proteomics (i.e. LC-SRM) for the verification phase is that, unlike the antibody-based approaches, we can measure the concentration levels of several candidates simultaneously (up to 54 in our case) increasing with that the likelihood of finding a clinically useful biomarker and allowing the evaluation of different combinations among all the studied proteins, enabling the generation of protein panels. Nevertheless, all this processing (from the isolation of exosomes to the sample preparation for proteomics analysis) is tedious and time-costly. In line with this, automated sample processing for proteomics is already available and many proteomics platforms have standardized the sample preparation in a robotic fashion. Although a recent work in our group has reported a good diagnostic signature in the whole fluid fraction of UAs, our work on exosomes explores a different, complementary and interesting source of biomarkers. New instruments are emerging in the field of exosomes that will facilitate the current issues regarding quickness and reproducibility. The company *Exosomes Diagnostics Inc.* have recently developed a point of care protein detection instrument that achieves high sensitivity by selectively targeting disease-specific exosomes. They describe the instrument as a powerful technology for discovering, assessing and validating clinical biomarkers. This and upcoming technical improvements on the field will ensure that studies like ours make an impact in clinical application shortly.

Despite these promising and encouraging results, further validation steps must be performed. The reasonable next step of this study should be the evaluation of the panels by immunoassays, such as ELISA.

In summary, the work presented in this thesis confirmed the existence of ELVs in the fluid fraction of UAs, and explored the suitability of this matrix as a source of biomarkers for EC. The development of the Super-SILAC approach combined with the in-gel digestion for LC-MS/MS analysis unveiled the proteome of ELVs derived from control and EC patients and allowed the identification of a high number of proteins with great potential to discriminate between cancer and non-cancer samples and between EEC and SEC, which were verified by LC-SRM. This targeted proteomic approach allowed for the development of a 2-protein signature

composed by AGRIN and CD81 that permits EC diagnosis (ROC AUC=0.94); and a 2-protein signature composed by CLD6 and RAB8A that can differentiate EEC versus SEC histologies (ROC AUC=0.93).

This study represents an important contribution to the field of proteomic biomarker identification in UAs and has important implications in improving the detection of EC and in patient stratification. Hopefully, these results will be further validated in order to achieve clinical utility shortly, improving the current diagnostic process as well as helping clinicians to accurately define the surgical treatment of EC patients.

# CONCLUSIONS

---



The main conclusions derived from this thesis are:

1. ELVs exist in the fluid fraction of uterine aspirates and can be isolated by a standard protocol based on differential ultracentrifugation, giving sufficient proteomic material for further analyses.
2. The proteomic content of ELVs was very heterogeneous among different patients, so the use of gel fractionation approach allowed for a better coverage of the ELVs proteome. A total of 2,138 proteins were identified in ELVs from UAs.
3. The proteomic study of 5 EC cell lines permitted the development of a Super-SILAC mix combining ELVs derived from the AN3CA, KLE and RL95.2 cell lines. This combination permitted to quantify 43% of the patients' proteome in the discovery study.
4. The Gene Ontology analysis of the label-free data showed that biological processes such as adhesion, mitosis and exocytosis were overrepresented in EC ELVs. Moreover, cellular processes enhanced in EC ELVs included cell cycle, cell proliferation, cell communication, cytokinesis, chromosome segregation and cellular component movement, all of them related to the hallmarks of cancer.
5. The discovery study permitted the generation of a list of 54 candidates, being 50 potential diagnostic biomarkers and 15 prognostic biomarkers. This list was generated combining (i) the relative quantification analysis, (ii) the presence/absence study, and (iii) biological criteria.
6. The use of LC-SRM in the verification phase enabled the accurate detection and quantification of 83.3% (45/54) of candidates in ELVs from UAs of an independent cohort of 107 patients, that include EC1, EC2 and Control patients.
7. The LC-SRM study revealed that 89.6% of the diagnostic and 26.6% of the prognostic candidate biomarkers were verified, even though no sample fractionation was performed this time. This led us conclude that the study of



proteins in ELVs have great potential for the identification of EC biomarkers, and presumably, this could be expanded to other gynecological diseases.

8. The LC-SRM study of the 54 candidates in the whole fluid fraction of UAs, permitted to understand that whilst some proteins are highly specific of ELVs and are not detected in UAs (29.5%), some others might have also a diagnostic and prognostic potential if analyzed in UAs. This opens an avenue to explore the insights of highly specific ELVs proteins, but also, to explore the use of UAs as a direct source of biomarkers.

9. Thanks to our results, we developed a diagnostic panel composed of AGRIN and CD81 that exhibited 89% sensitivity and 88% specificity for detecting EC cases in UAs ELVs. If validated, this panel is expected to improve the diagnostic power of UAs, leading to a reduction of invasive samplings in the diagnostic process.

10. Finally, we also developed a prognostic panel composed of CLD6 and RAB8A that exhibited 95% sensitivity and 87% specificity for the discrimination of the two main histological subtypes studied (i.e. EEC and SEC). If validated, this panel is expected to improve the prognostic power of UAs, leading to a better patient's stratification and helping to predict the optimal surgical treatment for EC patients.

# **JOURNAL PUBLICATIONS**

---



This thesis led to the publication (Paper A) and the preparation (Paper B) of the following manuscripts:

**[Paper A]** I. Campoy\*, L. Lanau\*, T. Altadill, T. Sequeiros, S. Cabrera, M. Cubo-Abert, A. Pérez-Benavente, A. Garcia, S. Borrós, A. Santamaria, J. Ponce, X. Matias-Guiu, J. Reventós, A. Gil-Moreno, M. Rigau, and E. Colas, “Exosome-like vesicles in uterine aspirates: a comparison of ultracentrifugation-based isolation protocols,” *J. Transl. Med.*, vol. 14, no. 1, p. 180, Dec. 2016. (\* co-authors)

**[Paper B]** I. Campoy et al. “Identification of proteomic signatures for endometrial cancer diagnosis and prognosis in exosome-like vesicles of uterine aspirates”, in preparation for submission to a high impact factor journal.

In the course of this thesis, collaborations with other projects focused on endometrial cancer research have also been done. Although these works are not part of the thesis, they are listed in the following articles:

**[Paper C]** T. Altadill, I. Campoy, L. Lanau, K. Gill, M. Rigau, A. Gil-Moreno, J. Reventós, S. Byers, E. Colas, and A. K. Cheema, “Enabling metabolomics based biomarker discovery studies using molecular phenotyping of exosome-like vesicles,” *PLoS One*, vol. 11, no. 3, 2016.

**[Paper D]** A. Mota, E. Colás, P. García-Sanz, I. Campoy, A. Rojo-Sebastián, S. Gatiús, Á. García, L. Chiva, S. Alonso, A. Gil-Moreno, X. González-Tallada, B. Díaz-Feijoo, A. Vidal, P. Ziober-Malinowska, M. Bobiński, R. López-López, M. Abal, J. Reventós, X. Matias-Guiu, and G. Moreno-Bueno, “Genetic analysis of uterine aspirates improves the diagnostic value and captures the intra-tumor heterogeneity of endometrial cancers,” *Mod. Pathol.*, vol. 30, no. 1, pp. 134–145, Jan. 2017.

**[Paper E]** N. Pedrola, L. Devis, M. Llauradó, I. Campoy, E. Martínez-García, M. García, L. Muínelo-Romay, L. Alonso-Alconada, M. Abal, F. Alameda, G. Mancebo, R. Carreras, J. Castellví, S. Cabrera, A. Gil-Moreno, X. Matias-Guiu, J. L. Iovanna, E. Colas, J. Reventós, and A. Ruiz, “Nidogen 1 and Nuclear Protein 1: novel targets of ETV5 transcription factor involved in endometrial cancer invasion,” *Clin. Exp. Metastasis*, vol. 32, no. 5, pp. 467–478, 2015.

**[Paper F]** E. Colas, N. Pedrola, L. Devis, T. Ertekin, I. Campoy, E. Martínez, M. Llauradó, M. Rigau, M. Olivan, M. García, S. Cabrera, A. Gil-Moreno, J. Xercavins, J. Castellvi, A. García, S. Santiago Ramon Y Cajal, G. Moreno-Bueno,

X. Dolcet, F. Alameda, J. Palacios, J. Prat, A. Doll, X. Matias-Guiu, M. Abal, and J. Reventos, "The EMT signaling pathways in endometrial carcinoma," *Clin. Transl. Oncol.*, vol. 14, no. 10, pp. 715–720, 2012.

# BIBLIOGRAPHY

---



1. M. S. Baggish, R. F. Valle, and H. Guedj, *Hysteroscopy: visual perspectives of uterine anatomy, physiology and pathology*. Wolters Kluwer Health/Lippincott Williams & Wilkins, 2007.
2. K. M. Van De Graaff, *Human anatomy*, 6th ed. 2001.
3. L. Cabero Roura, E. Cabrillo Rodríguez, and J. Bajo Arenas, *Tratado de ginecología y obstetricia*. Panamericana, 2013.
4. R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer Statistics, 2017.," *CA. Cancer J. Clin.*, vol. 67, no. 1, pp. 7–30, 2017.
5. J. Ferlay, I. Soerjomataram, R. Dikshit, et al., "Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012," *Int. J. Cancer*, vol. 136, no. 5, pp. E359–E386, Mar. 2015.
6. Sociedad Española de Oncología Médica, "Las Cifras del Cáncer en España 2016," 2016.
7. R. Marcos-Gragera, M. Loreto, V. Gil, et al., "El Càncer a Girona 2007-2009. Projeccions de la Incidència 2013-14," 2013.
8. "MEMÒRIA 2013 FUNCA (Fundació lliga per a la Investigació i Prevenció del Càncer)," Tarragona, 2013.
9. Generalitat de Catalunya. Departament de Salut, "Estrategia y prioridades en cáncer en Cataluña. Plan Director de Oncología: Objetivos 2010," Barcelona, 2009.
10. American cancer Society, "Cancer Facts & Figures 2017," 2017.
11. J. Ponce, R. Torrejón and M. Barahona. "Oncoguía SEGO: Cáncer de Endometrio 2016. Guías de práctica clínica en cáncer ginecológico y mamario", *Publicaciones SEGO*, vol. Oncoguías, 2016.
12. G. Emons, G. Fleckenstein, B. Hinney, et al., "Hormonal interactions in endometrial cancer," *Endocr. Relat. Cancer*, vol. 7, pp. 227–242, 2000.
13. J. L. Hecht and G. L. Mutter, "Molecular and pathologic aspects of endometrial carcinogenesis.," *J. Clin. Oncol.*, vol. 24, no. 29, pp. 4783–91, Oct. 2006.
14. F. Amant, P. Moerman, P. Neven, et al., "Endometrial cancer.," *Lancet*



