Universitat Autònoma de Barcelona

Metabolomic study for the identification of diagnostic markers and the characterization of the dissemination process of endometrial cancer

Tatiana Altadill Vallespí

Metabolomic study for the identification of diagnostic markers and the characterization of the dissemination process of endometrial cancer

PhD Thesis 2018
Tatiana Altadill Vallespí

Metabolomic study for the identification of diagnostic markers and the characterization of the dissemination process of endometrial cancer

Doctoral thesis presented by

Tatiana Altadill Vallespí

to obtain the degree of

Doctor for the Universitat Autònoma de Barcelona (UAB)

Doctoral thesis performed at Vall d'Hebron Research Institute, in the Group of Biomedical Research in Gynecology, under the supervision of

Dr. Amrita K Cheema, Dr. Eva Colás and Dr. Antonio Gil

Thesis affiliated to the Department of Cell Biology, Physiology and Immunology from the Faculty of Medicine at the UAB, in the PhD program of Cell Biology, under the tutoring of

Dr. Rosa Miró

Universitat Autònoma de Barcelona, January 2018

Dra. Amrita K Cheema (director) Dra. Eva Colás (director) Dr. Antonio Gil (director)

Dra. Rosa Miró (tutor) Tatiana Altadill (student)

A la meua família,

a la meua gent,

a la meua terra.

This thesis work has been carried out between September 2014 and February 2018 in the Group of Biomedical Research in Gynecology at Vall d'Hebron Research Institute. It has been supported by the "Agència de Gestió d'Ajuts Universitaris i de Recerca, Generalitat de Catalunya" thanks to the achievement of the scholarship "Programa de formació del personal investigador" (REF-2015 FI_B 00703). Moreover, this thesis was supported in part by the International Research Staff Exchange Scheme (IRSES) - Marie Curie Actions - and by the Metabolomics Facility of Georgetown University (USA) for a long stage fellowship at the Lombardi Comprehensive Cancer Center (Georgetown, USA) from September 2014 to December 2015. This work has been supervised by Dr. Amrita K Cheema, Dr. Eva Colás and Dr. Antonio Gil.

ACKNOWLEDGEMENTS

Al final de tot, ha acabat passant. Dies més bons i altres de més dolents han resultat en el dipòsit d'aquesta tesi. Res, però, no hagués tingut sentit ni sigut possible sense les persones que m'han acompanyat de ben prop en aquest viatge. Persones que han permès omplir el "camí cap a la defensa" d'experiències que m'han fet crèixer no només com a científica, sinó com a persona. Persones a les que m'agradaria agraïr que hagin transformat el que podia haver sigut una etapa dura i amarga, en una vivència enriquidora i que recordaré de bona gana.

Per començar, m'agradaria agraïr al Dr. Jaume Reventós que m'obrís les portes del seu laboratori i em permetés embarcar-me en mil aventures al seu costat (TransBioMed, Washington DC, etc). I agraïr-te, Jaume, la teva proximitat així com que m'acabessis convencent per matricular-me al doctorat. També voldria donar les gràcies a les persones que van ajudar-me a donar els primers passos dins el laboratori, en especial a tu, Marina. Sempre has estat un punt de referència per mi.

Per descomptat, agraïr també als meus directors de tesi Dra. Eva Colás, Dr. Antonio Gil i Dr. Amrita K Cheema i a la meva tutora Dra. Rosa Miró que m'hagin donat aquesta oportunitat. Eva (mamá moderna!), gracias por apoyarme y aguatarme siempre, aún con mi temperamento. Por el cariño que me has demostrado. Gracias por la confianza que me has dado para hablarte siempre claro. Sabes que has sido mucho más que una directora de tesis para mi, verdad? Antonio, me gustaría agradecer tu predisposición a ayudarnos así com tu confianza en nuestras capacidades y proyectos. Dr. Cheema, I would like to thank you for guiding me along this thesis. For making possible the realization of the metabolomics projects but also for the warm reception and relationship you offered me when I was so far away from home. I would like to thank also Dr. Stephen Byers for giving me this opportunity.

Many thanks also to all the lab mates at Georgetown for your help and the great moments we shared inside and outside the lab. Also, to all the people I met in DC and made me feel at home. It was a privilege to meet Julian and Michelle and to have them by my side. Y a toda la trupe española i catalana. A Pedro; eso no fue más que un principio. Y a mis "cerdis"! Cuanto os echo de menos!

ACKNOWLEDGEMENTS

Gràcies al servei de Ginecologia i Patologia de l'Hospital Vall d'Hebron. Així com també a tot el personal clínic i als pacients que fan possible la investigació translacional que duem a terme al laboratori.

No podria faltar el meu agraïment als de la casa: Eulàlia, Rosa, Pilar, Rai, Paco, Tau i Rachidia. La vostra presència, la vostra actitud i el "bon dia" que no falta mai fa que l'edifici Collserola no sigui un annex més del VHIR, sinó un lloc que sentim com casa nostra.

M'agradaria també agraïr els consells del Cristian, la Mire, l'Anna Santamaria, el Pep Roma, el Miguel Segura, l'Aroa, la Maite, la Maria Antonia, la Rosanna... i de tots aquells qui saben més i que en algun moment de la tesi ens han donat un cop de mà quan estàvem estancats entre experiments que no hi havia manera que sortissin bé.

I seria impossible no relacionar "la Era tesi" en tots valtres. En els que vau formar part del començament, que ja no esteu, però vau ser tan importants (Marina, Tamara, Raúl, Marta Llauradó, Marta Garcia, Núria Pedrola, Melània, Tugçe, Gabriel, Anna Almazan...); en la gent nova que s'ha anat incorporant (Eva, Carlos, Cristian, Alfonso, Núria, Melissa, Irina, Patri) i en los que heu estat sempre aquí (Blanca, Laura, Lucía, Isaac, Elena, Irene, Júlia, Berta, Eli). No us quedareu a la meua memòria només en la imatge d'una pipeta a la mà; sinó que me'n vaig en lo record de les mil risses, (i plors...), els sopars, els mil viatges (Lux, Istambul, USA, Grècia, escapades a cases rurals), les esquiades, sortides en bici, sing stars ♫♪ i un munt de coses més que podria escriure per acabar fent una llista infinita. Estic molt contenta d'acabar ja en tot lo mal de cap i burocràcia del dipòsit i la defensa, però a la vegada me fa molta pena deixar tot això atràs...no me'n faig a la idea! Encara que sé que fora d'aquestes parets blanques encara mos queda temps per escriure moltes històries més!

I algunes mencions especials... A tu Blanca, perquè ets una crack! Perquè t'admiro des del moment zero. Perquè tu vas ser el punt d'entrada al 209. Gràcies per tots estos anys al meu costat i per l'amistat que m'has demostrat sempre. A la Ire, la Helen i la Lau; sense valtres lo lab no hagués tingut gens però que gens de salsa! Quan penso en valtres se'm dibuixa un somriure a la cara. A la nostra secre molonga! Gràcies per estar sempre en tot, Eli! A la Júlia;

poques paraules tinc per descriure la sort d'haver-te trobat! Recordo lo teu començament "marginal" ☺ i com poc a poc te mos has anat guanyant a totes en la teua màgia i positivisme. A tu Lu! Lo meu llucet!! Un pilar base al 209 i fora d'ell. Per tot lo que s'ha creat entre naltres i per tot lo temps que durarà! A les noves incorporacions potents d'endometri: Eva i Carlos. Valeu molt! Trobaré a faltar los nostres caferets post-dinar. I a la meva ultra mega maxi pivona!!! La que arbolitza i ictusa pel 209 continuament. Quan me despertava les últimes setmanes sense més energia per vindre al lab (duríssim! fatalíssim!), pensava en tu (a gaaaas!) i ja no me costava aixecar-me del llit (...del pal/ (barra)/ tipu.....què farà la meua nena pija sense mi -alias dismi-???....). Gràcies, xurri, pels èpics moments al meu costat!

I sas allò que vas més o menos organitzant los agraïments per grups de gent. I arribo a tu, Isaac. Quin problema...te poso al grup de gent del lab (el tècnic més guaperas del VHIR!), d'ex-parelles yogus horribles ;), de col·lega d'escalada i de súper entrenament, de futurs companys de viatge?? On te poso? A tu Isaac, gràcies perquè has sigut lo descobriment més gran fruit de la tesi. Per estar sempre allà quan calia (des dels clonatges de FOXM1 hasta quan me vas ensenyar a conduir la moto per Barna), per desaparèixer quan va fer falta, i pel nostre gran retrobament ♥♥. T'estimo, guapo!

I a la meua gran familia aquí a la city (i a alguns de l'extra-radi): a tots los "Chatarreros del Segre", a Arturo (per tot lo que em compartit i per tot lo que mos queda, *pineipol*!!), a la Marta (gràcies per ajudar-me en lo disseny de la portada!), a l'Ángel, a la Laia, als swingueros rotos i mosquetros del swing (Rafa, Nahat!), a la Iris... Als que un dia vau estar tan prop i als relleus que van arribant → A ti, Stefano, por recordarme que es estar otra vez motivada por las cosas que me gusta hacer y darme el refuerzo positivo que ha sido tan importante en estas últimas semanas ♥. Cada un de valtres és una petita part de mi. Però per davant de tot, a qui dec les gràcies més grans és als de casa: Arkaitz, Pablo i Andrea. A valtres que sou lo millor que m'envolta. Los que heu estat en lo bo (i molt bo), però també en lo dolent; i és que sou qui m'heu donat la mà per aixeca'm d'anterra tantes vegades com ha calgut. Sóc la més afortunada del món de tindreu's al costat!

ACKNOWLEDGEMENTS

A la meua família de sang, no m'imagino una vida sense un coixí com el que tinc la sort de tindre. Des de muns iaios, hasta tots los tiets i tietes, padrins i padrines i cosins i cosines. Gràcies per haver-me estimat sempre. Als petits de la casa (o ja no tan petits) Ari i Joel, Ferran, Damià i Regi. Als segons cosins, especialment a tu, Àxel (perreteee!). A Simó i a la Maria (ets una floreta brillant in crescendo! ☼), per la sort que hem tingut de crèixer junts com a cosins, amics i tot lo que se pugo ser. M'heu donat la confiança que m'ha fet falta per arribar aquí!

I a muns pares, que m'ho heu donat tot sempre des de l'amor. Que vau crear la base de lo que sóc avui. Una base que me vau ensenyar a remodelar en criteri i quan calgués per arribar a ser millor persona. A valtres, que m'heu ensenyat a caminar per la vida en ganes i sense por a res; sempre des del respecte. A valtres que heu fet possible que arriba a ser doctora. I a valtres, que m'heu donat la cosa que estimo més de l'univers: l'Andrea. Us estimo!

A tu, Andrea. Per ser tant especial. Per compartir cada moment en mi. Per estima'm *hasta l'* ∞ i fe'm tocar de peus anterra quan ha fet falta. Per creure sempre en mi. Per dona'm una visió tant diferent de la vida, de les situacions i de les persones. I per ser la força que fa possible que aconseguixa les metes que me poso i la meua energia per tirar sempre avant.

INDEX

INDEX

LIST OF FIGURES

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ADAR: adenosine deaminase acting on RNA

ANOVA: analysis of variance

AUC: area under the curve

BT: brachytherapy

CA125: cancer antigen 125

CTNNB1: catenin (cadherine-associated protein), beta 1

DA: discriminant analysis

DDA-MS: data dependent acquisition MS

DNA: deoxyribonucleic acid

EC: endometrial cancer

EEC: endometrioid endometrial cancer

EIN: endometrial intraepithelial neoplasia

ELISA: enzyme-linked immunosorbent assay

ELV: exosome-like vesicles

EMT: epithelial to mesenchymal transition

EpCAM: epithelial cell adhesion molecule

ESI: electrospray ionization

EVs: extracellular vesicles

FACS: fluorescence-activated cell sorting

FIGO: Federation of Ginecology and Obstetrics

G: grade

GC: gas chromatography

HCA: hierarchical clustering analysis

HILIC: hydrophilic-interaction chromatography

HMDB: human metabolome database

HMDB: human metabolome database

HMP: human metabolome project

HPLC: high-performance liquid chromatography

HRMS: high resolution mass spectrometry

IHQ: immunohistochemistry

iTRAQ: isobaric tags for relative and absolute quantitation

LC: liquid chromatography

LO: large oncosomes

LVSI: lymphovascular space invasion

m/z: mass-to-charge ratio

miRNA: micro RNA

MMCD: Madison metabolomics consortium database

MRI: magnetic resonance imaging

ABBREVIATIONS

MRM: multiple reaction monitoring

mRNA: messenger RNA

MS/MS: tandem mass spectrometry

MS: mass spectrometry

MSe: mean squared error

MSEA: metabolite set enrichment analysis

MVBs: multivesicular bodies

MVs: microvesicles

NEEC: non-endometrioid endometrial cancer

NMR: nuclear magnetic resonance

NPF: negative predictive factor

OPLS: orthogonal projections to latent structures

P: plasma

PC: principal component

PCA: principal component analisis

PLS: partial least squares

POLE: DNA polymerase epsilon catalytic subunit

PRM: parallel reaction monitoring

PTEN: phosphatase and tensin homolog

QC: quality control

QIT: quadrupole-ion trap

QQQ: triple quadrupole

Q-TOF: quadrupole time-of-flight

RNA: ribonucleic acid

RT: radiotherapy

S1P: sphingosine-1-phosphate

SELDI-MS: surface-enhanced laser desorption/ionization MS

SPHK: sphingosine kinase

SVM: support vector machine

TCGA: The Cancer Genome Atlas

TMA: tissue microarray

TP53: tumor protein p53

UA: uterine aspirate

UHPLC: ultra high performance liquid chromatography

UPLC: ultra-performance liquid chromatography

WB: western blot

INTRODUCTION

1. ENDOMETRIAL CANCER

1.1. ANATOMY OF THE UTERUS AND THE ENDOMETRIUM

The **uterus** is an intrapelvic muscular organ of the feminine reproductive system. Its main function is the fetus development during pregnancy. It is located in front of the rectum and is connected to the vagina and Fallopian tubes. The uterus is divided into 4 regions from the upper to the lowest part: the **fundus**, next to the connection point of the Fallopian tubes; the **body**, the largest part; the **isthmus**, a constriction that separates the body from the cervix; and the **cervix**, that is the part that opens to the vagina.

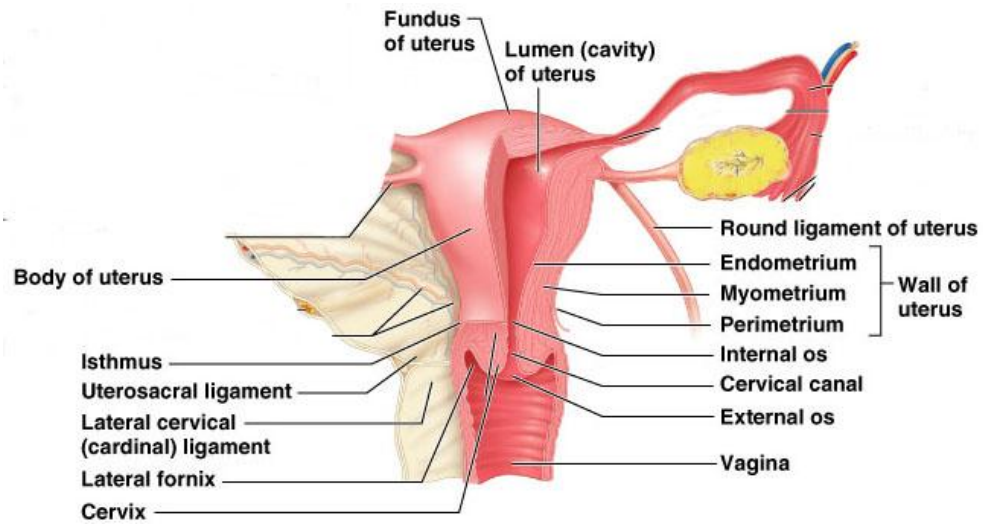


Figure 1. The human uterus. Adapted from http://www.apsubiology.org/

The uterus wall is composed by three tissue layers (see Figure 1):

- The **perimetrium** or serosa layer: outer serosa layer that is the continuation of the peritoneum.
- The **myometrium** or muscular layer: middle smooth muscle stratum that is divided into 3 parts: *stratum submucosum*, *stratum vasculate,* and *stratum supravasculate,* from the inner to the outer myometrium. Changes in these layers are observed during pregnancy.

- The **endometrium** or mucosa layer: inner glandular layer. It consists of a thick connective stromal tissue, where glands and other type of cells (like immune cells) are also present, and a simple columnar epithelium of cells, partially cialiated. The endometrium can be subdivided into 2 parts: the *stratum basalis*, maintained quite stable along the menstrual cycle; and the *stratum superficialis*, secretory part that is lost and regenerated in each menstrual cycle in response to hormones [1].

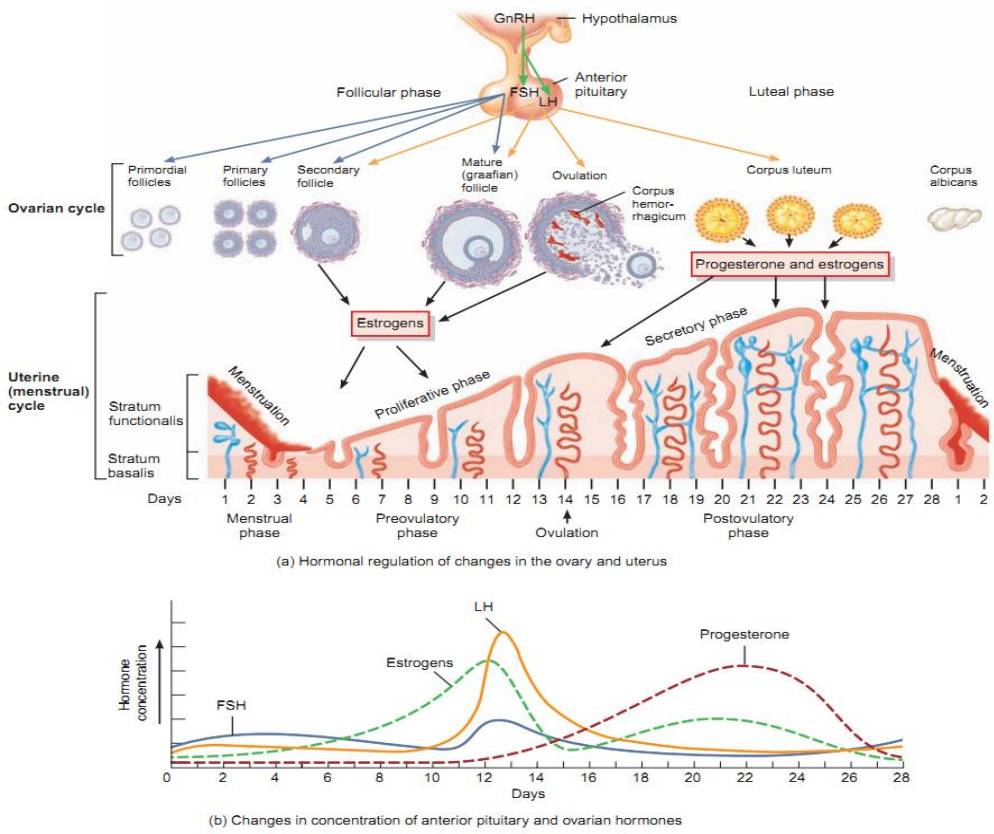


Figure 2. Changes during the menstrual cycle. Adapted from Gerard J. Tortora. Principles of Anatomy and Physiology [2].

From puberty and during the reproductive life of women the endometrium is functional but after menopause it becomes atrophic. When the uterus is functional, it experiences changes according to the reproductive cycle (Figure 2). Cyclic alterations in the gonadotropin and steroid release are observed as well as complex interactions between the hypothalamus, pituitary gland and ovaries [2].

1.2. DISORDERS OF THE ENDOMETRIUM

Different types of endometrial lesions can be developed along the lifetime, from precancerous to malignant disorders.

1.2.1. Benign Endometrial Hyperplasia

The benign endometrial hyperplasia is a precancerous lesion that appears as a response of the endometrium to an abnormal estrogenic stimulus resulting in large polyclonal proliferations [3].

Endometrial hyperplasia is characterized by the remodeling of the glands (gland dilatation and packing), irregular cellular restructuration, vascular thrombi, stromal breakdown, and disordered proliferative tissue. Although this lesion is mostly encountered around the time of menopause or in post-menopause, the endometrium is still functional and responding to the abnormal hormonal environment.

Traditionally, we can divide the hyperplasia into 2 types based on the pathological grade (simple and complex) and the presence of atypia (Figure 3). The progression to endometrial cancer (EC) will appear in 1% of patients with simple hyperplasia without atypia, 3% of patients with complex hyperplasia without atypia, 8% of patients with simple hyperplasia with atypia, and 29% of patients with complex hyperplasia with atypia [4].

In 2000, the International Endometrial Collaborative Group introduced a new and simpler classification system to better predict progression to cancer, the endometrial intraepithelial neoplasia (EIN), that defined 3 groups: non-EIN or benign hyperplasia (polyclonal hormone dependent diffuse lesion), EIN (a neoplasic monoclonal lesion that can be localized initially but later become diffuse) and EC. The detection of EIN increases 45-fold times the risk of developing EC in one year [5,6].

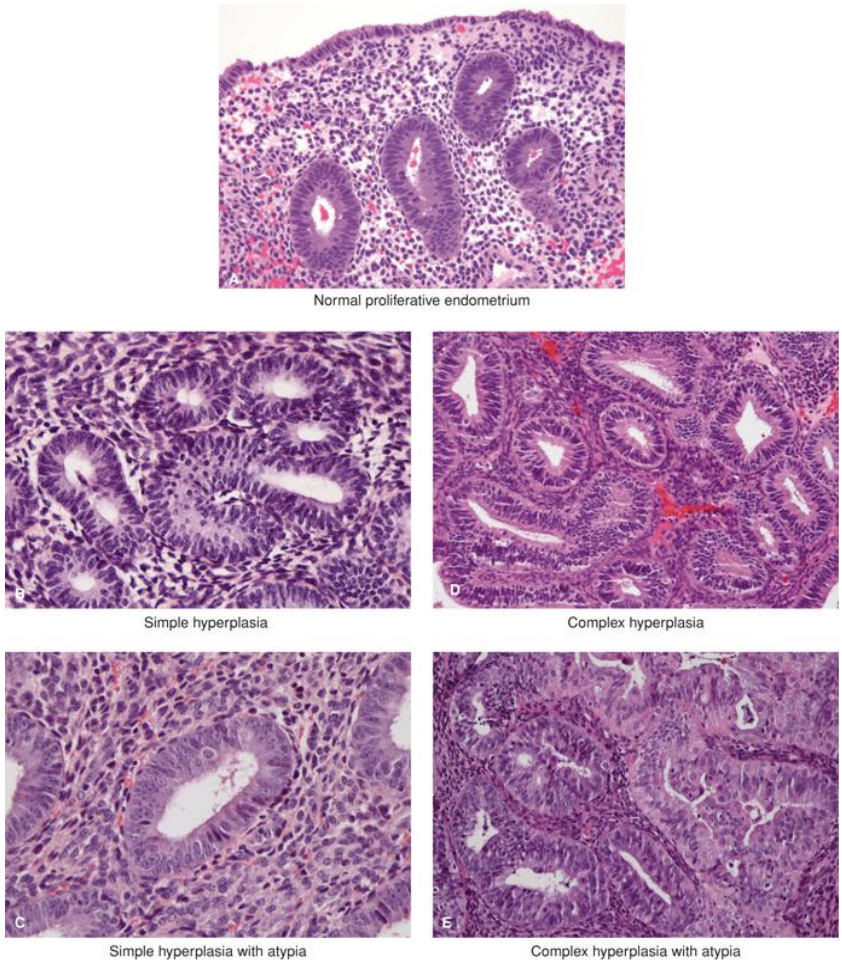


Figure 3. Endometrial hyperplasia types [4].

Endometrial hyperplasia is detected usually by the pathological evaluation of an endometrial biopsy performed to patients with a specific symptomatology (i.e. presence of abnormal vaginal bleeding) or accidentally due to an exploration related with another disorder. Neither PAP smear nor standard screening programs are effective in order to diagnose this lesion.

1.2.2. Endometrial polyps

Endometrial polyps (Figure 4) are the most common benign endometrial lesions in reproductive women, but the etiology is unclear. They can be sessile or pedunculated and the associated risk to a neoplasic transformation is very low.

Sometimes they are detected by an abnormal vaginal bleeding (10-20% of the cases) and treated easily and quickly by resection [7].

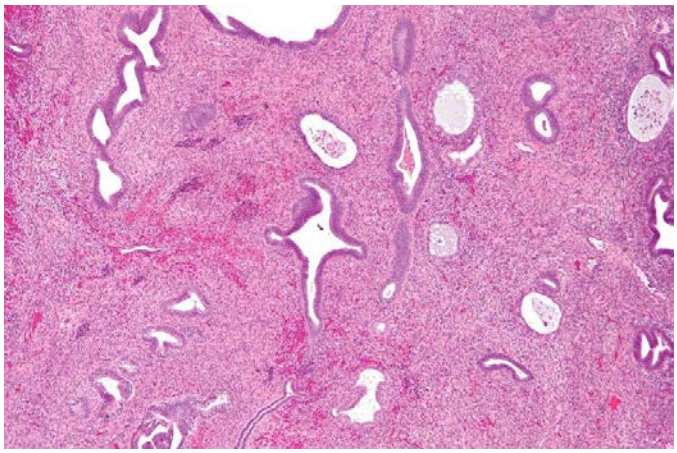


Figure 4. Endometrial benign polyp. Adapted from https://en.wikipedia.org/

1.3. ENDOMETRIAL CANCER

1.3.1. Epidemiology

EC has an incidence of 61,380 new cases per year (2017), representing a 7% of the total new cases of cancers. It accounts for 4% of the estimated deaths per year, being the fourth most common cancer among females in developed countries (Figure 5). An important problem of EC is that the number of deaths per year in the last decades has been increasing rapidly (from 2,900 deaths in 1987 to 10,920 in 2017 in USA) [8].

Estimated New Cases

Males			Females		
Prostate	161,360	19%	Breast	252,710	30%
Lung & bronchus	116,990	14%	Lung & bronchus	105,510	12%
Colon & rectum	71,420	9%	Colon & rectum	64,010	8%
Urinary bladder	60,490	7%	Uterine corpus	61,380	7%
Melanoma of the skin	52,170	6%	Thyroid	42,470	5%
Kidney & renal pelvis	40,610	5%	Melanoma of the skin	34,940	4%
Non-Hodgkin lymphoma	40,080	5%	Non-Hodgkin lymphoma	32,160	4%
Leukemia	36,290	4%	Leukemia	25,840	3%
Oral cavity & pharynx	35,720	4%	Pancreas	25,700	3%
Liver & intrahepatic bile duct	29,200	3%	Kidney & renal pelvis	23,380	3%
All Sites	**836,150**	**100%**	**All Sites**	**852,630**	**100%**

Estimated Deaths

Males			Females		
Lung & bronchus	84,590	27%	Lung & bronchus	71,280	25%
Colon & rectum	27,150	9%	Breast	40,610	14%
Prostate	26,730	8%	Colon & rectum	23,110	8%
Pancreas	22,300	7%	Pancreas	20,790	7%
Liver & intrahepatic bile duct	19,610	6%	Ovary	14,080	5%
Leukemia	14,300	4%	Uterine corpus	10,920	4%
Esophagus	12,720	4%	Leukemia	10,200	4%
Urinary bladder	12,240	4%	Liver & intrahepatic bile duct	9,310	3%
Non-Hodgkin lymphoma	11,450	4%	Non-Hodgkin lymphoma	8,690	3%
Brain & other nervous system	9,620	3%	Brain & other nervous system	7,080	3%
All Sites	**318,420**	**100%**	**All Sites**	**282,500**	**100%**

Figure 5. Ten Leading Cancer Types for the Estimated New Cancer Cases and Deaths by Sex, United States, 2017. Estimates are rounded to the nearest 10 and cases exclude basal cell and squamous cell skin cancers and in situ carcinoma except urinary bladder. Adapted from Siegel et al. [8].

Most EC patients (67%) are diagnosed when the tumor is still confined to the uterus (stages I and II according to the F.I.G.O. classification. See section 1.3.4 for EC classification) and the 5-year survival rate associated to this group is 95%. In those cases diagnosed with high grade EC or when the tumor is already spread the 5-year survival rate to 35-45% [8] (Figure 6).

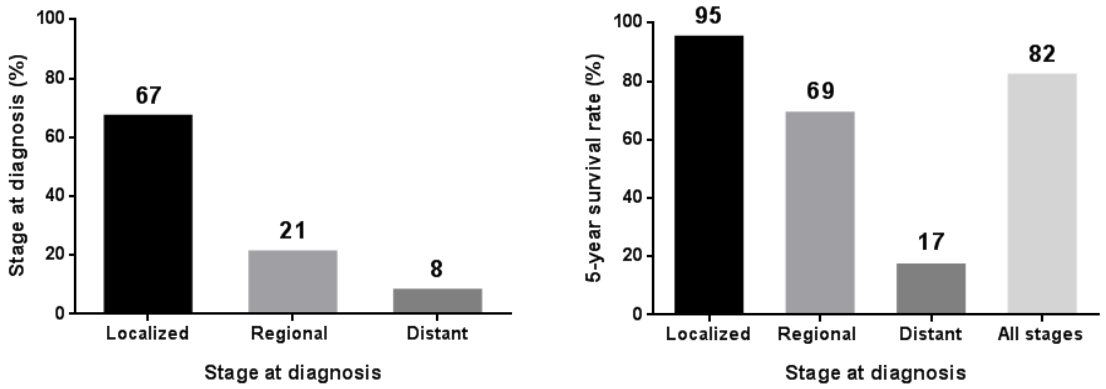


Figure 6. Stage distribution and 5-year relative survival rates of EC by stage at diagnosis (data from United States, 2006 to 2012). Stage categories do not sum to 100% because sufficient information is not available to stage all cases. Adapted from Siegel et al. [8].

Many studies have tried to elucidate the main causes for EC, and up to date we know that about 5% of ECs are caused by hereditary susceptibility. Moreover, some **risk factors** have been reported, mainly associated with endometrioid EC (EEC) rather than with the non-endometrioid EC (NEEC) (see clinicopathological classification section for a detailed description of these EC subtypes). The most important risk factors for EEC are obesity and excessive/unbalanced estrogen exposure. The excess of estrogen exposure can be due to exogenous (i.e. tamoxifen therapy for breast cancer treatment) or endogenous factors (i.e. early menarche, late menopause, nulliparity, infertility and chronic anovulation). Diabetes, high dietary fat intake, metabolic disorders (lipid and carbohydrate catabolism), hypertension and age have been also associated with increasing risk for EC. It has been also proved that elevate intake of acid folic and derivates is associated with higher risk to develop NEEC.

Some **preventive factors** to avoid the risk of developing EC have also been identified. For example, women that use oral contraceptive pills or are pregnant most of their lives have associated less risk to develop EC. Both situations increase the progesterone action and have the opposite effects compared to an excessive exposure to estrogens. Smoking and physical activity also increase the protection against EC [9].

1.3.2. Clinical presentation

EC is usually diagnosed at early stages since around 90% of patients present abnormal vaginal bleeding. The detection of these symptoms, especially when observed in post-menopausal women, should be considered suspicious for EC. Just a 15% of post-menopausal women with uterine bleeding will be diagnosed with EC. The remaining causes of that bleeding are intake of exogenous estrogens (30%), atrophic endometritis and vaginitis (30%), presence of polyps (10%), endometrial hyperplasia (5%) and others (10%).

Abdominal pain, changes in bladder functions, abdominal distension, alterations in vaginal discharge and anemia can be also observed, generally at advanced stages of EC [10].

1.3.3. Diagnosis

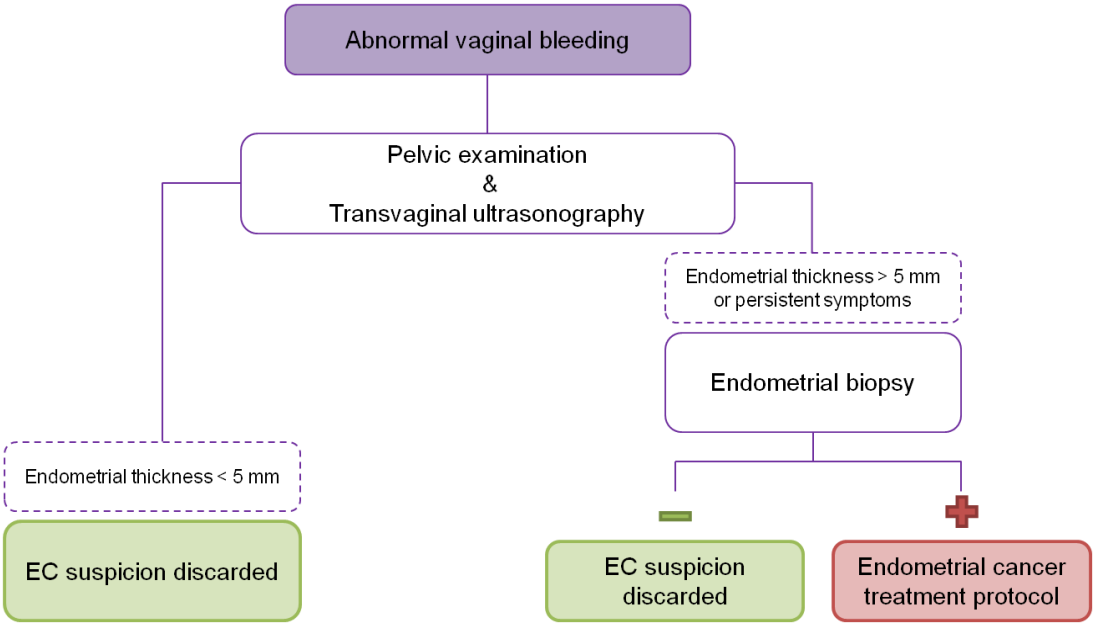


Figure 7. Diagnostic procedure.

A patient suspicious of having EC will undergo a pelvic examination and a transvaginal ultrasonography. Both methodologies are useful tools in order to better study the size and shape of the organs and thickness of the endometrial line. Nevertheless, a definitive diagnosis usually requires an endometrial biopsy. At early stages of EC, it is difficult to observe alterations in the morphology by palpation. However, the imaging technique is very effective for symptomatic patients to discard the presence of polyps and other benign pathologies. This technique has a sensitivity of an 90% and 54% of specificity. When thickening of endometrial line is detected (> 5 mm) or if the symptomatology persists, patient is suspicious for EC and should continue with further tests (Figure 7) [11].

A histopatologic examination is commonly needed for the final diagnosis through an endometrial biopsy. A small sample of the uterus is taken and observed by a pathologist under the microscope. As the first method of choice, the biopsy is taken by aspiration and gives a final diagnosis of an 80% of patients that presented vaginal bleeding or other symptoms. If it is not possible to get the biopsy by aspiration, the biopsy guided by hysteroscopy and or by dilation & curettage are other options with a greater statistical outcome as a diagnostic tool but more invasive. Along with EC diagnosis, the EC type and grade will be determined by the pathologist in order to guide the primary treatment, i.e. surgery (see section 1.3.6.).

Sometimes, along with the biopsy, Doppler technique test is also performed. It gives the resistance index of the arteries and smaller vases: low index corresponds to higher EC risk and is due to neoformation of vases with less muscle presence during carcinogenesis. In the case that EC is diagnosed, magnetic resonance and computed tomographic scan of the abdominal region are also commonly used in order to determine the grade of myometrial and cervical invasion as well as to analyze the peritoneal ganglia and the ascites presence. Moreover, when a patient is diagnosed with high grade EC, it can also be useful to follow up the tumor biomarker CA125 that can be assessed as a recurrence indicative [12–14].

1.3.4. EC classification

1.3.4.1. Clinicopathological classification and the dualistic model

Regarding the clinicopathological features, Bokman et al. [15] proposed the dualistic model in 1983. According to this model, EC can be divided into type I or EEC and type II or NEEC (Table 1).

	EEC	NEEC
Incidence	80%	20%
Development	Perimenopausal or early postmenopausal age	Elderly women
Histological subtypes	Endometrioid and mucinous	Papillary serous and cell clear
Tumoral grade	Mostly low (G1 and G2), but also G3	High (G3)
Molecular characteristics	PTEN (40-80%) K-ras (20-35%) Beta-catenin (40%) Microsatellite instability (20-40%)	TP53 (90%) PTEN (10%) P19 (40%) E-catherin (80-90%)
Hormone dependency	Yes	No
Estrogen exposure	Associated	Non-associated
Course of the disease	Slow and stable	Agressive
Origin	Endometrial hyperplasia	Polyps or precancerous lesions
Diagnosis	Early stages	Late stages
Prognosis	Good	Poor
Treatment	Hormone therapy sensitive	Chemotherapy
Recurrence	Low rate	High rate

Table 1. Main characteristics of EEC (type I) and NEEC (type II).

EEC or endometrioid adenocarcinoma accounts for about an 80% of EC cases. It also includes mucinous, villoglandular and squamous features. This type of tumor is usually associated to peri-menopausal and post-menopausal women, estrogen exposition, obesity, and endometrial hyperplasia. It is a hormone dependent cancer. Around an 80% of the EEC cases are detected when the tumor is still confined to the uterus. EEC shows a differentiated histological pattern with glandular and papillary structures and low myometrial invasion. It use to present a better prognosis in comparison to NEEC. A smaller proportion of EECs (20%) present a more undifferentiated tumor, with less presence of glandules, more solid, and atypical cell nucleus [16,17].

NEEC or non endometrioid adenocarcinoma represents around 20% of the total EC cases. The papillary serous and cell clear subtype are included in this group among other histological subtypes that are less prevalent. They generally are a different clinical and histological entity, less differentiated, more aggressive. They present worse prognosis and use to appear at elder ages (post-menopause). It is a non-estrogen associated lesion that does not appear accompanied by endometrial hyperplasia but can develop from polyps or precancerous disorders. NEEC are usually diagnosed at advanced stages and are usually associated to recurrence [17].

The **tumor grade** is a good indicative to predict the dissemination of the disease and its outcome. Well-differentiated tumors have a probability of 90% of not having proximal lymph node affection. Also, the 5-year survival rate varies depending on the grade [1]. As explained in the next section, EEC can present several grades while NEEC are always associated to high grade tumors.

1.3.4.2. Histological classification of EC

Regarding the histology of the tumor and the characteristics of the cancer cells, we can distinguish between the following EC types (Table 2) [18]:

Subtype	Arquitectural grade	Incidence
Endometrioid adenocarcinoma:		
➢ Villoglandular (papillary) ➢ Secretory ➢ Ciliated cell ➢ Adenocarcinoma with squamous differentiation	• G1 (glandular normal characteristics with less than a 5% of solid non-squamous regions). • G2 (6 to 50% of solid non-squamous areas). • G3 (more than 50% of the tumor composed by non-squamous tumoral regions).	80-90%
Mucinous carcinoma		1-9%
Papillary serous adenocarcinoma	• G3 by definition	5-10%
Clear cell adenocarcinoma	• G3 by definition	1-5%
Undifferentiated carcinoma		
Mixed carcinoma		
Neuroendocrine tumors		

Table 2. Histological classification of the endometrial cancer.

Endometrioid adenocarcinoma

It is the most common type of EC (80%). It is usually presented as a well-differentiated tumor, with the presence of small and rounded glands without the stromal intervention. Several complexity levels can be observed in this type of EC and also several grades. Low grade tumors are formed by epithelial cells with uniform nuclei, no presence of atypia and small nucleoli. Cellular axes are perpendicular to the basement membrane, with or without stratification. High grade EECs show more atypia, pleomorphism and alterations in the nucleoli. Necrosis is not or slightly observed and stromal foam cells can be found [1].

Endometrioid adenocarcinoma includes (Figure 8):

- **Villoglandular carcinoma**: papillary architecture with stalks, cuboidal or columnar cells, nuclear pleomorphism and absence or slight stratification.
- **Secretory carcinoma**: with an incidence around 1%, it is characterized by the formation of glands similar to the ones observed in the secretory phase of a normal endometrium. Vacuolated cytoplasm in columnar cells is also observed.

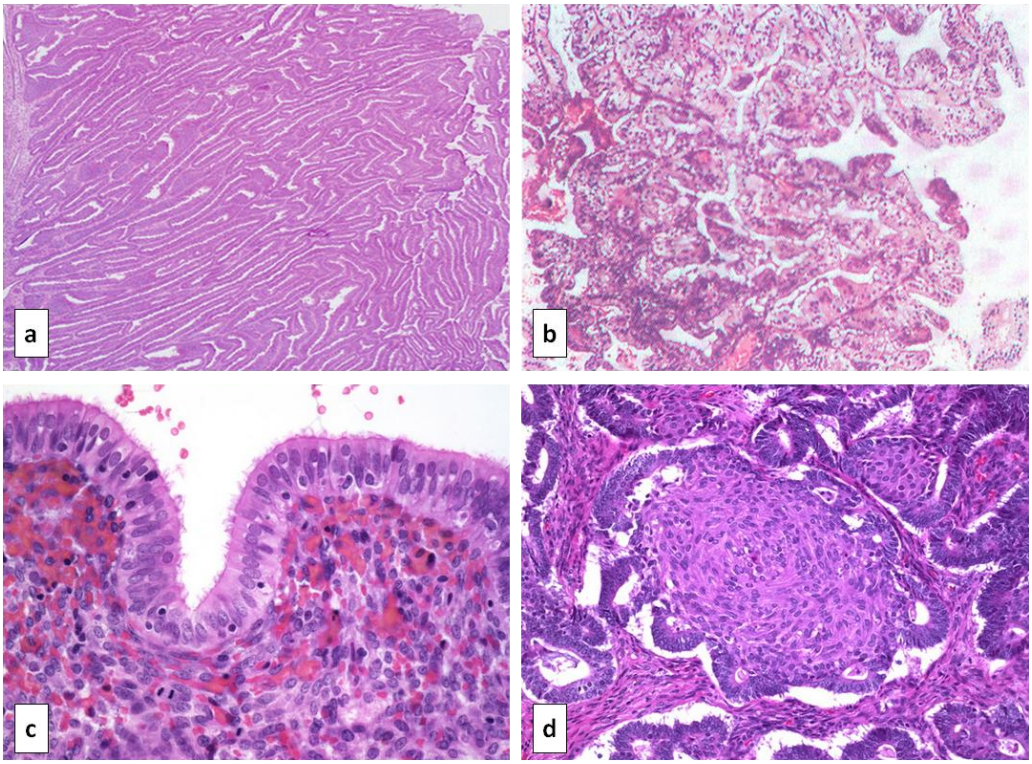


Figure 8. Endometrioid adenocarcinoma histology. **a**) Long papillary stalks lined by endometrioid-type cells are characteristic of the villoglandular or papillary variant of endometrial adenocarcinoma. **b**) Secretory carcinoma characterized by prominent cytoplasmic vacuoles and intraluminal secretions. **c**) Ciliated cell carcinoma. Most of the neoplastic glands are lined by ciliated cells showing mild to moderate nuclear atypia. **d**) Squamous differentiation in endometrial adenocarcinoma is frequent. Adapted from http://www.jpma.org.pk, http://www.webpathology.com and https://basicmedicalkey.com

- **Ciliated cell carcinoma**: it is a very rare type that is characterized by the presence of ciliated cells without nuclear atypia and eosinophilic cytoplasm surrounded by glands.
- **Adenocarcinoma with squamous differentiation variants**: representing about a 25% of the total number of EECs. The main features observed are the formation of polygonal cells with eosinophilic cytoplasm in between neoplasic glands. The cells form intercellular bridges [19].

Mucinous adenocarcinoma

It is not a very common type of EC (9%) and it is usually associated to low grade tumors. They can be included as type I tumors. The main feature of this histological type is the presence of mucin in the columnar cells cytoplasm. Nuclear middle atypia is observed and mitoses are unusual. Generally, the glands form papillary structures and they are cystically dilated with abundant intraluminal mucin (Figure 9).

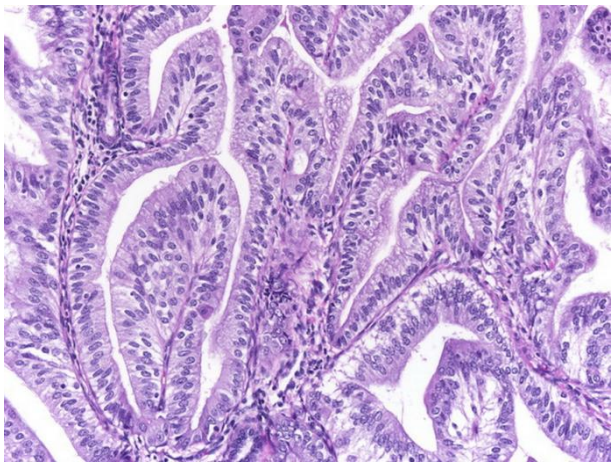


Figure 9. Mucinous carcinoma. It is generally well differentiated and is characterized by columnar cells with basally located nuclei and cytoplasm rich in mucin. Adapted from http://www.webpathology.com

The following types of ECs are considered NEEC tumors. They are aggressive and although they represent just about a 29% of the total number of ECs they have associated the lowest proportion of survivors [19].

Serous adenocarcinoma

Serous adenocarcinoma is a very aggressive type of EC with a 5-10% of incidence. It is usually presented with high myometrial and vascular invasion and associated to EC stages II and III. They usually spread to the peritoneum,

liver and lung. Patients with serous adenocarcinoma are usually diagnosed at advances ages, having a very poor prognosis associated. Histologically, usually presents a papillary structure with cellular budding, but lesions may grow in a solid or acinar pattern. Cell nuclei are not well differentiated [1,20] (Figure 10).

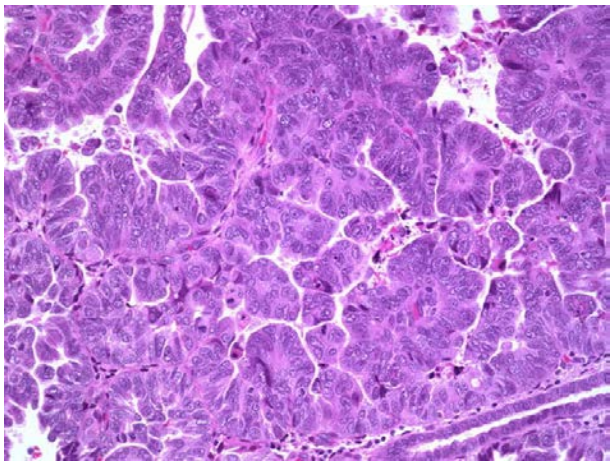


Figure 10. Serous carcinoma. Presents complex arborization of papillary structures. Cellular budding and tufting can often be appreciated under low power. Adapted from http://www.webpathology.com

Clear cell adenocarcinoma

Clear cell carcinoma represents a 1-5% of the total cases of EC. It is less common than the serous type but it also has a very low 5-year survival rate associated. It usually appears in postmenopausal women as a high grade disease. Presence of clear cells (because of the accumulation of glycogen in the cytoplasm of the cell), but also other cells like cuboidal cells are commonly found. The main patterns observed under the microscope are the papillary, tubulocystic, glandular and solid [19] (Figure 11).

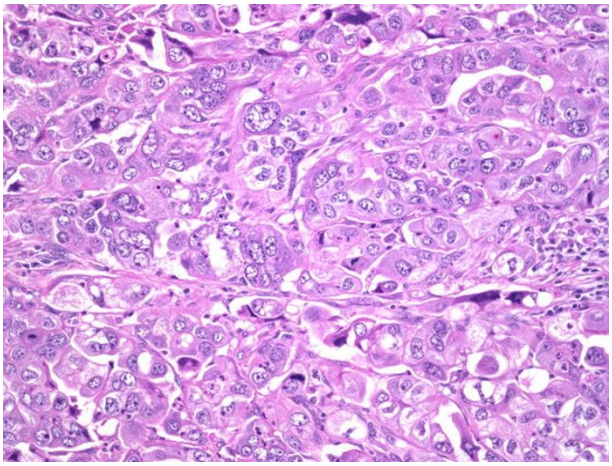


Figure 11. Clear cell carcinoma may exhibit a solid pattern consisting of sheets of clear cells separated focally by thin fibrous bands. Adapted from http://www.webpathology.com

Undifferentiated carcinoma

Undifferentiated carcinomas are a group of ECs that include those tumors that do not show glandular neither squamous differentiation. They include large and small cell carcinomas, characterized histologically by the observation of multinucleated giant cells or diffuse sheets of undifferentiated cells respectively (Figure 12). Both have a very poor associated survival rate.

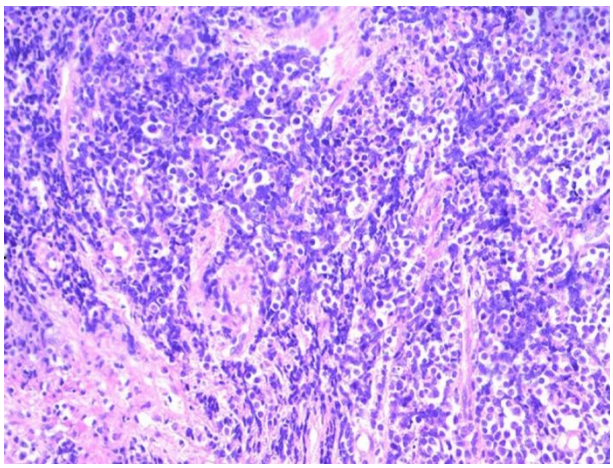


Figure 12. Small cell carcinoma. Adapted from Chen et al. [21].

Mixed carcinoma

Mixed carcinomas are tumors that contain different type of cells (each one accounting at least for a 10% of the total tumor). Not much is known about these tumors but they use to present a behavior similar to high grade ECs.

1.3.4.3. FIGO staging

Nowadays, EC is classified by stages according to the Federation of Gynecology and Obstetrics (FIGO). The FIGO staging is based on information retrieved by the uterine examination after its resection. The parameters that are assessed are the percentage of myometrial and cervical invasion, histological grade, metastasis to proximal organs and pelvic lymph nodes and distal dissemination. This classification, together with the clinic-pathological information, permits to classify patients based on their risk of recurrence [22–24].

FIGO defines staging criteria since 1988, but already in the 1950s FIGO elaborated a first set of rules for the classification of the gynecological cancers.

The last modification of FIGO staging was done in 2009 and published in 2010 (Table 3) [25].

Stage I*	**Tumor confined to the uterine corpus but not to the uterine serosa**		
	IA		No presence or < 50% of myometrial invasion.
	IB		≥ 50% of myometrial invasion.
Stage II*	**Tumoral invasion of the cervical stroma, but not extended beyond the uterus**		
Stage III*	**Local and/or regional dissemination of the tumor**		
	IIIA		Invasion of the serous layer of the uterine corpus and/or adnexae.
	IIIB		Vaginal dissemination.
	IIIC		Metastasis to pelvic and/or para-aortic lymph nodes.
		IIIC1	Positive pelvic nodes.
		IIIC2	Positive para-aortic nodes with or without implication of the pelvis nodes.
Stage IV*	**Bladder or intestinal mucosa invasion of the tumor and/or distal metastasis**		
	IVA		Bladder or intestinal mucosa involvement.
	IVB		Distal metastasis including intra-abdominal, extra-abdominal dissemination and/or inguinal lymphatic nodes.

*They can be grade 1, 2 or 3.

Table 3. FIGO endometrial cancer staging. Adapted from Pecorelli et al. 2009 [25].

1.3.4.4. Molecular classification: the TCGA model

Although there were already evidences of the molecular alterations associated to each EC type of the dualistic model (as seen in Table 1), The Cancer Genome Atlas (TCGA) has recently developed a classification system that includes more extensively the molecular alterations that characterize EC types.

Four molecular groups are defined [26]:

- **POLE ultramutated**. It is the smallest group representing about 10% of EECs. They are characterized by the presence of mutations in the exonuclease domain of POLE. They usually present good prognosis. Among the POLE ultramutated carcinomas, 60% are high grade EECs.
- **Microsatellite instability hypermutated**. This group presents a very high mutation rate and is characterized by high MH1 promoter mutilation events. It involves EEC tumors.
- **Copy-number-low**. This group is associated to low mutation rates (mainly affecting PTEN), frequent (52%) mutations in CTNNB1 and presents intermediate prognosis.
- **Copy-number-high**. Group composed mostly of serous-like tumors. They present genomic instability and TP53 mutations (>90% of the cases).

1.3.5. Prognostic factors

Generally, EC presents a favorable prognosis since it is usually detected at early stages. However, some patients recur when the tumor is still confined to the uterus. This is because there are relevant subgroups of patients that present poor prognosis. For this reason, it is very important to establish accurate predictive and prognostic factors in order to identify those patients presenting poor prognosis in order to be able to select the best primary and adjuvant treatment.

The EC prognostic factors are commonly divided into uterine and extra-uterine factors. Uterine factors include histological grade, histological type, depth of myometrial invasion, presence of atypical endometrial hyperplasia, vascular invasion, cervical affection, DNA ploidy and S-phase fraction and hormone receptors expression. On the other hand, extra-uterine factors include positive peritoneal cytology, adnexal involvement, pelvic and/or paraaortic lymph node metastasis, and peritoneal metastasis. The most significant prognostic factors are the histological grade and type, the percentage of myometrial invasion and the affection of the lymph nodes [27].

The FIGO stage is the most important individual prognostic factor and, for this reason, it is the most used as reference. Advanced FIGO stages are associated

to significant reduction of the survival. The involvement of regional lymph nodes is also an important factor to consider although the benefits of a lymphadenectomy in women presenting early stage tumors is still controversial.

It is necessary to assess the prognostic factors before and after the surgery treatment in order to define the stage of the tumor properly and the risk of recurrence that each patient has associated [28].

In clinics, all this information is needed to classify patients into different recurrence risk (Table 4). The 5-year risk of recurrence for patients included in the intermediate and high-intermediate risk groups is established to be between 20 and 25%. Differently, the high and advanced risk group present around 30-65% of recurrence risk. Interestingly, NEEC only represent a 10% of diagnosed ECs but they account for more than 50% of total recurrences and deaths [29].

Risk group	Description
Low risk	· Stage IA G1-2 with no LVSI (Type I)
Intermediate risk	· Stage IB G1-2 with no LVSI (Type I)
High-intermediate risk	· Stage IA G3 regardless of LVSI (Type I) · Stage I G1-2, LVSI unequivocally positive, regardless of depth of invasion (Type I)
High risk	· Stage IB G3 regardless of LVSI (Type I) · Stage II · Stage III (Type I), no residual disease · All Type II
Advanced	· Stage III residual disease · Stage IVA
Metastatic	· Stage IVB

Table 4. Risk stratification used to guide adjuvant treatment, Colombo et al. [29]. (FIGO 2009 staging used; molecular factors were considered but not included; tumor size was considered but not included; nodal status may be considered for treatment recommendations (G: grade, LVSI, lymphovascular space invasion).

1.3.6. Treatment

1.3.6.1. Preoperative risk assessment

According to the recommendations adopted by the ESGO-ESMO-ESTRO consensus conference, an extensive evaluation is mandatory before surgery. It must include family history, list of comorbidities, geriatric assessment, clinical examination, transvaginal or transrectal ultrasound, pathology assessment (type and grade of the tumor) of an endometrial biopsy or curettage sample. Endometrial biopsies are obtained by aspiration o guided hysteroscopy and serve as a confirmatory diagnostic sample and also to assess the tumor grade and histological type [29].

The pre-operative staging (defined by the type and grade of the tumor, the percentage of myometrial invasion and the cervical involvement) and the medical condition of the patient will guide the extent of the surgery, which is the primary EC treatment. The magnetic resonance imaging (MRI) is considered the preferred imaging technique for preoperative staging, especially to determine myometrial invasion, lymphatic metastases and cervical involvement. The definitive staging of the tumor will be determined after surgery and it is known as clinical staging [30].

The main objective of achieving a proper risk assessment is to classify correctly patients into those groups of risk for lymph nodes dissemination and disease recurrence to define the best surgical treatment.

1.3.6.2. Surgery

The first step in the EC treatment is a surgical intervention based in a **hysterectomy** with bilateral salpingo-oophorectomy. Depending on the tumor characteristics and the stage of the disease, a full pelvic and para-aortic lymphadenectomy will be also required.

The surgical procedures that are recommended depending on the tumor stage are listed in Table 5:

Preoperative staging		Recommended surgical procedure
Stage I	· IA; G1–G2	Hysterectomy with bilateral salpingo-oophorectomy
	· IA; G3	Hysterectomy with bilateral salpingo-oophorectomy ± bilateral pelvic-para-aortic lymphadenectomy
	· IB; G1-G2-G3	Hysterectomy with bilateral salpingo-oophorectomy ± bilateral pelvic-para-aortic lymphadenectomy
Stage II		Radical hysterectomy with bilateral salpingo-oophorectomy and bilateral pelvic-para-aortic lymphadenectomy
Stage III		Maximal surgical cytoreduction with a good performance status
Stage IV	· IVA	Anterior and posterior pelvic exenteration
	· IVB	Systemic therapeutical approach with palliative surgery
Serous and cell clear		Hysterectomy with bilateral salpingo-oophorectomy, bilateral pelvic-para-aortic lymphadenectomy, omentectomy, appendectomy and peritoneal biopsies

Table 5. Surgical treatment based on tumoral staging. Adapted from Colombo et al. [31]. (G: grade).

1.3.6.3. Adjuvant treatment

Most of the EC patients are diagnosed with low risk of recurrence and are just treated with surgery. However, the disease stage and the recurrence risk of the patient will define the recommended adjuvant treatment (Table 6).

