

DEPARTAMENT DE PEDIATRIA, OBSTETRICIA I
GINECOLOGIA

DESARROLLO DE UNA HERRAMIENTA MOLECULAR
DIAGNÓSTICA DE LA RECEPTIVIDAD ENDOMETRIAL.

PATRICIA DÍAZ GIMENO

UNIVERSITAT DE VALÈNCIA
Servei de Publicacions
2011

Aquesta Tesi Doctoral va ser presentada a València el dia 22 de juliol de 2011 davant un tribunal format per:

- Dr. Antonio Pellicer Martínez
- Dr. Francisco José Esteban Ruiz
- Dr. Juan Antonio García Velasco
- Dra. Ana Victoria Conesa Cegarra
- Dr. Manuel Pérez Alonso

Va ser dirigida per:

Dr. Carlos Antonio Simón Vallés

Dr. José Antonio Horcajadas Almansa

©Copyright: Servei de Publicacions
Patricia Díaz Gimeno

I.S.B.N.: 978-84-370-8519-7

Edita: Universitat de València

Servei de Publicacions

C/ Arts Gràfiques, 13 baix

46010 València

Spain

Telèfon:(0034)963864115



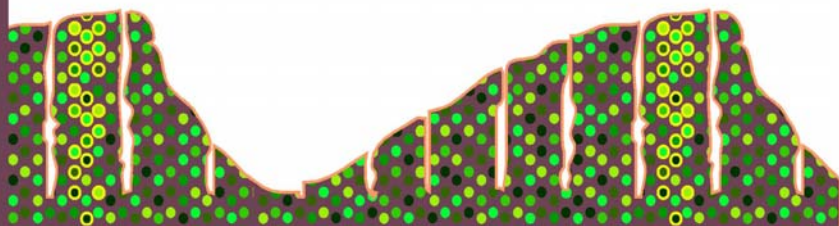
Facultad de Medicina y Odontología
Departamento de Pediatría, Obstetricia y Ginecología

Valencia 2011

TESIS DOCTORAL

DESARROLLO DE UNA HERRAMIENTA
MOLECULAR DIAGNÓSTICA
DE LA RECEPTIVIDAD ENDOMETRIAL

Patricia Díaz Gimeno
Licenciada en Biología





Facultad de Medicina y Odontología
Departamento de Pediatría, Obstetricia y Ginecología

TESIS DOCTORAL

“DESARROLLO DE UNA HERRAMIENTA MOLECULAR DIAGNÓSTICA DE LA RECEPTIVIDAD ENDOMETRIAL”

Autora:

Patricia Díaz Gimeno

Licenciada en Biología

Directores:

Prof. Carlos Antonio Simón Vallés

Dr. José Antonio Horcajadas Almansa

Valencia 2011



VNIVERSITAT
DE VALÈNCIA

Prof. Carlos Antonio Simón Vallés, Catedrático de Pediatría, Obstetricia y Ginecología de la Universidad de Valencia y Director Científico del Instituto Valenciano de Infertilidad (IVI) y del Centro de Investigaciones Príncipe Felipe (CIPF).

CERTIFICA:

Que el trabajo de investigación titulado: “**DESARROLLO DE UNA HERRAMIENTA MOLECULAR DIAGNÓSTICA DE LA RECEPTIVIDAD ENDOMETRIAL**” ha sido realizado íntegramente por Dña. Patricia Díaz Gimeno bajo mi dirección. Dicha memoria está concluida y reúne todos los requisitos para su presentación y defensa como TESIS DOCTORAL ante un tribunal.

Y para que así conste a los efectos oportunos, firmo la presente certificación en Valencia a 12 de Mayo de 2011.

Fdo. Prof. Carlos Antonio Simón Valles

Dr. José Antonio Horcajadas Almansa, Doctor en Biología,
Investigador de la Fundación IVI y Director científico de iGenomix S.L.

CERTIFICA:

Que el trabajo de investigación titulado: **“DESARROLLO DE UNA HERRAMIENTA MOLECULAR DIAGNÓSTICA DE LA RECEPTIVIDAD ENDOMETRIAL”** ha sido realizado íntegramente por Dña. Patricia Díaz Gimeno bajo mi dirección. Dicha memoria está concluida y reúne todos los requisitos para su presentación y defensa como TESIS DOCTORAL ante un tribunal.

Y para que así conste a los efectos oportunos, firmo la presente certificación en Valencia a 12 de Mayo de 2011.

Fdo. Dr. José Antonio Horcajadas Almansa

Agradecimientos

Un guerrero de la luz nunca olvida la gratitud.

Durante la lucha, fue ayudado por los ángeles; las fuerzas celestiales colocaron cada cosa en su lugar y permitieron que él pudiera dar lo mejor de sí.

Los compañeros comentan: "¡Qué suerte tiene!". Y el guerrero a veces consigue mucho más de lo que su capacidad permite.

Por eso, cuando el sol se pone, se arrodilla y agradece el Manto Protector que lo rodea.

Su gratitud, no obstante, no se limita al mundo espiritual; él jamás olvida a sus amigos, porque la sangre de ellos se mezcló con la suya en el campo de batalla.

Un guerrero no necesita que nadie le recuerde la ayuda de los otros; él se acuerda solo y reparte con ellos la recompensa.

Manual del Guerrero de la Luz. Paulo Coelho

Este pasaje del guerrero de la luz es una manera metafórica de expresar mi vivencia y agradecimiento. Al fin y al cabo la realización de una tesis doctoral es una batalla, una carrera de fondo...

Escribir los agradecimientos de esta tesis significa haber concluido un intenso trabajo por el cual estoy principalmente agradecida a mí misma...Porque sólo del amor propio puede nacer el amor y agradecimiento a los demás.

Fruto de mi satisfacción quiero mostrar mi gratitud a todas aquellas personas que, en mayor o menor medida, de manera directa o indirecta, estén mencionadas de forma explícita o no, han hecho posible esta tesis.

En primer lugar, quiero agradecer al Profesor Carlos Simón la posibilidad de comenzar mi carrera investigadora en la fundación IVI, así como

ser mi tutor y director junto al Doctor José Antonio Horcajadas. Gracias por darme la oportunidad, la guía y la idea original para esta Tesis.

A Antonio Pellicer y a José Remohí por permitirme realizar este trabajo en la Fundación IVI y en el Instituto Valenciano de Infertilidad (IVI).

A las donantes y pacientes que han participado en este estudio con sus biopsias endometriales.

Al equipo de ginecólogos y biólogos del IVI, que me han ayudado en la parte clínica y en la recogida de muestras, y en especial a todos los miembros de IVI Valencia y de Equipo IVI. Quiero hacer especial mención a Pilar Alamá, Nuria Bosch, Inma Sánchez, Gema Castellón, Nicolás Garrido, Manuel Fernández, Cristina González y Juancho García Velasco.

A la Generalitat Valenciana por la beca de formación predoctoral con la que he podido realizar mi doctorado.

A Josón por ser la persona en la que creí para entregarme a este proyecto. Aunque el camino ha sido a veces tortuoso, gracias por tu ayuda, por tus consejos y por darme tu apoyo.

A Nacho por ser mi bio_ángel_informático, por su ayuda desinteresada, por su profesionalidad y por ser esa guía que ha trabajado conmigo y que me ha enseñado tanto de predicción computacional.

A Pichi por su trabajo y asesoramiento en parte de esta tesis, por su calidad humana, por sus aportaciones científicas y literarias, por ofrecerme todos sus medios, por creer en mí y por sus enseñanzas.

A todos mis compañeros de la Fundación IVI, de iGenomix, de la facultad de Medicina y del departamento de DPG del IVI porque además de con su ayuda, con su cariño y convivencia he sido feliz día a día (almuerzos, debates, confesiones, apoyo...). Tendría mucho que agradecerlos y de forma muy personalizada, pero me tendré que conformar con nombrarlos, aunque cada uno ya sabe lo significa para mí y lo que hemos compartido. A Marcos, Isabel Galarza, Jaime, Isabel Gómez, Carmen, Ali, Sebas, Laura, Tamara,

Horten, Claudia, Aymara, Isabel Herrero, Leslie, Irene, Paco, Carlos Estella, Oscar, Raúl, Josean, María, Sandra, Gustavo, Juanma, Julio, Anita, Vicente, Aylim, Mariaje, Mila, Pere, Mónica, Nuria y como no, Concha. Gracias a todos por vuestro cariño, por ayudarme, quererme y aceptarme como soy.

A todos los miembros del grupo de investigación de Carlos Simón del CIPF que, aunque no he vivido con ellos un día a día, también han sido partícipes de mi formación predoctoral en reuniones y demás eventos. Gracias a: Diana, María Eugenia, Cristóbal, Amparo Galán, Anabel, Eva Gómez, Eva Sánchez, Sonia, Marcia, Jose Vicente, Vero, Amparo y Ana.

Al departamento de Bioinformática del CIPF de Joaquín Dopazo y en especial a Ana Conesa por sus aportaciones.

A mis amigos, cada uno sabe lo que significa para mí y la energía positiva que recibo porque existís en mi vida. No puedo nombraros a todos, pero me gustaría mencionar en especial a Cris por su amistad y ayuda en la bibliografía, impresión y empuje para terminar la tesis. A Ali, Tamara, Laura y Marcos por comprenderme, estar a mi lado y levantarme en los momentos de zambullida. A Isa y Alba por convivir conmigo estos años, compartir tantas cosas y respaldarme. Y a Luisa por su cariño y complicidad. Gracias a todos porque tanto en los buenos como en los malos momentos habéis estado a mi lado.

A Charo, por ayudarme a dar dinamismo, comprensión, fluidez y diseño a esta tesis y a mi vida. Gracias por ser esa luz que me ha dado empuje y por el diseño de la portada.

Finalmente a mi familia porque con su amor he podido sentirme libre de ser quien soy. Gracias por aceptarme y creer en mí, porque ese sustento es el que me ha hecho afrontar el mundo creyendo en mis posibilidades. A mi madre Inma, a mi padre Ángel, a mi teta Inma, a mis *aiios* Paco y Concha, y a mis tíos Paco y Fina.

Gracias a todos.

El presente trabajo de tesis doctoral ha sido realizado en los laboratorios de la Fundación IVI, así como en los laboratorios de iGenomix, gracias a la infraestructura del Instituto Valenciano de Infertilidad (IVI) y a la beca de formación predoctoral de la Generalitat Valenciana (GVA FPI BFPI/2007/106-AFPI/2007/106).

1.INTRODUCCIÓN	1
1.1. El endometrio	3
1.1.1. Definición	3
1.1.2. Anatomía	3
1.1.3. El ciclo menstrual	6
1.2. Receptividad endometrial	7
1.2.1. Definición	7
1.2.2. Marcadores de receptividad endometrial.....	12
1.2.2.1. Marcadores morfológicos	13
1.2.2.2. Marcadores moleculares clásicos	21
1.2.2.3. Marcadores moleculares “ómicos”	22
1.3. Transcriptómica: <i>microarrays</i> de expresión	23
1.3.1. <i>Microarrays</i> en investigación	27
1.3.2. Herramientas bioinformáticas	30
1.3.3. Transcriptómica en el endometrio	33
1.4. Contexto científico actual	37
2. HIPÓTESIS	39
3. OBJETIVOS	43
3.1. Objetivos generales	45
3.2. Objetivos específicos	45
4. DISEÑO EXPERIMENTAL	47
5. MATERIAL Y MÉTODOS	53
5.1. Muestras biológicas y criterios de inclusión	55
5.2. Obtención y procesado de las muestras biológicas	58
5.2.1. Obtención de las muestras	58
5.2.2. Extracción del ARN total y medición de la calidad	59
5.2.3. Etiquetado de la muestra e hibridación del <i>microarray</i>	59
5.2.4. Procesado y dataje endometrial por histología	62
5.3. Selección de genes	62

5.4. Diseño del <i>microarray</i> de receptividad endometrial (<i>Endometrial Receptivity Array, ERA</i>)	64
5.5. Preprocesado y normalización	65
5.6. Validación de la técnica de <i>microarrays</i>	66
5.7. Comportamiento de los genes del <i>ERA</i> en el ciclo menstrual	68
5.8. Estudios de expresión génica diferencial	69
5.9. Definiendo la firma transcriptómica de receptividad endometrial (<i>Endometrial Receptivity Signature, ERS</i>)	70
5.10. Métodos exploratorios de análisis en las muestras de entrenamiento del predictor	71
5.11. Entrenamiento del predictor	72
5.12. Herramienta molecular diagnóstica de la receptividad endometrial	74
5.13. Diagnóstico de la herramienta molecular en la población de pacientes	75
5.14. Comparativa dataje histológico vs. dataje predictivo	75
6. RESULTADOS . . .	79
6.1. Selección de genes	81
6.2. Diseño del <i>ERA</i>	91
6.3. Validación de la técnica de <i>microarrays</i>	91
6.4. Comportamiento de los genes del <i>ERA</i> en el ciclo menstrual	93
6.5. Definiendo la <i>ERS</i>	100
6.6. Análisis de agrupamiento jerárquico	104
6.6.1. Modelo de dataje endometrial	104
6.6.2. Modelo de clasificación patológica	105
6.7. Análisis de componentes principales (<i>Principal Component Analysis, PCA</i>) ..	108
6.7.1. Modelo de dataje endometrial	108
6.7.2. Modelo de clasificación patológica	109
6.8. Expresión génica diferencial del grupo control R-FI (receptivas con fallo de implantación)	112
6.9. Entrenamiento del predictor	115

6.9.1. Modelo de dataje endometrial	115
6.9.2. Modelo de clasificación patológica	115
6.10. Herramienta molecular diagnóstica de la receptividad endometrial	118
6.11. Diagnóstico de la herramienta molecular en la población de pacientes	120
6.12. Comparativa dataje histológico vs. dataje predictivo	120
7. DISCUSIÓN . ..	123
8. CONCLUSIONES .	143
9. BIBLIOGRAFÍA . ..	149
10. ANEXOS .	163

SIGLAS, ABREVIATURAS, ACRÓNIMOS Y ANGLICISMOS

- ACC:** **ACC**uracy (precisión).
- ADN:** **Á**cido **D**esoxirribo**N**ucleico.
- ADNc:** **Á**cido **D**esoxirribo**N**ucleico complementario.
- ARN:** **Á**cido **R**ibo**N**ucleico.
- ARNm:** **Á**cido **R**ibo**N**ucleico mensajero.
- ARNc:** **Á**cido **R**ibo**N**ucleico complementario.
- AUC:** **A**rea **U**nder **C**urve (área bajo la curva ROC).
- BP:** *gene ontology* **B**iological **P**rocess (procesos biológicos).
- CC:** *gene ontology* **C**ellular **C**omponents (componentes celulares).
- CNIO:** **C**entro **N**acional de **I**nvestigaciones **O**ncológicas.
- DIU:** **D**ispositivo **I**ntra**U**terino.
- ERA:** **E**ndometrial **R**eceptivity **A**rray (*Microarray* de receptividad endometrial).
- ERS:** **E**ndometrial **R**eceptivity **S**ignature (firma transcriptómica de receptividad endometrial).
- E2:** **E**stradiol.
- FC:** **F**old **C**hange (tasa de cambio, número de veces de cambio de una situación con respecto a otra).
- FDR:** **F**alse **D**iscovery **R**ate (p-valor corregido por el tamaño muestral).
- FI:** **F**allo de **I**mplantación.
- FISH:** **F**luorescence **I**n **S**itu **H**ybridization (hibridación in situ fluorescente).
- FIV:** **F**ecundación **I**n **V**itro.
- GEPAS:** **G**ene **E**xpression **P**attern **A**rray **S**uit (plataforma para analizar el patrón de expresión génica de análisis de *microarrays*).
- GO:** **G**ene **O**ntology (Ontología génica).
- H&E:** **H**ematoxilina y **E**osina.
- ICSI:** **I**ntra**C**itoplasmic **S**perm **I**njection (inyección intracitoplasmática de espermatozoides).
- IMC:** **I**ndice de **M**asa **C**orporal.
- IVI:** **I**nstituto **V**alenciano de **I**nfertilidad.
- KNN:** **K**-**N**earest **N**eighbours (algoritmo K de vecinos cercanos).

SIGLAS, ABREVIATURAS, ACRÓNIMOS Y ANGLICISMOS

- LH:** *Luteal Hormone* (hormona luteinizante).
- MAQC:** **MicroArray Quality Control** (control de calidad de *microarrays*).
- MCC:** **Matthews Correlation Coefficient** (coeficiente de correlación de Matthews)
- MF:** *gene ontology Molecular Function* (funciones moleculares).
- MM:** **Material y Métodos**.
- MMLV-RT:** **Moloney Murine Leukemia Virus Retro-Transcriptase** (retrotranscriptasa del virus murino de la leucemia Moloney).
- N:** número de muestras.
- PCA:** **Principal Component Analysis** (análisis de componentes principales).
- PC1:** **Principal Component 1** (componente principal primero).
- PC2:** **Principal Component 2** (componente principal segundo).
- PF:** muestra tomada en fase **ProliFerativa** y analizadas con el *ERA (Microarray de receptividad endometrial)*.
- PR:** muestra tomada en fase **Pre-Receptiva** y analizadas con el *ERA (Microarray de receptividad endometrial)*.
- PS:** muestra tomada en fase **PoS**t-receptiva y analizadas con el *ERA (Microarray de receptividad endometrial)*.
- R:** muestra tomada en fase **Receptiva** y analizadas con el *ERA (Microarray de receptividad endometrial)*.
- RF:** **Random Forest** (algoritmo de clasificación).
- R-FI:** **Receptiva con Fallo de Implantación** por causa endometrial.
- RIN:** **RNA Integrity Number** (número que valora la integridad del ARN).
- RMSE:** **Root Mean Squared Error** (error cuadrático medio).
- ROC:** **Receiver Operating Characteristic** (características operativas del receptor; curva de rendimiento diagnóstico).
- RT-PCR:** **Real Time Polymerase Chain Reaction** (reacción en cadena de la polimerasa a tiempo real).
- SAM:** **Significance Analysis of Microarrays** (análisis significativo de *microarrays*).
- SEM:** **Scanning Electron Microscopy** (microscopía electrónica de barrido).
- SVM:** **Support Vector Machine** (máquina de soporte vectorial).

SIGLAS, ABREVIATURAS, ACRÓNIMOS Y ANGLICISMOS

TEM: Transmission Electron Microscopy (microscopía electrónica de transmisión).

TRA: Tratamiento de Reproducción Asistida.

Vid.: véase.

Vs.: del latín *versus*.

WOI: Window Of Implantation (ventana de implantación).

\cap : intersección entre conjuntos.

\cup : unión entre conjuntos.

ANGLICISMOS:

Batch Effect: tendencia en los resultados debida al efecto de la tanda experimental, es decir, a las condiciones del experimento.

Cross-linking: entrecruzamiento.

Crossvalidation: validación cruzada. Método de estimación del error del algoritmo de predicción.

Fold change (FC): tasa de cambio.

Forward: secuencia del primer 5'-3' que se une en la cadena que va en dirección 3'-5' del ADN.

Gene Ontology (GO): es un proyecto que provee de un vocabulario controlado que describe el gen y los atributos del producto génico en cualquier organismo. En realidad son tres ontologías, cada cual representa un concepto clave en biología molecular: la de los procesos biológicos (BP) en los que está implicado dicho gen; la función molecular de los productos génicos (MF) y su localización en la célula (CC).

Gold standard: patrón de oro. Es el sistema tomado como referencia por excelencia.

Heatmap: es la representación natural de las matrices de expresión obtenidas mediante la tecnología de *microarrays*, ya que emula las fluorescencias estimuladas para detectar el nivel de transcripción. Un *heatmap* es una representación 2D de la matriz, donde las filas representan genes y las columnas las distintas muestras. Cada nivel de expresión se representa como un cuadrado de color en la posición correspondiente a su gen y su muestra.

SIGLAS, ABREVIATURAS, ACRÓNIMOS Y ANGLICISMOS

Microarray de expresión génica: micro matriz. Matriz bidimensional de soporte sólido en las que están impresas miles de sondas que detectan los genes expresados con una secuencia conocida.

Nombre de los genes: ej. PAEP, GPX3, GAPDH, MT1G, etc.

Pathway: ruta de señalización.

Primer: cebador. Secuencia corta de ácido nucleico que contiene un grupo 3'hidroxilo libre que forma pares de bases con una hebra molde complementaria y actúa como punto de inicio para la adición de nucleótidos con el fin de copiar la hebra molde. Se necesitan dos para la reacción de PCR, uno en el extremo 3' y el otro complementario para la otra hebra. Son de aproximadamente 20 nucleótidos, porque es la cantidad necesaria para que de manera probable coincida en un sitio de la cadena de DNA.

Reverse: secuencia del primer 3'-5' que se une en la cadena que va dirección 5'-3' del ADN.

Software: programa informático.

Splicing: proceso de edición post-transcripcional en el que se eliminan los intrones y queda el transcrito con los exones que se van a traducir.

Spot: cada uno de los puntos de luz del *microarray* que se corresponde a la medida de expresión de un oligonucleótido.

Training set: set de entrenamiento. Conjunto de muestras para entrenar al predictor.

Window of implantation (WOI): ventana de implantación. Días del ciclo menstrual en los que el endometrio posee un fenotipo receptivo que le permite que el blastocisto implante.



*“Lo que ahoga no es zambullirse,
sino permanecer bajo el agua”*

Adaptación del *Manual del Guerrero de la Luz*. Paulo Coelho

Introducción

1. INTRODUCCIÓN

1.1. El endometrio

1.1.1. Definición

El endometrio humano es el tejido que tapiza la cavidad uterina (figura 1). En la actualidad, se le considera como un órgano hormonalmente regulado que sufre cambios periódicos que son la base del ciclo menstrual. Tradicionalmente se ha considerado al endometrio como la membrana de un receptáculo inerte que constituye la parte pasiva de la gestación. Sin embargo, el endometrio es un órgano activo y dinámico entre cuyas funciones básicas se encuentra la de desarrollar temporalmente su capacidad de adhesión para posibilitar la implantación del embrión. El endometrio, además, permite que se produzca la posterior invasión controlada, la cual dará lugar a la placentación, y con ello al desarrollo fetal en los mamíferos. Es, por tanto, un órgano clave en la vida reproductiva de la mujer (Domínguez y cols., 2009).

1.1.2. Anatomía

El endometrio humano se encuentra constituido por los compartimentos epitelial, estromal y vascular, con la existencia, además, de una población de células inmunes residentes. Todo ello se distribuye en dos regiones denominadas *funcionalis* y *basalis*. La *funcionalis* se transforma y regenera cada mes durante la menstruación, mientras que la *basalis* permanece formando la base para regenerar cíclicamente la *funcionalis* (figura 1).

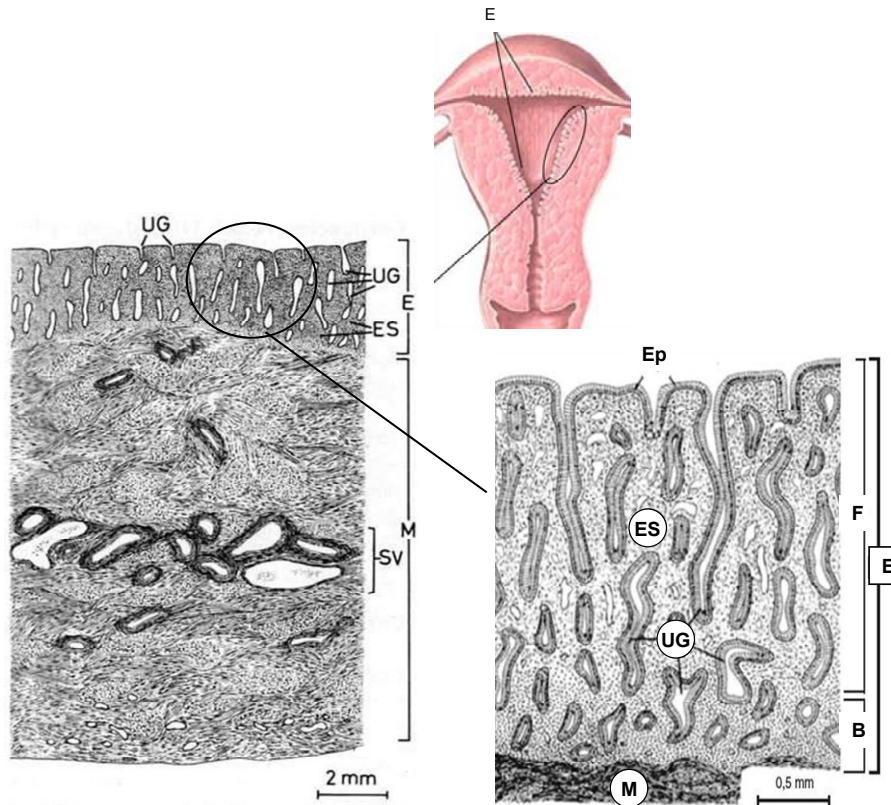


Figura 1. Endometrio Humano. Parte superior: localización del endometrio en el útero. Parte izquierda: esquema histológico del endometrio dentro del contexto del útero. Parte derecha: detalle del esquema histológico del endometrio. E: endometrio; Ep: Epitelio endometrial; UG: glándulas uterinas; ES: estroma endometrial; M: miometrio; SV: estrato vascular; B: *basalis*; F: *funcionalis* (Simón, 2009; URL1).

- **Compartimento epitelial**

El epitelio endometrial es una monocapa de células cuboidales polarizadas que tapizan el interior de la cavidad uterina (figura 1). Está constituido por un componente luminal y otro glandular (Martín y cols., 2002). Esta monocapa, como el resto de las mucosas, actúa como barrera para proporcionar protección contra los patógenos que logran acceder hasta la

cavidad endometrial, pero también debe permitir la adhesión del embrión humano que es, en esencia, la función primordial del endometrio.

El epitelio endometrial está regulado por las hormonas esteroideas ováricas que, directamente o a través del estroma (Cooke, 1995), inducen cambios morfológicos y bioquímicos cíclicos. Estos cambios ayudan a preparar el microambiente adecuado para la adhesión del embrión, dónde el epitelio actúa también como primer mediador del diálogo entre el embrión y el endometrio materno (Domínguez y cols., 2005).

El epitelio luminal varía su morfología según la fase del ciclo menstrual y el epitelio glandular forma las glándulas endometriales propiamente dichas (figura 1). Estas glándulas proliferan formando glándulas largas y convolutas durante la fase secretora temprana, que van aumentando a medida que la fase secretora avanza. La función del epitelio glandular es la producción y secreción de moléculas necesarias para la nutrición e implantación del blastocisto.

- **Compartimento estromal**

El estroma endometrial es un tejido conectivo compuesto por células y matriz extracelular (figura 1). El tipo celular que compone mayoritariamente el estroma es el fibroblasto, y está implicado en la remodelación de la matriz extracelular a lo largo del ciclo menstrual. Esta remodelación ocurre principalmente en la fase lútea mediante el proceso de decidualización. Dicho proceso comienza a partir del día 6-7 después de la aparición de la progesterona y se caracteriza por una serie de cambios morfológicos, bioquímicos y génicos de los fibroblastos. En respuesta a la exposición a los estrógenos y progesterona los fibroblastos se convierten en células

decidualizadas. La función de estos fibroblastos decidualizados es el control de la invasión del trofoblasto para que se produzca una correcta placentación.

- **Células inmunes residentes**

El endometrio humano contiene células del sistema inmunológico que son relevantes para la fisiología endometrial, especialmente en la regulación de la respuesta inmune local. El conjunto de células inmunes residentes está formado fundamentalmente por células *natural killer* uterinas (NKu), macrófagos y linfocitos T. Su función primordial es la de proteger el tracto genital frente a infecciones y evitar el rechazo inmunológico durante la implantación embrionaria. La población leucocitaria encontrada en el endometrio normal, supone entre un 10-15% de la población celular del estroma, fluctúa cíclicamente y es máxima en la fase secretora tardía y premenstrual (Bulmer y Johnson, 1985).

1.1.3. El ciclo menstrual

El ciclo menstrual es exclusivo de primates, el resto de mamíferos poseen un ciclo estral que se caracteriza por la reabsorción de su endometrio. En cambio, en el ciclo menstrual, el endometrio es destruido y expulsado en lo que conocemos como menstruación (Jolivet y Gautray, 1978; Jabbour y cols., 2006). El endometrio humano posee un ciclo menstrual, cuyo aspecto cíclico es debido a la exposición periódica de ambas hormonas esteroideas ováricas, los estrógenos y la progesterona (Critchey y cols., 2000).

El ciclo menstrual se divide en dos fases: proliferativa y secretora (figura 2). En la fase proliferativa el endometrio pasa de un grosor de 2mm hasta 12mm que puede llegar a alcanzar en la fase secretora.

En la fase proliferativa dominan los estrógenos, que regulan la proliferación de las células endometriales y su vascularización (Critchey y cols., 2000). Por el contrario, en la fase secretora, aparece la progesterona, aunque también siguen existiendo niveles elevados de estrógenos. Los progestágenos contrarrestan esas acciones proliferativas de los estrógenos por varias vías. Una de estas vías es la reducción del nivel de sus receptores, otra el aumento de la tasa del metabolismo del estradiol (E2) a compuestos inactivos (estrona y metabolitos sulfatados), y otra es la interferencia en sus acciones transcripcionales (Strauss III y Gurpide, 1993). Por ello, en la fase secretora los componentes celulares sufren diferenciación, produciéndose la secreción glandular. La fase secretora se divide a su vez en secretora temprana, media y tardía.

El ciclo menstrual y el ciclo ovárico se encuentran acoplados por la acción cíclica de las hormonas esteroideas. De modo que, atendiendo a las fases del ciclo ovárico, la fase proliferativa del endometrio se corresponde con la folicular ovárica, y la fase secretora del endometrio con la lútea.

1.2. Receptividad endometrial

1.2.1. Definición

Conocemos como receptividad endometrial al proceso biológico que permite que el endometrio alcance un estado receptivo que posibilita la implantación del embrión. Este proceso se caracteriza por producir toda una serie de cambios que se materializan a nivel histológico, celular y molecular, de modo que el endometrio adopta un fenotipo receptivo.