- (London, England), vol. 366, no. 9484, pp. 491–505, Aug. 2005.
15. B. S. Evans-Metcalf ER, Brooks SE, Reale FR, “Profile of women 45 years of age and younger with endometrial cancer. - PubMed - NCBI,” *Obstet. Gynecol.*, 1998.
  16. L. R. Duska, A. Garrett, B. R. Rueda, et al., “Endometrial Cancer in Women 40 Years Old or Younger,” *Gynecol. Oncol.*, vol. 83, no. 2, pp. 388–393, 2000.
  17. L. A. Brinton, M. L. Berman, R. Mortel, et al., “Reproductive, menstrual, and medical risk factors for endometrial cancer: results from a case-control study.,” *Am. J. Obstet. Gynecol.*, vol. 167, no. 5, pp. 1317–25, Nov. 1992.
  18. M. Fujita, T. Tase, Y. Kakugawa, et al., “Smoking, earlier menarche and low parity as independent risk factors for gynecologic cancers in Japanese: a case-control study.,” *Tohoku J. Exp. Med.*, vol. 216, no. 4, pp. 297–307, Dec. 2008.
  19. A. Zucchetto, D. Serraino, J. Polesel, et al., “Hormone-related factors and gynecological conditions in relation to endometrial cancer risk.,” *Eur. J. Cancer Prev.*, vol. 18, no. 4, pp. 316–21, Aug. 2009.
  20. L. A. Brinton, C. L. Westhoff, B. Scoccia, et al., “Causes of infertility as predictors of subsequent cancer risk.,” *Epidemiology*, vol. 16, no. 4, pp. 500–7, Jul. 2005.
  21. A. T. Ali, “Reproductive Factors and the Risk of Endometrial Cancer,” *Int. J. Gynecol. Cancer*, vol. 24, no. 3, pp. 384–393, Mar. 2014.
  22. N. Reis and N. K. Beji, “Risk factors for endometrial cancer in Turkish women: results from a hospital-based case-control study.,” *Eur. J. Oncol. Nurs.*, vol. 13, no. 2, pp. 122–7, Apr. 2009.
  23. J. A. Barry, M. M. Azizia, and P. J. Hardiman, “Risk of endometrial, ovarian and breast cancer in women with polycystic ovary syndrome: a systematic review and meta-analysis,” *Hum. Reprod. Update*, vol. 20, no. 5, pp. 748–758, Sep. 2014.
  24. M. Gottschau, S. K. Kjaer, A. Jensen, et al., “Risk of cancer among women with polycystic ovary syndrome: A Danish cohort study,” *Gynecol. Oncol.*,

- vol. 136, no. 1, pp. 99–103, Jan. 2015.
25. E. Weiderpass, H. O. Adami, J. A. Baron, et al., “Risk of endometrial cancer following estrogen replacement with and without progestins.,” *J. Natl. Cancer Inst.*, vol. 91, no. 13, pp. 1131–7, Jul. 1999.
  26. J. D. Woodruff and J. H. Pickar, “Incidence of endometrial hyperplasia in postmenopausal women taking conjugated estrogens (Premarin) with medroxyprogesterone acetate or conjugated estrogens alone. The Menopause Study Group.,” *Am. J. Obstet. Gynecol.*, vol. 170, no. 5 Pt 1, pp. 1213–23, May 1994.
  27. I. Schiff, H. K. Sela, D. Cramer, et al., “Endometrial hyperplasia in women on cyclic or continuous estrogen regimens.,” *Fertil. Steril.*, vol. 37, no. 1, pp. 79–82, Jan. 1982.
  28. P. Razavi, M. C. Pike, P. L. Horn-Ross, et al., “Long-term postmenopausal hormone therapy and endometrial cancer.,” *Cancer Epidemiol. Biomarkers Prev.*, vol. 19, no. 2, pp. 475–83, Feb. 2010.
  29. D. Grady, T. Gebretsadik, K. Kerlikowske, et al., “Hormone replacement therapy and endometrial cancer risk: A meta-analysis,” *Obstet. Gynecol.*, vol. 85, no. 2, pp. 304–313, Feb. 1995.
  30. J. V Lacey, L. A. Brinton, J. H. Lubin, et al., “Endometrial carcinoma risks among menopausal estrogen plus progestin and unopposed estrogen users in a cohort of postmenopausal women.,” *Cancer Epidemiol. Biomarkers Prev.*, vol. 14, no. 7, pp. 1724–31, Jul. 2005.
  31. J. A. Doherty, K. L. Cushing-Haugen, B. S. Saltzman, et al., “Long-term use of postmenopausal estrogen and progestin hormone therapies and the risk of endometrial cancer,” *Am. J. Obstet. Gynecol.*, vol. 197, no. 2, p. 139.e1-139.e7, Aug. 2007.
  32. J. Marjoribanks, C. Farquhar, H. Roberts, et al., “Long-term hormone therapy for perimenopausal and postmenopausal women,” in *Cochrane Database of Systematic Reviews*, vol. 1, C. Farquhar, Ed. Chichester, UK: John Wiley & Sons, Ltd, 2017, p. CD004143.
  33. H. Mouridsen, A. Giobbie-Hurder, A. Goldhirsch, et al., “Letrozole Therapy

- Alone or in Sequence with Tamoxifen in Women with Breast Cancer,” *N. Engl. J. Med.*, vol. 361, no. 8, pp. 766–776, Aug. 2009.
34. M. J. Mourits, E. G. De Vries, P. H. Willemse, et al., “Tamoxifen treatment and gynecologic side effects: a review.,” *Obstet. Gynecol.*, vol. 97, no. 5 Pt 2, pp. 855–66, May 2001.
  35. R. Kaaks, A. Lukanova, and M. S. Kurzer, “Obesity, endogenous hormones, and endometrial cancer risk: a synthetic review.,” *Cancer Epidemiol. Biomarkers Prev.*, vol. 11, no. 12, pp. 1531–43, Dec. 2002.
  36. R. E. Schmandt, D. A. Iglesias, N. N. Co, and K. H. Lu, “Understanding obesity and endometrial cancer risk: opportunities for prevention,” *Am. J. Obstet. Gynecol.*, vol. 205, no. 6, pp. 518–525, Dec. 2011.
  37. A. N. Fader, L. N. Arriba, H. E. Frasure, and V. E. von Gruenigen, “Endometrial cancer and obesity: Epidemiology, biomarkers, prevention and survivorship,” *Gynecol. Oncol.*, vol. 114, no. 1, pp. 121–127, Jul. 2009.
  38. L. R. Nelson and S. E. Bulun, “Estrogen production and action.,” *J. Am. Acad. Dermatol.*, vol. 45, no. 3 Suppl, pp. S116-24, Sep. 2001.
  39. M. Pollak, “Insulin and insulin-like growth factor signalling in neoplasia,” *Nat. Rev. Cancer*, vol. 8, no. 12, pp. 915–928, Dec. 2008.
  40. E. Friberg, N. Orsini, C. S. Mantzoros, and A. Wolk, “Diabetes mellitus and risk of endometrial cancer: a meta-analysis,” *Diabetologia*, vol. 50, no. 7, pp. 1365–1374, Jun. 2007.
  41. K. Lindemann, L. J. Vatten, M. Ellstrøm-Eng, and A. Eskild, “Body mass, diabetes and smoking, and endometrial cancer risk: a follow-up study.,” *Br. J. Cancer*, vol. 98, no. 9, pp. 1582–5, May 2008.
  42. R. T. Jung, “Obesity as a disease.,” *Br. Med. Bull.*, vol. 53, no. 2, pp. 307–21, 1997.
  43. E. E. Calle and R. Kaaks, “Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms,” *Nat. Rev. Cancer*, vol. 4, no. 8, pp. 579–591, Aug. 2004.
  44. M. Soler, L. Chatenoud, E. Negri, et al., “Hypertension and hormone-related neoplasms in women.,” *Hypertens. (Dallas, Tex. 1979)*, vol. 34, no. 2, pp.

- 320–5, Aug. 1999.
45. F. Parazzini, C. La Vecchia, E. Negri, et al., “Diabetes and endometrial cancer: an Italian case-control study.,” *Int. J. cancer*, vol. 81, no. 4, pp. 539–42, May 1999.
  46. L. L. Holman and K. H. Lu, “Genetic risk and gynecologic cancers.,” *Hematol. Oncol. Clin. North Am.*, vol. 26, no. 1, pp. 13–29, Feb. 2012.
  47. H. T. Lynch, M. W. Shaw, C. W. Magnuson, A. L. Larsen, and A. J. Krush, “Hereditary factors in cancer. Study of two large midwestern kindreds.,” *Arch. Intern. Med.*, vol. 117, no. 2, pp. 206–12, Feb. 1966.
  48. C. A. Shu, M. C. Pike, A. R. Jotwani, et al., “Uterine Cancer After Risk-Reducing Salpingo-oophorectomy Without Hysterectomy in Women With BRCA Mutations,” *JAMA Oncol.*, vol. 2, no. 11, p. 1434, Nov. 2016.
  49. M. M. de Jonge, A. L. Mooyaart, M. P. G. Vreeswijk, et al., “Linking uterine serous carcinoma to BRCA1/2-associated cancer syndrome: A meta-analysis and case report,” *Eur. J. Cancer*, vol. 72, pp. 215–225, 2017.
  50. D. W. Cramer, “The epidemiology of endometrial and ovarian cancer.,” *Hematol. Oncol. Clin. North Am.*, vol. 26, no. 1, pp. 1–12, Feb. 2012.
  51. D. W. Kaufman, S. Shapiro, D. Slone, et al., “Decreased Risk of Endometrial Cancer among Oral-Contraceptive Users,” *N. Engl. J. Med.*, vol. 303, no. 18, pp. 1045–1047, Oct. 1980.
  52. G. Sahin Ersoy, Y. Zhou, H. İnan, et al., “Cigarette Smoking Affects Uterine Receptivity Markers,” *Reprod. Sci.*, p. 193371911769712, Mar. 2017.
  53. C. Lawrence, I. Tessaro, S. Durgerian, et al., “Smoking, body weight, and early-stage endometrial cancer.,” *Cancer*, vol. 59, no. 9, pp. 1665–9, May 1987.
  54. E. M. John, J. Koo, and P. L. Horn-Ross, “Lifetime Physical Activity and Risk of Endometrial Cancer,” *Cancer Epidemiol. Biomarkers Prev.*, vol. 19, no. 5, pp. 1276–1283, May 2010.
  55. A. Goodman, “Abnormal genital tract bleeding.,” *Clin. Cornerstone*, vol. 3, no. 1, pp. 25–35, 2000.