Preoperative staging		Recommended adjuvant treatment
Stage I	· IA; G1–G2	Observation
	· IA; G3	Observation or vaginal BT (if NPF: pelvic RT and/or adjunctive chemotherapy could be considered)
	· IB; G1-G2	Observation or vaginal BT (if NPF: pelvic RT and/or adjunctive chemotherapy could be considered)
	· IB; G3	Pelvic RT (if NPF: combination of radiation and chemotherapy could be considered)
Stage II		Pelvic RT and vaginal BT · If grade 1–2 tumor, myometrial invasion <50%, negative LVSI and complete surgical staging: BT alone. · If NPF: chemotherapy + RT
Stage III-IV		Chemotherapy · If positive nodes: sequential radiotherapy · If metastatic disease: chemotherapy – RT for palliative treatment

Table 6. Recommended adjuvant treatment based on tumoral staging. Adapted from Colombo et al. [31]. (G: grade, NPF: negative predictive factor, BT: brachytherapy, RT: radiotherapy, LVSI, lymphovascular space invasion).

2. BIOMARKERS FOR ENDOMETRIAL CANCER

2.1. BIOMARKER DEFINITION AND PIPELINE

The working Group of the National Institute of Health defined a **biomarker** as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention". An ideal biomarker should be easily obtained with minimum discomfort or risk to the patient, specific, sensitive, reproducible, objective, quantifiable and economical [32]. In spite of the high and increasing incidence of EC, current tools to detect the disease at early stages are insufficient to manage it and to decrease its associated mortality rate; thus, molecular **diagnostic biomarkers** (preferably in non-invasive samples) are needed [33] (Box 1). The limiting factor in the use of biomarkers in the diagnosis of EC is their lack of specificity [34]. Biomarkers might include any type of bioactive material, such as DNA, RNA, proteins and metabolites.

Box 1. Biomarker classification based on their application:

Predisposition or screening biomarker: is a predictor of the probability of developing a specific disease.

Diagnostic biomarker: indicates if an individual has or not a specific disease.

Prognostic biomarker: indicative of the disease course in a patient already diagnosed with a specific pathology, usually after standard treatments.

Predictive or therapeutic biomarker: helps determining which patients are most likely to benefit from a specific treatment option. It has a potential value for improving new and more personalized therapies.

The procedure for a molecular biomarker panel development includes 3 phases: discovery, verification and validation. Following this pipeline, the final goal is to

translate a specific biomarker or panel of biomarkers for clinical use for improving patient management and clinical outcomes. Through the biomarker development pipeline, while the number of candidates is invariably reduced to a few validated biomarkers, the number of human samples increases significantly in order to test and establish clinical efficacy of the biomarker panels (Figure 14).

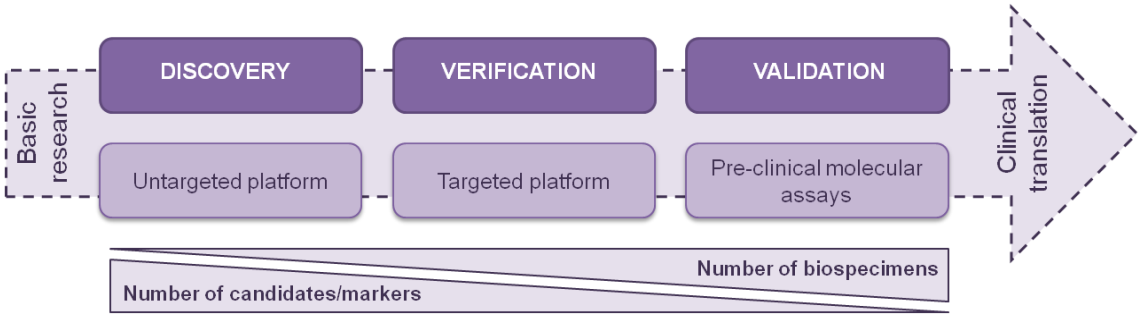


Figure 14. Biomarker development pipeline.

2.2. ENDOMETRIAL CANCER DIAGNOSTIC BIOMARKERS

Over the last decades, many studies have been performed in order to discover diagnostic biomarkers for EC. Transcriptomics and proteomics have been the most explored and promising fields until the moment for the biomarker identification. From 1978 to 2017, over 700 proteins have been identified as candidate biomarkers for EC. The most robust candidate biomarkers would be those that are identified and validated in an independent cohort of samples and by different research groups. Hence, 10 proteins have been described as possible EC diagnostic biomarkers in 3 or more validation studies: HE4, CA125, SLC2A1, MMP9, FOLR1, VEGFA, CD44, SAA1, TP53 and BCL2. The two most investigated proteins as EC biomarkers have been HE4 and CA125 [35].

Some of these studies demonstrated possible molecular diagnosis applicability. Unfortunately, none of the studies reported presented candidates with enough sensitivity, specificity and reproducibility compared to the current clinical diagnostic tools.

2.3. BIOMARKER SOURCES: CLINICAL SAMPLES

Most of the discoveries in EC have been performed using tissue biopsies as a biomarker source, followed by plasma/serum [35]. The main advantage of using tissues for a discovery phase of a biomarker pipeline is that any potential biomarker in the altered tissue will present a higher concentration compared to any other biofluid. However, profiling approaches of tissues are subject to sampling bias since they only provide a snapshot of the tumoral heterogeneity and the collection of the sample requires invasive procedures for the patient [36].

In contrast, **liquid biopsies** (containing circulating cells, nucleics acids, proteins, lipids, extracellular vesicles and other molecules) can be used for clinical oncology studies since they can capture better the molecular heterogeneity of primary and metastatic cancers and the dynamic information of the tumors at the time of blood drawing. Blood serum or plasma are the most used clinical fluid samples for the identification of biomarkers since they are collected routinely in the laboratory in a rapid and minimally invasive way. In addition to blood, many other human body fluids containing tumoral information can be used, such us saliva, ascites fluid, urine, pleural effusions, cerebrospinal fluid, amniotic fluid, etc [37]. In EC research, the use of a proximal biofluid such as **endometrial aspirates** (also called pipelle biopsies or uterine aspirates) could be more appropriate for a discovery since they are in direct contact with the site of disease, the endometrium in this case. The proximal biofluids are likely to be enriched in potential biomarkers coming from the altered tissue and represent an alternative matrix for biomarker discoveries [36,38]. Moreover, endometrial aspirates can be obtained quickly and routinely in the doctor's office with minimal risk for the patient (minimally-invasive sample). They also present some limitations including small sample volume, higher heterogeneity and diluted concentration of molecules and possible blood contamination. In the literature, just nine studies analyzing endometrial aspirates for discovery purposes in EC have been published (see Table 7).

Publication title	Authors	Year	Journal	PMID
Diagnosis of endometrial cancer in patients with postmenopausal bleeding by analysis of the lactate dehydrogenase isoenzyme activity profile in uterine fluid	Niklasson O et al. [39]	2004	Gynecological Oncology	15099950
Transvaginal ultrasound and lactate dehydrogenase isoenzyme activity profile in uterine aspirate for diagnosis of endometrial carcinoma in women with postmenopausal bleeding	Niklasson O et al. [40]	2007	International Journal of Gynecological Cancer	17367317
Comprehensive proteomic analysis of human endometrial fluid aspirate	Casado-Vela J et al. [41]	2009	Journal of Proteome Research	19670903
Digital morphometry of cytologic aspirate endometrial samples	Mahovlić V et al. [42]	2010	Collegium Antropologicum	20437635
Molecular markers of endometrial carcinoma detected in uterine aspirates	Colas E et al. [38]	2011	International Journal of Cancer	21207424
Molecular diagnosis of endometrial cancer from uterine aspirates	Perez-Sanchez C et al. [38]	2013	International Journal of Cancer	23649867
Development of a sequential workflow based on LC-PRM for the verification of endometrial cancer protein biomarkers in uterine aspirate samples	Martinez-Garcia E et al. [43]	2016	Oncotarget	27447978
Targeted proteomics identifies proteomic signatures in liquid biopsies of the endometrium to diagnose endometrial cancer and assist in the prediction of the optimal surgical treatment	Martinez-Garcia E et al. [44]	2017	Clinical Cancer Research	28790116
Genetic analysis of uterine aspirates improves the diagnostic value and captures the intra-tumor heterogeneity of endometrial cancers	Mota A et al. [36]	2017	Modern Pathology	27586201

Table 7. Publications in endometrial aspirates associated to endometrial cancer discoveries.

2.4. EXTRACELLULAR VESICLES AND EXOSOMES

Extracellular vesicles (EVs) are present in liquid biopsies; they are released by many cell types (including cancer cells) into the extracellular milieu. EVs are phospholipid bilayer-enclosed vesicles that have emerged as key mediators of cell-to-cell communication in either physiological or pathological situations via the horizontal transference of biologically active cargo (i.e. proteins, nucleic acids, enzymes, signaling molecules, mRNA, miRNAs, long non-coding RNAs, lipids, sugars, oncomolecules) suggesting that they can modulate the activity of the recipient cell (Figure 15). In physiological conditions, they can act as immune-modulators, and can participate in specific processes like programmed cell death, angiogenesis, protein trafficking, inflammation and coagulation. Over the years, researchers have indicated also their role as potential circulating biomarkers. Moreover, the role of EVs promoting epithelial to mesenchymal transition (EMT) processes, the formation of pre-metastatic niches and inducing metastasis has increased significantly [37,45] . They also have capabilities modifying the metabolic environment, remodeling the vascular organization, modulating immune suppression, regulating tumoral transformation and growth, acquiring therapy resistance and determining the organotropic homing [46,47].

EVs are classified into two groups: **microvesicles** (MVs) and **exosomes**. MVs shed directly from the cellular plasma membrane by budding and they range between 100 and 1500 nm; whilst exosomes are produced by budding of the late endosome or multivesicular bodies (MVBs) that fuse with the plasma membrane realizing their vesicular content as exosomes. Exosomes range between 30 and 120 nm [45,48]. There are still limitations to confirm the intra-luminal origin of the isolated vesicles (i.e. mainly lack of specific biomarkers). Thus, enriched vesicles having similar morphology and size and unknown biogenesis are defined as Exosome-Like Vesicles (ELVs). In Chapter I of the Results section we will use this nomenclature.

EVs can be isolated from body fluids (i.e. plasma, urine, pleural effusions, breast milk, and saliva) by centrifugation-based techniques such differential

centrifugation procedures or density gradient centrifugation protocols (i.e. sucrose gradient) coupled with ultracentrifugation. An important disadvantage of

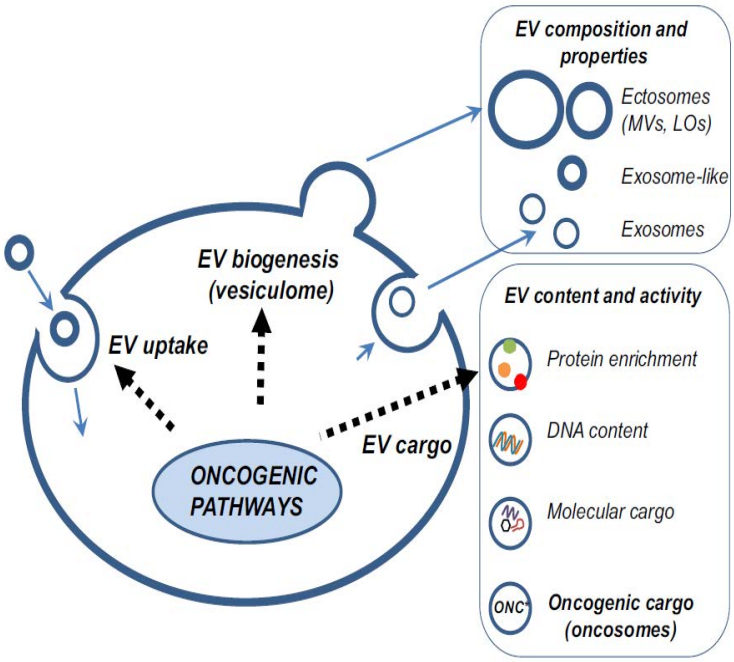


Figure 15. Impact of oncogenic transformation on EVs mediated communication in cancer. Oncogenic pathways affect a wide spectrum of genes, and through this influence, the production, biogenetic pathways (vesiculome), molecular content, protein loading, wider molecular profile, oncogenic cargo and the uptake of EVs in cancer. Cancer-related distortion of cellular signalling influences the output of major classes of EVs, such as ectosomes, MVs, large oncosomes (LO), and exosomes. This effect also often extends to protein enrichment of cancer EVs, their molecular composition, abnormal emission of DNA and exit of oncogenic macromolecules themselves, as oncosomes. Adapted from Choi et al. [49].

using centrifugation coupled to ultracentrifugation-based protocols is the inability to separate subtypes of EVs (i.e. sedimentation of MVs together with apoptotic bodies). The addition of filtration steps or the use of density gradient methods enhances the purity of the exosomal population, but these techniques require more sample volume and are time consuming. In an effort to find an optimal isolation procedure, some other alternative methods for exosomal enrichment have emerged such as immunoaffinity capture, polymer-mediated precipitation and filtration-based protocols. Moreover, many commercial kits

have been developed during the last few years in order to isolate EVs in a quick and easy way (i.e. ExoQuick and bead-based protocols). These kits have associated certain limitations since they are quite expensive and most of them co-precipitate EVs with other protein complexes. There is no consensus yet for a standard isolation method of EVs and exosomes, for this reason, the isolation protocol has to be elected regarding the scientific purpose.

Exosomes are considered to be valuable sources for biomarkers due to their presence and stability in most body fluids. Specifically, exosomes derived from tumoral cells likely serve as biomarker for early detection of cancer as they transport the cargo that reflects genetic and signaling alterations in the original cancer cells. As a result, exosome-based diagnostics are expected to provide higher sensitivity and specificity over conventional biopsy or liquid biopsy biomarkers. In consequence, exosomes are qualified as minimally invasive biomarkers for early detection, diagnosis and prognosis of cancer [50]. Some of the most relevant pre-clinical and clinical studies that identified diagnostic, prognostic or predictive biomarkers present in patient exosomes are listed in Table 8.

Cancer type	Exosomal protein	Body fluid	References (PMID)
Breast	CD24, EpCAM, EDIL3, Fibronectin, Survivin 2B, Survivin, CEA, Tumor antigen 15-3	Blood, ascites fluid	21601258, 26603257, 27250024, 24620748, 18507061, 24620748
Prostate	Survivin, PCA-3, TMPRSS2, ß-catenin, PSA, PSMA, ITGA3, ITGB1, PTEN	Blood, urine	23091600, 19401683, 25688242
Pancreatic	Glypican-1, MIF	Blood	26106858, 25985394
Ovarian	TGFß1, MAGE 3/6, CD24, EpCAM, CA125, Claudin4, L1CAM, ADAM10, EMMPRIN	Blood	19619303, 24466501
Glioblastoma	EGFR, CD63	Blood	23142818
Colorectal	CD147,CD9	Blood	24710016
Bladder	TACSTD2, EDIL3, Mucin4, EPS8L2, α6integrin, MUC1, Basigin	Urine	23082778
Lung	EpCAM, EGFR, CEA, LGR1	Blood	25735706, 21557262, 24400444
Melanoma	CD63, Caveolin1, TYRP2, VLA4, HSP70	Blood	22635005, 19381331, 24952934
Nasopharyngea	LMP1, Galectin-9, BARF1	Blood, saliva	17156439
Renal cell	MMP9, EMMPRIN, carbonic anhydrase	Urine	23511837
Spongioblastoma	PDPN, IDH1, EGFR-VII, EGFR	Blood	24952934, 19011622, 23142818
Stomach	HER2, CCR6	Blood	20043223

Table 8. Exosomal protein biomarkers from body fluids of patients in pre-clinical and clinical studies. Adapted from Soung et al. [50] and Li et al. [51].

3. CLINICAL METABOLOMICS

Metabolomics is a relatively new "omics" science that was recognized at the end of 1990s [52]. Like the other classic "omics" (genomics, transcriptomics and proteomics) it is considered to be an important tool in order to be able to analyze and understand the complexity of the physiological situation compared to the older reductionist approaches that used to study individual genes, transcripts or proteins [53]. The field of metabolomics is centered in the quantitative analysis of a large set of small molecules (<10 KDa) that are the final products of the cell activity giving as a result an instantaneous snapshot of the complex physiological status of a biological system (Figure 16). It is a phenotypic readout of the changes produced in a cell, tissue, organ or organism as a response to internal or external stimuli [54].

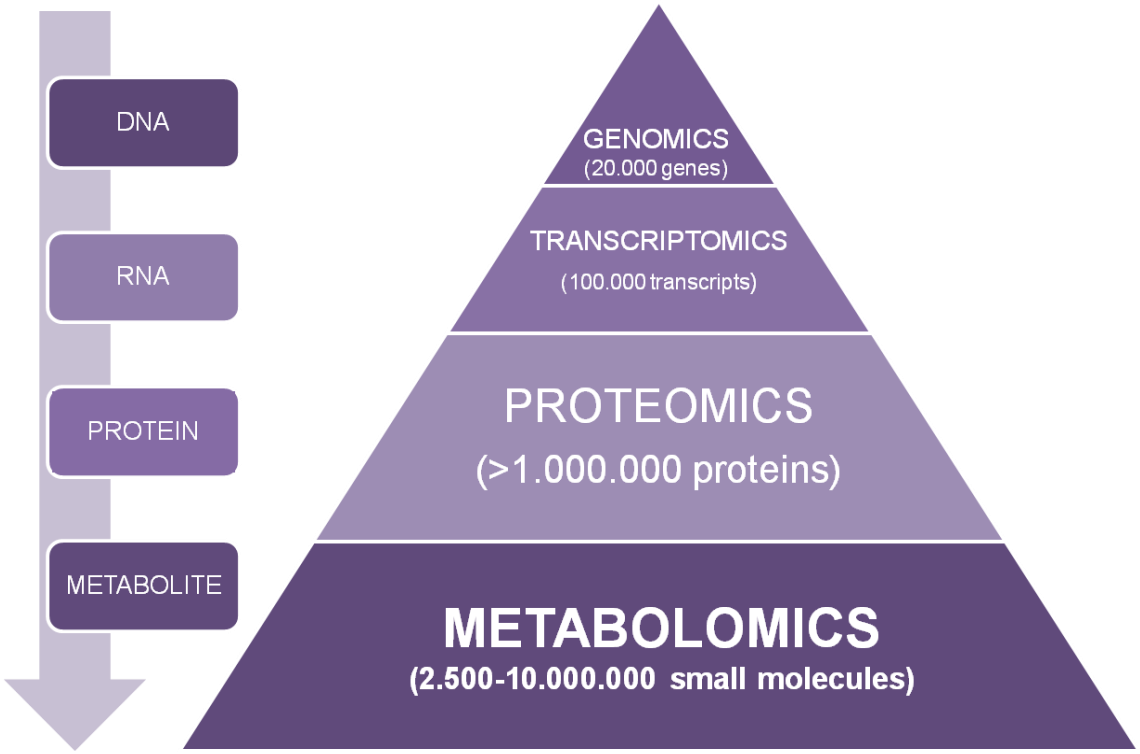


Figure 16. Omics pyramid of life. Alterations at the genome and proteome level (physiological changes) can be detected in the metabolome. The metabolome also responds to environmental changes.

The **metabolome** is the integration of a cell/organism genetic, nutritional, pharmacological and environmental status [55]. Thus, the metabolome can be fractionated into (Figure 17):

- Endogenous metabolome: set of metabolites characteristic of a specific biological system in a specific physiological situation.
- Exogenous metabolome: metabolites derived from the microflora (microbial metabolome) and from drugs, pollutants and nutritional factors (xenometabolome) [56].

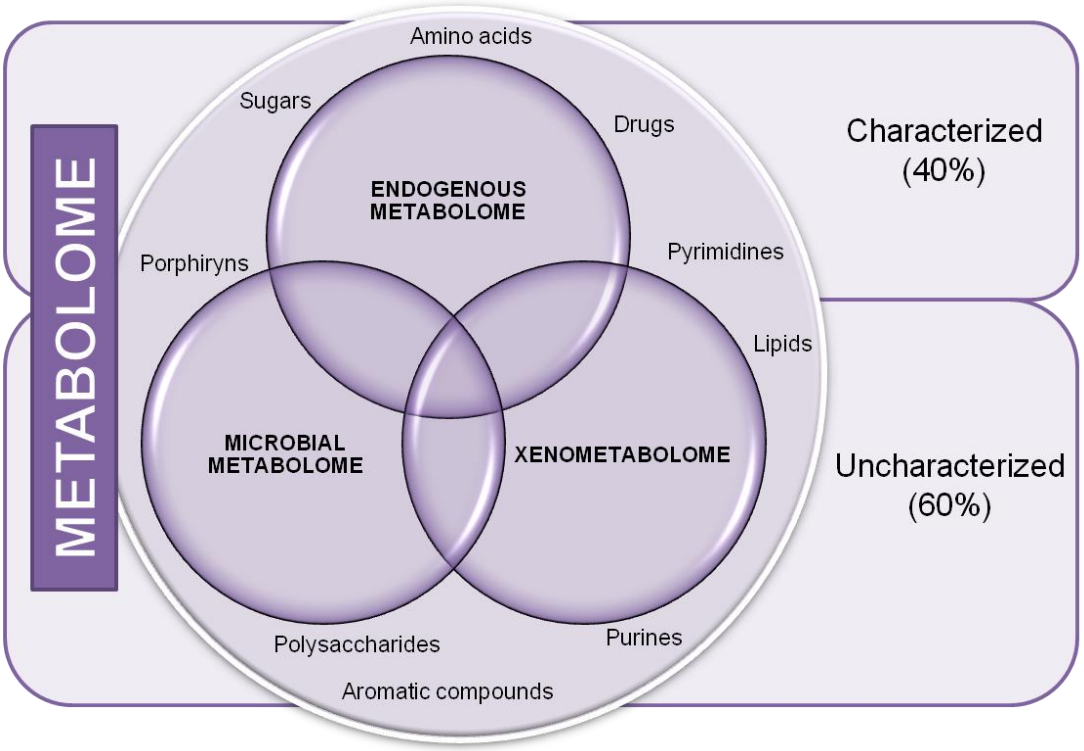


Figure 17. Human metabolome composition.

One of the advantages of the metabolomics is that the basic metabolic pathways and metabolites are much more conserved among different species than the genome, transcriptome and proteome. Hence, metabolomics can be considered a standard language among organisms integrating the complexity of life [57]. Moreover, subtle alterations at gene, transcript or protein level can result

in big changes in the metabolome offering a potential advantage of this field in terms of sensitivity and specificity compared to conventional clinical approaches [58].

The analysis of the metabolome in a specific physiological status such health or disease, will be unique for each individual or system offering the possibility to detect potential risks for a biological disorder that will finally help developing advances in terms of individualized treatments [58].

Biomedicine has been the area that has taken more advantage of metabolomics advances. The investigation in the metabolomics field in the context of human disorders have been focused in: a) the understanding of the molecular basis and metabolic pathway alterations of human disease; b) the identification of diagnostic and prognostic biomarkers that offer higher sensitivity and specificity; c) the translational applications in the metabolomic field including the metabolome reprogramming, defined as the use of metabolomic techniques for preventive or disease treatment purposes through the reconstitution of perturbed metabolic pathways [57].

In a biological system, the chemical transformations that generate and module the metabolome are catalyzed by enzymes. They are the responsible of the metabolic cascades that define each metabolic pathway maintaining the cellular homeostasis. Metabolites are the end products of the metabolism. The regulation of the metabolism can be performed at a genetic, transcriptional, post-transcriptional and allosteric level. Exogenous crucial metabolites can reprogram the metabolism since the abundance of a determined metabolite modulate the activity of several enzymes, affecting the metabolic pathways and networks [57].

3.1. METABOLOMIC APPROACHES

The two approaches commonly used for the recovery and identification of metabolites are targeted or untargeted (also known as global) mass-spectrometry-based metabolomics (Figure 18):

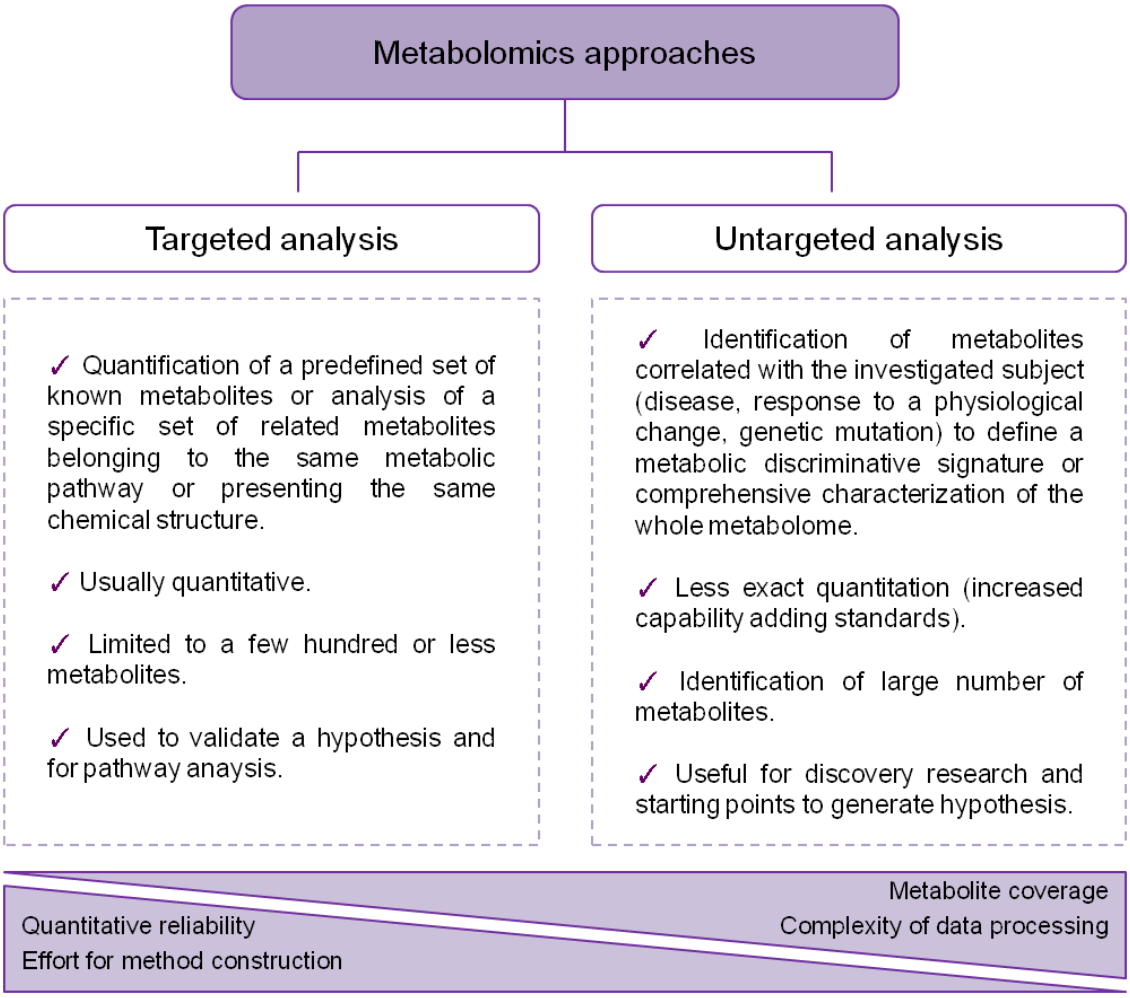


Figure 18. Metabolomics approaches. Comparison of targeted and untargeted analysis.

- **Untargeted or global metabolomics** is focused on the analysis of a wide range of small molecules that are not previously known. The biochemical and physical properties of the recovered molecules will be

determined by the extraction and separation methods used. The large amount of data obtained will require complex bioinformatics tools. This technique is very effective to detect novel perturbations. Some small unknown molecules detected may remain unannotated in the metabolomics databases.

- **Targeted metabolomics** measures a set of metabolites previously known with a higher sensitivity and specificity. The quantification can be much more accurate. The concentration of the analyte in a biological sample is extrapolated from a standard curve for the metabolite of interest. This approach is an important part of the metabolomics workflow since it provides analytical validation for the biomarkers identified by untargeted techniques [59].

3.2. METABOLOMICS WORKFLOW

The use of metabolomics in biomedicine assumes that one alteration in a biosystem (i.e. disease) will disturb the pathway homeostasis resulting in changes in the metabolomic phenotype that will result in different metabolomic signatures compared to the control situation (i.e. healthy subject). To identify this differential profile of the metabolome and be able to validate it for further clinical applications, an adequate, standard and robust metabolomic workflow must be followed. Moreover, due to the differences in the physic-chemical properties of the metabolites, metabolomic approaches will always have a certain grade of bias. This is another reason why the selection, extraction, detection, and analysis of the metabolites must be optimized to draw accurate conclusions [60].

A metabolomic approach includes several steps that are represented in Figure 19:

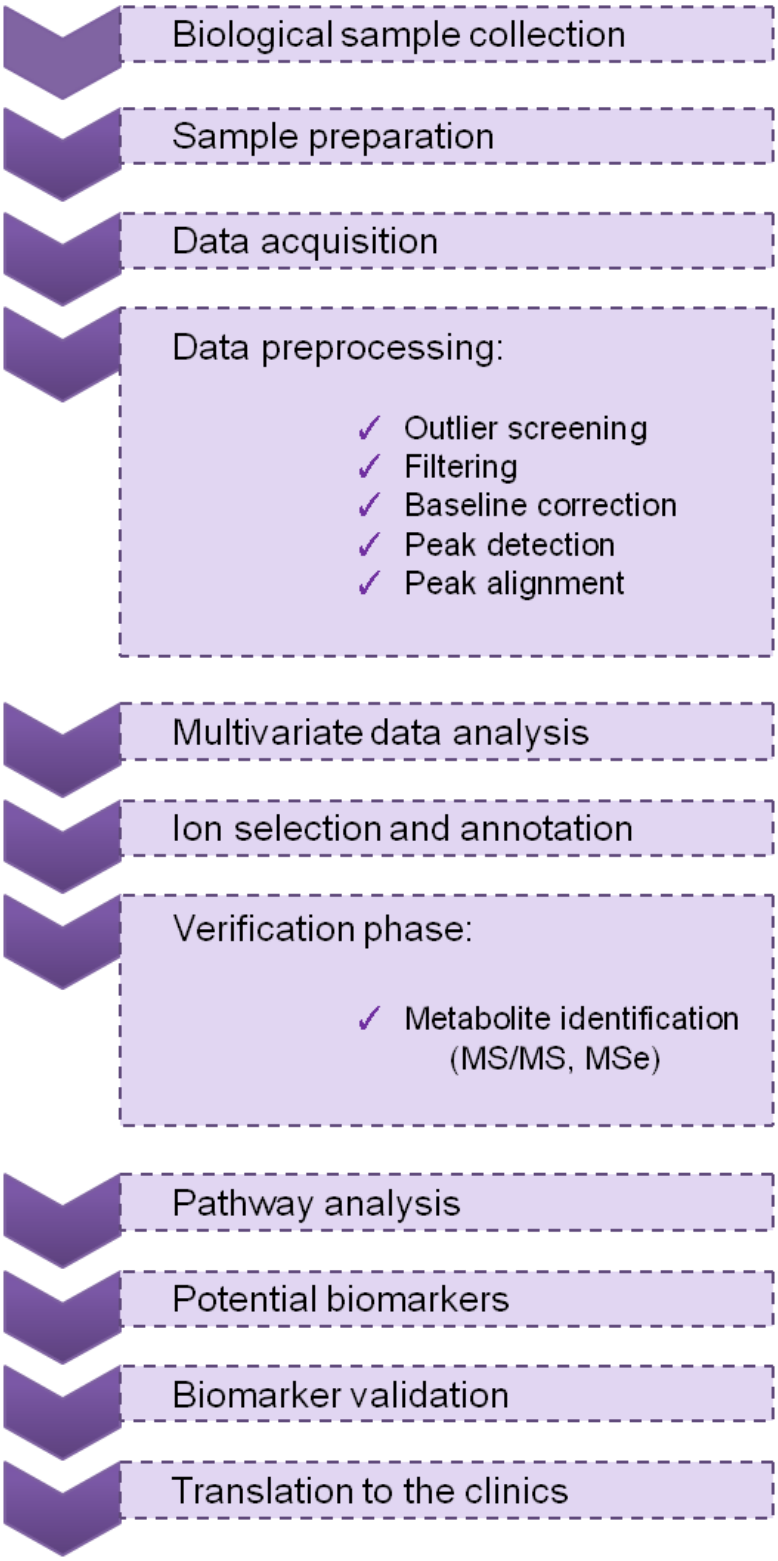


Figure 19. Metabolomics workflow.

3.2.1. Sample collection and preparation

There are many types of samples that can be analyzed using metabolomic approaches including whole organisms (i.e. bacteria), tissues, biofluids and cells. As cited in Section 2, liquid biopsies and EVs are considered valuable sources for metabolomic studies. Samples must be properly labeled and stored (preferably at - 80 °C).

Sample collection, preparation and storage have to be consistent in order to prevent biased analysis, degradation of metabolites or appearance of contaminants external to the original sample. This is also important to allow the comparison of the abundance of the metabolites of interest among samples.

Sample preparation requires the extraction of metabolites of a specific sample into a liquid form since it is not possible to analyze any sample without a previous preparation (but always trying to work with minimal pretreatment of the sample) [61]. In order to avoid degradation or variations of the original sample metabolome and depending on the type of sample, several stabilization methods will be used. These quenching methods include freezing the sample or the use of chemicals like nitric acid (for protein precipitation), sodium azide (to prevent bacterial growth) and antioxidants.

Due to the heterogeneity of the metabolites, it is impossible to establish a standard protocol or platform to cover the entire metabolome. The **platform** is defined by the sample preparation method and the chromatographic technique. Usually combinations of several extraction protocols (for polar/hydrophilic and non-polar/lipophilic metabolites) are followed for each sample in order to get the best possible coverage. During metabolomics experimental procedure, small molecules losses may occur due to protein co-precipitation, low solubility with the solvent or solvent saturation effects. These events are matrix-dependent and this is the main reason why a specific biological sample is usually prepared by different metabolomic extraction methods. While each metabolomic platform has different sample pretreatment requirements, deproteinization is required for all methods because the protein presence may strongly influence precision, accuracy and instrument lifetime [62]. For either targeted or untargeted approaches, generally a final step of the sample preparation requires

evaporation to concentrate the molecules and to be able to resuspend them with the desired solvent right before the metabolomic analysis [63].

For the extraction of **polar metabolites**, the most commonly used extraction method is organic solvent-based protein precipitation followed by centrifugation or membrane techniques (ultracentrifugation). For the extraction of **non-polar metabolites**, mostly lipids, it is commonly used a mixture of chloroform/methanol (in a ratio 2:1 or 1:2) with subsequent addition of 1 volume of chloroform and 1 volume of water. In both cases, for studies with large amount of samples automated protocols using 96 well plate formats are also commercially available [63].

3.2.2. Data acquisition

There are two main methods used in metabolomics to obtain the data from the prepared biological samples, **nuclear magnetic resonance** (NMR) spectroscopy and **mass spectrometry** (MS). NMR was used for the first experiments, but from 2005, MS techniques have been chosen [64]. Both are semi or totally quantitative techniques that measure the concentration of small molecules.

NMR spectroscopy is based on the exploitation of magnetic properties allowing the identification of different atomic nuclei based on their resonant frequencies, which are dependent on their locations in the molecule. It is associated to quick and easy sample preparation procedures. It is a rapid, non-destructive and non-invasive technique that provides highly quantitative and reproducible (> 98%) results. The observed unique pattern of resonance peaks can be reliably assigned to a specific metabolite; and the area of each peak is representative of the relative concentration of each nuclei resonating at a particular frequency. NMR has the ability to analyze liquid but also solid samples and, moreover, provides structural information. NMR spectroscopy is used to determine the effect of drugs, toxins, to study diseases and to trace metabolic pathways. However, compared to MS-based approaches, NMR is less sensitive resulting in difficulties to detect low-abundance metabolites. For this reason, this technique requires large sample volumes and involves more expensive

instrumentation. These are the main reasons why metabolomics community shifted to MS techniques [64–66].

MS is a highly sensitive technique that allows the detection, quantitation and structure elucidation of metabolites. This technique allows the detection of thousand small metabolites in a single experiment. It is an excellent analytical platform (Box 2) for metabolomic studies since it provides very reproducible and versatile data. MS-based experiments usually require more extensive sample preparation procedures compared to NMR. It measures the masses of molecules and their fragments to determine their identity by measuring the mass-to-charge ratio (*m*/*z*) of ions that are formed by inducing the loss or gain of a charge from neutral species. The prepared sample is injected into the mass spectrometer in liquid phase. Although direct sample infusion can also be implemented (i.e. high-throughput experiments), most MS projects require a prior chromatographic separation method before entering the mass spectrometer that minimizes signal suppression and allows for greater sensitivity; and by providing a retention time identifier it can further aid metabolite identification. This separation reduces the complexity of the mass spectra due to a separation of metabolites in a time dimension (Figure 20). The main separation methods coupled to MS and used in the literature are gas chromatography (GC) and liquid chromatography (LC) [59,66].

- **Gas Chromatography-Mass Spectrometry (GC-MS)** allows the restricted detection of compounds that are volatile or those which can be volatilized. One limitation of GC-MS techniques is that they require a proper derivatization procedure that is time-consuming, costly and may cause artifacts. The derivatization decreases the polarity and increase the volatility of compounds. Moreover, improves chromatographic resolution and peak shape and increases thermal stability of some compounds. However, an advantage that GC-MS has compared to LC-MS is the existence of extensive databases based on fragmentation patterns that help the identification of metabolites [64,67].

- **Liquid Chromatography-Mass Spectrometry (LC-MS)** is the technique predominantly used as it enables the detection of the highest number of metabolites (hydrophilic and hydrophobic) simultaneously, requires minimal amounts of sample and a sample derivatization step is usually not needed [68]. The most common LC-based method used for global metabolomics is the reversed-phase liquid chromatography-MS (RP-LC-MS) using C18 columns because results in the separation and detection of both polar and non-polar molecules. However, in order to avoid the elution of highly polar compounds in the void volume (a limitation of RP columns), hydrophilic-interaction chromatography (HILIC) is also applied in metabolomics experiments [61].

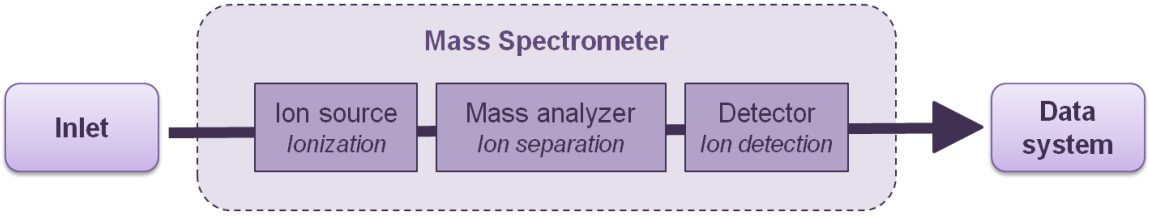


Figure 20. Schematic of a mass spectrometer.

In LC-MS techniques, the ionization efficiency depends on the physicochemical characteristics of the molecule, the ion source and the ionization mechanism (ionization method in which energetic electrons interact with atoms or molecules to produce ions). LC-MS approaches use commonly electrospray ionization (ESI) that has a typical disadvantage effect called ion suppression. Ion suppression refers to reduced detector response as a manifested effect of competition for ionization efficiency in the ionization source between the analyte of interest and other species which have not been removed from the sample matrix during sample preparation [69].

Box 2. Mass Analyzers

Triple quadrupole (QQQ)

Three quadrupole devices coupled in a linear array. In the single reaction monitoring mode (SRM), selected parent ions pass through the first quadrupole and enter a second quadrupole device used as a collision cell to generate product ions. The third quadrupole serves as a selective mass filter for the determination of the resulting 'daughter ions'. It is a low resolution analyzer. It can operate in the multiple reaction monitoring (MRM) mode too, usually used for targeted approaches. It permits fast scan speed and polarity switching and high sensitivity but does not resolve MS spectra well.

Quadrupole-ion trap (QIT)

QIT is also a low resolution instrument that collects and stores ions by forcing them into stable orbits and subsequently releases them mass-selectively. Collected and stored parent ions can also be fragmented and thus daughter ions can be analyzed.

Quadrupole time-of-flight (Q-TOF)

Ions are simultaneously accelerated resulting in the same kinetic energy for any given ion. Along an evacuated flight tube with a fixed distance ions are separated by their mass-to-charge ratio and velocity, respectively. It is a high resolution MS instrument commonly used for untargeted metabolomics. Can provide high quality MS/MS spectra but scan speed and polarity switching is low.

Adapted from Weckwerth and Morgenthal [76].

Multidimensional separation techniques consisting in the use of sequential GC or LC columns for the separation are becoming more commonly used in metabolomics experiments since they provide high peak resolution and spectral purity [70].

In addition to *m*/*z* and retention time information, the identification of an ion is facilitated by the fragmentation pattern information that is acquired by tandem mass spectrometry (MS/MS) [71].

3.2.3. Data preprocessing

Data preprocessing is a very critical step in the overall workflow. There is no universal standard software or procedure to be applied in this step, and the repercussion of the procedure used impacts the final report generated. An enormous volume of data will be generated when following an untargeted approach. These data cannot be manually deconvoluted and the spectra generated must be corrected for all variations caused by experimental variables [70]. LC-MS data preprocessing pipelines include multiple stages such as filtering, feature detection, peak normalization (adjusting peak intensities and reducing analytical drift), peak alignment (retention time alignment across multiple samples), baseline correction to remove background noise, deconvolution of peaks to allow detection of trace co-eluted compounds and signal extraction (peak detection and quantification) [72,73]. Several data preprocessing tools are currently available including vendor software (i.e. MarkerLynx, MarkerView, SIEVE), software from independent developers (i.e. GeneData), open access software (i.e. XCMS, MetAlign) and script platforms (i.e. R, Matlab). The final reports comprise extracted *m/z* values, their retention times and the intensities corresponding to each peak. Before the multivariate analysis of the data analysis workflow, an initial data **normalization** step might be an important part of the procedure and consists of an intensity adjustment step needed to overcome day-to-day differences in instrument sensitivity (see Box 3). In LC-MS techniques the best way to perform this normalization is the use of internal standards (spiked into the sample during the extraction protocol) [63,64].

Box 3. Quality control

Several issues related to changes in instrument sensitivity are usually observed during a metabolomic or lipidomic profiling analysis including degradation of the extracts, contamination of an analytical column an ion source by non volatiles and retention-time shifts.

Following points are considered as a minimum standard for keeping high quality control in metabolomics and lipidomics when running a sequence:

- randomization of samples across the whole run;
- regular analysis of quality control (QC) samples;
- use of internal standards spiked into the samples during the extraction and/or added prior to LC−MS analysis;
- analysis of procedural or method blanks within each batch of actual samples to monitor intralaboratory contamination;
- regular checking of the quality of solvents used for mobile phases;
- carryover monitoring by means of LC−MS runs without actual injection of the sample;
- maintenance and replacement of columns and peripheral consumables;
- mass spectrometer maintenance including ion-source cleaning.

QC samples are typically pooled QCs obtained by mixing small aliquots of each (or a few) biological sample to be studied or commercially available QC samples with representative composition. With respect to the frequency of QC injections, 3−25 injections of samples between each QC injection are recommended. QC samples are very useful for measurement of within-series and between-series repeatability in order to remove metabolic features with excessive signal drift prior to statistical analysis.

Further, QC samples are usually used:

- to equilibrate an analytical platform after routine maintenance;
- to check the metabolite profile and signal intensity if they are within control limits before starting the sequence;
- to calculate technical precision within each analytical batch;
- to calculate signal correction for normalization of data within and between analytical series;
- for standardization if well-defined QC samples are included in the sequence.

Adapted from Cajka et al. [63]

3.2.4. Statistical analysis

The simple statistics used to discriminate control and affected groups is based on univariate statistics with determination of differences based on a t-test, a Wilcoxon rank test or an ANOVA. Univariate analyses are useful to identify individual biomarkers but insufficient in terms of performance. Hence, the need to build predictive models based on multiple biomarkers necessitates the use of **multivariate methods**. Multivariate methods are also very useful for data-dimension reduction and necessary for the high complex and voluminous data sets generated in MS untargeted approaches. Multivariate methods allow extracting the maximum information of the transformed data being able to process up to thousands of *m/z* and their corresponding peak intensities [66]. Multivariate statistical methods can be grouped in two main categories: supervised methods and unsupervised methods. Supervised multivariate methods need some initial information about the identity of the samples. The scope of these tests is to develop a model based on the information contained in the samples and are commonly employed in biomarker discovery studies. The discriminant analysis (DA) is the most common supervised method used in metabolomics. It includes PLS-DA and OPLS-DA. The unsupervised methods are based on the separation of classes without the need for initial information on the nature of the samples and the aim is to identify natural groupings among the samples. The most common unsupervised methods used are hierarchical clustering analysis (HCA) and principal component analysis (PCA). They provide a reduction of the MS data and reveal the inherent structure. HCA creates a clustering into different groups represented by dendrograms. PCA reduces data complexity into few key space components or principal components (PC) [74,75].

3.2.5. Ion selection and annotation

Once we have the list of features (specific *m/z* with a retention time associated of the mass spectral ion) that are interesting for a determined study, the next step is to search the accurate masses against metabolomic databases for putative identifications. Thus, success of untargeted metabolomics profiling is dependent on the availability of extensively annotated databases. Moreover,

these databases need to be updated frequently since human metabolome is not yet a well-defined entity and will continue to be populated. The challenge for metabolomics is not only to elucidate unknown chemical structures but also to put meta-information, such as sample origin, tissue and experimental conditions, into an accessible format [76]. In 2004, the Human Metabolome Project (HMP) was launched and as a result the Human Metabolome Database (HMDB) was developed. Some other available databases for metabolomics online are METLIN, NIST, GMD and MassBank. Furthermore, in order to be able to understand the meaning of changes in a set of metabolites it is essential to integrate metabolic data with systems biology and place the molecules in their context or metabolic pathways [77].

3.2.6. Metabolite verification

Tandem mass spectrometry experiments can be carried out on one metabolite with a specified *m/z*, followed by matching with an authentic standard, in order to obtain characteristic fragments and retention time information to distinguish the ion from structural isomers. If the fragmentation pattern of the isolated ion matches with the fragmentation pattern of the authentic standard (MS/MS) or *in silico* (MSe), then the hypothetical identification for that feature is verified. Nevertheless, some features cannot be assigned to an identified structural molecule because they are not yet annotated or known [59].

3.2.7. Interpretation of metabolomic data. Pathway analysis

Perhaps the largest challenge that metabolomic approaches face is relating the identified metabolites to their biological roles, which is a necessary step for moving beyond biomarkers and towards mechanisms [59]. The understanding of the flux of metabolites through a network of pathways is important in order to reveal alterations in enzymatic reactions, which in-turn affect regulatory processes and putative drug targets in metabolism. This is possible by using multiple metabolomic datasets originated from different experimental conditions and biological sources. Intrinsic biological fluctuation of independent samples results in altered patterns of observed correlations. The observation of correlated metabolites offers the chance to investigate whole metabolite network dynamics based on correlation network topologies. Changes in the

network topology point to regulatory hubs in the biochemical network because the correlation matrix of metabolite pairs is a fingerprint of the enzymatic and regulatory reaction network. Most pairs of metabolites neighbored in the reaction network show a low correlation, whereas other metabolite pairs that are far apart from each other in the reaction network exhibit a strong correlation. Primarily regulatory properties, especially the modulation of enzyme activity, serve as a source of changes in the topology of the correlation network. Any alteration in the reaction network will result in different correlation matrices. Because we are looking for differences in correlations, these phenomena can be analyzed using multivariate statistical analysis. Furthermore, the approach can be extended to system analysis looking for co-regulation of metabolites, proteins, transcripts or any other external or internal parameter measurable in the system to reveal a holistic picture of the metabolism [76].

There are several pathway-mapping tools that help placing metabolites into context with upstream genes and proteins. Some well-known network resources are KEGG, Recon1 and Biocyc. There are also new developed programs, such as mummichog and metabolite set enrichment analysis (MSEA), that find pathway connectivity. In addition, stable isotope metabolomics and omics-scale big data integration can reveal interconnectivity between metabolites and their relationship with genes and proteins [59].

3.2.8. Selection of potential biomarkers and validation

From the verification phase and the pathway integrated study a list of potential biomarkers will be obtained. A validation of the pathway can be performed at a molecular level (i.e. immunohistochemistry for the molecules or enzymes of the selected network) but a quantitative study of the metabolites can also be validated using MS platforms.

3.2.8.1. Multiple reaction monitoring (MRM)

Multiple reaction monitoring (MRM) during liquid chromatography coupled to a triple quadrupole mass spectrometer (QQQ) has been the standard workhorse in the quantitation of small molecules and metabolites, as it offers good sensitivity, reproducibility and broad dynamic range. It also has a fast scan

speed and fast positive to negative ion mode switching associated. In MRM the collision energy and product ion *m/z* are pre-optimized for each analyte of interest to give the best signal. The instrument presents three quadrupoles arranged in series. The first quadrupole (Q1) selects the ion of interest (parent or precursor ion); the second quadrupole (Q2) works as a collision cell to fragment the parent ion; and the third quadrupole (Q3) isolates the preselected product ion (Figure 21). The main disadvantage of targeted assays using MRM techniques is that they may yield low resolution and insufficient MS/MS spectra information. For these reasons, several strategies such as chemical standards with or without isotopic labeling can be used to improve identification confidence for the MRM mode [78–80].

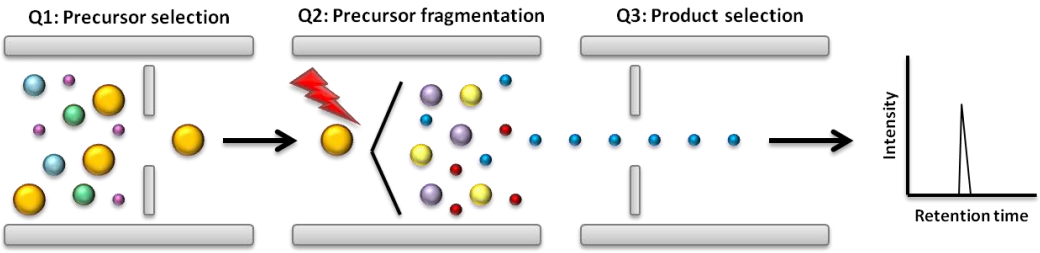


Figure 21. Diagram of a Multiple Reaction Monitoring (MRM) experiment.

3.2.8.2. Parallel reaction monitoring (PRM)

High resolution MS instruments (HRMS) such as Q-TOF or Orbitrap are gold standards for untargeted metabolomics profiling since they can simultaneously collect high resolution and high mass accuracy full-scans to analyze small molecules in complex samples. However, HRMS can also be implemented for targeted approaches. This strategy is known as parallel reaction monitoring (PRM). In this approach, a specific *m/z* or parent ion is selected in the Q1, then fragmented in the next cell (CID) and finally, the fragments or products are detected in parallel in a HRMS analyzer in a unique scan. This technique allows the identification of the metabolites by high mass accuracy MS/MS spectra and there is less dependence on chemical standards for the assay method construction [80].

3.3. ENDOMETRIAL CANCER AND METABOLOMICS

It is crucial to understand the pathways that can impact endometrial function in order to develop approaches for diagnosis, prognosis and treatment for EC. Extensive efforts (from the first histological studies [81] to the new integrated omics technologies [82]) have been applied to characterize and understand the biology of the receptive endometrium, but few publications analyzing the metabolic profile of the healthy or diseased endometrium have been performed.

There are a couple of studies published on metabolomics in healthy human endometrium and benign pathology. The first one, published by Vouk et al. in 2012 [83], is the first report identifying a panel of biomarkers from plasma samples that were able to diagnose endometriosis. They described elevated levels of sphingomyelins and ether-phospholipids in endometriosis compared to plasma from control subjects. The second study, published by Viella et al. in 2013 [84], is focused on the novel lipidome profiling of human endometrium, showing a significant increase of the lipid concentration in endometrial fluid at the embryo implantation moment (associated with the embryo receptivity capacity).

Few papers in the area of EC metabolomics have been published. The most initial papers in the field used targeted approaches. Knapp et al. in 2010 [85] outlined for the first time that the sphingolipid metabolism in human EC tissues is markedly augmented as compared to the normal endometrium. In their investigation, they analyzed the content of a set of some targeted metabolites (sphinganine, dihydroceramide, ceramide, sphingosine and sphingosine-1-phosphate (S1P)) using HPLC. They also analyzed the activity of some enzymes related to the sphingolipid pathway. Elevated levels of ceramide resulting from the activation of *de novo* synthesis pathway, accumulation of sphinganine and dihydroceramide, as well as upregulation of the serine palmitoyltransferase enzyme were detected in EC tissues compared to the healthy endometrium. Moreover, S1P concentration was also increased in pathological samples due to an overactivation of the enzyme sphingosine kinase (SPHK). As described in this investigation, these important changes in the sphingolipid metabolism observed in EC tissues likely contribute to its

progression and chemoresistance [85]. In the same year, Bufa et al. [86] also pointed out altered urinary profiles of endogenous steroids in postmenopausal women with EC. They used a quantitative GC-MS method in order to analyze the concentration of 23 androgen, progesterone and corticoid metabolites in urine samples collected from 13 EC patients and 10 control subjects. Also, Trousil et al. in 2014 [87] performed a targeted study in EC biopsies and confirmed alterations in the choline phospholipid metabolism. These alterations were caused due to an overactivation of the deacylation pathway and an increased expression of the enzyme choline kinase alpha.

In 2013, there was a study in which the metabolism of 2-Deoxy-D-glucose (2DG) in a human EC cell line was described in a targeted metabolomics profiling using TOF-MS. They characterized the role of 2DG as an inhibitor of tumor progression and described the impact of 2DG in EC cell lines. A group of preselected metabolites associated to 2DG action affected cell viability and may have therapeutic implications in EC [88]. Schuler et al [89] published a study related to the metabolic effects of a drug as anti-proliferative chemical for EC in 2015. They performed a global untargeted metabolomics analysis of serum pre- and post-metformin treatment and matched tumoral tissues of 20 obese women with EC. They concluded that metformin reduced tumor proliferation inhibiting the mTOR pathway. Due to the drug effect, they observed significant changes in lipid and glycogen metabolism. Their work underlined the potential benefit of metformin in cancer patients, including EC subjects.

Shao et al. [90] published another important study on the area of identifying biomarkers for EC using an untargeted metabolomics profiling approach in 2016. In this study, they analyzed the urine of a set of 35 control (25 healthy subjects and 10 patients with endometriosis) and 25 EC samples using a UPLC-Q-TOF-MS method in order to search for a group of potential diagnostic biomarkers for EC. Multivariate analysis was carried out to obtain a set of potential markers that were validated using the predictive model of the support vector machine (SVM). Shao et al. proposed a panel of 5 diagnostic biomarkers for EC, including 3 up-regulated (N-acetylserine, urocanic acid and isobutyrylglycine) and 2 down-regulated markers (porphobilinogen and acetylcysteine) that have an accuracy rate of 89.29% [90].

Finally, the most relevant and unique work in the metabolomics EC field aiming *de novo* identification of diagnostic biomarkers using human tissues was published in 2016 by Jové et al. [91]. Using a LC-Q-TOF-MS/MS platform they described a metabolomic signature that is specific for EEC compared to the corresponding controls. Moreover, they elucidated for the first time the involvement of the endocannabinoid system in EC pathogenesis. Additionally, they also depicted the implication of the purine pathway in tumor myometrial invasion.

In conclusion, the metabolomics and lipidomics research associated to EC is limited and needs expanded and concerted efforts. Since metabolite abundance patterns reflect events downstream of gene expression and it is considered to be closer to the actual phenotype than the other classic omics; it offers an important challenge in order to elucidate alterations in the physiological pathways of pathological conditions. Moreover, the integration of all the omics information (including genomics, epigenomics, transcriptomics, proteomics, glycomics, microbiomics and exposomics) in the research for EC and an improvement in the computational biology techniques are necessary for a better understanding of the disease that will result in the development of improved diagnostic and prognostic tools as well as individualized therapy options. In addition, several biomarker sources for the study of EC have not been analyzed yet (i.e. exosomes). However, since it is an emerging field and due to the complexity of datasets obtained, very few studies have been conducted focused on the metabolomic profiling of human cancer samples.

For all these reasons, any proposed metabolomics and lipidomics profiling studies with the goal of identifying, verifying and validating new specific and sensitive metabolites functioning as biomarkers for EC as well as the understanding of the metabolic pathways dysregulated in EC offer an appealing approach.

OBJECTIVES

EC represents the most frequent gynecological cancer in developed countries. Its diagnosis relies on the evaluation of an endometrial biopsy, which is a minimally-invasive method. Moreover, the management of patients is based on the assessment of the clinic-pathological features of the tumor, which is not completely accurate since some cases still recur unpredictably.

The main objectives of the thesis are focused on the identification of diagnostic markers and the characterization of the dissemination process of EC. We will achieve our goal by using metabolomic-based approaches in order to improve patient care and overcome the mortality rate of EC. Thus, this thesis has been structured in 3 specific objectives:

1. Optimization of the extraction methods for the metabolomic analysis of extracellular vesicles (EVs) contained in biofluids by using liquid chromatography mass spectrometry (LC-MS) techniques.

2. Identification, verification and validation of EC diagnostic markers from human biofluids. This strategy will be implemented following two specific sub-objectives related to each other:

2.1. Use of non-fractionated plasma and uterine aspirate samples from patients with EC and from control subjects.

2.2. Use of EVs isolated from plasma and uterine aspirates from the patients analyzed in Section 2.1.

3. Metabolomic profiling of EC tissues to elucidate the molecular alterations that take place in the initiation and progression of the disease.

3.1. Metabolomic study of different EC tissues including different F.I.G.O. stages in order to identify the main metabolic pathways altered in EC initiation and progression.

3.2. Molecular validation and functional characterization of the metabolic pathways altered in the carcinogenic process via *in vitro* studies.

RESULTS

The results generated during this thesis work are divided into 3 chapters, which corresponds each one to a paper. Chapters are preceded by a short summary. Chapter I and Chapter III are published papers and contain the original manuscript with the supplementary information and the references included at the end of the chapter. Chapter II is a manuscript in preparation and it also includes supplementary information and references attached at the end of the chapter.