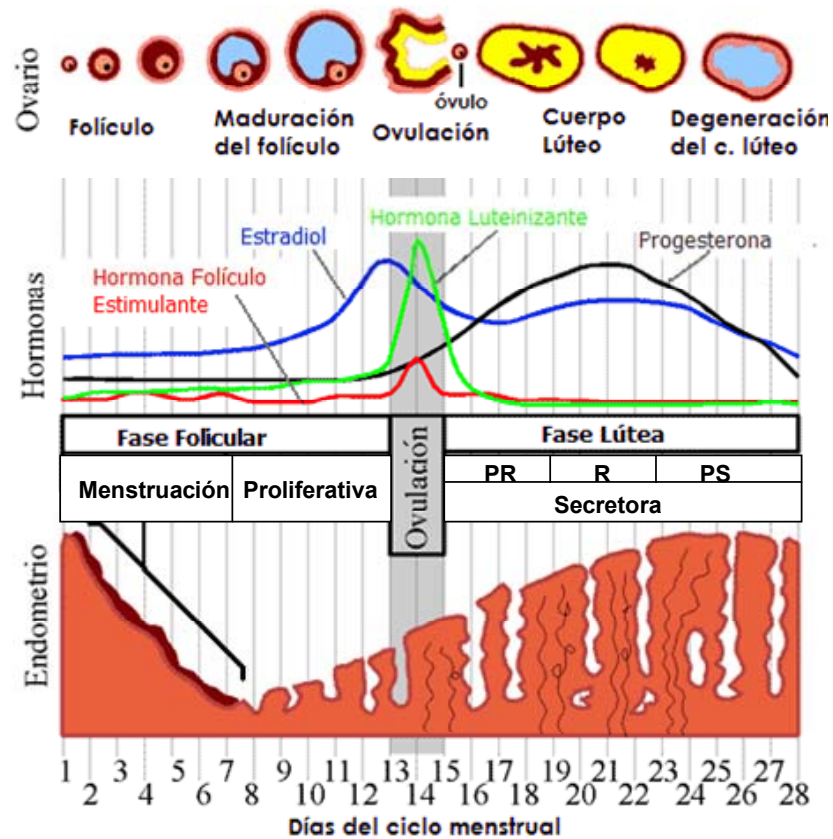


Figura 2. Ciclo menstrual y ovárico. En la parte superior de esta figura se pueden observar los cambios que suceden en el ciclo ovárico en función de los niveles hormonales. En la parte intermedia se muestran los niveles hormonales que acoplan el ciclo ovárico con el ciclo menstrual. En la zona inferior se indican los cambios que acontecen en el endometrio en cuanto al grosor, desarrollo glandular y vascular. Se muestran las fases del ciclo ovárico (folicular y lútea) y las del ciclo menstrual (proliferativa y secretora). Atendiendo a la receptividad del endometrio, podemos distinguir en la fase secretora, un periodo pre-receptivo (PR), el receptivo (R), y el post-receptivo (PS) que culmina con la menstruación. En la parte de debajo del endometrio se marcan los días del ciclo, considerando día 1 el día de comienzo de la regla (Aghajanova y cols., 2008b; Simón, 2009; URL2).

La receptividad endometrial se produce con cada ciclo menstrual, independientemente de que haya posibilidad de implantación (figura 2). Dicha receptividad acontece durante un periodo concreto de tiempo de la fase secretora que es conocido como ventana de implantación (WOI) (Wilcox y cols., 1999). La WOI en humanos se abre entorno al día 19 y se cierra en el 21 de un ciclo menstrual ideal (Bergh y Navot, 1992; Psychoyos, 1993).

En torno al día 13 del ciclo se produce un pico de la hormona luteinizante (LH) y 36 horas después ocurre la ovulación (figura 2). De este modo, se denomina día LH 0 al día que se produce el pico, y siguiendo con esta nomenclatura, el endometrio pre-receptivo correspondería al intervalo de LH+1 a LH+5 (días 14 al 18 del ciclo menstrual); el endometrio receptivo, correspondiente a la WOI, al intervalo de LH+6 a LH+8 (días 19 al 21 del ciclo); y el endometrio post-receptivo al intervalo de LH+9 a LH+13 (días 22 al 26 del ciclo).

Durante la WOI, el epitelio luminal adopta una serie de modificaciones a nivel celular que ocurren mediante un proceso denominado “transformación de la membrana plasmática” (Murphy y Shaw, 1994), en el que se originan cambios estructurales que modifican dicha membrana celular (Murphy y Shaw, 1994; Murphy, 2004) y el citoesqueleto (Murphy, 1995; Thie y cols., 1995; Martin y cols., 2002).

En respuesta a los estrógenos propios de la fase proliferativa, la membrana plasmática de la célula epitelial desarrolla unas microvellosidades largas y unas cortas uniones estrechas en la membrana lateral más apical. En la fase proliferativa el cinturón de actina es prominente, y los desmosomas son comparativamente más numerosos (figura 3).

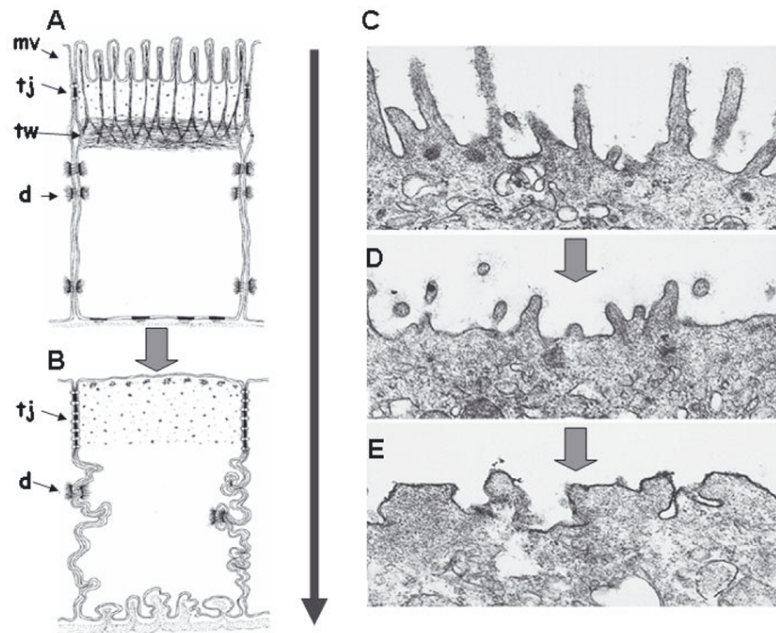


Figura 3. Transformación de la membrana plasmática del epitelio endometrial. Diagrama (A, B) y Microscopía Electrónica de Transmisión (Transmission Electron Microscopy, TEM) (C, D, E). Las imágenes A y C representan el epitelio no receptivo, la imagen D el momento de transición, mientras que B y E representan el epitelio receptivo. Parte izquierda: se muestra el esquema de la transformación de la membrana plasmática y del citoesqueleto, de la situación no receptiva (A) a la receptiva (B). Parte derecha: observamos el proceso de cambio en la superficie de la membrana plasmática, desde la situación no receptiva (C) a la receptiva (E), mostrándose la transición (D) (Murphy, 2004). Microvellosidades (mv); uniones estrechas (tj); cinturón de actina (tw); desmosomas (d).

Durante la fase secretora, la membrana plasmática pasa de la situación no receptiva (figura 3A, figura 3C) a la receptiva (figura 3B, figura 3E). En esta fase, el cinturón de actina se desorganiza, las microvellosidades desaparecen, las uniones estrechas se desarrollan y se sellan más las células, la membrana lateral y basal se vuelve más sinuosa, la frecuencia de desmosomas

disminuye, y los hemidesmosomas de la lámina basal desaparecen. Tras todo este proceso las células adoptan lo que se denomina el fenotipo receptivo (figura 3).

Mediante microscopía electrónica de barrido (Scanning Electron Microscopy, SEM) se observan unas proyecciones ectoplásmicas denominadas pinópodos o uterodomas (Martel y cols., 1987; Murphy y Shaw, 1994; Murphy, 1995; Bentin-Ley y cols., 1999; Murphy, 2004) (figura 4).

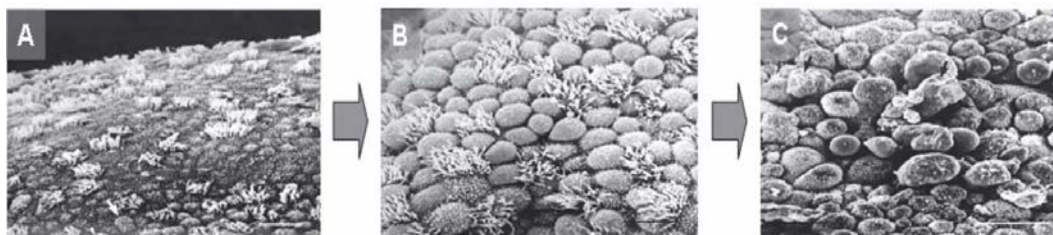


Figura 4. Proceso de aparición de los pinópodos. Microscopía electrónica de barrido (Scanning Electron Microscopy, SEM) del epitelio endometrial. **A.** LH+2 (fase pre-receptiva) donde las células ciliadas simples están separadas por células epiteliales con superficies lisas y no se observan pinópodos aún. **B.** LH+5 (fase pre-receptiva avanzada) momento de transición donde la superficie endometrial presenta células que empiezan a desarrollar pinópodos, entre las cuales se intercalan células ciliadas simples. **C.** LH+7 (fase receptiva) donde se muestran células con pinópodos desarrollados que son las células con superficie lisa arrugada (células señaladas con flecha). Las células que tienen superficie lisa y no se percibe la rugosidad, es porque sus pinópodos están en desarrollo (barra de escala= 40 μm) (Bentin-Ley y cols., 1999).

Estas formas redondeadas son microscópicas protuberancias del epitelio endometrial, como diminutos pomos, rodeadas de líquido rico en sustancias nutritivas. A pesar de desconocerse la función real de los pinópodos, parecen actuar en la modulación del ambiente uterino, absorbiendo por endocitosis de

forma activa líquido endometrial del lumen para hacer coalescer las paredes endometriales. Otra función que se les ha asignado, es la de estar relacionados con secreción (Kabir-Salmani y cols., 2005).

Estas proyecciones ectoplásmicas son estructuras celulares que permiten el intercambio de fluidos, electrolitos y proteínas de peso molecular bajo. Por otra parte, también se ha visto que actúan como sistemas de anclaje para el blastocisto en la fase de aposición y que a su vez el blastocisto tiene en su superficie una zona adaptada para tal efecto, que es el polo embrionario (Bentin-Ley y cols., 1999). La aparición de pinópodos está regulada hormonalmente, ya que se ve afectada por niveles elevados de E2 durante la fase proliferativa.

1.2.2. Marcadores de receptividad endometrial

Denominamos marcadores de receptividad endometrial a todos aquellos parámetros fáciles de detectar que nos permiten caracterizar a un endometrio como receptivo (Achache y Revel, 2006). Es por ello que los marcadores pueden ser utilizados para datar el endometrio en fase receptiva y evaluar el estatus de receptividad endometrial.

Existen distintos tipos de marcadores atendiendo al nivel de estudio en el que se obtiene información. De este modo se habla de marcadores de receptividad endometrial morfológicos cuando es la morfología la que nos informa; de marcadores de receptividad moleculares clásicos cuando son las moléculas de forma aislada las que nos caracterizan el endometrio; y de marcadores moleculares “ómicos” cuando son las moléculas en su conjunto las que nos permiten caracterizar dicho proceso biológico (Aghajanova y cols., 2008b).

1.2.2.1. Marcadores morfológicos

- **Dataje endometrial basado en los cambios cíclicos histológicos: criterios de Noyes**

Los cambios en la morfología del tejido endometrial han sido descritos mediante la observación por microscopía óptica de cortes histológicos de endometrio. Estos cortes histológicos se obtienen mediante inclusión en parafina y tinción con Hematoxilina-Eosina (H&E) del tejido endometrial. Durante la fase lútea del ciclo natural pueden observarse cambios morfológicos en el endometrio que, día a día, permiten su dataje. Actualmente la anatomía patológica rutinaria data el endometrio según los criterios de Noyes que fueron obtenidos a partir de 8.000 biopsias de endometrio de pacientes estériles (Noyes y cols., 1950 y 1975).

Los criterios de Noyes consideran ocho características histológicas básicas del endometrio: grado de mitosis glandular, pseudoestratificación de los núcleos, vacuolas basales, secreción, edema del estroma, reacción pseudodecidual, mitosis del estroma e infiltración leucocitaria (figura 5) (Noyes y cols., 1950 y 1975).

La precisión del dataje endometrial depende de múltiples factores. Estos factores son el método de determinación del día de la ovulación; el momento de toma de la biopsia; el estado de la biopsia; las técnicas de fijación y la interpretación del patólogo.

La determinación del momento exacto de la ovulación es una tarea ingrata de la rutina cotidiana, pero es esencial para la interpretación del dataje endometrial. La precisión del dataje es mayor si se toma como referencia la

determinación del pico LH, antes que referenciarse con la temperatura corporal basal o con la menstruación previa (Noyes y Haman, 1953; Acosta y cols., 2000).

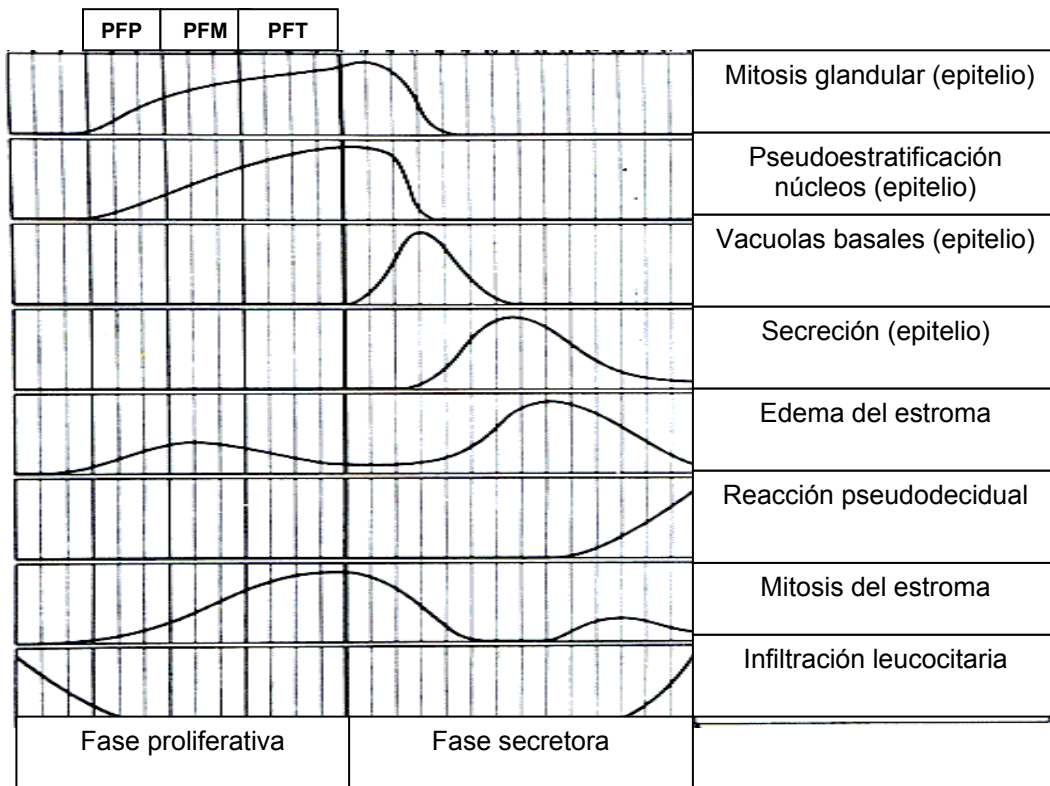


Figura 5. Criterios de Noyes. Se muestran las ocho características histológicas consideradas por Noyes para datar el endometrio (eje de ordenadas). Se puede observar como los niveles de estas características varían a lo largo del ciclo menstrual (eje de abcisas). PFP: fase proliferativa precoz; PFM: fase proliferativa media; PFT: fase proliferativa tardía. Figura adaptada a partir de Noyes y cols., 1950 y 1975.

La determinación más precisa de la ovulación se realiza midiendo el pico de LH en suero. La ovulación se produce, aproximadamente, 30 horas después del pico de LH, habiendo un desfase de tiempo entre el pico y el momento

exacto de la ovulación (figura 6). No obstante, la ovulación puede confirmarse con la ecografía ovárica y también con el aumento de progesterona que se produce el día después de ovular.

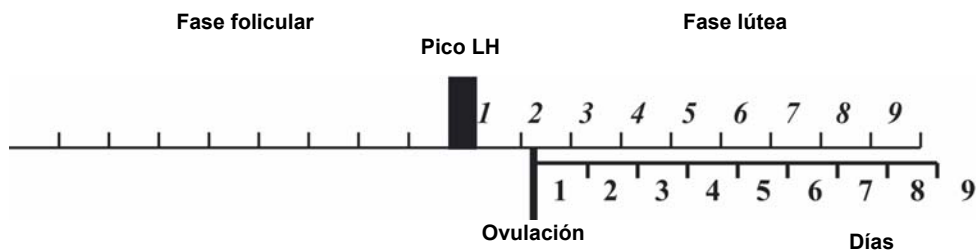


Figura 6. Desfase entre el pico de LH y la ovulación. En el eje superior se muestran los días de la fase lútea según el pico de LH, y en el eje inferior los días según la ovulación. El dataje endometrial es diferente según una u otra referencia. La ovulación ocurrirá entre uno y dos días después del pico de la LH (Bourgain y cols., 2009).

El algoritmo más sencillo es determinar primero si el endometrio es proliferativo o secretor. Si el endometrio está en la fase proliferativa muestra glándulas pequeñas, rectas y algunas algo más tortuosas. El epitelio muestra mitosis y el acúmulo de núcleos le dan un aspecto de pseudoestratificación. Las células del estroma también contienen figuras mitóticas y un edema de magnitud variable según el momento de la fase proliferativa. Esta fase también puede ser subdividida en proliferativa precoz, media y tardía. El aspecto proliferativo permanecerá sin alteraciones hasta uno a dos días después del aumento de progesterona (Bourgain y cols., 2009).

En cambio, si el endometrio está en la fase secretora es más fácil de identificar por la presencia de vacuolas y mitosis epiteliales, así como la existencia de edema y de cambios deciduales en el estroma.

Uno o dos días después del aumento de progesterona, el endometrio morfológicamente entra en la fase secretora temprana (pre-receptiva). Entre la ovulación y los cambios secretores tempranos, la morfología endometrial es reminiscente de la fase proliferativa tardía. Por lo tanto, un error de dos días puede ocurrir en el dataje de biopsias tomadas durante este período. La fase secretora temprana está caracterizada por la presencia de vacuolas en las células epiteliales, que se trasladan progresivamente de la zona subnuclear a la zona apical, y por la disminución de la actividad mitótica. El epitelio ahora presenta una única capa de células. El estroma se parece al de la fase proliferativa tardía. A mitad de la fase secretora (receptiva), que es también la WOI, la actividad mitótica cesa tanto en las glándulas como en el estroma. Las glándulas son ahora tortuosas y las vacuolas epiteliales desaparecen. El estroma se edematiza y los núcleos adquieren el llamado aspecto “desnudo”.

En toda la fase secretora media, el endometrio tendrá un aspecto muy similar y apenas es posible una diferenciación día a día. Al final de la fase secretora (post-receptiva), la actividad mitótica estromal reaparece en las células predeciduales. Los cambios deciduales se extienden progresivamente desde el área perivascular a todo el estroma. Los vasos sanguíneos se hacen más prominentes, en parte debido a la decidualización perivascular.

Finalmente, cuando disminuye la progesterona en ausencia del embarazo, la menstruación tiene lugar posteriormente a la observación de trombos en los capilares (Bourgain y cols., 2009).

A la hora de datar el endometrio basándonos en los criterios de Noyes, el dataje se debe realizar únicamente sobre un endometrio funcional, sin lesiones, carente de todos los elementos del estrato basal y del segmento uterino inferior. Los pólipos, endometritis o el tratamiento hormonal hacen imposible

aplicar correctamente los criterios clásicos de Noyes. Además, la estimulación ovárica, tan empleada en los tratamientos de reproducción asistida (TRA), produce modificaciones en el proceso de maduración endometrial si comparamos estos endometrios con endometrios en ciclo natural. De modo que, los endometrios de estas pacientes estimuladas difícilmente pueden ser explicados con los criterios de Noyes (Papanikolaou y cols., 2005).

Otra de las limitaciones de este sistema de dataje endometrial es la subjetividad que supone la observación visual. Existe una variabilidad interobservador, intraobservador e interciclo que altera la consistencia de los resultados obtenidos (Li y cols., 1989; Shoupe y cols., 1989; Balasch y cols., 1992; Batista y cols., 1993; Creus y cols., 2002; Ordi y cols., 2003; Coutifaris y cols., 2004; Murray y cols., 2004).

La precisión y relevancia funcional de este sistema, basado exclusivamente en observaciones histológicas interpretadas por uno o varios patólogos, ha sido cuestionado tanto en estudios clínicos retrospectivos (Shoupe y cols., 1989; Balasch y cols., 1992; Batista y cols., 1993), prospectivos (Li y cols., 1989; Creus y cols., 2002; Ordi y cols., 2003), así como recientemente en estudios randomizados (Coutifaris y cols., 2004; Murray y cols., 2004). Murray en su estudio sobre la precisión del dataje histológico, observó que más de un 20% de las biopsias endometriales en la fase lútea temprana fueron datadas con una diferencia de al menos dos días entre los distintos patólogos.

Con respecto a las variaciones interciclos, éstas llegaron a ser del 60% en la fase lútea media (Murray y cols., 2004) (figura 7).

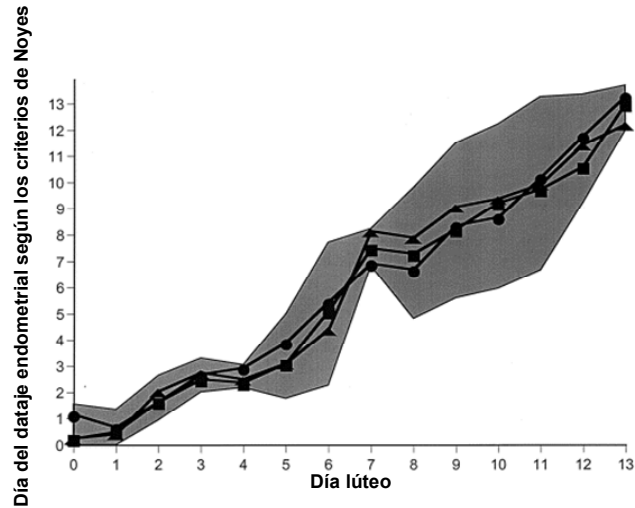


Figura 7. Variabilidad dataje histológico. En el eje de ordenadas se representa el día asignado por el patólogo de dataje histológico de acuerdo a los criterios de Noyes (Noyes y cols., 1950 y 1975). En el de abscisas se indica el día del ciclo real de la fase secretora (lútea ovárica) al que pertenece la muestra. Cada una de las tres tendencias (círculo, cuadrado y triángulo) corresponde a un patólogo y la comparativa ilustra la variabilidad del dataje entre ellos (variabilidad interobservador). Se puede apreciar como el día asignado por los patólogos está retrasado en un día con respecto al día del ciclo real al que pertenece la muestra. Cada una de las tendencias está confeccionada a partir de la media resultante de las muestras datadas pertenecientes a un mismo día del ciclo. El area sombreada representa la variabilidad intrasujeto calculada mediante la desviación típica de los datajes de muestras en un mismo día del ciclo. En la fase secretora temprana la variabilidad intrasujeto es estrecha, pero observamos la máxima variabilidad a partir de la fase secretora media y tardía (Murray y cols., 2004).

En este mismo trabajo se demostró que el dataje de la fase lútea es más exacto en la primera mitad y al final de ésta, ya que es cuando permite una valoración tanto de las características celulares glandulares como del estroma.

La precisión es más baja en las biopsias de la fase secretora media-tardía (figura 7). Por otra parte, también se demostró que el dataje histológico no está relacionado con el estatus de fertilidad (Coutifaris y cols., 2004).

Los criterios de Noyes han sido considerados, hasta el momento, el *gold standard* de la evaluación endometrial en la práctica diaria a pesar de las limitaciones expuestas en la introducción de la presente tesis doctoral. Estas variaciones descritas sugieren que los criterios tradicionales no son precisos y que se requieren nuevas tecnologías para datar e identificar funcionalmente el endometrio (Achache y Revel, 2006).

- **Pinópodos**

Se han intentado buscar otros marcadores morfológicos que permitan determinar el estado receptivo. Como hemos visto, la aparición de los pinópodos coincide con la apertura de la WOI. Por ello, inicialmente fueron propuestos considerándolos unos marcadores morfológicos del período de receptividad uterina (Nikas, 1999).

Sin embargo, además de tratarse de un método invasivo, hay otros dos motivos más que los hacen ser desestimados. En primer lugar, se trata de estructuras características del fenotipo receptivo que se observan mediante SEM (figura 4) (Bentin-Ley y cols., 1999), no resultando un marcador práctico para uso rutinario, ya que esta técnica de detección es de difícil disponibilidad por el manejo complicado y coste económico elevado.

En segundo lugar, se ha visto que los pinópodos permanecen en el endometrio post-receptivo no siendo representativos únicamente del estatus de receptividad (Quinn y cols., 2009).

- **Ecografía**

La ecografía constituye un método no invasivo, relativamente barato y que además se hace necesario a la hora de estimar el tamaño folicular. De este modo, mediante este método, la paciente no precisaría de una exploración “extra” para determinar la receptividad endometrial.

A lo largo de la historia se han identificado múltiples variables para determinar la receptividad endometrial y se han intentado correlacionar con la mayor o menor tasa de gestación en TRA. Estas variables se relacionan con el espesor y patrón ecográfico del endometrio, así como con la perfusión endometrial estudiada a distintos niveles del árbol vascular: arterias uterinas, arcuatas, radiales y espirales, zapping endometrial y subendometrial con *Doppler color y power Doppler* y, más recientemente, estudio de la vascularización endometrial en su conjunto mediante ecografía tridimensional *power Doppler* (Puente y cols., 2009).

Debido a que el incremento de vascularización, que aparece en el endometrio coincidente con el inicio de la WOI, se ha analizado en muy pocos casos, y a que los resultados de analizar el espesor, volumen y textura endometrial como parámetros clásicos pronóstico de gestación en TRA (Dickey y cols., 1992; Noyes y cols., 1995) han sido contradictorios, este tipo de análisis no han servido para evaluar la receptividad endometrial.

Por otra parte, algunos autores han encontrado una relación entre espesor endometrial y tasa de gestación en pacientes sometidas a fecundación in vitro (FIV) (Oliveira y cols., 1997; Raga y cols., 1999; Zhang y cols., 2005; McWilliams y Frattarelli, 2007) mientras que otros no encuentran tal asociación (Remohí y cols., 1997; García-Velasco y cols., 2003; Puerto y cols., 2003).

Asimismo Soares y cols., después de analizar un total de 3.089 ciclos, no encontraron diferencia estadística entre el espesor endometrial y la tasa de gestación (Soares y cols., 2005). Debido a la controversia mostrada no se ha llegado a un acuerdo, no obstante, actualmente se continúa trabajando en la búsqueda de factores pronóstico mediante la ecografía.

1.2.2.2. Marcadores moleculares clásicos

Con el desarrollo de la Biología Molecular que ha permitido profundizar en el nivel de estudio de los procesos biológicos tanto fisiológicos como patológicos, se han llevado a cabo estudios con objeto de analizar posibles marcadores bioquímicos de receptividad endometrial (Giudice y Saleh, 1995).

Existen multitud de publicaciones en las que se documenta la expresión de determinadas moléculas en el endometrio humano, en los distintos compartimentos celulares y durante las diferentes fases del ciclo menstrual (Aghajanova y cols., 2008a). De hecho, y como consecuencia de algunos de estos estudios, se ha ido a la búsqueda de marcadores moleculares de receptividad endometrial.

En la actualidad hay disponibles dos tests moleculares para la dataje endometrial basados en el análisis de una molécula o un grupo de ellas, como es el caso del *E-tegrity Test* que utiliza anticuerpos frente a la integrina $\beta 3$ (Lessey y cols., 1995) o el *EFT Quick Guide* que evalúa la expresión de ciclinas a lo largo del ciclo menstrual (Kliman y cols., 2006).

Sin embargo, estos métodos moleculares no han sustituido al método histológico basado en los criterios morfológicos de Noyes, no siendo utilizados de forma rutinaria como evaluación endometrial.

1.2.2.3. Marcadores moleculares “ómicos”

Con la secuenciación del genoma humano y el desarrollo de nuevas tecnologías, se han desarrollado lo que se ha denominado ciencias “ómicas” (Quackenbush, 2006b). Esto ha supuesto un cambio drástico en el enfoque de las investigaciones. Se ha pasado del análisis molécula a molécula a estudiar todo el conjunto de genes de un organismo en sólo un experimento.

La Genómica estudia el conjunto y ordenación de los genes en el genoma, mientras que la Genómica Funcional o Transcriptómica estudia la expresión de los genes a nivel de ARN mensajero (ARNm). Es decir, estudia la expresión de aquellos genes que realmente se expresan en el tejido o tipo celular investigado. El fundamento de la Genómica Funcional reside en el hecho de que todas las células poseen la misma información genética. Sin embargo, en función del tejido, del tipo celular o del proceso biológico que se esté llevando a cabo, la célula utiliza selectivamente ese material hereditario. El resultado es que hay una expresión génica diferencial si comparamos una población celular con respecto a otra.

Existen genes que se expresan de manera constitutiva en todos los tejidos, son los denominados genes *housekeeping*. Hay otros genes que en algunas células no se expresan (genes *off*) y en otras sí (genes *on*), y otros que se expresan pero en mayor o menor cantidad en función de la situación. De este modo, la transcriptómica o estudio del transcriptoma en su conjunto, proporciona información sobre la implicación de estos genes que se expresan de forma diferencial en un proceso biológico concreto o enfermedad.