56. J. R. Albers, S. K. Hull, and R. M. Wesley, "Abnormal uterine bleeding.," *Am. Fam. Physician*, vol. 69, no. 8, pp. 1915–26, Apr. 2004.
57. J. N. Bakkum-Gamez, J. Gonzalez-Bosquet, N. N. Laack, et al., "Current issues in the management of endometrial cancer.," *Mayo Clin. Proc.*, vol. 83, no. 1, pp. 97–112, Jan. 2008.
58. T. Gredmark, S. Kvint, G. Havel, and L.-A. Mattsson, "Histopathological findings in women with postmenopausal bleeding," *BJOG An Int. J. Obstet. Gynaecol.*, vol. 102, no. 2, pp. 133–136, Feb. 1995.
59. M. G. Munro, "Classification of menstrual bleeding disorders," 2012.
60. N. Colombo, C. Creutzberg, F. Amant, et al., and ESMO-ESGO-ESTRO Endometrial Consensus Conference Working Group, "ESMO-ESGO-ESTRO Consensus Conference on Endometrial Cancer: Diagnosis, Treatment and Follow-up.," *Int. J. Gynecol. Cancer*, vol. 26, no. 1, pp. 2–30, Jan. 2016.
61. N. Colombo, E. Preti, F. Landoni, et al., "Endometrial cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up," *Ann. Oncol.*, vol. 24, no. SUPPL.6, 2013.
62. R. A. Smith, V. Cokkinides, and O. W. Brawley, "Cancer screening in the United States, 2009: A review of current American Cancer Society guidelines and issues in cancer screening," *CA. Cancer J. Clin.*, vol. 59, no. 1, pp. 27–41, Jan. 2009.
63. R. Manchanda, E. Saridogan, A. Abdelraheim, et al., "Annual outpatient hysteroscopy and endometrial sampling (OHES) in HNPCC/Lynch syndrome (LS)," *Arch. Gynecol. Obstet.*, vol. 286, no. 6, pp. 1555–1562, Dec. 2012.
64. H. F. A. Vasen, I. Blanco, K. Aktan-Collan, et al., "Revised guidelines for the clinical management of Lynch syndrome (HNPCC): recommendations by a group of European experts," *Gut*, vol. 62, no. 6, pp. 812–823, Jun. 2013.
65. R. Smith-Bindman, K. Kerlikowske, V. A. Feldstein, et al., "Endovaginal ultrasound to exclude endometrial cancer and other endometrial

- abnormalities.," *JAMA*, vol. 280, no. 17, pp. 1510–7, Nov. 1998.
66. J. K. Gupta, P. F. W. Chien, D. Voit, T. J. Clark, and K. S. Khan, "Ultrasonographic endometrial thickness for diagnosing endometrial pathology in women with postmenopausal bleeding: a meta-analysis.," *Acta Obstet. Gynecol. Scand.*, vol. 81, no. 9, pp. 799–816, Sep. 2002.
67. A. Timmermans, B. C. Opmeer, K. S. Khan, et al., "Endometrial Thickness Measurement for Detecting Endometrial Cancer in Women With Postmenopausal Bleeding," *Obstet. Gynecol.*, vol. 116, no. 1, pp. 160–167, Jul. 2010.
68. D. Abramson and S. G. Driscoll, "Endometrial aspiration biopsy.," *Obstet. Gynecol.*, vol. 27, no. 3, pp. 381–91, Mar. 1966.
69. T. J. Clark, C. H. Mann, N. Shah, et al., "Accuracy of outpatient endometrial biopsy in the diagnosis of endometrial cancer: A systematic quantitative review," *BJOG An Int. J. Obstet. Gynaecol.*, vol. 109, no. 3, pp. 313–321, 2002.
70. L. D. Bradley, "Complications in hysteroscopy: prevention, treatment and legal risk.," *Curr. Opin. Obstet. Gynecol.*, vol. 14, no. 4, pp. 409–15, Aug. 2002.
71. N. P. Polyzos, D. Mauri, S. Tsioras, et al., "Intraperitoneal dissemination of endometrial cancer cells after hysteroscopy: a systematic review and meta-analysis.," *Int. J. Gynecol. Cancer*, vol. 20, no. 2, pp. 261–7, Feb. 2010.
72. W. M. Burke, J. Orr, M. Leitao, et al., "Endometrial cancer: A review and current management strategies: Part I," *Gynecol. Oncol.*, vol. 134, no. 2, pp. 385–392, 2014.
73. M. Lieng, O. Istre, and E. Qvigstad, "Treatment of endometrial polyps: a systematic review," *Acta Obstet. Gynecol.*, vol. 89, pp. 992–1002, 2010.
74. S. C. Lee, A. M. Kaunitz, L. Sanchez-Ramos, and R. M. Rhatigan, "The Oncogenic Potential of Endometrial Polyps," *Obstet. Gynecol.*, vol. 116, no. 5, pp. 1197–1205, Nov. 2010.
75. K.-R. Kim, R. Peng, J. Y. Ro, and S. J. Robboy, "A diagnostically useful histopathologic feature of endometrial polyp: the long axis of endometrial

- glands arranged parallel to surface epithelium.," *Am. J. Surg. Pathol.*, vol. 28, no. 8, pp. 1057–62, Aug. 2004.
76. V. Chandra, J. J. Kim, D. M. Benbrook, A. Dwivedi, and R. Rai, "Therapeutic options for management of endometrial hyperplasia," 2005.
  77. M. Hannemann, H. Alexander, N. Cope, and N. Acheson, "Endometrial hyperplasia: a clinician's review." 2007.
  78. F. A. Tavassoli and P. Devilee, "Pathology and Genetics of Tumours of the Breast and Female Genital Organs," *World Health Organization Classification Tumours series, IARC press*, 2003. .
  79. B. G. Skov, H. Broholm, U. Engel, et al., "Comparison of the reproducibility of the WHO classifications of 1975 and 1994 of endometrial hyperplasia.," *Int. J. Gynecol. Pathol.*, vol. 16, no. 1, pp. 33–7, Jan. 1997.
  80. Robbins and Cotran, *Patología estructural y funcional*. Elsevier, 2015.
  81. R. J. Kurman, P. F. Kaminski, and H. J. Norris, "The behavior of endometrial hyperplasia. A long-term study of &quot;untreated&quot; hyperplasia in 170 patients.," *Cancer*, vol. 56, no. 2, pp. 403–12, Jul. 1985.
  82. B. S. Kendall, B. M. Ronnett, C. Isacson, et al., "Reproducibility of the diagnosis of endometrial hyperplasia, atypical hyperplasia, and well-differentiated carcinoma.," *Am. J. Surg. Pathol.*, vol. 22, no. 8, pp. 1012–9, Aug. 1998.
  83. R. J. Zaino, J. Kauderer, C. L. Trimble, et al., "Reproducibility of the diagnosis of atypical endometrial hyperplasia," *Cancer*, vol. 106, no. 4, pp. 804–811, Feb. 2006.
  84. M. M. Hannemann, H. M. Alexander, N. J. Cope, N. Acheson, and A. Phillips, "Endometrial hyperplasia: a clinician's review," *Obstet. Gynaecol. Reprod. Med.*, vol. 20, no. 4, pp. 116–120, 2010.
  85. R. A. Owings and C. M. Quick, "Endometrial Intraepithelial Neoplasia," *Arch Pathol Lab Med*, vol. 1385858, 2014.
  86. "Endometrial Intraepithelial Neoplasia Committee on Gynecologic Practice Society of Gynecologic Oncology," 2015.

87. J. V Bokhman, "Two pathogenetic types of endometrial carcinoma," *Gynecol. Oncol.*, vol. 15, no. 1, pp. 10–7, Feb. 1983.
88. L. Catasus, P. Machin, X. Matias-Guiu, and J. Prat, "Microsatellite instability in endometrial carcinomas: clinicopathologic correlations in a series of 42 cases.," *Hum. Pathol.*, vol. 29, no. 10, pp. 1160–4, Oct. 1998.
89. G. L. Mutter, "Pten, a protean tumor suppressor.," *Am. J. Pathol.*, vol. 158, no. 6, pp. 1895–8, Jun. 2001.
90. T. Okuda, A. Sekizawa, Y. Purwosunu, et al., "Genetics of endometrial cancers.," *Obstet. Gynecol. Int.*, vol. 2010, p. 984013, 2010.
91. X. Matias-Guiu and J. Prat, "Molecular pathology of endometrial carcinoma," *Histopathology*, vol. 62, pp. 111–123, 2013.
92. P. Morice, A. Leary, C. Creutzberg, N. Abu-Rustum, and E. Darai, "Endometrial cancer," *Lancet*, vol. 387, no. 10023, pp. 1094–1108, 2016.
93. A. Yeramian, G. Moreno-Bueno, X. Dolcet, et al., "Endometrial carcinoma: molecular alterations involved in tumor development and progression," *Oncogene*, vol. 32, no. 4, pp. 403–413, Jan. 2013.
94. S. F. Lax, B. Kendall, H. Tashiro, R. J. Slebos, and L. Hedrick, "The frequency of p53, K-ras mutations, and microsatellite instability differs in uterine endometrioid and serous carcinoma: evidence of distinct molecular genetic pathways.," *Cancer*, vol. 88, no. 4, pp. 814–24, Feb. 2000.
95. N. Buza, D. M. Roque, and A. D. Santin, "HER2/neu in Endometrial Cancer: A Promising Therapeutic Target With Diagnostic Challenges," *Arch. Pathol. Lab. Med.*, vol. 138, no. 3, pp. 343–350, Mar. 2014.
96. O. Fadare, K. Gwin, M. M. Desouki, et al., "The clinicopathologic significance of p53 and BAF-250a (ARID1A) expression in clear cell carcinoma of the endometrium.," *Mod. Pathol.*, vol. 26, no. 8, pp. 1101–10, Aug. 2013.
97. H. C. & Y. R. Kurman RJ, Carcangiu ML, *WHO Classification of Tumours of Female Reproductive Organs.*, 4th editio. WHO, 2014.
98. P. B. Clement and R. H. Young, "Endometrioid carcinoma of the uterine corpus: a review of its pathology with emphasis on recent advances and