CHAPTER I

Enabling Metabolomics Based Biomarker Discovery Studies Using Molecular Phenotyping of Exosome-Like Vesicles

The following published manuscript comprises the results presented in this chapter:

Enabling Metabolomics Based Biomarker Discovery Studies Using Molecular Phenotyping of Exosome-Like Vesicles

Altadill T, Campoy I, Lanau L, Gill K, Rigau M, Gil-Moreno A, Reventos J, Byers S, Colas E, Cheema AK

PLoS One

doi: 10.1371/journal.pone.0151339

SUMMARY

Background

Endometrial cancer (EC) is the most frequent gynecological cancer diagnosed in the United States, accounting for 76,000 deaths annually. Mortality is associated with presentation of poor prognostic factors and/or advanced disease at diagnosis. Exosome-like vesicles (ELVs) obtained from easy-to-access biofluids are an untapped resource for discovery and validation of clinically relevant biomarkers in health and disease. Although the proteomic and transcriptomic profile of ELVs have been broadly described, the metabolomic profile is uncharacterized.

Objective

The goal of this study was to describe methodologies for UPLC-ESI-MS based small molecule profiling of ELVs from human plasma and cell culture media.

Methods

Plasma samples from 35 EC patients were collected and processed in Vall Hebron Hospital (Barcelona, Spain) in accordance with approved institutional consent and reviewed protocols. In parallel, PANC-1 cell culture media containing ELVs was compiled. ELVs were isolated from plasma and cell media by standard ultracentrifugation, characterized by immunoblot against known exosomal markers and size and concentration estimated using Nanoparticle Tracking Analysis. We analyzed the metabolomic and lipidomic profile of the ELVs by ultra-performance liquid chromatography coupled with electro-spray quadrupole time of flight mass spectrometry (UPLC-ESI-QTOF-MS). Data were pre-processed using the XCMS software while the database search was performed using the Madison Metabolomics Consortium Database (MMCD), the Human Metabolome Database (HMDB), LIPID MAPS and Metlin for putative metabolite identification. A multivariate analysis using Metaboanalyst 3.0 web tool was carried out. A subset of metabolites was validated by tandem mass spectrometry.

Results

As a first step, we confirmed the isolation and enrichment of ELVs in plasma and cell culture media by NTA and by determining the expression of known exosomal markers, such as Flotilin 1, TSG101, CD63, CD9, Rab5 and CD81. Secondly, we optimized the starting volume of biofluid for ELVs isolation that would yield high quality spectral data. Thus, ELVs from different starting volumes ranging from 500 to 2000µL of plasma were isolated. A comparative analysis showed that 500µL of plasma as a starting volume was sufficient for ELVs isolation to produce high quality metabolomic and lipidomic data. The metabolome and lipidome of ELVs derived from cell culture media and plasma samples was characterized and some metabolites were validated. Finally, we present the utility of using MS based techniques in order to reveal the differences in ELVs metabolomic profiles of EMT-induced PANC1 cell line and EC samples compared to their corresponding controls.

Conclusions

This research indicates the potential use of ELVs as a biomarker source for the identification of metabolomic signatures that could be used for risk stratification and diagnosis of EC patients.

Novel aspects

When we scanned the literature, we found few studies which have performed a systematic optimization on the use of ELVs for metabolomics studies (including optimization of biofluid volumes and the characterization of the nature of the ELVs metabolome) which prompted us to take a lead. Thus, this is the first report that determines the plasma volume needed to obtain high-quality metabolomics and lipidomics data from ELVs and attempts to examine the metabolomics profile of plasma ELVs. Moreover, we present results detailing metabolite extraction methods and metabolic characteristics with wide applicability that would be useful for a broad audience engaged in biomarker discovery and validation for a plethora of patho-physiologies. We believe that the methodology presented here will provide an impetus to the growing field of metabolomics underscoring its ability to delineate biomarkers that can be used

for pre-clinical detection as well as for following the natural history of disease progression.

PLOS ONE

RESEARCH ARTICLE

Enabling Metabolomics Based Biomarker Discovery Studies Using Molecular Phenotyping of Exosome-Like Vesicles

Tatiana Altadill[1], Irene Campoy[1], Lucia Lanau[1], Kirandeep Gill[2], Marina Rigau[3], Antonio Gil-Moreno[4], Jaume Reventos[3], Stephen Byers[2], Eva Colas[1,5], Amrita K. Cheema[2]*

1 Biomedical Research Group in Ginecology, Hospital Universitari Vall d'Hebron, Institut de Recerca (VHIR), Universitat Autònoma de Barcelona, Barcelona, Spain, **2** Departments of Oncology and Biochemistry, Molecular and Cellular Biology, Lombardi Comprehensive Cancer Center at Georgetown University Medical Center, Washington, D.C., United States of America, **3** Institut d'Investigació Biomedica de Bellvitge (IDIBELL), Barcelona, Spain, **4** Gynecological Department, Vall Hebron University Hospital, Universitat Autònoma de Barcelona, Barcelona, Spain, **5** Department of Pathology and Molecular Genetics/Oncologic Pathology Group, Hospital Universitari Arnau de Vilanova, Universitat de Lleida, IRBLleida, Lleida, Spain

* akc27@georgetown.edu

OPEN ACCESS

Citation: Altadill T, Campoy I, Lanau L, Gill K, Rigau M, Gil-Moreno A, et al. (2016) Enabling Metabolomics Based Biomarker Discovery Studies Using Molecular Phenotyping of Exosome-Like Vesicles. PLoS ONE 11(3): e0151339. doi:10.1371/journal.pone.0151339

Editor: Petras Dzeja, Mayo Clinic, UNITED STATES

Received: November 24, 2015

Accepted: February 26, 2016

Published: March 14, 2016

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the Spanish Ministry of Health (RD12/0036/0035), the Spanish Ministry of Economy and Competitivy (PI14/02043), the AECC (Grupos Estables de Investi-gacion 2011 - AECC- GCB 110333 REVE), the Fundació La Marató TV3 (2/C/2013), the CIRIT Generalitat de Catalunya (2014 SGR 1330) and the European Commission, 7th Framework Programe, IRSES (PROTBIOFLUID –269285) – Belgium. The authors would like to acknowledge the Proteomics and Metabolomics Shared Resource partially supported by Cancer Center Support Grant NIH/NCI grant P30-CA051008.

Abstract

Identification of sensitive and specific biomarkers with clinical and translational utility will require smart experimental strategies that would augment expanding the breadth and depth of molecular measurements within the constraints of currently available technologies. Exosomes represent an information rich matrix to discern novel disease mechanisms that are thought to contribute to pathologies such as dementia and cancer. Although proteomics and transcriptomic studies have been reported using Exosomes-Like Vesicles (ELVs) from different sources, exosomal metabolome characterization and its modulation in health and disease remains to be elucidated. Here we describe methodologies for UPLC-ESI-MS based small molecule profiling of ELVs from human plasma and cell culture media. In this study, we present evidence that indeed ELVs carry a rich metabolome that could not only augment the discovery of low abundance biomarkers but may also help explain the molecular basis of disease progression. This approach could be easily translated to other studies seeking to develop predictive biomarkers that can subsequently be used with simplified targeted approaches.

Introduction

Most mammalian cell types secrete three types of extracellular vesicles either constitutively or in a regulated manner: exosomes, that are 35–150 nm diameter vesicles; ectosomes (also called microvesicles), from 100 to 1000 nm; and apoptotic bodies, from 500 to 2000 nm. Exosomes are formed from intraluminal vesicles and are delivered from multivesicular bodies to the outside of the cell by fusion with the extracellular membrane (endolysosomal vesicles) [1–3].

Competing Interests: The authors have declared that no competing interests exist.

Microvesicles and apoptotic bodies originate by budding and fission of the plasma membrane. Exosomes are found in different biofluids including plasma, urine, cerebrospinal fluid, and uterine aspirates and were first described in 1983 by Pan BT et al. and Harding C et al. [4–5]. Exosomes contain proteins, nucleic acids, lipids, RNAs and small RNAs and metabolites and it is thought that their principal function is to facilitate cell-to-cell communication under normal and diseased conditions [6, 7]. They are rich in cell surface molecules that facilitate their fusion with the receptor membrane and release their cargo in the cytoplasm [8] and are constantly released into circulation or proximal biofluids under normal and diseased conditions, affecting either proximal or distant cells. Since exosomes can be easily enriched from biofluids and provide a fingerprint of their cell of origin, there is a growing interest in using exosomes for the identification of novel and specific biomarkers with potential utility for diagnosis and prognosis of different cancer types [9].

Although there is a general consensus in the scientific community about the use of serial ultracentrifugation as the method of choice to isolate exosomes, there are still limitations to confirm the intra-luminal origin of the isolated vesicles (i.e. mainly lack of specific biomarkers). Thus, enriched vesicles having similar morphology and size and unknown biogenesis are defined as Exosome-Like Vesicles (ELVs).

Groundbreaking research over last decade has delineated biomarkers that can be used for early detection of cancer, dementia as well as those that can be used for monitoring response to therapy [10, 11]. However, identification of biomarkers with high sensitivity and specificity for a given disease type, still remains a major challenge in the field [12]. Biofluid molecular profiling based approaches have intrinsic limitations for detection of low abundant biomarkers which are obscured by the presence of high abundance molecules in the matrix. Exosomes, on the other hand, offer promise as an untapped biomarker resource; given that enrichment of the exosome fraction is likely to alleviate the dynamic range issue that is a common analytical problem across a broad range of biomarker identification and characterization studies. Moreover, the exosome cargo is protected from nuclease and protease activity by a lipid bilayer, resulting in increased stability of the sample [13].

Several studies have described biomarkers associated with cancer cell related ELVs [14–16]. However, most comparative exosomal profiling studies with a case-control study design have focused on transcriptomic and proteomic techniques. Given that the ELVs membranes have a rich lipid and metabolite content, characterizing ELVs metabolomes from different biofluids is likely to provide new information that could be used for identification sensitive and specific biomarkers that would also serve as a phenotypic readout since metabolites represent the end point of cellular processes. A recently published proteoglycan study of the serum exosomal fraction has shown the value of this matrix as novel biomarker source with potential clinical utility [17].

Metabolomics is an emerging "omics" field that enables the identification and quantitation of a wide variety of small molecules that are indicative of metabolic, nutritional and physiological status of the patient. This analytical tool allows for the analysis of a large number of samples in a high-throughput manner, and consequently, permits the understanding of current molecular response of a biological system to any perturbation in its microenvironment [18]. Furthermore, combining chromatography coupled to electrospray ionization mass spectrometry and multivariate statistical analysis enables the detection of differential abundance of metabolites between two conditions and adds value to clinical and translational studies focusing on cost-effective, high through put biomarker development. Metabolomic profiling of human biofluids using the enriched ELVs fraction represents a surrogate for tissue biopsy as a non-invasive tool for low abundance biomarker discovery and validation studies. However, there are few metabolic characterization studies reported in literature augmenting these investigations.

In this study, we have characterized ELVs metabolomes derived from human plasma samples which is a widely used matrix for biomarker studies. In addition, we also analyzed the metabolome of ELVs isolated from cell culture media since cell lines are widely used as a model system in biomedical research. Moreover, we report the differences on the ELVs metabolomics profile when comparing two conditions: endometrial cancer (EC) patient plasma samples versus control subjects; TGF-β treated pancreatic cancer human cell lines versus matched controls as two proof of principle studies. To our knowledge, this is the first report describing the presence and composition of metabolite cargo of ELVs derived from plasma and cell culture media using a high resolution mass spectrometry approach. This method development effort represents a first yet critical step for biomarker discovery using exosomal metabolomics with wide applicability. In addition to performing extensive characterization of metabolomic content of ELVs in plasma, we also present an approach that can be generically used for biomarker discovery and validation studies.

Materials and Methods

Patients

Participants in the study attended the Department of Gynecologic Oncology at the Hospital Vall Hebron in Barcelona, Spain. None of the patients included in the study received treatment prior to the collection of the biofluids. Final diagnosis was performed in the Department of Pathology at the same hospital. All patients participating in the study signed an informed consent. The Clinical Research ethics Committee at the Hospital Universitari Vall d'Hebron (Barcelona, Spain) approved the study. Collection of biofluids for the study included 10 mL of blood from postmenopausal cancer patients, with endometrioid adenocarcinoma, and control subjects. Samples were drawn from patients under sterile conditions, de-identified for research purposes and were processed, aliquoted and frozen at -80°C within four hours of collection. A full description of the clinic pathologic features of the patients included in both studies are detailed in S1 Table.

Plasma collection

Blood samples were collected in 10 mL EDTA-tubes (Cat# 367525. Pulmolab, CA, USA) and inverted 10 times at room temperature. Protease inhibitors (1:200, Cat# P8340. Sigma Aldrich, MO, USA.) were added to the sample. The sample was then centrifuged at 1,300 g during 15 min at room temperature. The supernatant was transferred to a clean tube and centrifuged at 3,000 g for 15 min at 4°C. The supernatant (plasma) was aliquoted and frozen at -80°C.

Cell culture

Human pancreatic carcinoma epithelial-like cell line PANC1 (ATCC® CRL-1469™) was obtained from the Tissue Culture Shared Resource of Georgetown University (Washington DC, USA). PANC1 cells were grown in DMEM media supplemented with 10% fetal bovine serum (HyClone, Logan, UT, USA) and penicillin/streptomycin. Reagents were obtained from HyClone (Logan, UT, USA). Cells were grown in a 5% CO2 incubator at 37°C.

Cells treatment: Cell lines were grown to 60% confluence, serum starved for 24 h and treated with 10 ng/mL of TGF-β (Cat # P01137, R&D Systems, MN, USA) for 48 h along with matched controls. Cells then were harvested and collected in 1.5 mL tubes. Samples were centrifuged at 13,000 rpm for 5 min, supernatant was removed, and cells were kept at -80°C.

Real-Time Reverse Transcription PCR

RNA was purified using the detailed protocol of RNeasy Mini Kit (Cat # 74104, Qiagen, CA, USA) and quantified by NanoDrop (NanoDrop 2000c, Thermo Scientific, NY, USA). Reverse-

Transcription Reaction was performed following the standard protocol of RT2 First Strand Kit (Cat # 330401, Qiagen, CA, USA). The cDNA obtained was analyzed by Real Time PCR (Cat # 4309155, SYBR Green, Applied Biosystems, CA, USA). Primers used for the Real Time PCR: GAPDH-F 3acggatttggtcgtattggg5, GAPDH-R 3tgattttggagggatctcgc5, E-cadherin-F 3tgcccagaaaatgaaaaagg5, E-cadherin-R 3gtgtatgtgg caatgcgttc5, N-cadherine-F 3ggacagttcctgagggatca5, N-cadherine-R 3ggattgccttccatgtctgt5, Vimentin-F 3ggctcagattcaggaacagc5, Vimentin-R 3gcttcaacggcaaagttctc5.

ELVs isolation

For ELVs isolation from cell culture, 30 mL of media was centrifuged (for three biological replicates), for 5 min at 300 g in order to remove the cell debris followed by another centrifugation at 2500 g for 20 min. ELVs isolation was performed for cell media and plasma using a modification of the protocol published by Thery et al [19]. Briefly, samples were centrifuged at 16,500 g at 4°C for 20 min to remove cell debris and fine particulates; subsequently the supernatant was placed in an ultracentrifuge tube (Cat#0314, Thermo Scientific, NY, USA) and centrifuged at 100,000 g at 4°C for 2 h. The pellet obtained was washed carefully with Tris Buffered Saline (TBS)-Ca2+ and centrifuged again at 100,000 g at 4°C for 1 h. Finally ELVs (pellet) were re-suspended in 60 μL PBS 1X buffer and frozen at -80°C. Although, exosome isolation can be performed using commercially available kits; many of these use Polyethylene Glycol (PEG) which interferes with downstream mass spectrometry based analysis.

Nanoparticle Tracking Analysis

The size and number of ELVs was determined using a Nanosight LM10 instrument with Nanoparticle Tracking Analysis (NTA, Malvern Instruments, UK). ELVs were diluted with Milli-Q water (Milli-Q Synthesis, Merck Millipore, Massachusetts, USA) and the analysis was performed following manufacturer's instructions. The videos were recorded for 1 min. Measurements were repeated at least 3 times per sample. The data were analyzed using the version 2.3 of the NTA-software. The size of the particles was calculated automatically based on the Brownian motion rate.

Western blot

Protein extraction of total plasma and plasma ELV fraction was performed by adding RIPA buffer, incubating at 4°C for 1 h followed by sonication. The supernatant containing exosomal proteins was collected after 15 min of centrifugation at 15,000 rpm. Protein concentration was determined using the standard protocol for the Bradford colorimetric method. Samples were loaded onto SDS-PAGE gels and transferred to a PVDF membrane. Membranes were blocked with 5% non-fat dried milk in TBS-Tween (0.01%) for 45 min and incubated overnight at 4°C with the primary antibodies, washed with TBS1x, incubated at room temperature with the secondary antibody and revealed.

Reagents. Primary antibodies: Flotillin-1 (1:250, Cat# 610821, BD Biosciences, San Jose, CA, USA), TSG101 (1:500, Cat#ab83. Abcam, MA, USA), CD63 (1:1000, Cat#OP171. Calbiochem, Merck Millipore, Massachusetts, USA), CD9 (1:250, Cat#555370. BD Biosciences, San Jose, CA, USA), Rab5 (1:2000, Cat#ab13253, Abcam, MA, USA), CD81 (1:1000, Cat#sc-166028. Santa Cruz, Texas, USA) and Haptoglobin (1:1000, Cat#ab131236, Abcam, MA, USA). Secondary antibodies: Goat anti-rabbit, Cat# P0448; rabbit anti-mouse, Cat# P0260, both 1:2000. Dako, CA, USA). RIPA buffer (5nM EDTA, 150mM NaCl, 1% Triton, 20nM Tris pH8 and

1:200 protein inhibitors). Immobilon Western Chemiluminiscent (Cat#WBKLS0100. Merck Millipore, Massachusetts, USA).

Metabolomic analyses of ELVs and database search

ELVs isolated from plasma and cell culture media were processed for mass spectrometric analysis as described by Sheikh *et al* with minor modifications to the original protocol [20]. Initially, samples were thawed on ice and vortexed. For metabolite extraction, 55 μL (v/v) of ELVs containing approximately 10^{11-12}particles/mL were mixed with 95 μL of water. The tubes were placed on dry ice for 30 sec followed by 90 sec incubation in a 37°C water bath. The samples were sonicated for 30 sec. A total of 600 μL of chilled methanol containing internal standards (5 μL of debrisoquine at a concentration of 1mg/mL and 25 μL of 4-nitrobenzoic acid at a concentration of 1mg/mL) were added to each sample. The coefficient of variation (% CV) of each internal standard for each matrix and condition was calculated (S2 Table), which ranged between 5.5 to 13.83%. Following vortexing, samples were incubated on ice for 15 min and 600 μL of chloroform were added to each tube, followed by centrifugation at 13,000 rpm at 4°C for 10 min. The supernatant was transferred to a fresh tube and 600 μL of chilled ACN were added. Samples were incubated at -20°C overnight and centrifuged at 13,000 rpm at 4°C for 10 min. The supernatant was transferred to a fresh tube and dried under vacuum. This sequential extraction strategy [going from water (aqueous) to methanol (semi-polar) to non-polar solvent extraction (chloroform)] allows for a broad range extraction of metabolites. Dried samples were resuspended in 200 μL of 50% water and 50% methanol followed by UPLC-ESI-Q-TOF-MS analysis. The solvent composition supports the dissolution and MS analysis of a broad class of compounds. A total of 5 μL of each sample were injected onto a reverse phase Acquity UPLC CSH C18 1.7 μm, 2.1 x 100 mm column (Waters Corp.). Two different gradients were used based on the sample matrix (plasma and cell culture media). MS data acquisition was performed using ESI-QTOF MS within the mass range of 50 to 1200 mass-to-charge ratio (m/z) in positive and negative electrospray ionization modes on a (Xevo G2 Q-TOF; Waters Corp.). MS data were pre-processed using the XCMS software [21]. The Madison Metabolomics Consortium Database (MMCD) [22], the Human Metabolome Database (HMDB) [23], LIPID MAPS [24] and Metlin [25] were used for accurate mass based putative identification of metabolites. We performed a multivariate data analysis using Metaboanalyst 3.0 web tool [26]. Identities for a sub-set of metabolites were validated by tandem mass spectrometry.

Reagents for MS analysis: Chloroform, LC/MS-grade acetonitrile (ACN), water and methanol were purchased from Fisher Optima grade, Fisher Scientific (New Jersey, USA). PBS was purchased from Invitrogen (Carlsbad, CA, USA). High purity formic acid (99%) was purchased from Thermo-Scientific (Rockford, IL, USA). Ammonium formate, debrisoquine and 4-Nitrobenzoic acid (4-NBA) were purchased from Sigma- Aldrich (St. Louis, MO, USA).

Results

Characterization of plasma ELVs

Procuring human plasma samples from biorepositories for biomarker discovery efforts is often constrained by sample volumes and associated costs of the study; hence our initial goal was to optimize minimum volume of plasma for ELVs isolation suitable for downstream metabolomic analysis using high resolution mass spectrometry. For this purpose, we isolated ELVs from different volumes of human plasma (500, 1000 and 2000 μL) as described in methods section [19]. For all samples, we determined size distribution and concentration of the ELVs enriched fraction by NTA, protein concentration and total protein amount. The average size of plasma ELVs was approximately 150 nm (Fig 1A). The final concentration of ELVs correlated with the

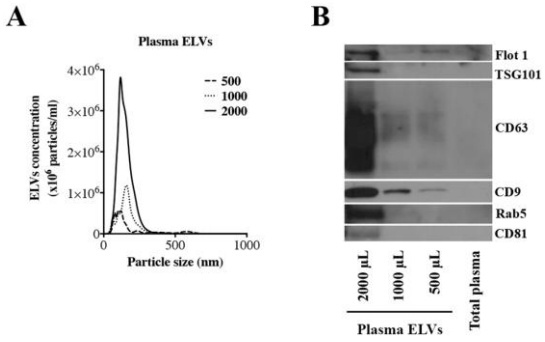


Fig 1. Nanoparticle Tracking Analysis (NTA). Determination of the concentration and particle size of human from 500, 1000 and 2000 μL of human plasma **(Panel A).** Size distribution of the nanovesicles was around 150 nm. **Western blot analysis to confirm the expression of exosome specific markers (Panel B).** Human plasma ELVs enrichment confirmed in exosomes isolated from different volumes of human plasma (2000, 1000 and 500 μL) and no presence of exosomal specific markers when loading total plasma (non-enriched). Lane A through C represent ELVs isolated from 2000, 1000 and 500 μL of plasma respectively; Lane D represents non-enriched total plasma.

doi:10.1371/journal.pone.0151339.g001

starting volume, ranging from $1x10^{12}$ particles per mL for 500 μL samples to $7.5x10^{12}$ particles per mL for 2000 μL of plasma. These results are consistent with previous studies that have reported similar size and number of particles per mL of plasma [27, 28]. Equal volumes of protein extract corresponding to ELVs purified from 2000, 1000 and 500 μL of plasma were subjected to SDS-PAGE. The gel was also loaded with protein extracted directly from total plasma (Fig 1B) in order to show the enrichment of exosome specific protein markers in the purified samples. For the total plasma we loaded the same amount of protein that we loaded for the protein extract corresponding to ELVs purified from the higher volume (2000 μL). The non-enriched plasma sample does not show the presence of ELVs specific markers while the enriched fractions exhibit the presence of these markers in a concentration dependent manner which in turn was dependent on the initial volume of plasma that was used for enrichment of the ELVs fraction. These include Flotillin 1, TSG101, CD63, CD9, Rab5 and CD81 (Fig 1B). However, as seen in the figure, the concentration of ELVs in non-purified plasma samples is too low for detection of these markers by immunoblotting. Uncropped western blots are shown in S1A and S1B Fig. In order to understand the amount of co-pelleted artifacts at each step of the isolation method, we performed an immunoblot against haptoglobin (S1C Fig), a soluble protein found in plasma, with MVs and ELVs isolated from 2000, 1000 and 500 μL of plasma. As we expected, we detected the presence of haptoglobin in MVs, since they are collected after the first round of centrifugation (16,500 g for 20 min) where we get rid of the large vesicles and molecules contained in the sample. However, the following steps of ultracentrifugation permitted to rule out most of the contaminants. We could also detect the presence of haptoglobin in ELVs isolated from large volumes of plasma (2000 μL), but not from 1000 and 500 μL. These results strongly suggest the enrichment of ELVs from plasma that were ready for metabolomics characterization using high resolution mass spectrometry.

The Total Ion Chromatograms (TICs) observed in the negative ionization mode for the metabolites extracted from the ELVs of 500, 1000 and 2000 μL of plasma are shown in Fig 2. The ratio of relative intensity values for some features detected are shown in S3 Table. Although with 1 or 2 mL of plasma, a small fraction of metabolites showed improved ion intensity, a large number of detected metabolites did not show any changes in relative ratios or improvement

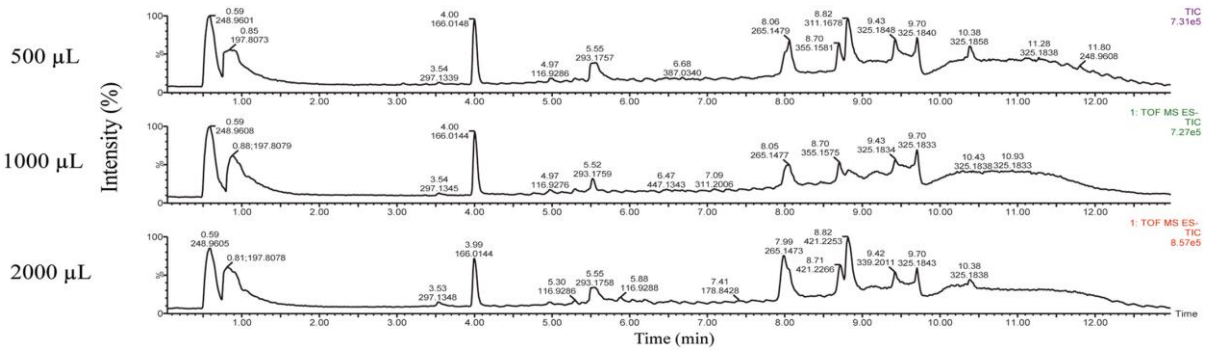


Fig 2. Total Ion Chromatograms (TIC) of human plasma Exosome-Like Vesicles (EVLs) in the negative electrospray ionization mode. ELVs were isolated from different volumes of human plasma (500, 1000 and 2000 µL) and subjected to TOF-MS analysis. The X-axis represents the chromatographic retention time while the Y-axis represents the relative intensity.

doi:10.1371/journal.pone.0151339.g002

signal to noise ratio for individual metabolites suggesting that 500 µL of plasma was optimal for performing metabolomics profiling studies using matrices of human plasma samples.

Hence further metabolomics characterization was performed with ELVs isolated from 500 µL of plasma samples that were analyzed using quadrupole time of flight mass spectrometry using three biological replicates. The total number of features detected in plasma derived ELVs samples was 840 and 2198 in the negative and positive ionization mode, respectively. Next, we annotated the features by performing an accurate mass based database search and an enrichment analysis using Metlin and Human Metabolome Databases (S4A and S5 Tables). Our results showed that the metabolome of ELVs derived from plasma maintained a high prevalence of glycerophospholipids (possibly representing exosomal membrane contents) and shingolipids, representing 29% of the total metabolome composition, such as PI (16:0/22:4), PE (22:2/16:1), GalCer (d18:2/16:0), GPCho (18:0/14:0) or TG (12:0/12:0/20:5). The fatty acid esters, amides and alcohols (oleamide, malonyl-CoA among others) represented 16% of the total metabolites detected in this analytical run while the nucleotides, nucleosides and their derivatives comprised 17% of the metabolome. Peptides, organic acids and derivatives were less abundant in the total ELVs metabolite content (Fig 3A). Some of the interesting

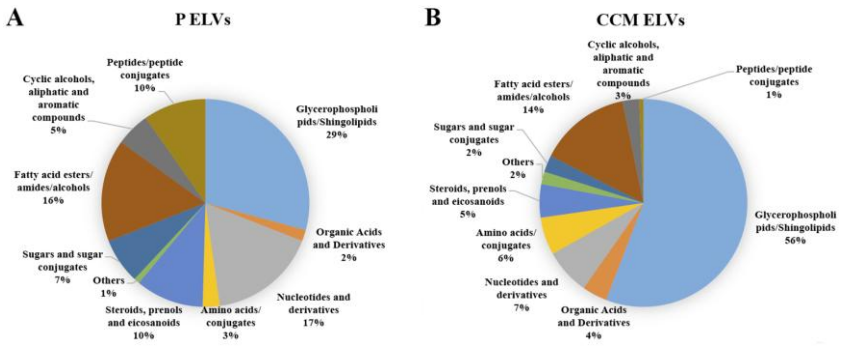


Fig 3. Metabolomic ELVs composition. Characterization of the metabolite content of human plasma Exosome-Like Vesicles (P ELVs) **(Panel A)** or PANC 1 cell culture media (CCM ELVs) **(Panel B)**.

doi:10.1371/journal.pone.0151339.g003

Table 1. Validated metabolites.

Metabolite	Metabolite ID	m/z	Mass error (ppm)	Retention time (min)	Matrix	Mode	Major CID fragments
PG(16:0/16:0)	123064911 (Pubchem SID)	707.520	0.002	10.129	P ELVs	Pos.	663.4540, 607.3931, 551.3286, 495.2679, 439.2032
N-Arachidonoyl-L-Serine	74380323 (Pubchem SID)	392.280	0.001	9.252	P ELVs	Pos.	149.0277, 123.0987, 106.0971, 95.0809, 69.0711, 57.0866
SM(d18:1/16:0)	7850646 (Pubchem SID)	703.574	0.001	9.993	P ELVs	Pos.	184.074
Oleamide	C19670	282.279	0.000	8.020	P and CCM ELVs	Pos.	97.1005, 83.0866
Coenzyme Q10	C11378	861.677	0.000	10.301	P ELVs	Neg.	724.8752, 588.8987, 520.9148, 452.9229, 384.9373, 316.9473, 248.9610, 180.9727, 112.9871
Malonyl-CoA	cq_00054	853.593	0.001	10.026	P ELVs	Neg.	808.727
All-trans-4-oxoretinoic acid	HMDB06285	315.195	0.001	4.327	CCM ELVs	Pos.	203.1072, 187.1123
Psychosine sulfate	5280538 (Pubchem SID)	542.301	0.001	4.049	CCM ELVs	Pos.	427.2240, 358.1841, 255.1232, 249.1212, 155.1072, 123.0446

List of metabolites contained in plasma Exosome-Like Vesicles (P ELVS) and PANC 1 cell culture media Exosome-Like Vesicles (CCM ELVs) confirmed by MS/MS.

doi:10.1371/journal.pone.0151339.t001

metabolites identified in plasma ELVs included coenzyme Q10, ubiquinone 9, palmitoyl glucuronide, PG (16:0/16:0), 25-hydroxy-hexadehydrovitaminD3, 10-formyldihydrofolate, acetyl glucosamine bisphosphate, malonyl-CoA, cytidine-5'-monophosphate, N-Arachidonoyl-L-Serine, tyrosyl-AMP, oleamide, deoxyuridine-diphosphate, picolinic acid and deoxyvitamin D3 and SM (d18:1/16:0) indicating a wide class of metabolites. A subset of these metabolites was validated using tandem mass spectrometry (Table 1). Some of these metabolites fall below the detection limits in routine untargeted metabolomics profiling of plasma samples, possibly because of ion suppression in presence of other high abundant metabolites that get eluted in the same chromatographic time scale.

Characterization of ELVs from cell culture media

In order to characterize ELVs in an *in vitro* model system, we isolated ELVs from cell culture media as described in the methods section. The ELVs were characterized using NTA, western blotting and finally subjected to high resolution mass spectrometry based analysis as described for plasma ELVs.

We analyzed the bio-composition of ELVs that were isolated from cell culture media obtained by growing PANC1, a human pancreatic cell line as described in the methods section. We used three biological replicates for these analyses. The ELVs for each sample were isolated from 30 mL of media. Quantitative analysis of the exosomes isolated using NTA revealed an average size of 95 nm and average concentration of 1.6 $x10^{10}$ particles per mL consistent with the size of ELVs previously reported. The Total Ion Chromatogram (TIC) observed in the positive ionization mode for PANC1 is shown in Fig 4. Data were pre-processed using XCMS software that yielded 596 and 2926 features in the negative and positive electrospray modes respectively.

Database search of ELVs metabolome obtained from cell culture media was performed (S4B and S5 Tables) and dominated by a majority of glycerophospholipids and shingolipids (56%)

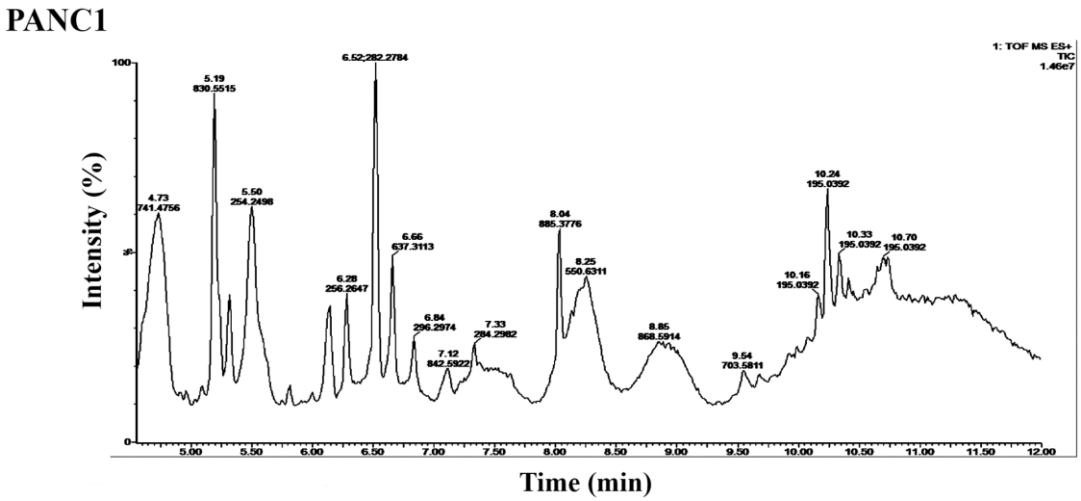


Fig 4. Total Ion Chromatograms (TIC) of PANC1 pancreatic cancer cell culture media Exosome-Like Vesicles (ELVs) in the positive electrospray ionization mode. ELVs were isolated from pancreatic cell line PANC 1 media and subjected to TOF-MS analysis. The X-axis represents the chromatographic retention time while the Y-axis represents the relative intensity.

doi:10.1371/journal.pone.0151339.g004

and a significant representation of fatty acid esters, amides and alcohols (14%) and nucleotides and derivatives (7%). We also found a representation of sugars, cyclic alcohols, aromatic compounds, steroids and organic acids which are generally suppressed when whole cell extracts are used for molecular profiling (Fig 3B). In addition we were able to detect Adenine, aminoadipic acid, enol-Phenylpyruvate, oleamide, 15-HETrE, 3-Dehydrosphinganine, all-trans-4-oxoretinoic acid, N,N-dimethylsphingosine, docosanamide, psychosine sulfate, pentaglutamyl folate and fructose 1,6-bisphosphate. Some of these metabolites were unambiguously identified by comparing fragmentation patterns of metabolites acquired in the elevated collision energy mode (MS^E) using mass fragment calculator (Table 1). These results emphasize the presence of a broad class of metabolites with potential to be used as biomarkers.

ELVs metabolome as a biomarker source

Having characterized the metabolome of ELVs isolated from plasma and cell culture media, our next goal was to test the possibility of using comparative ELV metabolomic/lipidomic profiling to distinguish between two or more groups. We used two proof of principle studies; the first study was designed to delineate differences in the ELV profiles enriched from PANC1 cells treated with transforming growth factor beta (TGF-β) to those that were untreated while in the other we compared the metabolomic profiles of ELVs isolated from plasma of EC patients to those obtained from healthy volunteers.

Comparative Profiling of TGF-β treated PANC1 cells. It has been demonstrated that epithelial to mesenchymal transition (EMT) contributes to metastasis and treatment resistance of pancreatic cancer cells [29]. It is also well known that TGF-β acts as a driver in cancer progression through induction of EMT [30]. Therefore, we treated PANC1 cells with TGF-β, while the control cells were treated with DMSO and interrogated the metabolomic alterations in ELVs isolated from spent media under these two experimental conditions. We used three

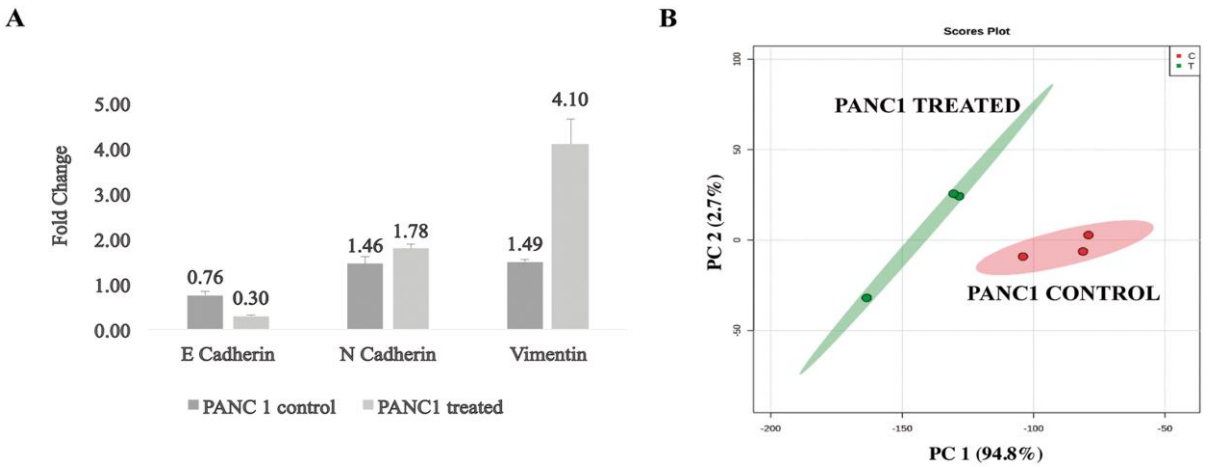


Fig 5. Panel A. TGF-beta induces epithelial to mesenchymal transition (EMT) in PANC1 cells. Real Time PCR analysis showed a decrease of the epithelial marker E-cadherin (p-v<0.05) and an increase of the mesenchymal markers N-Cadherin and Vimentin (p-v<0.05) expression after TGF-β treatment in PANC1 cells. **Panel B. Principal Component Analysis (PCA) of exosomal metabolome for MS negative ionization mode showing the separation between the two study groups**. TGF-β treated (T) and control (C) PANC1 cells. We analyzed three replicates per condition. The x-axis shows interclass separation while y-axis illustrates the intra-class variability on Y-axis.

doi:10.1371/journal.pone.0151339.g005

biological replicates for each condition (treated and untreated). The cells and media were collected after treatment (see methods). The cells were tested for EMT induction by TGF-β while the media were used for ELVs enrichment for subsequent metabolomic analyses. As expected, at RNA level we observed downregulation of the epithelial marker E-cadherin and the overexpression of the mesenchymal marker Vimentin in TGF-β treated cells compared to the controls (Fig 5A). Next we isolated ELVs from cell culture media and using an untargeted metabolomic approach we were able to elucidate a variation of the ELVs metabolome in response to TBF-β treatment. The principal component analysis (PCA) obtained shows a clear separation of the metabolomic composition of the ELVs due to the treatment (Fig 5B). Metabolites showing significant differences (fold change (FC) ≥ 2 and p-value ≤ 0.05) are listed in S6 Table. UDP-D-glucosamine (m/z: 550,084) and DG (20:2)/18:1/0:0) (m/z: 647,544) were two of the metabolites identified to be significantly up-regulated after TGF-β treatment.

Metabolomics of plasma ELVs reveals differences between controls and EC patients. We hypothesized that a compendium of molecular profiles (metabolomics) would provide insight into alterations that underscore the tumor phenotype in EC patients. Therefore, as a proof of principle, we compared metabolomic profiles of ELVs isolated from plasma of 13 control subjects (healthy patients) and 19 EC patients that were analyzed by UPLC-ESI-Q--TOF-MS. Multivariate analysis was performed in order to elucidate the differences in the ELVs metabolome composition in the two study groups. The differential abundance of metabolites was visualized as a heat map (Fig 6A) and PCA (Fig 6B) which indicated a differential pattern of the metabolites found in the ELVs isolated from EC plasma patients compared to the controls. Several significant metabolites did not yield an accurate mass based putative identification when searched against several databases which remains a major challenge in the field. The differentially abundant metabolites with putative IDs included substituted sugars and amino acids. Further characterization and validation of these findings is ongoing in our laboratory.

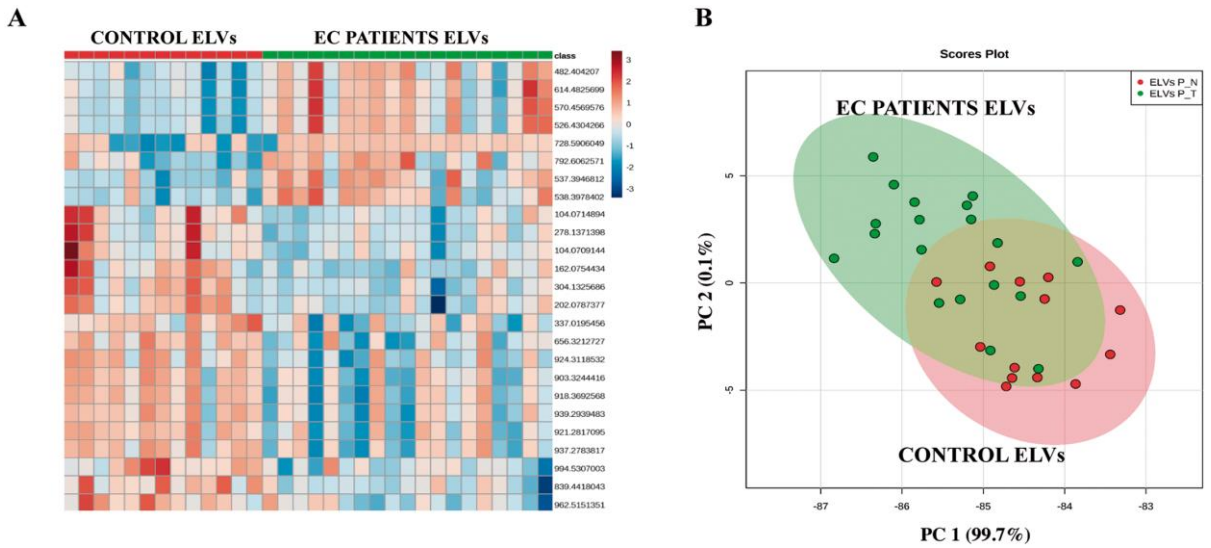


Fig 6. Multivariate analysis reveals distinct metabolic changes in plasma derived Exosome-Like Vesicles (ELVs) isolated from endometrial cancer (EC) patients compared to the control subjects. Panel A. Heat map visualization of ion rankings of volcano plot based m/z, corresponding to their relative levels (intensity) in plasma ELVs isolated from EC patients and control subjects for MS positive ionization mode. Each row on the heat map represents a unique feature with a characteristic mass to charge ratio and retention time while each column represents one subject. **Panel B**. Principal Component Analysis (PCA) plot for MS positive ionization mode showing the separation between EC and control plasma ELVs. The x-axis shows interclass separation while y-axis illustrates the intra-class variability.

doi:10.1371/journal.pone.0151339.g006

Discussion

In this study, we present a broadly applicable approach for metabolomic profiling of ELVs isolated from human plasma as well as cell culture media. We chose these two matrices for optimizing metabolomics methodologies since these are widely used in studies with clinical and basic science focus respectively. Furthermore, the use of ELVs as a biomarker resource is useful since they can be isolated from most bodily fluids and their rich content is protected from degradation, allowing for circulation as well as for cell-to-cell communication of molecules and metabolites that would not be stable in free plasma and other biofluids [31]. It has also been reported that the ELVs secretion rate derived from tumoral cells is much higher than the rate for healthy cells; being involved in signal transmission in tumoral microenvironment [32].

We started out by confirming isolation and enrichment of ELVs fraction from human plasma and cell culture media by immunoblotting. Next, we characterized their size and concentration by Nanoparticle Tracking Analysis. A set of common and well established ELVs markers were identified in the extractions enriched in vesicles around 95–150 nm confirming the specific enrichment of ELVs. We also determined that 500 μL volume of plasma is optimal for generating high quality MS data for profiling the metabolomic and lipidomic content of the ELVs; increasing plasma volumes further does not add significant value to the number and signal to noise ratio of the detected metabolites.

A sub-set of the metabolites identified in plasma ELVs were validated including PG (16:0/16:0), N-arachidonoyl-L-serine, SM (d18:1/16:0), coenzyme Q10 and malonyl CoA. PG (16:0/16:0) is a glycerophospholipd precursor of cardiolipin. Cardiolipin is found in the inner mitochondrial membrane and a change in its concentration and distribution in this organelle is known to cause several diseases including cancer and aging [33]. It has also been reported that

the endocannabinoid N-arachidonoyl-L-serine can have a neuroprotective effect in maintaining the undifferentiated state of some cells [34] and can promote cell proliferation, migration and angiogenesis [35]. Moreover, recent studies have revealed that Coenzyme Q10, apart from its antioxidant functions, also regulates the expression of genes involved in cell signaling and transport. The incorporation of Coenzyme Q10 in experimental lipovesicles enhanced their cell uptake [36] and an increase of Coenzyme Q10 concentration in vesicles prolonged their circulation in blood [37].

Furthermore, we analyzed the biochemical composition of cell culture media ELVs and we were able to identify some metabolites such aminoadipic acid that is an intermediate metabolite in the lysine pathway. It acts also as an antagonist of the neuroexcitatory activity and inhibits the production of kynurenic acid in some tissues [38]. Enol-phenylpyruvate was also present in cell culture ELVs and is related to phenylalanine and tyrosine metabolism pathways as well as to macrophage induced inflammatory processes [39]. We found also 15-HETrE, which is a polyunsaturated fatty acid involved in tumorigenesis and in the modulation of arachidonic acid metabolism. It also regulates the activity of cyclooxygenase-2 by inhibiting its expression [40]. Finally, all-trans-4-oxoretinoic acid was also present in the ELVs. It is an isomer of retinoic acid that is actively involved in cell-to-cell communication by modulating gap junctional activity [41]. There was an overlap in the different classes of metabolites obtained from plasma and cell culture media ELVs although the distribution across different classes was matrix dependent (Fig 3).

As expected, the major class of metabolites were glycerophospholipids, which are integral components of the exosomal membranes and could majorly contribute to high phospholipid content associated with human plasma samples. Interestingly, we also found metabolites that belong to different classes including organic acids (several glycolytic intermediates), cyclic alcohols, steroids, prenols and amino acid conjugates as well as sugar and sugar conjugates (S4A and S4B Table) both in plasma and cell culture media ELVs. In our experience with routine plasma profiling experiments many of these metabolites fall below the instrumental limit of detection. Hence we compared the number of features obtained from regular plasma profiling with those obtained from ELVs fraction enriched from 500 μl of plasma (Fig 7). There was little overlap of detected metabolites between the two matrices most probably because of sensitivity range of the instrument for detecting metabolites from 25 μl of human plasma that we routinely use for untargeted metabolomics profiling [42]. This underscores the importance of enriching this fraction for low abundance biomarker discovery using plasma samples. Thus, detection of these compounds with significant implications in cell-to-cell communication and signaling, regulation and metabolic status is of critical importance for novel, low abundance biomarker discovery. We further propose that these compounds can be further characterized from whole plasma extracts using multiple reaction monitoring based targeted mass spectrometry thus dramatically increasing the sensitivity and specificity of the assay as well as the flexibility of multiplexing such that multiple metabolites could be assayed in a single injections. This combinatorial approach (Fig 8) can also be used for testing clinical utility of biomarkers in clinical samples as well as to compare different conditions when performing *in vitro* or *in vivo* experiments. Furthermore, we present the utility of using MS in order to reveal the differences in ELV metabolomic profiles of EMT-induced PANC1cell line (as compared to control) and EC samples (as compared to controls). We further propose that the low abundance biomarkers discovered using ELVs from bodily fluids could then be analyzed using targeted mass spectrometry from plasma directly since the sensitivity would be greatly enhanced. The approach described here can be generically used for biomarker identification of cancer or other pathologies thus furthering the precision medicine paradigm. Thus findings from our study have great value given that not only do these present a standardized approach for mass

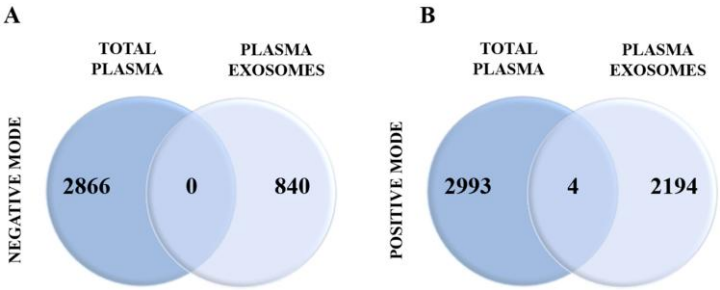


Fig 7. Comparison of the number of features detected in total plasma (non-purified) or plasma ELVs. The number of common and unique features for each type of sample is represented for the negative (Panel A) and positive (Panel B) electrospray ionization mode.

doi:10.1371/journal.pone.0151339.g007

spectrometry based metabolomics profiling, but also provide evidence for using ELVs metabolomics (derived from human fluids as well as from cell culture media) as a rich matrix for biomarker discovery for studies with basic science, clinical or translational focus.

Conclusions

Biomarker detection and characterization from complex biological matrices, especially for markers that are low in abundance, remains a challenge. There is increasing evidence that underscores the importance of ELVs as carriers of important cellular information that can be

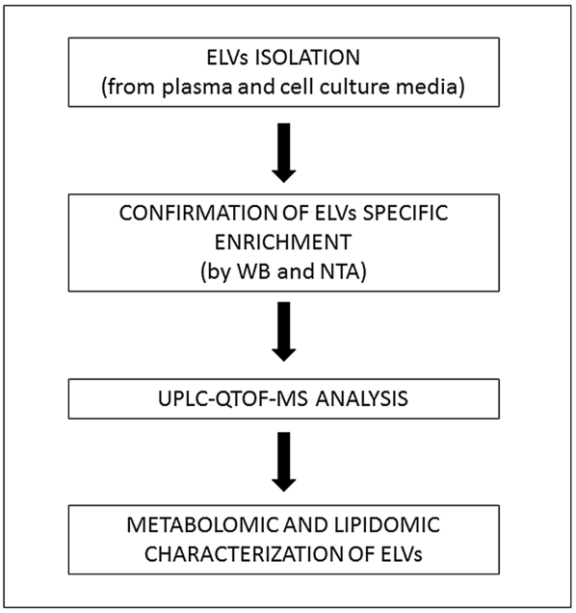


Fig 8. Workflow. Schematic showing experimental design in order to analyze the composition of exosome-like vesicles (ELVs) derived from plasma samples and cell culture media. (WB = western blot, NTA = nanoparticle tracking analysis).

doi:10.1371/journal.pone.0151339.g008

used for defining specific changes in health and disease. Herein, we present an experimental pipeline that can be broadly applied for identification of low abundance biomarkers and characterization in clinical samples. Finally, we demonstrate that ELVs represent an untapped source for metabolic and lipidomic biomarker discovery with high clinical and translational relevance. We believe that the methodology presented here will provide an impetus to the growing field of metabolomics underscoring its ability to delineate biomarkers that can be used for pre-clinical detection as well as for following the natural history of disease progression.

Supporting Information

S1 Fig. Analysis of protein expression by western blot of Exosome-Like Vesicles (ELVs) and microvesicles (MVs) isolated from human plasma (2000, 1000 and 500 μL). Expression was also analyzed in non-purified total plasma (T Plasma). Markers CD81, TSG101, Rab5 and Flotillin 1 were blotted in the same membrane (Panel A) and CD63 and CD9 were blotted in an independent membrane (Panel B). Expression of the soluble protein haptoglobin was analyzed in MVs and ELVs isolated from different volumes of T Plasma.
(TIF)

S1 Table. Patient clinic-pathological information.
(XLSX)

S2 Table. Coefficient of variation of Internal Standards in each group of samples.
(XLSX)

S3 Table. Comparative analyses of putative features identified in human plasma ELVs. Ratios of relative intensities for a selection of metabolites is represented.
(XLSX)

S4 Table. Metabolome profiling of ELVs isolated from human plasma samples and cell culture media. Table A. Putative identifications in plasma ELVs. **Table B**. Putative identifications in PANC 1 cell culture media ELVs. The metabolite ID, mass to charge ratio (m/z), mass error (ppm), retention time (min) and electrospray ionization mode are detailed.
(XLSX)

S5 Table. Unassigned metabolites IDs found in plasma and cell culture media ELVs.
(XLSX)

S6 Table. Significant m/z detected in ELVs isolated from PANC1 cell culture media. Mass to charge ratio (m/z), fold change (FC), p-value (p-v) and electrospray ionization mode are detailed.
(XLSX)

Acknowledgments

This work was supported by the Spanish Ministry of Health (RD12/0036/0035), the Spanish Ministry of Economy and Competivity (PI14/02043), the AECC (Grupos Estables de Investigación 2011—AECC- GCB 110333 REVE), the Fundació La Marató TV3 (2/C/2013), the CIRIT Generalitat de Catalunya (2014 SGR 1330) and the European Commission, 7th Framework Programe, IRSES (PROTBIOFLUID –269285)–Belgium. The authors would like to acknowledge the Proteomics and Metabolomics Shared Resource partially supported by Cancer Center Support Grant NIH/NCI grant P30-CA051008. We would like to thank Mr. Tyrone Dowdy and Ms Kirandeep Gill for their valuable help developing the metabolomics data. We also acknowledge technical assistance provided by Irene Campoy and Lucia Lanau, and all

clinicians that have participated in the recruitment of clinical samples. We thank volunteers for their willingness to participate in the study.

Author Contributions

Conceived and designed the experiments: AKC EC MR TA. Performed the experiments: TA IC. Analyzed the data: TA KG. Wrote the paper: AKC EC TA JR SB. Were responsible for clinical sample collection and processing: IC LL AGM.

References

1. Pfeifer Philipp, Werner Nikos, Jansen Felix. Role and Function of MicroRNAs in Extracellular Vesicles in Cardiovascular Biology. Biomed Research International. 2015; 2015: 161393. doi: 10.1155/2015/161393 PMID: 26558258
2. Pocsfalvi G, Stanly C, Vilasi A, Fiume I, Capasso G, Turiák L, et al. Mass spectrometry of extracellular vesicles. Mass Spectrom Reviews. 2016; 35: 3–21.
3. Pan BT, Johnstone RM. Fate of the transferrin receptor during maturation of sheep reticulocytes in vitro: selective externalization of the receptor. Cell. 1983; 33 (3): 967–78. PMID: 6307529
4. Harding C, Heuser J, Stahl P. Receptor-mediated endocytosis of transferrin and recycling of the transferrin receptor in rat reticulocytes. The Journal of cell biology. 1983; 97 (2): 329–39. PMID: 6309857
5. Stoorvogel W, Kleijmeer MJ, Geuze HJ, Raposo G. The biogenesis and functions of exosomes. Traffic. 2002; 3 (5): 321–30. PMID: 11967126
6. Bang C, Thum T. Exosomes: new players in cell-cell communication. Int J Biochem Cell Biol. 2012; 44 (11): 2060–4. doi: 10.1016/j.biocel.2012.08.007 PMID: 22903023
7. Ludwig AK, Giebel B. Exosomes: small vesicles participating in intercellular communication. Int J Biochem Cell Biol. 2012; 44 (1): 11–5. doi: 10.1016/j.biocel.2011.10.005 PMID: 22024155
8. van Niel G, Porto-Carreiro I, Simoes S, Raposo G. Exosomes: a common pathway for a specialized function. J Biochem. 2006; 140 (1): 13–21. PMID: 16877764
9. Properzi F, Logozzi M, Fais S. Exosomes: the future of biomarkers in medicine. Biomarkers in medicine. 2013; 7 (5): 769–78. doi: 10.2217/bmm.13.63 PMID: 24044569
10. Mabert K, Cojoc M, Peitzsch C, Kurth I, Souchelnytskyi S, Dubrovska A. Cancer biomarker discovery: current status and future perspectives. International journal of radiation biology. 2014; 90 (8): 659–77. doi: 10.3109/09553002.2014.892229 PMID: 24524284
11. Colas E, Perez C, Cabrera S, Pedrola N, Monge M, Castellvi J, et al. Molecular markers of endometrial carcinoma detected in uterine aspirates. International journal of cancer. International journal of cancer. 2011; 129 (10): 2435–44. doi: 10.1002/ijc.25901 PMID: 21207424
12. Diamandis EP. Present and future of cancer biomarkers. Clinical chemistry and laboratory medicine. 2014; 52 (6): 791–4. doi: 10.1515/cclm-2014-0317 PMID: 24803613
13. Keller S, Ridinger J, Rupp AK, Janssen JW, Altevogt P. Body fluid derived exosomes as a novel template for clinical diagnostics. J Transl Med. 2011; 9:86. doi: 10.1186/1479-5876-9-86 PMID: 21651777
14. Mitchell PJ, Welton J, Staffurth J, Court J, Mason MD, Tabi Z, et al. Can urinary exosomes act as treatment response markers in prostate cancer? J Transl Med. 2009; 7: 4. doi: 10.1186/1479-5876-7-4 PMID: 19138409
15. Mizutani K, Terazawa R, Kameyama K, Kato T, Horie K, Tsuchiya T, et al. Isolation of prostate cancer-related exosomes. Anticancer Res. 2014; 34 (7): 3419–23. PMID: 24982349
16. Szajnik M, Derbis M, Lach M, Patalas P, Michalak M, Drzewiecka H, et al. Exosomes in Plasma of Patients with Ovarian Carcinoma: Potential Biomarkers of Tumor Progression and Response to Therapy. Gynecol Obstet (Sunnyvale). 2013; Suppl 4: 3.
17. Melo SA, Luecke LB, Kahlert C, Fernandez AF, Gammon ST, Kaye J, et al. Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. Nature. 2015; 523 (7559): 177–82. doi: 10.1038/nature14581 PMID: 26106858
18. Wei R, Li G, Seymour AB. Multiplexed, quantitative, and targeted metabolite profiling by LC-MS/MRM. Methods Mol Biol. 2014; 1198: 171–99. doi: 10.1007/978-1-4939-1258-2_12 PMID: 25270930
19. Thery C, Amigorena S, Raposo G, Clayton A. Isolation and characterization of exosomes from cell culture supernatants and biological fluids. Current protocols in cell biology. 2006; Chapter 3, Unit 3 22.