La Proteómica estudia la colección de proteínas y su expresión y la Metabolómica estudia todos aquellos metabolitos no incluidos en la categoría

de genes, RNAm o proteínas. Por último, las disciplinas antes citadas se caracterizan por la ingente cantidad de información que producen, por ello, la Bioinformática hace que los datos obtenidos por las “-ómicas” puedan ser analizados e interpretados (Quackenbush, 2006b).

1.3. Transcriptómica: *Microarrays* de expresión

La expresión génica se empezó a medir gracias al *northern blot* que, de la misma manera que el *southern blot* con el ADN y que el *western blot* con las proteínas, permitía obtener, secuenciar y cuantificar el ARNm. En la actualidad, la técnica de la reacción en cadena de la polimerasa a tiempo real (RT-PCR) (Costa, 2004) es mucho más precisa y sensible para medir la expresión génica de un gen o de un conjunto de genes. Sin embargo, estas técnicas son limitadas para analizar la expresión génica de todo el genoma. La tecnología que permite identificar el transcriptoma, o perfil de expresión génica de todo el conjunto de genes, es la tecnología de *microarrays* de expresión génica (Schena y cols., 1995; Schena, 1996).

Los *microarrays* son matrices sólidas en las que están impresas miles de sondas de oligonucleótidos que detectan genes con una secuencia conocida. Cada sonda contiene una secuencia específica de sólo un gen, aunque para un mismo gen suele haber más de una sonda. De este modo, cada punto del *microarray* (*spot*) contiene un conjunto de sondas iguales, con una sola secuencia, con moléculas idénticas suficientes como para poder cuantificar la expresión del gen al que representan (figura 8A).

De esta manera, en solamente un experimento se puede analizar la expresión de todos los genes contenidos en el genoma mediante la cuantificación de su ARNm. Los eventos post-transcripcionales, así como la

traslocación de proteínas no pueden ser evaluados con esta técnica (Mata y cols., 2005). La tecnología de *microarrays* permite el análisis comparativo de los patrones de expresión génica de cada uno de los genes presentes en las distintas muestras.

Existe un amplio rango de *microarrays* de expresión génica caseros y comercialmente disponibles. También se encuentran disponibles en muchos formatos distintos (figura 8A, figura 8B) que se dividen principalmente en dos categorías: *microarrays* de ADN complementario (ADNc) y *microarrays* de oligonucleótidos sintéticos de alta densidad que son los que más se emplean en la actualidad (Barrett y Kawasaki, 2003). A partir de este momento se hará referencia a estos por ser los utilizados también en esta investigación.

Hay disponibles *microarrays* para un gran número de organismos que se usan como modelo de experimentación, desde bacterias a humanos, pasando por algas, levaduras y mamíferos inferiores. El número de puntos de un *microarray* puede variar desde un conjunto de genes relativos a un proceso hasta contener todo el genoma completo. El número de puntos necesarios depende del número de genes, del número de sondas por gen que queramos analizar, de las réplicas, del número de controles que incluyamos, así como si se pretende identificar los distintos transcritos resultado del *splicing* alternativo de un mismo ARNm, etc.

Para la fabricación del *microarray* se sintetiza el material genético en forma de ADNc u oligonucleótidos que se inserta de forma automatizada en una capa de cristal, nylon o plástico, colocándose en unas casillas que actúan a modo de microtubo de ensayo. Los *microarrays* de oligonucleótidos antes eran insertados por robots mediante fotolitografía o impresión piezoeléctrica, actualmente se sintetizan *in situ*. El proceso de fabricación está automatizado y

normalizado, lo cual permite insertar miles de sondas por cm^2 y minuto (figura 8A).

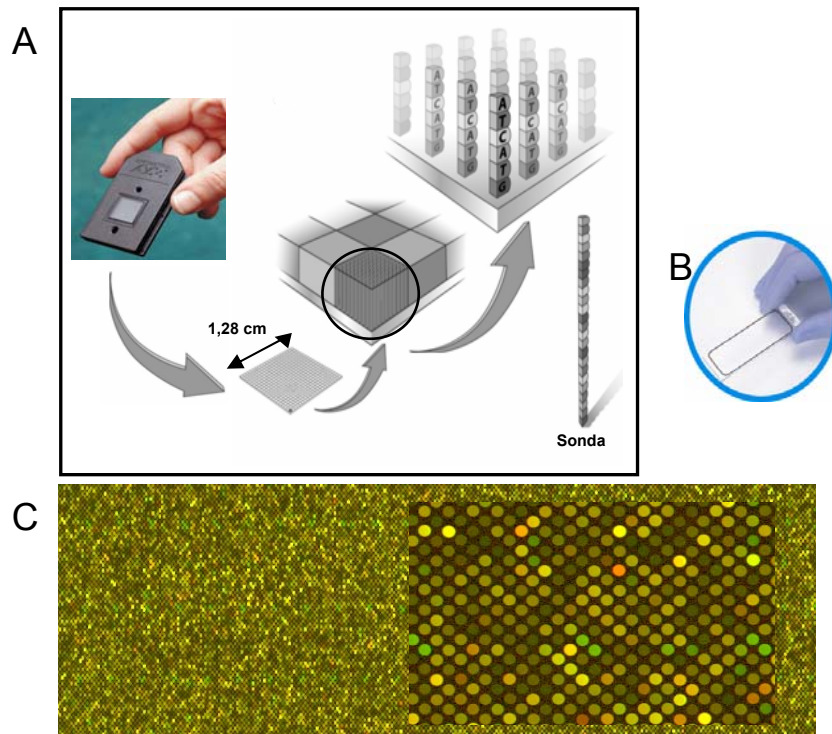


Figura 8. Microarrays. A) *Microarray* Affymetrix. Soporte que posee unas dimensiones de 1,28 cm de lado. En esta figura se observa el formato, así como la distribución de las sondas en el *microarray*. Un conjunto de sondas con la misma secuencia se sitúa en el mismo punto del soporte. En este caso, las sondas poseen 25 oligonucleótidos de longitud. B) *Microarray* de Agilent. Soporte con un formato donde los puntos de sondas se sitúan sobre un portaobjetos. C) Imagen resultante de escanear un *microarray* de expresión. Se observan, a distintos aumentos, las diferentes intensidades de fluorescencia de cada uno de los puntos del *microarray* que representan la cantidad de expresión de cada uno de los genes (Díaz-Gimeno y Horcajadas, 2009).

La primera parte del experimento consiste en tener un buen diseño experimental que nos permita disponer de un número de muestras apropiado en distintas condiciones predeterminadas. Una vez que las muestras son tomadas, se procede a la extracción del ARN total, donde está incluido el ARNm. Este paso es muy importante, ya que gran parte del éxito del análisis va a depender de la calidad del ARNm.

Una vez purificado, el ARNm es transcrito a ADNc al mismo tiempo que es marcado con nucleótidos que portan fluorescencia o que están biotinilados, según el sistema elegido. El ADNc marcado es hibridado sobre el *microarray* durante horas y posteriormente el *microarray* es lavado para eliminar los ADNc no unidos de forma específica (sólo las moléculas que hibridan permanecerán unidas a su sonda específica en el *microarray*).

El revelado se realiza mediante un escáner óptico de alta resolución que permite leer hasta imágenes de 5 μm . El resultado es una matriz de puntos que emiten más o menos fluorescencia en función de la cantidad de expresión del gen (figura 8C).

El estudio de la expresión génica se puede realizar con un marcaje de dos colores o de sólo un color. El de dos colores consiste en marcar una población de ARNm con un fluoróforo de color verde (Cy5) y la otra con un fluoróforo de color rojo (Cy3). Ambas poblaciones se hibridan al mismo tiempo en sólo un *microarray* y compiten por los sitios de unión. De modo que la intensidad de fluorescencia de cada *spot* dependerá de la cantidad de expresión relativa de ese gen en una muestra con respecto de la otra.

En el caso de marcajes con un único color, se utilizan *microarrays* distintos, uno para cada muestra, de manera que en la unión no hay

competencia. Con este sistema de sólo un color se puede comparar la expresión génica diferencial de más de dos muestras.

1.3.1 *Microarrays* en investigación

La tecnología de *microarrays* ha evolucionado de forma exponencial desde su invención hace ya 16 años (Schena y cols., 1995). Ofrece una nueva dimensión para el análisis, el diagnóstico objetivo y la búsqueda de un tratamiento etiológico de algunas enfermedades, entre las que han tenido la prioridad las oncológicas y las endocrinológicas (Quackenbush, 2006b; Medina y cols., 2007; Nevins y Potti, 2007).

El diseño experimental, la recogida y la selección de muestras son críticos en esta tecnología (Churchill, 2002). Los criterios de inclusión de las pacientes en cada estudio son fundamentales. Además, en el caso de los estudios en endometrio, es necesario asegurarse que dichas pacientes no sufren ninguna patología médica o quirúrgica conocida (ej. endometriosis) o que no estén tomando medicaciones concurrentes (ej. esteroides contraceptivos, terapia sustitutiva hormonal, antiinflamatorios no esteroideos, etc), que pueden afectar la expresión de los genes en el tejido endometrial. Las muestras de tejido que no cumplen con todos los requerimientos indicados deben ser retiradas para evitar falsas interpretaciones o artefactos.

El diseño de un experimento de *microarray* competente para obtener un poder estadístico suficiente requiere de un diseño exhaustivo, de conocimientos y experiencia en esta tecnología.

Por último, un experimento bien diseñado debe incluir una validación independiente (Rockett y cols., 2004) usando otras técnicas para medir la

expresión tales como la RT-PCR, hibridación *in situ*, o inmunohistoquímica. No obstante, esta necesidad de validación está empezando a ser cuestionada, debido a la evolución de los análisis bioinformáticos, a la experiencia, al perfeccionamiento y a la consolidación de esta tecnología en los últimos años.

Según la publicación MAQC (*MicroArray Quality Control*) (Shi y cols., 2010a) se muestra la consistencia de estos estudios entre distintas plataformas, así como el elevado nivel de concordancia dentro de una misma plataforma en los genes identificados como genes diferencialmente expresados. Este estudio proporciona un recurso que representa un primer paso importante hacia el establecimiento de un marco para la utilización de los *microarrays* en contextos clínicos y normativos.

También es importante tener en cuenta la variación biológica entre las muestras y las complicaciones inherentes a la técnica (degradación de ARN, calidad de marcaje, eficiencia de hibridación, etc).

- ***Microarrays* personalizados**

La tecnología de *microarrays* está en constante perfeccionamiento. A día de hoy se pueden diseñar *microarrays* “a la carta”, que incluyan los genes y sondas que sean interesantes para estudiar un proceso biológico determinado o enfermedad de una manera totalmente personalizada. Este tipo de servicio nos permite eliminar variables que no sean de interés, así como abaratar costes.

Las casas comerciales como *Agilent* poseen programas informáticos *on-line* (earray 4.5) donde nos ofrecen la posibilidad de elegir tanto formato, como

número de réplicas por punto, como las sondas por gen, ya sea diseñando las sondas como escogiéndolas del catálogo de la casa comercial...etc.

A la hora de diseñar un *microarray* personalizado hay que tener en cuenta el proceso biológico o enfermedad, el criterio de selección de genes a emplear, así como cuestiones técnicas como son los fenómenos de entrecruzamiento entre sondas (*cross-linking*).

En otros trabajos, se utilizan *microarrays* de genoma completo y luego se usan programas para analizar la expresión génica de aquellos genes que son de interés en la investigación, como es el caso de Khan J y cols. (Khan y cols., 2001) donde se reduce la dimensionalidad del *microarray* mediante análisis de componentes principales (PCA).

- **Firma transcriptómica**

Esta tecnología, que permite medir el perfil de expresión génica de toda la población de ARNm, ha originado una concepción diferente de marcador molecular. Ahora ya se puede definir el conjunto de genes con sus respectivos niveles de expresión que caracterizan un proceso biológico o enfermedad.

Con ello, ha surgido el concepto de firma transcriptómica o de expresión génica de un proceso o enfermedad. Según JR. Nevins y A. Potti el concepto de firma de expresión génica (*gene signature*) es una representación simplificada del estado biológico en forma de un patrón de expresión de genes que es único en esta circunstancia.

La firma transcriptómica representa una oportunidad para reflejar fenotipos biológicos (Nevins y Potti, 2007).

1.3.2. Herramientas bioinformáticas

- **Análisis de datos**

Tras el análisis de microarrays hace falta un análisis computacional (Quackenbush, 2001). Los resultados obtenidos del estudio de la expresión de miles de genes simultáneamente generan matrices de datos que requieren de complejos programas bioinformáticos para ser analizados. Estos programas procesan de forma estadística las variabilidades intra- e inter-ensayo para generar resultados estadísticamente significativos y biológicamente coherentes.

Los resultados de los análisis de expresión de microarrays son una matriz de datos con tantas variables como genes y tantas columnas como muestras. Los datos brutos que lee el escáner del sistema de detección del microarray (Axon Genepix 4100A) son transformados de forma matemática para eliminar el ruido de fondo de la técnica, y normalizados para poder comparar las distintas muestras (Quackenbush, 2002).

Tanto estos procesos de preprocesado, como la posterior interpretación de forma biológica, se pueden llevar a cabo mediante el desarrollo de sistemas de análisis. Estos programas informáticos están diseñados para ayudar al investigador a interpretar los datos en términos de sus implicaciones biológicas más que como una mera comparación de listas de genes (Al-Shahrour y cols., 2004; Al-Shahrour y Dopazo, 2005; Azuaje y Dopazo, 2005; Quackenbush, 2006a).

En la propuesta del 2001 denominada Minimal Information About Microarray Experiments (MIAME) (Brazma y cols., 2001) se establece un estándar de utilización de la tecnología de microarrays. Esta propuesta recoge

seis aspectos de actuación que son la base para estandarizar los experimentos de *microarrays*, con objeto de poder generar bases de datos públicas y desarrollar herramientas de análisis.

Existen distintos paquetes de programas informáticos ajustados para procesar los datos de expresión génica, como son el *Microarray Software Suite TM4* (Saeed y cols., 2003 y 2006) o el *Gene Expression Profile Analysis Suite (GEPAS)* (URL3). El GEPAS es uno de los programas más utilizados y ha sido diseñado para proporcionar una herramienta *on-line* que ofrece diversas opciones de análisis. Permite analizar desde los pasos tempranos de pre-procesamiento, hasta el paso final de la anotación funcional del experimento (usando *Gene Ontology*, rutas de señalización, etc), e incluye diferentes posibilidades de agrupamiento, selección génica, predicción de clases, etc. (Montaner y cols., 2006), que han sido actualmente ampliadas y actualizadas en la nueva versión de BABELOMICS (Medina y cols., 2010).

- **Predictores**

La tecnología de *microarrays* puede ser usada para identificar biomarcadores asociados con procesos biológicos o enfermedades mediante la realización de análisis de expresión génica diferencial, e incluso para encontrar patrones de expresión génica que pueden servir para distinguir subclases como se ha señalado con el concepto de firma transcriptómica (Quackenbush, 2006b; Medina y cols., 2007; Nevins y Potti, 2007).

Los predictores son algoritmos matemáticos que nos permiten clasificar categóricamente una muestra en función de las múltiples variables que se están midiendo. Se trata de algoritmos que, según el principio matemático que incluyan, dividen nuestro conjunto de datos en distintas clases en el espacio.

Una vez que el predictor ha aprendido a clasificar muestras conocidas, es capaz de predecir la clase a la que pertenece una muestra de iguales características pero de categoría desconocida.

Los primeros predictores se desarrollaron en los años 70 y fueron aplicados en distintos campos como es el reconocimiento de la escritura, la predicción del tiempo, el reconocimiento facial, etc. La predicción se puede utilizar para clasificar y predecir clases en muestras desconocidas a partir de datos de expresión génica de muestras conocidas. Es decir, utilizar esas firmas transcriptómicas que caracterizan los procesos para distinguir fenotipos.

La predicción computacional es el objetivo en muchos de los estudios de expresión génica. Cada muestra pertenece a una, y sólo a una, de las clases, pudiendo ser clasificadas en clases discretas o categorías, como son estados de enfermedades, diferentes tipos de tumores, diferentes estados del endometrio (Medina y cols., 2007). La predicción a partir de los datos de expresión de *microarrays* ha sido definida y estandarizada en MAQC-II (Shi y cols., 2010a), donde se ha estipulado qué algoritmos matemáticos son los más adecuados para realizar estos modelos predictivos a partir de esta tecnología.

El predictor utiliza un conjunto de muestras modélicas denominadas set de entrenamiento (*training set*) que son muestras control representativas de la población a estudiar. De estas muestras se conoce su clase real. El perfil de expresión génica del set de entrenamiento servirá para entrenar a los distintos predictores a distinguir entre las diferentes clases. El programa predictor calcula el error con el que cada algoritmo de predicción clasifica las muestras del set de entrenamiento. Este error es calculado en función de lo que coincide la predicción computacional con la clase real conocida. Para ello, se utiliza una

táctica denominada validación cruzada (*crossvalidation*), donde se estima la capacidad del predictor de asignar correctamente la categoría.

1.3.3. Transcriptómica en el endometrio

El nivel de estudio que proporciona la genómica funcional del endometrio constituye un enfoque más profundo sobre el conocimiento de la función endometrial, tanto a nivel fisiológico como patológico.

Se ha utilizado esta tecnología para analizar el transcriptoma del tejido endometrial a lo largo del ciclo menstrual, con especial atención en el periodo de receptividad endometrial (WOI).

El transcriptoma del endometrio en el ratón ha sido analizado en ciclos naturales (Yoshioka y cols., 2000; Reese y cols., 2001; Tan y cols., 2003) y con tratamientos para la estimulación ovárica (Cheon y cols., 2002; Watanabe y cols., 2003; Hewitt y cols., 2003; Hong y cols., 2004). Estudios similares han sido realizados en ratas (Naciff y cols., 2002; Wu y cols., 2003), y vacas (Ishiwata y cols., 2003). No obstante, también han sido llevados a cabo en primates como el *Macacus rhesus* (Ace y Okulicz, 2004; Tynan y cols., 2005). Existe una detallada revisión publicada en 2005 (White y cols., 2005) y más tarde ampliada (Horcajadas y cols., 2007) sobre los estudios realizados de la expresión génica del endometrio en humanos.

En el ámbito de la medicina reproductiva es de gran interés conocer el perfil de expresión génica del endometrio humano. Por ello, se han desarrollado modelos humanos *in vitro* e *in vivo*. Se ha estudiado la expresión génica de las células estromales durante el proceso de decidualización (Popovici y cols., 2000; Brar y cols., 2001; Tierney y cols., 2003), así como la

expresión génica que tienen estas células del estroma en respuesta a la progesterona (Okada y cols., 2003). También ha sido estudiado el patrón de expresión génica diferencial entre los compartimentos epitelial y estromal durante la fase proliferativa del ciclo menstrual en humanos, utilizando la técnica de microdissección por *laser capture* (Yanaihara y cols., 2005), así como en ratón (Niklaus y cols., 2006).

Se han encontrado evidencias moleculares convincentes sobre la identificación de perfiles de expresión génica del endometrio humano, no sólo en la fase de receptividad, sino a través de las diferentes fases del ciclo menstrual (Ponnampalam y cols., 2004; Punyadeera y cols., 2005; Talbi y cols., 2006).

El grupo de Peter Rogers en Australia fue el primero en proponer la caracterización genómica del endometrio humano a lo largo del ciclo menstrual usando la tecnología de *microarrays* (Ponnampalam y cols., 2004). Este grupo concluyó que es posible clasificar con exactitud el endometrio humano mediante su perfil molecular sin tener en cuenta su apariencia morfológica. Incluso demostraron que, en algunas muestras, el patrón de expresión génica corrige el criterio histopatológico convencional (Noyes y cols., 1950 y 1975) sugiriendo su utilización como marcador biológico del endometrio (Ponnampalam y cols., 2004). Aún más, establecieron la existencia de grupos de genes característicos de las diferentes fases del ciclo, destacando el potencial del perfil de expresión génica para el desarrollo de herramientas moleculares en la evaluación del estado endometrial.

Este estudio sobre el perfil de expresión génica del endometrio ha sido confirmado y ampliado por el grupo de Linda Giudice en la Universidad de California San Francisco (Talbi y cols., 2006) que ha analizado con detalle el

perfil molecular del endometrio humano a lo largo de todas las fases del ciclo menstrual. Sus resultados subrayan el potencial de los perfiles de expresión génica para desarrollar diagnósticos moleculares y poder distinguir entre endometrios funcionalmente normales y anormales con idéntica morfología.

- **Estudios de genómica funcional en la WOI**

¿Qué diferencias en la expresión génica se producen en el endometrio para que las células adquieran un fenotipo receptivo?. Esta es una cuestión vital para el proceso de la receptividad endometrial que ha sido estudiada mediante esta tecnología. Se ha conseguido establecer qué moléculas son las esenciales en el desarrollo del estatus receptivo del endometrio (Carson y cols., 2002; Borthwick y cols., 2003; Kao y cols., 2003; Riesewijk y cols., 2003; Mirkin y cols., 2005). Todos ellos son trabajos independientes sobre el análisis genómico funcional durante la WOI en la especie humana, que utilizan *microarrays* de expresión que identifican variaciones en el genoma completo (en aquel momento unos 12.000 genes conocidos, *Human Genome U95A Array*, *Affimetrix GeneChip Array*).

Kao y cols. compararon la fase folicular con distintos días de la fase receptiva entre LH+8 y LH+10, describiendo más de 150 genes regulados al alza y 350 regulados a la baja (Kao y cols., 2003). Carson y cols, compararon endometrios obtenidos en las fases secretora inicial (pre-receptiva) vs. media (receptiva). Describieron la existencia de 370 genes regulados a la baja, mientras que aproximadamente 320 genes aumentaban su expresión (Carson y cols., 2002). Asimismo, Borthwick y cols., mediante la utilización de un *microarray* de más de 60.000 oligonucleótidos, describen 149 transcritos, diferencialmente expresados entre endometrios proliferativos y secretores de distintas pacientes (Borthwick y cols., 2003).

En el trabajo de Riesewijk y cols., para comprender de forma global la base molecular de la receptividad endometrial en humanos, se comparó los perfiles de expresión génica del endometrio pre-receptivo (día LH+2) frente al receptivo (LH+7) de la misma mujer fértil (n=5), en el mismo ciclo menstrual (Riesewijk y cols., 2003). Una revisión más detallada de estos estudios puede ser encontrada en Horcajadas y cols. (Horcajadas y cols., 2004). Por último, el trabajo de Mirkin y cols., comparó genómicamente la fase secretora temprana frente a la tardía, utilizando el mismo *microarray* de Affymetrix y llegando a similares conclusiones que los trabajos anteriores (Mirkin y cols., 2005).

Aunque sólo uno de los genes, la osteopontina, se encontraba inducida en los cinco trabajos citados, hay varias moléculas importantes que han sido destacadas por su presencia en cuatro de los cinco artículos. Algunas de ellas son moléculas y proteínas previamente identificadas en el endometrio con o sin una función descrita. Genes implicados en el metabolismo lipídico (apolipoproteína D), respuesta inmune (*decay accelerating factor for complement*, serín o cisteín-proteasas, interleucina (IL) -15), regulación del ciclo celular (*growth arrest and DNA-damage-inducible, alfa*), unión a iones (anexina IV) o enzimas con diferentes funciones en distintos tejidos (monoamina oxidasa A).

Los resultados obtenidos por todos estos laboratorios han demostrado que la receptividad endometrial es un proceso equilibrado, complejo y activo, que implica cientos de genes inducidos y reprimidos. Aunque algunas moléculas son más relevantes que otras en el desarrollo de la receptividad endometrial, hay un pequeño cluster de 25 genes que son esenciales en la regulación de la WOI (Horcajadas y cols., 2007). Todos estos resultados también llevan a pensar que no existe una molécula clave con la capacidad de regular la receptividad endometrial por sí misma (Aghajanova y cols., 2008a).

1.4. Contexto científico actual

La aplicación de los avances tecnológicos y mejora del conocimiento en otras áreas de la medicina reproductiva contrasta con la ausencia de grandes desarrollos en referencia al endometrio humano. En la actualidad, se siguen utilizando criterios diagnósticos exclusivamente morfológicos para el dataje endometrial definidos hace más de 60 años (Noyes y cols., 1950 y 1975). También desconocemos la etiología de la endometriosis que fue descubierta hace 150 años (Von Rokitansky, 1860).

El carcinoma endometrial se ha convertido en el más común de los cánceres del aparato genital femenino en Europa y EE.UU (Evans, 1995) y es el cuarto en incidencia en la mujer después del cáncer de mama, pulmón y colon. Además, solamente del 20-30% de los embriones transferidos al útero en TRA consigue implantarse, y aunque en un buen número de casos el problema radica en el propio embrión, en otros, la causa es la infertilidad femenina debida a fallo de implantación, donde la receptividad endometrial es la responsable de muchos de estos fallos (Salamonsen y cols., 2009).

Por otra parte, la tendencia actual en los TRA es la de reducir el número de embriones transferidos y lograr la transferencia de un único embrión.

Por todo este contexto mencionado, se hace necesario cada vez más, la identificación de factores pronósticos tanto de origen embrionario como de origen endometrial (Puente y cols., 2009).

De este modo, la utilización de las nuevas tecnologías existentes puede dar nuevas posibilidades para mejorar la capacidad diagnóstica y terapéutica de los marcadores pronósticos de receptividad uterina.

Los distintos trabajos realizados del transcriptoma endometrial no han encontrado un consenso unánime entre ellos, lo cual puede ser debido a que tanto el diseño experimental, como la recolección de muestras, así como la selección de los genes, son cruciales a la hora de sacar conclusiones (Horcajadas y cols., 2007). No obstante, con el conocimiento que han aportado, se deduce que el transcriptoma parece ser un candidato firme para obtener un biomarcador de receptividad endometrial que permita caracterizar la firma transcriptómica de dicho proceso y obtener un método objetivo para la evaluación de la receptividad uterina.

Por lo tanto, se considera necesario tener disponible un *microarray* personalizado que englobe una selección de genes que sean específicos del endometrio, tanto como definir un perfil de expresión que sirva para diagnosticar y determinar si el estado de un endometrio particular corresponde al estado de receptividad/ no receptividad u otros estados alterados del endometrio, simplemente con una muestra del mismo.

Finalmente, la disponibilidad de herramientas bioinformáticas capaces de clasificar y predecir clases fenotípicas a partir de los *microrrays* de expresión, permiten el contexto científico y tecnológico idóneo para realizar el trabajo de investigación de la presente tesis doctoral.



*“La verdad en ciencia puede ser definida como
la hipótesis de trabajo que mejor se ajusta
para abrir el camino
a la siguiente mejor ajustada”*

Konrad Lorenz

Hipótesis

2. HIPÓTESIS

El transcriptoma es una propiedad inherente a cada tejido en un proceso biológico concreto que puede ser caracterizado y utilizado como biomarcador molecular. El patrón de expresión génica del tejido endometrial relacionado con la adquisición de un fenotipo receptivo puede ser caracterizado y utilizado para crear una herramienta efectiva para diagnosticar el estado de receptividad endometrial de forma objetiva.



“Cuanto más grande es el reto, más glorioso es el triunfo”

El circo de las mariposas. Joshua Weigel

Objetivos

3. OBJETIVOS

3.1. Objetivos generales

- Caracterizar y definir la firma transcriptómica de la receptividad endometrial humana (*Endometrial Receptivity Signature, ERS*).
- Crear una herramienta diagnóstica molecular de la receptividad endometrial, basada en el transcriptoma de los genes relacionados con dicho proceso, para trasladar a la clínica.

3.2. Objetivos específicos

- a) Identificar los genes que estén implicados en el proceso de la receptividad endometrial.
- b) Diseñar un *microarray* específico de receptividad endometrial con dichos genes (*Endometrial Receptivity Array, ERA*).
- c) Analizar los genes y el patrón de expresión del *ERA* durante la WOI, más específicamente, y en las otras fases del ciclo menstrual, para poder caracterizar a nivel biológico la *ERS*.
- d) Desarrollar un predictor basado en el transcriptoma *ERA* que permita evaluar de forma cuantitativa y objetiva el estado receptivo endometrial *in vivo*.
- e) Generar una herramienta diagnóstica molecular basada en el *ERA* y el predictor que contenga todos los pasos necesarios para realizar la evaluación de la receptividad endometrial.
- f) Comparar el dataje histológico basado en los criterios de Noyes con el método de predicción de la herramienta molecular diagnóstica de la receptividad endometrial.



“No hay viento favorable para el que no sabe donde va”

Séneca

Diseño Experimental

4. DISEÑO EXPERIMENTAL

Se seleccionaron los genes relativos al proceso de la receptividad endometrial a partir del estudio de expresión génica diferencial de genoma completo entre muestras control obtenidas en estado pre-receptivo y muestras control obtenidas en receptivo (figura 9). Con los genes seleccionados se diseñó un *microarray* personalizado al que se denominó *ERA*.

El transcriptoma *ERA* se analizó en otro conjunto de muestras control normales, a lo largo de todo el ciclo menstrual, con la finalidad de definir la *ERS* y de construir un modelo de predicción capaz de datar y evaluar el endometrio receptivo a partir del perfil de transcripción del *microarray* personalizado creado (figura 9).

En tercer lugar, se analizó el transcriptoma *ERA* en un conjunto de muestras endometriales diagnosticadas como fallo de implantación (FI) (figura 9). Estas muestras se utilizaron como control de FI y, junto con las muestras control receptivas normales, se generó un segundo modelo de predicción que fuera capaz de distinguir entre receptivas normales y muestras tomadas en fase receptiva pero con problemas de implantación.

Una vez establecidos ambos modelos de predicción y comprobada su capacidad predictiva, se generó la herramienta molecular diagnóstica de la receptividad endometrial. Esta herramienta funciona mediante un programa informático que es capaz de realizar todo el proceso de análisis desde la matriz de datos bruta del *ERA*, hasta el diagnóstico de receptividad mediante el sistema de predicción establecido (figura 9).