- problematic aspects.," *Adv. Anat. Pathol.*, vol. 9, no. 3, pp. 145–84, May 2002.
99. M. E. Sherman, P. Bitterman, N. B. Rosenshein, G. Delgado, and R. J. Kurman, "Uterine serous carcinoma. A morphologically diverse neoplasm with unifying clinicopathologic features.," *Am. J. Surg. Pathol.*, vol. 16, no. 6, pp. 600–10, Jun. 1992.
100. G. Plataniotis, M. Castiglione, and ESMO Guidelines Working Group, "Endometrial cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up," *Ann. Oncol.*, vol. 21, no. Supplement 5, pp. v41–v45, May 2010.
101. The Cancer Genome Atlas Research Network, "Integrated genomic characterization of endometrial carcinoma," 2013.
102. R. Murali, R. A. Soslow, and B. Weigelt, "Classification of endometrial carcinoma: more than two types," *Lancet Oncol.*, vol. 15, no. 7, pp. e268–e278, Jun. 2014.
103. M. Fung-Kee-Fung, J. Dodge, L. Elit, H. Lukka, A. Chambers, T. Oliver, and Cancer Care Ontario Program in Evidence-based Care Gynecology Cancer Disease Site Group, "Follow-up after primary therapy for endometrial cancer: A systematic review," *Gynecol. Oncol.*, vol. 101, no. 3, pp. 520–529, Jun. 2006.
104. F. Vidal and A. Rafii, "Lymph node assessment in endometrial cancer: towards personalized medicine.," *Obstet. Gynecol. Int.*, vol. 2013, p. 892465, 2013.
105. J. D. Wright, N. I. Barrena Medel, J. Sehouli, K. Fujiwara, and T. J. Herzog, "Contemporary management of endometrial cancer.," *Lancet (London, England)*, vol. 379, no. 9823, pp. 1352–60, Apr. 2012.
106. P. B. Panici, S. Basile, F. Maneschi, et al., "Systematic Pelvic Lymphadenectomy vs No Lymphadenectomy in Early-Stage Endometrial Carcinoma: Randomized Clinical Trial," *JNCI J. Natl. Cancer Inst.*, vol. 100, no. 23, pp. 1707–1716, Dec. 2008.
107. ASTEC study group, H. Kitchener, A. M. C. Swart, Q. Qian, C. Amos, and

- M. K. B. Parmar, "Efficacy of systematic pelvic lymphadenectomy in endometrial cancer (MRC ASTEC trial): a randomised study," *Lancet*, vol. 373, no. 9658, pp. 125–136, Jan. 2009.
108. A. J. Fakiris and M. E. Randall, "Endometrial carcinoma: The current role of adjuvant radiation," *J. Obstet. Gynaecol. (Lahore)*, vol. 29, no. 2, pp. 81–89, Jan. 2009.
109. M. Hendrickson, J. Ross, P. Eifel, A. Martinez, and R. Kempson, "Uterine papillary serous carcinoma: a highly malignant form of endometrial adenocarcinoma.," *Am. J. Surg. Pathol.*, vol. 6, no. 2, pp. 93–108, Mar. 1982.
110. M. E. Randall, V. L. Filiaci, H. Muss, N. M. Spirtos, et al., and Gynecologic Oncology Group Study, "Randomized Phase III Trial of Whole-Abdominal Irradiation Versus Doxorubicin and Cisplatin Chemotherapy in Advanced Endometrial Carcinoma: A Gynecologic Oncology Group Study," *J. Clin. Oncol.*, vol. 24, no. 1, pp. 36–44, Jan. 2006.
111. X. Wang, Z. Pan, X. Chen, W. Lü, and X. Xie, "Accuracy of tumor grade by preoperative curettage and associated clinicopathologic factors in clinical stage I endometrioid adenocarcinoma.," *Chin. Med. J. (Engl)*, vol. 122, no. 16, pp. 1843–6, Aug. 2009.
112. A. Karateke, N. Tug, C. Cam, S. Selcuk, M. R. Asoglu, and S. Cakir, "Discrepancy of pre- and postoperative grades of patients with endometrial carcinoma.," *Eur. J. Gynaecol. Oncol.*, vol. 32, no. 3, pp. 283–5, 2011.
113. T. P. Batista, C. L. C. Cavalcanti, A. A. G. Tejo, and A. L. R. Bezerra, "Accuracy of preoperative endometrial sampling diagnosis for predicting the final pathology grading in uterine endometrioid carcinoma.," *Eur. J. Surg. Oncol.*, vol. 42, no. 9, pp. 1367–71, Sep. 2016.
114. G. H. Eltabbakh, J. Shamonki, and S. L. Mount, "Surgical stage, final grade, and survival of women with endometrial carcinoma whose preoperative endometrial biopsy shows well-differentiated tumors.," *Gynecol. Oncol.*, vol. 99, no. 2, pp. 309–12, Nov. 2005.
115. L. Helpman, R. Kupets, A. Covens, et al., "Assessment of endometrial

- sampling as a predictor of final surgical pathology in endometrial cancer.," *Br. J. Cancer*, vol. 110, no. 3, pp. 609–15, Feb. 2014.
116. N. C. M. Visser, M. C. Breijer, M. C. Herman, et al., "Factors attributing to the failure of endometrial sampling in women with postmenopausal bleeding.," *Acta Obstet. Gynecol. Scand.*, vol. 92, no. 10, pp. 1216–22, Oct. 2013.
117. S. Nofech-Mozes, N. Ismiil, V. Dubé, et al., "Interobserver agreement for endometrial cancer characteristics evaluated on biopsy material.," *Obstet. Gynecol. Int.*, vol. 2012, p. 414086, 2012.
118. A. Mota, E. Colás, P. García-Sanz, et al., "Genetic analysis of uterine aspirates improves the diagnostic value and captures the intra-tumor heterogeneity of endometrial cancers," *Mod. Pathol.*, vol. 30, no. 1, pp. 134–145, Jan. 2017.
119. L. Savelli, M. Ceccarini, M. Ludovisi, et al., "Preoperative local staging of endometrial cancer: transvaginal sonography vs. magnetic resonance imaging," *Ultrasound Obstet. Gynecol.*, vol. 31, no. 5, pp. 560–566, May 2008.
120. S. Cabrita, H. Rodrigues, R. Abreu, et al., "Magnetic resonance imaging in the preoperative staging of endometrial carcinoma.," *Eur. J. Gynaecol. Oncol.*, vol. 29, no. 2, pp. 135–7, 2008.
121. O. Akbayir, A. Corbacioglu, C. Numanoglu et al., "Preoperative assessment of myometrial and cervical invasion in endometrial carcinoma by transvaginal ultrasound," *Gynecol. Oncol.*, vol. 122, no. 3, pp. 600–603, Sep. 2011.
122. S. Jantarsaengaram, N. Praditphol, T. Tansathit, C. Vipupinyo, and K. Vairojanavong, "Three-dimensional ultrasound with volume contrast imaging for preoperative assessment of myometrial invasion and cervical involvement in women with endometrial cancer," *Ultrasound Obstet. Gynecol.*, vol. 43, no. 5, pp. 569–574, May 2014.
123. C.-H. Lai, G. Lin, T.-C. Yen, and F.-Y. Liu, "Molecular imaging in the management of gynecologic malignancies," *Gynecol. Oncol.*, vol. 135, no.

- 1, pp. 156–162, Oct. 2014.
124. M. R. Ponisio, K. J. Fowler, and F. Dehdashti, “The Emerging Role of PET/MR Imaging in Gynecologic Cancers,” *PET Clin.*, vol. 11, no. 4, pp. 425–440, Oct. 2016.
125. O. Nørregaard Jensen, “Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry,” *Curr. Opin. Chem. Biol.*, vol. 8, no. 1, pp. 33–41, Feb. 2004.
126. P. Legrain, R. Aebersold, A. Archakov, et al., “The Human Proteome Project: Current State and Future Direction,” *Mol. Cell. Proteomics*, vol. 10, no. 7, p. M111.009993, Jul. 2011.
127. N. L. Anderson and N. G. Anderson, “The human plasma proteome: history, character, and diagnostic prospects.,” *Mol. Cell. Proteomics*, vol. 1, no. 11, pp. 845–67, Nov. 2002.
128. N. Rifai, M. A. Gillette, and S. A. Carr, “Protein biomarker discovery and validation: the long and uncertain path to clinical utility.,” *Nat. Biotechnol.*, vol. 24, no. 8, pp. 971–83, Aug. 2006.
129. P. Teng, N. W. Bateman, B. L. Hood, and T. P. Conrads, “Advances in proximal fluid proteomics for disease biomarker discovery.,” *J. Proteome Res.*, vol. 9, no. 12, pp. 6091–100, Dec. 2010.
130. E. Aasebø, J. A. Opsahl, Y. Bjørlykke, K.-M. Myhr, A. C. Kroksveen, and F. S. Berven, “Effects of blood contamination and the rostro-caudal gradient on the human cerebrospinal fluid proteome.,” *PLoS One*, vol. 9, no. 3, p. e90429, Mar. 2014.
131. A. P. Drabovich, E. Martínez-Morillo, and E. P. Diamandis, “Toward an integrated pipeline for protein biomarker development,” *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1854, no. 6, pp. 677–686, Jun. 2015.
132. C. Perez-Sanchez, E. Colas, S. Cabrera, et al., “Molecular diagnosis of endometrial cancer from uterine aspirates,” *Int. J. Cancer*, vol. 133, no. 10, pp. 2383–2391, Nov. 2013.
133. J. Casado-Vela, E. Rodriguez-Suarez, I. Iloro, et al., “Comprehensive proteomic analysis of human endometrial fluid aspirate,” *J. Proteome Res.*,

- vol. 8, no. 10, pp. 4622–4632, Oct. 2009.
134. A. Ametzazurra, R. Matorras, J. A. García-Velasco, et al., “Endometrial fluid is a specific and non-invasive biological sample for protein biomarker identification in endometriosis,” *Hum. Reprod.*, vol. 24, no. 4, pp. 954–965, 2009.
  135. L. A. Salamonsen, T. Edgell, L. J. F. Rombauts, et al., “Proteomics of the human endometrium and uterine fluid: a pathway to biomarker discovery.,” *Fertil. Steril.*, vol. 99, no. 4, pp. 1086–92, Mar. 2013.
  136. G. Raposo and W. Stoorvogel, “Extracellular vesicles: exosomes, microvesicles, and friends.,” *J. Cell Biol.*, vol. 200, no. 4, pp. 373–83, Feb. 2013.
  137. M. Simons and G. Raposo, “Exosomes-vesicular carriers for intercellular communication.,” *Curr. Opin. Cell Biol.*, vol. 21, no. 4, pp. 575–81, Aug. 2009.
  138. C. Ciardiello, L. Cavallini, C. Spinelli, et al., “Focus on Extracellular Vesicles: New Frontiers of Cell-to-Cell Communication in Cancer.,” *Int. J. Mol. Sci.*, vol. 17, no. 2, p. 175, Jan. 2016.
  139. C. Harding, J. Heuser, and P. Stahl, “Receptor-mediated endocytosis of transferrin and recycling of the transferrin receptor in rat reticulocytes.,” *J. Cell Biol.*, vol. 97, no. 2, pp. 329–39, Aug. 1983.
  140. B. T. Pan and R. M. Johnstone, “Fate of the transferrin receptor during maturation of sheep reticulocytes in vitro: selective externalization of the receptor.,” *Cell*, vol. 33, no. 3, pp. 967–78, Jul. 1983.
  141. G. Raposo, H. W. Nijman, W. Stoorvogel, et al., “B lymphocytes secrete antigen-presenting vesicles.,” *J. Exp. Med.*, vol. 183, no. 3, pp. 1161–72, Mar. 1996.
  142. L. Zitvogel, A. Regnault, A. Lozier, et al., “Eradication of established murine tumors using a novel cell-free vaccine: dendritic cell-derived exosomes.,” *Nat. Med.*, vol. 4, no. 5, pp. 594–600, May 1998.
  143. C. Théry, M. Ostrowski, and E. Segura, “Membrane vesicles as conveyors of immune responses.,” *Nat. Rev. Immunol.*, vol. 9, no. 8, pp. 581–93, Aug.