20. Sheikh KD, Khanna S, Byers SW, Fornace A, Cheema AK. Small molecule metabolite extraction strategy for improving LC/MS detection of cancer cell metabolome. J Biomol Tech. 2011; 22 (1): 1–4. PMID: 21455475

21. Gowda H, Ivanisevic J, Johnson CH, Kurczy ME, Benton HP, Rinehart D, et al. Interactive XCMS Online: simplifying advanced metabolomic data processing and subsequent statistical analyses. Analytical chemistry. 2014; 86 (14): 6931–9. doi: 10.1021/ac500734c PMID: 24934772

22. Cui Q, Lewis IA, Hegeman AD, Anderson ME, Li J, Schulte CF, et al. Metabolite identification via the Madison Metabolomics Consortium Database. Nature biotechnology. 2008; 26 (2): 162–4. doi: 10.1038/nbt0208-162 PMID: 18259166

23. Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, et al. HMDB 3.0—The Human Metabolome Database in 2013. Nucleic acids research. 2013; 41: D801–7. doi: 10.1093/nar/gks1065 PMID: 23161693

24. Sud M, Fahy E, Cotter D, Brown A, Dennis EA, Glass CK, et al. LMSD: LIPID MAPS structure database. Nucleic acids research. 2007; 35: D527–32. PMID: 17098933

25. Tautenhahn R, Cho K, Uritboonthai W, Zhu Z, Patti GJ, Siuzdak G. An accelerated workflow for untargeted metabolomics using the METLIN database. Nature biotechnology. 2012; 30 (9): 826–8.

26. Xia J, Psychogios N, Young N, Wishart DS. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. Nucleic acids research. 2009; 37: W652–60. doi: 10.1093/nar/gkp356 PMID: 19429898

27. Mazzeo C, Canas JA, Zafra MP, Marco AR, Fernandez-Nieto M, Sanz V, et al. Exosome secretion by eosinophils: A possible role in asthma pathogenesis. The Journal of allergy and clinical immunology. 2015; pii: S0091-6749(14)01724-2.

28. Franquesa M, Hoogduijn MJ, Ripoll E, Luk F, Salih M, Betjes MG, et al. Update on controls for isolation and quantification methodology of extracellular vesicles derived from adipose tissue mesenchymal stem cells. Frontiers in immunology. 2014; 5: 525. doi: 10.3389/fimmu.2014.00525 PMID: 25374572

29. Krantz SB, Shields MA, Dangi-Garimella S, Bentrem DJ, Munshi HG. Contribution of Epithelial-Mesenchymal transition to pancreatic cancer progression. Cancers. 2010; 2: 2084–2097. doi: 10.3390/cancers2042084 PMID: 24281219

30. Katsun Y, Lamouille S, Derunck R. TGF-β signaling and epithelial-mesenchymal transition in cancer progression. Current Opinion in Oncology. 2013;, 25(1): 76–84. PMID: 23197193

31. van der Pol E, Boing AN, Harrison P, Sturk A, Nieuwland R. Classification, functions, and clinical relevance of extracellular vesicles. Pharmacological reviews. 2012; 64 (3): 676–705. doi: 10.1124/pr.112.005983 PMID: 22722893

32. Yang Y, Yang X, Yang Y, Zhu H, Chen X, Zhang H, et al. Exosomes: a promising factor involved in cancer hypoxic microenvironments. Current medicinal chemistry. 2015 Aug 25.

33. Zeczycki TN, Whelan J, Hayden WT, Brown DA, Shaikh SR. Increasing levels of cardiolipin differentially influence packing of phospholipids found in the mitochondrial inner membrane. Biochemical and biophysical research communications. 2014; 450 (1): 366–71. doi: 10.1016/j.bbrc.2014.05.133 PMID: 24905496

34. Cohen-Yeshurun A, Willner D, Trembovler V, Alexandrovich A, Mechoulam R, Shohami E, et al. N-arachidonoyl-L-serine (AraS) possesses proneurogenic properties in vitro and in vivo after traumatic brain injury. Journal of the International Society of Cerebral Blood Flow and Metabolism. 2013; 33 (8): 1242–50.

35. Ho WS. Angiogenesis: a new physiological role for N-arachidonoyl serine and GPR55? British journal of pharmacology. 2010; 160 (7): 1580–2. doi: 10.1111/j.1476-5381.2010.00788.x PMID: 20649562

36. Makabi-Panzu B, Sprott GD, Patel GB. Coenzyme Q10 in vesicles composed of archaeal ether lipids or conventional lipids enhances the immuno-adjuvanticity to encapsulated protein. Vaccine. 1998; 16 (16): 1504–10. PMID: 9711796

37. Omri A, Makabi-Panzu B, Agnew BJ, Sprott GD, Patel GB. Influence of coenzyme Q10 on tissue distribution of archaeosomes, and pegylated archaeosomes, administered to mice by oral and intravenous routes. Journal of drug targeting. 2000; 7 (5): 383–92. PMID: 10721800

38. Guidetti P, Schwarcz R. Determination of alpha-aminoadipic acid in brain, peripheral tissues, and body fluids using GC/MS with negative chemical ionization. Molecular brain research. 2003; 118 (1–2): 132–9. PMID: 14559362

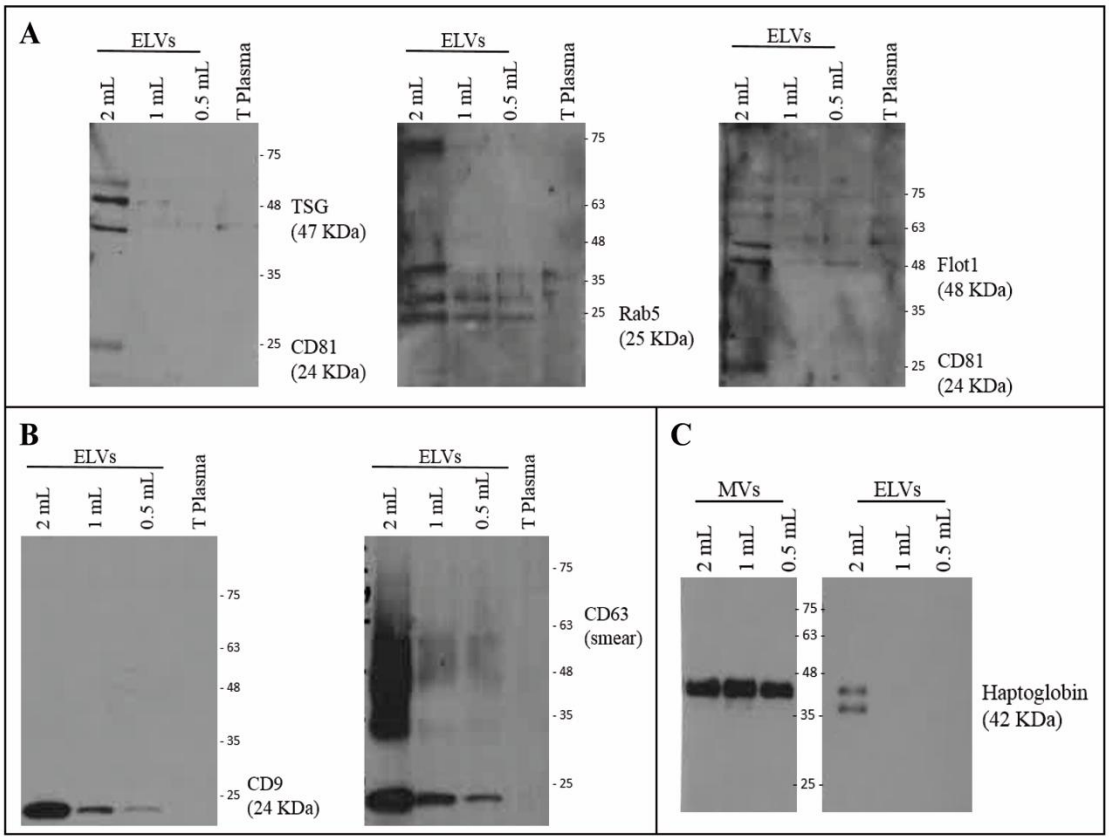
39. Healy ZR, Liu H, Holtzclaw WD, Talalay P. Inactivation of tautomerase activity of macrophage migration inhibitory factor by sulforaphane: a potential biomarker for anti-inflammatory intervention. Cancer epidemiology, biomarkers & prevention. 2011; 20 (7): 1516–23.

40. Pham H, Banerjee T, Ziboh VA. Suppression of cyclooxygenase-2 overexpression by 15S-hydroxyeicosatrienoic acid in androgen-dependent prostatic adenocarcinoma cells. International journal of cancer. 2004; 111 (2): 192–7.

41. Stahl W, Hanusch M, Sies H. 4-oxo-retinoic acid is generated from its precursor canthaxanthin and enhances gap junctional communication in 10T1/2 cells. Advances in experimental medicine and biology. 1996; 387: 121–8. PMID: 8794203

42. Kaur P, Rizk N, Ibrahim S, Luo Y, Younes N, Perry B, et al. Quantitative metabolomic and lipidomic profiling reveals aberrant amino acid metabolism in type 2 diabetes. Mol Biosyst. 2013; 9(2):307–17. doi: 10.1039/c2mb25384d PMID: 23247761

SUPPLEMENTARY FIGURE 1. Analysis of protein expression by western blot of Exosome-Like Vesicles (ELVs) and microvesicles (MVs) isolated from human plasma (2000, 1000 and 500 µL). Expression was also analyzed in non-purified total plasma (T Plasma). Markers CD81, TSG101, Rab5 and Flotillin 1 were blotted in the same membrane (Panel A) and CD63 and CD9 were blotted in an independent membrane (Panel B). Expression of the soluble protein haptoglobin was analyzed in MVs and ELVs isolated from different volumes of T Plasma.



SUPPLEMENTARY TABLE 1. Patient clinic-pathological information.

SUPPLEMENTARY TABLE 1. Patient clinic-pathological information.

Samples used for the "Characterization of Plasma ELVs" Results section:

Diagnosis	Number of patients	Grade	F.I.G.O.	Type I/II	Pre/Postm enopausal	Plasma
Endometrial adenosquamous carcinoma	1	G3	IIIC1	I	Post	10 ml
Endometrioid adenocarcinoma	1	G2	IA	I	Post	10 ml
Endometrioid adenocarcinoma	1	G2	IA	I	Post	10 ml

Samples used for the "ELVs metabolome as a biomarker source" Results section:

Diagnosis	Number of patients	Grade	F.I.G.O.	Type I/II	Pre/Postm enopausal	Plasma
Healthy endometrium	13	-	-	-	Post	10 ml
Endometrioid adenocarcinoma	10	G1 and 2	IA	I	Post	10 ml
Endometrioid adenocarcinoma	9	G1 and 2	IB	I	Post	10 ml

SUPPLEMENTARY TABLE 2. Coefficient of variation of Internal Standards in each group of samples (biological replicates).

SUPPLEMENTARY TABLE 2. Coefficient of variation of Internal Standards in each group of samples (biological replicates).

	Plasma ELVs pool EC patients	Plasma ELVs Normal Patients	Plasma ELVs Tumor Patients	PANC1 Control ELVs	PANC 1 treated ELVs
Internal Standards	CV %	CV %	CV %	CV %	CV %
4-Nitrobenzoic acid (negative mode)	7,23	5,5	9,8	6,06	8,55
Debrisoquine (positive mode)	4,61	13,2	12,21	4,17	13,83
Number of samples (n)	3	13	19	3	3

SUPPLEMENTARY TABLE 3. Comparative analyses of putative features identified in human plasma ELVs. Ratios of relative intensities for a selection of metabolites is represented.

SUPPLEMENTARY TABLE 3. Comparative analyses of putative features identified in human plasma ELVs. Ratios of rela-tive intensities for a selection of metabolites is represented.

m/z	RT	Metabolite ID	Ionization mode	Ratio 1000/500	Ratio 2000/1000
395,2219	5,3950	PA(15:0/0:0)	Negative	0,7783	1,1731
460,3302	9,4409	D-Glucosylsphingosine	Negative	0,9124	0,9952
637,5808	10,3574	DG(22:0/15:0/0:0)	Negative	0,8908	1,3497
668,5319	10,2023	DG(20:4/0:0/20:4)	Negative	0,8063	1,1711
718,5728	10,3213	PE(18:0/17:0)	Negative	0,8904	1,7432
760,6202	10,3574	PC(O-20:0/15:0)	Negative	0,9541	1,3554
831,6495	10,4462	TG(12:0/17:2/22:6)	Negative	0,8242	1,9401
878,6634	10,4673	PE(22:4/24:0)	Negative	0,9837	1,3979
920,7110	10,5165	PC(22:4/24:0)	Negative	0,8946	1,4059
997,8225	0,5978	TG(19:0/22:4/22:5)	Negative	0,8945	1,1034
1004,8050	0,5951	PC(32:0/20:4)	Negative	0,8773	1,0879
1031,8104	0,5904	TG(22:3/22:4/22:6)	Negative	0,9650	1,1154
256,2630	7,8447	Palmitamide	Positive	0,9979	1,3969
282,2779	7,2447	Oleamide	Positive	0,8253	0,9336
288,2899	4,6437	C17 Sphinganine	Positive	1,2345	1,7391
291,2317	7,1328	4-Androstenediol	Positive	0,9206	1,1744
300,2901	5,7272	Sphingosine	Positive	2,7210	0,9200
312,1293	8,0042	1,7-Dimethylguanosine	Positive	1,9035	1,0859
495,1368	11,2462	N-Adenylyl-L-phenylalanine	Positive	0,8223	1,2546
510,4905	10,1660	Cer(d18:1/14:0)	Positive	0,9844	1,1879
569,3800	9,9857	PG(22:0/0:0)	Positive	1,0261	0,9884
594,5826	10,3552	Cer(d14:1/24:0)	Positive	0,8370	1,1296
623,3210	10,1978	PI(20:3/0:0)	Positive	0,9290	1,2723
638,6108	10,1346	Cer(d16:1/24:0)	Positive	0,8504	1,3676
664,4566	10,1916	PS(P-16:0/12:0)	Positive	0,9114	1,3743
667,6264	10,2314	DG(17:0/22:0/0:0)	Positive	0,9878	1,2708
675,5319	10,2413	PA(16:0/19:1)	Positive	0,9719	2,1148
692,4872	7,4109	PS(16:0/14:0)	Positive	0,9503	1,2294
701,5598	10,1920	SM(d18:2/16:0)	Positive	0,9333	3,0731
705,4710	7,1219	PG(14:1/17:1)	Positive	0,8225	2,3582
733,5388	10,1930	PG(P-18:0/16:1)	Positive	0,9378	1,9442
736,5146	7,4079	PS(14:0/18:0)	Positive	0,9658	1,4603
743,6180	10,2878	TG(12:0/14:0/18:4)	Positive	0,9339	3,3803
748,5303	9,3035	PE(22:6/P-16:0)	Positive	0,7806	1,4810
765,5610	10,1723	PG(14:0/21:0)	Positive	0,9586	3,8208
769,6372	10,5314	TG(14:1/14:1/18:3)	Positive	1,1214	0,6144

SUPPLEMENTARY TABLE 4. Metabolome profiling of ELVs isolated from human plasma samples and cell culture media. Table 4.A. Putative identifications in plasma ELVs.

SUPPLEMENTARY TABLE 4. Metabolome profiling of ELVs isolated from human plasma samples and cell culture media. Table 4.A. Putative identifications in plasma ELVs.

Metabolite	m/z	RT (min)	Mass error (ppm)	Mode	MMCD ID	KEGG ID	PubChem SID	PubChem CID	HMDB ID
Glycerophospholipids/Shingolipids									
18:1-18:2-MGDG	781,582	10,171	0,000	Pos.		-	-	-	-
DG(16:0/17:0/0:0)	583,529	10,105	0,001	Pos.		-	14710864	-	-
DG(18:0/19:0/0:0)	637,580	10,337	0,002	Neg.		-	14710917	-	-
DISTEAROYL-MONOGALACTOSYL-DIGLYCERIDE	787,628	10,208	0,001	Pos.		-	-	-	-
GalCer(d18:2/16:0)	696,541	10,231	0,001	Neg.		-	123068846;123068819	-	-
Glucosylceramide (d18:1/16:0)	700,570	10,162	0,003	Pos.		-	-	6475228	HMDB04971
GPCho(18:0/14:0)	735,576	10,144	0,001	Pos.		-	-	3082164	HMDB05261
LysoPE(20:4/0:0)	500,280	6,104	0,002	Neg.		-	-	53480952	HMDB11518
MG(0:0/18:0/0:0)	359,314	8,814	0,001	Pos.		-	-	79075	HMDB11535
PA(21:0/18:1)	743,559	10,191	0,000	Neg.		-	123066845	-	-
PA(22:6/0:0)	483,251	5,534	0,000	Pos.		-	123067334	-	-
PC(20:4/32:0)	1004,806	0,583	0,001	Neg.		-	7983542	-	-
PC(P-18:1/20:1)	798,635	10,202	0,002	Pos.		-	-	53480813	HMDB11315
PC(P-19:1/0:0)	520,376	8,735	0,000	Pos.		-	74380419	-	-
PE(12:0/16:0)	636,459	7,123	0,001	Pos.		-	123062508	-	-
PE(22:2/16:1)	768,551	10,178	0,004	Neg.		C00350	123062324;-	52924779	HMDB09551
PE-Cer(d14:2/23:0)	701,556	9,748	0,003	Pos.		-	160780231	-	-
PG(16:0/16:0)	707,520	10,129	0,002	Pos.		-	123064911	-	-
PG(16:0/20:0)	777,563	10,092	0,002	Neg.		-	123064796	-	-
PHOSPHONYL CHOLINE	637,468	9,396	0,000	Pos.		-	-	-	-
phytosphingosine	614,572	10,358	0,000	Pos.		-	-	-	-
PI(16:0/22:4)	885,550	10,107	0,000	Neg.		C00626	123065958;-	-	HMDB09794
PI(P-16:0/18:4)	815,507	7,091	0,000	Pos.		-	123066090	-	-
PI-Cer(d20:0/18:0)	836,602	10,334	0,000	Neg.		-	160780305	-	-
PS(16:0/13:0)	680,483	7,114	0,003	Pos.		-	123063650	-	-
PS(21:0/0:0)	566,345	6,657	0,002	Neg.		-	123063853	-	-
PS-NAc(18:1/18:1/19:0)	1066,806	0,589	0,001	Neg.		-	49703606	-	-
Psychosine sulfate	542,302	10,116	0,002	Pos.		C02744	5704	5280538	-
SLBPA(54:3)	1039,795	0,593	0,002	Pos.		-	49703613	-	-
SM(d18:1/16:0)	703,574	9,993	0,001	Pos.		-	7850646	-	-
Sulfogalactosylceramide	890,636	8,786	0,003	Pos.		C06125	49703648;-	24779587	HMDB12318
TG(12:0/12:0/20:5)	741,601	8,815	0,002	Pos.		-	135638336	-	-
TG(13:0/18:3/22:6)	857,668	10,277	0,001	Neg.		-	135639669	-	-
Fatty acid esters/amides/amines/alcohols									
arachidonoyl serine	392,280	9,252	0,001	Pos.		-	74380323;7850527	-	-
9-Oxoasimicinone	637,468	9,396	0,001	Pos.		-	-	-	HMDB32949
Biotinyl-CoA			0,002	Pos.	cq_01214				
BUTAN-AMINE			0,001	Pos.	cq_11560				
Dimethyl-oxooctyl-CoA			0,003	Pos.	cq_19613				
Dodecanamide	200,201	5,785	0,000	Pos.		C13831	157540	14256	-
Hexadecenoyl-CoA			0,001	Pos.	cq_02941				
Hydroxydecanoyl-CoA			0,004	Pos.	cq_02933				
Hydroxydodecanoyl-CoA			0,004	Pos.	cq_02931				
Hydroxyhexanoyl-CoA			0,003	Pos.	cq_02937				
Hydroxyisopentyl-CoA			0,003	Pos.	cq_03374				
hydroxyphenylpropionyl-CoA			0,005	Pos.	cq_18165				
Malonyl-CoA	853,593	10,026	0,001	Neg.	cq_00054				
N-Oleoylethanolamine	326,306	8,578	0,000	Pos.		-	-	5283454	HMDB02088
Oleamide	282,279	8,020	0,000	Pos.		C19670	49703514;-;135637086	5283387	HMDB02117
Tetradecenoyl-CoA			0,004	Pos.	cq_03890				
Virodhamine	348,291	8,073	0,001	Pos.		-	123060035;-	5712057	HMDB13655
PHENYLETHYLAMINE			0,002	Pos.	cq_13599				
Organic Acids and Derivatives									
Hexacosanedioic acid	427,379	9,958	0,001	Pos.		-	14710839	-	-
Phosphopantothenate			0,000	Pos.	cq_02083				
Nucleotides and derivatives									
ADENOSINE			0,001	Pos.	cq_16117				
ADENOSINE 5'-MONOPHOSPHATE	405,819	0,787	0,001	Neg.	cq_11699				
CDP-N-methylethanolamine			0,001	Pos.	cq_02081				
CoA			0,000	Pos.	cq_00007				
CYTIDINE 5'-MONOPHOSPHATE	395,249	7,612	0,002	Neg.	cq_12390				
DEOXYURIDINE-5'-DIPHOSPHATE	387,155	8,072	0,001	Neg.	cq_13046				
Diadenosine heptaphosphate	1076,953	10,305	0,002	Pos.		-	-	53477728	HMDB01433
DIMETHYL-AMINO-ADENOSINE-MONOPHOSPHATE			0,000	Pos.	cq_11436				
DIMETHYLGUANOSINE-MONOPHOSPHATE			0,001	Pos.	cq_14205				
FLUORO-DEOXYTHYMIDINE MONOPHOSPHATE	323,194	8,793	0,001	Neg.	cq_13235				
GUANOSINE-5'-MONOPHOSPHATE	421,155	5,477	0,001	Neg.	cq_13383				
GUANOSINE-5'-TRIPHOSPHATE			0,001	Pos.	cq_14756				
METHYLTHIO-ADENOSINE-5'-MONOPHOSPHATE			0,002	Pos.	cq_10977				
PROPYL THYMIDINE-5-MONOPHOSPHATE			0,000	Pos.	cq_14827				
Ribose triphosphate			0,000	Pos.	cq_01737				
Se-Adenosylselenomethionine			0,002	Pos.	cq_03206				
THYMIDINE-5'-MONOPHOSPHATE	426,125	5,109	0,002	Pos.		-	-	-	-
UDP-AAGM-DIAMINOHEPTANEDIOATE			0,002	Pos.	cq_10399				
DIDEOXYCYTOSINE-DIPHOSPATE			0,000	Pos.	cq_12850				
Amino acids/conjugates									
L-histidine			0,000	Pos.	cq_02573				
N-oleoyl glutamic acid	412,307	5,233	0,001	Pos.		-	123060503	-	-
tyrosyl-AMP	527,128	5,479	0,001	Pos.		C11462	13633	443217	-
Peptides/peptide conjugates									
Ala Lys Thr Arg	473,282	1,354	0,002	Neg.			-	-	-
CoA-glutathione			0,001	Pos.	cq_00628				
Cys Phe Ser Lys	484,221	4,336	0,001	Pos.			-	-	-
Lys Lys Pro Leu	485,347	7,151	0,002	Pos.			-	-	-
Phe Trp Phe Arg	655,336	8,052	0,001	Pos.			-	-	-
Pro Asn Phe Gly	432,188	0,838	0,001	Neg.			-	-	-
Tyr Phe Phe Trp	662,299	8,056	0,002	Pos.			-	-	-
Tyr Val Arg Tyr	600,313	7,998	0,002	Pos.			-	-	-
Val Lys Lys Lys	502,373	7,149	0,002	Pos.			-	-	-
S-nitrosoglutathione			0,000	Pos.	cq_19046				
Nitro-glutathionyl-hydroxy-dihydronaphthalene			0,001	Pos.	cq_16463				
Sugars and sugar conjugates									
Acetyl glucosamine bisphosphate	381,291	8,638	0,001	Neg.	cq_02586				
anhydro-deoxy-galactopyranosyl-D-arabino-Hex-1-enitol	366,140	6,514	0,001	Pos.		-	-	-	HMDB02282
BUTANOYL-AMINO-DEOXY-GLUCOPYRANOSIDE			0,001	Pos.	cq_12179				
BUTANOYL GLUCOSAMINE			0,002	Pos.	cq_02737				
D-Glucosaminide			0,002	Pos.	cq_17588				
D-Mannosylglycoprotein	902,827	0,588	0,002	Neg.	cq_01971				
Fucosyllactose	489,182	4,330	0,001	Pos.		-	-	21771334	HMDB06620
Glycoprotein phospho-D-mannose			0,004	Pos.	cq_02353				
Cyclic alcohols, aliphatic and aromatic compounds									
10-formyldihydrofolate	472,159	5,484	0,002	Pos.		C03204	-	168809	HMDB06485
hydroxymethylbilane			0,003	Pos.	cq_19312				
PHENYLLACTATE			0,001	Pos.	cq_11841				
POLYPROPYLENE GLYCOL	612,469	8,610	0,000	Pos.		-	-	-	-
Urobilinogen			0,001	Pos.	cq_03242				
Picolinic acid	122,024	4,043	0,001	Neg.		C10164	-	1018	HMDB02243
Steroids, prenols and eichosanoids									
15-HETE-Gly	378,265	5,335	0,001	Pos.		-	126522571	-	-
25-hydroxy-hexadehydrovitamin D3	393,280	8,787	0,000	Neg.		-	14715099	-	-
3-Deoxyvitamin D3	369,350	10,181	0,002	Pos.		-	49703700	-	-
CE(20:5)	669,563	10,397	0,001	Neg.		C02530	24702211;-	53477889	HMDB06731
Coenzyme Q10	861,677	10,301	0,000	Neg.		C11378;C00399	4266323	5926222;5281915;-	HMDB02139;HMDB01072;HMDB06709
dihydroxy-oxavitamin D3	419,316	9,655	0,000	Pos.		-	14715071	-	-
Fluoro-hydroxyhydrocortisone	397,202	7,978	0,000	Pos.		C14638	-	-	-
hydroxy-dihydrovitamin D3	433,332	9,833	0,000	Pos.		-	14715203;14715202	-	-
hydroxy-norvitamin D3	441,298	9,691	0,000	Pos.		-	14715032	-	-
Nutriacholic acid	391,283	9,262	0,002	Pos.		-	380848	313030	HMDB00467
trifluoro-hydroxyvitamin D3	455,311	9,809	0,002	Pos.		-	14715138	-	-
Ubiquinone-9	793,610	10,139	0,004	Neg.		C01967	5068	5280473	-
Others									
Palmitoyl glucuronide	417,287	6,764	0,001	Neg.		C03033	-	161223	HMDB10331

SUPPLEMENTARY TABLE 4. Metabolome profiling of ELVs isolated from human plasma samples and cell culture media. Table 4.B. Putative identifications in PANC-1 cell culture media ELVs.

SUPPLEMENTARY TABLE 4. Metabolome profiling of ELVs isolated from human plasma samples and cell culture media. Table 4.B. Putative identifications in PANC-1 cell culture media ELVs.

Name	m/z	Mass error (ppm)	Mode	KEGG ID	PubChem SID	PubChem CID	HMDB ID
Glycerophospholipids/Shingolipids							
(3'-sulfo)Galbeta-Cer(d18:1/24:1(15Z)(2OH))	906,6315022	0,001953125	Pos.	-	24779588	-	-
(4E,8Z,d18:2) sphingosine	298,2743381	0,000274658	Pos.	-	11993697;11220228	-	-
16:4(4Z,7Z,10Z,13Z)	249,1845266	3,51E-04	Pos.	-	5312433	-	-
1-O-eicosanoyl-Cer(d18:1/16:0)	832,8123017	6,71E-04	Pos.	-	-	-	-
1-STEAROYL-2-OLEOYL-SN-GLYCERO-3-PHOSPHATE	702,5226965	0,003295898	Pos.	-	-	-	-
3-Dehydrosphinganine	300,2892092	4,88E-04	Pos.	C02934	5855	439853	HMDB01480
CDP-DG(18:1(11Z)/20:4(5Z,8Z,11Z,14Z))	1028,540369	0,003173828	Pos.	-	53477934	-	HMDB06991
Cer(d14:1(4E)/26:0(2OH))	638,6069683	0,001159668	Pos.	-	70698960	-	-
DG(16:0/22:3(10Z,13Z,16Z)/0:0)[iso2]	647,5590049	0,001831055	Pos.	-	9543781	-	-
iso (4E,15-methyl-d16:1) sphingosine	286,2743382	0,000274658	Pos.	-	42608345	-	-
LacCer(d18:0/18:0)	892,6748602	0,002929688	Pos.	-	52931265	-	-
LacCer(d18:1/22:0)	946,7217046	0,002807617	Pos.	-	44260159;44260143;44261959	-	-
LysoPE(0:0/18:1(11Z))	478,2919262	0,001983643	Neg.	-	53480924;-	53480924	HMDB11475
LysoPE(0:0/20:4(5Z,8Z,11Z,14Z))	500,2757868	0,002502441	Neg.	-	53480936;-	53480936	HMDB11487
LysoPE(0:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	526,2942355	0,001464844	Pos.	-	53480945	-	HMDB11496
N,N-dimethylsphingosine	328,3203729	0,000579834	Pos.	-	5282309	-	-
PA(15:1(9Z)/0:0)	395,2203355	0,001068115	Pos.	-	52929755	-	-
PA(19:1(9Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	761,512688	0,001159668	Pos.	-	52929101	-	-
PA(20:4(5Z,8Z,11Z,14Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	769,4795955	6,71E-04	Pos.	-	52929239	-	-
PA(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/19:1(9Z))	761,512688	0,001159668	Pos.	-	52929426	-	-
PA(O-16:0/16:1(9Z))	633,4852947	6,10E-05	Pos.	-	52929571	-	-
PA(O-20:0/0:0)	453,3362216	0,002258301	Pos.	-	52929769	-	-
PC(0:0/16:0)	496,3389828	7,63E-04	Pos.	-	15061532	-	-
PC(17:0/22:4(7Z,10Z,13Z,16Z))	822,6039162	0,00213623	Neg.	-	52922519	-	-
PC(17:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	820,5864184	0,001342773	Pos.	-	24778795	-	-
PC(18:0/24:1(15Z))	872,7059533	0,004333496	Pos.	-	24778881	-	HMDB08059
PC(18:1(11Z)/P-18:0)	772,6219955	5,49E-04	Pos.	-	70698809	-	-
PC(18:4(6Z,9Z,12Z,15Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	826,541397	0,003234863	Pos.	-	52922939	-	-
PC(19:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	848,6124677	0,003845215	Pos.	-	52922989	-	-
PC(19:3(10Z,13Z,16Z)/0:0)	532,3401855	4,27E-04	Pos.	-	24779452	-	-
PC(20:0/26:0)	930,7890269	4,27E-04	Pos.	-	24779059	-	-
PC(20:4(5Z,8Z,11Z,14Z)/32:0)	1004,80449	8,54E-04	Neg.	-	24779078	-	-
PC(20:5(5Z,8Z,11Z,14Z,17Z)/0:0)	542,3249609	9,16E-04	Pos.	-	11757087	-	HMDB10397
PC(O-16:0/O-1:0)	496,376635	0,000488281	Pos.	-	127599	-	-
PC(P-20:0/16:1(9Z))	772,6219955	0,000549316	Pos.	-	52923990	-	-
PE(14:1(9Z)/14:1(9Z))	632,4269266	0,001647949	Pos.	-	52924935	-	-
PE(16:0/0:0)	454,2940624	0,001312256	Pos.	-	9547069	-	HMDB11503
PE(18:4(6Z,9Z,12Z,15Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	784,4873289	0,003845215	Pos.	-	52924469	-	-
PE(22:4(7Z,10Z,13Z,16Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	840,5574337	0,003662109	Pos.	-	52923868	-	-
PE(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/0:0);LysoPE(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/0:0)	524,2772358	0,001037598	Neg.	-	52925132;-	52925132	HMDB11526
PE(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/18:4(6Z,9Z,12Z,15Z))	784,4873289	0,003845215	Pos.	-	52924848	-	-
PE(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/20:0)	820,5864184	0,001342773	Pos.	-	52924851	-	-
PE(24:1(15Z)/18:3(9Z,12Z,15Z))	822,6039162	0,002075195	Neg.	C00350	-	53480027	HMDB09755
PG(12:0/0:0)	429,2268203	0,002044678	Pos.	-	52927434	-	-
PG(12:0/18:4(6Z,9Z,12Z,15Z))	687,4227584	4,27E-04	Pos.	-	52926321	-	-
PG(13:0/0:0)	443,2406232	1,83E-04	Pos.	-	42607482	-	-
PG(13:0/22:2(13Z,16Z))	761,5356875	0,003051758	Pos.	-	52926355	-	-
PG(14:0/20:5(5Z,8Z,11Z,14Z,17Z))	741,4690493	0,001037598	Pos.	-	52927187	-	-
PG(15:0/20:5(5Z,8Z,11Z,14Z,17Z))	755,4862178	4,27E-04	Pos.	-	52926420	-	-
PG(18:1(9Z)/0:0)	511,3038566	7,93E-04	Pos.	-	9547135	-	-
PG(19:0/0:0)	527,3319051	0,002441406	Pos.	-	52927454	-	-
PG(19:1(9Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	835,5468304	0,001525879	Pos.	-	52926775	-	-
PG(20:1(11Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	849,5612849	0,002685547	Pos.	-	52926822	-	-
PG(20:5(5Z,8Z,11Z,14Z,17Z)/0:0)	531,2733265	0,001525879	Pos.	-	52927453	-	-
PG(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/0:0)	557,2890534	0,001708984	Pos.	-	52927439	-	-
PG(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/19:1(9Z))	835,5468304	0,001525879	Pos.	-	52927099	-	-
PG(8:0/8:0)	499,2644088	0,002258301	Pos.	-	9547125	-	-
PI(10:0/16:0)	725,4229571	0,001647949	Neg.	-	70698913	-	-
PI(12:0/15:0)	741,4578214	0,002990723	Pos.	-	52927468	-	-
PI(14:0/0:0)	545,2735774	0,001403809	Pos.	-	52928603	-	-
PI(17:1(10Z)/0:0)	585,3022509	0,001220703	Pos.	-	42607494	-	-
PI(19:0/0:0)	613,336679	8,54E-04	Neg.	-	52928622	-	-
PI(O-16:0/20:5(5Z,8Z,11Z,14Z,17Z))	843,5358716	0,002258301	Pos.	-	52928489	-	-
PI(P-18:0/18:4(6Z,9Z,12Z,15Z))	843,5358716	0,002258301	Pos.	-	52928551	-	-
PI-Cer(d18:0/20:0(2OH))	852,5982073	0,001098633	Neg.	-	70699104	-	-
PI-Cer(d18:1/22:0)	862,6190738	0,001159668	Neg.	-	9547204	-	-
PI-Cer(t18:0/24:0(2OH))	926,6698145	6,10E-04	Pos.	-	70699112	-	-
PI-Cer(t20:0/18:0)	852,5982073	0,001098633	Neg.	-	70699091	-	-
PI-Cer(t20:0/22:0(2OH))	926,6698145	0,000610352	Pos.	-	70699117	-	-
PS(12:0/18:0)	708,4822126	0,001281738	Pos.	-	52926048	-	-
PS(12:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	752,4483401	0,001281738	Pos.	-	52926044	-	-
PS(13:0/17:0)	708,4783318	0,002624512	Pos.	-	52926042	-	-
PS(14:0/16:0)	708,4822126	0,001281738	Pos.	-	52926089	-	-
PS(14:1(9Z)/20:5(5Z,8Z,11Z,14Z,17Z))	752,4483401	0,001281738	Pos.	-	52925242	-	-
PS(18:3(6Z,9Z,12Z)/20:5(5Z,8Z,11Z,14Z,17Z))	804,4842722	0,003295898	Pos.	-	52925500	-	-
PS(18:3(9Z,12Z,15Z)/20:5(5Z,8Z,11Z,14Z,17Z))	804,4842722	0,003295898	Pos.	-	52925528	-	-
PS(22:0/22:1(11Z))	902,6861844	0,001708984	Pos.	-	52925832	-	-
PS(P-16:0/17:2(9Z,12Z))	730,5022978	5,49E-04	Pos.	-	52926193	-	-
Psychosine sulfate	542,3010029	0,001708984	Pos.	C02744	5704	5280538	-
Sphingosine	300,2891064	5,80E-04	Pos.	C00319	5280335	-	HMDB00252
TG(12:0/17:0/22:3(10Z,13Z,16Z))[iso6]	843,7455949	0,00201416	Pos.	-	56937182	-	-
TG(12:0/20:0/22:0)[iso6]	889,8218397	0,001098633	Neg.	-	56937377	-	-
TG(12:0/22:0/22:4(7Z,10Z,13Z,16Z))[iso6]	911,8042609	0,00189209	Pos.	-	56937444	-	-
TG(17:0/22:3(10Z,13Z,16Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))[iso6]	971,8042499	0,00189209	Pos.	-	9545990	-	-
TG(18:0/18:0/18:0)	889,8218397	0,001159668	Neg.	-	-	11146	HMDB05393
TG(19:0/22:5(7Z,10Z,13Z,16Z,19Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))[iso6]	995,8018995	0,004272461	Pos.	-	9546416	-	-
Fatty acid esters/amides/amines/alcohols							
2-Phenylacetamide	136,075016	0,000686646	Pos.	C02505	150762	-	-
3,8-dioxooct-5-enoyl-CoA	920,1684177	0,001464844	Pos.	-	56927763	-	-
Anandamide (18:4, n-3)	320,2568919	0,001464844	Pos.	-	5283448	-	-
Dodecanamide	200,200698	1,98E-04	Pos.	C13831	157540	14256	-
Oleamide	282,2805083	0,001342773	Pos.	-	5353370;5283387	-	-
Linoleamide	280,2641591	0,000701904	Pos.	-	6435901	-	-
m-Tyramine	138,0916713	3,36E-04	Pos.	-	-	11492	HMDB04989
Myristoyl-EA	272,2583911	3,05E-05	Pos.	-	8890	-	-
N-Acetylserotonin	219,1134948	6,87E-04	Pos.	C00978	213206	-	HMDB01238
N-arachidonoyl GABA	390,3002011	6,10E-05	Pos.	-	16759310	-	-
N-hydroxy arachidonoyl amine	320,2568919	0,001464844	Pos.	-	5283414	-	-
N-oleoyl ethanolamine	326,3061263	7,63E-04	Pos.	-	5283454	-	HMDB02088
N-oleoyl GABA	368,3170876	0,001220703	Pos.	-	16759340	-	-
O-glutarylcarnitine	276,1442025	3,05176E-05	Pos.	-	71464488;53481622	-	HMDB13130
O-methylmalonylcarnitine	262,128991	5,49E-04	Pos.	-	71464478;53481628	-	HMDB13133
O-succinylcarnitine	262,128991	0,000549316	Pos.	-	71464481	-	-
Palmitoleamide	254,2488906	0,00100708	Pos.	-	56936054	-	-
Palmitoleoyl-EA	298,2743381	0,000274658	Pos.	-	9835868	-	-
Palmitoyl-EA	300,2894194	0,000244141	Pos.	-	4671	-	HMDB02100
Pentadecanoyl-EA	286,2743382	2,75E-04	Pos.	-	14455157	-	-
Stearoyl-EA	328,3203729	5,80E-04	Pos.	-	27902	-	-
Tridecanamide	214,2165188	6,10E-05	Pos.	-	4167542	-	-
Organic Acids and Derivatives							
Allantoate	177,061704	0,00012207	Pos.	C00499	208923	-	HMDB01209
Aminoadipic acid	162,0758805	1,98E-04	Pos.	-	469	-	HMDB00510
enol-Phenylpyruvate	165,0547143	9,16E-05	Pos.	C02763	5719	641637	HMDB01237
Homocarnosine	241,1292891	2,29E-04	Pos.	C00884	675516	89235	HMDB00745
Tetradecanedioic acid	259,1901419	2,14E-04	Pos.	-	-	13185	HMDB00872
2-amino-isobutyric acid	104,0701151	0,000465393	Pos.	-	6119	-	HMDB01906
Nucleotides and derivatives							
3'-THIO-THYMIDINE-5'-PHOSPHATE	339,0398195	0,001190186	Pos.	-	-	-	-
5'-Methylthioadenosine	298,0969767	0,000152588	Pos.	C00170	3470	439176	HMDB01173
6- FLAVINE MONONUCLEOTIDE	681,2939175	0,003295898	Neg.	-	-	-	-

N-oleoyl ethanolamine	326,3061263	7,63E-04	Pos.	-		5283454	- HMDB02088
N-oleoyl GABA	368,3170876	0,001220703	Pos.	-		16759340	- -
O-glutarylcarnitine	276,1442025	3,05176E-05	Pos.	-	71464488;53481622	-	HMDB13130
O-methylmalonylcarnitine	262,128991	5,49E-04	Pos.	-	71464478;53481628	-	HMDB13133
O-succinylcarnitine	262,128991	0,000549316	Pos.	-		71464481	- -
Palmitoleamide	254,2488906	0,00100708	Pos.	-		56936054	- -
Palmitoleoyl-EA	298,2743381	0,000274658	Pos.	-		9835868	- -
Palmitoyl-EA	300,2894194	0,000244141	Pos.	-		4671	- HMDB02100
Pentadecanoyl-EA	286,2743382	2,75E-04	Pos.	-		14455157	- -
Stearoyl-EA	328,3203729	5,80E-04	Pos.	-		27902	- -
Tridecanamide	214,2165188	6,10E-05	Pos.	-		4167542	- -
Organic Acids and Derivatives							
Allantoate	177,061704	0,00012207	Pos.	C00499		208923	- HMDB01209
Aminoadipic acid	162,0758805	1,98E-04	Pos.	-		469	- HMDB00510
enol-Phenylpyruvate	165,0547143	9,16E-05	Pos.	C02763		5719	641637 HMDB01237
Homocarnosine	241,1292891	2,29E-04	Pos.	C00884		675516	89235 HMDB00745
Tetradecanedioic acid	259,1901419	2,14E-04	Pos.	-	-		13185 HMDB00872
2-amino-isobutyric acid	104,0701151	0,000465393	Pos.	-		6119	- HMDB01906
Nucleotides and derivatives							
3'-THIO-THYMIDINE-5'-PHOSPHATE	339,0398195	0,001190186	Pos.	-	-	-	-
5'-Methylthioadenosine	298,0969767	0,000152588	Pos.	C00170		3470	439176 HMDB01173
6- FLAVINE MONONUCLEOTIDE	681,2939175	0,003295898	Neg.	-	-	-	-
6-AMINO-1-METHYLPURINE	151,0847336	0,000518799	Pos.	-	-	-	-
7-METHYL-6-THIO-GUANOSINE	314,0915715	1,53E-04	Pos.	-	-	-	-
Adenine	136,0622285	4,58E-04	Pos.	C00147		149245	- HMDB00034
dUMP	307,0324934	0,001190186	Neg.	C00365	-		65063 HMDB01409
N1-methyladenine	151,0847336	5,19E-04	Pos.	-	-	-	-
S-Adenosyl-L-methionine	399,1450817	5,80E-04	Pos.	C00019		176451	34755 HMDB01185
S-ADENOSYLMETHIONINE METHYL ESTER	415,1770317	0,001220703	Pos.	-	-	-	-
UDP-D-glucosamine	550,084437	0,001037598	Pos.	-	-		6857448 -
Amino acids/conjugates							
4-Oxoproline	130,0501397	2,75E-04	Pos.	C01877		728486	107541 HMDB03578
D-Arginine	175,11935	3,97E-04	Pos.	C00792		4050	71070 HMDB03416
L-Arginine	175,11935	0,000396729	Pos.	C00062		149262	- -
L-Phenylalanine	166,0865722	3,20E-04	Pos.	C00079		149037	- HMDB00159
GLYCINE	201,1225583	8,09E-04	Pos.	-	-	-	-
N-oleoyl tryptophan	469,3444663	0,001983643	Pos.	-		52922061	- -
N-palmitoyl tryptophan	443,3252979	0,001464844	Pos.	-		10321153	- -
L-SERINE	309,1285938	6,41E-04	Pos.	-	-	-	-
(2R,4S)-2,4-Diaminopentanoate	133,0971697	1,52588E-05	Pos.	C03943		6667	440169 -
Peptides/peptide conjugates							
Homocarnosine	241,1292891	0,000228882	Pos.	C00884		675516	89235 HMDB00745
Sugars and sugar conjugates							
(S)-malyl alpha-D-glucosaminide	296,0962432	0,001312256	Pos.	-		57339274	- -
4-AMINO-2-DEOXY-2,3-DEHYDRO-N-NEURAMINIC ACID	291,1184238	2,14E-04	Pos.	-	-	-	-
D-Ribose 1-diphosphate	292,9844405	0,00112915	Neg.	C06248		8487	440963 -
Fructose 1,6-bisphosphate	322,9934885	3,66E-04	Neg.	C00354; C05378	-		445555 HMDB01058
Cyclic alcohols, aliphatic and aromatic compounds							
5,6-Dihydroxyindole	150,0550344	7,63E-05	Pos.	C05578		691592	114683 HMDB04058
Indole-3-acetaldehyde	160,0760248	0,000320435	Pos.	C00637		765782	- HMDB01190
Pentaglutamyl folate	958,3170378	3,05E-04	Pos.	-	-		73374 HMDB06487
Thiamin triphosphate	502,9964309	2,14E-04	Neg.	C03028		161947	18989 HMDB01512
Steroids, prenols and eicosanoids							
12R-HETrE	323,2573853	0,000671387	Pos.	-		5283149	- -
15-HETrE	323,2573853	6,71E-04	Pos.	-	44322431;5283145	-	HMDB05045
all-trans-4-oxoretinoic acid	315,1954569	0	Pos.	-		6437063	- HMDB06285
Coenzyme Q4	455,3141991	0,001342773	Pos.	-		5283545	- HMDB06709
PGE2alpha dimethyl amine	368,3170876	0,001220703	Pos.	-		5283101	- -
11-dehydro-2,3-dinor-TXB2	339,1797467	0,001586914	Neg.	-		35024530	- -
PGE2alpha dimethyl amine	368,3170876	0,001220703	Pos.	-		5283101	- -
PGF2alpha-11-acetate	395,2429838	9,46E-04	Neg.	-		5283088	- -
Others							
Dihydroxyacetone Phosphate Acyl Ester	196,9850873	5,65E-04	Neg.	C03372	-		6857386 HMDB11750
(R)(-)-Allantoin	159,0509586	0,000305176	Pos.	C02348		5396	439713 -
molybdenum cofactor (sulfide)	537,8588862	0,00177002	Pos.	-	-	-	-

SUPPLEMENTARY TABLE 5. Unassigned metabolites IDs found in plasma and cell culture media ELVs. This table is not included due to its extension.

SUPPLEMENTARY TABLE 6. Significant *m/z* detected in ELVs isolated from PANC1 cell culture media.

SUPPLEMENTARY TABLE 6.
Significant m/z detected in ELVs isolated from PANC1 cell culture media.

m/z	FC	p-v	Mode
355,166	2,214	0.004	Negative
647,089	2,389	0.005	Negative
835,386	2,036	0.008	Negative
579,103	2,089	0.014	Negative
585,121	2,035	0.019	Negative
549,071	2,028	0.020	Negative
579,104	2,136	0.021	Negative
449,149	2,166	0.023	Negative
374,340	2,053	0.024	Negative
449,149	2,176	0.024	Negative
517,136	2,355	0.024	Negative
517,136	2,261	0.024	Negative
807,371	2,416	0.024	Negative
549,071	2,404	0.029	Negative
355,166	2,661	0.031	Negative
385,177	2,669	0.032	Negative
508,279	6,874	0.033	Negative
594,068	2,183	0.037	Negative
242,175	2,876	0.045	Negative
385,177	2,958	0.047	Negative
228,438	2,525	0,004	Positive
451,166	2,202	0,004	Positive
550,084	2,285	0,004	Positive
244,264	2,622	0,005	Positive
945,303	3,361	0,005	Positive
459,349	7,492	0,005	Positive
395,103	2,256	0,006	Positive
939,276	3,877	0,006	Positive
175,016	2,739	0,007	Positive
507,085	2,198	0,008	Positive
509,081	2,404	0,009	Positive
961,263	4,369	0,009	Positive
227,048	2,741	0,009	Positive
339,040	2,517	0,009	Positive
275,069	2,874	0,012	Positive
302,266	2,504	0,014	Positive
647,544	3,681	0,016	Positive
163,077	2,221	0,019	Positive
395,733	2,046	0,023	Positive
149,095	2,295	0,024	Positive
344,254	2,613	0,025	Positive
901,462	2,789	0,025	Positive
275,068	2,079	0,025	Positive
272,295	2,719	0,025	Positive
149,061	2,873	0,026	Positive
901,330	3,977	0,026	Positive
180,173	3,148	0,026	Positive
377,091	2,287	0,027	Positive
365,137	2,464	0,027	Positive
243,210	2,012	0,028	Positive
337,106	3,059	0,029	Positive
954,242	9,699	0,029	Positive
921,465	2,041	0,035	Positive
647,559	2,097	0,038	Positive
580,455	3,801	0,042	Positive
1051,539	2,644	0,048	Positive

CHAPTER II

Integrative Analysis of the Lipidome and Metabolome of Biofuids and Extracellular Vesicles Yields Biosignatures to Diagnose Endometrial Cancer

Manuscript in preparation

INTEGRATIVE ANALYSIS OF THE LIPIDOME AND METABOLOME OF BIOFLUIDS AND EXTRACELLULAR VESICLES YIELDS BIOSIGNATURES TO DIAGNOSE ENDOMETRIAL CANCER

ABSTRACT

Endometrial cancer (EC) is the fourth most common cancer in women in developed countries. Although pioneering efforts have been made into discovering new diagnostic biomarkers for EC, none of the candidates have thus far met the criteria for specificity and sensitivity that would augment translation for potential clinical use. In order to overcome this problem, we sought to identify a panel of diagnostic markers by using a multi-pronged approach; this included analysis of 4 different human biofluids: plasma, uterine aspirates (UAs) and extracellular vesicles (EVs) derived from plasma or from UA, in order to develop small molecule biosignatures. Metabolomic profiling provides insights into the changes associated with a diseased state and can be useful for developing novel non-invasive diagnostic tests. Thus, in this study, we initially used an untargeted profiling approach (discovery phase) using a clinical cohort that included 127 patient samples that resulted in the identification of a set of 77 candidate metabolite markers that were verified by MS/MS. These data were used to build classifiers and biomarker efficiency was evaluated using ROC analysis. We developed 4 discrete biomarker panels, one for each type of matrix with potential for clinical classification of EC. The best predictive value was obtained for a 13 metabolite panel verified in plasma samples with an AUC of 0.986 that could be developed as a minimally invasive biomarker test following large scale validations.

To our knowledge, this is the first report that integrates biomarker information derived from the combination of 4 different matrices for EC diagnosis. Results from this study lay the foundation for future large scale validation studies for delineating robust biomarker panels with clinical potential. These efforts will improve the early diagnosis of EC, thus improving survival of patients.

INTRODUCTION

Endometrial cancer (EC) is the fourth most common cancer in women in developed countries and the most common malignancy within the female reproductive system and its incidence and mortality rate are still increasing [92]. Current methods of diagnosis include minimally-invasive diagnostic tests (pipelle biopsies, which have a 22% of failure associated, and/or more risky and invasive approaches (such as curettage, biopsy and hysteroscopy) [93]. In the last decade, many studies sought to identify EC diagnostic markers to achieve a non-invasive diagnosis; however those panels lacked specificity and sensitivity [34]. Hence, the identification of novel sensitive and specific biomarkers for a non-invasive diagnosis of EC is still required.

Tumor biomarkers are described as quantifiable molecules produced and released by cells representing a pathological condition in a subject compared to normal physiological state [94]. For diagnostic purposes, the detection of biomarkers in human biofluids is highly recommended since they are readily accessible and can be generally collected in larger quantities compared to tissues [95].

Our approach for identifying EC biomarkers was based in a comparison between EC patients and control subjects. The most common biofluid used in discovery studies is plasma, since it is a sample that permits to achieve an easy and non-invasive diagnosis. In our study, a part from plasma, we included 3 additional matrices. Thus, the analysis was performed using plasma, uterine aspirate (UA), and extracellular vesicles (EVs) isolated from plasma (P EVs) and from UA (UA EVs). On one hand, UAs are endometrial biopsies obtained by aspiration from the uterine cavity [96]. Since they are in direct contact with the uterus, they are considered as a proximal biofluid. UA reflect the molecular status of the growing endometrial tumor and thus can be used as an information rich matrix for the detection and identification of specific and sensitive biomarkers. They contain epithelial secretions of the uterus, compounds filtered selectively and endometrial and blood cells. Some research has been done using UAs as an informative source for diagnostic purposes [38,96–99] but this is the first attempt to characterize the metabolome of UAs. On the other hand,

EVs are heterogeneous nanovesicles (from 30 to 10,000 nm) limited by lipid membranes that originate from the plasma membrane or from the multivesicular bodies [100]. They were first described by Pan and Johnstone in 1983 [101] and they are known to play a fundamental role in cellular communication [102,103] and in cancer [102,104,105]. EVs can be isolated from many biofluids [106] and can be used as an informative sample for biomarker discovery studies [107–109], among others.

Metabolomics and ultra high performance liquid chromatography mass spectrometry (UHPLC-MS) analysis are the method of choice for this study since they offer a good promise as an effective and quickly diagnostic technique in several diseases, including cancer [110]. This new field is growing as one of the more efficacious in the discovery for diagnostic biomarkers [111,112]. MS-based metabolomic research allows the separation and identification of an extensive number of endogenous metabolites [113] and has been used to characterize changes in low molecular weight compounds associated to genesis and progression of human pathologies [114].

Since clinical biofluids are routinely collected in an easy, non-invasive, and inexpensive way, development of biofluid based clinical assays is a pragmatic approach. In our study, we present a standard strategy for biomarker development [115] including a discovery phase by using UHPLC-MS, and a tandem mass spectrometry (MS/MS) verification phase in order to develop a biomarker panel allowing the early and non-invasive diagnostic of EC. A validation phase by multiple reaction monitoring (MRM) with an independent cohort of 121 patients to evaluate the specificity of the panel is in progress.

MATERIAL AND METHODS

Study population

All patients included in the study underwent surgery at Vall Hebron University Hospital (VHUH) in Barcelona, Spain, and signed an informed consent. The Clinical Research ethics Committee at VHUH approved the study (approval number: PR_AMI_50-2012). The final diagnosis of each patient was determined

in the VHUH Pathology Department. A description of the samples in the study is detailed in Table 1.

Samples used for the discovery and MSMS verification:

Diagnosis	F.I.G.O.	Type I/II	Pre/Post-menopausal	Number of samples used per matrix: Plasma	Uterine aspirate	Plasma Evs	Uterine aspirate Evs
Control (14 patients)			Post	10	12	13	14
Endometrial adenosquamous carcinoma (12 patients)	IA	I	Post	11	10	12	10
Endometrial adenosquamous carcinoma (10 patients)	IB	I	Post	10	10	7	8

Samples for MRM validations:

Diagnosis	F.I.G.O.	Type I/II	Pre/Post-menopausal	Matrix	Number of patients
Control			Post	Plasma	41
Endometrial adenosquamous carcinoma	IA	I	Post	Plasma	38
Endometrial adenosquamous carcinoma	IB	I	Post	Plasma	18
Endometrial adenosquamous carcinoma	II	I	Post	Plasma	16
Endometrial adenosquamous carcinoma	III	I	Post	Plasma	4
Endometrial adenosquamous carcinoma	IV	I	Post	Plasma	4

Table 1. Clinical information of the study cohort.

Sample collection and processing

Blood and UA were collected under sterile conditions in the surgery room and rapidly transported to the laboratory to be processed.

UA samples were collected by aspiration using an endometrial suction curette (Cat# 4164. Gynetics Medical Products, Lommel, Belgium). PBS1X was added to each sample in a 1:1 ratio. UA were vortexed and centrifuged for 20 min at 2.500 g at 4 ºC. The supernatant was then transferred to a clean 1.5 ml tube and frozen at -80 ºC.

Blood samples were collected in tubes containing EDTA (Cat# 367525. Pulmolab, Los Angeles, California). After inverting the sample, protease inhibitors in a ratio 1:200 were added at room temperature. The samples were centrifuged for 15 min and the supernatant was placed in a new tube and centrifuged again. The final supernatant or plasma was properly labeled and stored at -80 ºC.

EVs isolation

EVs were isolated from 400 µL of UA supernatant and 250 µL of plasma since in previous studies we already optimized the minimum starting volume of

biofluid needed to obtain quality data [109]. For EVs purification we followed a modification of the protocol published by Thery et al [116]. After centrifuging the fluids at 16,500 g at 4 ºC for 20 min, the supernatant was transferred to an ultracentrifuge tube (Cat#03141. Thermo Scientific, NY, USA) and centrifuged again at 100,000 g at 4 ºC for 2 h. The pellet obtained was washed with Tris Buffered Saline-Calcium and centrifuged again at 100,000 g at 4 ºC for 1 h. Finally the pellet (containing the EVs) was collected in 60 µL PBS 1X buffer and stored at -80 ºC for further analysis.