Posteriormente, otro conjunto de muestras nuevo se analizó con objeto de testar en la población de pacientes el diagnóstico de la herramienta molecular de receptividad endometrial. Este grupo de muestras estuvo constituido por biopsias endometriales tomadas en fase receptiva de pacientes que se iban a someter a un TRA. Tras obtener el transcriptoma del ERA, la matriz de datos generada, se implementó en la herramienta molecular de receptividad endometrial dando un diagnóstico transcriptómico de dataje (endometrio receptivo o no receptivo) y otro de normalidad o patología (figura 9).

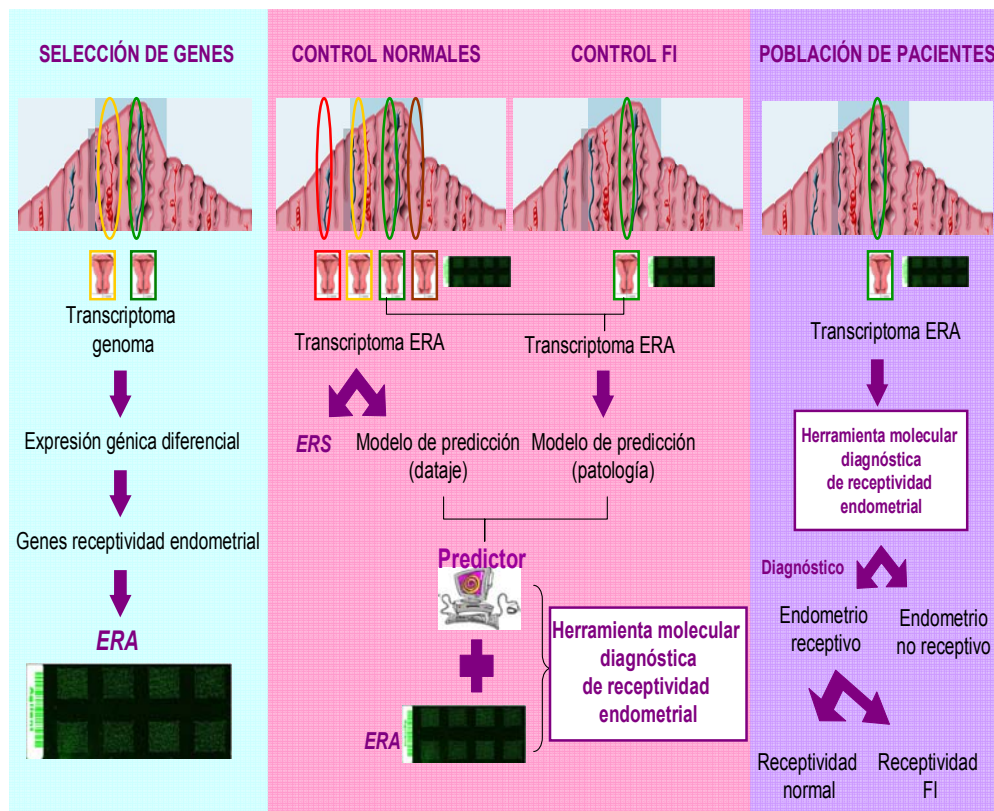


Figura 9. Diagrama de flujo del procedimiento experimental. El color de fondo indica cada una de las fases en el proceso experimental. Con fondo azul celeste

(izquierda) se incluyen los pasos seguidos para la obtención del *microarray* personalizado *ERA*. Con fondo rosa (parte central) se muestra el procedimiento seguido para obtener la *ERS* y el predictor. Por último, con fondo morado (derecha) se esquematiza el procedimiento seguido para el diagnóstico, empleando la herramienta molecular diagnóstica de receptividad endometrial obtenida gracias a las fases previas. En la parte superior de cada una de las fases se representa la imagen de un endometrio a lo largo del ciclo menstrual. El sombreado azul más oscuro del endometrio señala la fase secretora; y la pequeña franja sombreada previa muestra el momento de la ovulación. Las elipses indican el momento del ciclo en el que las biopsias fueron tomadas: la elipse de color rojo la fase proliferativa endometrial; la elipse amarilla la fase secretora temprana (pre-receptiva); la elipse de color verde la fase secretora media (receptiva) y la elipse marrón la fase secretora tardía (post-receptiva). FI: Fallo de Implantación.



“El fin no justifica los medios”

Anónimo

Material y Métodos

5. MATERIAL Y MÉTODOS

5.1. Muestras biológicas y criterios de inclusión

El material de estudio empleado fue un total de 150 biopsias endometriales de mujeres diferentes (muestras independientes), en distintos días del ciclo menstrual, siempre en ciclo natural no habiendo tomado anticonceptivos en al menos un mes previo a la toma de la biopsia.

Las mujeres reclutadas eran de raza caucásica, con un índice de masa corporal normal (IMC) de entre 19–25 kg/m² y con criterios de inclusión añadidos en función del objeto de cada parte del estudio (figura 9).

- **Muestras para la selección de genes**

Este set de muestras control sirvió para seleccionar los genes relacionados con el proceso de la receptividad endometrial para diseñar el *ERA*.

Se consideró como grupo control a mujeres de raza caucásica, sanas y fértiles con edad comprendida entre 23 y 39 años y ciclos menstruales regulares. Se asumió como fertilidad probada el haber tenido embarazos previamente.

Se calculó el pico de LH en orina y se tomaron muestras en día LH+1, LH+3, LH+5 todas ellas como fase pre-receptiva y LH+7 como fase receptiva. Se tomaron cinco muestras en cada uno de los días (n=20).

Siguiendo los criterios de Noyes (Noyes y cols., 1950 y 1975), se realizó el dataje histológico de forma ciega y por dos patólogos distintos. Sólo fueron incluidas las muestras cuyo diagnóstico de dataje era coincidente entre los dos patólogos y, a su vez, con la determinación del día del ciclo según la LH.

Posteriormente, a estas muestras incluidas se les realizó el análisis de expresión génica de todo el genoma.

- **Muestras control normales para analizar con el ERA**

Se consideró como grupo control a mujeres de raza caucásica, donantes, sanas y fértiles (haber tenido embarazos) con edad comprendida entre 19 y 40 años y con ciclos menstruales regulares. Un total de 79 muestras formaron este grupo.

La LH fue medida en sangre tomándose muestras en fase receptiva (R), día LH+7(n=39), en fase pre-receptiva (PR), días LH+1-LH+4 (n=14), y en fase post-receptiva (PS), días LH+11-LH+13 (n=12). Las muestras previas al pico de LH fueron tomadas en fase proliferativa (PF), día 8-12 (n=14), cuyo día del ciclo se determinó según la referencia al primer día de regla y con control ecográfico.

- **Muestras control con fallo de implantación (R-FI) para analizar con el ERA.**

Se consideró a mujeres de raza caucásica con edades comprendidas entre 30 y 42 años, a las que se les diagnosticó FI por haberse realizado al menos 3 FIV y/o ICSI fallidos, es decir, sin implantación y con embriones normales (programa de donación de óvulos o óvulos propios pero que con la

técnica de hibridación in situ fluorescente (FISH) más del 50% de embriones sean normales).

Se tomaron 11 biopsias endometriales en ciclo natural en LH+7 (fase receptiva) de mujeres con FI (R-FI), y el pico de LH se determinó en orina. A todas las muestras control (PF, PR, R, PS y R-FI), se les realizó el análisis transcriptómico con el *ERA*, y fueron etiquetadas según la tanda experimental, con el fin de poder detectar que no hubiera efectos en los resultados debidos al orden en que se realizó el análisis experimental (*batch effect*).

- **Muestras de la población de pacientes para testar con la herramienta molecular de receptividad endometrial**

Las muestras biológicas procedían de mujeres que acudían a la clínica para realizarse un TRA, con edades comprendidas entre 23 y 43 años. Se obtuvieron 40 biopsias endometriales en LH+7 y en ciclo natural.

El pico de LH fue detectado en sangre, de forma preferente, o en orina si a la paciente le era imposible acudir a las mediciones. A estas muestras biológicas también se les realizó el análisis transcriptómico con el *ERA* y, posteriormente, fueron testadas por la herramienta molecular de receptividad endometrial.

Consentimiento informado

Antes de la colección de cualquier tipo de muestra, se obtuvo el consentimiento escrito de la paciente previa aprobación por el comité ético del Instituto Valenciano de la Infertilidad (IVI). Anexo I.

Obtención del pico de la hormona luteinizante (LH)

A partir del día 10 del ciclo menstrual se realizó el control ecográfico. Con folículos de 16 mm se inició la determinación de LH en suero o en orina diaria. Se asumió un pico de LH según los valores que se obtuvieron durante las distintas determinaciones y se consideró el día del pico como día 0. A partir de entonces se tomó la biopsia, X días después, siendo considerado ese día como LH+X.

Para la medición de la LH urinaria se empleó un kit comercial predictor de la ovulación (Donacheck ovulación, Novalab Ibérica, S.A.L, Coslada, Madrid, España), y para la medición de la LH sérica se utilizó otro kit comercial (LH reactiva abbot laboratorios axsym system).

5.2. Obtención y procesado de las muestras biológicas

5.2.1. Obtención de las muestras

Todas las biopsias endometriales fueron tomadas del fundus uterino usando una *pipelle de Cornier* (Genetics, Namont-Achel, Belgium) en condiciones de esterilidad.

Una vez limpias de moco y sangre, las biopsias endometriales fueron congeladas a -80°C para la posterior extracción del ARN total para el análisis de *microarrays*. A las biopsias endometriales a las que se les iba a realizar el análisis histológico, antes de ser congeladas, se les tomó 1/3 de la biopsia con este fin.

5.2.2. Extracción del ARN total y medición de su calidad

La extracción de ARN total de todas las muestras fue realizada mediante el “método del trizol” de acuerdo con el protocolo recomendado por la casa comercial (Life Technologies, Inc., USA). La concentración de ARN total se midió por espectrofotometría (aproximadamente de 1-2 µg del ARN total fue obtenido por mg de tejido). La calidad del ARN fue medida con el Nanochip de ARN de Agilent (ARN 6000 Nano labChip Kit) a partir de muestras con una concentración de ARN entre 50-500 ng y fue analizada en el bioanalizador A2100 (Agilent Technologies, Germany). Sólo las muestras con un RIN mayor de 7 (*RNA Integrity Number*: número que valora la integridad del ARN) fueron válidas para analizar su transcriptoma.

5.2.3. Etiquetado de la muestra e hibridación del *microarray*

- **Microarray de genoma completo**

Las muestras utilizadas para realizar la selección de genes (vid. MM 5.1), fueron hibridadas en un *microarray* de expresión de genoma completo *Affymetrix HG-U133A chip* (Affymetrix, Inc., Santa Clara, CA) y procesadas por *Gene Logic an Ocimum Biosolutions Co.* (Gaithersburg, MD) como se describe en Tackels-Horne y cols., 2001. El resultado fue una matriz de datos con las muestras como columnas y los genes del genoma como filas.

- **Microarray *ERA***

Las muestras control (PF, PR, R, PS y R-FI), y las muestras de la población de pacientes para testar (vid. MM 5.1) fueron hibridadas con el *ERA*.

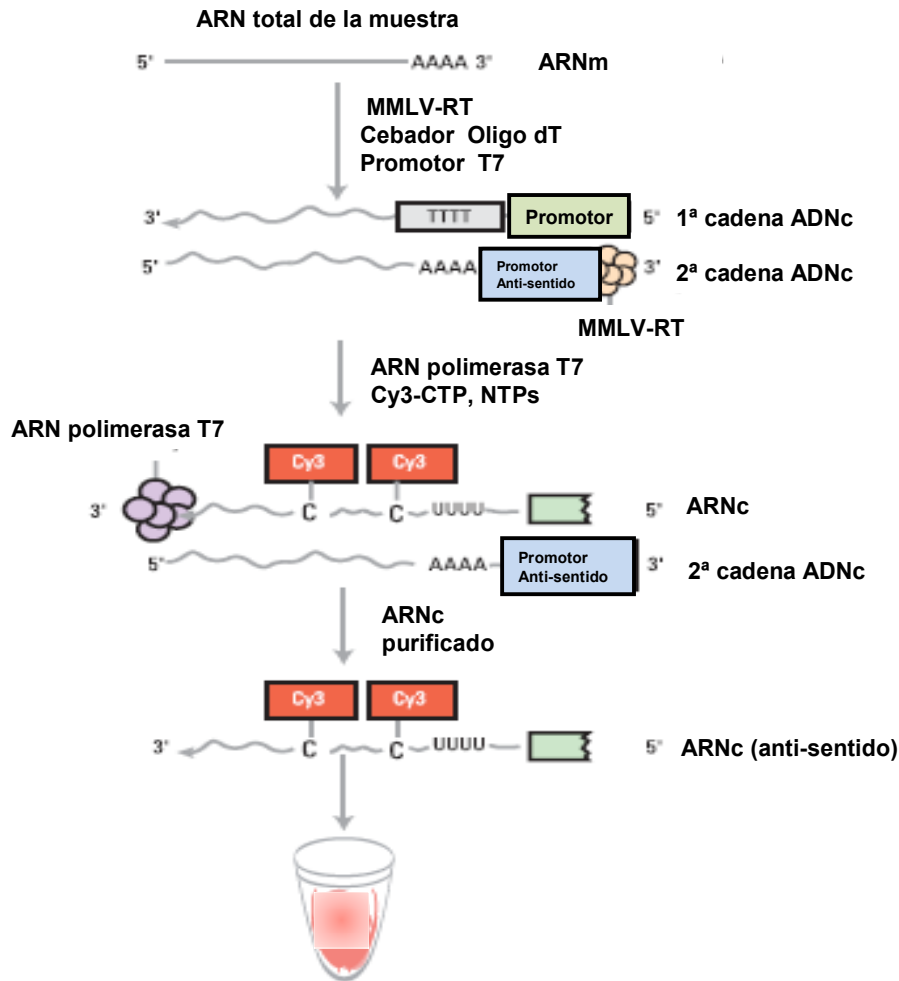
La preparación de la muestra y la hibridación con el *microarray* fue realizada según el protocolo de un color de *Agilent* (figura 10). El ADNc de simple cadena fue transcrito a partir de 1 µg del ARN total. Para ello se utilizó como cebador el promotor T7-Oligo (dT). Las muestras fueron transcritas *in vitro* y al mismo tiempo etiquetadas con el fluoróforo Cy3 (Quick-AMP labeling kit (Agilent technologies, Madrid, Spain)).

El ARN transcrito a partir del ADNc se denomina ARNc (ARN complementario). La síntesis genera entre 10 y 15 µg de ARNc. Seguidamente, se realizó una ronda de limpieza con el kit de QIAGEN (QIAGEN, Barcelona, Spain), y el ARNc se fragmentó en trozos con un rango de 35 a 200 bases, lo cual fue confirmado por el bioanalizador *Agilent 2100 Bioanalyzer technology*.

Las muestras de ARNc fragmentado (1,65 µg) fueron hibridadas con el *microarray* por incubación a 65°C durante 17h con rotación constante. Después los *microarrays* fueron lavados, en dos pasos de un minuto, en dos tampones de lavado (*Agilent technologies, Madrid, Spain*). Los *microarrays* hibridados con los ARNc marcados, fueron escaneados en el escáner Axon 4100A (Molecular Devices, Sunnyvale, CA, USA), y los datos fueron extraídos con el *software* GenePix Pro 6.0 (Molecular Devices, Sunnyvale, CA, USA).

Únicamente se incluyeron aquellas muestras que incorporaron el fluoróforo con una actividad mayor de 9.

Realizar el protocolo de un color implica que se hibridó cada muestra en un *microarray* independiente. Las muestras fueron hibridadas en tandas de 16 *microarrays* (dos portaobjetos de ocho muestras cada uno).



5.2.4. Procesado y dataje endometrial por histología

Las muestras para la selección de genes y un subconjunto de las muestras control normales fueron procesadas por histología. Aproximadamente 1/3 de la biopsia endometrial fue destinada para el estudio de anatomía patológica (100-200 mg de tejido). La muestra fue fijada en formol al 4% (p/v), incluida en parafina y teñida con H&E para su posterior dataje. El dataje según Noyes y cols. (Noyes y cols., 1950 y 1975) fue realizado de forma ciega por dos patólogos distintos.

Todas las muestras empleadas para seleccionar los genes relacionados con la receptividad endometrial tenían un dataje endometrial coincidente entre los dos patólogos y con respecto a la LH. El dataje de las muestras control (PF, PR, R y PS) sirvió para comparar el método histológico clásico con el modelo de dataje endometrial de predicción computacional que incluía la herramienta molecular diagnóstica de la receptividad endometrial (vid. punto 6.10).

5.3. Selección de genes

Se eligió un método y se estableció un criterio para identificar los genes que se encontraban regulados en el proceso de la receptividad endometrial, los cuales iban a formar parte del *microarray* personalizado.

- **Criterio estadístico**

La selección de genes fue realizada basándonos en las diferencias del perfil de expresión génica del estado pre-receptivo, representado por biopsias endometriales de mujeres en LH+1, LH+3 y LH+5, vs. biopsias de mujeres en fase receptiva, representado por LH+7 (vid. MM 5.1). El transcriptoma de estas

muestras se obtuvo a partir de un *microarray* de expresión de oligos de genoma completo de más de 22.000 genes, el *microarray Affymetrix HG-U133A chip* (Affymetrix, Inc., Santa Clara, CA).

Se realizó un *T-Test* entre el transcriptoma del grupo de muestras pre-receptivo vs. receptivo. Los genes que mostraron diferencias de expresión significativas en estas dos situaciones, con un *False Discovery Rate* (FDR) <0,05 y con un *Fold Change* (FC) >3, fueron seleccionados. El tamaño muestral empleado (vid. MM 5.1) fue considerado en el análisis estadístico, ya que el FDR calcula el p-valor en función del número de muestras.

El análisis estadístico fue realizado por distintos programas bioinformáticos: el T-Rex de la plataforma GEPAS, el SAM (*Significance Analysis of Microarrays*) y el *packaged multtest* de Bioconductor. Se realizó un diagrama de Venn utilizando la aplicación VENNY (Oliveros, 2007 URL4) con los grupos de genes obtenidos con cada programa estadístico, con objeto de seleccionar aquellos genes consenso.

- **Análisis funcional**

Con la finalidad de estudiar el *gene ontology* (GO) de los genes seleccionados para ser incluidos en el *ERA*, se analizó la relevancia biológica de la selección estadística. Para ello, se estudió qué términos estaban representados de forma significativa más de lo que cabría esperar con respecto al conjunto de todos los genes del genoma humano. Para tal efecto, se utilizó la herramienta bioinformática FatiGO (Al-Shahrour y cols., 2004) de Babelomics Suite versión 3.5 (Al-Shahrour y cols., 2005). Se seleccionó un análisis no inclusivo donde se estudiaron todos los términos del GO de los procesos biológicos (BP), de las funciones moleculares, (MF), y de los componentes

celulares (CC) entre los niveles del GO del 3-12 que poseyeran un número de genes comprendido entre 5-1000. Los términos del GO con un p-valor ajustado por Benjamin-Hochberg (Benjamini, 1995) por debajo de 0,05, se consideraron significativos.

Además de estudiar el GO, también se analizó si había alguna ruta de señalización representada de forma significativa (KEGG *pathways*), o alguna interacción entre las proteínas codificadas por el transcriptoma del *ERA* (interpro Swissprot).

5.4. Diseño del *microarray* de receptividad endometrial (*Endometrial Receptivity Array, ERA*)

El *microarray* fue diseñado y elegido con el soporte de la tecnología de *Agilent* (*Agilent technologies*, Madrid, España) a partir del programa earray 4.5 (URL5). El *ERA* es un *microarray* de expresión de sondas de oligonucleótidos compuestas por 60 bases, con un formato de 15Kx8 (ocho *microarrays* independientes del orden de 15.000 puntos impresos en un mismo portaobjetos). Cada *microarray* contiene 15.744 puntos e incluye las sondas correspondientes a los genes seleccionados para formar parte del *ERA* (un total de ocho réplicas por sonda); así como puntos de controles internos propios de *Agilent* y un conjunto de sondas para normalizar.

Las sondas fueron diseñadas y testadas por la tecnología de *Agilent*. De su catálogo de genes de genoma humano únicamente se escogieron aquellas sondas representativas de los genes seleccionados. De este modo, se utilizaron sondas testadas que daban la seguridad de que iban a funcionar correctamente, evitándose así problemas de entrecruzamiento.

5.5. Preprocesado y normalización

Tanto los *microarrays* de genoma completo de *Affymetrix*, como los del *ERA*, fueron preprocesados y normalizados del mismo modo.

El preprocesado se realizó con *preprocessCore* 1.2.1 y consistió en restar a la media de intensidad de cada punto, la mitad del valor de la mediana de la intensidad del ruido de fondo. La normalización entre los *microarrays* fue realizada utilizando el método de los cuantiles, implementado con *Bioconductor* (URL6) *limma packaged* versión 2.14.7, en R *software* versión 2.7.1 (URL7).

Para la normalización del *ERA* se utilizaron todos los puntos del microarray, tanto los de control interno y de normalización que incluía el diseño del *microarray* personalizado, así como el resto de puntos. Se analizó la distribución de la expresión de todos los puntos dentro de cada *microarray* (muestra). Se obtuvo la distribución de intensidades representando la intensidad en el eje de las abscisas y el número de puntos en el eje de ordenadas, de manera que se visualizaba la distribución de qué cantidad de puntos tenían intensidades altas, qué cantidad medias y qué cantidad bajas.

A partir de cada distribución, se confeccionó el diagrama de cajas para cada muestra, de modo que estas distribuciones fueron comparadas entre las distintas muestras del *ERA*, comprobándose que los diferentes *microarrays* a normalizar seguían una distribución parecida y por tanto era adecuado aplicar el método de los cuantiles.

Aquellas muestras que no dieron señal o en su distribución se vió que las intensidades no se ajustaban a lo esperado en una hibridación correcta, fueron eliminadas.

Posteriormente, se calculó la media de las ocho réplicas de la misma sonda y, luego, se realizó la media de las sondas pertenecientes a un mismo gen, obteniendo así, una matriz de datos con tantas columnas como muestras, y tantas filas como genes. Posteriormente, se aplicó una transformación a escala logarítmica (\log_2), originándose una matriz final lista para ser utilizada en los sucesivos estudios.

5.6. Validación de la técnica de *microarrays*

- **Validación del microarray de genoma completo por RT-PCR y mediante inmunohistoquímica**

La validación mediante RT-PCR fue realizada para cuatro genes seleccionados: el gen que codifica para la proteína de unión IGF (IGFBP-3), el gen que codifica para la glutatión peroxidasa-3 (GPX3), el gen *solute carrier family 1 (neuronal/epithelial high-affinity glutamate transporter system Xag) miembro 1* (SLC1A1) y la glicodelina [Progestagen-Associated Endometrial Protein (PAEP); PEG; placental protein 14 (PP14)]. Estos genes fueron escogidos por su relevancia en el proceso de la receptividad endometrial. Los niveles de expresión relativa de cada gen en el endometrio fueron determinados por la RT-PCR utilizando *primers* específicos para cada gen.

La inmunohistoquímica fue realizada para tres genes distintos y a diferentes días del ciclo menstrual en ciclo natural: IGFBP-7, *growth arrest and DNA damage inducible gene-45*, y la glicodelina (PAEP). Datos de la validación publicados en Horcajadas y cols., 2008.

- **Validación del ERA mediante RT-PCR**

Se validó la expresión de cuatro genes que en el *microarray* de genoma completo estaban aumentados de forma significativa en fase receptiva (LH+7), con lo que eran relevantes en el proceso, y de los cuales se tenía la seguridad de que los primers diseñados amplificaban bien en la RT-PCR. Estos genes fueron: GPX3, FXYD2 (*FXYD containing ion transport regulator 2, transcript variant a*), SPP1 (*phosphoprotein secreted 1* o Osteopontina), MT1G (*Metalotionin 1G*). Como gen *housekeeping* de control interno para normalizar se utilizó la GAPDH (*Glyceraldehyde 3-phosphate dehydrogenase*).

La expresión del *microarray* fue validada para un subconjunto de muestras (n=29) del grupo control normal analizado con el ERA (vid. MM 5.1). Este set estaba compuesto por 13 muestras en fase no receptiva (PF n=8, PR n=5) y 16 muestras en fase receptiva (R n=16). Estas muestras, de forma más detallada, estaban constituidas por: muestras PF en día 8 n=2, en día 9 n=2, en día 10 n=2, en día 11 n=1 y en día 14 n=1; muestras PR en día 16 (LH+2) n=5 y muestras R en día 21(LH+7) n=16. Se realizó la media de la expresión génica para muestras pertenecientes a un mismo día.

Los valores de expresión génica de los cuatro genes seleccionados obtenidos por RT-PCR se compararon con los obtenidos para estos mismos genes con el ERA. El FC fue calculado haciendo la media de la expresión de las muestras pertenecientes a un mismo día del ciclo menstrual frente a la media de las muestras en estado receptivo. Los *primers* (tabla 1) se diseñaron utilizando el programa *GeneFisher* (URL8) (forma *forward* y *reverse*) y luego fueron lanzados mediante la aplicación del *BLAST* para detectar homologías de secuencia (URL9).

Gen	Forward	Reverse
SPP1	TAAACCCTGACCCATCTCAGA	TGAGACTCATCAGACTGGTGA
FXYD2	AATGACTGGGTTGTCGATGGA	ACAGCGGAATCTTCTGCTGA
MT1G	TCCTGCAAGTGCAAAGAGTG	GGAATGTAGCAAAGGGGTCA
GPX3	AGGTGGAGGCTTTGTCCCTA	TATACCATCTGGCCCCACCA
GAPDH	GAAGGTGAAGGTCGGAGTC	GAAGATGGTGATGGGATTC

Tabla 1. Primers. La secuencia *forward* y la secuencia *reverse* de cada uno de los *primers* utilizados en la RT-PCR para cuantificar la expresi3n de los cuatro genes elegidos.

5.7. Comportamiento de los genes del ERA en el ciclo menstrual

Con objeto de caracterizar los genes seleccionados, se estudi3 los diferentes patrones de comportamiento de la expresi3n de los genes del ERA a lo largo del ciclo menstrual. Para ello, se utiliz3 la herramienta *maSigPro* 1.12.0 (Conesa y cols., 2006) de la plataforma GEPAS (Vaquerizas y cols., 2005). Los par3metros escogidos fueron: el n3mero de grupos en los que estudiar las tendencias comunes de comportamiento de los genes, eligiendo un n3mero de nueve grupos, que incluyeron los nueve comportamientos m3s característicos en los que se podía caracterizar a los genes del ERA; un p-valor para la selecci3n de genes de 0,2 y un p-valor para la variable del modelo de 0,2.

Las variables utilizadas fueron la expresi3n de los genes y el día del ciclo menstrual. Las muestras fueron etiquetadas atendiendo a un día del ciclo menstrual perteneciente a la fase en las que habían sido tomadas. De este modo, en los grupos de comportamiento, el endometrio proliferativo fue etiquetado como 7, el pre-receptivo como 15, el endometrio receptivo como 21 y el post-receptivo como 25. Utilizamos los datos de expresi3n del ERA de las 79 muestras controles normales (PF=14, PR=14, R=39 y PS=12) (vid. MM 5.1).

Para completar el análisis del comportamiento de los grupos obtenidos se realizó un análisis funcional de aquellos que se consideraron relevantes. El análisis fue llevado a cabo con los mismos parámetros que el análisis funcional de los genes del *ERA* (vid. MM 5.3). Del mismo modo, se utilizó la herramienta bioinformática FatiGO (Al-Shahrour y cols., 2004) de *Babelomics Suite* versión 3.2 (Al-Shahrour y cols., 2005). Se seleccionó un análisis no inclusivo donde se estudiaron todos los términos del GO de los procesos biológicos (BP), de las funciones moleculares, (MF), y de los componentes celulares (CC) entre los niveles del GO del 3-12 que poseyeran un número de genes comprendido entre 5-1000. Se consideraron significativos los términos del GO con un p-valor ajustado por Benjamin-Hochberg (Benjamini, 1995) (FDR) por debajo de 0,05.

Además de estudiar el GO, también se comprobó si había alguna ruta de señalización representada de forma significativa (KEGG *pathways*) o alguna interacción entre las proteínas codificadas por el transcriptoma del *ERA* (*interpro Swissprot*).

5.8. Estudios de expresión génica diferencial

Con el fin de analizar los genes diferencialmente expresados sólo en fase receptiva, se calculó el *T-Test* entre las muestras control normales R y el resto de grupos control normal de muestras analizados con el *ERA* (PF, PR y PS). Las combinaciones realizadas fueron: R vs. P, R vs. PR y R vs. PS. Al mismo tiempo, se confeccionó un diagrama de Venn con la aplicación VENNY (Oliveros, 2007 URL4) entre todos los grupos de genes diferencialmente expresados en cada comparación en primera instancia y, posteriormente entre estos mismos genes con respecto del total de genes del *ERA*. Esto último se realizó para confirmar que todos los genes del *ERA* estaban expresados

diferencialmente. Este análisis se completó con un ANOVA entre todos los grupos controles sanos (PF, PR, R y PS).

Otro análisis de expresión génica diferencial fue realizado con la finalidad de estudiar las diferencias entre las muestras control R y las control R-FI. Finalmente, se analizaron las diferencias de expresión génica dentro del grupo control R-FI.

Para los análisis de expresión génica diferencial, se utilizó la herramienta de expresión génica diferencial de la plataforma GEPAS (T-Rex) con un *T-Test* y la significatividad medida con un p-valor ajustado por Benjamin-Hochberg (FDR) (Benjamini, 1995) por debajo de 0,05.

5.9. Definiendo la firma transcriptómica de receptividad endometrial (*Endometrial Receptivity Signature, ERS*)

Para la obtención de la firma transcriptómica de la receptividad endometrial se seleccionaron aquellos genes específicos de la fase receptiva del ciclo menstrual. Para ello, se utilizó el estudio de expresión génica diferencial de la expresión del *ERA* entre R vs. PF, R vs. PR, y R vs. PS del apartado anterior.

Tras realizar el diagrama de Venn entre los grupos de genes diferencialmente expresados en cada comparación (vid. MM 5.8) seleccionamos aquellos genes expresados diferencialmente en fase receptiva con respecto al resto de fases del ciclo menstrual, como los genes que definen la *ERS*.