- 2009.
144. A. Bobrie, M. Colombo, G. Raposo, and C. Théry, "Exosome Secretion: Molecular Mechanisms and Roles in Immune Responses," *Traffic*, vol. 12, no. 12, pp. 1659–1668, Dec. 2011.
  145. M. Yáñez-Mó, P. R.-M. Siljander, Z. Andreu, et al., "Biological properties of extracellular vesicles and their physiological functions.," *J. Extracell. vesicles*, vol. 4, p. 27066, 2015.
  146. R. J. Simpson and S. Mathivanan, "Extracellular Microvesicles: The Need for Internationally Recognised Nomenclature and Stringent Purification Criteria," *J. Proteomics Bioinform.*, vol. 5, no. 2, 2012.
  147. S. J. Gould and G. Raposo, "As we wait: coping with an imperfect nomenclature for extracellular vesicles.," *J. Extracell. vesicles*, vol. 2, 2013.
  148. H. Kalra, G. Drummen, and S. Mathivanan, "Focus on Extracellular Vesicles: Introducing the Next Small Big Thing," *Int. J. Mol. Sci.*, vol. 17, no. 2, p. 170, Feb. 2016.
  149. C. Théry, S. Amigorena, G. Raposo, and A. Clayton, "Isolation and characterization of exosomes from cell culture supernatants and biological fluids," *Curr. Protoc. cell Biol. / Editor. board, Juan S Bonifacino [et al]*, vol. Chapter 3, p. Unit 3.22, 2006.
  150. H. Kalra, C. G. Adda, M. Liem, et al., "Comparative proteomics evaluation of plasma exosome isolation techniques and assessment of the stability of exosomes in normal human blood plasma," *Proteomics*, vol. 13, no. 22, 2013.
  151. B. J. Tauro, D. W. Greening, R. A. Mathias, et al., "Comparison of ultracentrifugation, density gradient separation, and immunoaffinity capture methods for isolating human colon cancer cell line LIM1863-derived exosomes," *Methods*, vol. 56, no. 2, pp. 293–304, Feb. 2012.
  152. A. Cvjetkovic, J. Lötvall, and C. Lässer, "The influence of rotor type and centrifugation time on the yield and purity of extracellular vesicles.," *J. Extracell. vesicles*, vol. 3, no. 1, p. 23111, Jan. 2014.
  153. D. D. Taylor, W. Zacharias, and C. Gerceel-Taylor, "Exosome Isolation for

- Proteomic Analyses and RNA Profiling,” in *Methods in molecular biology (Clifton, N.J.)*, vol. 728, 2011, pp. 235–246.
154. R. Grant, E. Ansa-Addo, D. Stratton, et al., “A filtration-based protocol to isolate human plasma membrane-derived vesicles and exosomes from blood plasma,” *J. Immunol. Methods*, vol. 371, no. 1–2, pp. 143–51, Aug. 2011.
  155. S. Boukouris and S. Mathivanan, “Exosomes in bodily fluids are a highly stable resource of disease biomarkers,” *Proteomics. Clin. Appl.*, vol. 9, no. 3–4, pp. 358–67, Apr. 2015.
  156. E. Willms, H. J. Johansson, I. Mäger, et al., “Cells release subpopulations of exosomes with distinct molecular and biological properties,” *Sci. Rep.*, vol. 6, p. 22519, Mar. 2016.
  157. S. Fais, F. E. Borrás, E. Buzas, et al., “Evidence-Based Clinical Use of Nanoscale Extracellular Vesicles in Nanomedicine,” *ACS Nano*, vol. 10, pp. 3886–3899, 2016.
  158. C. Soekmadji, P. J. Russell, and C. C. Nelson, “Exosomes in prostate cancer: putting together the pieces of a puzzle,” *Cancers (Basel)*, vol. 5, no. 4, pp. 1522–44, Nov. 2013.
  159. R. A. Dragovic, C. Gardiner, A. S. Brooks, et al., “Sizing and phenotyping of cellular vesicles using Nanoparticle Tracking Analysis,” *Nanomedicine*, vol. 7, no. 6, pp. 780–8, Dec. 2011.
  160. B. György, T. G. Szabó, M. Pásztói, et al., “Membrane vesicles, current state-of-the-art: emerging role of extracellular vesicles,” *Cell. Mol. Life Sci.*, vol. 68, no. 16, pp. 2667–2688, Aug. 2011.
  161. E. van der Pol, A. N. Böing, P. Harrison, A. Sturk, and R. Nieuwland, “Classification, functions, and clinical relevance of extracellular vesicles,” *Pharmacol. Rev.*, vol. 64, no. 3, pp. 676–705, Jul. 2012.
  162. C. Théry, S. Amigorena, G. Raposo, and A. Clayton, “Isolation and characterization of exosomes from cell culture supernatants and biological fluids,” *Curr Protoc Cell Biol*, vol. Chapter 3, p. Unit 3 22, 2006.
  163. J. C. Akers, D. Gonda, R. Kim, B. S. Carter, and C. C. Chen, “Biogenesis of

- extracellular vesicles (EV): exosomes, microvesicles, retrovirus-like vesicles, and apoptotic bodies.,” *J. Neurooncol.*, vol. 113, no. 1, pp. 1–11, May 2013.
164. M. S. Pols and J. Klumperman, “Trafficking and function of the tetraspanin CD63,” *Exp. Cell Res.*, vol. 315, no. 9, pp. 1584–1592, May 2009.
165. C. Raiborg and H. Stenmark, “The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins,” *Nature*, vol. 458, no. 7237, pp. 445–452, Mar. 2009.
166. J. H. Hurley, “The ESCRT complexes,” *Crit. Rev. Biochem. Mol. Biol.*, vol. 45, no. 6, pp. 463–487, Dec. 2010.
167. M. F. Baietti, Z. Zhang, E. Mortier, et al., “Syndecan–syntenin–ALIX regulates the biogenesis of exosomes,” *Nat. Cell Biol.*, vol. 14, no. 7, pp. 677–685, Jun. 2012.
168. J. F. Nabhan, R. Hu, R. S. Oh, S. N. Cohen, and Q. Lu, “Formation and release of arrestin domain-containing protein 1-mediated microvesicles (ARMMs) at plasma membrane by recruitment of TSG101 protein,” *Proc. Natl. Acad. Sci.*, vol. 109, no. 11, pp. 4146–4151, Mar. 2012.
169. S. Stuffers, C. Sem Wegner, H. Stenmark, and A. Brech, “Multivesicular Endosome Biogenesis in the Absence of ESCRTs,” *Traffic*, vol. 10, no. 7, pp. 925–937, Jul. 2009.
170. G. van Niel, S. Charrin, S. Simoes, et al., “The tetraspanin CD63 regulates ESCRT-independent and -dependent endosomal sorting during melanogenesis.,” *Dev. Cell*, vol. 21, no. 4, pp. 708–21, Oct. 2011.
171. H. Matsuo, J. Chevallier, N. Mayran, et al., “Role of LBPA and Alix in Multivesicular Liposome Formation and Endosome Organization,” *Science (80-. )*, vol. 303, no. 5657, pp. 531–534, Jan. 2004.
172. D. J. Gillooly, I. C. Morrow, M. Lindsay, et al., “Localization of phosphatidylinositol 3-phosphate in yeast and mammalian cells,” *EMBO J.*, vol. 19, no. 17, pp. 4577–4588, Sep. 2000.
173. T. Kobayashi, E. Stang, K. S. Fang, P. de Moerloose, R. G. Parton, and J. Gruenberg, “A lipid associated with the antiphospholipid syndrome



- regulates endosome structure and function.,” *Nature*, vol. 392, no. 6672, pp. 193–7, Mar. 1998.
174. K. Trajkovic, C. Hsu, S. Chiantia, et al., “Ceramide Triggers Budding of Exosome Vesicles into Multivesicular Endosomes,” *Science (80-. )*, vol. 319, no. 5867, pp. 1244–1247, Feb. 2008.
175. N. Bishop and P. Woodman, “ATPase-defective mammalian VPS4 localizes to aberrant endosomes and impairs cholesterol trafficking.,” *Mol. Biol. Cell*, vol. 11, no. 1, pp. 227–39, Jan. 2000.
176. X. Du, A. S. Kazim, I. W. Dawes, A. J. Brown, and H. Yang, “The AAA ATPase VPS4/SKD1 Regulates Endosomal Cholesterol Trafficking Independently of ESCRT-III,” *Traffic*, vol. 14, no. 1, pp. 107–119, Jan. 2013.
177. X. Du, A. S. Kazim, A. J. Brown, and H. Yang, “An essential role of Hrs/Vps27 in endosomal cholesterol trafficking.,” *Cell Rep.*, vol. 1, no. 1, pp. 29–35, Jan. 2012.
178. A. de Gassart, C. Geminard, B. Fevrier, G. Raposo, and M. Vidal, “Lipid raft-associated protein sorting in exosomes,” *Blood*, vol. 102, no. 13, pp. 4336–4344, Dec. 2003.
179. J. R. Edgar, E. R. Eden, and C. E. Futter, “Hrs- and CD63-dependent competing mechanisms make different sized endosomal intraluminal vesicles.,” *Traffic*, vol. 15, no. 2, pp. 197–211, Feb. 2014.
180. S. Keerthikumar, D. Chisanga, D. Ariyaratne, et al., “ExoCarta: A Web-Based Compendium of Exosomal Cargo,” *J. Mol. Biol.*, Oct. 2015.
181. A. Michael, S. Bajracharya, P. Yuen, et al., “Exosomes from human saliva as a source of microRNA biomarkers,” *Oral Dis.*, vol. 16, no. 1, pp. 34–38, Jan. 2010.
182. M. W. Graner, O. Alzate, A. M. Dechkovskaia, J et al., “Proteomic and immunologic analyses of brain tumor exosomes.,” *FASEB J.*, vol. 23, no. 5, pp. 1541–57, May 2009.
183. F. Andre, N. E. C. Scharz, M. Movassagh, et al., “Malignant effusions and immunogenic tumour-derived exosomes.,” *Lancet (London, England)*, vol. 360, no. 9329, pp. 295–305, Jul. 2002.