Nanoparticle tracking analysis

EVs were firstly analyzed using a Nanoparticle Tracking Analysis (NTA, Nanosight LM10, Malvern Instruments, UK) according to the standard protocol provided by the manufacturer. EVs were diluted with Milli-Q water (Milli-Q Synthesis, Merck Millipore, Massachusetts, USA) and videos were recorded for 1 min (at least 3 measurements per sample were taken). The number and size of the particles was calculated automatically based on the Brownian motion rate. The data were analyzed using the software version 2.3.

Sample preparation for the metabolome LC-MS analysis

A volume of 25 µL of plasma was added to 175 µL of a 40% ACN: 25% methanol: 35% water solution containing the internal standards (IS). Samples were vortexed and incubated for 10 min on ice, centrifuged again and the supernatant was dried under vacuum and stored and -80 ºC for the analysis.

EVs isolated from both fluids and UAs were processed following a modified protocol previously described by Sheikh et al [117] . Samples were thawed on ice and vortexed and, for metabolite extraction, 55 µL (v/v) of UA supernatant or EVs were adjusted up to 150 µL of water and placed on dry ice for 30 sec followed by incubation in a 37 ºC water bath for 90 sec (heat shock). After sonicating the samples for 30 sec, 600 µL of chilled methanol containing IS were added and the tubes were incubated on ice for 15 min. A total of 600 µL of chloroform was added to each tube followed by centrifugation at 13,000 at 4 ºC for 10 min. The supernatant was transferred to a new tube containing 600 µL of ACN and tubes were incubated overnight at -20 ºC. The tubes were centrifuged

at 13,000 at 4 ºC for 10 min and the supernatant was dried under vacuum and stored and -80ºC for the analysis. This sequential extraction method (using first aqueous, followed by semi-polar, and finally non-polar solvents) permits the extraction of a broad range of metabolites. Dried samples were resuspended in 200 µL of 50% water and 50% methanol followed by UPLC-ESI-Q-TOF-MS analysis. Solvents chosen for the analysis allows

IS: 10 µL of debrisoquine (1 mG/mL) and 50 µL of 4-nitrobenzoat (1 mG/mL) for each 10 mL of solution

Reagents for UPLC-MS Analysis

Chloroform, LC/MS-grade acetonitrile (ACN), water and methanol were purchased from Fisher Optima grade, Fisher Scientific (New Jersey, USA). PBS was purchased from Invitrogen (Carlsbad, CA, USA). Ammonium formate, debrisoquine, 4-Nitrobenzoic acid (4-NBA) were purchased from Sigma- Aldrich (St. Louis, MO, USA). High purity formic acid (99%) was purchased from Thermo-Scientific (Rockford, IL, USA).

Ultra Performance Liquid Chromatography coupled to Quadrupole Time-Of-Flight Mass Spectrometry (UPLC-QTOF-MS)

5 µL of each sample were injected onto a 1.7 µm, 2.1 x 100 mm reverse phase Acquity UPLC CSH C18 column (Waters Corp.). The mobile phase was composed by 0.1% formic acid in water (Solvent A), 0.1% formic acid in ACN (Solvent B), IPA/ACN (90:10) containing 0.1% formic acid and 10mM ammonium formate (Solvent D). The resolution was performed at a flow rate of 0.4 mL/min for 13 min. The gradient consisted of 97% Solvent A and 3% Solvent B for 0.5 min, then a ramp of curve 6 to 60% Solvent B from 0.5 minute to 4 min. From min 4 to 8 a ramp of cure 6 increases Solvent B to 98%. At min 9 Solvent B decreases to 5% and Solvent D increases to 95% and it is hold until min 10. At 11 min at ramp of curve 6 the solvent composition goes to 25% Solvent A, 25% Solvent B and 50% Solvent D followed by a calibration to starting gradient of 97% Solvent A and 3% Solvent B. The capillary voltage used was 3.2 KV and a sampling cone voltage of 30 V in negative mode and 20 V in positive mode. The desolvation gas flow was set to 750 L/h, and the

temperature was set to 350 °C. The cone gas flow was 25 L/h, and the source temperature was 120 °C. Accurate mass was maintained by introduction of LockSpray interface of Leucine Enkaphalin (556.2771 [M + H] + or 554.2615 [M - H]-). Data were acquired in TOF MS centroid mode from 50 to 1200 mass-to-charge ratio (*m/z*) in MS scanning at a rate of 0.3 seconds. MS data acquisition was performed using ESI-QTOF MS within the mass range of 50 to 1200 *m/z* in positive and negative electrospray ionization (ESI) modes on a Xevo G2 Q-TOF (Waters Corp.).

The data were pre-processed using the XCMS software [118]. Molecular identification was determined by comparing the chromatographic characteristics (*m/z*, retention time, spectra) with reported mass spectral libraries and databases: HMDB (http://www.hmdb.ca/), MMCD (http://mmcd.nmrfam.wisc.edu/), LIPID MAPS (http://www.lipidmaps.org/), KEGG (http://www.genome.jp/kegg/) and METLIN (https://metlin.scripps.edu/). A set of resulting identities were verified by MS/MS comparing with the fragmentation spectra of the authentically commercial standards using the same gradient on a SYNAPT G2-Si (Waters Corp.) and also using MassFragment software (Waters Corp.). Targeted analysis for the biomarker panel validation were performed with a triple quadrupole MS operating in the MRM mode.

Statistical analyses

A multivariate data analysis using Metaboanalyst 3.0 web tool [119] was performed for the MS data. Metabolites with an adjusted p-value of less than 0.05 were considered significant for constructing receiver operating characteristic (ROC) analysis. Metabolite abundance patters were visualized as box plots using Metaboanalyst 3.0 web tool. The area under the ROC curve (AUC) for each metabolite and for combination of metabolites were calculated to determine the specificity and sensitivity of the markers.

RESULTS

Discovery phase

To the date, most of the reported EC biomarker studies used tissues or blood as a sample matrix, however, there are no EC biomarkers implemented in clinics yet [35]. In order to expand the scope of biomarker investigation using a clinical cohort study, we performed an untargeted metabolomic discovery using UPLC-ESI-TOF-MS. A total of 36 patients were used to obtain plasma, UA, P EVs and UA EVs samples. The cohort included in the discovery phase comprised control subjects (N=14), and EC type I patients at early stages of the disease -FIGO stages IA and IB- (N=22). All patients included in the study were postmenopausal women and did not receive any pre-surgery treatment. Patient information is detailed in Table 1.

The use of plasma, UA and EVs derived from human samples for based biomarker discovery studies has been previously studied in our group [96,109]. Following the same methodology, samples were collected and EVs were isolated. Size distribution and concentration of the EVs enriched fraction were determined by NTA (Supplementary figure 1). The population of isolated EVs followed a uniform size distribution with a unique peak corresponding to a mode of 161 ± 6 nm for P EVs and 157 ± 5 nm for UA EVs.

In the discovery phase, a total of 958 features were detected for the positive ESI mode and 755 features for the negative ESI mode in the entire set of samples. Afterwards, we compared for each matrix separately, control versus EC samples (see Table 1) using an MS approach. In order to identify the differences of the metabolome composition between the two study groups (control subjects and EC patients) a multivariate analysis was performed. The number of significant features (adjusted p-value ≤ 0.05) encountered to be differential between the two groups for each matrix is shown in Figure 1, Panel A. They accounted for 683 in plasma; 612 for UA EVs; 276 for P EVs, and 234 for UA samples.

We then checked the overlap of significantly dysregulated features when comparing control to EC samples among the four matrices (Figure 1, Panel B).

Interestingly, just five common features were significantly dysregulated in the four matrices, but no possible annotation for them was found. Plasma samples presented a 25.7% of unique features, but showed an overlap of 139 significant features with P EVs, 236 common features with UA EVs and 69 with UA. Similarly, UA samples presented a big percentage (20.8%) of significant features exclusively detected in UA. UA presented an overlap of 112 significant features with UA EVs and 29 with P EVs. A total of 90 common features were found to be significantly dysregulated in tumor samples in both UA EVs and P EVs.

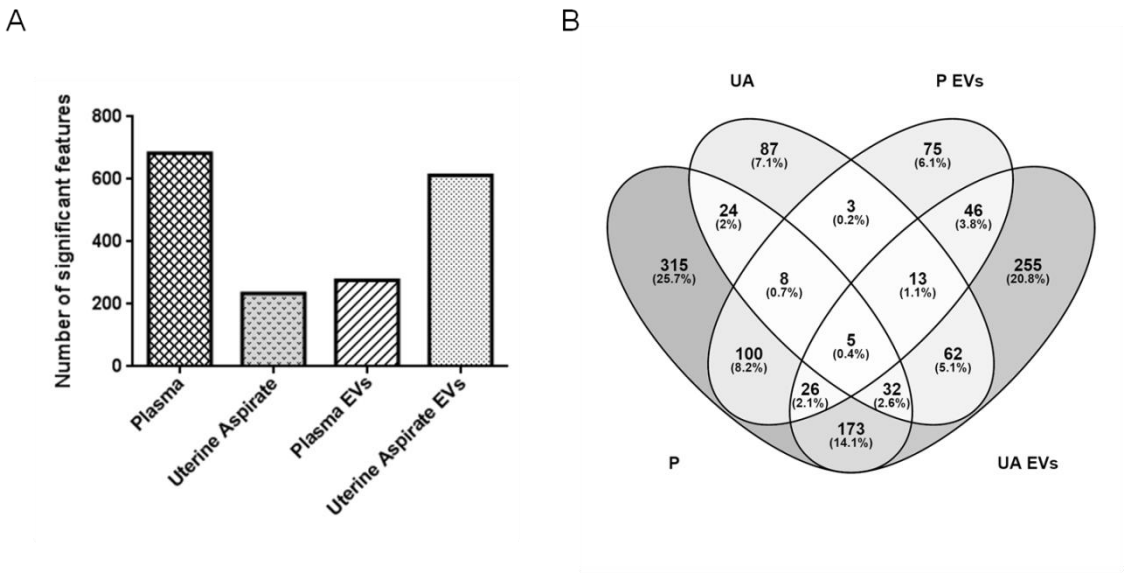


Figure 1. Number of significant features detected in the discovery phase in each human matrix when comparing the control group against the tumoral samples (Panel A) and overlap of significant features among the different matrices. Extracelular vesicles (EVs), plasma (P), uterine aspirate (UA).

The differential abundance of metabolites was plotted as a PCA showing the separation between the two study groups (Figure 2). Although a similar number of significant features were detected in plasma and in UA EVs, the biggest separation between the two groups was given for the UA EVs since the features detected in this matrix showed more robust concentration differences, in terms of p-value.

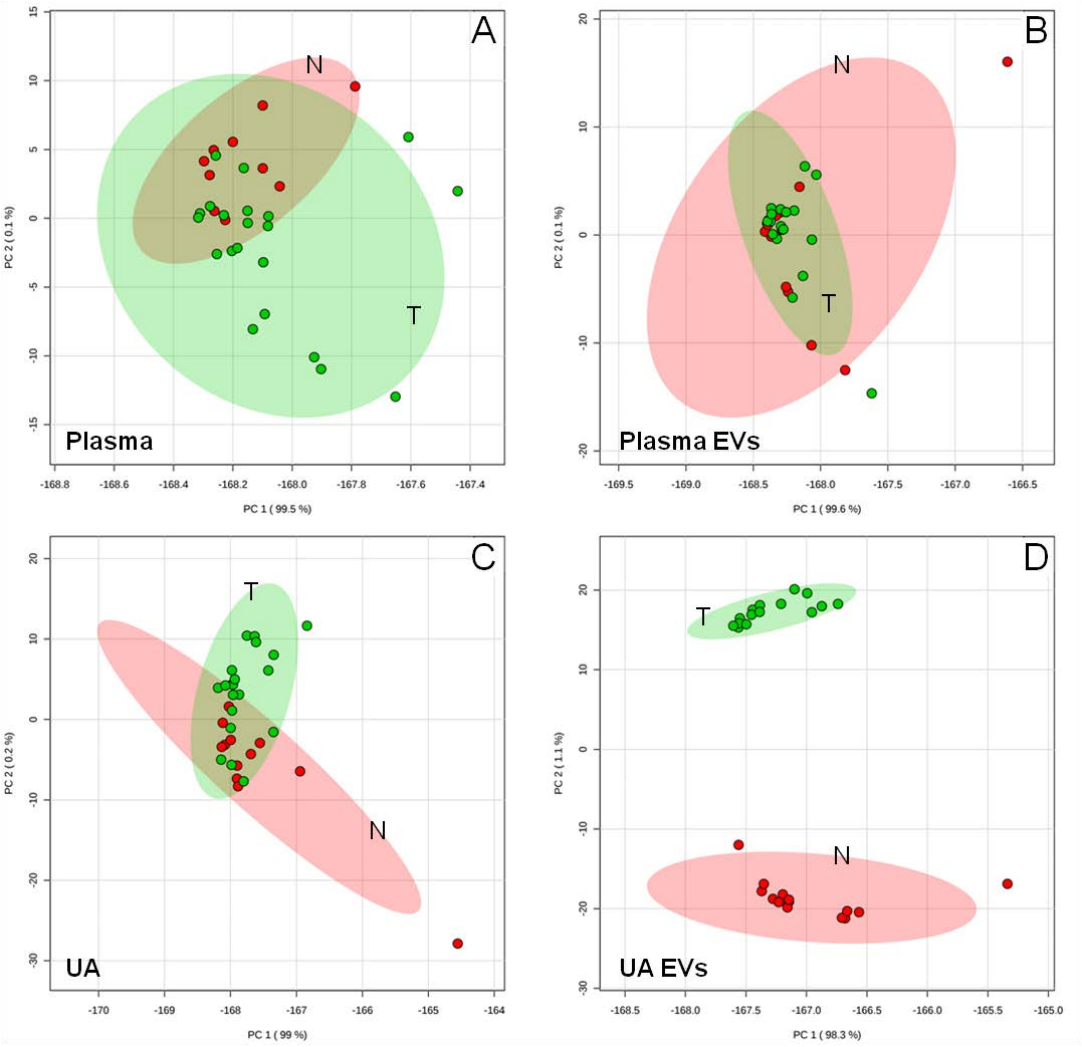


Figure 2. Principal Component Analysis (PCA) plots for the positive ESI mode showing the separation between control subjects (N) and EC patients (T) in plasma (Panel A), P EVs (Panel B), UA (Panel C) and UA EVs (Panel D).

Verification phase

The putative identity of a subset of significant features identified in UA and UA EVs was confirmed by MS/MS (Table 2). A total of 8 features were verified for the negative and 15 for the positive ESI mode. The fragmentation pattern of one of the verified features is shown as an example in Supplementary figure 2. Some interesting metabolites such as lactic acid, picolinic acid and lipoxin A4 were verified.

m/z	dppm	RT (min)	FC (T/N)	Adj. p-v	ID	Formula	Metlin ID	Matrix	Mode	Major CID fragments
283.2622	3	8.5266	1.3740	2.74E-02	Oleic acid	C18H34O2	45812	UA	Negative	137.506, 179.025, 181.028, 207.008, 247.013, 265.002
188.1284	1	2.6261	0.5738	2.91E-02	KAPA (8-amino-7-Oxononanoic acid)	C9H17NO3	3322	UA	Positive	39.871, 91.051, 105.543, 128.026, 170.065
478.2933	1	6.0954	2.1508	2.52E-02	Lyso PE (18:1)	C23H46NO7P	40778	UA Evs	Negative	78.955, 140.007, 152.991, 196.032, 214.043, 281.245, 282.249, 435.169, 460.046
483.2505	0	5.0060	0.5006	3.04E-05	PA(22:6/0:0)	C25H39O7P	82338	UA	Positive	137.097, 139.074, 309.206
526.2936	1	5.6651	0.6341	1.67E-02	PE(22:6/0:0)	C27H44NO7P	62284	UA Evs	Positive	269.14, 283.12, 311.15
863.5657	0	10.0322	2.5842	1.78E-02	PI(14:0/22:1)	C45H85O13P	80096	UA	Negative	279.233, 291.953, 339.064
864.5763	0	10.0302	2.1787	1.51E-02	PS(20:1/22:4)	C48H84NO10P	78220	UA	Negative	297.044, 510.129, 528.946, 574.994
89.0242	2	0.4841	2.2644	6.69E-03	Lactic acid	C3H6O3	116	UA	Negative	45.2708, 71.974
186.0400	4	0.4380	1.9772	4.86E-02	1-cyclopropanecarboxylic acid	C7H9NO5	C01215	UA	Negative	45.181, 56.005, 59.579, 71.014, 87.007, 102.052, 128.035
315.0809	4	0.4841	2.0523	1.86E-02	Ser-Ser-OH	C12H14N2O8	65143	UA	Positive	80.9485, 135.002, 271.040, 268.121
335.1914	3	3.8185	0.8408	1.77E-02	Ser Lys Thr	C13H26N4O6	17321	UA	Positive	69.069, 84.962, 216.181, 317.292
359.1702	3	3.4146	0.2114	1.66E-09	Pro Gly Trp	C18H22N4O4	15856	UA Evs	Positive	127.07, 155.07, 231.11, 262.39, 313.34, 341.51
418.2425	4	4.8570	0.7069	2.50E-03	Asp Lys Arg	C16H31N7O6	16445	UA	Positive	70.888, 107.085, 121.099, 232.959, 286.112, 357.202
425.2157	3	4.5664	0.8143	2.89E-02	Ser Arg Tyr	C18H28N6O6	15751	UA	Positive	82.0766, 123.041, 137.061, 157.110
436.2552	0	4.6339	0.1395	2.21E-10	Ala Ala Phe Lys	C21H33N5O5	103568	UA Evs	Positive	70.00, 99.04, 277.15, 373.90
438.1990	1	5.0229	0.4817	2.19E-05	Tyr Gln Gln	C19H27N5O7	20796	UA	Positive	99.050, 135.070, 205.085, 223.133
477.2357	0	4.1804	0.1991	1.08E-10	Tyr Ser Leu Pro	C23H34N4O7	261671	UA Evs	Negative	41.54, 107.01, 116.92, 212.98
637.3066	4	7.5089	1.5709	7.79E-03	Glu Phe Arg Trp	C31H40N8O7	129378	UA Evs	Positive	83.05, 147.118, 163.01, 239.04, 295.108, 337.02, 393.07, 469.12, 525.17, 581.23
641.3524	0	5.3379	0.1728	1.09E-10	Phe Arg Arg Tyr	C30H44N10O6	141378	UA Evs	Positive	122.07, 136.05
358.2627	2	5.0808	0.6760	3.14E-02	Lipoxin A4	C20H27D5O5	96397	UA	Positive	117.0582, 164.937, 175.152, 199.209, 204.904, 276.124, 286.789
288.2903	2	4.9141	0.4697	3.40E-04	Sphinganine	C18H39NO2	41558	UA	Positive	88.0757, 106.087, 125.071, 191.118, 209.129, 218.129, 270.279, 274.195
298.2748	2	6.0583	1.7518	4.18E-02	Sphingosine	C18H35NO2	46564	UA	Positive	55.089, 69.070, 71.086, 79.056, 81.076, 207.181, 219.210, 263.238, 281.250
122.0245	2	3.1930	1.1240	2.93E-05	Picolinic Acid	C6H5NO2	44803	UA Evs	Negative	81.53, 87.12, 105.02, 107.01

Table 2. Confirmed markers identified in UA and UA EVs by MS/MS.

In this phase, a 13% of fatty acyls/acids and conjugates, 47,8% of amino acids and peptides, a 21,7% of glicerophospholipids, and a 17,4% of organonitrogen compounds/pyridines/steroids were confirmed (Supplementary figure 3, Panel A). UA EVs presented a very high percentage of possible annotations corresponding to short peptides. Most of these peptides were not able to be proved by MS/MS and this resulted in a lower percentage of verified features for this sample type.

Heat maps of confirmed metabolites for UA and for UA EVs were generated (Figure 3). Again, as seen in PCA plots, UA EVs gives a better separation between the control and the EC group.

Moreover, ROC curves for the best panel of metabolites for each matrix were also generated. A 7-metabolite panel for UA gave an AUC of 0.933 (Figure 4, Panel A). The metabolites included in this panel are the lactic acid, sphinganine,

Asp Lys Arg, Tyr Gln Gln, PA(22:6/0:0), PI(14:0/22:1) and PS(20:1/22:4). For UA EVs we also generated ROC analysis and we obtained an AUC of 0.936 for an 8-metabolite panel (Figure 4, Panel B). The metabolites included in this panel were the Pro Gly Trp Ala, Ala Ala Phe Lys, Tyr Ser Leu Pro, Lyso PE (18:1), PE(22:6/0:0), Glu Phe Arg Trp and Phe Arg Arg Tyr.

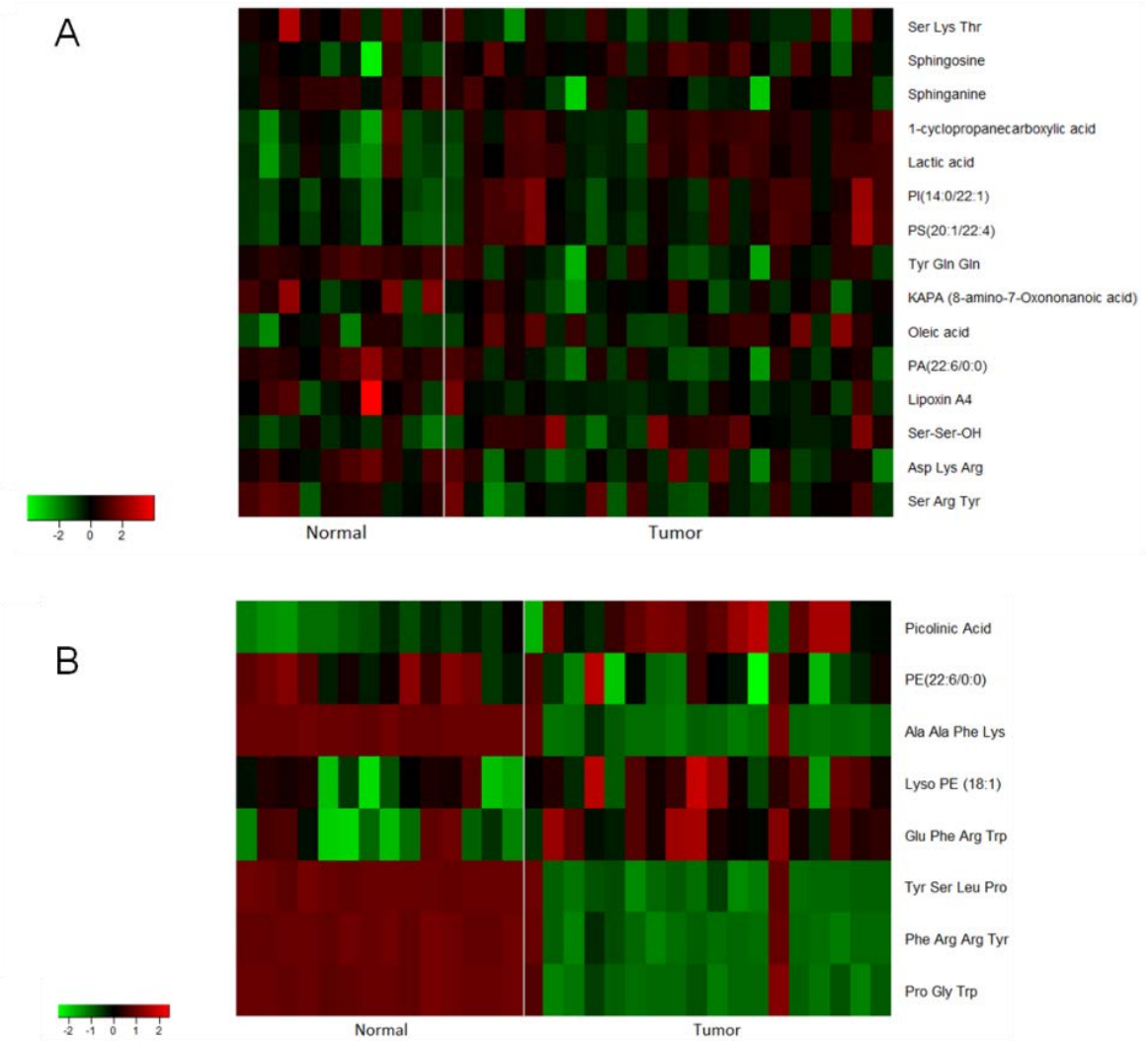


Figure 3. Heat map showing the metabolites verified by MS/MS in UA (Panel A) and UA EVs (Panel B) in metabolomic analysis. Individual scale maps for heat intensity are represented for each matrix.

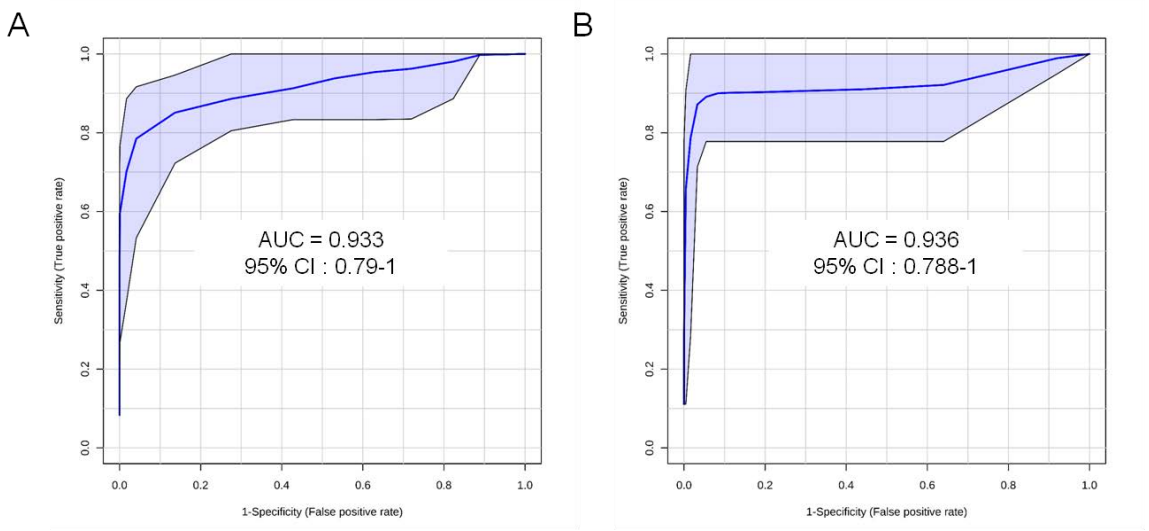


Figure 4. ROC analysis representing the predictive power of a 7-metabolite (lactic acid, sphinganine, Asp Lys Arg, Tyr Gln Gln, PA(22:6/0:0), PI(14:0/22:1) and PS(20:1/22:4)) panel for UA samples (Panel A) and a 8-metabolite (the Pro Gly Trp Ala, Ala Ala Phe Lys, Tyr Ser Leu Pro, Lyso PE (18:1), PE(22:6/0:0), Glu Phe Arg Trp and Phe Arg Arg Tyr) panel for UA Evs (Panel B) in EC.

In Table 3, we show the list of confirmed identities that were identified in plasma samples and P EVs. A total of 22 features were verified for the negative and 32 for the positive ESI mode. Many of the markers detected in P EVs to be significant in the discovery phase could not be verified since they could not be associated to any ID.

A big percentage (54.5%) of glycerophospholipids was confirmed by MS/MS and also an important number of amino acids/peptides or analogues (25.4%) (see Supplementary figure 3, Panel B).

The heat maps generated for the metabolites verified in plasma and in P EVs are plotted in Figure 5. As seen in the table and the figure, a larger number of small molecules were verified in plasma compared to P EVs.

m/z	dppm	RT (min)	FC (T/N)	Adj. p-v	ID	Formula	Metlin ID	Matrix	Mode	Major CID fragments
131.0825	0	0.4698	0.7693	5.81E-04	Ornithine	C5H12N2O2	C01602	Plasma	Negative	43.017, 115.007, 116.874
232.1542	0	3.3585	0.8015	1.74E-03	Butyryl-carnitine	C11H21NO4	964	Plasma	Positive	85.02, 144.00, 173.08
255.2305	9	7.8317	1.3430	4.09E-02	Palmitic acid	C16H32O2	187	Plasma	Negative	99.927, 117.915, 131.474, 167.906, 211.100, 215.001
280.2640	2	6.8046	5.4550	4.44E-02	Linoleamide	C18H33NO	43435	Plasma	Positive	95.086, 133.102, 219.212, 245.227, 263.238
297.2413	3	5.7240	0.5652	8.92E-05	Linoleic acid	C18H32O3	35632	Plasma	Positive	43.732, 83.083, 95.087, 111.121, 137.013, 235.429
159.1024	1	0.4834	0.7735	7.42E-04	Hydroxyoctanoic acid	C8H16O3	2988	Plasma	Negative	71.015, 85.027, 92.927, 101.9509, 115.092
258.1103	0	0.4851	0.6516	3.99E-03	Glycerophosphocholine	C8H20NO6P	370	Plasma	Positive	60.0850, 86.0955, 104.11, 124.99, 166.06, 184.07
371.3165	2	8.7095	0.8245	2.67E-03	Prostaglandin G1	C20H34O6	35989	Plasma	Positive	73.040, 208.973, 267.012, 281.070
436.2822	2	6.1165	0.7604	2.08E-02	PE(P-16:0/0:0)	C21H44NO6P	46719	Plasma	Negative	44.420, 142.9978, 223.006, 315.708
464.3131	3	6.1299	0.7342	2.31E-02	PE(O-18:1/0:0)	C23H48NO6P	46718	Plasma	Negative	325.195, 343.965, 405.332, 423.858, 449.773
468.3086	0	5.2991	0.4740	5.92E-04	LysoPC(14:0)	C22H46NO7P	40278	Plasma	Positive	86.102, 104.108, 166.065, 184.0737, 450.321
480.3089	1	5.9324	0.7206	3.14E-03	LysoPC(15:0)	C23H48NO7P	61691	Plasma	Negative	96.960, 197.138, 283.254, 299.145, 317.0023, 395.005
494.3240	0	5.5117	0.6117	6.01E-03	LysoPC(16:1)	C24H48NO7P	40287	Plasma	Positive	60.0815, 86.0971, 104.1076, 184.0740, 258.112
496.3401	0	5.9836	0.6306	1.53E-04	PC(16:0/0:0)	C24H50NO7P	182	Plasma	Positive	60.0815, 86.0970, 104.1076, 184.0738, 239.2405, 258.1104, 313.2751, 330.4574
504.3090	1	5.6568	0.6722	6.28E-03	PE(20:2/0:0)	C25H48NO7P	77687	Plasma	Negative	291.072, 365.0661, 463.0309
506.3244	1	6.1156	0.7367	1.67E-02	LysoPE(0:0/20:1)	C25H50NO7P	62270	Plasma	Negative	113.014, 140.005, 224.069, 242.084
520.3403	1	5.7123	0.5731	8.97E-04	1-Linoleoylglycerophosphocholine	C26H50NO7P	1	Plasma	Positive	60.0845, 86.0973, 104.1073, 184.0739, 240.1030, 258.1016, 322.970, 337.268, 502.3290
522.3556	0	6.1641	0.6430	3.03E-03	PC(18:1/0:0)	C26H52NO7P	184	Plasma	Positive	41.9141, 86.095, 104.106, 155.0121, 184.074, 504.3617
524.3707	0	6.7599	0.5699	5.37E-05	Lyso PC (18:0)	C26H54NO7P	61694	Plasma	Positive	60.081, 86.098, 104.11, 125.001, 166.064, 184.083, 240.10, 258.111, 285.279, 311.295, 341.308, 447.287, 506.362
528.3081	2	5.6570	0.7312	1.87E-02	PE(22:4/0:0)	C27H48NO7P	77680	P Evs	Negative	44.79, 269.226, 287.25, 315.00
538.3133	4	5.4578	0.6711	2.42E-02	PS (19:0/0:0)	C23H41N9O6	78855	Plasma	Negative	253.217, 281.172, 373.994, 478.298
542.3219	4	5.7122	0.6160	9.62E-04	PC(20:5/0:0)	C28H48NO7P	40316	Plasma	Positive	86.094, 104.1092, 146.982, 184.072, 243.782, 284.140, 337.276, 359.255, 483.255
544.3400	0	5.6961	0.7192	2.80E-02	PC(20:4/0:0)	C28H50NO7P	40314	Plasma	Positive	60.081, 86.096, 104.107, 184.074, 258.107, 361.270, 440.219, 484.247, 526.331
552.3073	4	5.6397	0.7840	2.45E-03	LysoPE(0:0/24:6)	C29H48NO7P	62313	P Evs	Negative	309.16, 311.17, 413.07, 431.00
566.3457	1	6.1145	0.7480	2.80E-02	PS(21:0/0:0)	C27H54NO9P	78853	Plasma	Negative	265.414, 281.249, 327.163, 383.981
568.3262	3	4.5971	0.7957	6.30E-04	PS(10:0/10:0)	C26H50NO10P	40807	Plasma	Positive	57.7046, 86.096, 229.182, 481.294, 551.338
730.5730	2	9.7840	1.2026	1.78E-02	PC(15:0/P-18:1)	C41H80NO7P	46707	P Evs	Positive	88.11, 99.04, 104.10, 184.0751, 207.10, 462.29
744.5534	1	9.9457	1.2057	2.95E-02	PC(15:1/18:0)	C41H80NO8P	75722	Plasma	Negative	104.559, 180.991, 222.992, 237.025, 281.251, 460.295
765.5764	3	10.0065	0.8542	1.34E-03	PA(P-20:0/22:4)	C45H81O7P	82325	Plasma	Positive	116.145, 250.146, 352.198, 361.266
786.5998	1	9.9710	1.2012	4.90-02	PC(18:0/18:2)	C44H84NO8P	39360	P Evs	Positive	88.11, 184.07, 265.20, 325.18, 341.12
804.5737	2	9.9622	1.2399	1.50E-02	PS(16:0/21:0)	C43H84NO10P	77862	Plasma	Negative	86.122, 257.241, 281.242, 480.313, 504.294, 621.023
806.5681	1	9.7755	1.2149	5.93E-03	PC(18:2/20:4)	C46H80NO8P	39390	P Evs	Positive	88.11, 184.07, 224.02, 439.86, 623.50
832.6109	1	10.0564	1.1302	4.71E-02	PC (18:0/22:6)	C48H82NO8P	60173	P Evs	Negative	146.986, 184.073, 221.085, 281.051, 355.070, 505.098, 579.114, 651.518, 775.521
862.5604	0	9.9658	1.1654	2.12E-02	PS(20:2/22:4)	C48H82NO10P	78472	P Evs	Negative	263.99,309.00, 556.94
166.0868	3	0.4974	0.6612	2.68E-04	Phenylalanine	C9H11NO2	28	Plasma	Positive	41.00, 77.03, 91.05, 93.06, 103.01
247.0930	2	0.4868	0.7784	1.17E-03	Glu Thr	C9H16N2O6	23793	Plasma	Negative	75.010, 88.040, 115.052, 145.093, 159.113
322.1881	2	3.9359	0.6835	1.03E-02	Arginyl-Phenylalanine	C15H23N5O3	85631	Plasma	Positive	65.292, 79.420, 118.914, 127.992
345.1583	4	4.1937	0.9173	4.72E-02	Trp Ala Ala	C17H22N4O4	18222	P Evs	Negative	187.00, 204.98, 230.86, 258.97, 301.07
352.1982	0	3.4914	0.8144	1.42E-02	His Pro Val	C16H25N5O4	16618	Plasma	Positive	154.003, 310.997, 334.857
389.1844	3	3.9182	0.8862	3.32E-02	Gly Pro Ala Phe	C19H26N4O5	148283	P Evs	Negative	96.02, 100.04,198.99, 373.90
408.2244	0	3.9132	0.7437	2.28E-04	Phe Asn Lys	C19H29N5O5	16043	Plasma	Positive	77.7849, 133.096, 261.134, 371.7392
414.2702	2	8.6413	0.8407	3.10E-03	Gly Ile Pro Lys	C19H35N5O5	106937	Plasma	Positive	171.000, 227.172, 285.078
441.2115	3	4.6528	0.7598	5.45E-05	Phe Gln Phe	C23H28N4O5	16088	Plasma	Positive	79.0513, 120.0925, 148.950, 203.037, 231.039
442.3023	0	9.0649	0.8000	9.68E-05	Lys Pro Val Val	C21H39N5O5	172636	Plasma	Positive	55.077, 169.055, 209.111, 396.7895
521.2601	2	4.3673	0.8790	2.17E-02	His Asn Pro Arg	C21H34N10O6	156171	P Evs	Negative	95.0504, 108.99, 129.0572, 207.14, 224.99
338.3424	2	8.6839	0.8152	2.02E-05	Docosenamide	C22H43NO	64926	Plasma	Positive	83.087, 97.065, 114.085, 163.142, 303.302, 321.307
168.0336	5	8.6398	0.8143	3.77E-03	Quinolinic acid	C7H5NO4	330	Plasma	Positive	78.950, 94.935, 122.020, 132.862, 141.915
441.2986	2	9.0652	0.8162	3.39E-04	trifluoro-25-hydroxy-27-norvitamin D3	C26H39F3O2	41965	Plasma	Positive	83.086, 105.067, 123.821, 223.917, 235.123, 261.079
482.2614	9	3.98	0.68	4,528E-06	N-acetylleukotriene E4	C25H39NO6S		Plasma	Positive	60.0828, 85.100, 86.097, 104.1076, 184.073, 185.077
518.3224	3	5.99	0.67	6.502E-03	PC(18:3/0:0)	C26H48NO7P	40303	Plasma	Positive	86.097, 104.107, 184.073, 335.250, 415.224, 459.247
523.3596	7	6.17	0.65	8.689E-03	2-Oleoylglycerophosphocholine	C26H53NO7P		Plasma	Positive	86.0922, 184.0745, 265.1021, 283.122, 339.314 , 437.810, 486.005, 492.388, 504. 334
617.1813	5	4.7	0.5	1.119E-02	Heme	C34H32FeN4O4	3680	Plasma	Positive	117.057, 497.170, 512.189, 543.568, 553.075, 557.145, 558.170, 579.276, 582.180, 591.322
381.1748	8	7.938	1.77	6.503E-04	Lys-Lys-OH	C17H26N4O6	65269	Plasma	Negative	116.928, 118.930, 127.962, 309.190, 337.182
425.3622	3	8.28	0.58	2.517E-02	Hexacosanedioic acid	C26H50O4	35992	Plasma	Negative	223.883, 321.473, 339.018, 363.360, 381.681

Table 3. Confirmed identities identified in plasma and P EVs by MS/MS.

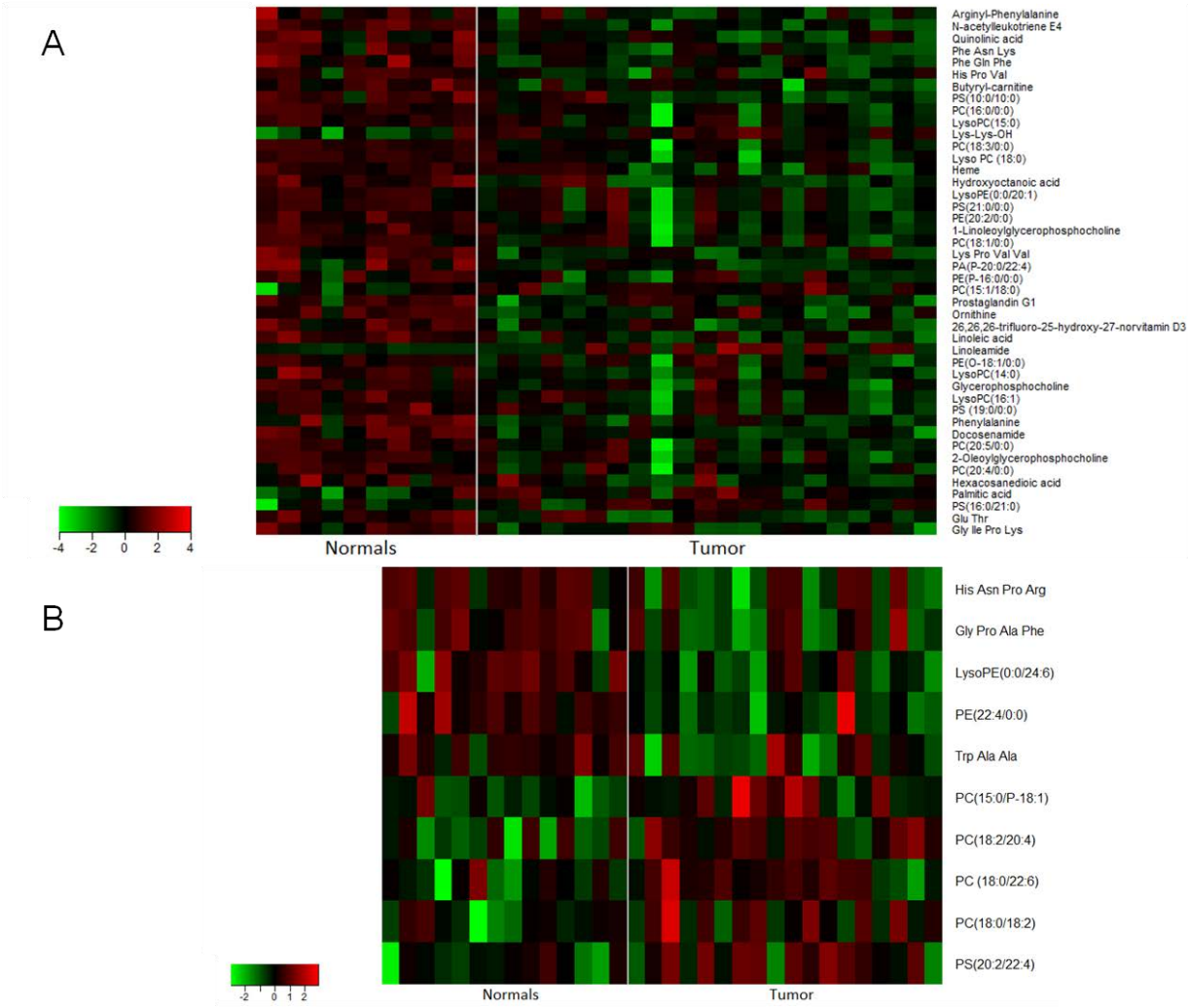


Figure 5. Heat map showing the metabolites verified by MS/MS in plasma (Panel A) and P Evs (Panel B) in metabolomic analysis. Individual scale maps for heat intensity are represented for each matrix.

Additionally, we generated ROC curves for evaluation of biomarker performance. A panel of 13 metabolites (including ornithine, hydroxyoctanoic acid, phenylalanine, linoleic acid, docosenamide, Lys-Lys-OH, Phe Asn Lys, Phe Gln Phe, 26,26,26-trifluoro-25-hydroxy-27-norvitamin D3, Lys Pro Val Val, LysoPC(14:0), N-acetylleukotriene E4 and PS(10:0/10:0)) for plasma samples resulted in an AUC of 0.986 (Figure 6 A). The results for P EVs showed lower efficacy since the best panel with 4 metabolites (His Asn Pro Arg, PE(22:4/0:0), LysoPE(0:0/24:6), PC(18:2/20:4)) exhibited an AUC of 0.767 (Figure 6 B). This could be attributed in part to the fact that several metabolites that showed

dysregulation in the plasma EV fraction were not validated using tandem mass spectrometry.

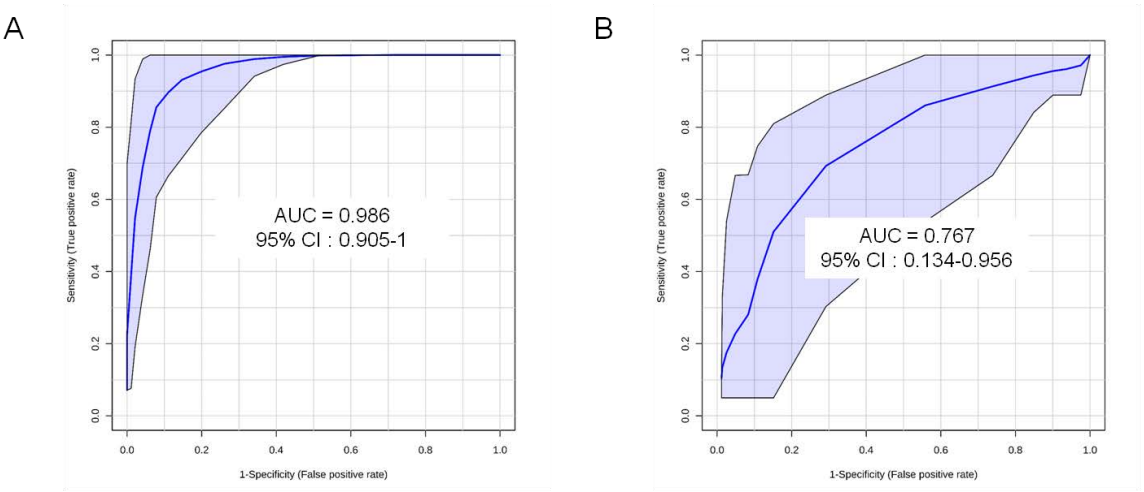


Figure 6. ROC analysis representing the predictive power of a 13-metabolite panel (ornithine, hydroxyoctanoic acid, phenylalanine, linoleic acid, docosenamide, Lys-Lys-OH, Phe Asn Lys, Phe Gln Phe, 26,26,26-trifluoro-25-hydroxy-27-norvitamin D3, Lys Pro Val Val, LysoPC(14:0), N-acetylleukotriene E4 and PS(10:0/10:0)) for plasma samples (Panel A) and a 4-metabolite panel (His Asn Pro Arg, PE(22:4/0:0), LysoPE(0:0/24:6), PC(18:2/20:4)) for P Evs (Panel B) in EC.

Regarding all the results, the AUC for plasma samples gave the best sensitivity and specificity. For each metabolite included in the panel we generated box plots expressing their concentration (in terms of normalized intensity) for the control group compared to the EC group (Figure 7).

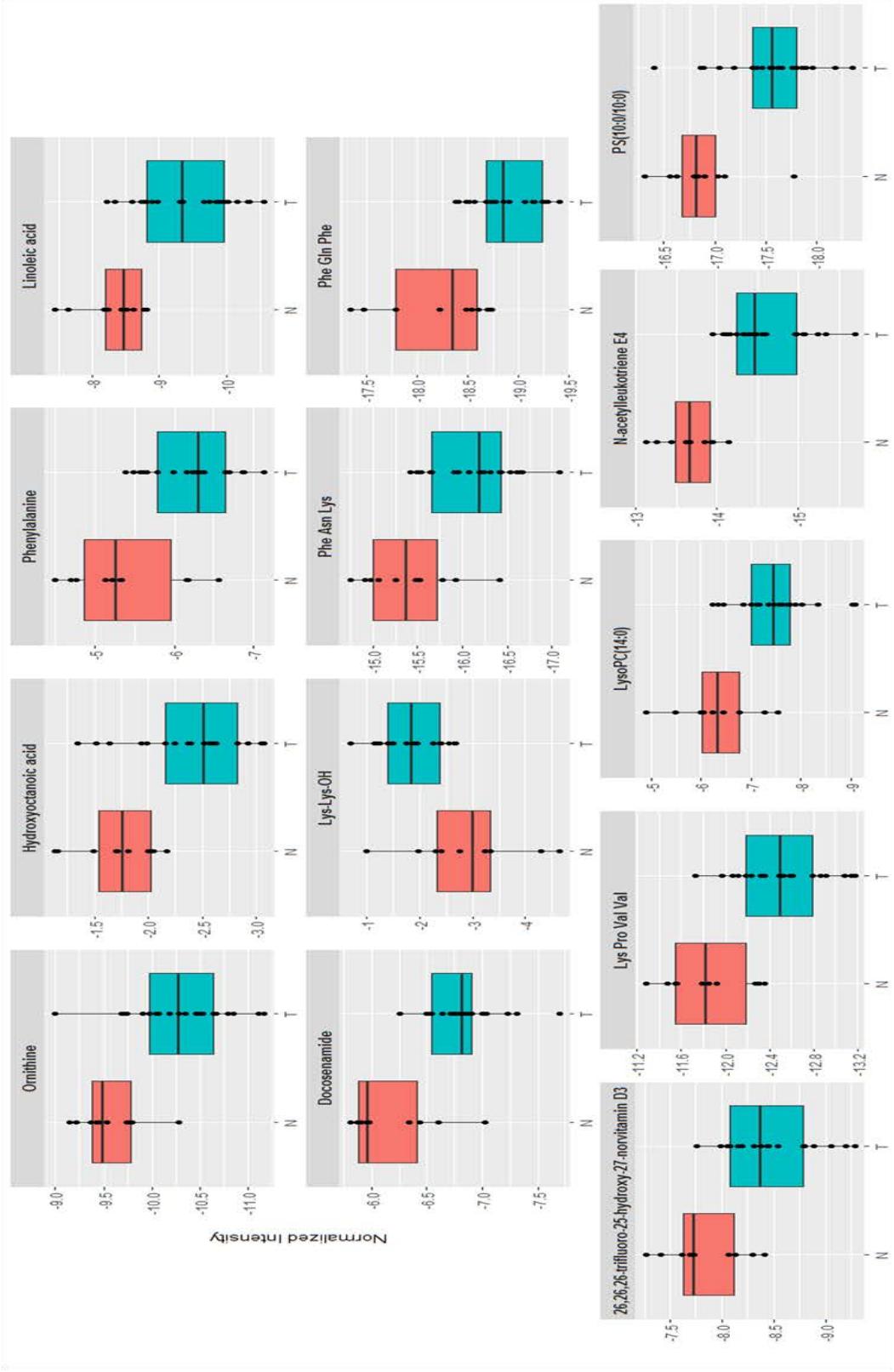


Figure 7. Box plots showing the differential concentration (normalized intensity) of 13 metabolites verified in plasma samples in the EC group (T) compared to the controls (N).

DISCUSSION

Over the last decade many investments have been made to find cancer biomarkers without much success [35]. To overcome this clinical limitation, in this study we aimed to identify a biomarker panel for EC diagnosis by analyzing the metabolome of four different human sample types: plasma, P EVs, UA, and UA EVs. Here we present for the first time a discovery combining four different human matrices in which we included 36 patients, followed by a verification phase.

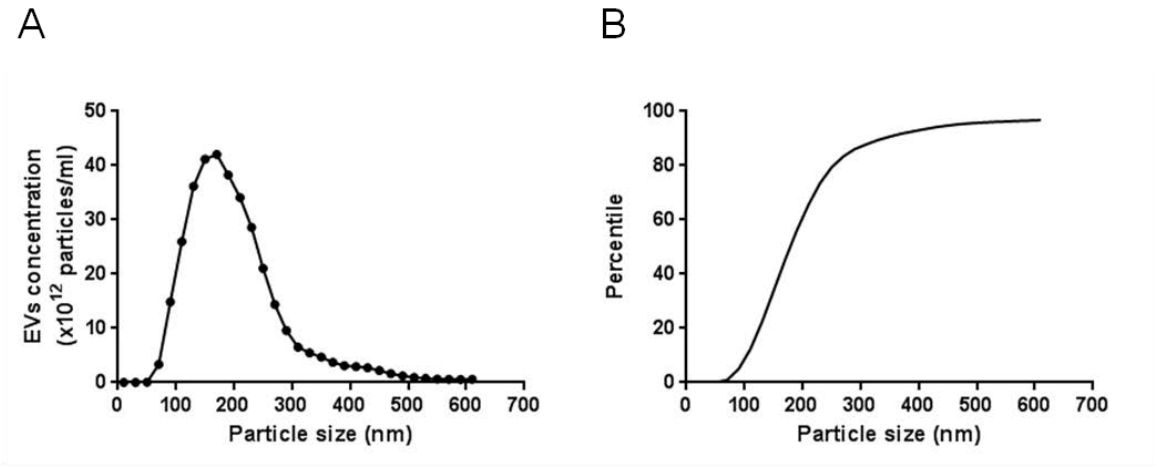
For each matrix, we described a set of features that were differentially expressed when comparing the tumor group versus the control one. Plasma and UA EVs gave the largest number of significant metabolites. As shown in PCA plots, UA EVs showed the biggest separation between the two groups of study.

Independently of the matrix used, all metabolites entered in a verification phase. We were able to confirm 77 tumor biomarkers. We observed that plasma, UA and UA EVs are excellent matrices for the discovery of specific and sensible biomarkers. The plasma based biomarker study yielded a panel with the highest specificity and sensibility; we generated a panel of 13 metabolites (ornithine, hydroxyoctanoic acid, phenylalanine, linoleic acid, docosenamide, Lys-Lys-OH, Phe Asn Lys, Phe Gln Phe, 26,26,26-trifluoro-25-hydroxy-27-norvitamin D3, Lys Pro Val Val, LysoPC(14:0), N-acetylleukotriene E4 and PS(10:0/10:0)) that gave an AUC of 0.986. Moreover, we generated a panel of 7 biomarkers (lactic acid, sphinganine, Asp Lys Arg, Tyr Gln Gln, PA(22:6/0:0), PI(14:0/22:1) and PS(20:1/22:4)) with an AUC of 0.933 for UA and a panel of 8 biomarkers (the Pro Gly Trp Ala, Ala Ala Phe Lys, Tyr Ser Leu Pro, Lyso PE (18:1), PE(22:6/0:0), Glu Phe Arg Trp and Phe Arg Arg Tyr) with an AUC of 0.936 for the UA EVs. However, the results were very disparate among matrices. Just five common features were significantly dysregulated in the four matrices, but no possible annotation for them was found. Additionally, the metabolites verified in UA and UA EVs do not align with the ones confirmed in plasma probably because the expression of some of the biomarkers was lost in circulation compared to their higher concentration in UA. In UA and UA EVs we validated a large number of peptides while in plasma and P EVs the metabolites most

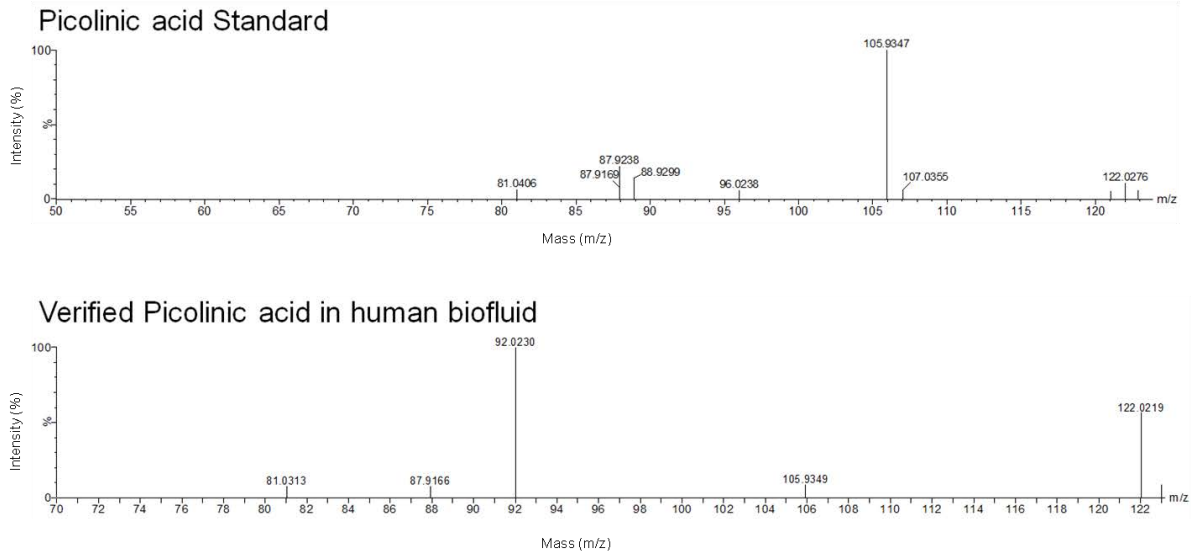
abundant were glycerophospholipids. Plasma EVs showed several markers of interest, however, many of these were not verified since they could not be associated to any ID and had little overlap with the metabolites detected and verified in plasma. Thus, although P EVs remains an untapped resource for low abundance biomarkers, the findings are difficult to translate for clinical classification since many of these metabolites are unknown. Thus, presumably the development of EC diagnostic biomarker panels needs to be initiated and validated in the same matrices.

Hence, our results show the value of using a high throughput metabolomics approach for delineating biomarker panels that can be used to diagnose EC in bodyfluids. These biomarker panels as standalone tests or in conjunction with existing clinical methods can be used for patient management with high accuracy. However, the ultimate transition of these biomarker panels for clinical use will require validation with large and diverse cohorts and prioritization and updating these panels to achieve higher specificity and sensitivity in order to minimize false positives. As a future step of our project, we will perform a validation phase using a targeted multiple reaction monitoring (MRM) approach that allows for the quantification of known metabolites even when they are present at a low concentration. For this purpose, plasma is an easier matrix to use for clinics diagnostics routine. Hence, a new cohort of patients was recruited for the validation phase and a total of 121 plasma samples will be analyzed (Table 1). We will use stable isotope labeling and multiple reaction monitoring (SID-MRM) in order to evaluate the biomarker panel specificity for EC diagnosis.