Con el objetivo de caracterizar la *ERS* se realizó el análisis funcional de los genes que contenía la firma frente al resto del genoma. Para ello, se utilizó la misma herramienta y parámetros que en el resto de estudios funcionales: FatiGO (Al-Shahrour y cols., 2004) de *Babelomics Suite* versión 3.2 (Al-Shahrour y cols., 2005). Se seleccionó un análisis no inclusivo donde se estudiaron todos los términos del GO de BP, MF y CC entre los niveles del GO del 3-12 que poseyeran un número de genes comprendido entre 5-1000. Se consideraron significativos los términos del GO con un p-valor ajustado por Benjamin-Hochberg (Benjamini, 1995) (FDR) por debajo de 0,05.

5.10. Métodos exploratorios de análisis en las muestras de entrenamiento del predictor

Estos métodos exploratorios fueron utilizados para analizar la variabilidad en el transcriptoma *ERA* dentro de cada clase y entre las distintas clases; para saber si teníamos un tamaño muestral adecuado, según la variabilidad biológica subyacente, y para saber si se tenía un set de entrenamiento representativo para generar el predictor.

Además, estos análisis también sirvieron para visualizar y comprobar que otras co-variables no estaban influyendo en el agrupamiento. Por otra parte, también permitieron detectar el posible *batch effect*, observando si las muestras analizadas en un mismo experimento se agrupaban juntas.

De esta manera, se aplicaron los métodos exploratorios utilizando el lenguaje R v2.8.1 para analizar las muestras que se iban a utilizar para el entrenamiento del predictor para el modelo de dataje endometrial y para el modelo de clasificación patológica (vid. MM 5.1). Los métodos exploratorios empleados fueron el análisis jerárquico de agrupamiento y el PCA.

- **Análisis de agrupamiento**

El agrupamiento jerárquico en función de los genes del *ERA* se realizó para ver como se agrupaban las muestras control normal, por una parte, y por otra, las muestras R-FI junto a las R. El método de agrupamiento empleado fue el de UPGMA con el coeficiente de correlación de Pearson como función de la distancia entre las muestras.

- **PCA**

Como segundo método exploratorio se realizaron distintos PCA para ver como se agrupaban las muestras en el espacio en función de la expresión de los genes del *ERA*. Los distintos PCA implementados fueron: uno con la expresión del *ERA* de todas las muestras control normal, otro con todas estas muestras más las control R-FI, otro sólo con las control R y las R-FI; y por último, un PCA sólo con las muestras R-FI (vid. MM 5.1). Se obtuvo el peso de las dos primeras componentes principales para cada uno de los PCA.

5.11. Entrenamiento del predictor

Para obtener un sistema que nos permitiera clasificar las muestras en función del transcriptoma *ERA* se utilizaron distintos algoritmos matemáticos. El objetivo era elegir el mejor método para distinguir clases dentro de la variabilidad biológica subyacente.

Los predictores se implementaron con Weka 3.7.0. y se entrenaron con aquellos algoritmos que estaban más aceptados por la comunidad científica (MAQC-II) como los que mejor clasifican los resultados multivariantes de análisis de *microarrays* (Shi y cols., 2010a).

Se utilizó el método *support vector machine* (SVM) (Vapnik, 1999) con la función *linear kernel*, el *K-nearest neighbours* (KNN) (Hastie y Friedman, 2001; Ripley, 1996) con una $K=5$ y el *Random Forest* (RF) (Breiman, 2001).

Para estimar el error de clasificación se realizó diez veces una validación cruzada de cinco particiones estratificadas, lo que conllevaba dejar fuera el 20% de las muestras como testado y hacer 10 combinaciones manteniendo las proporciones originales. La precisión (ACC), el área bajo la curva ROC (AUC), el error medio cuadrático (RMSE) y el *Matthews correlation coefficient* (MCC) fueron calculados y comparados entre los predictores, con objeto de evaluar la eficiencia de la predicción. La matriz de confusión y los términos de especificidad y sensibilidad fueron también comparados entre predictores.

Con estos parámetros se generaron dos modelos de predicción distintos:

- **El modelo de dataje endometrial**

Su finalidad fue la de datar las muestras como receptivas frente a las distintas fases del ciclo menstrual. Las muestras control utilizadas para entrenar a los clasificadores fueron las PF, PR, R y PS que se describen en el punto MM 5.1. Se consideró como resultado positivo el estado receptivo normal.

- **El modelo de clasificación patológica**

Su finalidad fue la de distinguir las muestras receptivas normales de las receptivas con problemas de FI por causa endometrial. Las muestras control que se utilizaron para generar el modelo de clasificación fueron el conjunto de

muestras R-FI (n=11) y el de muestras R (n=39), estas últimas también utilizadas para entrenar el modelo de dataje endometrial. Se consideró como resultado positivo el estado receptivo con FI.

Ambos modelos de clasificación se entrenaron con el valor de expresión de los genes incluidos en el *ERA*. También se entrenaron sólo con el valor de expresión de los genes que definieron la *ERS* con el fin de estudiar el poder predictivo del fenotipo de receptividad establecido tras la presente tesis doctoral.

5.12. Herramienta molecular diagnóstica de la receptividad endometrial

A partir de los datos de expresión del *ERA* y de los modelos de predicción escogidos se diseñó el procedimiento de la herramienta molecular diagnóstica de la receptividad endometrial. Dicha herramienta está en uso clínico actualmente y está constituida por el *ERA* y por el predictor. Para su función como diagnóstico clínico se emplea un programa informático que es capaz de datar el endometrio en fase receptiva, y de detectar si existe alguna alteración en el perfil de expresión normal de receptividad.

Este programa contiene todo el proceso, desde el preprocesado y normalización de los datos de expresión del *ERA*, hasta el diagnóstico final, tras aplicar los distintos modelos de predicción establecidos con la presente tesis doctoral. La normalización se realiza siguiendo el procedimiento definido (vid. MM 5.5) y en función del conjunto de muestras control (set de entrenamiento) con el que se ha construido el predictor (vid. MM 5.1 y 5.11).

Con esta manera de normalizar se consigue que las nuevas muestras a testar estén escaladas y tengan la misma distribución que las muestras con las que se ha definido la clasificación. Para tal función se utilizaron los paquetes de *Bioconductor*, *limma* y *preproccesCore* y para la implementación de los algoritmos de predicción se utilizó *weka* v3.7.0.

El diagnóstico de cada algoritmo es calculado con un valor de confianza que estima con qué probabilidad se da la predicción. Además, el programa contiene la posibilidad de re-entrenar los modelos de predicción con muestras nuevas. Así como generar nuevos modelos de predicción si aparecieran nuevos perfiles de expresión asociados a un fenotipo que distinguiera clases.

5.13. Diagnóstico de la herramienta molecular en la población de pacientes

Una vez generada la herramienta molecular diagnóstica de la receptividad endometrial, se analizó el transcriptoma *ERA* de 40 biopsias endometriales en fase receptiva (LH+7) de mujeres que acudían a la clínica para someterse a un TRA (vid. MM 5.1). Esto se realizó con objeto de testar a la población de pacientes.

Las matrices de datos resultantes del análisis *ERA* de estas muestras se lanzaron con el programa predictor de la herramienta molecular obteniendo un diagnóstico de receptividad y patología.

5.14. Comparativa dataje histológico vs. dataje predictivo

Con el objetivo de comparar el predictor del nuevo método diagnóstico, generado en la presente tesis doctoral, con el clásico método de dataje de

Noyes (Noyes y cols., 1950 y 1975) se utilizó un subconjunto de las muestras control normales (vid. MM 5.1). Un total de 49 muestras formadas por: PF (n=6), R (n=8), R (n=24) y PS (n=11).

Las biopsias endometriales fueron procesadas para la preparación de cortes histológicos (vid. MM 5.2.4) y analizadas según los criterios de Noyes (Noyes y cols., 1950 y 1975) por dos patólogos diferentes en las distintas fases del ciclo menstrual y de forma ciega. A su vez, estas mismas 49 muestras fueron datadas de forma ciega según el modelo de predicción de dataje endometrial que mejores resultados de predicción obtuvo. Se fijó como medida de referencia el pico de LH, y el día del ciclo menstrual a partir del primer día de regla para las mujeres en fase proliferativa.

Asumimos que, en los controles, el ciclo menstrual y con ello el estado del endometrio, está acoplado al ovárico, ya que son mujeres sanas, de fertilidad probada y con ciclos menstruales regulares.

Para comparar ambos métodos se estableció un criterio de correspondencia del dataje que permitiera evaluar la variable categórica (fase del ciclo menstrual) de una manera objetiva y por igual. Se etiquetó como endometrio proliferativo a aquellas muestras datadas del día 8 a 14, como endometrio pre-receptivo a aquellas entre el día 15 al día 18 inclusive, como endometrio receptivo a aquellas del día 19 al día 23, incluyendo las etiquetadas como día 19-20 y día 23-24, y como endometrio post-receptivas a aquellas entre el día 25 al 27, incluyendo las datadas como día 24-25.

La concordancia de métodos de dataje (predictor y anatomía patológica) en relación con el pico de LH y entre patólogos fue analizado estadísticamente

por el análisis del índice de concordancia Kappa ponderado por pesos cuadráticos (Domenech, 2006), usando el paquete estadístico SPSS.

Los valores del índice Kappa entre 0,61-0,80 son considerados como “buena concordancia”, mientras que valores entre 0,81-1,00 como “muy buena concordancia”. También se realizó una comparación de los tres datajes (patólogo 1, patólogo 2 y predictor) en cada fase del ciclo menstrual, midiendo el % de los aciertos y de fallos.



*“Lo importante en ciencia no es tanto obtener nuevos hechos
como descubrir nuevas formas de pensar sobre ellos”*

William Lawrence Bragg

RESULTADOS

6. RESULTADOS

6.1. Selección de genes

- **Criterio estadístico**

Tras realizar un análisis de expresión génica diferencial, entre el transcriptoma de genoma completo de los grupos de muestras pre-receptivas vs. receptivas (vid. MM 5.1), se seleccionaron los genes que mostraron diferencias de expresión entre estas dos situaciones con un $FDR < 0,05$ y un $FC > 3$. El análisis estadístico fue realizado por distintos programas bioinformáticos: el T-Rex de la plataforma GEPAS, el SAM, y el *packaged multtest*. El número de genes que finalmente contuvo el *ERA* fue el resultado de la selección obtenida mediante la unión de conjuntos entre GEPAS y SAM ($GEPAS \cup SAM$) con la intersección, del conjunto resultante de dicha unión, con el multtest ($(GEPAS \cup SAM) \cap multtest$). Se realizó como se esquematiza en el diagrama de Venn (figura 11). Tras el proceso de selección se encontraron 238 genes diferencialmente expresados (143 incrementados y 95 disminuidos) (tabla 2).

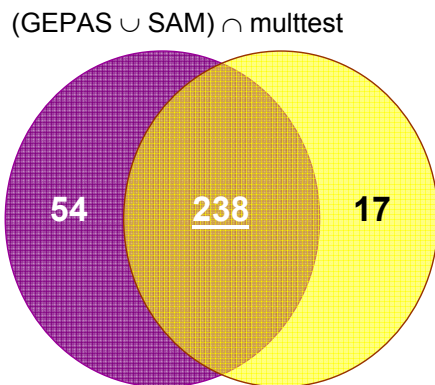


Figura 11. Selección de genes. Diagrama de Venn resultado de la unión de genes entre la selección resultante de utilizar GEPAS (T-Rex) y SAM (suman un total de 292 genes), con la intersección con multtest (contiene una selección de 255 genes). El resultado fue los 238 genes que incluye el *ERA*.

Desarrollo de una Herramienta Molecular Diagnóstica de la Receptividad Endometrial

Símbolo del gen	Nombre del gen	FC	n° sondas
GPX3*	Glutathione peroxidase 3 (plasma).	35,49	2
PAEP	Progestagen-associated endometrial protein (placental protein 14, pregnancy-associated endometrial alpha-2-globulin, alpha uterine protein) (PAEP), transcript variant 2.	31,43	1
COMP*	Cartilage oligomeric matrix protein.	30,95	2
SLC1A1	Solute carrier family 1 (neuronal/epithelial high affinity glutamate transporter, system Xag), member 1.	17,57	3
LIF	Leukemia inhibitory factor (cholinergic differentiation factor).	15,03	3
TCN1*	Transcobalamin I (vitamin B12 binding protein, R binder family).	14,76	1
CXCL14	Chemokine (C-X-C motif) ligand 14.	14,02	2
C4BPA	Complement component 4 binding protein, alpha.	13,14	2
TSPAN8*	Tetraspanin 8.	12,90	2
LAMB3	Laminin, beta 3, transcript variant 2.	11,32	1
MAOA	Monoamine oxidase A, nuclear gene encoding mitochondrial protein.	9,39	2
SOD2	Superoxide dismutase 2, mitochondrial, nuclear gene encoding mitochondrial protein, transcript variant 2.	9,06	2
GADD45A	Growth arrest and DNA-damage-inducible, alpha.	8,25	1
MUC16	Mucin 16, cell surface associated.	8,01	8
THBD	Thrombomodulin.	7,84	3
NNMT	Nicotinamide N-methyltransferase.	7,74	2
DPP4	Dipeptidylpeptidase 4 (CD26, adenosine deaminase complexing protein 2).	7,72	3
SCGB2A2	Secretoglobin, family 2A, member 2.	7,43	2
S100P*	S100 calcium binding protein P.	6,95	1
SNX10	Sorting nexin 10.	6,56	2
CP	Ceruloplasmin (ferroxidase).	6,34	2
G0S2	Putative lymphocyte G0/G1 switch gene.	6,20	2
C4.4 ^a	GPI-anchored metastasis-associated protein homolog.	6,03	1
ANG	Angiogenin, ribonuclease, RNase A family, 5.	5,98	2
ABCC3	ATP-binding cassette, sub-family C (CFTR/MRP), member 3.	5,98	1
XCL1	Chemokine (C motif) ligand 1.	5,80	3
ADRA2A	Adrenergic, alpha-2A-, receptor.	5,78	2
EFNA1*	Ephrin-A1, transcript variant 1.	5,77	3
KLRC1	killer cell lectin-like receptor subfamily C, member 1, transcript variant 2.	5,75	2
TAGLN	Transgelin.	5,71	3
SLC15A1	Solute carrier family 15 (oligopeptide transporter), member 1.	5,59	2
IGFBP1	Insulin-like growth factor binding protein 1.	5,35	3
PTGER2	Prostaglandin E receptor 2 (subtype EP2), 53kDa.	5,18	2
THBS2	Thrombospondin 2.	5,17	3
HPSE*	Heparanase.	5,17	1
SERPING1	Serpin peptidase inhibitor, clade G (C1 inhibitor), member 1, (angioedema, hereditary), transcript variant 1.	5,16	2
CFD	D component of complement (adipsin) (DF).	5,13	2
CRISP3	Cysteine-rich secretory protein 3.	5,09	1
C14orf161	Chromosome 14 open reading frame 161.	5,07	2

Tabla 2. Genes que componen el ERA.

Símbolo del gen	Nombre del gen	FC	n° sondas
RPRM	Reprimo, TP53 dependant G2 arrest mediator candidate.	5,03	3
GAST	Gastrin.	5,00	2
EDNRB	Endothelin receptor type B, transcript variant 2.	4,89	4
RARRES1	Retinoic acid receptor responder (tazarotene induced) 1, transcript variant 2.	4,87	1
ABLIM3	Actin binding LIM protein family, member 3.	4,87	4
GLIPR1	GLI pathogenesis-related 1 (glioma).	4,83	2
AOX1*	Aldehyde oxidase 1.	4,82	2
CYP3A5	Cytochrome P450, family 3, subfamily A, polypeptide 5.	4,82	3
CTSW	Cathepsin W.	4,78	3
DHRS3*	Dehydrogenase/reductase (SDR family) member 3.	4,72	1
MYL9*	Myosin, light polypeptide 9, regulatory, transcript variant 2.	4,72	3
CLU	Clusterin (CLU), transcript variant 2.	4,70	2
IER3	Immediate early response 3, transcript variant long.	4,69	4
GPRC5C	G protein-coupled receptor, family C, group 5, member C, transcript variant 1.	4,69	6
C3	Complement component 3.	4,67	3
AQP3	Aquaporin 3 (Gill blood group).	4,63	2
S100A4*	S100 calcium binding protein A4, transcript variant 1.	4,59	2
GNLY	Granulysin (GNLY), transcript variant NKG5.	4,53	1
INDO	Indoleamine-pyrrole 2,3 dioxygenase.	4,49	1
NKG7	Natural killer cell group 7 sequence.	4,42	1
CORO1A*	Coronin, actin binding protein, 1A.	4,40	1
SLPI	Secretory leukocyte protease inhibitor (antileukoproteinase).	4,39	3
ARG2*	Arginase, type II, nuclear gene encoding mitochondrial protein.	4,35	1
ANXA4*	Annexin A4.	4,27	1
ARID5B	AT rich interactive domain 5B (MRF1-like).	4,19	6
DEFB1*	Homo sapiens defensin, beta 1.	4,17	2
<u>SPP1</u>	<u>Secreted phosphoprotein 1 (osteopontin, bone sialoprotein I, early T-lymphocyte activation 1).</u>	4,15	2
DKK1*	Dickkopf homolog 1 (Xenopus laevis).	4,10	1
HABP2*	Hyaluronan binding protein 2.	4,09	2
GZMA	Granzyme A (granzyme 1, cytotoxic T-lymphocyte-associated serine esterase 3).	4,07	1
TBC1D2*	TBC1 domain family, member 2.	4,07	2
CDA	Homo sapiens cytidine deaminase (CDA), mRNA [NM_001785]	4,00	2
C10orf10	Chromosome 10 open reading frame 10.	3,98	3
SLC38A1	Solute carrier family 38, member 1.	3,97	2
FLJ20152	Hypothetical protein FLJ20152, transcript variant 1.	3,94	2
STAR	Steroidogenic acute regulator, nuclear gene encoding mitochondrial protein, transcript variant 1.	3,93	2
LMOD1	Leiomodin 1 (smooth muscle).	3,92	2
HAND2	Heart and neural crest derivatives expressed 2.	3,91	2
IMPA2	Inositol(myo)-1(or 4)-monophosphatase 2.	3,91	2
CXCL13*	Chemokine (C-X-C motif) ligand 13 (B-cell chemoattractant).	3,90	3

Desarrollo de una Herramienta Molecular Diagnóstica de la Receptividad Endometrial

Simbolo del gen	Nombre del gen	FC	n° sondas
NP	Nucleoside phosphorylase.	3,85	1
DEPDC6	DEP domain containing 6.	3,80	1
RARRES3	Retinoic acid receptor responder (tazarotene induced) 3.	3,79	1
XCL2	Chemokine (C motif) ligand 2.	3,73	4
MT2A	Metallothionein 2A.	3,72	2
TIMP3	Tissue inhibitor of metalloproteinase 3 (Sorsby fundus dystrophy, pseudoinflammatory).	3,71	2
RRAS	Related RAS viral (r-ras) oncogene homolog.	3,70	3
CLDN10	Claudin 10, transcript variant 1.	3,68	2
AGR2	Anterior gradient 2 homolog (Xenopus laevis).	3,67	1
PROM1	Prominin 1.	3,65	3
COTL1	Coactosin-like 1 (Dictyostelium).	3,58	5
ACTA2	Actin, alpha 2, smooth muscle, aorta.	3,58	2
ID4*	Inhibitor of DNA binding 4, dominant negative helix-loop-helix protein.	3,56	1
FAM59A	Family with sequence similarity 59, member A.	3,54	2
GAS1	Growth arrest-specific 1.	3,53	2
FGB	Fibrinogen beta chain.	3,53	2
LMCD1	LIM and cysteine-rich domains 1.	3,53	2
SYNE2	Spectrin repeat containing, nuclear envelope 2, transcript variant 5.	3,52	6
TRA@	T cell receptor alpha locus, mRNA (cDNA clone MGC:70922 IMAGE:5228329), complete cds.	3,52	5
MRPS2	Mitochondrial ribosomal protein S2, nuclear gene encoding mitochondrial protein.	3,49	1
KCNK7	Potassium channel, subfamily K, member 7, transcript variant C.	3,49	2
KRT7*	Keratin 7.	3,48	2
GDF15	Growth differentiation factor 15.	3,47	2
MFAP5*	Microfibrillar associated protein 5.	3,42	2
IGF2*	Insulin-like growth factor 2 (somatomedin A), transcript variant 1.	3,41	3
<u>MT1G</u>	<u>Metallothionein 1G.</u>	3,41	3
VCAM1	Vascular cell adhesion molecule 1, transcript variant 1.	3,41	1
CLDN4	Claudin 4.	3,41	3
PTPRR*	Protein tyrosine phosphatase, receptor type, R , transcript variant 1.	3,40	4
SCYE1	Small inducible cytokine subfamily E, member 1 (endothelial monocyte-activating).	3,40	2
DDX52	DEAD (Asp-Glu-Ala-Asp) box polypeptide 52, transcript variant 2.	3,38	3
HAL	Histidine ammonia-lyase.	3,37	4
LEPREL1	Leprecan-like 1.	3,33	2
ADAMTS1	A disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motif, 1.	3,32	3
TH	Tyrosine hydroxylase (TH), transcript variant 3.	3,31	4
PROS1	Homo sapiens protein S (alpha).	3,28	5
EMCN	Homo sapiens endomucin.	3,27	1
FOSL2	FOS-like antigen 2.	3,26	5
PSMB10	Proteasome (prosome, macropain) subunit, beta type, 10.	3,22	2

Tabla 2. Genes que componen el ERA.

Simbolo del gen	Nombre del gen	FC	n° sondas
ACADSB	Acyl-Coenzyme A dehydrogenase, short/branched chain (ACADSB), nuclear gene encoding mitochondrial protein.	3,22	3
S100A1	S100 calcium binding protein A1.	3,20	1
AMIGO2*	Adhesion molecule with Ig-like domain 2.	3,20	1
HRASLS3*	HRAS-like suppressor 3.	3,18	1
HMHA1	Histocompatibility (minor) HA-1.	3,17	1
POLD4	Polymerase (DNA-directed), delta 4.	3,16	3
CD7	CD7 antigen (p41).	3,15	3
ARHGDIB	Rho GDP dissociation inhibitor (GDI) beta.	3,14	2
LRFN4	Leucine rich repeat and fibronectin type III domain containing 4.	3,14	2
GBP2*	Guanylate binding protein 2, interferon-inducible.	3,14	2
GABARAPL1	GABA(A) receptor-associated protein like 1.	3,13	3
MT1H	Metallothionein 1H.	3,12	1
FXYP2*	FXYP domain containing ion transport regulator 2, transcript variant b.	3,09	2
RNASE4	Ribonuclease, RNase A family, 4, transcript variant 1.	3,09	2
PPARGC1A	Peroxisome proliferative activated receptor, gamma, coactivator 1, alpha.	3,09	4
CYBRD1	Cytochrome b reductase 1.	3,08	3
ENPEP	Glutamyl aminopeptidase (aminopeptidase A).	3,05	2
CES1	Carboxylesterase 1 (monocyte/macrophage serine esterase 1), transcript variant 3.	3,04	1
ATP6V1A	ATPase, H+ transporting, lysosomal 70kDa, V1 subunit A.	3,04	3
ASS1*	Argininosuccinate synthetase (ASS), transcript variant 2.	3,03	4
IL15	Interleukin 15, transcript variant 1.	3,03	1
BCL6*	B-cell CLL/lymphoma 6 (zinc finger protein 51), transcript variant 2.	3,02	2
EVC	Tissue-type heart Ellis-van Creveld syndrome protein (EVC) mRNA, complete cds. [AF216184]	3,02	2
IL2RB	Interleukin 2 receptor, beta.	3,01	3
EFEMP1	EGF-containing fibulin-like extracellular matrix protein 1, transcript variant 1.	3,00	2
TRH	Thyrotropin-releasing hormone.	-21,69	2
HLA-DOB	Major histocompatibility complex, class II, DO beta.	-11,06	5
ATP6V0E2	ATPase, H+ transporting V0 subunit e2.	-9,82	2
CSRP2	Cysteine and glycine-rich protein 2.	-9,62	2
OLFM4	Olfactomedin 4.	-9,35	2
SLC15A2	Solute carrier family 15 (H+/peptide transporter), member 2.	-8,44	2
CALB2	Calbindin 2, 29kDa (calretinin), transcript variant CALB2.	-8,40	1
SFRP4*	Secreted frizzled-related protein 4.	-8,33	1
CTNNA2	Catenin (cadherin-associated protein), alpha 2.	-8,02	3
NR4A2	Nuclear receptor subfamily 4, group A, member 2, transcript variant 1.	-7,05	3
DUOX1	Dual oxidase 1, transcript variant 1.	-6,61	2
KIF20A	kinesin family member 20A.	-6,33	1
PENK	Proenkephalin.	-6,16	3
POSTN	Periostin, osteoblast specific factor.	-6,04	2

Simbolo del gen	Nombre del gen	FC	n° sondas
LRP4	Low density lipoprotein receptor-related protein 4.	-5,87	2
SLC16A6	Solute carrier family 16 (monocarboxylic acid transporters), member 6.	-5,61	3
GPR64	G protein-coupled receptor 64.	-5,52	4
RANBP17*	RAN binding protein 17.	-5,50	2
EDN3	Endothelin 3, transcript variant 3.	-5,30	4
C10orf3	Chromosome 10 open reading frame 3.	-5,24	1
CENPE	Centromere protein E, 312kDa.	-5,20	2
KIAA0888*	KIAA0888 protein.	-5,17	5
COL16A1*	Collagen, type XVI, alpha 1.	-4,89	2
GALNT12*	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 12 (GalNAc-T12).	-4,85	3
ANK3	Ankyrin 3, node of Ranvier (ankyrin G), transcript variant 1.	-4,79	6
CAPN6*	Calpain 6 (CAPN6).	-4,76	2
HPGD	Hydroxyprostaglandin dehydrogenase 15-(NAD).	-4,72	3
SLC7A1	Solute carrier family 7 (cationic amino acid transporter, y+ system), member 1.	-4,72	2
MMP26*	matriz metalloproteinase 26.	-4,66	1
LRR17*	Leucine rich repeat containing 17, transcript variant 2.	-4,64	1
KCNG1	Potassium voltage-gated channel, subfamily G, member 1, transcript variant 1.	-4,63	5
CYP2J2	Cytochrome P450, family 2, subfamily J, polypeptide 2.	-4,54	1
MFAP2	Microfibrillar-associated protein 2, transcript variant 1.	-4,48	2
ALPL	Alkaline phosphatase, liver/bone/kidney.	-4,44	4
CDC20	CDC20 cell division cycle 20 homolog (S. cerevisiae).	-4,39	3
GREM2	Gremlin 2 homolog, cysteine knot superfamily (Xenopus laevis).	-4,30	5
SOX17*	SRY (sex determining region Y)-box 17.	-4,24	3
TACC3	Transforming, acidic coiled-coil containing protein 3.	-4,08	1
HEY2*	Hairy/enhancer-of-split related with YRPW motif 2.	-4,07	4
PAQR4*	Progesterone and adipoQ receptor family member IV.	-4,06	4
MGC11242*	Hypothetical protein MGC11242.	-4,03	1
HSD11B2	Hydroxysteroid (11-beta) dehydrogenase 2.	-4,03	1
KIF11	kinesin family member 11.	-4,01	2
NRG2*	Neuregulin 2, transcript variant 5.	-4,00	5
OLFM1	Olfactomedin 1, transcript variant 3.	-3,99	4
COBL	Cordon-bleu homolog (mouse).	-3,99	2
TOP2A	Topoisomerase (DNA) II alpha 170kDa.	-3,98	1
KNTC2	kinetochore associated 2.	-3,88	3
KIAA0802*	KIAA0802 protein.	-3,85	3
SERPINA5*	Serpin peptidase inhibitor, clade A (alpha-1 antitrypsin, antitrypsin), member 5.	-3,84	2
DLG7	Discs, large homolog 7 (Drosophila).	-3,84	1
ASPM	Asp (abnormal spindle)-like, microcephaly associated (Drosophila).	-3,79	4
SPDEF	SAM pointed domain containing ets transcription factor.	-3,78	1
KMO	kynurenine 3-monooxygenase (kynurenine 3-hydroxylase).	-3,78	3

Tabla 2. Genes que componen el ERA.