184. J. H. Hurley and P. I. Hanson, "Membrane budding and scission by the ESCRT machinery: it's all in the neck.," *Nat. Rev. Mol. Cell Biol.*, vol. 11, no. 8, pp. 556–66, Aug. 2010.
185. S. I. Buschow, J. M. P. Liefhebber, R. Wubbolts, and W. Stoorvogel, "Exosomes contain ubiquitinated proteins," *Blood Cells, Mol. Dis.*, vol. 35, no. 3, pp. 398–403, Nov. 2005.
186. D. Perez-Hernandez, C. Gutiérrez-Vázquez, I. Jorge, et al., "The intracellular interactome of tetraspanin-enriched microdomains reveals their function as sorting machineries toward exosomes.," *J. Biol. Chem.*, vol. 288, no. 17, pp. 11649–61, Apr. 2013.
187. A. Calzolari, C. Raggi, S. Deaglio, et al., "TfR2 localizes in lipid raft domains and is released in exosomes to activate signal transduction along the MAPK pathway.," *J. Cell Sci.*, vol. 119, no. Pt 21, pp. 4486–98, Nov. 2006.
188. M. Ostrowski, N. B. Carmo, S. Krumeich, et al., "Rab27a and Rab27b control different steps of the exosome secretion pathway.," *Nat. Cell Biol.*, vol. 12, no. 1, pp. 19–30–13, 2010.
189. A. Savina, M. Vidal, and M. I. Colombo, "The exosome pathway in K562 cells is regulated by Rab11.," *J. Cell Sci.*, vol. 115, no. Pt 12, pp. 2505–15, Jun. 2002.
190. C. Hsu, Y. Morohashi, S.-I. Yoshimura, et al., "Regulation of exosome secretion by Rab35 and its GTPase-activating proteins TBC1D10A-C.," *J. Cell Biol.*, vol. 189, no. 2, pp. 223–32, Apr. 2010.
191. G. Hu, A.-Y. Gong, A. L. Roth, et al., "Release of Luminal Exosomes Contributes to TLR4-Mediated Epithelial Antimicrobial Defense," *PLoS Pathog.*, vol. 9, no. 4, p. e1003261, Apr. 2013.
192. K. Koles, J. Nunnari, C. Korkut, et al., "Mechanism of Evenness Interrupted (Evi)-Exosome Release at Synaptic Boutons," *J. Biol. Chem.*, vol. 287, no. 20, pp. 16820–16834, May 2012.
193. L. A. Mulcahy, R. C. Pink, and D. R. F. Carter, "Routes and mechanisms of extracellular vesicle uptake," *J. Extracell. Vesicles*, vol. 3, no. 1, p. 24641, Jan. 2014.

194. G. Van Niel, J. Mallegol, C. Bevilacqua, et al., "Intestinal epithelial exosomes carry MHC class II/peptides able to inform the immune system in mice.," *Gut*, vol. 52, no. 12, pp. 1690–7, Dec. 2003.
195. A. V. Vlassov, S. Magdaleno, R. Setterquist, and R. Conrad, "Exosomes: Current knowledge of their composition, biological functions, and diagnostic and therapeutic potentials," *Biochim. Biophys. Acta - Gen. Subj.*, vol. 1820, no. 7, pp. 940–948, Jul. 2012.
196. C. Corrado, S. Raimondo, A. Chiesi, et al., "Exosomes as Intercellular Signaling Organelles Involved in Health and Disease: Basic Science and Clinical ApplicationsExosomes as Intercellular Signaling Organelles Involved in Health and Disease: Basic Science and Clinical Applications," *Int. J. Mol. Sci*, vol. 14, pp. 5338–5366, 2013.
197. A. Lakkaraju and E. Rodriguez-Boulan, "Itinerant exosomes: emerging roles in cell and tissue polarity.," *Trends Cell Biol.*, vol. 18, no. 5, pp. 199–209, May 2008.
198. J. Gobbo, G. Marcion, M. Cordonnier, et al., "Restoring Anticancer Immune Response by Targeting Tumor-Derived Exosomes With a HSP70 Peptide Aptamer.," *J. Natl. Cancer Inst.*, vol. 108, no. 3, p. djv330, Mar. 2016.
199. D. Yu, Y. Wu, H. Shen, et al., "Exosomes in development, metastasis and drug resistance of breast cancer," *Cancer Sci.*, vol. 106, no. 8, pp. 959–964, Aug. 2015.
200. P. Kharaziha, S. Ceder, Q. Li, and T. Panaretakis, "Tumor cell-derived exosomes: a message in a bottle.," *Biochim. Biophys. Acta*, vol. 1826, no. 1, pp. 103–11, Aug. 2012.
201. A. S. Azmi, B. Bao, and F. H. Sarkar, "Exosomes in cancer development, metastasis, and drug resistance: a comprehensive review," *Cancer Metastasis Rev.*, vol. 32, no. 3–4, pp. 623–642, Dec. 2013.
202. D. Hanahan and R. A. Weinberg, "Hallmarks of Cancer: The Next Generation," *Cell*, vol. 144, no. 5, pp. 646–674, 2011.
203. M. Logozzi, A. De Milito, L. Lugini, et al., "High levels of exosomes expressing CD63 and caveolin-1 in plasma of melanoma patients.," *PLoS*

- One*, vol. 4, no. 4, p. e5219, Apr. 2009.
204. P. Sánchez-Vela, N. A. Garcia, M. Campos-Segura, and J. Forteza-Vila, "Liquid biopsy and tumor derived exosomes in clinical practice," *Rev. Española Patol.*, vol. 49, no. 2, pp. 106–111, Apr. 2016.
205. C. Rolfo, M. Castiglia, D. Hong, et al., "Liquid biopsies in lung cancer: the new ambrosia of researchers.," *Biochim. Biophys. Acta*, vol. 1846, no. 2, pp. 539–46, Dec. 2014.
206. S. Boukouris and S. Mathivanan, "Exosomes in bodily fluids are a highly stable resource of disease biomarkers.," *Proteomics. Clin. Appl.*, vol. 9, no. 3–4, pp. 358–67, Apr. 2015.
207. G. Pocsfalvi, C. Stanly, A. Vilasi, I. Fiume, G. Capasso, L. Turiák, E. I. Buzas, and K. Vékey, "Mass spectrometry of extracellular vesicles," *Mass Spectrom. Rev.*, vol. 35, no. 1, pp. 3–21, Jan. 2016.
208. S. Kreimer, A. M. Belov, I. Ghiran, S. K. Murthy, D. a. Frank, and A. R. Ivanov, "Mass-Spectrometry-Based Molecular Characterization of Extracellular Vesicles: Lipidomics and Proteomics," *J. Proteome Res.*, p. 150514113430005, 2015.
209. F. Simona, S. Laura, T. Simona, and A. Riccardo, "Contribution of proteomics to understanding the role of tumor-derived exosomes in cancer progression: State of the art and new perspectives," *Proteomics*, vol. 13, pp. 1581–1594, 2013.
210. C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis.," *Nat. Methods*, vol. 9, no. 7, pp. 671–5, Jul. 2012.
211. L. L. Manza, S. L. Stamer, A.-J. L. Ham, S. G. Codreanu, and D. C. Liebler, "Sample preparation and digestion for proteomic analyses using spin filters," *Proteomics*, vol. 5, no. 7, pp. 1742–1745, May 2005.
212. H. Mi, A. Muruganujan, J. T. Casagrande, and P. D. Thomas, "Large-scale gene function analysis with the PANTHER classification system," *Nat. Protoc.*, vol. 8, no. 8, pp. 1551–1566, Jul. 2013.
213. M. Pathan, S. Keerthikumar, C-S. Ang, et al., "FunRich: An open access

- standalone functional enrichment and interaction network analysis tool,” *Proteomics*, vol. 15, no. 15, pp. 2597–2601, Aug. 2015.
214. I. Campoy, L. Lanau, T. Altadill, et al., “Exosome-like vesicles in uterine aspirates: a comparison of ultracentrifugation-based isolation protocols,” *J. Transl. Med.*, vol. 14, no. 1, p. 180, Dec. 2016.
215. T. Geiger, J. Cox, P. Ostasiewicz, J. R. Wisniewski, and M. Mann, “Super-SILAC mix for quantitative proteomics of human tumor tissue.,” *Nat. Methods*, vol. 7, no. 5, pp. 383–5, May 2010.
216. R. Etzioni, C. Kooperberg, M. Pepe, R. Smith, and P. H. Gann, “Combining biomarkers to detect disease with application to prostate cancer.,” *Biostatistics*, vol. 4, no. 4, pp. 523–38, Oct. 2003.
217. E. Martinez-Garcia, A. Lesur, L. Devis, et al., “Development of a sequential workflow based on LC-PRM for the verification of endometrial cancer protein biomarkers in uterine aspirate samples,” *Oncotarget*, vol. 7, no. 33, pp. 53102–53115, 2016.
218. N. Van Hanegem, M. C. Breijer, K. S. Khan, et al., “Diagnostic evaluation of the endometrium in postmenopausal bleeding: An evidence-based approach,” *Maturitas*, vol. 68, pp. 155–164, 2010.
219. N. C. M. Visser, J. Bulten, A. A. M. van der Wurff, et al., “PIpelle Prospective ENDOmetrial carcinoma (PIPENDO) study, pre-operative recognition of high risk endometrial carcinoma: a multicentre prospective cohort study.,” *BMC Cancer*, vol. 15, p. 487, Jan. 2015.
220. H. Saji, K. Kurose, K. Sugiura, et al., “Endometrial aspiration cytology for diagnosis of peritoneal lesions in extrauterine malignancies.,” *Acta Cytol.*, vol. 51, no. 4, pp. 533–40.
221. I. Kinde, C. Bettgowda, Y. Wang, et al., “Evaluation of DNA from the Papanicolaou Test to Detect Ovarian and Endometrial Cancers,” *Sci. Transl. Med.*, vol. 5, no. 167, p. 167ra4-167ra4, Jan. 2013.
222. E. Colas, C. Perez, S. Cabrera, et al., “Molecular markers of endometrial carcinoma detected in uterine aspirates,” *Int. J. Cancer*, vol. 129, no. 10, pp. 2435–2444, Nov. 2011.