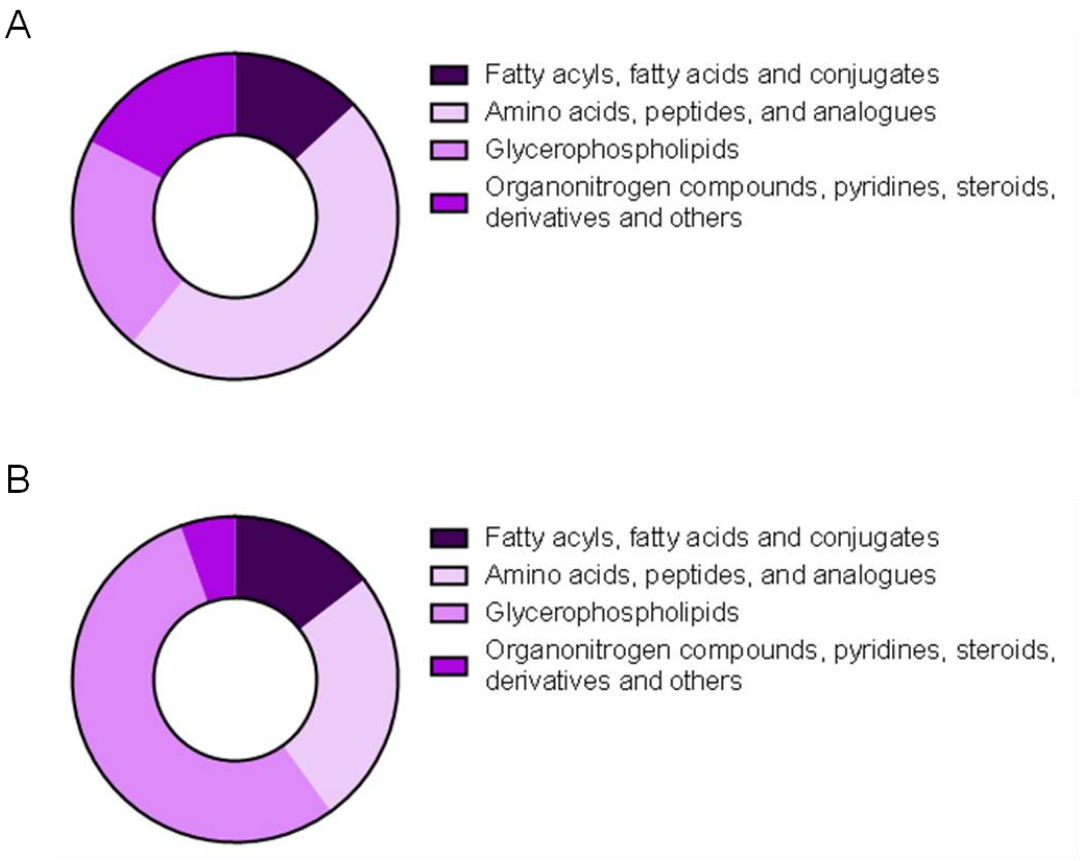
SUPPLEMENTARY FIGURES



Supplementary figure 1. Concentration (Panel A) and size distribution (Panel B) of EVs isolated from human biofluids was determined by NTA. Data are representative of three independent replicates of EVs isolated from a specific EC patient UA sample.



Supplementary figure 2. MS/MS structural verification of picolinic acid. Confirmation of identification of the metabolite picolinic acid with an *m/z* of 122.0245 in the electrospray negative ionization mode [M-H]-. Comparison to the fragmentation pattern of commercial standard of picolinic acid (40V).



Supplementary figure 3. Class distribution of the metabolites found in UA and UA EVs (Panel A) and plasma and P EVs (Panel B) that were confirmed by MS/MS.

CHAPTER III

Metabolomic and Lipidomic Profiling Identifies the Role of the RNA Editing Pathway in Endometrial Carcinogenesis

The following published manuscript comprises the results presented in this chapter:

Metabolomic and Lipidomic Profiling Identifies the Role of the RNA Editing Pathway in Endometrial Carcinogenesis

Altadill T, Dowdy TM, Gill K, Reques A, Menon SS, Moiola CP, Lopez-Gil C, Coll E, Matias-Guiu X, Cabrera S, Garcia A, Reventos J, Byers SW, Gil-Moreno A*, Cheema AK*, Colas E* (*senior authors)

Scientific Reports

doi: 10.1038/s41598-017-09169-2

SUMMARY

Background

Endometrial cancer (EC) is the most common malignancy of the female genital tract in developed countries, and both its incidence and mortality are increasing in the last years. Extensive research performed using high-throughput technologies including genomics, transcriptomics and proteomics has increased our knowledge associated with human endometrium and the onset and progression of EC. However, the metabolomic study of EC is still a significant gap.

Objective

The overall goal of this study was to discover novel metabolic pathways altered in EC carcinogenesis and progression.

Methods

For the discovery phase of the study, we used a high-resolution mass spectrometry based molecular phenotyping approach to characterize the metabolomics profile of 39 human EC and 17 healthy endometrial tissue samples. A set of identified metabolites was verified by MS/MS. Moreover, an independent cohort of 183 human EC tissues and matched controls was analyzed by immunohistochemistry (IHQ) in order to validate the results. Finally, *in vitro* studies were performed in 3 EC cell lines (HEC-1A, RL95-2 and Ishikawa) to confirm the metabolic pathway alterations in EC.

Results

We observed a dysregulation of several interesting metabolic pathways such as the kynurenine pathway, the endocannabinoids signaling pathway, the lipid metabolism, and the RNA editing pathway. The dysregulation of the RNA editing pathway was further investigated and confirmed by IHQ. We found that ADAR1 and ADAR2 were overexpressed in EC in a manner that correlates with the tumor histological type and grade. Furthermore, silencing of ADAR2 in 3 EC cell lines resulted in a decreased proliferation rate, increased apoptosis and reduced migration capabilities *in vitro*. Taken together, our results suggested

that ADAR2 functions as an oncogene in endometrial carcinogenesis and could be a potential target for improving EC treatment strategies.

Conclusions

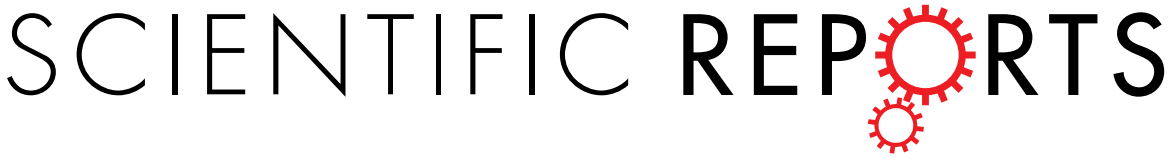
In this study, we contribute with three major findings that are expected to improve the comprehension of endometrial carcinogenesis: i) we elucidated the metabolomic and lipidomic profile of EC tissues by using ultra-performance liquid chromatography mass spectrometry (UP-LC-MS), and ii) we described metabolomic alterations that are related to cancer progression. Among all the identified alterations, iii) we profoundly characterized the role of a novel metabolic pathway altered in EC, the adenosine to inosine (A-to-I) editing pathway.

In conclusion, we believe that the methodology presented here will provide an impetus to the growing field of EC research to delineate new targets that can be used for individualized treatment of the patients.

Novel aspects

There are few papers analyzing the alterations in the lipidome and metabolome of EC subjects. Here in, we present a novel metabolic pathway altered in EC, the adenosine to inosine (A-to-I) editing pathway. For the first time, we describe that ADAR2, a main enzyme of the pathway, functions as an oncogene in endometrial carcinogenesis and could be a potential target for improving EC treatment strategies.

www.nature.com/scientificreports



OPEN

Metabolomic and Lipidomic Profiling Identifies The Role of the RNA Editing Pathway in Endometrial Carcinogenesis

Received: 3 May 2017
Accepted: 21 July 2017
Published online: 18 August 2017

Tatiana Altadill[1], Tyrone M. Dowdy[2], Kirandeep Gill[2], Armando Reques[3], Smrithi S. Menon[2], Cristian P. Moiola[1], Carlos Lopez-Gil[1], Eva Coll[1], Xavier Matias-Guiu[4], Silvia Cabrera[5], Angel Garcia[3], Jaume Reventos[1,6], Stephen W. Byers[2,7], Antonio Gil-Moreno[3,5], Amrita K. Cheema[2,7] & Eva Colas[1]

Endometrial cancer (EC) remains the most common malignancy of the genital tract among women in developed countries. Although much research has been performed at genomic, transcriptomic and proteomic level, there is still a significant gap in the metabolomic studies of EC. In order to gain insights into altered metabolic pathways in the onset and progression of EC carcinogenesis, we used high resolution mass spectrometry to characterize the metabolomic and lipidomic profile of 39 human EC and 17 healthy endometrial tissue samples. Several pathways including lipids, Kynurenine pathway, endocannabinoids signaling pathway and the RNA editing pathway were found to be dysregulated in EC. The dysregulation of the RNA editing pathway was further investigated in an independent set of 183 human EC tissues and matched controls, using orthogonal approaches. We found that ADAR2 is overexpressed in EC and that the increase in expression positively correlates with the aggressiveness of the tumor. Furthermore, silencing of ADAR2 in three EC cell lines resulted in a decreased proliferation rate, increased apoptosis, and reduced migration capabilities *in vitro*. Taken together, our results suggest that ADAR2 functions as an oncogene in endometrial carcinogenesis and could be a potential target for improving EC treatment strategies.

Endometrial cancer (EC) accounts for 7% of the new cases of female cancers in United States in 2017 and its incidence is increasing[1,2]. The most standardized classification divides EC in two different subtypes: type I or endometrioid carcinomas (EEC), which is the most frequent subtype; and type II or non-endometrioid carcinomas (NEEC), which are more aggressive tumors. Among NEEC subtypes, serous is the most prominent histology[3,4]. Moreover, EC tumors are classified according to the extent of tumor dissemination (International Federation of Gynecology and Obstetrics or FIGO staging) and histological grade[1]. About 20% of patients are diagnosed at an advanced stage and/or at a high histological tumor grade and have a low 5-year survival rate associated[5,6]. As such, availability of biomarkers for disease stratification and a thorough understanding of biochemical perturbations that underscore EC progression are likely to improve clinical outcomes.

Extensive research performed using high-throughput technologies including genomics, transcriptomics and proteomics has augmented characterization of molecular changes at different levels of cellular expression that are

[1]Biomedical Research Group in Gynecology, Vall Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, CIBERONC, Barcelona, Spain. [2]Department of Oncology, Georgetown University Medical Center, Washington D.C., USA. [3]Pathology Department, Vall Hebron University Hospital, Barcelona, Spain. [4]Pathological Oncology Group and Pathology Department, University Hospital Arnau de Vilanova, and University Hospital Bellvitge, IRBLLEIDA and Idibell, University of Lleida, CIBERONC, Lleida, Spain. [5]Gynecological Oncology Department, Vall Hebron University Hospital, Barcelona, Spain. [6]Basic Sciences Department, International University of Catalonia, CIBERONC, Barcelona, Spain. [7]Department of Biochemistry Molecular and Cellular Biology, Georgetown-Lombardi Comprehensive Cancer Center, Washington, D.C., United States. Amrita K. Cheema and Eva Colas contributed equally to this work. Correspondence and requests for materials should be addressed to E.C. (email: eva.colas@vhir.org)

Diagnosis	Number of samples	FIGO stage	Type I/II	Pre/Postmenopausal	Phase
Control tissue	17			Post	Discovery and verification
EC tissue	10	IA	I	Post	Discovery and verification
	9	IB	I	Post	Discovery and verification
	10	II	I	Post	Discovery and verification
	10	III	I	Post	Discovery and verification

Table 1. Clinical and histopathological information of the patients included in this study.

specifically associated with human endometrium[7] and onset and progression of EC[8,9]. Metabolomics defines the end point of cellular processes and hence provides a readout of current physiological status of the system[10]. Thus, metabolomics, lipidomics and glycomics have emerged as promising tools for clinical and translational research[11]. Rapid advancement of metabolomics technologies such as ultra-performance liquid chromatography mass spectrometry (UPLC-MS), enable comprehensive interrogation of the human metabolome and lipidome[12–15]. Bioinformatics analyses of these data is likely to augment the discovery of new clinical and pharmacological targets[16].

It is known that changes at the transcriptomic levels generate a new source of complexity that promotes initiation and progression of cancer and other diseases[17–19]. The most common RNA editing events are mediated post-transcriptionally by the Adenosine deaminases acting on RNA (ADAR) family of enzymes[20]. The ADAR gene family catalyzes the deamination of adenosine that is converted to an inosine creating a dysregulation of the adenosine/inosine (A/I) ratio in the cell. Adenosine to inosine (A-to-I) editing can lead to amino acid recoding events[21] since inosine is recognized as a guanosine. This family includes three enzymes, ADAR1 (UniProtKB P55265), ADAR2 (UniProtKB P78563) and ADAR3 (UniProtKB Q9NS39); also known as ADAR, ADARB1 and ADARB2, respectively[22]. ADAR1 and ADAR2 are ubiquitously expressed and, differently that the brain specific ADAR3, they show catalytic activity[23].

The overall goal of this study was to identify altered metabolic pathways in EC. Characterization of the metabolome and lipidome of EC tumors was performed using a high resolution mass spectrometry approach in conjunction with UPLC-MS. Pathway validation was performed with an independent set of samples that for the first time, yielded insights into dysregulation of the RNA editing pathway in endometrial tumors. Alterations of the RNA editing pathway correlated with high histological grade and EC serous subtypes that are indicators of poor prognosis in EC. Finally, the expression of ADAR editing enzymes was modulated *in vitro* to confirm an important oncogenic role of this pathway in EC cell proliferation, apoptosis and migration. These results are instructive of the role of this pathway in EC tumor progression.

Results

Untargeted metabolomics profiling of EC human tissue samples. To better understand alterations in metabolic pathways occurring in EC, we performed metabolomics/lipidomics untargeted discovery using UPLC-ESI-TOF-MS of tumor and non-tumoral tissue samples. The sample set included a total of 56 samples: 39 EEC tumors from different FIGO stages, including 10 stages IA, 9 stages IB, 10 stages II and 10 stages III; and 17 benign endometrial tissues (see patient details in Table 1). All patients included in the study were postmenopausal women and did not receive any treatment before surgery. Pre-processing of TOFMS data yielded a total of 8,146 features in the positive and 7,558 in the negative electrospray ionization mode, respectively.

In order to define a generic metabolomic profile of EC, initially, we combined all tumor samples from EEC patients in one group (n = 39) and compared them against matched control tissues (n = 17). Inherent differences in metabolomic profiles were visualized using descriptive Principal Component Analysis (PCA) plot that showed clear separation between tumors and controls (Fig. 1A). Subsequently, t-statistics was used to select 80 metabolites that showed significant variation (adjusted p-value < 0.05) between the two study groups and fold change (FC) values over 2 and below 0.5 (Supplementary Table 1). We confirmed the putative identity of a subset of 42 metabolites using tandem mass spectrometry (MS/MS) (Table 2). An example of the fragmentation pattern of two verified metabolites is shown in Supplementary Figures 1 and 2. We found a significant dysregulation in the lipid metabolism, with an important number of glycerophosphocholines (PCs), phosphatidylserine (PSs), phosphatidylethanolamines (PEs), phosphatidylinositols (PIs), and phosphatidylglycerol (PGs) that were upregulated in the endometrial tumor tissue (detailed in Table 2) including PC (14:0/18:2), PC (16:0/20:4), PC (16:0/20:5), PC (16:0/22:6), PC (18:0/20:2), PC (18:1/14:0), and PC (18:1/22:6). Moreover, a total of 9 PEs (PE (16:0/22:6), PE (16:1/P-18:1), PE (18:0/0:0), PE (18:0/18:3), PE (18:1/22:6), PE (18:1/16:0), PE (18:1/18:1), PE (18:4/P-18:1) and PE (P-16:0/0:0)) and 4 PIs (PI (14:0/22:1), PI (16:0/18:1), PI (16:0/22:3) and PI (18:1/18:1)) showed significant changes in the relative abundance in EC tumors as compared to the matched controls. Metabolites such as linoleic acid, 3-Deoxyvitamin D3, UDP-N-acetyl-D-galactosamine and 1-Palmitoyl-2-linoleoyl PE were observed to be also upregulated whereas peptide Glu Phe Arg Trp, some amides (palmic amide, stearamide and oleamide), PA (18:0/18:1), PE (20:1/22:6), PE (22:6/P-18:1), PG (19:0/22:4) and other important metabolites such inosine and picolinic acid showed lower abundance in the tumor tissue.

Few studies have used high-throughput approaches to study the changes in metabolomic and lipidomic profiles that underscore EC progression. Hence, in order to understand the metabolic phenotype associated with EC development, we interrogated profile differences in 29 tumor tissues restricted to the uterine cavity (FIGO stages I and II) compared to 10 tumors showing lymph node dissemination (FIGO stage III). We found a set of features to be significantly dysregulated (adjusted p-value < 0.05) among the different FIGO stages (Supplementary Table 2).

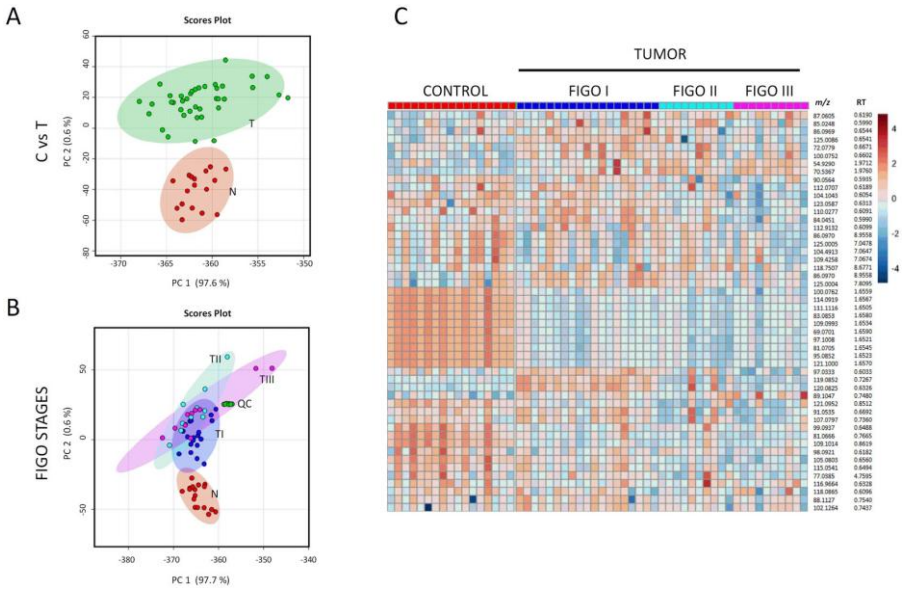


Figure 1. Multivariate analysis showing metabolic profiles in EC. Principal Component Analysis (PCA) plots showing separation between EC tissue samples (T, in green) and control tissue samples (N, in red) (Panel A) and separation between different EC FIGO stages (stage 1: TI, stage 2: TII, stage 3: TIII) (Panel B) for the positive MS ionization mode. X-axis shows interclass separation and Y-axis illustrates the intra-class variability. Panel C. Heat map of ion rankings, corresponding to their relative concentrations (intensity) for positive MS ionization mode, in the same cohort of subjects. Each row represents a unique feature with a specific *m/z* and RT while each column represents a unique subject. (*m/z*: mass to charge ratio; RT: retention time).

The PCA plots showing the segregation of the groups and the heat map showing the expression of several features in normal endometrium and in different FIGO stages of the tumor are represented in Fig. 1B and C. A subset of seven metabolites was verified by MS/MS (Table 3). The dysregulation observed in lipids (two PCs and three PEs) in tumor tissues compared to controls was significant along tumor progression. Additionally, we observed that arachidonic acid and UDP-N-acetyl-D-galactosamine are important contributors in the tumor progression, as they appeared to be upregulated in advanced compared to early FIGO stages (Supplementary Figure 3).

Increased expression of ADAR family of enzymes in human EC tumors. Among the metabolites identified in EC tissues, we were particularly interested in the dysregulation of nucleoside inosine since the relative abundance of this metabolite was significantly higher in EC tumors. To our knowledge, the underlying impact of alterations of endogenous levels of inosine has never been investigated in EC. Dysregulated levels of adenosine and inosine (A/I ratio) can be attributed to the modulation of the A-to-I editing pathway[24]. Consequently, we studied the status of this pathway in EC by analyzing the expression level of members of the A-to-I editing enzymes family (ADAR1 and ADAR2) in three independent sets of samples by immunohistochemistry (IHQ) (Table 4). The first set included the evaluation of 20 EEC samples and their corresponding paired healthy tissues; the second set included 36 EEC tumors diagnosed at different histological grades (low grade, n = 12; intermediate grade, n = 13; high grade, n = 11); and the third set allowed to differentiate between histological subtypes, and so, 78 EEC and 29 NEEC tissue samples were included.

Although ADAR1 and ADAR2 staining was clearly localized in the nucleus of epithelial, stromal, and endothelial cells, we specifically analyzed the staining of the epithelial tumor cells (Fig. 2B,D,F and Supplementary Figure 4B). Interestingly, no staining was observed in cells undergoing mitosis (Supplementary Figure 4A). Our results demonstrated that ADAR1 and ADAR2 were both significantly increased in tumor samples compared to healthy endometrial tissue (Fig. 2A), confirming that the RNA editing pathway is activated in EC carcinogenesis. ROC analysis for ADAR1 and ADAR2 expression yielded an AUC of 0.79 and 0.90 respectively, emphasizing the differences observed between control and EC samples (Supplementary Figure 5). Moreover, ADAR expression increased progressively with the presence of poor prognostic factors, such as tumors presenting high grade or a NEEC histology (Fig. 2C and E). These data suggested for the first time an activation of the RNA editing pathway in patients with EC, which positively correlates with aggressive disease, resulting in a dysregulation of the nucleosides/nucleotides balance in the tumor tissue.

Inhibition of ADAR2 reduces viability, increases apoptosis and reduces migration capabilities in human EC cell lines. Next, in order to determine the possible role of the RNA editing pathway in EC

Metabolite	*m/z*	Mass error (ppm)	RT (min)	In EC compared to Control	FC (T/C)	p-v	FDR	AUC	Formula	Ionization mode	Major CID fragments
Picolinic acid	122.02446	2	0.6499	↓	0.1834	3.07E-03	3.47E-03	0.8275	C6H5NO2	Negative	107.01, 105.02, 87.12, 81.53
Linoleic acid	279.2320	3	6.9054	↑	3.2817	2.33E-02	2.33E-02	0.6772	C18H32O2	Negative	240.99, 170.99, 96.96
Vaccenic acid	281.2482	1	7.3972	↓	0.12817	1.65E-07	4.73E-07	0.6352	C18H34O2	Negative	94.99, 96.96, 100.99, 126.90
Arachidonic Acid (peroxide free)	303.2325	1	6.8246	↓	0.29314	9,94E-04	1.30E-03	0.6439	C20H32O2	Negative	259.25, 216.99, 205.19
5,8,11-eicosatrienoic acid	305.2481	1	7.1599	↑	19.139	2.28E-05	4.09E-05	0.6189	C20H34O2	Negative	216.99, 287.00, 216.98, 1165.93, 78.98
PE(P-16:0/0:0)	436.2827	1	5.8143	↑	8.4545	1.82E-03	2.24E-03	0.7984	C21H44NO6P	Negative	78.96, 140.01, 196.03
PE(18:0/0:0)	480.3092	0	6.3049	↑	11.4850	7.95E-04	1.07E-03	0.8246	C23H48NO7P	Negative	96.96, 140.01, 196.04, 283.27
PG(22:6/0:0)	555.2716	2	5.6562	↑	5.2860	4.68E-03	5.16E-03	0.8024	C28H45O9P	Negative	170.98, 283.24, 327.23
UDP-N-acetyl-D-galactosamine	606.07389	0	0.5454	↑	59.8910	4.41E-06	9.15E-06	0.8695	C17H27N3O17 P2	Negative	111.0222, 158.92, 176.93, 282.04, 362.00, 385.00
PE(16:1/P-18:1)	698.5116	2	9.5703	↑	2.9631	9.80E-03	1.05E-02	0.7232	C39H74NO7P	Negative	78.96, 140.00, 253.21, 281.23
PA (18:0/18:1)	701.5236	0	9.8145	↓	0.1377	3.47E-06	7.85E-06	0.9056	C38H75N2O7P	Negative	281.25, 283.25, 419.25, 437.28
1-Palmitoyl-2-linoleoyl PE	714.5078	0	9.4018	↑	14.4588	1.16E-04	1.92E-04	0.8549	C39H74NO8P	Negative	140.01, 255.23, 279.22
PE(18:1/16:0)	716.5231	0	9.2768	↑	20.1049	6.31E-04	8.75E-04	0.8491	C39H76NO8P	Negative	140.01, 196.04, 281.25
PE(18:4/P-18:1)	720.4964	1	9.8168	↑	12.3568	1.50E-04	2.39E-04	0.859	C41H72NO7P	Negative	259.20, 581.49
PE(18:0/18:3)	740.5213	3	9.4586	↑	2.4862	8.10E-07	2.18E-06	0.8823	C41H76NO8P	Negative	140.01, 283.24, 482.32
PE(18:1/18:1)	742.5388	0	9.4276	↑	86.8130	2.11E-06	5.04E-06	0.9324	C41H78NO8P	Negative	140.01, 281.24, 460.28, 478.30
PG(O-16:0/20:1)	761.5677	3	9.1881	↑	2.4041	3.40E-04	5.04E-04	0.8392	C42H83O9P	Negative	152.99, 255.23, 391.20, 465.25
PE(16:0/22:6)	762.5106	3	9.9369	↑	2.5258	2.95E-04	4.53E-04	0.8322	C43H74NO8P	Negative	140.01, 255.23, 283.24, 327.23, 452.27
PE(22:6/P-18:1)	772.5278	1	9.3001	↓	0.2313	2.17E-03	2.59E-03	0.7861	C45H76NO7P	Negative	140.01, 283.24, 327.23, 444.28, 462.29
PS(13:0/22:1)	774.5303	1	8.9030	↑	76.5521	2.09E-09	1.12E-08	0.9307	C41H78NO10P	Negative	152.99, 491.31, 687.49
PE(18:1/22:6)	788.5258	2	9.1145	↑	8.9976	9.22E-05	1.59E-04	0.8939	C45H76NO8P	Negative	78.95 140.01, 152.99, 281.24, 283.26, 327.23, 460.26, 478.29, 506.26
PE(20:1/22:6)	816.5526	2	8.9397	↓	0.5173	2.94E-12	2.53E-11	0.979	C47H80NO8P	Negative	78.95, 152.99, 283.25, 309.27, 506.32
PI(16:0/18:1)	835.5348	0	9.7641	↑	20.4232	5.08E-06	9.93E-06	0.8881	C43H81O13P	Negative	78.95, 152.99, 281.25, 391.22, 553.29, 579.29
PG(19:0/22:4)	839.5768	4	9.8027	↓	0.4242	4.47E-06	9.15E-06	0.8753	C47H85O10P	Negative	78.95, 152.99, 331.26
PI(18:1/18:1)	861.5505	0	9.7810	↑	72.7961	1.00E-07	3.07E-07	0.9237	C45H83O13P	Negative	152.99, 223.00, 241.00, 579.30, 597.29
PI(14:0/22:1)	863.5645	1	9.8451	↑	60.5400	8.37E-08	2.88E-07	0.9161	C45H85O13P	Negative	78.96, 152.99, 241.01
PI(16:0/22:3)	887.5644	1	8.3686	↑	56.9735	6.90E-13	9.89E-12	0.7633	C47H85O13P	Negative	78.95, 152.99, 241.01, 553.24
Inosine	267.07301	1	0.4075	↓	0.1724	2.68E-03	3.11E-03	0.7984	C10H12N4O5	Negative	135.03, 108.02, 92.03
Continued											

Metabolite	*m/z*	Mass error (ppm)	RT (min)	In EC compared to Control	FC (T/C)	p-v	FDR	AUC	Formula	Ionization mode	Major CID fragments
Palmitic amide	256.2643	3	1.5642	↓	0.1359	4.10E-13	8.82E-12	0.8695	C16H33NO	Positive	158.154, 144.1381, 116.1068, 102.0914, 88.0753, 74.0598, 57.06
Oleamide	282.2804	4	1.6536	↓	0.9586	7.55E-09	3.61E-08	0.8403	C18H35NO	Positive	265.25, 247.242, 135.117, 97.102, 83.0853, 69.07
Stearamide	284.2957	3	2.21027	↓	0.8763	1.03E-14	4.43E-13	0.8648	C18H37NO	Positive	228.23, 200.20, 186.18, 172.17, 144.13, 116.10, 102.09, 88.07, 74.06
13Z-Docosenamide	338.3427	2	8.8864	↓	0.1942	1.94E-10	1.39E-09	0.8726	C22H43NO	Positive	321.3156, 303.3065, 149.1326, 135.117, 97.10, 83.08, 69.07
3-Deoxyvitamin D3	369.3527	3	12.7020	↑	14.4965	1.54E-06	3.90E-06	0.8619	C27H44	Positive	287.26, 233.22, 215.18, 161.13, 147.11, 133.10, 109.10, 81.06, 67.05
PC(16:0/0:0)	496.3409	2	5.6375	↑	91.6590	1.93E-02	1.98E-02	0.7051	C24H50NO7P	Positive	478.32, 184.07, 166.06, 125.00, 104.10, 86.09, 60.08
Glu Phe Arg Trp	637.3067	3	1.2381	↓	0.2823	7.91E-06	1.48E-05	0.866	C31H40N8O7	Positive	83.05, 147.11, 163.01, 239.04, 337.02, 393.07, 469.12, 525.17, 581.23
PC(14:0/18:2)	730.5416	4	8.9420	↑	286.1312	9.26E-10	5.69E-09	0.9312	C40H76NO8P	Positive	184.07, 242.11, 285.24
PC(18:1/14:0)	732.5552	1	7.6818	↑	4.9261	2.68E-12	2.53E-11	0.8275	C40H78NO8P	Positive	86.09, 184.07, 549.48
PC(16:0/20:5)	780.5545	0	8.8476	↑	40.7199	6.95E-08	2.72E-07	0.9161	C44H78NO8P	Positive	88.11, 184.07, 255.21, 393.24
PC(16:0/20:4)	782.5708	1	6.7811	↑	6.6011	1.17E-02	1.23E-02	0.6482	C44H80NO8P	Positive	86,09, 184.0, 125.00, 258.10, 313.27, 419.2485, 478.3295, 496.34, 526.32, 599.5016
PC(16:0/22:6)	806.5709	1	7.2133	↑	284.1474	1.52E-03	1.92E-03	0.7885	C46H80NO8P	Positive	184.07, 267.21, 478.32, 550.32, 623.50
PC(18:0/20:2)	814.6337	2	9.9895	↑	10.9063	3.56E-04	5.10E-04	0.7925	C46H88NO8P	Positive	184.07, 263.27, 341.30, 508.37
PC(18:1/22:6)	832.5861	1	9.8159	↑	9.9810	8.71E-08	2.88E-07	0.8922	C48H82NO8P	Positive	184.07, 522,35, 568,33

Table 2. Potential biomarkers confirmed by MS/MS. The table lists metabolites that showed significant change in the relative abundance in EC tissues compared to controls. (*m/z* = mass/charge; ppm = parts per million; RT = retention time; p-v = p-value; FC = fold change; FDR = false discovery rate; AUC = area under the curve).

progression, we transiently silenced the expression of the ADAR1 and ADAR2 enzymes in three EC cell lines (HEC-1A, Ishikawa and RL95-2). Transfection efficiency was determined by comparing each transfection against the corresponding negative control (NC) (Supplementary Figure 6). Gene silencing for both enzymes was confirmed by western blot analysis (WB) and immunofluorescence (IF) after 96 h of transfection (Supplementary Figure 7), time point in which the modulation of cell growth and cell viability in EC cell lines was assessed. The inhibition of ADAR2, but not ADAR1, resulted in a significant reduction of cell viability and proliferation compared to controls in the three cell lines used in this study (Fig. 3A). Furthermore, the ratio of apoptotic cells significantly increased in the three cell lines upon silencing of ADAR2. Similarly, proliferation assay performed using cells treated with siRNA-ADAR1, did not result in significant changes in apoptosis (Fig. 3B). Finally, the migration capabilities of the EC cell lines were interrogated after inhibiting ADAR1 and ADAR2 in a wound healing assay. Knockdown of ADAR2 expression resulted in significant reduction of the migration rate in HEC-1A and RL95-2 cell lines. Remarkably, similar changes were not observed when the same cell lines were treated with siRNA-ADAR1 (Fig. 3C).

Metabolite	*m/z*	Mass error (ppm)	RT (min)	Early vs late EC stages	FC (early/late)	p-value	FDR	AUC	Formula	Ionization mode	Column	Major CID fragments
Arachidonic Acid	303.2325	1	6.8246	↓	0.1402	6.24E-03	8.74E-03	0.778	C20H32O2	Negative	C18_BEH	259.25, 216.99, 205.19
PC(16:0/20:5)	780.5545	0	8.8476	↓	0.3311	3.13E-03	5.48E-03	0.750	C44H78NO8P	Positive	C18_BEH	88.11, 184.07, 255.21, 393.24
PC(16:0/22:6)	806.5709	1	9.0460	↑	3.0814	7.86E-04	5.48E-03	0.807	C46H80NO8P	Positive	C18_BEH	184.07, 267.21, 478.32, 550.32, 623.50
PE(16:0/22:6)	762.5106	3	9.0937	↑	2.0349	8.36E-03	9.75E-03	0.752	C43H74NO8P	Negative	C18_BEH	140.01, 255.23, 283.24, 327.23, 452.27
PE(18:1/22:6)	788.5258	2	9.1145	↑	15.2799	2.2E-02	2.20E-02	0.676	C45H76NO8P	Negative	CSH and C18_BEH	78.95 140.01, 152.99, 281.24, 283.26, 327.23, 460.26, 478.29, 506.26
PE(22:6/P-18:1)	772.5278	1	9.3001	↓	0.4065	1.86E-03	5.48E-03	0.800	C45H76NO7P	Negative	CSH and C18_BEH	140.01, 283.24, 327.23, 444.28, 462.29
UDP-N-acetyl-D-galactosamine	606.0739	0	0.5454	↓	0.3765	2.97E-02	5.48E-03	0.688	C17H27N3O17P2	Negative	C18_BEH	111.0222, 158.92, 176.93, 282.04, 362.00, 385.00

Table 3. Potential biomarkers confirmed using tandem mass spectrometry. List of metabolites that showed significant alterations in early vs late stages of EC. (*m/z* = mass/charge; ppm = parts per million; RT = retention time; p-v = p-value; FC = fold change; FDR = false discovery rate; AUC = area under the curve).

In order to further confirm the functional role of ADAR2 in EC cell lines, we also conducted functional assays silencing ADAR2 expression with two new and different siRNAs (siRNA-ADAR2_B and siRNA-ADAR2_C), independent from the siRNA-ADAR2 used in Fig. 3. We corroborate a significant decrease of cell viability, a significant increase in apoptosis rate and a significant decrease of wound healing rate of the 3 EC cell lines induced by the silencing of ADAR2 expression (Supplementary Figure 8). These results clearly demonstrate that changes in the expression of ADAR2 leads to significant changes in the aggressive behavior of EC cell lines, specifically on cell viability and apoptosis, and cell migration capabilities of EC cell lines.

Discussion

Recently, few studies have reported alterations in the metabolomic or proteomic phenotype of EC in serum[25], urine[26] or in other sample types[27,28], underscoring the clinical translational relevance of these technologies in furthering the personalized medicine initiative. In this study, we used a global metabolomics profiling approach in order to understand the metabolic changes that take place in EC carcinogenesis and during tumor progression (Fig. 4).

Our study reveals an array of metabolites dysregulated in EC, some of which have been previously described in endometrial carcinogenesis. Glycerophospholipid class of metabolites was found to be upregulated in tumor tissues, including PCs, PEs, and PIs. Lipid biosynthesis and catabolism is known to be altered in several diseases, including cancer[28–32]. Consistent with our findings, Trousil *et al.*[28] also observed an increase (up to 70%) in PC levels in EC tissues. We also observed a downregulation of the acylamido analogs of endocannabinoids such as palmitamide, stearamide and oleamide in EC. This downregulation has been previously reported[33]. We also observed a decreased proportion of picolinic acid in tumor samples. Picolinic acid and quinolinic acid are the end products of the Kynurenine pathway and have been shown to have anti-tumoral and pro-tumoral activity respectively[34]. Moreover, the activity of one of the main enzymes of the pathway, indoleamine 2,3-dioxygenase (IDO) has also been studied in EC[35]. Decreased levels of Glu Phe Arg Trp and inosine as well as upregulation of 3-Deoxyvitamin D3 and UDP-N-acetyl-D-galactosamine were also found in our EC tumor sample set compared to control tissues. We also analyzed metabolomic changes that underscore tumor progression. Our data suggest significant changes in the lipidome including PC (16:0/20:5), PC (16:0/22:6), PE (16:0/22:6), PE (22:6/P-18:1) and PE (18:1/22:6) at different stages of cancer progression, as well as a significant increase in the levels of

		Cohort 1			Cohort 2			Cohort 3		
		EEC (n = 20)	NEEC (n = 0)	Control (n = 20)	EEC (n = 36)	NEEC (n = 0)	Control (n = 0)	EEC (n = 78)	NEEC (n = 29)	Control (n = 0)
Age	> 50	20	—	20	36	—	—	78	29	—
	< 50	—	—	—	——	—	—	—	—	—
Collection center	VHUH	20	—	20	7	—	—	78	29	—
	Lleida	—	—	—	13	—	—	—	—	—
	Santiago	—	—	—	6	—	—	—	—	—
	Virgen Rocío	—	—	—	8	—	—	—	—	—
	MD Anderson	—	—	—	2	—	—	—	—	—
Uterine condition	Pre-menopausal	—	—	—	—	—	—	—	—	—
	Post-menopausal	20	—	20	36	—	—	78	29	—
FIGO stage	IA	10	—	—	13	—	—	17	7	—
	IB	6	—	—	10	—	—	31	5	—
	II	2	—	—	5	—	—	20	4	—
	III	2	—	—	8	—	—	7	11	—
	IV	—	—	—	—	—	—	3	2	—
Histologic grade	G1	4	—	—	12	—	—	13	0	
	G2	7	—	—	13	—	—	33	2	
	G3	9	—	—	11	—	—	32	27	
Format	TMA	—	—	—	36	—	—	78	29	—
	Individual slide	20	—	20	—	—	—	—	—	—

Table 4. Clinical and histological information of patients included in the validation set.

UDP-N-acetyl-D-galactosamine and arachidonic acid at advanced stages of EC. Further studies would be needed to fully understand the scope and impact of these alterations in cancer progression.

The dysregulation of inosine was further studied to dissect the functional implications of the RNA editing pathway in EC. The dysregulation of inosine is indicative of a possible imbalance in the I/A ratio, which was also reported by Trousil *et al.*[28] in studies with EC tissue compared to normal endometrium. The A-to-I conversion is the most common type of RNA editing found in mammals mediated by the ADAR enzymes. Although, to date, the RNA editing pathway and the expression and function of the ADAR gene family has not been interrogated in EC, it has been reported that the expression of ADAR enzymes is upregulated in many cancers[18,19], including breast[17] and esophageal squamous cell carcinoma[36,37]. The ADAR family is comprised of three members: ADAR1 and ADAR2, that are present in most human tissues; and ADAR3, that is brain specific[23]. Changes in editing frequencies have been described in other diseases including prostate, lung, kidney and testis tumors while reduced RNA levels of ADAR1, ADAR2 and ADAR3 have been observed in brain tumors[38,39]. ADAR enzymes are also involved in physiological events such neuronal development, immune response, cell response to viruses and regulation of miRNA expression among others[40,41].

Hence, we asked if the expression of ADAR enzymes had any correlation with EC initiation and progression. Our findings not only elucidate ADAR1 and ADAR2 enzymes to be significantly upregulated in EC tumor tissues compared to healthy endometrium, but also demonstrate a significant correlation between ADARs expression and the malignancy of the tumor. We found that the ADARs expression increased progressively with tumor grade. More importantly, our data showed a significant increase in the expression of ADAR1 and ADAR2 in the most aggressive subtype of EC, the NEEC, that have the worst predicted survival[3].

The role of the RNA editing pathway in EC was further investigated by knocking-down the expression of ADAR1 and ADAR2 in HEC-1A, RL95-2 and Ishikawa EC lines. Our results demonstrate the impact of decreased ADAR2 expression on an array of cellular functions in EC cell lines including a significant decrease of cell proliferation and viability, increased apoptosis rate, and reduced migration capabilities *in vitro*. Similar to our observations, silencing of ADARs in breast cancer cell lines led to less cell proliferation and more apoptosis[17] and overexpression of editing enzymes accelerated growth rate and colony formation in esophageal squamous cell carcinoma *in vitro*[36]. Taken together, our results strongly suggest, for the first time, that the RNA editing gene family, specifically ADAR2, may play an important role promoting EC carcinogenesis.

In conclusion, a global molecular profiling approach using high resolution mass spectrometry has been useful, not only to describe changes in the metabolome and lipidome in human endometrial carcinogenesis and EC progression but also led to the discovery of an important alteration of the RNA editing pathway in EC. We further demonstrated the role of this pathway in proliferation and viability, apoptosis, and migration of EC cells, leading us to conclude that the activation of the RNA editing pathway is an oncogenic process in EC. This study opens several avenues for further investigations of ADAR2 as possible target for the development of therapeutic approaches for the treatment of EC patients.

Methods

Patients and tissue collection. *Discovery and verification phase.* A total of 56 women (39 diagnosed with EC and 17 non-EC patients) were recruited at Vall Hebron University Hospital (VHUH) in Barcelona, Spain. All

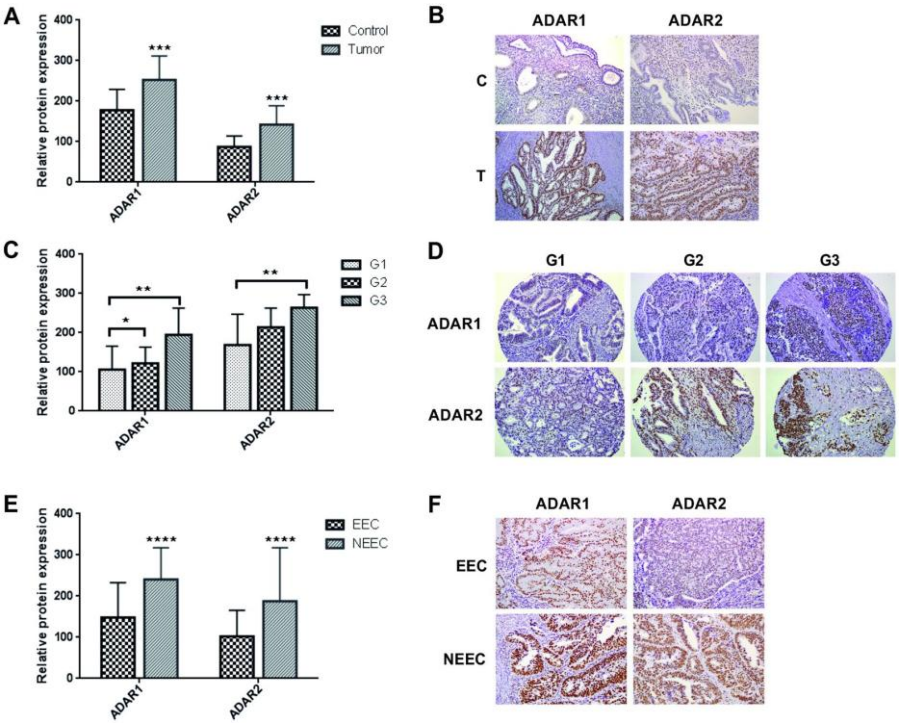


Figure 2. ADAR1 and ADAR2 proteins are overexpressed in EC. Panel A. ADAR1 and ADAR2 expression in EC tumors (T) compared to the paired control (C) tissues. Relative protein expression of each enzyme is plotted in the bar graphs. Panel B. Example of ADAR1 and ADAR2 staining levels in a matched EC tissue with the corresponding paired control. Panel C. ADAR1 and ADAR2 expression in EC tissues significantly correlate with the tumor grade. Panel D. Example of ADAR1 and ADAR2 staining levels in a set of 3 different EC tumor grades slides (grade 1, 2 and 3). Panel E. ADAR1 and ADAR2 levels are significantly increased in NEEC compared to EEC tumors. Panel F. Example of ADARs staining in EEC and NEEC tumors.

patients participating signed an informed consent and the study was approved by the Clinical Research Ethics Committee (CREC) at the VHUH (approval number: PR_AMI_50–2012). Endometrial tissues were collected at the histopathology department of VHUH after the surgical intervention. Each tissue piece (about 5 to 10 mg) was placed in a sterile and separate container, properly labeled and stored at −80 °C immediately. A description of the clinical and pathological characteristics of the tumors is detailed in Table 1. Inclusion criteria: post-menopausal women, > 50 years, no previous treatment for pelvic gynecological cancer, negative for HIV and hepatitis viruses.

Validation phase. Patients split in 3 new cohorts were enrolled in VHUH or in University Hospital Arnau de Vilanova of Lleida following the approval of the CREC at each participating institution. Tissue samples were embedded in paraffin block for individual slides or tissue microarray (TMA) construction in VHUH. A description of the clinical and pathological characteristics of the tissues is detailed in Table 4.

Tissue metabolomics using Ultra Performance Liquid Chromatography coupled to Quadrupole Time-Of-Flight Mass Spectrometry (UPLC-QTOF-MS). Reagents and chemicals: Solvents using chloroform, ACN, water and methanol were purchased from Fisher Optima grade, Fisher Scientific (New Jersey, USA). High purity formic acid (99%) was purchased from Thermo-Scientific (Rockford, IL, USA). Ammonium formate, debrisoquine and 4-nitrobenzoic acid (4-NBA) were purchased from Sigma- Aldrich (St. Louis, MO, USA).

For metabolomics analysis, endometrial tissue samples were prepared following the procedure previously described.[42] Fresh frozen tissue sections were homogenized on ice using a buffer containing 50% methanol and internal standards (1 mg/ml debrisoquine in distilled water, 1 mg/ml of 4-nitrobenzoic acid in methanol, 10 µg/ml of phosphatidic acid in 50% methanol-water, 0.1 µg/ml of lysophosphatidylcholine in 50% methanol-water). Protein precipitation was done by adding a 1:1 ratio of acetonitrile (ACN). Samples were centrifuged, the supernatant (Supernatant 1) was transferred to a fresh vial and dried under vacuum and the pellet was resuspended in prechilled dichloromethane:methanol (3:1). After sonication and centrifugation, the supernatant (Supernatant 2) was transferred to a fresh vial and dried under vacuum and the pellet was kept for protein estimation. Supernatant

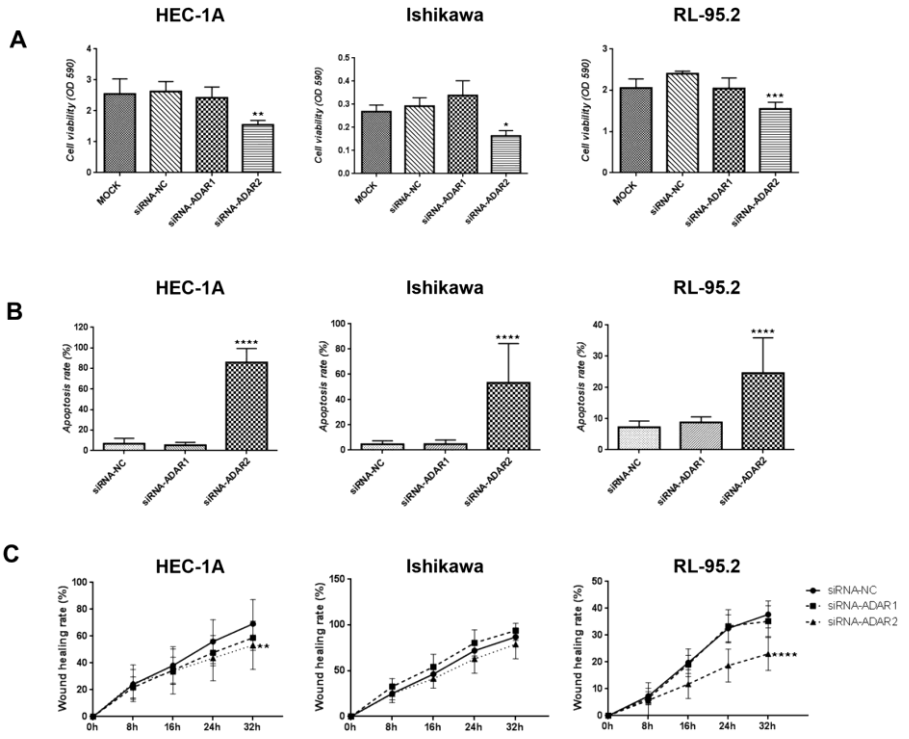


Figure 3. Functional assays revealing that ADAR2 presents oncogenic functions *in vitro*. HEC-1A, Ishikawa and RL95-2 EC cell lines were used for the functional assays. Panel A. Proliferation assay showing a significant decrease in cell viability (OD 590 nm) in the 3 cell lines when inhibiting ADAR2 expression. Panel B. Apoptosis assay showing a significant increase in apoptosis rate when silencing ADAR2. Panel C. Wound healing assay indicating a significant decrease in HEC-1A and RL95-2 migration capabilities (% of wound healing) when treating cells with siRNA-ADAR2. No significant changes were seen when inhibiting ADAR1.

1 and Supernatant 2 were finally resuspended in a buffer containing methanol:ACN:water in a ratio 50:25:25 for LC-MS analysis. The extraction procedure (from aqueous to semi-polar and lastly non-polar solvent) allows the isolation of a wide range of metabolites. In parallel, the pellet was resuspended with RIPA buffer and centrifuged in order to quantify the protein amount using Bradford method[43]. Resuspended pellets (5 µl) were injected onto Acquity UPLC CSH 1.7 µm, 2.1 × 100 mm column (Waters Corp.) or in a BEH C18 1.3 µm, 2.1 × 50 mm column. The mobile phase gradient consisted of ACN/water (60/40) containing 10 mM ammonium formate and 0.1% formic acid (Solvent A) and IPA/ACN (90/10) containing 10 mM ammonium formate and 0.1% formic acid (Solvent B). UPLC separation was performed at a flow rate of 0.4 ml/min for 20 min. Two different gradients were used in order to analyze the lipidome or the metabolome of the sample. MS data acquisition was performed using ESI-QTOF MS within the mass range of 50 to 1200 mass-to-charge ratio (m/z) in positive and negative electrospray ionization modes on a SYNAPT G2 Si (Waters Corporation, USA). The capillary voltage used was 3.2 kV and a sampling cone voltage of 30 V in negative mode and 20 V in positive mode. The desolvation gas flow was set to 750 l/h, and the temperature was set to 350 °C. The cone gas flow was 25 l/h, and the source temperature was 120 °C. Accurate mass was maintained by infusion of LockSpray interface with Leucine Enkaphalin (556.2771 $[M+H]^+$ and 554.2615 $[M-H]^-$). Data were acquired in TOF MS centroid mode and also in continuum mode for the mass range of 50 to 1200 mass-to-charge ratio (m/z) with MS scanning at a rate of 0.3 seconds. Total protein concentration for each sample was used to normalize any inconsistencies that would arise during tissue sampling. Subsequently, these data were normalized to internal standard to correct for analytical inconsistencies that could potentially occur during MS batch acquisition.

MS data were pre-processed using the XCMS software[44]. In order to determinate the identification of the metabolites based on the mass and charge, the following databases were used: Human Metabolome Database (www.hmdb.ca), Madison Metabolomics Consortium Database (mmcd.nmrfam.wisc.edu), LIPID MAPS (www.lipidmaps.org), KEGG (www.kegg.jp/kegg), and Metlin (metlin.scripps.edu). Multivariate data analysis was performed using Metaboanalyst 3.0 web tool[45,46] and a sub-set of metabolites were verified by tandem mass spectrometry (MS/MS) and using Mass Fragments software (Waters Corp.).

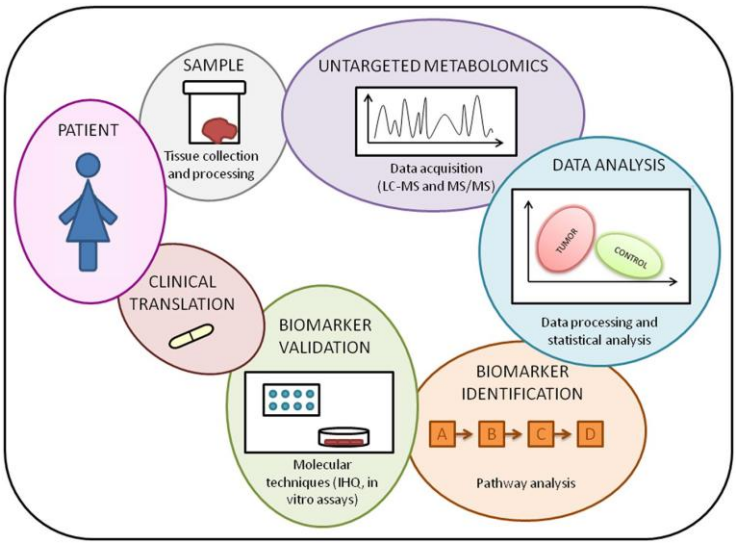


Figure 4. Project workflow summary.

Immunohistochemistry. A total of three TMAs were constructed as described previously[47]. Paraffin-embedded TMA or individual sections of the samples were mounted on slides, deparaffined and rehydrateted. Antigen retrieval was done during 4 min at 115 °C (pH 6) and subsequently blocked with peroxidase (3%) for 5 min and incubated with the primary antibody for 1 h and 30 min at room temperature in a humidified chamber. After blocking for 30 min with EnVision secondary antibody (Ref. K5007-100 ml, Agilent; Santa Clara, CA, USA) slides were treated with DAB (Ref. K3468, Agilent; Santa Clara, CA, USA) reagent and counterstained with hematoxylin for 30 sec. Finally, a dehytratation step and DPX mounting were performed. Pictures were taken using the Olympus BX41 microscope (20X). The following antibodies were tested: ADAR1 (Ref. AMAb90535, dilution 1:100, Sigma-Aldrich; St. Louis, MO, USA) and ADAR2 (Ref. sc-73409, dilution 1:50, Santa Cruz; Heidelberg, Germany, EU). A pathologist evaluated the expression using two criteria: the intensity of staining (0, no staining; 1, weak intensity; 2, moderate intensity; 3, high intensity) and the percentage of endometrial epithelial stained cells (0–100). The product of the two scores yielded final values on a sacale ranging from 0 to 300.

Cell lines. The following EC epithelial cell lines were used in this study. HEC-1A (Ref. HTB112, ATCC; Manassas, VA, USA) cell line was cultured in McCoys 5 A medium (Ref. 80014020, ThermoFisher; Waltham, MA, USA); Ishikawa (Ref. 99040201, Sigma-Aldrich; St. Louis, MO, USA) and RL95-2 (CRL-1671, ATCC; Manassas, VA, USA) were cultured in DMEM/F12 (Ref. 11320-033, ThermoFisher; Waltham, MA, USA). Both supplemented with penicillin/streptomycin (1%) and fetal bovine serum (10%). Incubated at 37 °C in a 5% CO2 humidified chamber. Cell lines were morphologically and genetically authenticated and tested for mycoplasma in accordance with AACR guides.

Western Blot. Protein extraction from cell lines was performed using RIPA buffer (5 nM EDTA, 150 mM NaCl, 1% Triton, 20 nM Tris pH 8 and 1:200 protein inhibitors). After Bradford colorimetric protein quantification samples were run on a SDS/PAGE acrylamide gel for 2 h and transferred to a PVDF membrane. Membranes were blocked for 1 h with 5% milk and incubated with the primary antibodies described above overnight and with the secondary antibodies -Goat anti-rabbit (Ref. P0448, dilution 1:2000, Agilent; Santa Clara, CA, USA) and rabbit anti-mouse (Ref. P0260, dilution 1:2000, Agilent; Santa Clara, CA, USA)- for 1 h at room temperature. Membranes were developed using Immobilon Immobilon Western Chemiluminiscent (Ref. WBKLS0100, Merck Millipore; Billerica, MA, USA). Membranes were finally stained with naphtol blue (Ref. N3393, Sigma-Aldrich; St. Louis, MO, USA) to detect all proteins transferred to the membranes in order to normalize the ADAR protein bands. The following antibodies were tested: ADAR1 (Ref. AMAb90535, dilution 1:100, Sigma-Aldrich; St. Louis, MO, USA) and ADAR2 (Ref. sc-73409, dilution 1:50, Santa Cruz; Heidelberg, Germany, EU).

Immunofluorescence. Cells were fixed for 10 min 96 h after transfection. Cells were blocked for 15 min with 5% BSA solution in 10 ml of 0.5% PBS-T. They were incubated with the primary antibody for 2 h at room temperature and with the secondary antibody for 45 min at room temperature in a dark chamber. Mounting and DAPI staining were performed with ProLong Gold Antifade reagent with DAPI (Ref. P36931, ThermoFisher; Waltham, MA, USA). Pictures were taken under the fluorescence microscope Nikon Eclipse TE2000-S. Primary antibodies: ADAR1 (Ref. AMAb90535, dilution 1:100, Sigma-Aldrich; St. Louis, MO, USA) and ADAR2 (Ref. sc-73409, dilution 1:50, Santa Cruz; Heidelberg, Germany, EU). Secondary antibody: Alexa Fluor 488 (Ref. A11017, Labeling and detection, ThermoFisher; Waltham, MA, USA).

Reverse Transfection of siRNAs. Desired number of cells according to the assay (see proliferation, apoptosis and wound healing assay sections for the exact number of cells) was seeded using media without antibiotic supplementation. Transfection mix was prepared in the following proportion: 25 µl of OptiMEM (Ref. 11058021, Invitrogen-ThermoFisher; Waltham, MA, USA), 0.2 µl of lipofectamine (Ref. 11668019, Invitrogen-ThermoFisher; Waltham, MA, USA), and 0.375 µl of siRNA (50 nM) and added in a 1:3 (v/v) to final volume of the well. After 24 h of transfection, media was changed. Pre-designed siRNAs for ADARs were purchased from Sigma-Aldrich: siRNA-ADAR1: SASI Hs01 00115047- GACUAUCUCUUCAAUGUGU; siRNA-ADAR2: SASI Hs01 00237747-GAGUGAUCGUGGCCUUGCA; siRNA-ADAR2_B: SASI Hs01 00222314- GAGUGAUCGUGGCCUUGCA; siRNA-ADAR2_C: SASI_Hs01_00135204-GAGUGAUCGUGGCCUUGCA; siRNA-NC: negative control, BLOCK-it Fluorescent Oligo, for lipid transfection (Ref. 2013, ThermoFisher; Waltham, MA, USA).