Símbolo del gen	Nombre del gen	FC	n° sondas
CREB3L1	cAMP responsive element binding protein 3-like 1.	-3,77	2
ATP1B1	ATPase, Na ⁺ /K ⁺ transporting, beta 1 polypeptide.	-3,73	4
PECI*	Peroxisomal D3,D2-enoyl-CoA isomerase, transcript variant 2.	-3,71	1
IDH1	Isocitrate dehydrogenase 1 (NADP ⁺), soluble.	-3,68	3
CBR3	Carbonyl reductase 3.	-3,66	3
C11orf8*	Chromosome 11 open reading frame 8.	-3,65	1
CKB	Creatine kinase, brain.	-3,61	3
ADAMTS8	A disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motif, 8.	-3,59	2
BIRC3*	Baculoviral IAP repeat-containing 3 (BIRC3), transcript variant 1.	-3,58	2
DFNB31	Deafness, autosomal recessive 31.	-3,54	3
KHDRBS3	KH domain containing, RNA binding, signal transduction associated 3.	-3,51	2
GALNT4*	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 4 (GalNAc-T4).	-3,50	4
TMEPAI	Transmembrane, prostate androgen induced RNA, transcript variant 1.	-3,50	3
KIF4A	kinesin family member 4A.	-3,50	1
KCNJ2	Potassium inwardly-rectifying channel, subfamily J, member 2.	-3,47	2
BARD1*	BRCA1 associated RING domain 1.	-3,47	2
MAP2K6	Mitogen-activated protein kinase kinase 6, transcript variant 1.	-3,45	2
OFD1	Oral-facial-digital syndrome 1.	-3,44	3
PLA1A	Phospholipase A1 member A.	-3,42	2
RAD54B	RAD54 homolog B (S. cerevisiae), transcript variant 1.	-3,42	3
PBK	PDZ binding kinase.	-3,42	2
NDRG2*	NDRG family member 2, transcript variant 1.	-3,40	2
ECM1	Extracellular matrix protein 1, transcript variant 2.	-3,34	3
PRC1	Protein regulator of cytokinesis 1 (PRC1), transcript variant 1.	-3,29	1
MSX1	Msh homeo box homolog 1 (Drosophila).	-3,26	3
CCNB2	Cyclin B2.	-3,23	3
LRRC1	Leucine rich repeat containing 1.	-3,23	3
SORD*	Sorbitol dehydrogenase.	-3,21	5
EPHB3	EPH receptor B3.	-3,20	2
TMSL8	Thymosin-like 8.	-3,20	2
RASSF2*	Ras association (RalGDS/AF-6) domain family 2, transcript variant 1.	-3,19	1
TTC21B	mRNA for KIAA1992 protein. [AB082523]	-3,17	2
OPRK1	Opioid receptor, kappa 1.	-3,17	4
TMEM16A*	Transmembrane protein 16 ^a	-3,13	4
CRABP2*	Cellular retinoic acid binding protein 2.	-3,10	3
FLJ10719	Hypothetical protein FLJ10719.	-3,10	1
PRKCQ	Protein kinase C, theta.	-3,08	1
CDC2	Cell division cycle 2, G1 to S and G2 to M, transcript variant 1.	-3,06	2
BUB1B	BUB1 budding uninhibited by benzimidazoles 1 homolog beta (yeast).	-3,05	2
STEAP4*	STEAP family member 4.	-3,04	2

Símbolo del gen	Nombre del gen	FC	n° sondas
HEY1	Hairy/enhancer-of-split related with YRPW motif 1.	-3,03	3

Tabla 2. Genes que componen el ERA. Se muestran los 238 genes seleccionados. Se indican el símbolo y el nombre del gen, la tasa de cambio (Fold Change, FC) según la matriz de datos original (transcriptoma de genoma completo) de las muestras para la selección de genes, así como el número de sondas presentes pertenecientes a cada gen (anexo II). Los genes validados por RT-PCR aparecen en subrayado. En **negrita** se marcan los genes comunes a los 25 genes incluidos como biomarcadores de receptividad endometrial, según el estudio comparativo de Horcajadas y cols., en el 2007. Los genes que forman parte de la firma transcriptómica de receptividad endometrial están indicados con un asterisco.

- **Criterio funcional**

Con la finalidad de estudiar el GO de los genes seleccionados para ser incluidos en el ERA, se analizó la relevancia biológica de la selección estadística. Para ello, se estudió qué términos estaban representados de forma significativa más de lo que cabría esperar con respecto al conjunto de todos los genes del genoma humano. Los términos más representados de forma significativa (p -valor $<0,05$) en la selección de genes del ERA (tabla 3), incluían en el BP los procesos relativos al sistema inmune como la respuesta al estrés, la respuesta defensiva, la respuesta inmune humoral e innata así como la respuesta a heridas. Otros BP significativos fueron la circulación, la respuesta a estímulos externos, el ciclo celular, y la adhesión celular. En la sección de MF, tres términos estaban más representados: la actividad oxido reductasa, carbohidratos de unión y receptores de unión. En la sección relativa a CC, el huso mitótico y más concretamente los microtúbulos del huso mitótico estaban más representados de lo que cabría esperar.

Términos más representados	Nivel GO	% genes	Símbolo del gen perteneciente al término	P-valor ajustado
Respuesta inmune (GO:0006955)	BP3	14,94	DPP4 OPRK1 C3 CRISP3 HLA-DOB C4BPA NR4A2 XCL2 LIF CXCL14 GZMA XCL1 CXCL13 KLRC1 IL15 INDO CD7 BCL6 SERPING1 CLU CTSW GBP2 PSMB10 DEFB1 CFD ARHGDIB	4,14E-04
Respuesta al estrés (GO:0006950)	BP3	15,52	ADRA2A C3 PROS1 AOX1 C4BPA HPSE DUOX1 SOD2 FGB SCYE1 CXCL14 MAP2K6 RAD54B THBD CXCL13 BARD1 IDH1 ANG BCL6 GPX3 GADD45A SERPING1 CLU PPARGC1A PENK CFD CDC2	9,49E-03
Mecanismos de defensa (GO:0006952)	BP3	10,34	C3 AOX1 CRISP3 C4BPA HPSE SCYE1 CXCL14 GNLY CXCL13 IL15 CD7 BCL6 SERPING1 CLU OFD1 DEFB1 PENK CFD	9,76E-03
Circulación (GO:0008015)	BP3	4,60	EDNRB FGB XCL2 CYP2J2 S100A1 SERPING1 EDN3 TH	9,76E-03
Respuesta a estímulos externos (GO:0009605)	BP3	10,34	C4BPA HPSE FGB XCL2 SCYE1 CXCL14 THBD XCL1 CXCL13 BCL6 SERPING1 CLU PPARGC1A DEFB1 CFD C3 PROS1 AOX1	9,76E-03
Comportamiento (GO:0007610)	BP3	6,32	OPRK1 XCL2 SCYE1 CXCL14 XCL1 CXCL13 MAOA TRH TH DEFB1 PENK	3,36E-02
Ciclo celular (GO:0007049)	BP3	12,07	HRASLS3 RASSF2 MAP2K6 RAD54B CCNB2 GAS1 PBK BARD1 G0S2 HPGD KIF11 PRC1 ASPM GADD45A IGF2 RPRM DLG7 CORO1A BUB1B CDC2 CDC20	3,59E-02
Adhesión celular (GO:0007155)	BP3	10,34	CRISP3 MUC16 HAMP2 CTNNA2 CLDN10 EMCN BCL6 CLDN4 VCAM1 COBL COMP AMIGO2 THBS2 POSTN SPP1 LAMB3 COL16A1 ARHGDIB	4,18E-02
Estructura anatómica del desarrollo (GO:0048856)	BP3	2,41	POSTN ID4 EVC OLFM1 SPP1 EDN3 LAMB3 TH CRABP2 NR4A2 NDRG2 DUOX1 EDNRB ALPL EFNA1 IGFBP1 CDA TAGLN LIF HAND2 ADAMTS1 PRKCQ CTNNA2 ENPEP CXCL13 DKK1 LRP4 ANG HEY2 IER3 IL15 EMCN BCL6 CSRP2 HEY1 COBL COMP IGF2 MSX1	4,52E-02
Respuesta inmune humoral (GO:0006959)	BP4	6,02	C3 C4BPA NR4A2 XCL1 KLRC1 SERPING1 CLU PSMB10 DEFB1 CFD	7,06E-03

Términos más representados	Nivel GO	% genes	Símbolo del gen perteneciente al término	P-valor ajustado
Señalización célula a célula (GO:0007267)	BP4	13,25	OPRK1 SLC1A1 HSD11B2 NR4A2 GDF15 EFNA1 XCL2 SCYE1 LIF CXCL14 XCL1 ENPEP CXCL13 NRG2 MAOA TRH IL15 GABARAPL1 DLG7 EDN3 TH PENK	7,06E-03
Respuesta inmune innata (GO:0045087)	BP4	4,22	SERPING1 CLU DEFB1 CFD C3 CRISP3 C4BPA	7,90E-03
Ciclo celular mitótico (GO:0000278)	BP4	6,63	DLG7 CORO1A BUB1B CDC2 CDC20 CCNB2 GAS1 PBK KIF11 PRC1 ASPM	2,99E-02
Respuesta a heridas (GO:0009611)	BP4	8,43	C4BPA HPSE FGB SCYE1 CXCL14 THBD CXCL13 BCL6 SERPING1 CLU CFD C3 PROS1 AOX1	3,17E-02
Actividad oxido reductasa (GO:0016491)	MF3	18,10	CBR3 SORD SOD2 IDH1 MAOA KMO CYP2J2 ACADSB INDO HSD11B2 LEPREL1 LAMB3 AOX1 TH CYBRD1 GPX3 HPGD DHRS3 CYP3A5 CP DUOX1	4,49E-02
Carbohidratos de unión (GO:0030246)	MF3	9,48	ADAMTS1 EMCN FAM59A THBD GALNT12 SERPINA5 HABP2 KLRC1 THBS2 ANG POSTN	4,62E-02
Receptores de unión (GO:0005102)	MF4	16,91	FGB EFNA1 XCL2 SCYE1 IGF2 ANG CXCL14 LIF TRH C3 ADAMTS1 EDN3 SPP1 PPARGC1A XCL1GABARAPL1 GDF15 CXCL13 GREM2 NRG2 PENK GAST DKK1	1,52E-02
Huso acromático (GO:0005819)	CC10	30,43	CDC2 CDC20 KIF11 PRC1 KIF4A DLG7 BUB1B	9,47E-05
Microtúbulos del huso acromático (GO:0005876)	CC11	30,77	CDC2 KIF11 PRC1 KIF4A	7,29E-04

Tabla 3. Análisis *Gene Ontology* (GO) de los genes del ERA. Se indican los términos más representados de lo que cabría esperar con respecto a su representación en el genoma completo; nivel del GO; porcentaje de genes del ERA presentes en cada término; símbolos de los genes; p-valor de la representación.

Además de estudiar el GO también se comprobó si había alguna ruta de señalización representada de forma significativa (KEEG *pathways*), o alguna interacción entre las proteínas codificadas por el transcriptoma *ERA* (interpro Swissprot) no observándose nada significativo.

6.2. Diseño del *ERA*

El *ERA* fue testado y diseñado con el soporte de la tecnología *Agilent*. Es un *microarray* de expresión de oligonucleótidos de 60 bases con un formato de ocho *microarrays* del orden de 15.000 sondas, cada uno, en un mismo portaobjetos (formato 8 x 15K). De modo que en cada portaobjetos se pueden hibridar ocho muestras distintas, una por *microarray*.

Cada *microarray* contiene 15.744 puntos e incluye las 569 sondas correspondientes a los 238 genes seleccionados replicadas ocho veces (un total de 4.552 puntos), así como 1.858 puntos de controles internos propios de *Agilent*. Los puntos incluidos totales fueron 6.410 y cubrieron el 40,7% de la capacidad del *microarray* personalizado quedando 9.334 puntos vacíos. Las 569 sondas que representan los 238 genes en el *microarray* se adjuntan en el anexo II.

6.3. Validación de la técnica de *microarrays*

El preprocesamiento de los *microarrays* ya constituyó una validación del correcto funcionamiento de la técnica. Se analizó la intensidad media de los puntos dentro de un mismo *microarray* (muestra) y se obtuvo la distribución. En la distribución de intensidades se visualizó qué cantidad de puntos tenían intensidades altas, qué cantidad medias y qué cantidad bajas. Debido a que el *ERA* posee un % de puntos vacíos, a los controles y a los genes incluidos que

variaban su intensidad de una fase del ciclo a otra (vid. punto 6.2), se esperaban muchos puntos sin intensidad, otros con una intensidad parecida en todos los arrays (los controles), y otros con intensidad variable (genes seleccionados diferencialmente expresados). En la representación del diagrama de cajas se comprobó que en todas las muestras, a excepción de tres que fueron eliminadas del estudio, las intensidades se distribuían de la forma esperada (figura 12). Todo este análisis fue indicativo de que la técnica se había producido correctamente.

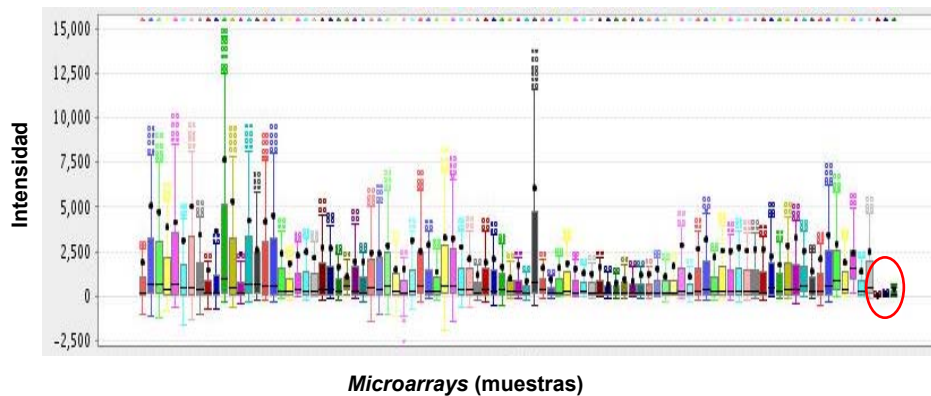


Figura 12. Distribución en cajas de las intensidades del *microarray ERA preprocesado*. Se representa en el eje de ordenadas la distribución de intensidad y en el eje de las abscisas se sitúan cada una de las muestras (*microarrays ERA*). La caja representa qué cantidad de puntos tienen más intensidad. El círculo rojo indica las tres muestras que fueron eliminadas del estudio por no poseer prácticamente señal.

- **Validación del *ERA* mediante RT-PCR**

La expresión del *ERA* quedó validada al comparar la expresión obtenida mediante RT-PCR en un subconjunto de 29 muestras control normales (vid. MM 5.1) para las que también se había realizado el *ERA* (figura 13). Este set estaba formado por 16 muestras en fase receptiva y 13 en fase no receptiva (PF n=8 y PR n=5). Estas muestras, de forma más detallada, estaban

constituidas por: muestras PF en día 8 n=2, en día 9 n=2, en día 10 n=2, en día 11 n=1 y en día 14 n=1; muestras PR en día 16 (LH+2) n=5 y muestras R en día 21(LH+7) n=16. Se realizó la media de la expresión génica para muestras pertenecientes a un mismo día.

La expresión de los 4 genes que fueron validados por ambas técnicas se muestra en la figura 13. Los genes GPX3, FXVD2, SPP1 y MT1G mostraron que aumentaban su expresión en fase receptiva con respecto a la fase proliferativa y pre-receptiva en ambas técnicas. De este modo, quedaron así validados los resultados obtenidos por la técnica de *microarrays*.

6.4. Comportamiento de los genes del ERA en el ciclo menstrual

Con objeto de caracterizar los genes seleccionados se estudió si existían grupos de genes con un mismo patrón de comportamiento en su expresión a lo largo del ciclo menstrual (figura 14). Se emplearon los datos de expresión del ERA de las 79 muestras del set de entrenamiento (PF=14, PR=14, R=39 y PS=12) (vid. MM 5.1). Las variables utilizadas fueron, la expresión de los genes ERA y el día del ciclo menstrual. El endometrio proliferativo fue etiquetado como 7, el endometrio pre-receptivo como 15, el endometrio receptivo como 21 y el endometrio post-receptivo como 25.

- **Grupos de comportamiento obtenidos**

De los nueve grupos que se establecieron, se detectaron tres tendencias generales de comportamiento (figura 14). El comportamiento seguido por los grupos 1, 4, 6, 8 y 9 fue el de genes que disminuían su expresión de la fase proliferativa a la fase pre-receptiva, donde mostraban su menor expresión, para luego aumentarla hacia la fase receptiva.

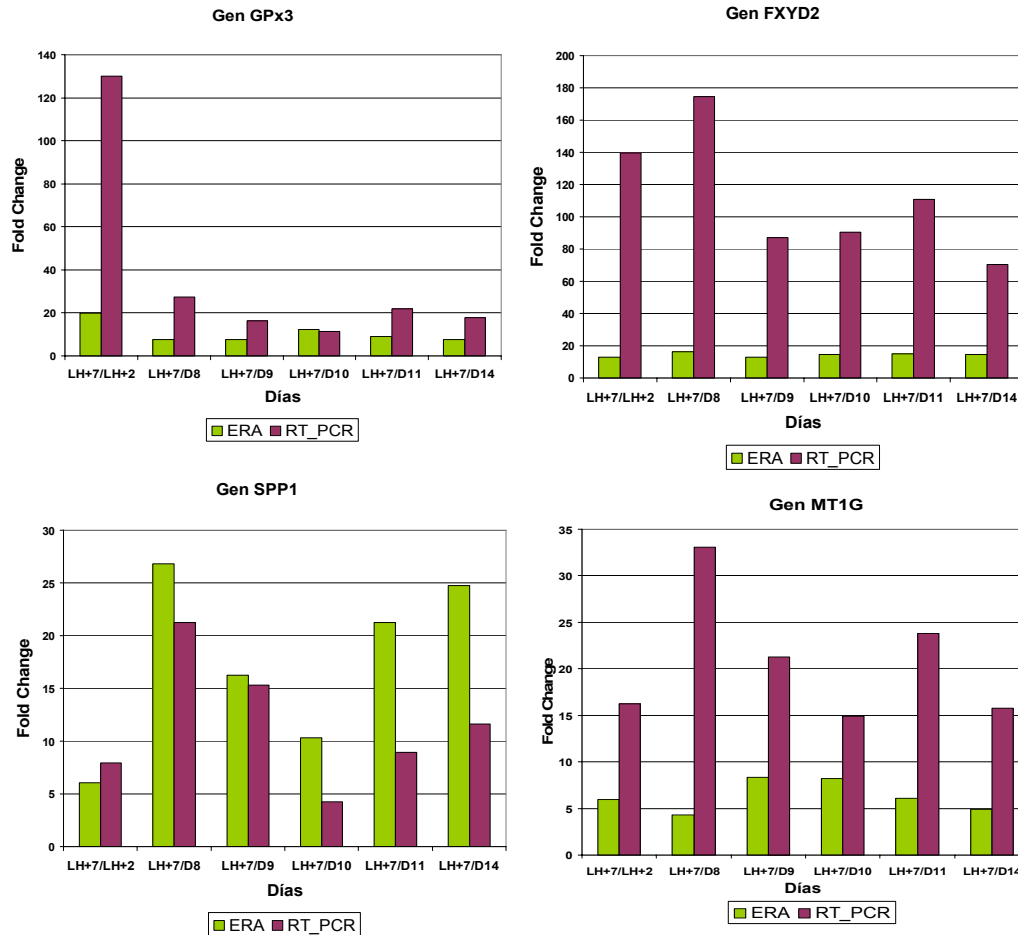


Figura 13. Validación del ERA mediante RT-PCR. Diagrama de barras que representa la tasa de cambio (*Fold Change* (FC)) (eje ordenadas) de la expresión de los cuatro genes seleccionados (GPX3, FXyD2, SPP1 y MT1G) en distintos días del ciclo menstrual, con respecto a la fase receptiva (LH+7). En el eje de abcisas se muestra la tasa de cambio de los datos de expresión del ERA (verde) y de los datos obtenidos para las mismas muestras y genes en la RT-PCR (morado). De izquierda a derecha, LH+7/LH+2, LH+7/D8, LH+7/D9, LH+7/D10, LH+7/D11 y LH+7/D14. D: día del ciclo menstrual (D8, D9, D10, D11 y D14 (fase proliferativa); LH+7: siete días después del pico de la hormona luteinizante (fase receptiva); LH+2: dos días después del pico de la hormona luteinizante (fase pre-receptiva); /: división.

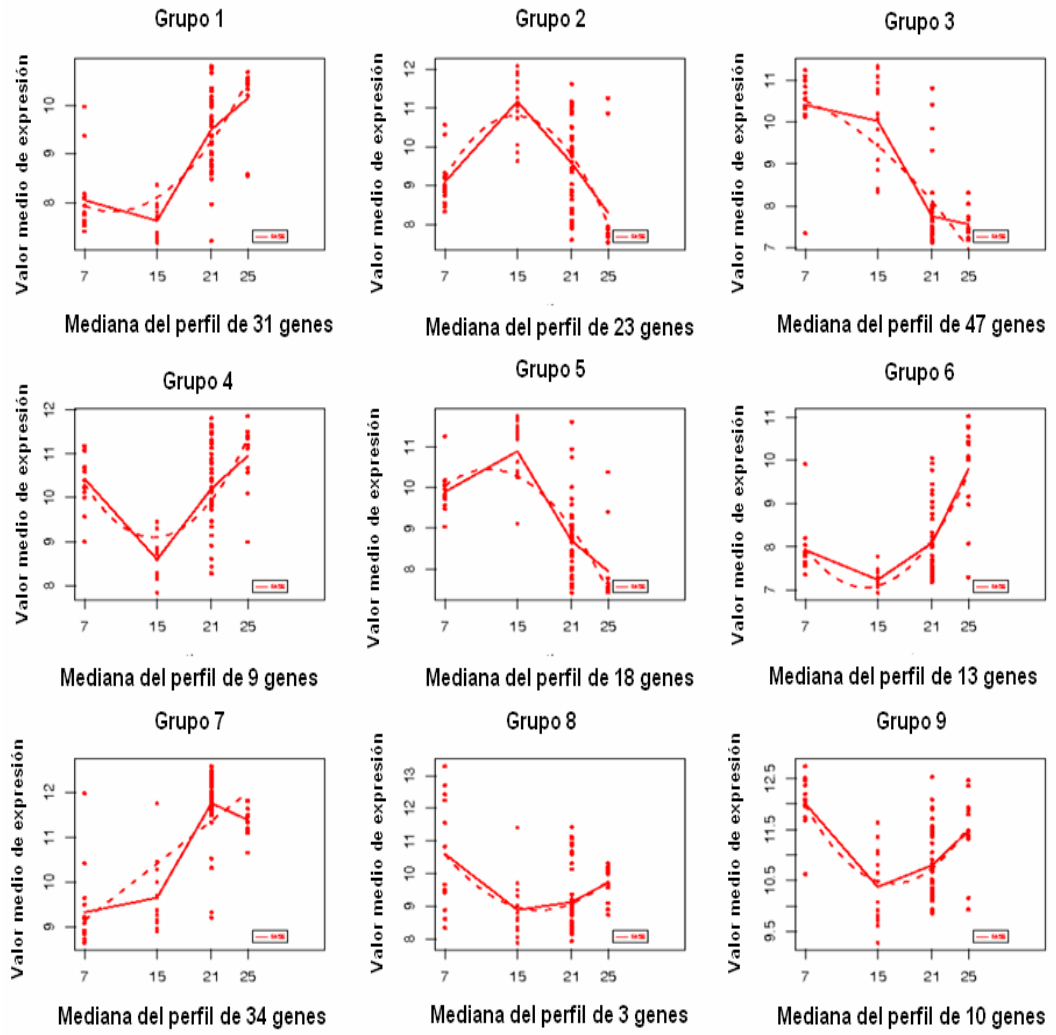


Figura 14. Comportamiento de los genes del ERA a lo largo del ciclo menstrual. El eje de las ordenadas representa el valor medio de la expresión de los genes que forman un grupo con el mismo comportamiento de expresión. En el eje de las abscisas, tenemos cada una de las fases: el día 7 como representativo de la fase proliferativa, el día 15 representando la expresión de los genes en fase pre-receptiva, el día 21 a la expresión en fase receptiva y el día 25 la expresión de los genes en fase post-receptiva.

En los grupos 1, 4, y 6 aumentaba la expresión hasta hacerla máxima en la fase post-receptiva. Mientras que en los grupo 8 y 9 se recuperaban los niveles de expresión iniciales de la fase proliferativa. La segunda tendencia general detectada fue la de los grupos 2, 3 y 5 que disminuían su expresión en la fase receptiva, siendo mínima en la fase post-receptiva. Por último, la tendencia del grupo 7, formado por genes que tienen su máxima expresión en fase receptiva, disminuía su expresión en la fase post-receptiva.

De forma más específica, el grupo 2 y el 5 poseían 23 y 18 genes respectivamente que tenían su máxima expresión en fase pre-receptiva (figura 14). El grupo 4 y el 6 estaban formados por 9 y 13 genes respectivamente con su mínima expresión en fase pre-receptiva.

Destacaron los genes del grupo 7 y los del grupo 3 por contener genes con máxima y mínima expresión en fase receptiva que es la que más interesaba por el objetivo que se perseguía. El grupo 7 estaba formado por 34 genes con su máxima expresión en fase receptiva disminuyendo esta en la fase post-receptiva. Mientras que el grupo 3 estaba constituido por 47 genes con una expresión muy baja en receptividad que se mantenía en la fase post-receptiva.

- **Análisis funcional de los grupos relevantes**

También se realizó un análisis funcional de aquellos grupos que se consideraron relevantes por estar más relacionados con la receptividad endometrial. El análisis fue llevado a cabo con los mismos parámetros que el análisis funcional de los genes del *ERA* (vid. MM 5.3) mediante la herramienta bioinformática FatiGO (Al-Shahrour y cols., 2004) de Babelomics Suite versión 3.2 (Al-Shahrour y cols., 2005) en cuyo análisis no inclusivo se estudiaron

todos los términos del GO de BP, MF y CC entre los niveles del GO del 3-12 que tuvieran anotados un número de genes comprendido entre 5-1000. Se consideraron significativos los términos del GO con un p-valor ajustado por Benjamin-Hochberg (FDR) (Benjamini, 1995) por debajo de 0,05.

Se escogieron como grupos más representativos de la receptividad endometrial: el grupo 7, con 34 genes que poseían su máxima expresión en receptividad con respecto al resto del ciclo menstrual, y el grupo 3, con 47 genes que poseían su menor expresión en receptividad y se mantenían con una expresión baja en la fase post-receptiva.

El análisis funcional del grupo 7, como genes más expresados en receptividad, y del grupo 3, como genes menos expresados en receptividad, con respecto al patrón de comportamiento que mostraban en las distintas fases del ciclo, reveló diversas funciones representadas de forma significativa (p -valor $<0,05$) (tabla 4 y 5 respectivamente).

Debido al escaso número de genes y al estado de la anotación de las bases de datos, el análisis funcional del grupo 7 no produjo unos resultados interpretables estadísticamente (tabla 4). De modo que no se encontraron términos significativos en BP, aunque sí que se detectaron algunas MF representadas más de lo que cabría esperar de forma significativa.

Las MF estuvieron relacionadas con unión a iones cobre y unión a iones cadmio, así como con los receptores de unión. También un CC quedó más representado de lo que cabría esperar de forma significativa, el componente extracelular (tabla 4).

Los genes del grupo 3 resultaron estar relacionados con los BP de ciclo celular y mitosis, así como con los BP de microtúbulos asociados del huso mitótico y la organización y biogénesis del citoesqueleto (tabla 5). La MF más representada de lo que cabría esperar fue la actividad motora de los microtúbulos, y los CC fueron los microtúbulos, el citoesqueleto y los complejos asociados a microtúbulos.

Términos más representados	Nivel GO	% genes	Símbolo del gen perteneciente al término	P-valor ajustado
Unión a iones de cobre (GO:0005507)	MF	16,0	ANG MT1H CP MT1G	5,15E-03
Receptores de unión (GO:0005102)	MF	32,0	EFNA1 ANG CXCL14 SPP1 GABARAPL1 GDF15 CXCL13 DKK1	1,54E-02
Unión a iones de Cadmio (GO:0046870)	MF	8,0	MT1G MT1H	3,65E-02
Espacio extracelular (GO:0005615)	CC	36,8	ANG C4BPA GDF15 CXCL14 SPP1 CP CXCL13	4,46E-02

Tabla 4. Análisis *Gene Ontology* (GO) de los genes del patrón de comportamiento del grupo 7. Se indican los términos más representados de lo que cabría esperar con respecto a su representación en el genoma completo; nivel del GO; porcentaje de genes del grupo 7 presentes en cada término; símbolos de los genes; p-valor de la representación.

Además de estudiar el GO de los grupos 7 y 3, también se analizó si había alguna ruta de señalización representada de forma significativa (KEEG *pathways*) o alguna interacción entre las proteínas codificadas por el transcriptoma *ERA* (interpro Swissprot), no observándose nada significativo.

Términos más representados	Nivel GO	% genes	Símbolo del gen perteneciente al término	P-valor ajustado
Ciclo celular mitótico (GO:000278)	BP	27,27	DLG7 BUB1B CDC2 CDC20 CCNB2 PBK KIF11 PRC1 ASPM	8,60E-6
Mitosis (GO:0007067)	BP	24,24	CCNB2 PBK KIF11 ASPM DLG7 BUB1B CDC2 CDC20	8,60E-6
División celular (GO:0051301)	BP	21,21	PRC1 ASPM BUB1B CDC2 CDC20 CCNB2 KIF11	3,49E-4
Ciclo celular (GO:0007049)	BP	33,33	RASSF2 CCNB2 PBK BARD1 KIF11 PRC1 ASPM DLG7 BUB1BCDC2 CDC20	8,13E-4
Procesos basados en microtúbulos (GO:0007017)	BP	18,18	KIF20A CENPE KIF11 PRC1 KIF4A BUB1B	1,92E-3
Organización y biogénesis del huso mitótico (GO:0007051)	BP	9,09	KIF11 PRC1 BUB1B	8,73E-3
Organización y biogénesis del citoesqueleto (GO:0007010)	BP	21,21	KIF20A CENPE KIF11 PRC1 KIF4A TMSL8 BUB1B	1,90E-2
Movimientos basados en microtúbulos (GO:0007018)	BP	12,12	KIF20A CENPE KIF11 KIF4A	1,90E-2
Regulación y progresión a través del ciclo celular (GO:0000074)	BP	21,21	DLG7 BUB1B CDC2 CDC20 RASSF2 CCNB2 BARD1	2,71E-2
Regulación ciclo celular (GO:0051726)	BP	21,2	DLG7 BUB1B CDC2 CDC20 RASSF2 CCNB2 BARD1	2,71E-2
Regulación de la mitosis (GO:0007088)	BP	9,1	DLG7 BUB1B CDC2	3,10E-2
Actividad motora microtúbulos (GO:0003777)	MF	12,9	KIF11 KIF20A KIF4A CENPE	3,24E-2

Términos más representados	Nivel GO	% genes	Símbolo del gen perteneciente al término	P-valor ajustado
Huso acromático (GO:0005819)	CC	28	CDC2 CDC20 KIF11 PRC1 KIF4A DLG7 BUB1B	2,56E-7
Microtúbulos del citoesqueleto (GO:0015630)	CC	40	CENPE KIF11 PRC1 KIF20A KIF4A DLG7 CCNB2 BUB1B CDC2 CDC20	8,84E-6

Tabla 5. Análisis *Gene Ontology* (GO) de los genes del patrón de comportamiento del grupo 3. Se indican los términos más representados de lo que cabría esperar con respecto a su representación en el genoma completo; nivel del GO; porcentaje de genes del grupo 3 presentes en cada término; símbolos de los genes; p-valor de la representación.