223. M. Esteller, A. García, J. M. Martínez-Palones, J. Xercavins, and J. Reventós, "Detection of clonality and genetic alterations in endometrial pipelle biopsy and its surgical specimen counterpart," *Lab. Invest.*, vol. 76, no. 1, pp. 109–116, Jan. 1997.
224. N. J. Hannan, C. J. Stoikos, A. N. Stephens, and L. A. Salamonsen, "Depletion of high-abundance serum proteins from human uterine lavages enhances detection of lower-abundance proteins," *J. Proteome Res.*, vol. 8, no. 2, pp. 1099–1103, Feb. 2009.
225. C. Harding and P. Stahl, "Transferrin recycling in reticulocytes: pH and iron are important determinants of ligand binding and processing.," *Biochem. Biophys. Res. Commun.*, vol. 113, no. 2, pp. 650–8, Jun. 1983.
226. R. M. Johnstone, M. Adam, J. R. Hammond, L. Orr, and C. Turbide, "Vesicle formation during reticulocyte maturation. Association of plasma membrane activities with released vesicles (exosomes).," *J. Biol. Chem.*, vol. 262, no. 19, pp. 9412–20, Jul. 1987.
227. C. Bang and T. Thum, "Exosomes: new players in cell-cell communication.," *Int. J. Biochem. Cell Biol.*, vol. 44, no. 11, pp. 2060–4, Nov. 2012.
228. A.-K. Ludwig and B. Giebel, "Exosomes: small vesicles participating in intercellular communication.," *Int. J. Biochem. Cell Biol.*, vol. 44, no. 1, pp. 11–5, Jan. 2012.
229. I. V. Miller and T. G. P. Grunewald, "Tumour-derived exosomes: Tiny envelopes for big stories.," *Biol. Cell*, vol. 107, no. 9, pp. 287–305, Sep. 2015.
230. N. Kosaka, H. Iguchi, and T. Ochiya, "Circulating microRNA in body fluid: a new potential biomarker for cancer diagnosis and prognosis.," *Cancer Sci.*, vol. 101, no. 10, pp. 2087–92, Oct. 2010.
231. A. Gámez-Valero, S. I. Lozano-Ramos, I. Bancu, R. Lauzurica-Valdemoros, and F. E. Borràs, "Urinary extracellular vesicles as source of biomarkers in kidney diseases," *Front. Immunol.*, vol. 6, p. 6, 2015.
232. C. Lässer, V. S. Alikhani, K. Ekström, et al., "Human saliva, plasma and

- breast milk exosomes contain RNA: uptake by macrophages.," *J. Transl. Med.*, vol. 9, p. 9, Jan. 2011.
233. M. F. Peterson, N. Otoc, J. K. Sethi, A. Gupta, and T. J. Antes, "Integrated systems for exosome investigation," *Methods*, Apr. 2015.
234. N. J. Hannan, A. N. Stephens, A. Rainczuk, et al., "2D-DiGE analysis of the human endometrial secretome reveals differences between receptive and nonreceptive states in fertile and infertile women," *J. Proteome Res.*, vol. 9, no. 12, pp. 6256–6264, Dec. 2010.
235. J. G. Scotchie, M. A. Fritz, M. Mocanu, B. A. Lessey, and S. L. Young, "Proteomic analysis of the luteal endometrial secretome," *Reprod. Sci.*, vol. 16, no. 9, pp. 883–893, Sep. 2009.
236. C. M. Boomsma, A. Kavelaars, M. J. C. Eijkemans, et al., "Endometrial secretion analysis identifies a cytokine profile predictive of pregnancy in IVF," *Hum. Reprod.*, vol. 24, no. 6, pp. 1427–1435, Jun. 2009.
237. F. Vilella, J. M. Moreno-Moya, N. Balaguer, et al., "Hsa-miR-30d, secreted by the human endometrium, is taken up by the pre-implantation embryo and might modify its transcriptome.," *Development*, vol. 142, no. 18, pp. 3210–21, Sep. 2015.
238. B. Levänen, N. R. Bhakta, P. Torregrosa Paredes, et al., "Altered microRNA profiles in bronchoalveolar lavage fluid exosomes in asthmatic patients.," *J. Allergy Clin. Immunol.*, vol. 131, no. 3, pp. 894–903, Mar. 2013.
239. J. Nilsson, J. Skog, A. Nordstrand, et al., "Prostate cancer-derived urine exosomes: a novel approach to biomarkers for prostate cancer.," *Br. J. Cancer*, vol. 100, no. 10, pp. 1603–7, May 2009.
240. P. Mittal, M. Klingler-Hoffmann, G. Arentz, et al., "Proteomics of endometrial cancer diagnosis, treatment, and prognosis," *PROTEOMICS - Clin. Appl.*, vol. 10, no. 3, pp. 217–229, Mar. 2016.
241. Y. Ueda, T. Enomoto, T. Kimura, et al., "Serum Biomarkers for Early Detection of Gynecologic Cancers," *Cancers (Basel)*, vol. 2, no. 2, pp. 1312–1327, Jun. 2010.

242. J. L. Powell, K. A. Hill, B. C. Shiro, S. J. Diehl, and W. H. Gajewski, "Preoperative serum CA-125 levels in treating endometrial cancer.," *J. Reprod. Med.*, vol. 50, no. 8, pp. 585–90, Aug. 2005.
243. H. Hareyama, N. Sakuragi, S. Makinoda, and S. Fujimoto, "Serum and tissue measurements of CA72-4 in patients with endometrial carcinoma.," *J. Clin. Pathol.*, vol. 49, no. 12, pp. 967–70, Dec. 1996.
244. R. Konno, T. Takano, S. Sato, and A. Yajima, "Serum soluble fas level as a prognostic factor in patients with gynecological malignancies.," *Clin. Cancer Res.*, vol. 6, no. 9, pp. 3576–80, Sep. 2000.
245. D. J. Brennan, A. Hackethal, A. M. Metcalf, et al., "Serum HE4 as a prognostic marker in endometrial cancer — A population based study," *Gynecol. Oncol.*, vol. 132, no. 1, pp. 159–165, Jan. 2014.
246. I. B. ENGELSEN, L. A. AKSLEN, and H. B. SALVESEN, "Biologic markers in endometrial cancer treatment," *APMIS*, vol. 117, no. 10, pp. 693–707, Oct. 2009.
247. A. G. Zeimet, D. Reimer, M. Huszar, et al., "L1CAM in early-stage type I endometrial cancer: results of a large multicenter evaluation.," *J. Natl. Cancer Inst.*, vol. 105, no. 15, pp. 1142–50, Aug. 2013.
248. M. Lambropoulou, N. Papadopoulos, G. Tripsianis, et al., "Co-expression of survivin, c-erbB2, and cyclooxygenase-2 (COX-2): prognostic value and survival of endometrial cancer patients.," *J. Cancer Res. Clin. Oncol.*, vol. 136, no. 3, pp. 427–35, Mar. 2010.
249. P. Dong, M. Kaneuchi, Y. Konno, et al., "Emerging Therapeutic Biomarkers in Endometrial Cancer," *Biomed Res. Int.*, vol. 2013, pp. 1–11, 2013.
250. H. Ge, A. J. M. Walhout, and M. Vidal, "Integrating 'omic' information: a bridge between genomics and systems biology.," *Trends Genet.*, vol. 19, no. 10, pp. 551–60, Oct. 2003.
251. W. Tian, Y. Zhu, Y. Wang, et al., "Visfatin, a potential biomarker and prognostic factor for endometrial cancer.," *Gynecol. Oncol.*, vol. 129, no. 3, pp. 505–12, Jun. 2013.
252. K. Tanabe, M. Matsumoto, S. Ikematsu, et al., "Midkine and its clinical



- significance in endometrial carcinoma.,” *Cancer Sci.*, vol. 99, no. 6, pp. 1125–30, Jun. 2008.
253. Z. Yurkovetsky, S. Ta’asan, S. Skates, et al., “Development of multimarker panel for early detection of endometrial cancer. High diagnostic power of prolactin.,” *Gynecol. Oncol.*, vol. 107, no. 1, pp. 58–65, Oct. 2007.
254. E. Isci Bostanci, A. Ugras Dikmen, G. Girgin, et al., “A New Diagnostic and Prognostic Marker in Endometrial Cancer,” *Int. J. Gynecol. Cancer*, vol. 27, no. 4, pp. 754–758, May 2017.
255. H. P. T. Nguyen, R. J. Simpson, L. A. Salamonsen, and D. W. Greening, “Extracellular Vesicles in the Intrauterine Environment: Challenges and Potential Functions.,” *Biol. Reprod.*, vol. 95, no. 5, p. 109, Nov. 2016.
256. G. Burns, K. Brooks, M. Wildung, et al., “Extracellular Vesicles in Luminal Fluid of the Ovine Uterus,” *PLoS One*, vol. 9, no. 3, p. e90913, Mar. 2014.
257. G. W. Burns, K. E. Brooks, and T. E. Spencer, “Extracellular Vesicles Originate from the Conceptus and Uterus During Early Pregnancy in Sheep,” *Biol. Reprod.*, vol. 94, no. 3, p. 56, Mar. 2016.
258. G. S. Griffiths, D. S. Galileo, K. Reese, and P. A. Martin-Deleon, “Investigating the role of murine epididymosomes and uterosomes in GPI-linked protein transfer to sperm using SPAM1 as a model.,” *Mol. Reprod. Dev.*, vol. 75, no. 11, pp. 1627–36, Nov. 2008.
259. Y. H. Ng, S. Rome, A. Jalabert, A. Forterre, H. Singh, C. L. Hincks, and L. A. Salamonsen, “Endometrial exosomes/microvesicles in the uterine microenvironment: a new paradigm for embryo-endometrial cross talk at implantation.,” *PLoS One*, vol. 8, no. 3, p. e58502, Jan. 2013.
260. A. G. Braundmeier, C. A. Dayger, P. Mehrotra, et al., “EMMPRIN is secreted by human uterine epithelial cells in microvesicles and stimulates metalloproteinase production by human uterine fibroblast cells.,” *Reprod. Sci.*, vol. 19, no. 12, pp. 1292–301, Dec. 2012.
261. L. A. Burnett, M. M. Light, P. Mehrotra, and R. A. Nowak, “Stimulation of GPR30 increases release of EMMPRIN-containing microvesicles in human uterine epithelial cells.,” *J. Clin. Endocrinol. Metab.*, vol. 97, no. 12, pp.