Cell proliferation assay. In order to determine cell proliferation and viability, a total of 4×10^3 cells (HEC-1A); 12×10^3 cells (Ishikawa), and 2×10^4 cells (RL95-2) per well were plated in a 96 well plate (6 replicates per condition). The cell growth rate was measured after 96 h of transfection. Cells were fixed to the plate with glutaraldehide 1% (Ref. G6257, Sigma-Aldrich; St. Louis, MO, USA) for 15 min and stained with Crystal Violet (Ref. C3886, Sigma-Aldrich; St. Louis, MO, USA). After 20 min, cells were washed with H_2O and Acetic Acid 15% (Ref. 0641, ThermoFisher; Waltham, MA, USA) was added. After 10 min shaking plates, they were read at 590 nm. Three independent experiments were carried out.

Apoptosis assay. In order to determine cell apoptosis rate, a total of 8×10^4 cells (HEC-1A); 5×10^4 cells (Ishikawa), and 2×10^5 cells (RL95-2) per well were seed in a p24 (8 replicates per condition). After 96 h of transfection, cells were stained with Hoechst (1:1000, Ref. H6024, Sigma-Aldrich; St. Louis, MO, USA) and after 24 h pictures were taken under the fluorescence microscope Nikon Eclipse TE2000-S. The number of nucleus with condensed chromatin and strongly staining of the DNA were quantified as apoptotic cells and divided by the total number of nucleus. Three independent experiments were carried out (4 camps per well were counted).

Wound healing assay. To evaluate cell migration capabilities, a total of 8×10^4 cells (HEC-1A); 5×10^4 cells (Ishikawa), and 2×10^5 cells (RL95-2) per well were seed in a p24 (3 replicates per condition). Wound was generated 72 h post-transfection. Pictures were taken every 8 h using an Olympus FSX100 microscope. Wound healing area was measured using Image J software at the different time points. Three experiments were carried out independently.

All experiments were performed in accordance with the relevant guidelines and regulations.

Statistical analysis. The SPSS statistical package version 23 for Windows® and the GraphPad Prism version 6 were used to perform the statistical analyses and ROC analysis. Each value represents the mean of at least 3 replicates with the corresponding standard deviation. We analyzed the normality of each data set and we used, according to the sample distribution, parametric or no parametric tests. For the IHQ analysis, a Wilcoxon signed rank test was used to compare tumors from controls; and a Kruskal-Wallis and Mann-Whitney tests were applied when comparing protein expression according to grades and histological subtypes, respectively. For the functional analysis, means of the different groups were compared by Kruskal-Wallis followed by Dunn's multiple comparisons test (in case of significance). We considered significant p-values < 0.05. (*p-v < 0.05; **p-v < 0.01; ***p-v < 0.001; ****p-v: 0.0001).

References

1. Morice, P., Leary, A., Creutzberg, C., Abu-Rustum, N. & Darai, E. Endometrial cancer. *Lancet (London, England)* **387**, 1094–108 (2016).
2. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2016. *CA Cancer J Clin* **66**, 7–30 (2016).
3. Piulats, J. M. *et al.* Molecular approaches for classifying endometrial carcinoma. *Gynecol. Oncol.* doi:10.1016/j.ygyno.2016.12.015 (2016).
4. Bokhman, J. V. Two pathogenetic types of endometrial carcinoma. *Gynecol. Oncol.* **15**, 10–17 (1983).
5. Miller, K. D. *et al.* Cancer treatment and survivorship statistics, 2016. *CA. Cancer J. Clin.* **66**, 271–89 (2016).
6. Prat, J., Gallardo, A., Cuatrecasas, M. & Catasús, L. Endometrial carcinoma: pathology and genetics. *Pathology* **39**, 72–87 (2007).
7. Altme, S. *et al.* Guidelines for the design, analysis and interpretation of 'omics' data: Focus on human endometrium. *Hum. Reprod. Update* **20**, 12–28 (2014).
8. Day, R. S. *et al.* Identifier mapping performance for integrating transcriptomics and proteomics experimental results. *BMC Bioinformatics* **12**, 213 (2011).
9. Zauber, P. *et al.* Strong correlation between molecular changes in endometrial carcinomas and concomitant hyperplasia. *Int. J. Gynecol. Cancer* **25**, 863–8 (2015).
10. Pappa, K. I. & Anagnou, N. P. Emerging issues of the expression profiling technologies for the study of gynecologic cancer. *American Journal of Obstetrics and Gynecology* **193**, 908–918 (2005).
11. Zhang, A., Sun, H., Yan, G., Wang, P. & Wang, X. Metabolomics for Biomarker Discovery: Moving to the Clinic. *Biomed Res. Int.* **2015**, 1–6 (2015).
12. Kalita-de Croft, P., Al-Ejeh, F., McCart Reed, A. E., Saunus, J. M. & Lakhani, S. R. 'Omics Approaches in Breast Cancer Research and Clinical Practice. *Adv. Anat. Pathol.* **23**, 356–367 (2016).
13. Luan, H. *et al.* Non-targeted metabolomics and lipidomics LC-MS data from maternal plasma of 180 healthy pregnant women. *Gigascience* **4**, 16 (2015).
14. Christinat, N., Morin-Rivron, D. & Masoodi, M. High-Throughput Quantitative Lipidomics Analysis of Nonesterified Fatty Acids in Human Plasma. *J. Proteome Res.* **15**, 2228–2235 (2016).
15. Tsutsui, H. *et al.* Biomarker discovery in biological specimens (plasma, hair, liver and kidney) of diabetic mice based upon metabolite profiling using ultra-performance liquid chromatography with electrospray ionization time-of-flight mass spectrometry. *Clin. Chim. Acta* **412**, 861–872 (2011).
16. Nordström, A. & Lewensohn, R. Metabolomics: Moving to the clinic. *Journal of Neuroimmune Pharmacology* **5**, 4–17 (2010).
17. Fumagalli, D. *et al.* Principles Governing A-to-I RNA Editing in the Breast Cancer Transcriptome. *Cell Rep.* **13**, 277–289 (2015).

18. Han, L. *et al.* The Genomic Landscape and Clinical Relevance of A-to-I RNA Editing in Human Cancers HHS Public Access. *Cancer Cell* **28**, 515–528 (2015).
19. Paz-Yaacov, N. *et al.* Elevated RNA Editing Activity Is a Major Contributor to Transcriptomic Diversity in Tumors. *Cell Rep.* **13**, 267–276 (2015).
20. Rayon-Estrada, V., Papavasiliou, F. N. & Harjanto, D. RNA Editing Dynamically Rewrites the Cancer Code. *Trends in Cancer* **1**, 211–212 (2015).
21. Meier, J. C., Kankowski, S., Krestel, H. & Hetsch, F. RNA Editing—Systemic Relevance and Clue to Disease Mechanisms? *Front. Mol. Neurosci.* **9**, 124 (2016).
22. Dominissini, D., Moshitch-Moshkovitz, S., Amariglio, N. & Rechavi, G. Adenosine-to-inosine RNA editing meets cancer. *Carcinogenesis* **32**, 1569–1577 (2011).
23. Chen, C. X. *et al.* A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains. *RNA* **6**, 755–67 (2000).
24. Huntley, M. A. *et al.* Complex regulation of ADAR-mediated RNA-editing across tissues. *BMC Genomics* **17**, 61 (2016).
25. Gaudet, M. M. *et al.* Analysis of serum metabolic profiles in women with endometrial cancer and controls in a population-based case-control study. *J. Clin. Endocrinol. Metab.* **97**, 3216–23 (2012).
26. Shao, X. *et al.* Screening and verifying endometrial carcinoma diagnostic biomarkers based on a urine metabolomic profiling study using UPLC-Q-TOF/MS. *Clin. Chim. Acta* **463**, 200–206 (2016).
27. Spratlin, J. L., Serkova, N. J. & Eckhardt, S. G. Clinical applications of metabolomics in oncology: a review. *Clin. Cancer Res.* **15**, 431–40 (2009).
28. Trousil, S. *et al.* Alterations of choline phospholipid metabolism in endometrial cancer are caused by choline kinase alpha overexpression and a hyperactivated deacylation pathway. *Cancer Res.* **74**, 6867–77 (2014).
29. Qiu, J.-F. *et al.* Abnormalities in Plasma Phospholipid Fatty Acid Profiles of Patients with Hepatocellular Carcinoma. *Lipids* **50**, 977–85 (2015).
30. Cífková, E. *et al.* Lipidomic differentiation between human kidney tumors and surrounding normal tissues using HILIC-HPLC/ESI-MS and multivariate data analysis. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* **1000**, 14–21 (2015).
31. Marien, E. *et al.* Non-small cell lung cancer is characterized by dramatic changes in phospholipid profiles. *Int. J. cancer* **137**, 1539–48 (2015).
32. Marien, E. *et al.* Phospholipid profiling identifies acyl chain elongation as a ubiquitous trait and potential target for the treatment of lung squamous cell carcinoma. *Oncotarget* **7**, 12582–97 (2016).
33. Jové, M. *et al.* Metabotyping human endometrioid endometrial adenocarcinoma reveals an implication of endocannabinoid metabolism. *Oncotarget* **7**, 52364–74 (2016).
34. Heng, B. *et al.* Understanding the role of the kynurenine pathway in human breast cancer immunobiology. *Oncotarget* **7**, 6506–20 (2016).
35. Ino, K. *et al.* Indoleamine 2,3-dioxygenase is a novel prognostic indicator for endometrial cancer. *Br. J. Cancer* **95**, 1555–61 (2006).
36. Qin, Y. R. *et al.* Adenosine-to-inosine RNA editing mediated by adars in esophageal squamous cell carcinoma. *Cancer Res.* **74**, 840–851 (2014).
37. Qiao, J.-J., Chan, T. H. M., Qin, Y.-R. & Chen, L. ADAR1: a promising new biomarker for esophageal squamous cell carcinoma? *Expert Rev. Anticancer Ther.* **14**, 1–4 (2014).
38. Paz, N. *et al.* Altered adenosine-to-inosine RNA editing in human cancer Altered adenosine-to-inosine RNA editing in human cancer. *Genome Res.* **17**, 1586–1595 (2007).
39. Zipeto, M. A., Jiang, Q., Melese, E. & Jamieson, C. H. M. RNA rewriting, recoding, and rewiring in human disease. *Trends in Molecular Medicine* **21**, 549–559 (2015).
40. Behm, M., Wahlstedt, H., Widmark, A., Eriksson, M. & Öhman, M. Accumulation of nuclear ADAR2 regulates A-to-I RNA editing during neuronal development. (2017).
41. Song, C., Sakurai, M., Shiromoto, Y. & Nishikura, K. Functions of the RNA Editing Enzyme ADAR1 and Their Relevance to Human Diseases. doi:10.3390/genes7120129.
42. Ghosh, S. P. *et al.* Metabolomic changes in gastrointestinal tissues after whole body radiation in a murine model. *Mol. BioSyst. Mol. BioSyst* **9**, 723–731 (2013).
43. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).
44. Mahieu, N. G., Genenbacher, J. L. & Patti, G. J. A roadmap for the XCMS family of software solutions in metabolomics. *Current Opinion in Chemical Biology* **30**, 87–93 (2016).
45. Xia, J., Psychogios, N., Young, N. & Wishart, D. S. MetaboAnalyst: A web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* **37**, (2009).
46. Xia, J., Sinelnikov, I. V., Han, B. & Wishart, D. S. MetaboAnalyst 3.0-making metabolomics more meaningful. *Nucleic Acids Res.* **43**, W251–W257 (2015).
47. Colas, E. *et al.* ETV5 cooperates with LPP as a sensor of extracellular signals and promotes EMT in endometrial carcinomas. *Oncogene* **31**, 4778–4788 (2012).

Acknowledgements

This work was supported by the Spanish Ministry of Health (RD12/0036/0035), the Spanish Ministry of Economy and Competitivy (PI14/02043), the AECC (Grupos Estables de Investigacion 2011 - AECC- GCB 110333 REVE), the Fundació La Marató TV3 (2/C/2013), the CIRIT Generalitat de Catalunya (2014 SGR 1330) and the European Commission, 7th Framework Program, IRSES (PROTBIOFLUID –269285) – Belgium. The Spanish Ministry of Economy and Competitiveness (IJCI-2015-25000) granted Dr. Colás and and the AGAUR Generalitat de Catalunya (2015FI_B00703) granted Tatiana Altadill. The authors would like to acknowledge the Proteomics and Metabolomics Shared Resource partially supported by Cancer Center Support Grant NIH/NCI grant P30-CA051008. The Institut de Salud Carlos III (FIS (PI13/01701)) also supported this project. Tissue samples were obtained with the support of "Xarxa Catalana de Bancs de Tumors" and "Plataforma de Biobancos" ISCIII (PT13/0010/0014). We would like to acknowledge experimental design guidance provided by Blanca Majem and all clinicians, Irene Campoy, Elena Martinez and Laura Devis for participating in the recruitment of clinical samples. We also acknowledge assistance in the statistical analysis provided by Ms Yiwen Wang. We finally thank volunteers that participated in the study.

Author Contributions

T.A., S.C. and A.G.-M. collected the samples. T.A. processed the samples. T.M.D., K.G. and S.M. collaborated analyzing the samples and interpreting the metabolomics data. A.R. and A.G. performed the histological

examination of the tissues. C.P.M., C.L. and E.C. contributed performing the *in vitro* assays. T.A., E.C. and A.K.C. wrote the manuscript. X.M.-G., J.R., S.W.B., A.K.C. and E.C. collaborated on designing the project.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-09169-2

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

SUPPLEMENTARY TABLES

Supplementary Table 1. List of metabolites putatively identified (based on accurate mass) in the discovery set that showed significant changes in their relative abundance in EC tissues compared to the controls. The identity of a sub-set of metabolites confirmed, using tandem mass spectrometry and matching the fragmentation pattern to a reference standard, are listed in Table 2.

Metabolite	*m/z*	Mass error (ppm)	RT (min)	FC (T/N)	p-value	Formula	Mode
Picolinic acid	122.0245	2	0.650	0.183	3.07E-03	C6H5NO2	Negative
Palmitic amide	256.2644	3	1.564	0.014	4.10E-13	C16H33NO	Positive
His Ile	267.1458	1	5.165	41.789	2.49E-02	C12H20N4O3	Negative
Linoleic acid	279.2320	3	6.905	3.282	2.33E-02	C18H32O2	Negative
Oleamide	282.2804	4	1.654	0.010	7.55E-09	C18H35NO	Positive
Stearamide	284.2957	3	2.210	0.088	1.03E-14	C18H37NO	Positive
3-Deoxyvitamin D3	369.3527	3	12.702	14.497	1.54E-06	C27H44	Positive
15beta-Hydroxy-7alpha-mercapto-pregn-4-ene-3,20-dione 7-acetate	403.1940	2	6.526	6.063	2.37E-03	C23H32O4S	Negative
Gly His Pro Val	409.2184	2	0.780	0.112	9.39E-06	C18H28N6O5	Positive
Lys Met His	415.2128	1	4.775	133.472	1.56E-05	C17H30N6O4S1	Positive
Gly Pro Arg Ser	416.2247	1	0.733	6.407	8.96E-11	C16H29N7O6	Positive
Ala Lys Asn Ser	417.2103	0	6.955	186.790	1.38E-03	C16H30N6O7	Negative
Ala Ala Lys Met	418.2134	0	6.329	23.056	1.16E-03	C17H33N5O5S	Negative
Ala Ala Ser Trp	432.1879	2	4.760	0.115	3.62E-03	C20H27N5O6	Negative
PE(P-16:0/0:0)	436.2827	1	5.814	8.455	1.82E-03	C21H44NO6P	Negative
Trp Met Asp	451.1664	3	0.652	0.370	7.81E-03	C20H26N4O6S1	Positive
Cys Lys Thr Cys	454.1790	0	0.713	52.837	1.28E-03	C16H31N5O6S2	Positive
1-O-alpha-D-glucopyranosyl	477.3807	4	0.733	129.167	1.13E-04	C26H52O7	Positive
PE(18:0/0:0)	480.3092	0	6.305	11.485	7.95E-04	C23H48NO7P	Negative
PC(16:0/0:0)	496.3409	2	5.637	91.659	1.93E-02	C24H50NO7P	Positive
Ala Glu Lys Arg	501.2812	4	5.336	0.386	3.21E-03	C20H38N8O7	Negative
Thr Thr Gly Leu Ile	502.2902	3	5.493	15.121	8.04E-04	C22H41N5O8	Negative
Ile Ile Met Arg	530.3152	4	6.199	26.818	3.44E-06	C23H45N7O5S	Negative
Glu Lys Lys Lys	532.3430	4	5.768	16.779	9.04E-03	C23H45N7O7	Positive
Asp His Lys Arg	553.2860	1	5.187	37.104	4.69E-04	C22H38N10O7	Negative
His Ile Met Arg	554.2877	0	5.143	10.703	4.84E-03	C23H41N9O5S	Negative
PG(22:6/0:0)	555.2716	2	5.656	5.286	4.68E-03	C28H45O9P	Negative
Arg Arg Met Leu	573.3305	0	6.330	49.118	4.38E-02	C23H46N10O5S	Negative
Lys Thr Glu Lys Ala	574.3226	3	6.329	22.446	1.27E-03	C24H45N7O9	Negative
Lys Lys Tyr Tyr	599.3197	0	6.653	50.352	1.79E-06	C30H44N6O7	Negative
UDP-N-acetyl-D-galactosamine	606.0739	0	0.545	59.891	4.41E-08	C17H27N3O17P2	Negative
His Gln Tyr Tyr	608.2452	3	9.007	0.014	3,96E-27	C29H35N7O8	Negative
Lys Thr Trp Trp	620.3220	4	6.141	34.551	2.34E-02	C32H41N7O6	Positive
Glu Phe Arg Trp	637.3067	3	1.238	0.028	7.91E-06	C31H40N8O7	Positive
SM(d18:1/14:0)	675.5442	1	8.327	6.491	2.30E-03	C37H75N2O6P	Positive
PE(16:1/P-18:1)	698.5116	2	9.570	2.963	9.80E-03	C39H74NO7P	Negative

SM(d16:1/18:1)	701.5599	0	8.436	4.410	8.92E-04	C39H77N2O6P	Positive
PA(O-16:0/21:0)	705.5816	3	7.054	0.132	8.23E-05	C40H81O7P	Positive
PE(16:0/17:0)	706.5414	4	7.300	382.658	4.95E-08	C38H76NO8P	Positive
PE(16:0/18:2)	714.5078	0	9.402	14.459	1.16E-04	C39H74NO8P	Negative
PE(18:1/16:0)	716.5231	0	9.277	20.105	6.31E-04	C39H76NO8P	Negative
PE(18:4/P-18:1)	720.4964	1	9.082	12.357	1.50E-04	C41H72NO7P	Negative
PC(14:0/18:2)	730.5389	1	8.942	58.473	4.66E-11	C40H76NO8P	Positive
SM(d18:1/18:0)	731.6070	1	8.726	0.016	4.59E-12	C41H83N2O6P	Positive
PC(18:1/14:0)	732.5552	1	7.682	4.926	2.68E-12	C40H78NO8P	Positive
PE(18:0/18:3)	740.5213	3	9.459	2.486	8.10E-07	C41H76NO8P	Negative
PE(18:1/18:1)	742.5388	0	9.428	86.813	2.11E-06	C41H78NO8P	Negative
PE(18:0/18:1)	746.5703	1	8.345	100.242	4.96E-08	C41H80NO8P	Positive
PE(O-16:0/22:5)	752.5607	2	9.745	0.139	5.42E-09	C43H78NO7P	Positive
PG(13:0/22:6)	753.4738	4	0.690	11.581	4.54E-05	C41H69O10P	Positive
SM(d18:1/20:0)	759.6383	1	9.807	0.116	1.21E-05	C43H87N2O6P	Positive
PG(O-16:0/20:1)	761.5677	3	9.188	2.404	3.40E-04	C42H83O9P	Negative
PE(16:0/22:6)	762.5106	3	9.094	2.526	2.95E-04	C43H74NO8P	Negative
PC(15:0/20:5)	764.5268	4	8.250	16.476	1.33E-05	C43H76NO8P	Negative
PE(22:6/P-18:1)	772.5278	1	9.300	0.231	2.17E-03	C45H76NO7P	Negative
PS(13:0/22:1)	774.5303	1	8.903	76.552	2.09E-09	C41H78NO10P	Negative
PA(20:2)/22:4)	775.5311	3	8.895	5.340	1.97E-12	C45H77O8P	Negative
PS(13:0/22:0)	776.5439	1	7.668	161.392	9.88E-14	C41H80NO10P	Negative
PA(20:1/22:4)	777.5470	3	7.668	12.800	1.24E-12	C45H79O8P	Negative
PC(16:0/20:4)	782.5708	1	6.781	6.601	1.17E-02	C44H80NO8P	Positive
SM(d18:1/22:0)	787.6701	1	10.615	0.017	1.75E-09	C45H91N2O6P	Positive
PE(18:1/22:6)	788.5258	2	9.115	8.998	9.22E-05	C45H76NO8P	Negative
PS(18:0/18:1)	790.5608	2	9.072	0.067	1.22E-09	C42H80NO10P	Positive
PS(15:0/22:2)	800.5437	1	8.995	5.224	5.79E-07	C43H80NO10P	Negative
PC(16:0/22:6)	806.5709	1	7.213	284.147	1.52E-03	C46H80NO8P	Positive
SM(d18:2/24:1)	811.6659	3	9.929	0.155	1.97E-10	C47H91N2O6P	Positive
PC(18:0/20:2)	814.6337	2	9.990	10.906	3.56E-04	C46H88NO8P	Positive
SM(d18:1/24:0)	815.7016	1	11.266	0.146	1.47E-05	C47H95N2O6P	Positive
PE(20:1/22:6)	816.5526	2	8.940	0.052	2.94E-12	C47H80NO8P	Negative
PS(17:0/22:4)	824.5433	1	8.838	5.989	9.13E-07	C45H80NO10P	Negative
PS(17:0/22:2)	828.5754	0	8.269	2.310	8.09E-05	C45H84NO10P	Negative
PC(18:1/22:6)	832.5861	1	9.082	9.981	8.71E-08	C48H82NO8P	Positive
PI(16:0/18:1)	835.5348	0	9.764	20.423	5.08E-06	C43H81O13P	Negative
PG(19:0/22:4)	839.5768	4	9.803	0.424	4.47E-06	C47H85O10P	Negative
Galbeta1-4Glcbeta-Cer(d18:1/16:0)	860.6069	4	8.906	4.897	1.22E-05	C46H87NO13	Negative
PI(18:1/18:1)	861.5505	0	9.781	72.796	1.00E-07	C45H83O13P	Negative
PI(14:0/22:1)	863.5645	1	9.845	60.540	8.37E-08	C45H85O13P	Negative
PI(16:0/22:3)	887.5644	1	8.369	56.974	6.90E-13	C47H85O13P	Negative
Inosine	267.07301	1	0.407	0.172	2.68E-03	C10H12N4O5	Negative
PC(16:0/20:5)	780.55447	0	8.848	40.720	6.95E-08	C44H78NO8P	Positive

Supplementary Table 2. List of features that showed significant dysregulation in the different FIGO stages of EC tissues that were interrogated in this study. The same input mass or metabolite can appear more than once in the table because it appears as significantly altered in more than one comparison (*m/zs* marked with an asterix). The identity of a sub-set of the metabolites was confirmed, using tandem mass spectrometry and matching the fragmentation pattern to a reference standard, are listed in Table 3.

m/z	Mass error (ppm)	FC	Comparison	p-value	Metabolite	Formula	Mode	Adduct
267.146*	1	2.080	Early (I and II)/Late (III)	2.88E-02	Leu His	C12H20N4O3	Negative	[M-H]-
303.233	1	0.140		6.24E-03	Arachidonic Acid	C20H32O2	Negative	[M-H]-
403.194*	1	4.152		7.90E-04	Glu Thr Arg	C15H28N6O7	Negative	[M-H]-
404.197	0	1.934		1.02E-03	Ala Gly Lys Met	C16H31N5O5S	Negative	[M-H]-
405.193	1	1.594		8.32E-04	Met Thr Arg	C15H30N6O5S1	Negative	[M-H]-
417.199	0	2.686		2.63E-04	Ala Leu Thr Asp	C17H30N4O8	Negative	[M-H]-
417.210*	0	1.118		3.13E-04	Ala Lys Asn Ser	C16H30N6O7	Negative	[M-H]-
418.208	3	1.748		3.59E-04	Trp Ser Lys	C20H29N5O5	Negative	[M-H]-
418.213*	0	9.460		4.17E-04	Ala Ala Lys Met	C17H33N5O5S	Negative	[M-H]-
431.179	0	3.203		7.69E-04	Ala Glu Val Asp	C17H28N4O9	Negative	[M-H]-
478.294	0	43.213		2.45E-02	PE(18:1/0:0)	C23H46NO7P	Negative	[M-H]-
480.309*	2	0.267		1.48E-02	PE(18:0/0:0)	C23H48NO7P	Negative	[M-H]-
485.282	4	3.028		2.28E-02	Ala Ile Gln Arg	C20H38N8O6	Negative	[M-H]-
524.278	0	0.106		3.33E-04	LysoPE(0:0/22:6)	C27H44NO7P	Negative	[M-H]-
526.293	1	0.390		2.44E-03	LysoPE(0:0/22:5)	C27H46NO7P	Negative	[M-H]-
566.346*	1	0.163		3.50E-02	PS(21:0/0:0)	C27H54NO9P	Negative	[M-H]-
606.074	0	0.377		2.98E-02	UDP-N-acetyl-D-galactosamine	C17H27N3O17P2	Negative	[M-H]-
665.314	1	1.636		9.71E-05	Lys Asp Tyr Glu Leu	C30H46N6O11	Negative	[M-H]-
746.513	0	2.387		2.78E-04	PE(20:5/P-18:1)	C43H74NO7P	Negative	[M-H]-
762.511*	3	2.035		8.36E-03	PE(16:0/22:6)	C43H74NO8P	Negative	[M-H]-
768.552	3	1.507		1.68E-02	PC(16:0/19:3)	C43H80NO8P	Negative	[M-H]-
772.528*	1	0.406		1.86E-03	PE(22:6/P-18:1)	C45H76NO7P	Negative	[M-H]-
774.542*	3	0.025		5.52E-03	PE(P-18:0/22:6)	C45H78NO7P	Negative	[M-H]-
779.563	4	0.144		5.74E-03	PA(20:0/22:4)	C45H81O8P	Negative	[M-H]-
788.526*	2	15.280		2.21E-02	PE(18:1/22:6)	C45H76NO8P	Negative	[M-H]-
803.563	4	11.292		3.35E-02	PA(22:0/22:6)	C47H81O8P	Negative	[M-H]-
824.543*	1	0.497		4.44E-02	PS(17:0/22:4)	C45H80NO10P	Negative	[M-H]-
828.575*	0	2.069		3.40E-04	PS(17:0/22:2)	C45H84NO10P	Negative	[M-H]-
851.562	4	0.127		4.82E-04	PI(13:0/22:0)	C44H85O13P	Negative	[M-H]-
852.574	2	1.570		5.41E-03	PS(19:0/22:4)	C47H84NO10P	Negative	[M-H]-
856.608	1	2.441		7.37E-04	PS(19:0/22:2)	C47H88NO10P	Negative	[M-H]-
858.660	0	3.101		1.13E-02	PS(P-20:0/22:0)	C48H94NO9P	Negative	[M-H]-
288.290	0	0.489		2.16E-04	C17 Sphinganine	C17H37NO2	positive	[M+H]+
415.213*	1	0.288		2.52E-05	Lys Met His	C17H30N6O4S1	positive	[M+H]+
454.179*	0	0.369		2.99E-05	Cys Cys Lys Thr	C16H31N5O6S2	positive	[M+H]+
665.419*	2	1.606		1.90E-03	PA(14:1/20:5)	C37H61O8P	positive	[M+H]+

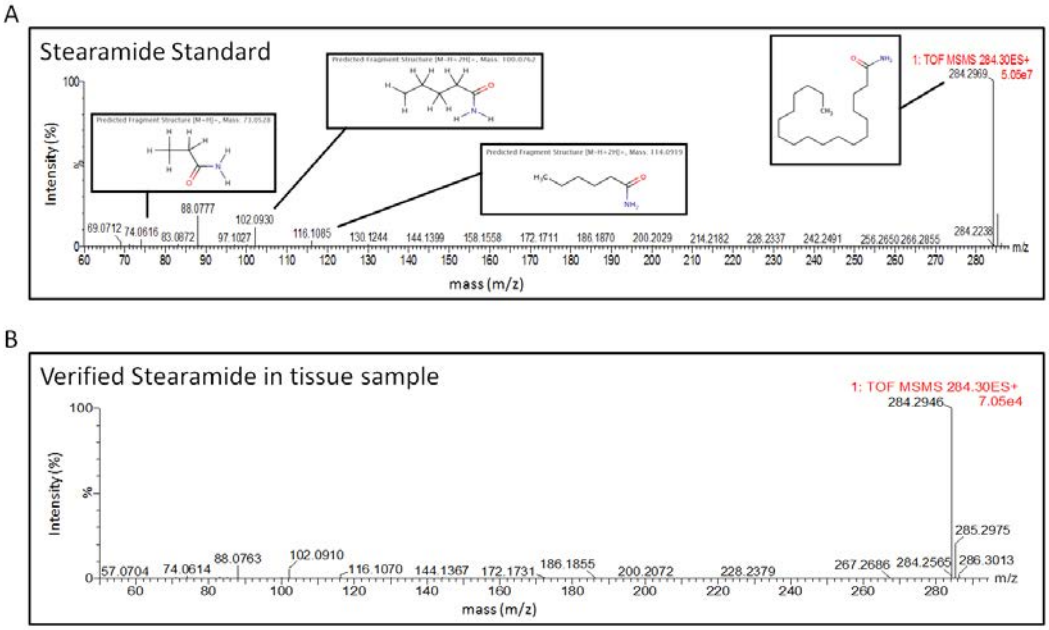
705.592	1	3.371		1.84E-02	SM(d18:0/16:0)	C39H81N2O6P	positive	[M+H]+
780.554*	0	0.331		3.13E-03	PC(16:0/20:5)	C44H78NO8P	positive	[M+H]+
806.571*	1	3.081		7.86E-04	PC(16:0/22:6)	C46H80NO8P	positive	[M+H]+
806.571*	1	0.019		5.34E-03	PE(22:6/19:0)	C46H80NO8P	positive	[M+H]+
305.248	1	19,139		2.77E-03	5,8,11-eicosatrienoic acid	C20H34O2	Negative	[M-H]-
436.283	1	16.302		1.53E-05	PE(P-16:0/0:0)	C21H44NO6P	Negative	[M-H]-
464.314	1	10.907		3.10E-02	PC(P-15:0/0:0)	C23H48NO6P	Negative	[M-H]-
480.309*	0	15.617		1.25E-05	PE(18:0/0:0)	C23H48NO7P	Negative	[M-H]-
502.290*	3	8.916		2.05E-04	Thr Thr Gly Leu Ile	C22H41N5O8	Negative	[M-H]-
506.323	4	9		2.56E-03	PC(17:1/0:0)	C25H50NO7P	Negative	[M-H]-
538.347*	0	0.236		4.53E-02	Lys Lys Lys His	C24H45N9O5	Negative	[M-H]-
541.333	4	14.347		2.20E-02	Val Ile Pro Lys Ser	C25H46N6O7	Negative	[M-H]-
553.286*	1	36.245		3.20E-02	Asp His Lys Arg	C22H38N10O7	Negative	[M-H]-
555.272	2	9.7976		3.35E-03	PG(22:6/0:0)	C28H45O9P	Negative	[M-H]-
566.346	1	7.742		1.31E-03	PS(21:0/0:0)	C27H54NO9P	Negative	[M-H]-
599.320	0	62.527		3.01E-06	Lys Lys Tyr Tyr	C30H44N6O7	Negative	[M-H]-
608.245*	3	0.014		2.28E-20	His Gln Tyr Tyr	C29H35N7O8	Negative	[M-H]-
714.508*	0	10.866		6.47E-05	PE(16:0/18:2)	C39H74NO8P	Negative	[M-H]-
720.496	1	24.311		8.18E-04	PE(18:4/P-18:1)	C41H72NO7P	Negative	[M-H]-
740.521*	3	17.02		9.06E-06	PE(18:0/18:3)	C41H76NO8P	Negative	[M-H]-
742.539*	0	87.086		1.76E-08	PE(18:1/18:1)	C41H78NO8P	Negative	[M-H]-
747.517	1	2.337		2.16E-04	PG(16:0/18:1)	C40H77O10P	Negative	[M-H]-
748.528*	0	0.200		2.79E-06	PE(20:4/P-18:1)	C43H76NO7P	Negative	[M-H]-
750.544	0	0.454	Early (I and II)/N	8.50E-10	PE(O-16:0/22:5)	C43H78NO7P	Negative	[M-H]-
762.511*	3	2.658		3.11E-05	PE(16:0/22:6)	C43H74NO8P	Negative	[M-H]-
764.527*	4	12.513		1.08E-05	PC(15:0/20:5)	C43H76NO8P	Negative	[M-H]-
772.528*	1	0.358		1.56E-02	PE(22:6/P-18:1)	C45H76NO7P	Negative	[M-H]-
774.530*	1	73.131		2.79E-09	PS(13:0/22:1)	C41H78NO10P	Negative	[M-H]-
776.544*	1	169.37		2.46E-13	PS(13:0/22:0)	C41H80NO10P	Negative	[M-H]-
777.547*	3	10.311		1.64E-11	PA(20:1/22:4)	C45H79O8P	Negative	[M-H]-
788.526*	2	9.8885		2.28E-06	PE(18:1/22:6)	C45H76NO8P	Negative	[M-H]-
788.544*	0	0.263		5.69E-10	PS(18:0/18:1)	C42H80NO10P	Negative	[M-H]-
800.544*	1	6.613		2.29E-08	PS(15:0/22:2)	C43H80NO10P	Negative	[M-H]-
816.553*	2	0.008		4.16E-11	PE(20:1/22:6)	C47H80NO8P	Negative	[M-H]-
824.543*	1	8.711		1.14E-08	PS(17:0/22:4)	C45H80NO10P	Negative	[M-H]-
828.575*	0	2.794		5.41E-06	PS(17:0/22:2)	C45H84NO10P	Negative	[M-H]-
835.535*	0	21.918		2.27E-06	PI(16:0/18:1)	C43H81O13P	Negative	[M-H]-
839.577*	4	0.424		2.53E-05	PG(19:0/22:4)	C47H85O10P	Negative	[M-H]-
861.551*	0	90.4		2.90E-08	PI(18:1/18:1)	C45H83O13P	Negative	[M-H]-
863.565*	1	75.341		7.75E-07	PI(14:0/22:1)	C45H85O13P	Negative	[M-H]-
887.564*	1	43.357		1.15E-09	PI(16:0/22:3)	C47H85O13P	Negative	[M-H]-
256.264*	3	0.0136		5.36E-13	Palmitic amide	C16H33NO	Positive	[M+H]+
282.280*	4	0.019		2.55E-10	Oleamide	C18H35NO	Positive	[M+H]+
284.296*	3	0.088		1.01E-14	Stearamide	C18H37NO	Positive	[M+H]+
338.343*	2	0.035		5.99E-08	13Z-Docosenamide	C22H43NO	Positive	[M+H]+

369.353	3	28.704		1.16E-04	3-Deoxyvitamin D3	C27H44	Positive	[M+H]+
409.218*	2	0.112		5.43E-05	Gly His Pro Val	C18H28N6O5	Positive	[M+H]+
415.213*	1	130.06		1.37E-20	Lys Met His	C17H30N6O4S1	Positive	[M+H]+
451.166	3	0.370		1.64E-02	Trp Met Asp	C20H26N4O6S1	Positive	[M+H]+
454.179*	0	13.624		4.05E-18	Cys Cys Lys Thr	C16H31N5O6S2	Positive	[M+H]+
477.381*	4	209.13		3.10E-08	1-O-alpha-D-glucopyranosyl-1,2-eicosandiol	C26H52O7	Positive	[M+H]+
496.341	2	141.65		5.38E-03	PC(16:0/0:0)	C24H50NO7P	Positive	[M+H]+
522.356	1	75.614		1.46E-03	PC(O-16:1/2:0)	C26H52NO7P	Positive	[M+H]+
637.307*	3	0.028		1.65E-03	Glu Phe Arg Trp	C31H40N8O7	Positive	[M+H]+
675.544*	1	8.3		1.02E-02	SM(d18:1/14:0)	C37H75N2O6P	Positive	[M+H]+
705.582*	3	0.015		7.23E-04	PA(O-16:0/21:0)	C40H81O7P	Positive	[M+H]+
706.541*	4	454.26		1.61E-06	PE(16:0/17:0)	C38H76NO8P	Positive	[M+H]+
720.555*	1	31.796		5.81E-07	PE(17:0/17:0)	C39H78NO8P	Positive	[M+H]+
730.542*	4	286.63		7.60E-10	PC(14:0/18:2)	C40H76NO8P	Positive	[M+H]+
731.607*	1	0.014		1.44E-09	SM(d18:1/18:0)	C41H83N2O6P	Positive	[M+H]+
732.555*	1	5.55		1.28E-11	PC(14:0/18:1)	C40H78NO8P	Positive	[M+H]+
746.570*	1	111.54		1.23E-09	PE(18:0/18:1)	C41H80NO8P	Positive	[M+H]+
752.560*	1	0.064		1.40E-10	PE(20:3/P-18:1)	C43H78NO7P	Positive	[M+H]+
759.638*	1	0.116		1.22E-04	SM(d18:1/20:0)	C43H87N2O6P	Positive	[M+H]+
780.554*	0	82.53		2.45E-10	PC(16:0/20:5)	C44H78NO8P	Positive	[M+H]+
782.571	1	12.62		3.80E-02	PE(22:4/17:0)	C44H80NO8P	Positive	[M+H]+
787.670*	1	0.024		1.12E-08	SM(d18:1/22:0)	C45H91N2O6P	Positive	[M+H]+
806.571*	1	367.48		1.28E-05	PC(16:0/22:6)	C46H80NO8P	Positive	[M+H]+
811.666*	3	0.011		2.02E-09	SM(d18:2/24:1)	C47H91N2O6P	Positive	[M+H]+
813.686*	1	0.281		7.98E-10	SM(d18:1/24:1)	C47H93N2O6P	Positive	[M+H]+
814.634	2	21.334		3.64E-04	PC(16:0/22:2)	C46H88NO8P	Positive	[M+H]+
815.701	1	0.279		5.56E-05	SM(d18:1/24:0)	C47H95N2O6P	Positive	[M+H]+
832.586	1	19.434		8.70E-08	PC(18:1/22:6)	C48H82NO8P	Positive	[M+H]+
834.601	0	50.046		1.93E-03	PC(18:0/22:6)	C48H84NO8P	Positive	[M+H]+
836.611	1	10.669		7.67E-04	LacCer(d18:0/14:0)	C44H85NO13	Positive	[M+H]+
267.146*	1	43.913	Late (III)/N	4.77E-05	Histidylleucine	C12H20N4O3	Negative	[M-H]-
281.248	1	0.128		1.53E-05	2Z-octadecenoic acid	C18H34O2	Negative	[M-H]-
303.232	2	0.293		9.94E-04	8,11-eicosadiynoic acid	C20H32O2	Negative	[M-H]-
355.199	3	29.133		2.87E-02	Ala Ala Pro Val	C16H28N4O5	Negative	[M-H]-
397.209	0	37.246		2.27E-02	Pro Pro Ser Val	C18H30N4O6	Negative	[M-H]-
399.224	1	31.782		1.65E-02	Ala Ile Pro Thr	C18H32N4O6	Negative	[M-H]-
403.194*	1	10.872		2.16E-16	Gly Lys Asn Ser	C15H28N6O7	Negative	[M-H]-
417.210*	0	363.29		1.03E-20	Ala Lys Asn Ser	C16H30N6O7	Negative	[M-H]-
418.213*	0	44.8		3.71E-22	Ala Ala Lys Met	C17H33N5O5S	Negative	[M-H]-
426.259	2	40.28		3.44E-03	Arg Pro Arg	C17H33N9O4	Negative	[M-H]-
432.188	2	0.019		4.67E-05	Ala Ala Ser Trp	C20H27N5O6	Negative	[M-H]-
443.253	3	32.89		2.53E-03	Ala Glu Ile Ile	C20H36N4O7	Negative	[M-H]-
449.150	0	0.032		2.31E-02	Trp Met Asp	C20H26N4O6S1	Negative	[M-H]-
471.280	1	13.68		1.68E-02	Ala Ala Arg Arg	C18H36N10O5	Negative	[M-H]-
480.309*	0	7.56		4.84E-02	PE(18:0/0:0)	C23H48NO7P	Negative	[M-H]-

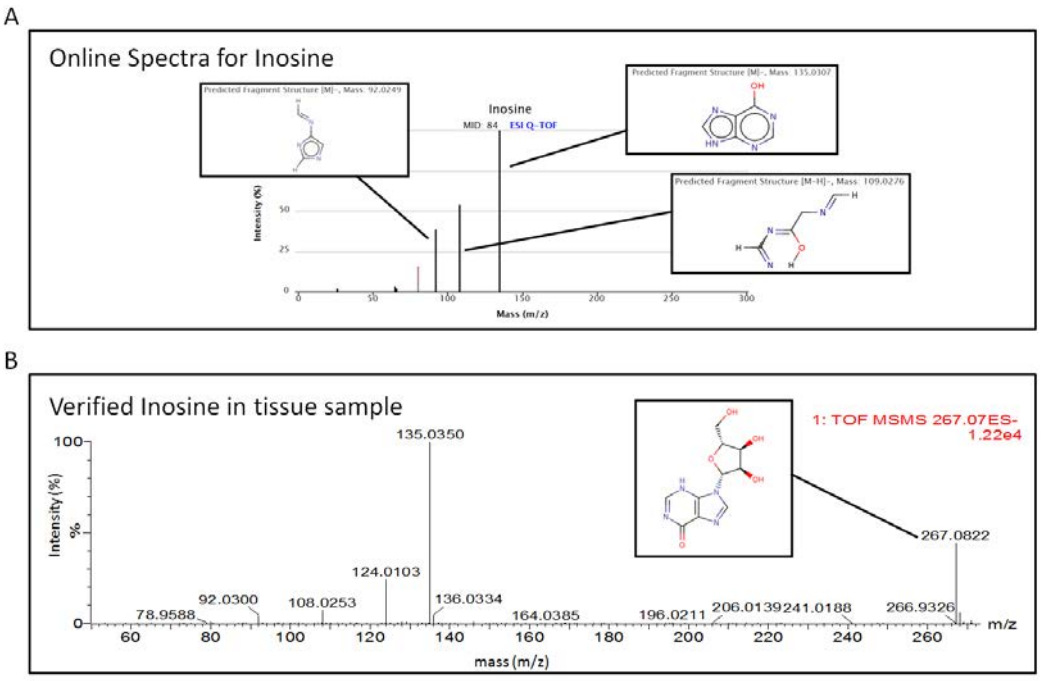
499.360	2	3.134		2.82E-03	Ile Ile Lys Lys	C24H48N6O5	Negative	[M-H]-
501.281	4	0.465		1.58E-02	Ala Glu Lys Arg	C20H38N8O7	Negative	[M-H]-
502.290*	3	21.16		1.51E-02	Thr Thr Gly Leu Ile	C22H41N5O8	Negative	[M-H]-
530.315	4	51.346		1.78E-07	Ile Ile Met Arg	C23H45N7O5S	Negative	[M-H]-
538.347*	0	0.236		9.44E-04	His Lys Lys Lys	C24H45N9O5	Negative	[M-H]-
553.286*	1	37.92		2.30E-06	Asp His Lys Arg	C22H38N10O7	Negative	[M-H]-
558.339	3	28.384		1.28E-04	Glu Lys Lys Arg	C23H45N9O7	Negative	[M-H]-
573.330	0	94.831		2.28E-04	Ile Met Arg Arg	C23H46N10O5S	Negative	[M-H]-
574.323	3	42.82		6.99E-05	Lys Thr Glu Lys Ala	C24H45N7O9	Negative	[M-H]-
599.319	0	38.788		9.18E-04	Lys Lys Tyr Tyr	C30H44N6O7	Negative	[M-H]-
601.389	2	56.529		1.76E-03	PA(12:0/17:2)	C32H59O8P	Negative	[M-H]-
608.245*	3	0.014		4.75E-34	His Gln Tyr Tyr	C29H35N7O8	Negative	[M-H]-
687.544	0	0.323		2.33E-04	SM(d16:1/17:0)	C38H77N2O6P	Negative	[M-H]-
714.508*	0	17.872		3.48E-03	PE(16:0/18:2)	C39H74NO8P	Negative	[M-H]-
716.523	0	21.399		3.52E-03	PE(18:1/16:0)	C39H76NO8P	Negative	[M-H]-
724.528	1	2.184		2.50E-07	PE(18:2/P-18:1)	C41H76NO7P	Negative	[M-H]-
736.528	1	2.436		8.68E-05	PC(18:4/P-16:0)	C42H76NO7P	Negative	[M-H]-
740.521*	3	17.11		8.94E-03	PE(18:0/18:3)	C41H76NO8P	Negative	[M-H]-
742.539*	0	86.553		1.12E-03	PE(18:1/18:1)	C41H78NO8P	Negative	[M-H]-
748.528*	0	0.233		2.13E-04	PE(20:4/P-18:1)	C43H76NO7P	Negative	[M-H]-
761.568	3	3.7379		1.47E-02	PG(O-16:0/20:1)	C42H83O9P	Negative	[M-H]-
762.511*	3	2.400		1.08E-02	PE(16:0/22:6)	C43H74NO8P	Negative	[M-H]-
764.527*	4	20.241		1.17E-03	PC(15:0/20:5)	C43H76NO8P	Negative	[M-H]-
768.553	2	4.6468		4.49E-02	PC(16:0/19:3)	C43H80NO8P	Negative	[M-H]-
772.528*	1	0.110		1.74E-03	PE(22:6/P-18:1)	C45H76NO7P	Negative	[M-H]-
774.530*	1	79.805		1.88E-06	PS(13:0/22:1)	C41H78NO10P	Negative	[M-H]-
774.542*	3	0.405		2.22E-03	PE(P-18:0/22:6)	C45H78NO7P	Negative	[M-H]-
775.531	3	9.4641		9.31E-08	PA(20:2/22:4)	C45H77O8P	Negative	[M-H]-
776.544*	1	153.81		9.32E-09	PS(13:0/22:0)	C41H80NO10P	Negative	[M-H]-
777.547*	3	15.165		4.10E-08	PA(20:1/22:4)	C45H79O8P	Negative	[M-H]-
788.526*	2	8.1507		1.30E-02	PE(18:1/22:6)	C45H76NO8P	Negative	[M-H]-
788.544*	0	0.494		7.83E-07	PS(18:0/18:1)	C42H80NO10P	Negative	[M-H]-
800.544*	1	3.9045		1.12E-03	PS(15:0/22:2)	C43H80NO10P	Negative	[M-H]-
816.553*	2	0.093		5.24E-09	PE(20:1/22:6)	C47H80NO8P	Negative	[M-H]-
824.543*	1	3.403		8.41E-04	PS(17:0/22:4)	C45H80NO10P	Negative	[M-H]-
835.535*	0	19.3		3.92E-04	PI(16:0/18:1)	C43H81O13P	Negative	[M-H]-
839.577*	4	0.424		3.19E-05	PG(19:0/22:4)	C47H85O10P	Negative	[M-H]-
860.607	4	8.6		4.22E-05	Galbeta1-4Glcbeta-Cer(d18:1/16:0)	C46H87NO13	Negative	[M-H]-
861.551*	0	56.74		1.68E-04	PI(18:1/18:1)	C45H83O13P	Negative	[M-H]-
863.565*	1	46.48		1.25E-05	PI(14:0/22:1)	C45H85O13P	Negative	[M-H]-
887.564*	1	69.912		5.25E-09	PI(16:0/22:3)	C47H85O13P	Negative	[M-H]-
256.264*	3	0.014		3.38E-10	Palmitic amide	C16H33NO	Positive	[M+H]+
282.280*	4	5.671E-04		3.11E-07	Oleamide	C18H35NO	Positive	[M+H]+
284.296*	3	0.088		4.15E-11	Stearamide	C18H37NO	Positive	[M+H]+
338.343*	2	4.402E-03		5.63E-16	13Z-Docosenamide	C22H43NO	Positive	[M+H]+

409.218*	2	0.112		2.55E-03	Gly His Pro Val	C18H28N6O5	Positive	[M+H]+
416.225	1	11.544		1.06E-06	Gly Pro Arg Ser	C16H29N7O6	Positive	[M+H]+
463.365	4	18.616		4.36E-03	1-O-alpha-D-glucopyranosyl-1,2-nonadecandiol	C25H50O7	Positive	[M+H]+
477.381*	4	53.203		1.70E-02	1-O-alpha-D-glucopyranosyl-1,2-eicosandiol	C26H52O7	Positive	[M+H]+
532.343	4	31.769		1.83E-03	Glu Lys Lys Lys	C23H45N7O7	Positive	[M+H]+
620.322	4	66.424		7.57E-04	Lys Thr Trp Trp	C32H41N7O6	Positive	[M+H]+
637.307*	3	0.028		1.25E-05	Glu Phe Arg Trp	C31H40N8O7	Positive	[M+H]+
665.419*	2	13.503		2.90E-03	PA(14:1/20:5)	C37H61O8P	Positive	[M+H]+
675.544*	1	5.551		3.37E-03	SM(d18:1/14:0)	C37H75N2O6P	Positive	[M+H]+
701.560	0	6.794		2.13E-03	SM(d16:1/18:1)	C39H77N2O6P	Positive	[M+H]+
705.582*	3	0.243		3.69E-04	PA(O-16:0/21:0)	C40H81O7P	Positive	[M+H]+
706.541*	4	314.64		5.80E-06	PE(16:0/17:0)	C38H76NO8P	Positive	[M+H]+
720.555*	1	4.2454		1.30E-04	PE(17:0/17:0)	C39H78NO8P	Positive	[M+H]+
730.542*	4	285.66		1.02E-06	PC(14:0/18:2)	C40H76NO8P	Positive	[M+H]+
731.607*	1	0.029		4.63E-08	SM(d18:1/18:0)	C41H83N2O6P	Positive	[M+H]+
732.555*	1	4.8033		8.81E-08	PC(14:0/18:1)	C40H78NO8P	Positive	[M+H]+
746.570*	1	89.507		4.05E-05	PE(18:0/18:1)	C41H80NO8P	Positive	[M+H]+
746.604	2	0.353		3.64E-02	PC(O-18:1/16:0)	C42H84NO7P	Positive	[M+H]+
752.560*	1	0.210		5.78E-06	PE(20:3/P-18:1)	C43H78NO7P	Positive	[M+H]+
752.561	2	0.070		2.03E-06	PE(O-16:0/22:5)	C43H78NO7P	Positive	[M+H]+
753.474	4	21.633		4.28E-06	PG(13:0/22:6)	C41H69O10P	Positive	[M+H]+
759.638*	1	0.116		2.35E-06	SM(d18:1/20:0)	C43H87N2O6P	Positive	[M+H]+
787.670*	1	9.620E-03		9.22E-07	SM(d18:1/22:0)	C45H91N2O6P	Positive	[M+H]+
794.606	0	0.408		2.35E-04	PC(O-18:1/20:4)	C46H84NO7P	Positive	[M+H]+
806.571*	1	2.334		3.59E-02	PC(16:0/22:6)	C46H80NO8P	Positive	[M+H]+
811.666*	3	0.291		1.14E-07	SM(d18:2/24:1)	C47H91N2O6P	Positive	[M+H]+
813.686*	1	0.465		4.38E-06	SM(d18:1/24:1)	C47H93N2O6P	Positive	[M+H]+
815.702	1	0.020		1.97E-04	SM(d18:1/24:0)	C47H95N2O6P	Positive	[M+H]+

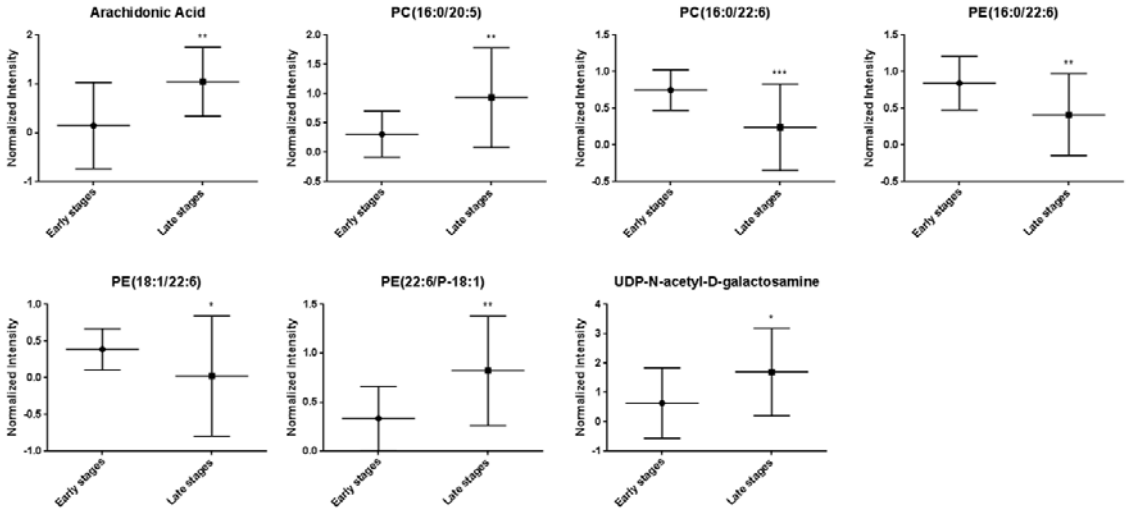
Supplementary figure 1. Verification of stearamide using tandem MS. The identification of the metabolite stearamide (284.2946) in the samples was confirmed by performing MS/MS in the ESI postive mode with a collision energy of 40 V; the resultant fragmentation spectra was compared to that of a commercially available.



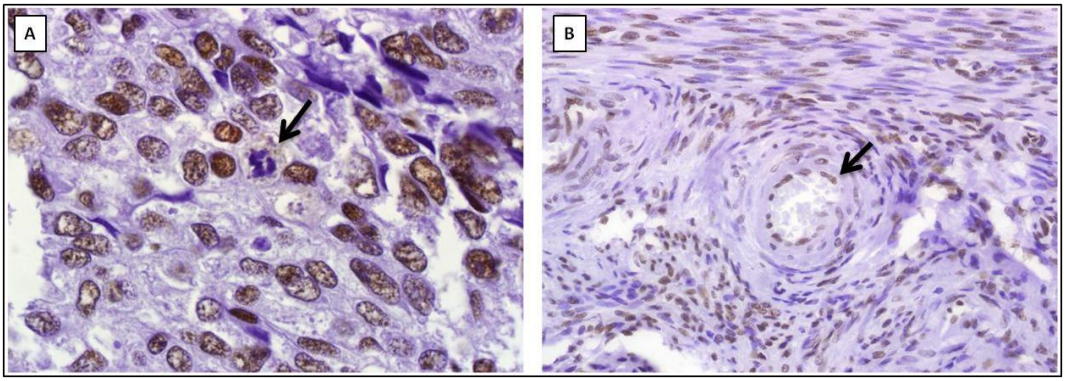
Supplementary figure 2. MS/MS structural verification of inosine. The identification of the metabolite inosine (267.0822) in the samples was confirmed by performing MS/MS in the ESI negative mode with a collision energy of 40 V; the resultant fragmentation spectra was compared to that available in Metlin web database.



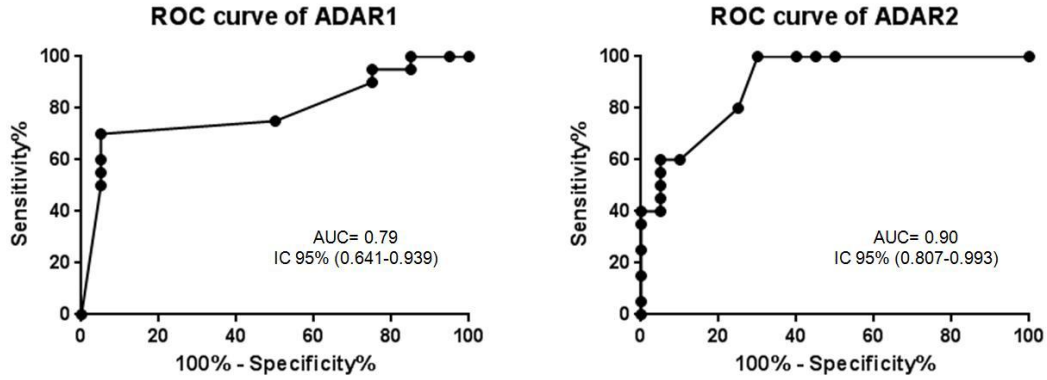
Supplementary figure 3. Metabolite panel as a predictor of EC progression. The realtive abundance of a panel of metabolites including arachidonic acid, PC (16:0/20:5), PE (22:6/P-18:1) and UDP-N-acetyl-D-galactosamine were significantly (p-v < 0.05) increased in late FIGO stages of EC compared to early stages while PC (16:0/22:6), PE (16:0/22:6) and PE (18:1/22:6) levels were significantly decreased.



Supplementary figure 4. Staining profiles of ADAR proteins in EC tissues. Panel A. Representative section showing that ADAR1 expression is not detected in mitotic cells (black arrow). **Panel B.** ADAR1 specific staining of endothelial cells (black arrow). The same pattern was observed for ADAR2 antibody (data not shown). (100X).