6.5. Definiendo la ERS

- **Criterio**

Para la obtención de la ERS se seleccionaron aquellos genes diferencialmente expresados específicos de la fase receptiva del ciclo menstrual (LH+7). Con este fin, se utilizó el estudio de expresión génica diferencial de la expresión del ERA entre R vs. PF; R vs. PR; y R vs. PS de las muestras del set de entrenamiento (PF=14, PR=14, R=39 y PS=12, vid. MM 5.1). De la comparación R vs. PF, se obtuvieron 176 genes diferencialmente expresados (92 aumentados y 84 disminuidos). De la comparación R vs. PR, resultaron 200 genes (110 al alza y 90 a la baja) y de la tercera comparación R vs. PS se obtuvieron 94 genes (56 genes aumentados y 38 disminuidos) (figura 15). Las muestras PR fueron las que más se distinguían de las R, seguidas por las PF. Las que más se parecían a las R fueron las PS.

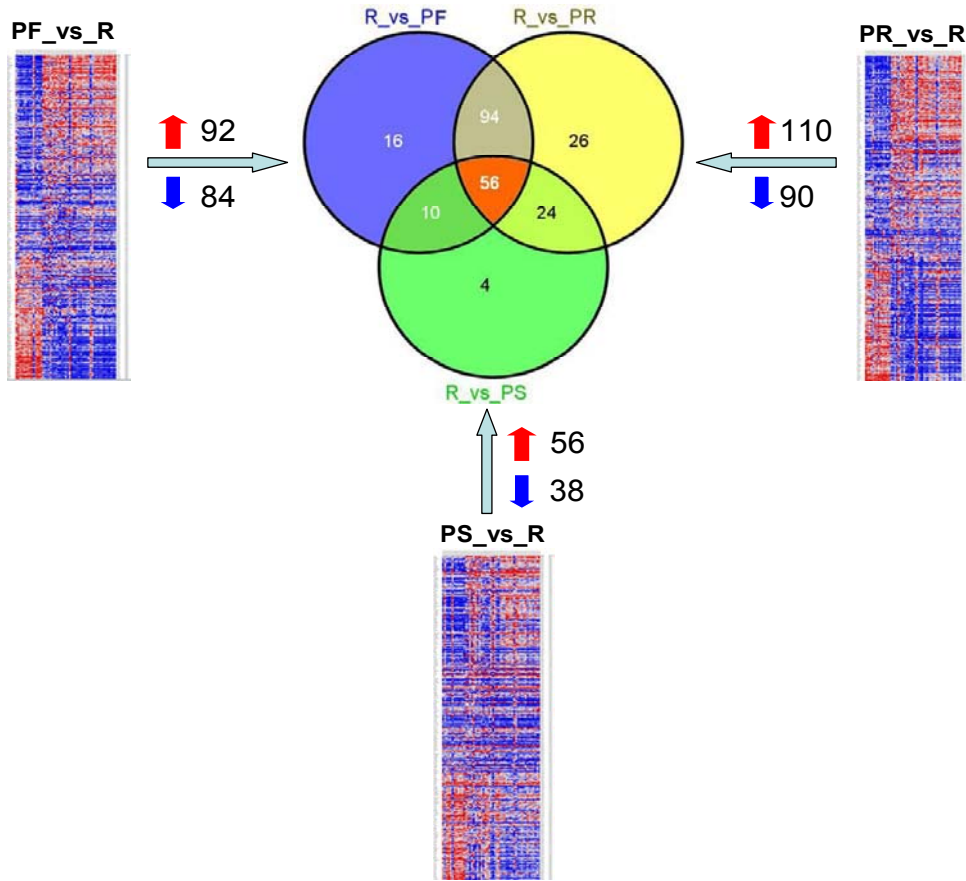


Figura 15. Firma transcriptómica de la receptividad endometrial (ERS). Transcriptoma del ERA de las muestras en fase proliferativa (PF); transcriptoma del ERA de las muestras en fase pre-receptiva (PR); transcriptoma del ERA de las muestras en fase receptiva (R); transcriptoma del ERA de las muestras en fase post-receptiva (PS). Tras realizar los estudios de expresión génica diferencial de la comparación R vs. PF, R vs. PR y de R vs. PS se indica el número de genes diferencialmente expresados al alza (flecha roja con punta hacia arriba) y los diferencialmente expresados a la baja (flecha azul con punta hacia abajo), así como el *heatmap* de genes diferencialmente expresados para cada comparación. En el *heatmap* el eje de ordenadas corresponde a cada uno de los genes diferencialmente expresados ordenados de mayor a menor por el valor del

estadístico; el eje de abscisas representa cada una de las muestras. Los cuadrados rojos indican la expresión mayor para el gen y los cuadrados azules la expresión menor. Del diagrama de Venn de los grupos de genes diferencialmente expresados obtenidos se generó el conjunto intersección común a todas las comparaciones (triángulo invertido naranja). 56 genes resultaron expresados diferencialmente en fase receptiva con respecto a todo el resto del ciclo menstrual y constituyen la *ERS*.

De los datos obtenidos, se seleccionaron aquellos genes expresados diferencialmente en fase receptiva con respecto a todas las fases del ciclo menstrual. Para ello, se escogieron los genes diferencialmente expresados que se obtuvieron de la intersección del diagrama de Venn entre todos los grupos de genes diferencialmente expresados en cada comparación (figura 15). Estos genes están etiquetados en la tabla 2 con un asterisco en la columna del símbolo del gen y son los que se han definido como la *ERS*.

- **Análisis funcional**

Se realizó el análisis funcional de la *ERS*, comparándola con el resto del genoma, utilizando la misma herramienta y parámetros que en el resto de estudios funcionales, pero no se encontró ningún término significativo por debajo de 0,05.

- **Genes del *ERA* expresados diferencialmente**

Con los genes diferencialmente expresados de las comparaciones R vs. P, R vs. PR y R vs. PS, también se realizó un diagrama de Venn con cada grupo y los genes que incluía el *ERA* (figura 16). Se comprobó que prácticamente todos los genes que formaban parte del *ERA* estaban diferencialmente expresados en la fase receptiva con respecto a alguna de las

otras fases del ciclo menstrual, a excepción de: C3, CES1, CRISP3, EMCN, FGB, GZMA, KIAA0888, LRFN4, PROS1 y TRA@.

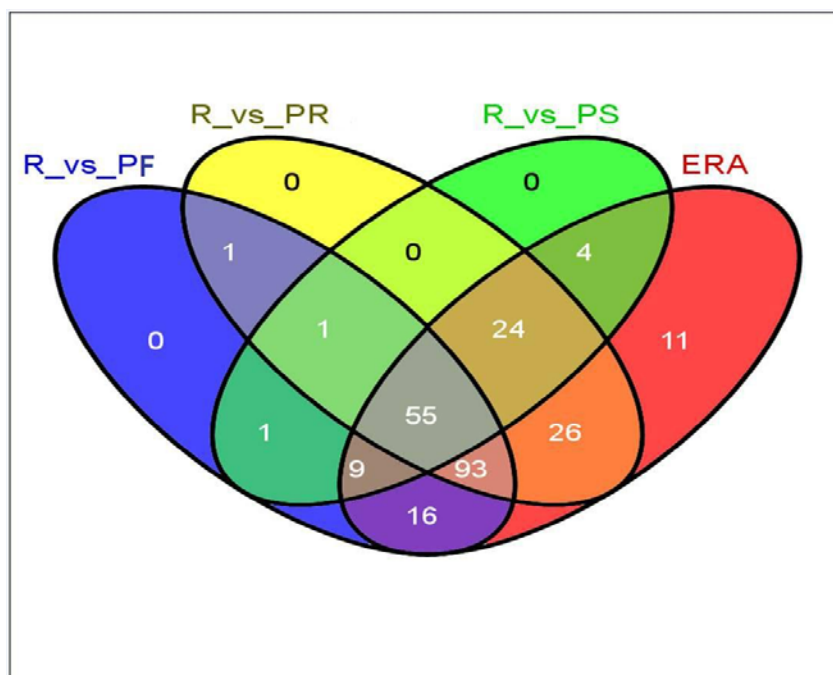


Figura 16. Genes diferencialmente expresados. Diagrama de Venn en el que se representan el grupo de genes del ERA y los tres grupos de genes diferencialmente expresados resultantes de los estudios de expresión génica diferencial entre las muestras receptivas (R) vs. proliferativas (PF), entre las R vs. las pre-receptivas (PR), y entre las R vs. las post-receptivas (PS).

También se realizó un ANOVA entre los cuatro grupos (PF, PR, R y PS) para la expresión de genes del ERA que mostró que todos los genes tenían una expresión génica diferencial significativa, a excepción de un conjunto de genes (tabla 6).

Símbolo del gen	Estadístico	p-valor ajustado
SLC15A1	2,627	0,059
EFEMP1	2,095	0,112
TRA@	1,431	0,248
PROS1	1,399	0,256
FGB	1,365	0,265
VCAM1	1,199	0,321
XCL1	1,097	0,360
KIAA0888	0,855	0,472
EMCN	0,198	0,901

Tabla 6. Genes no expresados diferencialmente tras el ANOVA. La tabla muestra los genes no expresados diferencialmente de forma significativa tras realizar un ANOVA entre la expresión del ERA de los cuatro grupos (PF, PR, R y PS). En **negrita** se distinguen aquellos genes que son comunes con respecto a los genes que no muestran diferencias tras el análisis de expresión génica diferencial (figura 16).

6.6. Análisis de agrupamiento jerárquico

6.6.1. Modelo de dataje endometrial

El agrupamiento jerárquico se realizó con las muestras control normales PF, PR, R y PS (vid. MM 5.1). En el análisis se observó que las muestras se separaban en bloques bien definidos en función de los genes (figura 17). Por tanto, parecían muestras bastante homogéneas dentro de cada clase, y que se diferenciaban entre las distintas clases.

Una agrupación correspondió al grupo “receptivas”, donde colocalizaron 36 de las 39 muestras R. Otras dos agrupaciones próximas denominadas “pre-receptivas”, por encontrarse 13 de las 14 muestras PR (7 en un grupo y 6 en el otro). Se obtuvo otro agrupamiento intercalado entre las “pre-receptivas”, correspondiente a “proliferativas”, por contener 12 de las 14 muestras PF totales. Por último, un grupo de “post-receptivas” que incluía 9 de las 12 muestras PS. Las tres muestras R que quedaron alejadas de su grupo “receptivo” se situaron entre la agrupación “pre-receptivas” y “proliferativas”.

Tanto la muestra PR como la PF que no se situó dentro de las suyas, lo hizo en el grupo “receptivas”. Las otras tres muestras PS se incluyeron dos dentro de la agrupación “receptivas” y una dentro de la agrupación “pre-receptivas” (figura 17).

En ninguno de los casos las muestras se agrupaban en función de la tanda experimental, sino en función de su fase del ciclo, comprobando que no existía un *batch effect* (figura 17).

6.6.2. Modelo de clasificación patológica

El análisis de agrupamiento jerárquico se realizó con las muestras control R-FI (n=11) con respecto a las muestras control R (n= 39) (vid. MM 5.1). El análisis mostró que no se separaban en dos bloques bien definidos en función de los genes del *ERA* (figura 18). Luego, según el agrupamiento no fue posible distinguir patología.

Hubo tres muestras R-FI (era7_S2; era3_S5 y era7_S1) que se agrupaban juntas y de forma alejada al resto de R-FI (figura 18). Las muestras

R-FI no se agrupaban en función de la tanda experimental comprobando que no existía un *batch effect* (figura 18).

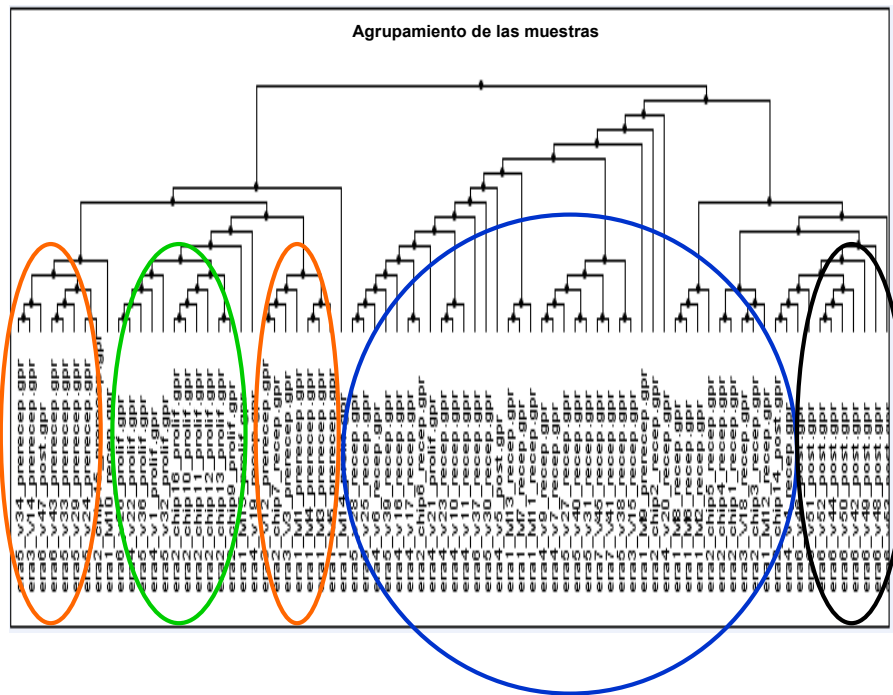


Figura 17. Análisis de agrupamiento jerárquico de las muestras control del ciclo menstrual. Se agrupan las distintas muestras que forman el set de entrenamiento para el dataje endometrial en función de la expresión de los 238 genes del ERA, y según el UPGMA con el coeficiente de correlación de Pearson como función de la distancia. Se observa un grupo correspondiente a las muestras “receptivas” compuesto por 36 muestras R, 2 muestras PS, 1 PF y 1 PR (círculo azul). Otros dos grupos próximos de “pre-receptivas”; un grupo que incluye 7 muestras PR con 1 muestra PS intercalada y el otro grupo con 6 muestras PR y 1 muestra R (círculos naranjas). También se encuentra otro conjunto en medio de las “pre-receptivas” correspondiente a la agrupación “proliferativas” con 12 muestras PF y ninguna de otra fase del ciclo menstrual (círculo verde). Por último, se observa el grupo “post-receptivas” constituido por 9 muestras PS

y ninguna de otra fase del ciclo (círculo negro). En el etiquetado de las muestras el número que va detrás de “era” (eraX_) indica la tanda experimental. Donde x= [1,7].

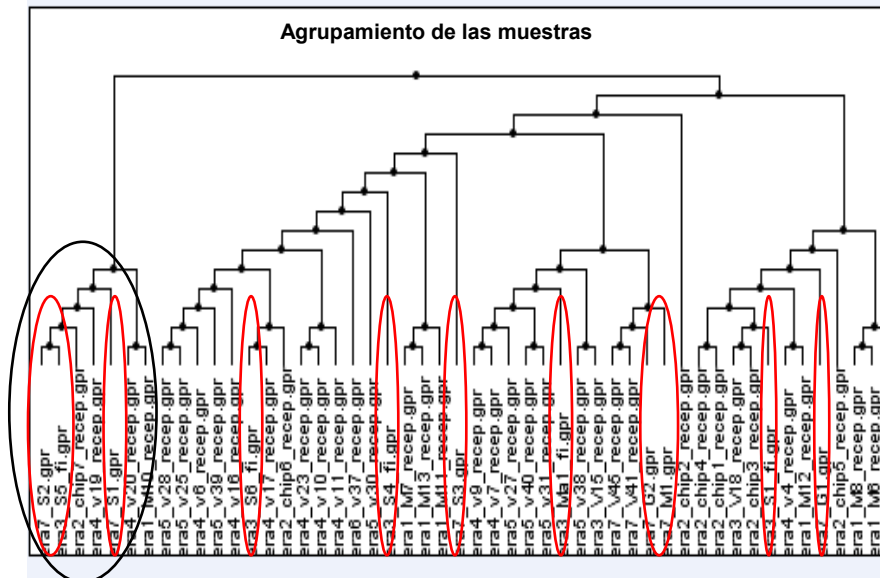


Figura 18. Análisis de agrupamiento jerárquico de las muestras control receptoras normales (R) y de las receptoras con fallo de implantación (R-FI). Se representa el agrupamiento de las muestras utilizadas para entrenar al predictor para el modelo de clasificación patológica. El agrupamiento se realizó en función de la expresión de los 238 genes del ERA, y con el UPGMA con el coeficiente de correlación de Pearson como función de la distancia. En rojo se hayan marcadas las muestras R-FI (era7_S2, era3_S5, era7_S1, era3_S6, era3_S4, era7_S3, era3_Ma1, era7_G2, era7_M1, era3_S1 y era7_G1), y se observa como se distribuyen entre las R sin formar un grupo localizado conjuntamente. La elipse de color negro indica las tres muestras R-FI que se hayan en un grupo alejado del resto de R-FI. En el etiquetado de las muestras, el número que va detrás de “era” (eraX_) indica la tanda experimental. Donde x= [1,7].

6.7. Análisis de componentes principales (*Principal Component Analysis, PCA*)

6.7.1. Modelo de dataje endometrial

Tras realizar el PCA de las muestras control para el modelo de dataje endometrial (vid. MM 5.1) con la expresión de los genes del *ERA* como variables, se observó como las muestras se separaban en el espacio de forma clara en cuatro grupos según la fase del ciclo menstrual (PF, PR, R y PS) (figura 19). Este resultado implicó que el transcriptoma del *ERA* permitía distinguir las muestras según la fase del ciclo menstrual y que cada una de las clases, sobretudo la receptiva, era bastante homogénea y se distinguía del resto.

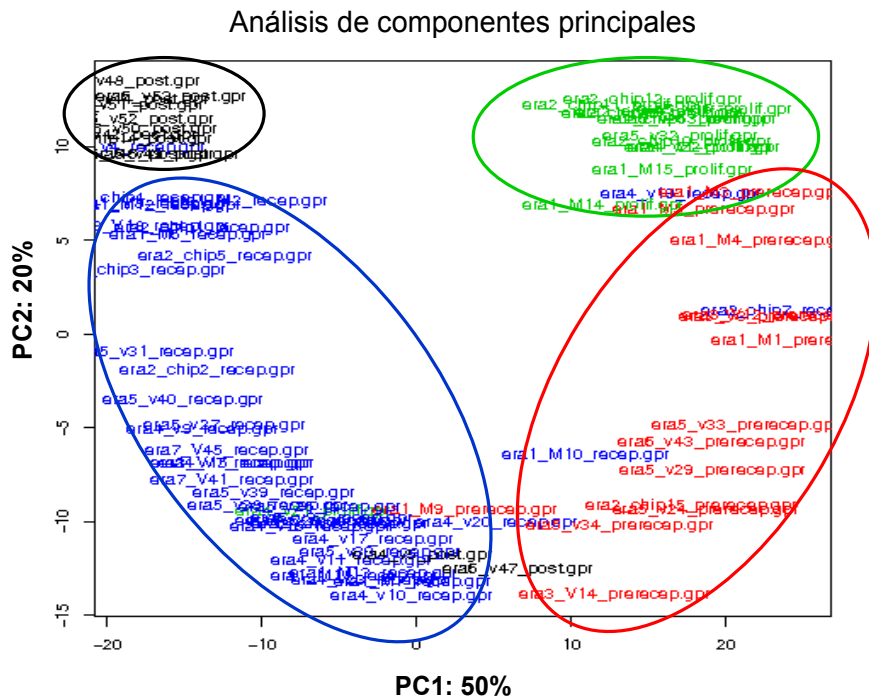


Figura 19. PCA de las muestras control del ciclo menstrual. Se muestra la

distribución en el espacio de la expresión del *ERA* en función de los dos componentes principales. El componente principal 1 (PC1 eje ordenadas) es sobre el que recae el 50% del peso de la separación entre las clases; y el componente principal 2 (PC2 eje abscisas), tiene el peso del 20%. El grupo de color negro representa las muestras “post-receptivas” mayoritariamente, el de color azul las muestras “receptivas”, el verde las muestras “proliferativas” y el rojo las “pre-receptivas”. También podemos observar como algunas muestras se sitúan fuera del grupo donde se encuentran mayoritariamente según la fase del ciclo.

6.7.2. Modelo de clasificación patológica

Se realizó un PCA con el transcriptoma *ERA* de las muestras R-FI junto con las muestras control PF (n=14), PR (n=14), R (n=39) y PS (n=12) (vid. MM 5.1), utilizadas anteriormente en el PCA. En la figura 20 podemos observar como se agrupan, sin distinguirse, todas las muestras receptivas juntas, las R-FI y las R. El resto de muestras (PF, PR y PS) se seguían agrupando diferencialmente en el espacio. Por lo que el PCA no distinguió en el espacio a las R de las R-FI. Por otra parte, parecía que se podían distinguir dos grupos de R-FI, uno más parecido a las R, y otro formado por tres muestras (era7_S2, era3_S5 y era7_S1) que se agrupaban juntas en el espacio más alejadas de las R y próximas a las PR (figura 20).

Se realizó otro PCA sólo con las muestras R y con las R-FI para asegurarnos de que las muestras PF, PR y PS no nos dificultaban la separación. Utilizando el componente principal primero (PC1) junto con el componente principal segundo (PC2), no sumaban el peso del 70%. El PC1 era del 44% corroborándose la dificultad de separar las muestras R de las R-FI por su similitud, y el agrupamiento en el espacio como en la figura 20. Concluyendo

que el PCA a partir del transcriptoma del *ERA* no distinguía a las R-FI de las muestras R.

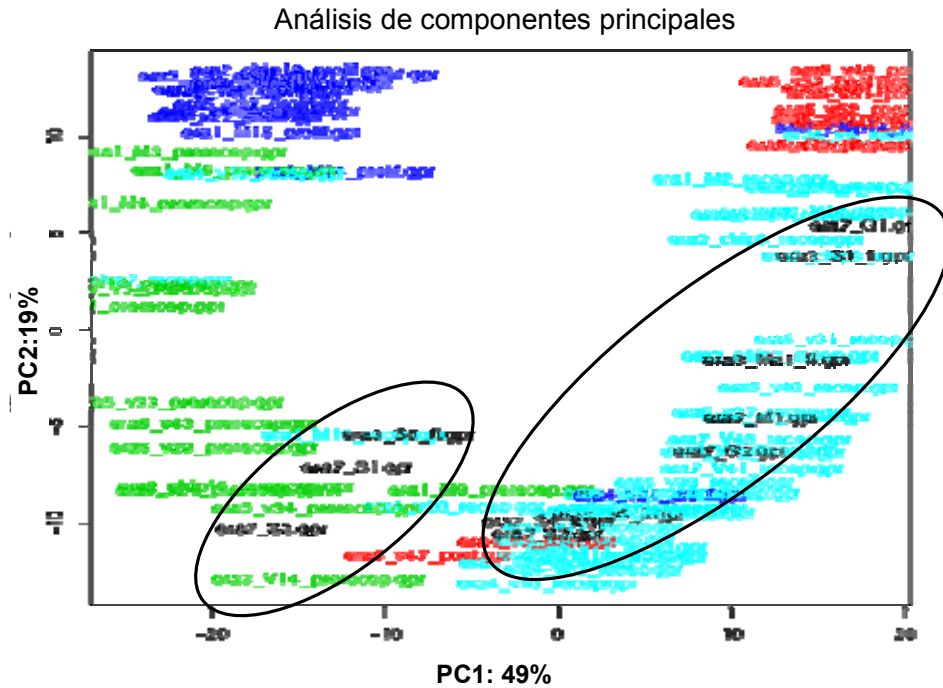


Figura 20. PCA de las muestras control PF, PR, R y PS junto con las R-FI. Se observa la distribución en el espacio de las muestras en función de los dos componentes principales. El componente principal 1 (PC1 eje ordenadas) es sobre el que recae el 49% del peso de la separación entre las clases y el componente principal 2 (PC2 eje abcisas) tiene el peso del 19%. Las muestras PF se representan en azul oscuro, las PR se muestran en verde, las R-FI en color negro, las R control en azul claro, y las PS en rojo. Las elipses negras indican los dos grupos de muestras R-FI encontrados.

Debido a la aparente heterogeneidad de las muestras R-FI por los resultados del análisis de los PCA (figura 20) reforzados por los análisis de agrupamiento jerárquico (figura 18), se realizó otro PCA sólo de las muestras R-FI que mostró la heterogeneidad dentro de ellas (figura 21).

El PC1 junto con el PC2 contenía el 73% de la variabilidad. El PC1 por sí solo ya explicaba el 61% de la variabilidad dentro de las muestras R-FI, esto ya indicaba la facilidad de separar las muestras debido a la heterogeneidad. Se observaron dos grandes grupos de muestras R-FI, uno compuesto por cinco muestras (era7_G1, era3_S1, era3_Ma1, era7_M1 y era7_G2) y otro por seis muestras (era7_S2, era3_S5, era7_S1, era3_S6, era3_S4 y era7_S3) (figura 21).

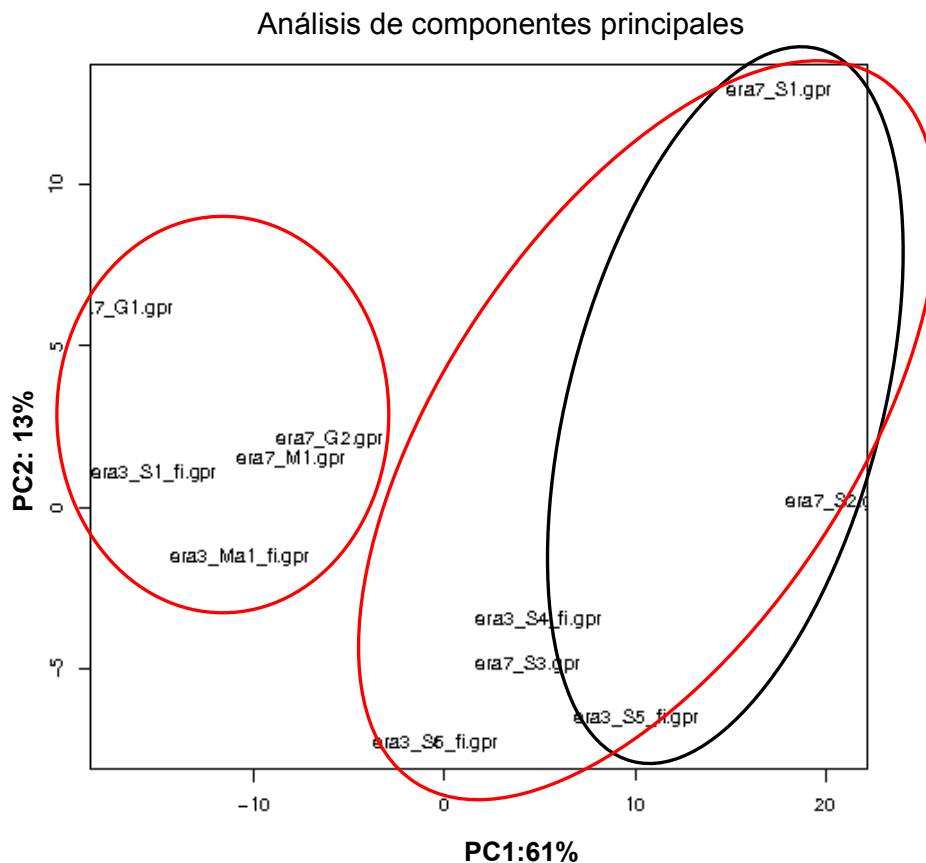


Figura 21. PCA de las muestras control R-FI. Se observa la distribución en el espacio de las muestras en función de los dos componentes principales. El componente principal 1 (PC1 eje ordenadas) es sobre el que recae el 61% del peso de la separación entre las muestras y el componente principal 2 (PC2 eje abscisas) tiene el peso del 13%.

Las muestras R-FI se representan escritas de color negro. La elipse negra indican las tres muestras R-FI (era7_S2, era3_S5 y era7_S1) que se agrupaban juntas en el espacio diferenciándose del resto de R-FI en el agrupamiento (figura 18) y en el PCA (figura 20). Las elipses rojas delimitan los dos grandes grupos de muestras R-FI que parece haber.

6.8. Expresión génica diferencial del grupo control R-FI (receptivas con fallo de implantación)

- **Expresión génica diferencial entre R vs. R-FI**

De los estudios de expresión génica diferencial entre las muestras R vs. R-FI, el menor p-valor registrado fue de 0,2, luego no se obtuvo ningún gen expresado diferencialmente de forma significativa. Este resultado corroboró que estos dos grupos eran muy parecidos como ya indicaba el agrupamiento (figura 18) y el PCA (figura 20).

- **Expresión génica diferencial dentro del grupo control R-FI**

Debido a la aparente heterogeneidad dentro del grupo R-FI por los resultados obtenidos en el agrupamiento de las muestras R-FI (figura 18) y por los PCA (figura 20 y figura 21), se decidió realizar un estudio de expresión génica diferencial entre las mismas muestras control R-FI. En primera instancia, se compararon las tres muestras R-FI que parecían formar un grupo aparte (era7_S2, era3_S5 y era7_S1) (figura 18 y figura 20) vs. el resto de muestras R-FI (n=8). En segunda instancia, debido a la distribución de las muestras R-FI en el PCA (figura 21), también se compararon entre sí los dos grandes grupos de R-FI (grupos definidos por las elipses rojas de la figura 21).

El resultado de la primera comparativa fueron 44 genes diferencialmente expresados, 16 genes regulados al alza y 28 a la baja (figura 22A). El *heatmap*, según estos genes, evidenció que estas tres muestras R-FI estaban muy diferenciadas del resto.

Se compararon los 44 genes diferencialmente expresados entre las tres muestras con respecto al resto de R-FI (figura 22A) con los 176 genes diferencialmente expresados de la comparación R vs. PF; los 200 genes de R vs. PR; y los 94 de R vs. PS (vid. punto 6.5). Se observó que 41 de los 44 genes estaban incluidos en los diferencialmente expresados obtenidos para la fase proliferativa, la totalidad de los 44 genes formaban parte de los genes diferencialmente expresados en fase pre-receptiva y 19 de los 44 genes estaban presentes en la comparativa de la fase post-receptiva. Estos resultados estuvieron en consonancia con que estas tres muestras estuvieran más próximas de las muestras pre-receptivas según indicaba el PCA de la figura 20.