- 4613–22, Dec. 2012.
262. D. W. Greening, H. P. T. Nguyen, K. Elgass, R. J. Simpson, and L. A. Salamonsen, “Human Endometrial Exosomes Contain Hormone-Specific Cargo Modulating Trophoblast Adhesive Capacity: Insights into Endometrial-Embryo Interactions.,” *Biol. Reprod.*, vol. 94, no. 2, p. 38, Feb. 2016.
263. A. Shenoy and T. Geiger, “Super-SILAC: current trends and future perspectives,” *Expert Rev. Proteomics*, vol. 12, no. 1, pp. 13–19, Jan. 2015.
264. A. Shevchenko, H. Thomas, J. Havli, J. V Olsen, and M. Mann, “In-gel digestion for mass spectrometric characterization of proteins and proteomes,” *Nat. Protoc.*, vol. 1, no. 6, pp. 2856–2860, Jan. 2007.
265. D. Hanahan and R. A. Weinberg, “Hallmarks of cancer: The next generation,” *Cell*, vol. 144, no. 5, pp. 646–674, 2011.
266. A. P. Novetsky, I. Zigelboim, D. M. Thompson, et al., “Frequent mutations in the RPL22 gene and its clinical and functional implications.,” *Gynecol. Oncol.*, vol. 128, no. 3, pp. 470–4, Mar. 2013.
267. K. Goudarzi and M. Lindström, “Role of ribosomal protein mutations in tumor development (Review),” *Int. J. Oncol.*, vol. 48, no. 4, pp. 1313–24, Feb. 2016.
268. A. Artero-Castro, J. Castellvi, A. García, et al., “Expression of the ribosomal proteins Rplp0, Rplp1, and Rplp2 in gynecologic tumors,” *Hum. Pathol.*, vol. 42, no. 2, pp. 194–203, Feb. 2011.
269. Rital B. Bhavsar, Leah N. Makley and Panagiotis A. Tsonis , “The other lives of ribosomal proteins.” *Human genomics*, vol.4, no. 5, pp. 327-344, June 2010.
270. Y. Zhu, X. Chen, Q. Pan, et al., “A Comprehensive Proteomics Analysis Reveals a Secretory Path- and Status-Dependent Signature of Exosomes Released from Tumor-Associated Macrophages,” *J. Proteome Res.*, vol. 14, no. 10, pp. 4319–4331, Oct. 2015.
271. G. Solinas, G. Germano, A. Mantovani, and P. Allavena, “Tumor-associated macrophages (TAM) as major players of the cancer-related inflammation.”

- J. Leukoc. Biol.*, vol. 86, no. 5, pp. 1065–73, Nov. 2009.
272. C. Guo, A. Buranych, D. Sarkar, P. B. Fisher, and X.-Y. Wang, “The role of tumor-associated macrophages in tumor vascularization,” *Vasc. Cell*, vol. 5, no. 1, p. 20, Dec. 2013.
273. S. Mathivanan, J. W. E. Lim, B. J. Tauro, H. Ji, R. L. Moritz, and R. J. Simpson, “Proteomics analysis of A33 immunoaffinity-purified exosomes released from the human colon tumor cell line LIM1215 reveals a tissue-specific protein signature.,” *Mol. Cell. Proteomics*, vol. 9, no. 2, pp. 197–208, Feb. 2010.
274. C. Théry, M. Boussac, P. Véron, et al., “Proteomic analysis of dendritic cell-derived exosomes: a secreted subcellular compartment distinct from apoptotic vesicles.,” *J. Immunol.*, vol. 166, no. 12, pp. 7309–18, Jun. 2001.
275. R. J. Burgess and Z. Zhang, “Histone chaperones in nucleosome assembly and human disease,” *Nat. Struct. Mol. Biol.*, vol. 20, no. 1, pp. 14–22, Jan. 2013.
276. J. Kowal, G. Arras, M. Colombo et al., “Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes.”
277. H. Peinado, M. Alečković, S. Lavotshkin, et al., “Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET,” *Nat. Med.*, vol. 18, no. 6, pp. 883–891, May 2012.
278. J. Palma, S. C. Yaddanapudi, L. Pigati, et al., “MicroRNAs are exported from malignant cells in customized particles,” *Nucleic Acids Res.*, vol. 40, no. 18, pp. 9125–9138, Oct. 2012.
279. G. Müller, M. Schneider, J. Gassenhuber, and S. Wied, “Release of exosomes and microvesicles harbouring specific RNAs and glycosylphosphatidylinositol-anchored proteins from rat and human adipocytes is controlled by histone methylation,” *Am. J. Mol. Biol.*, vol. 2, no. 3, pp. 187–209, 2012.
280. A. G. Paulovich, J. R. Whiteaker, A. N. Hoofnagle, and P. Wang, “The interface between biomarker discovery and clinical validation: The tar pit of

- the protein biomarker pipeline.” *Proteomics. Clin. Appl.*, vol. 2, no. 10–11, pp. 1386–1402, Oct. 2008.
281. V. Lange, P. Picotti, B. Domon, and R. Aebersold, “Selected reaction monitoring for quantitative proteomics: a tutorial.” *Mol. Syst. Biol.*, vol. 4, p. 222, 2008.
282. G. Bezakova and M. A. Ruegg, “New insights into the roles of agrin.” *Nat. Rev. Mol. Cell Biol.*, vol. 4, no. 4, pp. 295–308, Apr. 2003.
283. K. Baghy, P. Tátrai, E. Regős, and I. Kovalszky, “Proteoglycans in liver cancer.” *World J. Gastroenterol.*, vol. 22, no. 1, pp. 379–93, Jan. 2016.
284. L. Muinelo-Romay, E. Colas, J. Barbazan, et al., “High-risk endometrial carcinoma profiling identifies TGF- $\beta$ 1 as a key factor in the initiation of tumor invasion.” *Mol. Cancer Ther.*, vol. 10, no. 8, pp. 1357–66, Aug. 2011.
285. P. Tátrai, J. Dudás, E. Batmunkh, et al., “Agrin, a novel basement membrane component in human and rat liver, accumulates in cirrhosis and hepatocellular carcinoma.” *Lab. Invest.*, vol. 86, no. 11, pp. 1149–60, Nov. 2006.
286. S. Chakraborty, M. Lakshmanan, H. L. F. Swa, et al., “An oncogenic role of Agrin in regulating focal adhesion integrity in hepatocellular carcinoma,” *Nat. Commun.*, vol. 6, p. 6184, Jan. 2015.
287. R. Kawahara, D. C. Granato, C. M. Carnielli, et al., “Agrin and Perlecan Mediate Tumorigenic Processes in Oral Squamous Cell Carcinoma,” *PLoS One*, vol. 9, no. 12, p. e115004, Dec. 2014.
288. F. Vences-Catalán, C. Duault, C.-C. Kuo, R. Rajapaksa, R. Levy, and S. Levy, “CD81 as a tumor target,” *Biochem. Soc. Trans.*, vol. 45, no. 2, pp. 531–535, Apr. 2017.
289. F. Vences-Catalán, R. Rajapaksa, M. K. Srivastava, et al., “Tetraspanin CD81 promotes tumor growth and metastasis by modulating the functions of T regulatory and myeloid-derived suppressor cells.” *Cancer Res.*, vol. 75, no. 21, pp. 4517–26, Nov. 2015.
290. Y. Liu, X. Jin, Y. Li, et al., “DNA methylation of claudin-6 promotes breast cancer cell migration and invasion by recruiting MeCP2 and deacetylating

- H3Ac and H4Ac," *J. Exp. Clin. Cancer Res.*, vol. 35, no. 1, p. 120, Dec. 2016.
291. S. Yamazawa, T. Ushiku, A. Shinozaki-Ushiku, et al., "Gastric Cancer With Primitive Enterocyte Phenotype," *Am. J. Surg. Pathol.*, p. 1, May 2017.
292. P. Micke, J. S. M. Mattsson, K. Edlund, et al., "Aberrantly activated claudin 6 and 18.2 as potential therapy targets in non-small-cell lung cancer.," *Int. J. cancer*, vol. 135, no. 9, pp. 2206–14, Nov. 2014.
293. E. Rendón-Huerta, F. Teresa, G. M. Teresa, et al., "Distribution and expression pattern of claudins 6, 7, and 9 in diffuse- and intestinal-type gastric adenocarcinomas.," *J. Gastrointest. Cancer*, vol. 41, no. 1, pp. 52–9, Mar. 2010.
294. T. Ushiku, A. Shinozaki-Ushiku, D. Maeda, S. Morita, and M. Fukayama, "Distinct expression pattern of claudin-6, a primitive phenotypic tight junction molecule, in germ cell tumours and visceral carcinomas.," *Histopathology*, vol. 61, no. 6, pp. 1043–56, Dec. 2012.
295. C. R. Stadler, H. Bähr-Mahmud, L. M. Plum, et al., "Characterization of the first-in-class T-cell-engaging bispecific single-chain antibody for targeted immunotherapy of solid tumors expressing the oncofetal protein claudin 6," *Oncoimmunology*, vol. 5, no. 3, p. e1091555, Mar. 2016.
296. A. D. Santin, F. Zhan, S. Cane', et al., "Gene expression fingerprint of uterine serous papillary carcinoma: identification of novel molecular markers for uterine serous cancer diagnosis and therapy," *Br. J. Cancer*, vol. 92, no. 8, pp. 1561–1573, Apr. 2005.
297. A. D. Santin, S. Bellone, E. R. Siegel, et al., "Overexpression of Clostridium perfringens Enterotoxin Receptors Claudin-3 and Claudin-4 in Uterine Carcinosarcomas," *Clin. Cancer Res.*, vol. 13, no. 11, pp. 3339–3346, Jun. 2007.
298. J. Agola, P. Jim, H. Ward, S. BasuRay, and A. Wandinger-Ness, "Rab GTPases as regulators of endocytosis, targets of disease and therapeutic opportunities," *Clin. Genet.*, vol. 80, no. 4, pp. 305–318, Oct. 2011.
299. Y. Bie and Z. Zhang, "RAB8A a new biomarker for endometrial cancer?"

300. Y. Sun, P. J. Bilan, Z. Liu, and A. Klip, "Rab8A and Rab13 are activated by insulin and regulate GLUT4 translocation in muscle cells.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 46, pp. 19909–14, Nov. 2010.
301. K. Meehan and L. J. Vella, "The contribution of tumour-derived exosomes to the hallmarks of cancer.," *Crit. Rev. Clin. Lab. Sci.*, pp. 1–11, Oct. 2015.