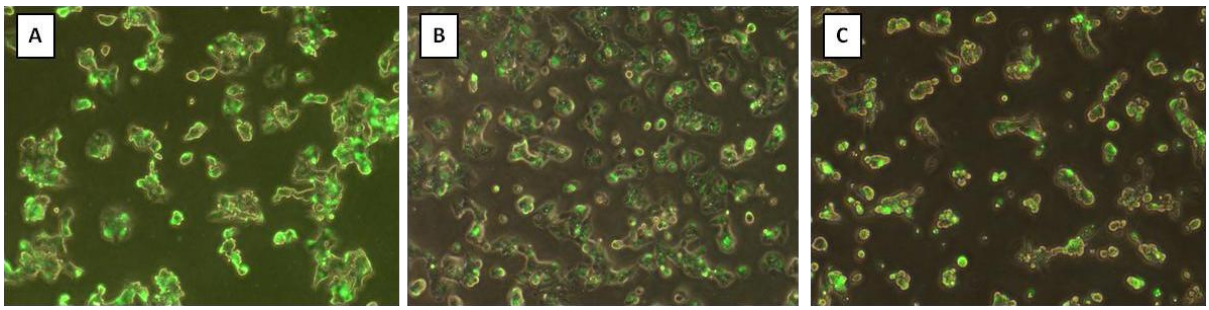


Supplementary figure 5. ROC curve analysis for IHQ data obtained for ADAR1 and ADAR2.



Supplementary figure 6. Representative graphic illustrating the transfection efficiency in three cell lines that was quantified as green fluorescence upon transfecting with siRNA-NC (**Panel A**) HEC-1A , (**Panel B**) Ishikawa and (**Panel C**) RL95-2 cells.

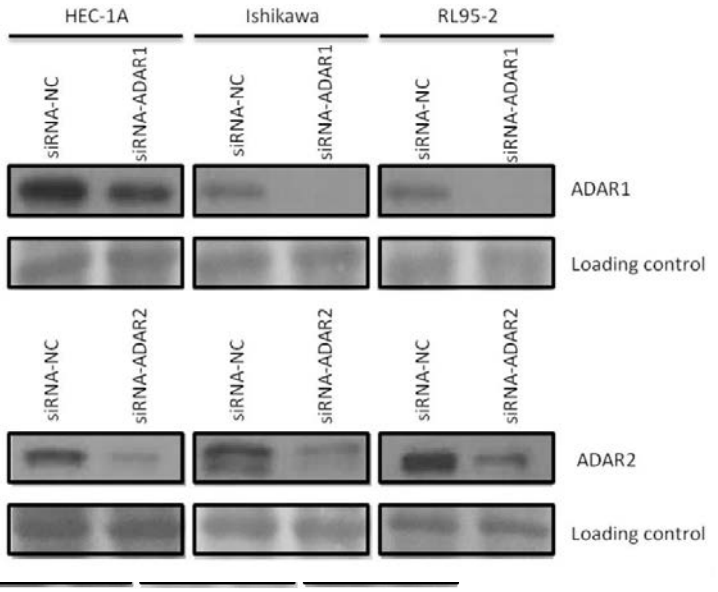
Supplementary figure 6. Representative graphic illustrating the transfection efficiency in three cell lines that was quantified as green fluorescence upon transfecting with siRNA-NC (**Panel A**) HEC-1A , (**Panel B**) Ishikawa and (**Panel C**) RL95-2 cells.



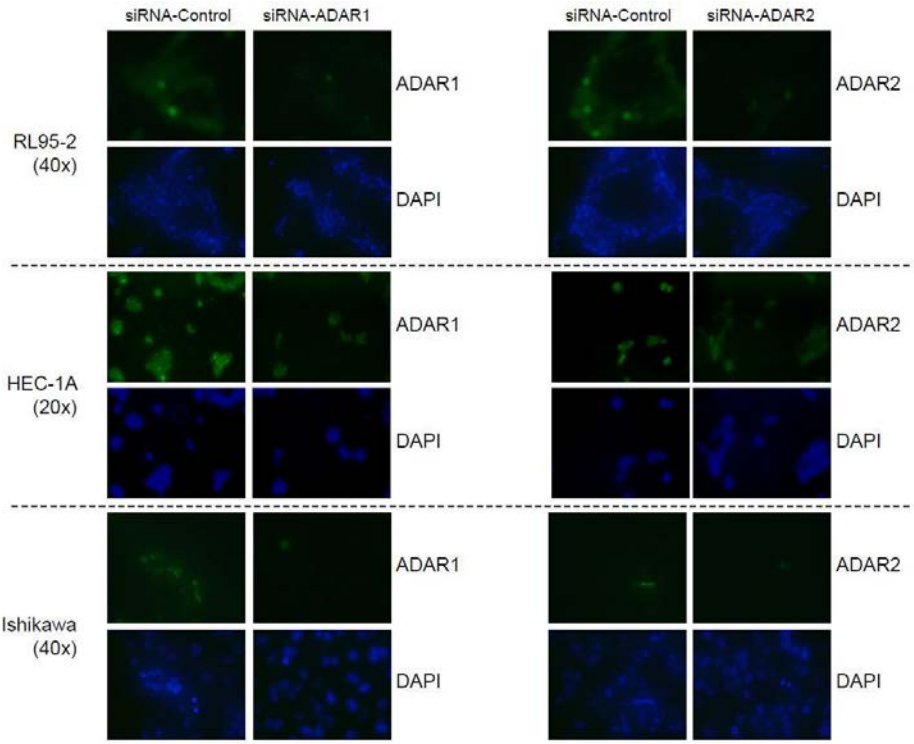
Supplementary figure 7. Panel A. Western Blot analysis showing the inhibition of ADARs protein expression. After 96 h of transfection of HEC-1A, Ishikawa and RL95-2 EC cell lines with siRNA-ADAR1, siRNA-ADAR2 or siRNA-NC, a decrease in protein abundance was observed in all cases. Naphtol blue staining of the membranes was used as a protein loading control. **Panel B. Immunofluorescence (IF) showing the decrease of ADARs protein expression**. IF was performed in the 3 cell lines after 96 h of transfection.

Supplementary figure 7. Panel A. Western Blot analysis showing the inhibition of ADARs protein expression. After 96 h of transfection of HEC-1A, Ishikawa and RL95-2 EC cell lines with siRNA-ADAR1, siRNA-ADAR2 or siRNA-NC, a decrease in protein abundance was observed in all cases. Naphtol blue staining of the membranes was used as a protein loading control. **Panel B. Immunofluorescence (IF) showing the decrease of ADARs protein expression**. IF was performed in the 3 cell lines after 96 h of transfection.

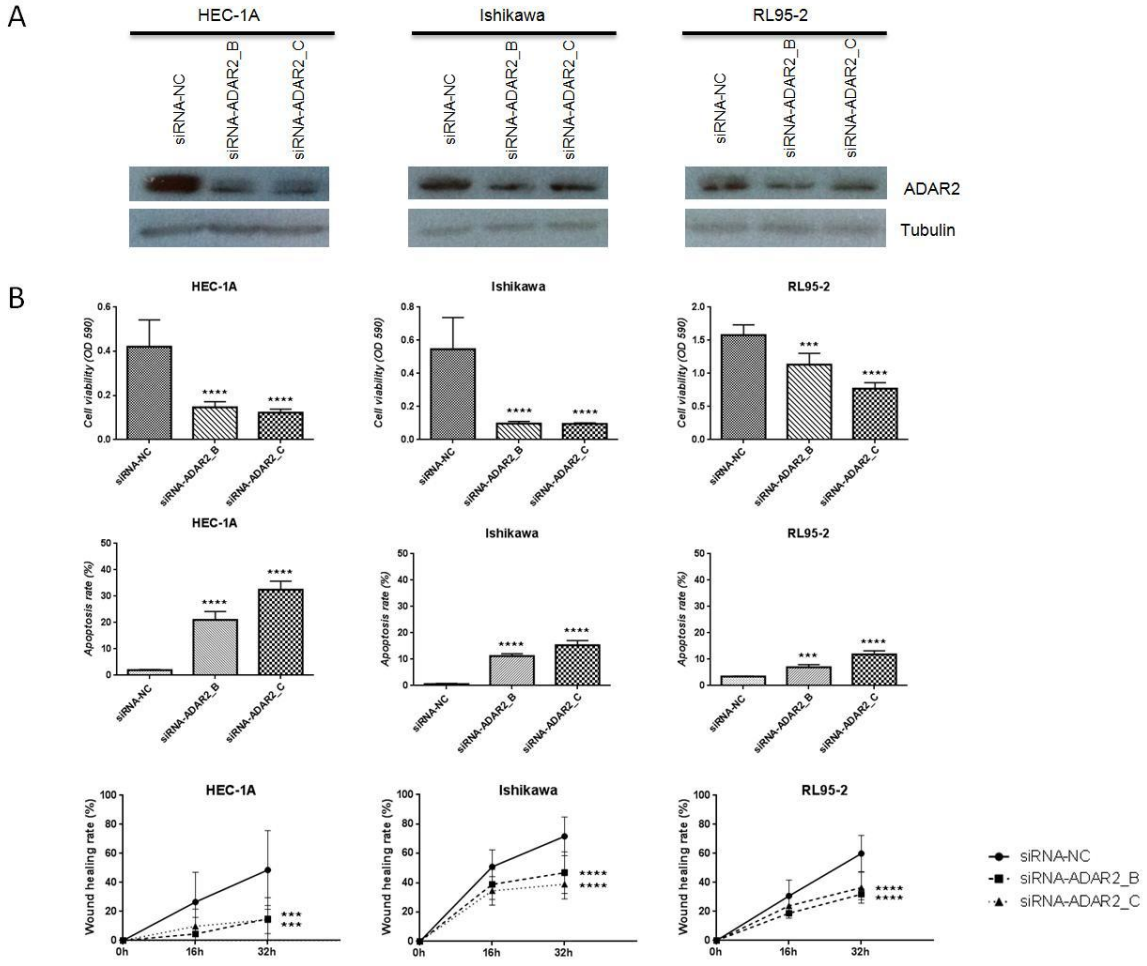
A



B



Supplementary figure 8. Panel A. Western Blot analysis showing the inhibition of ADAR2 protein expression. After 96 h of transfection of HEC-1A, Ishikawa and RL95-2 EC cell lines with siRNA-ADAR2_B, siRNA-ADAR2_C or siRNA-NC, a decrease in protein abundance was observed in all cases. Tubulin was used as a protein loading control. **Panel B. Functional assays revealing that ADAR2 presents oncogenic functions in vitro.** HEC-1A, Ishikawa and RL95-2 EC cell lines were used for the functional assays**.** Proliferation assay shows a significant decrease in cell viability (OD 590 nm) in the 3 cell lines when inhibiting ADAR2 expression with siRNA-ADAR2_B and siRNA-ADAR2_C. Apoptosis assay shows a significant increase in the apoptosis rate when silencing ADAR2 with both siRNAs. Wound healing assay indicating a significant decrease in the migration capabilities of the 3 EC cell lines (% of wound healing) when treating cells with siRNA-ADAR2_B and siRNA-ADAR2_C.



DISCUSSION

Cancer is an important health concern worldwide. In many developed countries, including Europe, cancer represents the second most common cause of death next to cardiovascular diseases. More than half of the patients diagnosed with cancer around the world succumb to it. Moreover, since the world population aging is increasing as well as the cancer associated lifestyle habits, cancer will promptly turn into the first cause of death in many parts of the world [120].

EC represents a set of tumors that are mostly generated in the endometrial glands of postmenopausal women. It is the most common gynecological malignancy in developed countries and the fourth most common cancer in women. EC is a gynecological disease that shows a continuous increasing incidence among older, but also younger, patients. Concerning this problematic, many new research lines focused on the **better understanding of the molecular changes associated to EC** have appeared the last decades.

Currently, endometrial biopsies are the most common surgical specimen used to diagnose EC and to assess the grade and histological type of the tumor. This preoperative information has a big impact on the therapeutic decision, however, the subjectivity of this analysis and the small amount of material contained in endometrial biopsies leads to an inaccurate prediction of the histological type and grade in many cases [121]. Therefore, an easy, quick, applicable, and accurate technique in order to properly diagnose EC and define the histological type and grade is urgently needed. For this purpose, several studies carried out during the last years based their efforts in **identifying novel tumor biomarkers in order to improve EC diagnosis, staging, prognosis and therapeutic response**. The translational impact that these findings will have into the clinics will not just facilitate the diagnosis and prognosis of the disease, but also provide new tools for individualized treatment of the patients [122,123].

In order to achieve the implementation of molecular tools in the clinical setting, new research projects have focused their efforts to the era of "big biology" and the integrative study of biological systems, also called "omics" technologies. From a recent past, various “omics” have been used for the research of cancer biomarkers. The “omics” approaches include the study of the entire genome, transcriptome, proteome and metabolome; thus, concentrating in the study of a

global view of an organism and not researching single genes, transcripts, proteins or metabolites. The integration of all “omics” information will permit a better understanding of human cancer and drive towards individualized treatments [124].

Over the past six decades the study of DNA, RNA and proteins has been the spotlight of cancer research. **Metabolomics has emerged as a potentially transformational avenue towards the identification of novel tumor biomarkers**. In this context, it is indispensable to comprehend changes that occur in the metabolome during tumor formation and cancer progression in order to provide new insights for the detection of disease biomarkers. The metabolome is defined by the complete collection of substrates and end products of the metabolism that drive essential cellular functions. These metabolites or low weight molecules reflect the physiological and pathological changes of a biological system [125]. Metabolomics-based biomarkers may be applied for disease diagnosis, patient stratification, drug discovery, targeted therapies efficacy assessment and monitoring the response to a specific treatment [126]. Despite its potential, **no metabolites have yet reached clinical EC diagnostic use neither therapeutic guidance application.**

Tissue-based tumor signatures are prone to bias since multiple sub-clonal populations of cells compete against each other in a tumor. Liquid biopsies have been reported to carry valuable information about tumor development with important implications for clinical oncology. They are being already used in research as samples for detecting tumor biomarkers and to monitor and improve treatment selection for the patient [37]. Although liquid biopsies provide a non-invasive substitute to solid biopsies, they have some limitations such us lack of consensus on the detection methods, small sample volume and diluted concentration of molecules, among others. Hence, no clinical sample can be established as a standard for a discovery project; depending on the biomarker scientific pipeline and the specific objectives of the study the optimum source of biomarkers will have to be chosen. For that reason, **the human sample type to use in the research for biomarkers and for the study of the dysregulated pathways in carcinogenesis still remains a challenge.**

In this concern, EVs and exosomes represent an innovative source of biomarkers containing specific cargo that resembles to their cells of origin. They can be easily isolated and studied as a mirror of genetic and signaling alterations of the parental cell. Moreover, they are present in most body fluids and their cargo stability is guaranteed since it is protected from enzymatic activity by a lipid bilayer. Besides, they alleviate the dynamic range issue that is a usual problem in molecular characterization studies based in body fluids, due to the limitation for detection of low abundant molecules which are hidden by the existence of high concentration molecules in the study sample [50].

In order to overcome the outlined clinical challenges, **the major goal of this thesis work was to identify a novel panel of EC diagnostic markers and to gain insights into altered metabolic pathways in the establishment and dissemination process of EC using metabolomic-based approaches in order to improve patient care and overcome the mortality rate of this malignancy.**

To achieve this objective, we divided the thesis work into three chapters each one of them associated to a specific objective: 1) Optimization of the extraction methods for the metabolomic analysis of extracellular vesicles (EVs) contained in human biofluids using liquid chromatography mass spectrometry (LC-MS) techniques; 2) Identification, verification, and validation of EC diagnostic biomarkers from human biofluids (use of plasma, uterine aspirates and EVs isolated from both biofluids); and 3) Metabolomic study of the EC initiation and dissemination processes at tissues level.

As this thesis is centered in the metabolomics study of EC, a first step consisted in setting up the methods to obtain high-quality data. The matrices that were analyzed along the thesis investigations were plasma, uterine aspirates, tissue biopsies and EVs isolated from human biofluids. The use of tissues, plasma samples, and endometrial aspirates for MS studies has been published by several groups [127–131] but not much was known about the EVs isolation and preparation for metabolomic and lipidomic profiling studies [132]. For this reason, in **Chapter 1- "Enabling metabolomics based biomarker discovery studies using molecular phenotyping of exosome-like vesicles (ELVs)"-** we

described methodologies for UPLC-ESI-MS based small molecule profiling of ELVs from biological samples, including plasma and cell culture media. We chose those matrices for optimizing metabolomics methodologies since those are used in our work, but also are widely used in other clinical and basic research studies. As a result of this work, we could confirm that the isolation and enrichment of ELVs from plasma and cell culture media was successful and we were able to characterize the size and concentration of the vesicles. Moreover, since procuring human plasma samples from biorepositories is often limited by the cost and sample volume, we determined the minimum volume of plasma needed to generate optimal MS data when profiling the metabolome of the ELVs, this was 500 µl. A set of metabolites that were found in plasma and cell culture media ELVs were also validated confirming the biochemical composition of these vesicles. Additionally, we highlighted the importance of enriching the ELVs fraction in contraposition to the use of the non-fractionated biofluid to achieve the discovery of low abundant metabolites. The use of ELVs allows the detection and identification of many metabolites that fall below the instrumental limit of detection when analyzing the non-fractionated matrix. Besides, we demonstrated the utility of ELVs in MS-based research in order to reveal changes in pathological situations compared to the control condition and the possibility to use targeted MS approaches to analyze the low abundant metabolites as biomarkers of this pathological situation. These findings describe a standardized methodology for MS based metabolomics profiling of ELVs and evidenced that the enriched cargo contained in ELVs offers new opportunities for the discovery of low abundant metabolites as disease biomarkers. This approach could be easily translated to other fields of biomarker research.

Hence, once we standardized the protocol and conditions to be used for our research, in **Chapter 2- "Integrative analysis of the lipidome and metabolome of biofluids and extracellular vesicles yields biosignatures to diagnose endometrial cancer"-** we aimed to identify a diagnostic biomarker panel by studying the metabolic and lipidomic profile of 4 different human biofluids: plasma, uterine aspirates (UA), extracellular vesicles derived from plasma (P EVs) and extracellular vesicles derived from uterine aspirates (UA EVs). As mentioned previously, many efforts and investments have been made

to find diagnostic cancer biomarkers over the last decade without success. In order to overcome this clinical limitation, here we present, for the first time, a discovery mode study that combined 4 different matrices derived from human body fluids. We developed a biomarker workflow starting with an untargeted discovery MS phase (global LC-MS) and verification phase (MS/MS) that will be soon followed by a validation phase using an independent cohort of plasma samples (targeted MRM).

The collection of novel strategies followed in this work was expected to result in advances in the clinics of EC diagnosis since: a) The integration of the profiling information obtained from 4 different sample types for each patient included in the discovery phase of the study allows a much better understanding of EC and offers higher opportunities to identify EC specific biomarkers; b) The use of proximal biofluids as source of biomarkers, in this case UAs, has a particular interest since a part from being a sample that is obtained by a minimally invasive collection method, it is in direct contact with the affected endometrium. It has been already proved that UAs reflect better EC changes since they are more enriched in molecules that are directly secreted by the diseased cells [99]; c) isolated EVs from human biofluids are contemplated as a promising source of biomarkers since their cargo is protected from degradation and, more importantly, they may help solve the dynamic-range problem of the abundant molecules present in human biofluids that often mask the detection of less abundant metabolites that are usually the most promising candidates for diagnostic purposes [105]; d) The untargeted profiling of this research line was carried out analyzing 4 different matrices in order to increase our chances to detect potential diagnostic biomarkers. However, the validation of the discovered biomarkers will be analyzed by targeted metabolomics only in plasma samples, aiming to achieve the most simple, quick and non-invasive diagnostic molecular signature. Since the MS-platforms used for discovery purposes allow the detection of a large number of metabolites but just those present at a minimum over a minimum concentration threshold, we hypothesized that this approximation would have allowed the identification of very low abundant biomarkers (overcoming the dynamic-range problem with the use of UAs and EVs) that would be much more specific for EC. These

biomarkers would be in a next step detectable in plasma samples since the targeted platform that will be used in the validation phase will allow the quantification of known biomarkers even when they are found at a very low concentration.

In our study, we observed that plasma, plasma EVs, UA and UA EVs are excellent matrices for the discovery of specific and sensible biomarkers. The plasma based biomarker study yielded a panel with the highest specificity and sensibility; we generated a panel of 13 metabolites that gave an AUC of 0.986. Moreover, we generated a panel of 7 biomarkers with an AUC of 0.933 for UA; a panel of 8 biomarkers with an AUC of 0.936 for the UA EVs; and a panel of 4 metabolites with an AUC of 0.767 for plasma EVs. For UA EVs, we observed the best separation between the control and EC group in the PCA plots and also in the heat maps. However, the results were very disparate among matrices. In fact, just five common features were significantly dysregulated in the four matrices, but no possible annotation for them was found. Additionally, the metabolites verified in UA and UA EVs do not align with the ones confirmed in plasma and plasma EVs probably because the expression of some of the biomarkers was lost in circulation compared to their higher concentration in UA. In UA and UA EVs we validated a large number of peptides while in plasma and P EVs the metabolites most abundant were glycerophospholipids. P EVs showed several markers of interest, however, many of these were not verified since no possible annotation for them was found. Thus, although P EVs remains an untapped resource for low abundance biomarkers, the findings are difficult to translate for clinical classification since many of these metabolites are still unknown. Presumably, the development of EC diagnostic biomarker panels needs to be initiated and validated in the same matrices.

Hence, our results show the value of using a high throughput metabolomics approach for delineating biomarker panels that can be used to diagnose EC in bodyfluids. These biomarker panels as standalone tests or in conjunction with existing clinical methods can be used for patient management with high accuracy. However, the ultimate transition of these biomarker panels for clinical use will require validation with large and diverse cohorts and prioritization and updating these panels to achieve higher specificity and sensitivity in order to

minimize false positives. As a future step of our project, we will perform a validation phase in plasma samples using a targeted multiple reaction monitoring (MRM) approach in order to evaluate the biomarker panel specificity for EC diagnosis.

Finally, in **Chapter 3- "Metabolomic and lipidomic profiling identifies the role of the RNA editing pathway in endometrial carcinogenesis"-** we focused our efforts on the study of the alterations that metabolic pathways suffer in the tumor tissue during EC carcinogenesis and dissemination. As previously mentioned in this thesis, several mass spectrometry proteomics studies have been performed in the EC field [131,133–137] but just one untargeted study interrogating the metabolomic profile of EC has been published up to date [91]. In our approach, we used a larger set of samples for the discovery and validation phase compared to the mentioned study. Moreover, we did not just validate our findings by IHQ but we also performed functional assays. Hence, we used a global metabolomic profiling approach to analyze a set of samples including 39 EC tissue specimens and 17 matched controls that revealed significant changes of about 80 metabolites in EC. A subset of 42 metabolites was validated by MS/MS revealing the dysregulation of several pathways in EC. We found lipid anabolism and catabolism altered in EC, with the upregulation of several PCs, PEs and PIs in the tumor tissue. Imbalances in the acyl analogs of endocannabinoids concentration and in the end products of the kynurenine pathway were also observed in the tumor. Some of these abnormalities have been already described in EC for other scientific groups; Trousil et al. performed a targeted metabolomic study in EC tissues and described metabolic alterations in EC including significant increased levels of cholines and lipids [87]. Also, Jové et al. described the endocannabinoid and purine metabolism to be involved in EC pathogenesis and myometrial invasion in endometrial tissues [91]. The role of the endocannabinoid system in gynecological cancer has become also focus of research in the last years by some investigators [138,139]. Moreover, we described significant changes in tumor progression such as alterations in the lipidome, and increased concentration values of UDP-N-acetyl-D-galactosamine and arachidonic acid at advanced EC stages. Aberrant cell-surface glycosylation

has been reported to happen in many human cancers; and targeting cancer-associated glycans may be a future anticancer therapeutic strategy [140,141].

A dysregulation of nucleoside inosine in EC was also observed in our discovery and could be an indicative of an imbalance in the adenosine versus inosine ratio (A/I), as reported also by Trousil et al. The underlying effect of alterations of the endogenous levels of inosine had never been deeply investigated in EC. Dysregulated levels of A/I ratio can be attributed to alterations in the adenosine to inosine enzymes of the RNA editing pathway, known as ADAR family of enzymes. Therefore, our group decided to focus on the study of the RNA editing pathway and the ADAR family of enzymes in EC. The complete biological connotation of ADARs function is still widely unknown. It has been reported the existence of aberrant ADAR activity in a variety of diseases including cancer [142,143], but no studies of ADARs in EC have been previously performed. Hence, we described for the first time that ADAR1 and ADAR2 enzymes show significant increased concentrations in EC tissue and that their levels increase progressively with the tumor grade and malignancy. Additionally, *in vitro* assays performed in 3 EC cell lines (HEC-1A, Ishikawa, RL95-2) confirmed that a silencing of ADAR2 expression resulted in significant changes in the oncogenic capabilities and in the aggressiveness behavior of EC cells lines including a drastic decrease of cell viability and migration potential and an increase in the apoptosis rate. We did not observe significant differences in the functional assays when silencing ADAR1. On the whole, our investigation proposes that ADAR2 acts as an oncogene in EC carcinogenesis and could be a potential target for improving therapy strategies. On that account, the use of mass spectrometry profiling approaches for the understanding of biochemical perturbation and metabolomic pathway imbalances of a human disease may have huge clinical translational relevance in furthering the personalized medicine leadership and improving clinical outcomes. Moreover, underscoring ADARs physiological and pathological role may provide mechanistic insights in the development of novel antitumor therapies.

In summary, in this thesis work we describe standardized metabolomic and lipidomic MS approaches that can have big impact in translational clinics and improving the outcome of EC patients, and also can be translated to other fields

of research. Importantly, we highlighted the importance of the use of proximal biofluids, specifically UA and EVs, in biomarker research and we opened a new avenue for the identification of sensitive and specific EC biomarkers in bodyfluids. Moreover, the data presented in this dissertation depicts a significant advance in the understanding of the metabolic alterations that take place in EC tumorigenesis. Apart from validating different metabolomics pathways already reported in the EC carcinogenic process, our results and *in vitro* models evidenced a novel oncogenic metabolomic pathway in EC, the RNA editing pathway. This suggests that a deeper inspection of the ADARs role in EC may represent an attractive strategy for the development of new therapeutic approaches.

CONCLUSIONS

The main conclusions derived from this thesis are:

1. LC-MS is an ideal platform for the metabolomic profiling of EVs isolated from human bodyfluids and cell culture media for studies with basic science, clinical or translational focus.

2. The minimum starting volume of plasma for EVs isolation to produce high quality metabolomic and lipidomic data is 500 µl.

3. Plasma, plasma EVs, UAs and UA EVs are excellent matrices for the identification of new disease biomarkers. A total of 45, 9, 15 and 8 EC markers were verified in each matrix, respectively.

4. Our approach demonstrated that high throughput metabolomics can be used for delineating EC diagnostic biomarker panels. The plasma based biomarker study yielded a panel of 13 metabolites with the highest specificity and sensibility (AUC of 0.986). Additional ROC analysis generated a 7 metabolite panel with an AUC of 0.933 for UAs; an 8 metabolite panel with an AUC of 0.936 for UA EVs; and a 4 metabolite panel with an AUC of 0.767 for plasma EVs. These panels should be now validated in a larger cohort of patients.

5. In an overall perspective, the metabolites identified and verified in UA-based matrices were discordant with the ones identified and verified in plasma-based matrices.

6. EVs derived from plasma showed several markers of interest in the discovery phase, however, many of these were not verified since they could not be annotated. Thus, although EVs derived from plasma remains an untapped resource for low abundance biomarkers, the findings are difficult to translate for clinical classification since many of these metabolites are still unknown.

7. The metaboloprofiling of EC and healthy tissue elucidated the dysregulation of the kynurenine pathway, the endocannabinoids signaling pathway, the lipid metabolism and the RNA editing pathway. These metabolomic alterations are associated to EC initiation and progression.

8. The immunohistochemical analysis on 183 tissue samples permitted to decipher that ADAR1 and ADAR2 are overexpressed in endometrial cancer in a manner that their expression correlates with the tumor histological type and grade. As increased is their expression, poorer histological type and grade presents the patient.

9. Silencing of ADAR2, but not ADAR1, in three endometrial cancer cell lines resulted in decreased proliferation rate, increased apoptosis, and reduced migration capabilities *in vitro*.

10. This study demonstrated that ADAR2 functions as an oncogene in endometrial carcinogenesis and could be a potential target for improving EC treatment startegies.

11. The data presented in this dissertation depicts a significant advance in the understanding of the metabolic and lipidomic alterations that take place in EC tumorigenesis and can have big impact in translational clinics.

LIST OF PUBLICATIONS

This thesis resulted in the publication of several manuscripts. Three of them are included as parts of the thesis, but other articles were published thanks to the collaboration in other research projects:

Publication A) Altadill T, Dowdy TM, Gill K, Reques A, Menon SS, Moiola CP, Lopez-Gil C, Coll E, Matias-Guiu X, Cabrera S, Garcia A, Reventos J, Byers SW, Gil-Moreno A, Cheema AK, Colas E. “Metabolomic and Lipidomic Profiling Identifies The Role of the RNA Editing Pathway in Endometrial Carcinogenesis”, Sci Rep. 2017 Aug 18;7(1):8803.

Publication B) Campoy I, Lanau L, **Altadill T**, Sequeiros T, Cabrera S, Cubo-Abert M, Pérez-Benavente A, Garcia A, Borrós S, Santamaria A, Ponce J, Matias-Guiu X, Reventós J, Gil-Moreno A, Rigau M, Colas E. “Exosome-like vesicles in uterine aspirates: a comparison of ultracentrifugation-based isolation protocols”, J Transl Med. 2016 Jun 18;14(1):180.

Publication C) Altadill T, Campoy I, Lanau L, Gill K, Rigau M, Gil-Moreno A, Reventos J, Byers S, Colas E, Cheema AK. “Enabling Metabolomics Based Biomarker Discovery Studies Using Molecular Phenotyping of Exosome-Like Vesicles”, PLoS One. 2016 Mar 14;11(3):e0151339.

Publication D) Llauradó M, Majem B, **Altadill T**, Lanau L, Castellví J, Sánchez-Iglesias JL, Cabrera S, De la Torre J, Díaz-Feijoo B, Pérez-Benavente A, Colás E, Olivan M, Doll A, Alameda F, Matias-Guiu X, Moreno-Bueno G, Carey MS, Del Campo JM, Gil-Moreno A, Reventós J, Rigau M. “MicroRNAs as prognostic markers in ovarian cáncer”, Mol Cell Endocrinol. 2014 Jun 5;390(1-2):73-84.

Publication E) Almacellas E, **Altadill T**, Escolano JC, Quesada López TP, Robles D, Sánchez-Úbeda S, Vázquez-Bernat N, Pulido M.” Unclear instructions to authors for length of original articles on the Journal's website”, Int J Stroke. 2014 Apr;9(3):E10.

REFERENCES

1. Ralph, T. & Carvajal, J. Manual de Obstetricia y Ginecología. *Man. Obstet. Y Ginecol.* 252 (2012).
2. Tortora, G. J. & Derrickson, B. *Principles of Anatomy and Physiology*. *Wiley.* (2014).
3. Mutter, G. L. Diagnosis of premalignant endometrial disease. *J. Clin. Pathol.* **55,** 326–31 (2002).
4. Hoffman, Barbara L; Schorge, J. O. in *Williams gynecology.* 969 (2012).
5. Yang, Y. F. *et al.* Prediction of coexistent carcinomas risks by subjective EIN diagnosis and comparison with WHO classification in endometrial hyperplasias. *Pathol. Res. Pract.* **208,** 708–712 (2012).
6. Li, X. C. & Song, W. J. Endometrial intraepithelial neoplasia (EIN) in endometrial biopsy specimens categorized by the 1994 world health organization classification for endometrial hyperplasia. *Asian Pacific J. Cancer Prev.* **14,** 5935–5939 (2013).
7. Gregoriou, O. *et al.* Clinical parameters linked with malignancy in endometrial polyps. *Climacteric.* **12,** 454–458 (2009).
8. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2016. *CA Cancer J Clin.* **66,** 7–30 (2016).
9. Di Cristofano, A. & Ellenson, L. H. Endometrial carcinoma. *Annu. Rev. Pathol.* **2,** 57–85 (2007).
10. Sonoda, Y. & Barakat, R. R. Screening and the prevention of gynecologic cancer: endometrial cancer. *Best Pract. Res. Clin. Obstet. Gynaecol.* **20,** 363–77 (2006).
11. Timmermans, A. *et al.* Endometrial Thickness Measurement for Detecting Endometrial Cancer in Women With Postmenopausal Bleeding A Systematic Review and Meta-Analysis. *Am. J. Obstet. Gynecol.* **116,** 160–167 (2010).
12. Denschlag, D., Ulrich, U. & Emons, G. The diagnosis and treatment of endometrial cancer: progress and controversies. *Dtsch. Arztebl. Int.* **108,** 571–7 (2010).
13. Amant, F. *et al.* Endometrial cancer. *Lancet (London, England).* **366,** 491–505 (2005).
14. Burke, W. M. *et al.* Endometrial cancer: A review and current management strategies: Part i. *Gynecologic Oncology.* **134,** 385–392

(2014).

15. Bokhman, J. V. Two pathogenetic types of endometrial carcinoma. *Gynecol. Oncol.* **15,** 10–17 (1983).
16. American Cancer Society. ¿Qué es el cáncer de endometrio? *www.cancer.org.* 5–6 (2016).
17. Bergeron, C. Histología y fisiología del endometrio normal. *EMC - Ginecol.* **42,** 1–8 (2006).
18. Kurman, R. J., Carcangiu, M. L., Herrington, C. S. & Young, R. H. *Pathology and Genetics of Tumours of the Breast and Female Genital Organs. World Health Organization Classification of Tumours* (2014).
19. Tavassoéli, F. & Devilee, P. *Tumours of the Breast and Female Genital Organs*. *Pathology and Genetics* (2003).
20. Gatius, S. & Matias-Guiu, X. Practical issues in the diagnosis of serous carcinoma of the endometrium. *Mod. Pathol.* **29,** S45–S58 (2016).
21. Chen, J. *et al.* Small cell carcinoma of the endometrium: a clinicopathological and immunohistochemical study. *Int. J. Clin. Exp. Pathol.* **7,** 8869–74 (2014).
22. Beddy, P. *et al.* FIGO Staging System for Endometrial Cancer: Added Benefits of MR Imaging. *Radiographics.* **32,** 241–254 (2012).
23. Lewin, S. N. Revised FIGO staging system for endometrial cancer. *Clin. Obstet. Gynecol.* **54,** 215–218 (2011).
24. Werner, H. M. J. *et al.* Revision of FIGO surgical staging in 2009 for endometrial cancer validates to improve risk stratification. *Gynecol. Oncol.* **125,** 103–108 (2012).
25. Pecorelli, S. Corrigendum to 'Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium'. *International Journal of Gynecology and Obstetrics.* **108,** 176 (2010).
26. Cancer Genome Atlas Research Network. Integrated genomic characterization of endometrial carcinoma. *Nature.* **497,** 67–73 (2013).
27. Vidal, F. & Rafii, A. Lymph node assessment in endometrial cancer: towards personalized medicine. *Obstet. Gynecol. Int.* **2013,** 1–8 (2013).
28. Wright, J. D., Medel, N. I. B., Sehouli, J., Fujiwara, K. & Herzog, T. J. Contemporary management of endometrial cancer. *The Lancet.* **379,** 1352–1360 (2012).

29. Colombo, N. *et al.* ESMO-ESGO-ESTRO Consensus Conference on Endometrial Cancer: diagnosis, treatment and follow-up. *Ann. Oncol.* **27,** 16–41 (2016).
30. Plataniotis, G. & Castiglione, M. Endometrial cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **21,** (2010).
31. Colombo, N. *et al.* Endometrial cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **24,** (2013).
32. Strimbu, K. & Tavel, J. a. What are Biomarkers? *Curr Opin HIV AIDS.* **5,** 463–466 (2011).
33. Werner, H. M. J. & Salvesen, H. B. Current status of molecular biomarkers in endometrial cancer. *Current Oncology Reports.* **16,** (2014).
34. Żyła, M. M. *et al.* Review paper The significance of markers in the diagnosis of endometrial cancer. **15,** 176–185 (2016).
35. Martinez-Garcia, E. *et al.* Advances in endometrial cancer protein biomarkers for use in the clinic. *Expert Rev. Proteomics.* 1–19 (2017).
36. Mota, A. *et al.* Genetic analysis of uterine aspirates improves the diagnostic value and captures the intra-tumor heterogeneity of endometrial cancers. *Mod. Pathol.* (2017).
37. Siravegna, G., Marsoni, S., Siena, S. & Bardelli, A. Integrating liquid biopsies into the management of cancer. *Nat. Rev. Clin. Oncol.* (2017).
38. Colas, E. *et al.* Molecular markers of endometrial carcinoma detected in uterine aspirates. *Int. J. Cancer.* **129,** 2435–2444 (2011).
39. Niklasson, O., Skude, G., Marsal, K. & Casslén, B. Diagnosis of endometrial cancer in patients with postmenopausal bleeding by analysis of the lactate dehydrogenase isoenzyme activity profile in uterine fluid. *Gynecol. Oncol.* **93,** 385–389 (2004).
40. Niklasson, O., Skude, G., Lingman, G., Casslén, B. & Maršál, K. Transvaginal ultrasound and lactate dehydrogenase isoenzyme activity profile in uterine aspirate for diagnosis of endometrial carcinoma in women with postmenopausal bleeding. *Int. J. Gynecol. Cancer.* **17,** 1322–1326 (2007).
41. Casado-vela, J. *et al.* Comprehensive Proteomic Analysis of Human Endometrial Fluid Aspirate research articles. *J. Proteome Res.* 4622–

4632 (2009).

42. Mahovlic, V. *et al.* Digital morphometry of cytologic aspirate endometrial samples. *Coll. Antropol.* **34,** 45–51 (2010).
43. Martinez-Garcia, E. *et al.* Development of a sequential workflow based on LC-PRM for the verification of endometrial cancer protein biomarkers in uterine aspirate samples. *Oncotarget.* **7,** 53102–53115 (2016).
44. Martinez-Garcia, E. *et al.* Targeted Proteomics Identifies Proteomic Signatures in Liquid Biopsies of the Endometrium to Diagnose Endometrial Cancer and Assist in the Prediction of the Optimal Surgical Treatment. *Clin. Cancer Res.* **23,** 6458–6467 (2017).
45. Zha, Q. Bin, Yao, Y. F., Ren, Z. J., Li, X. J. & Tang, J. H. Extracellular vesicles: An overview of biogenesis, function, and role in breast cancer. *Tumor Biol.* **39,** (2017).
46. Steinbichler, T. B., Dudás, J. & Riechelmann, H. The Role of Exosomes in Cancer Metastasis. *Semin. Cancer Biol.* (2017).
47. Lobb, R. J., Lima, L. G. & Möller, A. Exosomes: Key mediators of metastasis and pre-metastatic niche formation. *Semin. Cell Dev. Biol.* 1–8 (2017).
48. Gopal, S. K. *et al.* Extracellular vesicles: their role in cancer biology and epithelial–mesenchymal transition. *Biochem. J.* **474,** 21–45 (2016).
49. Choi, D. *et al.* Extracellular vesicle communication pathways as regulatory targets of oncogenic transformation. *Semin. Cell Dev. Biol.* 1–12 (2017).
50. Soung, Y., Ford, S., Zhang, V. & Chung, J. Exosomes in Cancer Diagnostics. *Cancers (Basel).* **9,** 8 (2017).
51. Li, W. *et al.* Role of exosomal proteins in cancer diagnosis. *Mol. Cancer.* **16,** 145 (2017).
52. Gomase V.S., Changbhale, S.S., Patil, S.A., Kale, K. V. Metabolomics. *Curr. Drug Metab.* **9,** 89–98 (2008).
53. Roessner, U. & Bowne, J. What is metabolomics all about? *BioTechniques.* **46,** 363–365 (2009).
54. Patel, S. & Ahmed, S. Emerging field of metabolomics: Big promise for cancer biomarker identification and drug discovery. *Journal of Pharmaceutical and Biomedical Analysis.* **107,** 63–74 (2015).
55. Goodacre, R., Vaidyanathan, S., Dunn, W. B., Harrigan, G. G. & Kell, D.

B. Metabolomics by numbers: Acquiring and understanding global metabolite data. *Trends in Biotechnology.* **22,** 245–252 (2004).

56. Johnson, C. H., Patterson, A. D., Idle, J. R. & Gonzalez, F. J. Xenobiotic metabolomics: major impact on the metabolome. *Annu. Rev. Pharmacol. Toxicol.* **52,** 37–56 (2012).
57. Peng, B., Li, H. & Peng, X. X. Functional metabolomics: from biomarker discovery to metabolome reprogramming. *Protein and Cell.* **6,** 628–637 (2015).
58. Holmes, E., Wilson, I. D. & Nicholson, J. K. Metabolic Phenotyping in Health and Disease. *Cell.* **134,** 714–717 (2008).
59. Johnson, C. H., Ivanisevic, J. & Siuzdak, G. Metabolomics: beyond biomarkers and towards mechanisms. *Nat. Rev. Mol. Cell Biol.* **17,** 451–9 (2016).
60. Putri, S. P. *et al.* Current metabolomics: Practical applications. *Journal of Bioscience and Bioengineering.* **115,** 579–589 (2013).
61. Yanes, O., Tautenhahn, R., Patti, G. J. & Siuzdak, G. Expanding coverage of the metabolome for global metabolite profiling. *Anal. Chem.* **83,** 2152–2161 (2011).
62. Kind, T. & Fiehn, O. Advances in structure elucidation of small molecules using mass spectrometry. *Bioanalytical Reviews.* **2,** 23–60 (2010).
63. Cajka, T. & Fiehn, O. Toward Merging Untargeted and Targeted Methods in Mass Spectrometry-Based Metabolomics and Lipidomics. *Analytical Chemistry.* **88,** 524–545 (2016).
64. Monteiro, M. S., Carvalho, M., Bastos, M. L. & Guedes de Pinho, P. Metabolomics analysis for biomarker discovery: advances and challenges. *Curr. Med. Chem.* **20,** 257–71 (2013).
65. Davis, V. W., Bathe, O. F., Schiller, D. E., Slupsky, C. M. & Sawyer, M. B. Metabolomics and surgical oncology: Potential role for small molecule biomarkers. *Journal of Surgical Oncology* **103,** 451–459 (2011).
66. Nagana Gowda, A. & Raftery, D. Biomarker Discovery and Translation in Metabolomics. *Curr. Metabolomics.* **1,** 227–240 (2013).
67. Dunn, W. B. & Ellis, D. I. Metabolomics: Current analytical platforms and methodologies. *TrAC - Trends Anal. Chem.* **24,** 285–294 (2005).
68. Mathew, A. K. & Padmanaban, V. C. Metabolomics: The apogee of the

omics trilogy. *International Journal of Pharmacy and Pharmaceutical Sciences.* **5,** 45–48 (2013).

69. Want, E. J., Nordström, A., Morita, H. & Siuzdak, G. From exogenous to endogenous: The inevitable imprint of mass spectrometry in metabolomics. *Journal of Proteome Research.* **6,** 459–468 (2007).
70. Wang, J. H., Byun, J. & Pennathur, S. Analytical approaches to metabolomics and applications to systems biology. *Semin. Nephrol.* **30,** 500–511 (2010).
71. Siuzdak, G. The Expanding Role of Mass Spectrometry in Biotechnology. *Expand. role mass Spectrom. Biotechnol.* **15,** 625 (2006).
72. Katajamaa, M. & Orešič, M. Data processing for mass spectrometry-based metabolomics. *Journal of Chromatography A.* **1158,** 318–328 (2007).
73. Boccard, J., Veuthey, J. L. & Rudaz, S. Knowledge discovery in metabolomics: An overview of MS data handling. *Journal of Separation Science.* **33,** 290–304 (2010).
74. Nunes de Paiva, M. J., Menezes, H. C. & de Lourdes Cardeal, Z. Sampling and analysis of metabolomes in biological fluids. *Anal.* **139,** 3683–94 (2014).
75. Zhang, T. & Watson, D. G. A short review of applications of liquid chromatography mass spectrometry based metabolomics techniques to the analysis of human urine. *Anal.* **140,** 2907–15 (2015).
76. Weckwerth, W. & Morgenthal, K. Metabolomics: From pattern recognition to biological interpretation. *Drug Discovery Today.* **10,** 1551–1558 (2005).
77. Wishart, D. S. *et al.* HMDB: A knowledgebase for the human metabolome. *Nucleic Acids Res.* **37,** (2009).
78. Lu, W., Bennett, B. D. & Rabinowitz, J. D. Analytical strategies for LC – MS-based targeted metabolomics. *J Chromatogr B Anal. Technol Biomed Life Sci.* **871,** 236–242 (2008).
79. Tsugawa, H. *et al.* MRM-DIFF: Data processing strategy for differential analysis in large scale MRM-based lipidomics studies. *Front. Genet.* **5,** 1–15 (2015).
80. Zhou, J. & Yin, Y. Strategies for large-scale targeted metabolomics quantification by liquid chromatography-mass spectrometry. *Anal.* **141,**

6362–6373 (2016).

81. Noyes, R. W., Hertig, A. T. & Rock, J. Dating the endometrial biopsy. *Am J Obstet Gynecol* **122,** 262–263 (1975).
82. Altmae, S. *et al.* Guidelines for the design, analysis and interpretation of 'omics' data: Focus on human endometrium. *Hum. Reprod. Updat.* **20,** 12–28 (2014).
83. Vouk, K. *et al.* Discovery of phosphatidylcholines and sphingomyelins as biomarkers for ovarian endometriosis. *Hum. Reprod.* **27,** 2955–2965 (2012).
84. Vilella, F., Ramirez, L. B. & Simón, C. Lipidomics as an emerging tool to predict endometrial receptivity. in *Fertility and Sterility.* **99,** 1100–1106 (2013).
85. Knapp, P. *et al.* Altered sphingolipid metabolism in human endometrial cancer. *Prostaglandins Other Lipid Mediat.* **92,** 62–6 (2010).
86. Bufa, A. *et al.* Altered urinary profiles of endogenous steroids in postmenopausal women with adenocarcinoma endometrii. *Gynecol. Endocrinol.* **3590,** (2010).
87. Trousil, S. *et al.* Alterations of choline phospholipid metabolism in endometrial cancer are caused by choline kinase alpha overexpression and a hyperactivated deacylation pathway. *Cancer Res.* **74,** 6867–77 (2014).
88. Urakami, K., Zangiacomi, V., Yamaguchi, K. & Kusuhara, M. Impact of 2-deoxy-D-glucose on the target metabolome profile of a human endometrial cancer cell line. *Biomed. Res.* **34,** 221–229 (2013).
89. Schuler, K. M. *et al.* Antiproliferative and metabolic effects of metformin in a preoperative window clinical trial for endometrial cancer. *Cancer Med.* **4,** 161–173 (2015).
90. Shao, X. *et al.* Screening and verifying endometrial carcinoma diagnostic biomarkers based on a urine metabolomic profiling study using UPLC-Q-TOF/MS. *Clin. Chim. Acta.* **463,** 200–206 (2016).
91. Jové, M. *et al.* Metabotyping human endometrioid endometrial adenocarcinoma reveals an implication of endocannabinoid metabolism. *Oncotarget.* **7,** 52364–74 (2016).
92. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2017. *CA. Cancer*

J. Clin. **67,** 7–30 (2017).

93. Bourgin, C. *et al.* Endometrial cancer in elderly women: Which disease, which surgical management? A systematic review of the literature. *European Journal of Surgical Oncology.* **42,** 166–175 (2016).
94. Gadducci, A., Cosio, S., Carpi, A., Nicolini, A. & Genazzani, A. R. Serum tumor markers in the management of ovarian, endometrial and cervical cancer. *Biomed. Pharmacother.* **58,** 24–38 (2004).
95. Xu, X. & Veenstra, T. D. Analysis of biofluids for biomarker research. *Proteomics - Clin. Appl.* **2,** 1403–1412 (2008).
96. Campoy, I. *et al.* Exosome-like vesicles in uterine aspirates: a comparison of ultracentrifugation-based isolation protocols. *J. Transl. Med.* **14,** 180 (2016).
97. Perez-Sanchez, C. *et al.* Molecular diagnosis of endometrial cancer from uterine aspirates. *Int. J. Cancer.* **133,** (2013).
98. Esteller, M., García, A., Martínez-Palones, J. M., Xercavins, J. & Reventós, J. Detection of clonality and genetic alterations in endometrial pipelle biopsy and its surgical specimen counterpart. *Lab. Invest.* **76,** 109–16 (1997).
99. Mota, A. *et al.* Genetic analysis of uterine aspirates improves the diagnostic value and captures the intra-tumor heterogeneity of endometrial cancers. *Mod. Pathol.* **30,** 134–145 (2017).
100. Abels, E. R. & Breakefield, X. O. Introduction to Extracellular Vesicles: Biogenesis, RNA Cargo Selection, Content, Release, and Uptake. *Cell. Mol. Neurobiol.* **36,** 301–12 (2016).
101. Pan, B. T. & Johnstone, R. M. Fate of the transferrin receptor during maturation of sheep reticulocytes in vitro: selective externalization of the receptor. *Cell* **33,** 967–78 (1983).
102. Choi, H. & Lee, D. S. Illuminating the physiology of extracellular vesicles. *Stem Cell Res. Ther.* **7,** 55 (2016).
103. Basso, M. & Bonetto, V. Extracellular Vesicles and a Novel Form of Communication in the Brain. *Front. Neurosci.* **10,** 127 (2016).
104. Yang, Q., Diamond, M. P. & Al-Hendy, A. The emerging role of extracellular vesicle-derived miRNAs: implication in cancer progression and stem cell related diseases. *J. Clin. epigenetics.* **2,** (2016).

105. Desrochers, L. M., Antonyak, M. A. & Cerione, R. A. Extracellular Vesicles: Satellites of Information Transfer in Cancer and Stem Cell Biology. *Dev. cell.* **37,** 301–309 (2016).
106. Foster, B. P. *et al.* Extracellular vesicles in blood, milk and body fluids of the female and male urogenital tract and with special regard to reproduction. *Critical Reviews in Clinical Laboratory Sciences.* **53,** 379–395 (2016).
107. Barreiro, K. & Holthofer, H. Urinary extracellular vesicles. A promising shortcut to novel biomarker discoveries. *Cell and Tissue Research.* **369,** 217–227 (2017).
108. Garcia-Contreras, M., Brooks, R. W., Boccuzzi, L., Robbins, P. D. & Ricordi, C. Exosomes as biomarkers and therapeutic tools for type 1 diabetes mellitus. *Eur. Rev. Med. Pharmacol. Sci.* **21,** 2940–2956 (2017).
109. Altadill, T. *et al.* Enabling metabolomics based biomarker discovery studies using molecular phenotyping of exosome-like vesicles. *PLoS ONE.* **11,** (2016).
110. Dudley, J. T. & Butte, A. J. Identification of discriminating biomarkers for human disease using integrative network biology. *Pac. Symp. Biocomput.* **38,** 27–38 (2009).
111. Sreekumar, A. *et al.* Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature.* **457,** 910–4 (2009).
112. DeSouza, L. V *et al.* Endometrial carcinoma biomarker discovery and verification using differentially tagged clinical samples with multidimensional liquid chromatography and tandem mass spectrometry. *Mol. Cell. proteomics.* **6,** 1170–1182 (2007).
113. Spratlin, J. L., Serkova, N. J. & Eckhardt, S. G. Clinical applications of metabolomics in oncology: A review. *Clin. Cancer Res.* **15,** 431–440 (2009).
114. Graham, S. F. *et al.* Use of NMR metabolomic plasma profiling methodologies to identify illicit growth-promoting administrations. *Anal. Bioanal. Chem.* **403,** 573–82 (2012).
115. Denoroy, L., Zimmer, L., Renaud, B. & Parrot, S. Ultra high performance liquid chromatography as a tool for the discovery and the analysis of biomarkers of diseases: A review. *J. Chromatogr. B Anal. Technol.*

Biomed. Life Sci. **927,** 37–53 (2013).

116. Théry, C., Amigorena, S., Raposo, G. & Clayton, A. Isolation and characterization of exosomes from cell culture supernatants and biological fluids. *Curr. Protoc. Cell Biol.* **Chapter 3,** Unit 3.22 (2006).
117. Sheikh, K. D., Khanna, S., Byers, S. W., Fornace, A. & Cheema, A. K. Small molecule metabolite extraction strategy for improving LC/MS detection of cancer cell metabolome. *J. Biomol. Tech.* **22,** 1–4 (2011).
118. Mahieu, N. G., Genenbacher, J. L. & Patti, G. J. A roadmap for the XCMS family of software solutions in metabolomics. *Curr. Opin. Chem. Biol.* **30,** 87–93 (2016).
119. Xia, J., Sinelnikov, I. V, Han, B. & Wishart, D. S. MetaboAnalyst 3.0--making metabolomics more meaningful. *Nucleic Acids Res.* **43,** W251–7 (2015).
120. Fitzmaurice, C. *et al.* The Global Burden of Cancer 2013. *JAMA Oncol.* **1,** 505 (2015).
121. Clark, T. J. *et al.* Accuracy of outpatient endometrial biopsy in the diagnosis of endometrial cancer: A systematic quantitative review. *BJOG: An International Journal of Obstetrics and Gynaecology.* **109,** 313–321 (2002).
122. Nastic, D. *et al.* A Selective Biomarker Panel Increases the Reproducibility and the Accuracy in Endometrial Biopsy Diagnosis. *Int. J. Gynecol. Pathol.* **36,** 339–347 (2017).
123. Taoussi, N., Alghamdi, A., Futyma, K. & Rechberger, T. Biological markers with potential clinical value in endometrial cancer — review of the literature. *Ginekol. Pol.* **88,** 331–336 (2017).
124. Judes, G. *et al.* High-throughput «Omics» technologies: New tools for the study of triple-negative breast cancer. *Cancer Letters.* **382,** 77–85 (2016).
125. Johnson, C. H., Ivanisevic, J. & Siuzdak, G. Metabolomics: beyond biomarkers and towards mechanisms. *Nat. Rev. Mol. Cell Biol.* **17,** 451–459 (2016).
126. Basetti, M. Special Issue: Cancer Metabolism. *Metabolites.* **7,** 41 (2017).
127. Ametzazurra, A. *et al.* Endometrial fluid is a specific and non-invasive biological sample for protein biomarker identification in endometriosis. *Hum. Reprod.* **24,** 954–965 (2009).

128. Chen, Q. *et al.* Label-free proteomics uncovers energy metabolism and focal adhesion regulations responsive for endometrium receptivity. *J. Proteome Res.* **14,** 1831–1842 (2015).
129. Subramani, E. *et al.* NMR-based metabonomics for understanding the influence of dormant female genital tuberculosis on metabolism of the human endometrium. *Hum. Reprod.* **31,** 854–865 (2016).
130. Xie, H. *et al.* Metabolic profiling and novel plasma biomarkers for predicting survival in epithelial ovarian cancer. *Oncotarget.* **8,** 32134–32146 (2017).
131. Yang, E. C. C. *et al.* Protein expression profiling of endometrial malignancies reveals a new tumor marker: Chaperonin 10. *J. Proteome Res.* **3,** 636–643 (2004).
132. Lydic, T. A. *et al.* Rapid and comprehensive 'shotgun' lipidome profiling of colorectal cancer cell derived exosomes. *Methods.* **87,** 83–95 (2015).
133. Qiu, F., Gao, Y. H., Jiang, C. G., Tian, Y. P. & Zhang, X. J. Serum proteomic profile analysis for endometrial carcinoma detection with MALDI-TOF MS. *Arch. Med. Sci.* **6,** 245–252 (2010).
134. Alonso-Alconada, L. *et al.* Annexin-A2 as predictor biomarker of recurrent disease in endometrial cancer. *Int. J. Cancer.* **136,** 1863–1873 (2015).
135. DeSouza, L. *et al.* Search for cancer markers from endometrial tissues using differentially labeled tags iTRAQ and cICAT with multidimensional liquid chromatography and tandem mass spectrometry. *J. Proteome Res.* **4,** 377–386 (2005).
136. Monge, M. *et al.* Proteomic approach to ETV5 during endometrial carcinoma invasion reveals a link to oxidative stress. *Carcinogenesis.* **30,** 1288–1297 (2009).
137. Li, Z. *et al.* Prognostic evaluation of epidermal fatty acid-binding protein and calcyphosine, two proteins implicated in endometrial cancer using a proteomic approach. *Int. J. Cancer.* **123,** 2377–2383 (2008).
138. Alpini, G. & DeMorrow, S. Chapter 18 Changes in the Endocannabinoid System May Give Insight into new and Effective Treatments for Cancer. *Vitamins and Hormones.* **81,** 469–485 (2009).
139. Ayakannu, T., Taylor, A. H., Willets, J. M. & Konje, J. C. The evolving role of the endocannabinoid system in gynaecological cancer. *Hum. Reprod.*

Update **21,** 517–535 (2014).

140. Banford, S. & Timson, D. J. UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase- 6 (pp-GalNAc-T6): Role in Cancer and Prospects as a Drug Target. *Curr. cancer drug targets.* **17,** 53–61 (2017).
141. Rodriguez, M. C. *et al.* Targeting cancer-specific glycans by cyclic peptide lectinomimics. *Amin. Acids.* **49,** 1867–1883 (2017).
142. Wang, C., Zou, J., Ma, X., Wang, E. & Peng, G. Mechanisms and implications of ADAR-mediated RNA editing in cancer. *Cancer Lett.* **411,** 27–34 (2017).
143. Gatsiou, A., Vlachogiannis, N., Lunella, F. F., Sachse, M. & Stellos, K. Adenosine-to-Inosine RNA Editing in Health and Disease. *Antioxid. Redox Signal.* (2017).


Georgetown | Lombardi
COMPREHENSIVE CANCER CENTER

UAB
Universitat Autònoma
de Barcelona

Vall d'Hebron
Institut de Recerca

Georgetown
University