El resultado del segundo análisis de expresión génica diferencial entre las muestras R-FI fue 85 genes expresados diferencialmente, 35 regulados al alza y 50 a la baja (figura 22B).

También se compararon estos 85 genes diferencialmente expresados entre los dos conjuntos de muestras R-FI (figura 22B) con los 176 genes diferencialmente expresados de la comparación R vs. PF; los 200 genes de R vs. PR; y los 94 de R vs. PS (vid. punto 6.5). Se observó que de la totalidad de los 85 genes, 66 formaban parte de los genes diferencialmente expresados en fase proliferativa, 81 de los 85 genes estaban incluidos en los diferencialmente expresados obtenidos para la fase pre-receptiva y 39 de los 85 genes estaban presentes en la comparativa de la fase post-receptiva.

Tras todos estos análisis de expresión génica diferencial se concluyó que el grupo control R-FI era un grupo heterogéneo como ya iban indicando los análisis exploratorios previos.

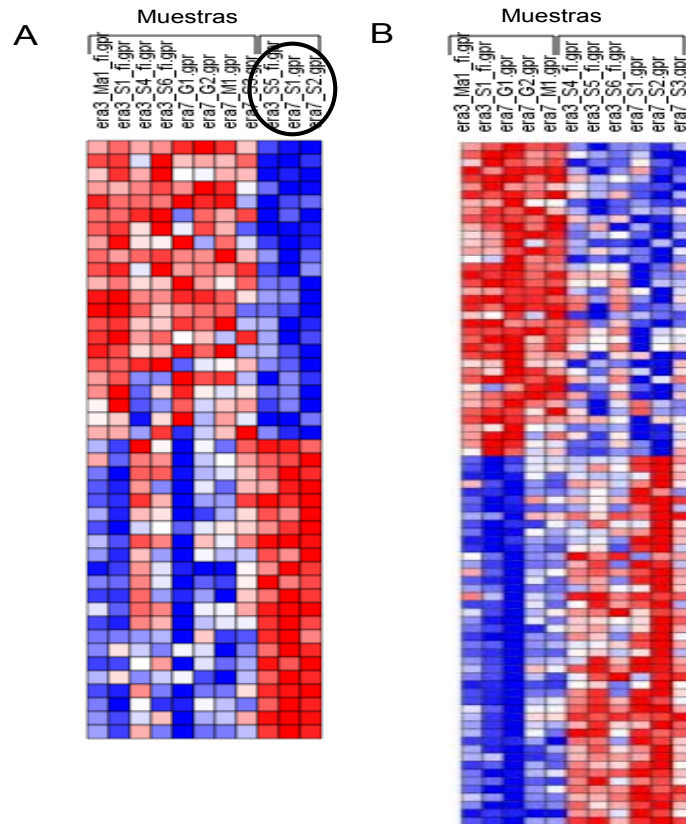


Figura 22. Heatmap de los genes diferencialmente expresados dentro de las muestras control R-FI. A) *Heatmap* del estudio de expresión génica diferencial entre las tres muestras R-FI (círculo negro) y el resto de muestras R-FI control. B) *Heatmap* del estudio de expresión génica diferencial entre los dos grandes grupos de R-FI. El eje de ordenadas corresponde a cada uno de los genes diferencialmente expresados ordenados de mayor a menor por el valor del estadístico. En el eje de abscisas se representan cada una de las muestras. Los cuadrados rojos indican la expresión mayor para el gen y los cuadrados azules la expresión menor.

6.9. Entrenamiento del predictor

6.9.1. Modelo de dataje endometrial

Los tres algoritmos empleados (SVM, KNN y RF) para clasificar el transcriptoma obtenido por el *ERA* en función del set de entrenamiento, es decir, PF, PR, R, y PS (vid. MM 5.1), mostraron un buen poder diagnóstico (tabla 7A y 7B). Sin embargo, el SVM, con la función linear Kernel, destacó tanto al utilizar la expresión de todos los genes del *ERA* (tabla 7A) como al utilizar sólo los 56 genes de la *ERS* (tabla 7B). Este algoritmo presentó los mejores valores en 3 de las 4 métricas usadas (ACC, RMSE y MCC) y la mejor sensibilidad (tabla 7A). El modelo utilizando todos los genes del *ERA* (tabla 7A) produjo unos resultados en la clasificación muy similares a los obtenidos con los 56 genes de la *ERS* (tabla 7B). Esto mostró el poder predictivo de la *ERS* y reforzó su representación fenotípica a la hora de datar las muestras como receptivas.

6.9.2. Modelo de clasificación patológica

En el modelo de predicción para clasificar muestras receptivas normales y muestras receptivas con fallo de implantación, se utilizó como set de entrenamiento las muestras R y las R-FI. El resultado de los tres algoritmos empleados (SVM, KNN y RF) se muestra en la tabla 7C y 7D. El SVM destacó frente al KNN y al RF que, a su vez, poseían valores muy parecidos. Aunque el KNN mostró menos clasificados incorrectos que el RF. El modelo utilizando todos los genes del *ERA* (tabla 7C) produjo unos resultados en la clasificación ligeramente menores que los obtenidos con los 56 genes de la *ERS* (tabla 7D). Esto mostró el poder predictivo de la *ERS* y reforzó su representación fenotípica a la hora de clasificar las muestras como receptivas normales.

A) Dataje endometrial con el ERA

Algoritmo	ACC	AUC	RMSE	MCC	C	I	E	S
SVM	0,915 +/-0,014	0,916 +/-0,014	0,180 +/-0,011	0,843 +/-0,024	73 (92,4%)	6 (7,6%)	0,83	1,00
KNN	0,868 +/-0,018	0,909 +/-0,010	0,245 +/-0,004	0,768 +/-0,022	69 (87,3%)	10 (12,7%)	0,90	0,86
RF	0,863 +/-0,011	0,924 +/-0,013	0,247 +/-0,004	0,733 +/-0,019	68 (86%)	11 (14%)	0,84	0,89

MATRIZ DE CONFUSIÓN

		PREDICCIÓN											
		R			PR			PF			PS		
		SVM	KNN	RF	SVM	KNN	RF	SVM	KNN	RF	SVM	KNN	RF
CLASE REAL	R	39	34	35	0	2	2	0	1	1	0	2	1
	PR	1	1	1	13	12	13	0	1	0	0	0	0
	PF	2	1	2	0	0	0	12	13	12	0	0	0
	PS	3	2	4	0	0	0	0	0	0	9	10	8

B) Dataje endometrial con la ERS

Algoritmo	ACC	AUC	RMSE	MCC	C	I	E	S
SVM	0,894 +/- 0,010	0,895 +/-0,010	0,210 +/-0,011	0,806 +/-0,017	71 (89,9%)	8 (10,1%)	0,79	1,00
KNN	0,889 +/-0,009	0,905 +/-0,015	0,232 +/-0,005	0,785 +/-0,013	71 (89,9%)	8 (10,1%)	0,90	0,88
RF	0,868 +/-0,019	0,916 +/-0,012	0,252 +/-0,008	0,741 +/-0,035	69 (87,3%)	10 (12,7%)	0,85	0,89

MATRIZ DE CONFUSIÓN

		PREDICCIÓN											
		R			PR			PF			PS		
		SVM	KNN	RF	SVM	KNN	RF	SVM	KNN	RF	SVM	KNN	RF
CASE REAL	R	39	35	35	0	2	2	0	1	1	0	1	1
	PR	3	1	1	11	13	13	0	0	0	0	0	0
	PF	2	1	1	0	0	0	12	13	13	0	0	0
	PS	3	2	4	0	0	0	0	0	0	9	10	8

C) Clasificación patológica con el ERA

Algoritmo	ACC	AUC	RMSE	MCC	C	I	E	S
SVM	0,908 +/-0,028	0,825 +/-0,050	0,241 +/-0,063	0,720 +/-0,091	45 (90%)	5 (10%)	0,65	0,98
KNN	0,784 +/-0,020	0,703 +/-0,044	0,392 +/-0,021	0,144 +/- 0,032	39 (78%)	11 (22%)	0,06	0,99
RF	0,754 +/-0,016	0,715 +/-0,043	0,399 +/-0,009	0,132 +/-0,109	38 (76%)	12 (24%)	0,06	0,95

MATRIZ DE CONFUSIÓN

CLASE REAL	PREDICCIÓN						
		R			R-FI		
		SVM	KNN	RF	SVM	KNN	RF
R	38	38	37	1	1	2	
R-FI	4	10	10	7	1	1	

D) Clasificación patológica con la ERS

Algoritmo	ACC	AUC	RMSE	MCC	C	I	E	S
SVM	0,942 +/-0,035	0,892 +/-0,061	0,144 +/-0,072	0,830 +/-0,103	47 (94%)	3 (6%)	0,77	0,99
KNN	0,772 +/-0,016	0,662 +/-0,038	0,416 +/-0,015	0,120 +/-0,094	38 (76%)	12 (24%)	0,11	0,96
RF	0,754 +/-0,016	0,623 +/-0,053	0,413 +/-0,009	0,039 +/-0,082	37 (74%)	13 (26%)	0,03	0,96

MATRIZ DE CONFUSIÓN

CLASE REAL	PREDICCIÓN						
		R			R-FI		
		SVM	KNN	RF	SVM	KNN	RF
R	39	37	37	0	2	2	
R-FI	3	10	11	8	1	0	

Tabla 7. Parámetros de la predicción. En las siguientes tablas se representan los tres algoritmos distintos con los que se han generado los modelos (SVM, KNN y RF) en las filas, y en las columnas, los distintos parámetros para medir la eficacia de la predicción con los que se ha valorado el error de cada modelo (precisión (ACC); área bajo la curva ROC (AUC); error cuadrático medio (RMSE); coeficiente de correlación de

Matthews (MCC); Correctos (C); Incorrectos (I); Especificidad (E) y Sensibilidad (S)). La tabla 7A muestra los resultados del entrenamiento del predictor para el modelo de dataje endometrial con los 238 genes del *ERA*. La tabla 7B contiene el entrenamiento del predictor para el dataje endometrial con los 56 genes de la *ERS*. La tabla 7C el entrenamiento del predictor para el modelo de clasificación patológica con los 238 genes del *ERA*. La tabla 7D presenta los parámetros resultantes del entrenamiento del predictor para el modelo de clasificación patológica con los 56 genes de la *ERS*.

6.10. Herramienta molecular diagnóstica de la receptividad endometrial

La herramienta molecular diagnóstica de la receptividad endometrial quedó constituida por el *microarray* personalizado *ERA* y por un predictor compuesto a su vez por dos modelos.

Con los resultados de poder predictivo obtenidos en los modelos de predicción del punto 6.9 se estableció como la herramienta molecular realizaría el diagnóstico.

El modelo de dataje endometrial quedó constituido por el algoritmo SVM con todos los genes del *ERA*. El método SVM, con la función linear kernel, se eligió por presentar los mejores valores en 3 de las 4 métricas usadas (ACC, RMSE y MCC) y porque además presentaba la mejor sensibilidad (tabla 7A).

Para reforzar la detección de las muestras FI, el modelo de clasificación patológica quedó constituido por el consenso de los dos algoritmos SVM y KNN con todos los genes del *ERA*. El método SVM, con la función linear kernel, se escogió por poseer el mejor valor en las cuatro métricas (ACC, AUC, MCC y RMSE) (tabla 7C). La sensibilidad y especificidad era muy parecida en los tres

modelos. El KNN fue el segundo método elegido por poseer unos resultados en la clasificación ligeramente mayores que el RF. En caso de discrepancia entre los algoritmos se optaba por el diagnóstico menos optimista.

Con toda esta información se generó un programa informático predictivo que funciona de la siguiente manera:

Una biopsia endometrial tomada en LH+7 se analiza con el *ERA*. La matriz de datos obtenida se implementa en el programa informático, donde se preprocesa y normaliza según el set de entrenamiento. El modelo de dataje endometrial data la muestra como receptiva o no receptiva según el modelo SVM. Sólo si en el primer modelo es diagnosticada como receptiva pasa a ser analizada por el segundo modelo. El segundo modelo (modelo de clasificación patológica) clasifica la muestra como receptiva o FI en función del consenso entre los algoritmos SVM y KNN. Si hay discrepancia entre ambos algoritmos se clasifica la muestra en función del diagnóstico menos optimista (FI).

La herramienta genera un diagnóstico con una probabilidad asociada. Cada muestra testada por el predictor se predice con un valor de confianza para la predicción en cada algoritmo. De modo que esta probabilidad asociada es el resultado de multiplicar el ACC del algoritmo por el valor de confianza diagnóstica para esa muestra.

Si una muestra es diagnosticada como no receptiva, ya no pasa al segundo modelo y queda con la probabilidad asociada del primer modelo. Sin embargo, las probabilidades de los dos modelos son independientes, y una vez que una muestra pasa al segundo modelo por ser receptiva, el modelo de clasificación patológica obtendrá la clasificación de la muestra para cada algoritmo (SVM y KNN) con una probabilidad asociada propia. Si los dos

algoritmos están de acuerdo se escoge la mejor probabilidad de las dos. Mientras que si están en desacuerdo se escoge el diagnóstico FI con los valores ACC y de confianza del algoritmo que lo haya diagnosticado (SVM o KNN).

En el anexo III se muestra con más detalle como funciona el programa informático de la herramienta molecular.

6.11. Diagnóstico de la herramienta molecular en la población de pacientes

Con la herramienta molecular se analizaron 40 biopsias endometriales en LH+7 de la población de pacientes de la clínica (vid. MM 5.1). Con el modelo de dataje endometrial las 40 muestras fueron diagnosticadas como receptivas. Posteriormente, con el modelo de clasificación patológica 38 se clasificaron como receptivas normales con una probabilidad de diagnóstico de 0,944 y 2 como receptivas con FI con una probabilidad de diagnóstico de 0,784.

6.12. Comparativa dataje histológico vs. dataje predictivo

Con el objetivo de comparar el método de dataje de la predicción de la herramienta molecular diagnóstica de la receptividad endometrial, con el clásico método de Noyes (Noyes y cols., 1950 y 1975), se realizaron ambos tipos del dataje en un subconjunto de las muestras control del ciclo menstrual (n=49) (vid. MM 5.1) formado por: PF n=6, PR n=8, R n=24, y PS n=11. Se eligió el algoritmo SVM para el dataje endometrial por ser el modelo que aplica la herramienta molecular diagnóstica de receptividad endometrial por mostrar los mejores resultados (tabla 7A).

La concordancia de ambos métodos, en relación con el pico de LH y entre patólogos, fue analizada estadísticamente por el análisis del índice de concordancia Kappa ponderado por pesos cuadráticos (tabla 8), mostrando que la comparación entre el pico de LH y los dos patólogos era de 0,618 (0,446-0,791) y 0,685 (0,545-0,824) respectivamente. Mientras que con el predictor la concordancia resultaba de 0,922 (0,815-1,000).

Métodos dataje	Índice Kappa	Intervalo de confianza 95%
LH vs P1	0,618	0,446 a 0,791
LH vs P2	0,685	0,545 a 0,824
LH vs PRE	0,922	0,815 a 1,020
P1 vs P2	0,622	0,435 a 0,809

Tabla 8. Comparativa de los métodos de dataje endometrial. Se muestra el índice de concordancia Kappa junto con el intervalo de confianza al 95% de los distintos métodos para datar el endometrio. LH es el método de detección del día de ovulación; PRE es el resultado de la predicción con el modelo SVM de dataje endometrial; P1 y P2 son cada uno de los distintos patólogos que basándose en los criterios de Noyes (Noyes y cols., 1950 y 1975) han datado el endometrio.

El estudio del porcentaje de aciertos y fallos de cada método de dataje, analizado de forma aislada en cada fase del ciclo menstrual, se muestra en la figura 23.

El predictor clasificaba mejor las muestras en fase pre-receptiva y receptiva que las muestras en fase proliferativa y post-receptiva (figura 23). El método de dataje histológico clasificaba mejor las muestras en fase proliferativa, seguidas de las muestras en fase pre-receptiva. Además, no había

discrepancia en el porcentaje de aciertos y fallos entre patólogos dentro de estas dos fases. Sin embargo, el dataje histológico clasificaba peor las muestras en fase receptiva seguidas de las muestras en fase post-receptiva que eran las fases donde peor databa y donde había más discrepancia entre los distintos patólogos.

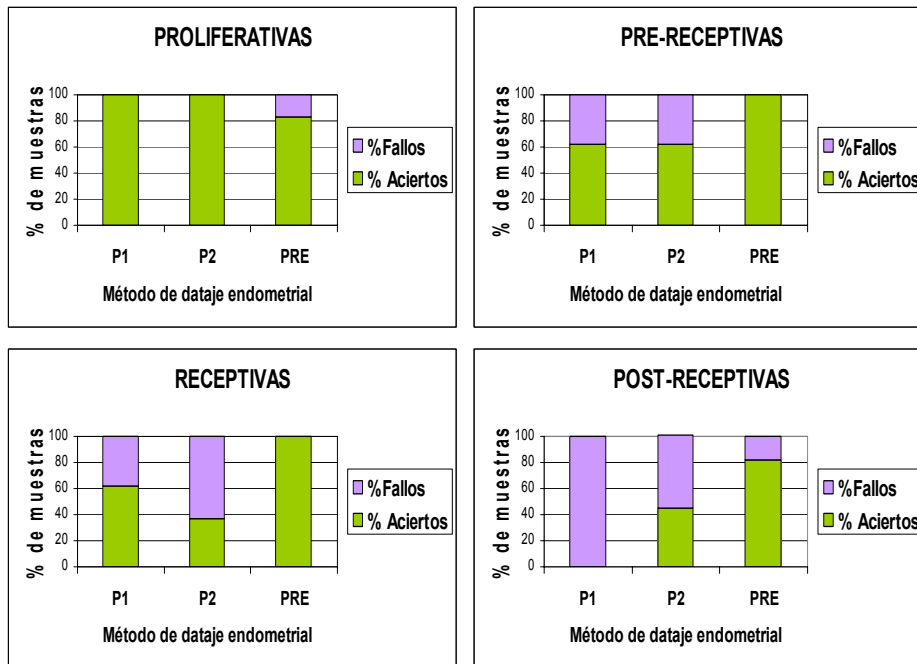


Figura 23. Dataje endometrial comparado para las distintas fases del ciclo menstrual. Se muestran los distintos métodos de dataje (P1: patólogo 1; P2: patólogo 2; PRE: predictor) (eje de abcisas) de forma independiente para cada fase del ciclo menstrual. El eje de ordenadas representa el porcentaje de aciertos (color verde) y fallos (color malva) de cada método con respecto al conjunto de muestras de una misma fase del ciclo menstrual.



“Porque soy fuerte puedo mostrar mis debilidades”

Anónimo

Discusión

7. DISCUSIÓN

- **Enfermedad y diagnóstico clínico**

La definición de enfermedad, así como su diagnóstico y tratamiento, ha cambiado paulatinamente a lo largo de los dos últimos siglos. Históricamente los médicos definían las enfermedades como un conjunto de síntomas a los que se fueron añadiendo criterios anatómicos, bioquímicos y, últimamente, genéticos, que modificaron la taxonomía médica. Por ejemplo, antes del siglo XX la esquizofrenia y la demencia sifilítica eran consideradas la misma enfermedad. En el comienzo de 1900 se separó la psicosis de la demencia, y desde 1960 la esquizofrenia está evolucionando desde una enfermedad psiquiátrica a una enfermedad con una posible etiología genética. Esta evolución ha sido posible gracias a disponer de herramientas diagnósticas mejores y más objetivas.

Con el avance de la ciencia y el desarrollo de la Biología Molecular y la Genética, nuevas tecnologías se han puesto al servicio de la Biomedicina. La tecnología de *microarrays* se ha aplicado en muchos ámbitos de la salud humana. Según se detalla en el informe de vigilancia tecnológica de Genoma España (Aplicaciones de los *Microarrays* y Biochips en Salud Humana), las aplicaciones de los *microarrays* en salud humana se han dividido en cinco sectores que se corresponden con la identificación de dianas terapéuticas, el descubrimiento y desarrollo de fármacos, el diagnóstico clínico, la farmacogenómica y farmacogenética, y la investigación básica.

Los *microarrays* de ADN han sido los más utilizados dentro del sector del diagnóstico clínico. Y una de sus aplicaciones dentro de este sector ha sido la identificación de patrones de expresión relacionados con diferentes estados

patológicos. El uso del perfil de expresión génica como biomarcador se ha desarrollado principalmente en el análisis de expresión génica en procesos tumorales. Para ciertos tipos de cáncer se ha elaborado una clasificación de los tumores, lo cual ha permitido un diagnóstico y pronóstico más preciso que el obtenido mediante otras técnicas moleculares (URL10). En el 2001 el centro nacional de investigaciones oncológicas (CNIO) desarrolló el OncoChip (URL11). Este biochip constituyó una plataforma de expresión génica múltiple específicamente dedicada a cáncer. El OncoChip estaba constituido por 9.300 genes relacionados con las formas más frecuentes de cáncer, como el cáncer de mama, el colorrectal, el de pulmón, linfomas, leucemia y tumores urinarios como el de riñón, vejiga o próstata.

El grupo de investigación de van 't Veer buscó en el perfil de expresión génica diagnosticar el cáncer de mama (van 't Veer y cols., 2002; van de Vijver y cols., 2002) hasta obtener la firma transcriptómica de 70 genes como diagnóstico de este tipo de cáncer, generándose así la herramienta molecular llamada *Mammaprint*. Esta herramienta también se diseñó con un *microarray* personalizado *Agilent* de formato 8x15k, el cual fue el mismo formato empleado en el *ERA* (Buyse y cols., 2006; Glas y cols., 2006). En esta misma línea el trabajo publicado por Tan y cols., en el 2008, supuso definir la firma transcriptómica para tres subtipos moleculares de cáncer de mama a partir del diseño de un *microarray* específico con 188 genes (Tan y cols., 2008). Todos estos trabajos mencionados muestran la aplicación diagnóstica del perfil de expresión génica para caracterizar fenotipos multifactoriales.

En el ámbito de la Ginecología el avance del conocimiento sobre el endometrio humano ha estado limitado durante mucho tiempo por la ausencia de progreso en el desarrollo de nuevos métodos de investigación y estudio. Un ejemplo claro es el método histológico basado en los criterios de Noyes (Noyes

y cols., 1950 y 1975), cuyos criterios morfológicos fueron fijados en 1950 y continúan aplicándose actualmente para el estudio del endometrio, evaluación de la receptividad endometrial y detección de anomalías endometriales.

De la misma manera, en medicina reproductiva se ha buscado aplicar estas nuevas tecnologías para identificar objetivamente y diagnosticar el estado de la receptividad endometrial durante la WOI. En los últimos años, diferentes trabajos han sido publicados intentando encontrar el conjunto de genes biomarcadores de receptividad endometrial (Haouzi y cols., 2009; Tseng y cols., 2009). Así como también se ha ido a la búsqueda de la firma transcriptómica de las alteraciones endometriales (Tseng y cols., 2009). El enfoque de estas investigaciones muestra que un análisis “ómico” está siendo una vía adecuada para el estudio del endometrio.

En esta vertiente “ómica”, la presente tesis doctoral se ha centrado en generar una herramienta molecular diagnóstica de aplicación clínica basada en el transcriptoma de la receptividad endometrial.

- **Muestras**

La selección de muestras para cada uno de los objetivos del estudio se hizo de la manera más rigurosa posible. La toma de biopsias se realizó midiendo el pico de LH como manera más fiable de saber en qué fase se encontraba el endometrio. Las muestras control poseían fertilidad probada, ciclos menstruales regulares, analíticas generales y cariotipo normal. No obstante, el pico de LH no se produce de la misma manera en todas las mujeres (Park y cols., 2007) y el ciclo menstrual de una misma mujer puede variar de un ciclo a otro, existiendo multitud de variables que pueden afectar. La edad, siempre y cuando no se haya producido la menopausia o no

sobrepase los 45 años, no se considera un factor que afecte al endometrio (Soares y cols., 2005). En cuanto al IMC, se observó que valores por encima de 25 Kg/m² podían afectar al endometrio (Bellver y cols., 2007). Es por ello, que se incluyeron muestras de mujeres que no superaran este IMC.

Sin embargo, aún seleccionando las muestras con los criterios de inclusión más rigurosos, podían existir otros parámetros que pudieran afectar al estudio, por lo que se comprobó que en las muestras seleccionadas no hubiera co-variables que alteraran los resultados. La existencia de estos posibles parámetros se analizó mediante los métodos exploratorios del análisis de agrupamiento jerárquico (figuras 17 y 18) y mediante el PCA (figuras 19 y 20).

Por otra parte, las muestras a analizar por la tecnología de microarrays pasaron distintos puntos control. De modo que aquellas muestras que no tuvieron un RIN>7 en la calidad de su ARN, o que no obtuvieron una actividad>9 a la hora de incorporar el fluoróforo, no fueron incluidas en el estudio. Con estos requisitos se garantizó que la medida de la expresión de un gen se correspondiera a la situación biológica estudiada.

- **Selección de genes**

El criterio establecido para la selección de genes aunó requerimientos estadísticos y biológicos, encontrando el sentido funcional de esta selección génica en el proceso biológico de la receptividad endometrial (tabla 3). Los valores escogidos de FDR y FC fueron valores estadísticamente adecuados que incluían un número manejable de 238 genes (tabla 2). Otros puntos de corte se podían haber realizado, pero se observó que excluían genes que se consideraban importantes en el proceso, como es el caso de la glicodelina (PAEP), una de las proteínas cuya población de ARNm aparece expresada al

alza con una de las mayores tasas de cambio en receptividad (Horcajadas y cols., 2004 y 2007) (tabla 2). También se podían haber seleccionado los genes relativos a la receptividad endometrial a partir de los trabajos publicados sobre transcriptómica del endometrio (vid. 1.3.3), pero debido a la falta de consenso entre los distintos estudios (Horcajadas y cols., 2007) se optó por obtenerla a partir de experimentos propios y del análisis de expresión génica diferencial.

En muchas investigaciones, incluida esta, se utilizó el método del doble filtro empleando el FC y el *T-Test* para la selección de genes diferencialmente expresados (Zhang y cols., 2009). Pero además, en este trabajo también se realizó un análisis funcional para conocer la relevancia biológica en el proceso de la receptividad endometrial de los genes escogidos, completándose así el sentido biológico de la selección.

Los genes que conformaron el *ERA* se compararon con la lista de los 25 genes diana propuestos por Horcajadas y cols. (Horcajadas y cols., 2007). Estos genes fueron seleccionados a partir de la comparativa de la transcriptómica de la WOI en ciclos naturales, en ciclos estimulados y en presencia de DIU. Se encontró que en la selección del *ERA*, a pesar de haber sido obtenida con un método independiente a los genes seleccionados por Horcajadas y cols., 18 de los 25 genes estaban presentes. Estos genes consenso se muestran en negrita en la tabla 2.

Por otra parte, también se comparó la selección del *ERA* con los 126 genes propuestos por Tseng como firma transcriptómica para el diagnóstico de alteraciones endometriales (Tseng y cols., 2009). Se encontró que 61 genes, de estos 126, estaban incluidos en la selección. Así mismo, la lista del *ERA* también incluía dos genes de los cinco recientemente publicados por Haouzi como potencialmente nuevos biomarcadores de la receptividad endometrial,

como son la laminina $\beta 3$ y la microfibrilina asociada a la proteína 5 (Haouzi y cols., 2009).

El análisis funcional por GO del *ERA* reveló el sentido biológico de la selección (tabla 3). La anotación más representada fue la presencia de BP relativos a la respuesta inmune, los cuales están implicados en la implantación embrionaria, como fue reportado previamente (Hannan y Salamonsen, 2007; Salamonsen y cols., 2007). Otro término destacado fue la adhesión celular tan necesaria para establecer los requerimientos físicos para la interacción epitelio-trofoblasto (van Mourik y cols., 2009), así como la mitosis, la señalización célula-célula, y las estructuras anatómicas del desarrollo. En relación a las MF, hubo tres términos diferentes en el nivel 4 que aparecieron más representados de lo que cabría esperar: la actividad oxidoreductasa, la unión a carbohidratos y los receptores de unión. Estas MF son funciones moleculares relevantes en actividades relacionadas con la adhesión celular, la transducción de señal y otras funciones cruciales en la implantación embrionaria.

Finalmente, se detectó que el huso mitótico y los microtúbulos del huso mitótico aparecían representados por encima de lo habitual dentro de la sección de CC, en los niveles específicos CC10 y CC11. Estos términos se relacionan principalmente con el ciclo mitótico y la proliferación celular que son procesos relacionados con el desarrollo endometrial.

Todos estos argumentos en relación al criterio de selección, y en relación a las distintas comparativas de los genes que incluye el *ERA* con otros genes seleccionados en otros trabajos, así como los argumentos sobre la relevancia funcional en el proceso de la receptividad endometrial, refuerzan el método mediante el que se realizó la selección de genes y los genes que incluye.

- **Diseño del ERA**

El diseño del ERA fue realizado con el soporte de la tecnología de *Agilent*, utilizando sondas diseñadas y testadas por ellos, así como sus puntos control. Además, distintas sondas representaban un mismo gen y cada una de las sondas que representaba a los 238 genes poseía ocho réplicas. Por todo lo comentado, se consideró que el ERA fue un *microarray* con un diseño adecuado y fiable.

- **Validación de la tecnología de *microarrays***

Por otra parte, la validación del correcto funcionamiento de la técnica de *microarrays* mediante el análisis de la distribución de intensidades ya fue un punto control en el proceso (figura 12). No obstante, se realizó una validación cuantitativa mediante la técnica RT-PCR, comprobando que los genes validados aumentaban su expresión con esta técnica al igual que con el *microarray ERA* (figura 13). Debido al distinto funcionamiento y sensibilidad entre la tecnología de *microarrays* y la RT-PCR los valores de expresión no se compararon de una manera cuantitativa. El objetivo de la validación fue comprobar que con ambas técnicas estos genes se expresaban más en receptividad.

- **Estudios de comportamiento de la expresión de los genes ERA**

Los resultados de los estudios del comportamiento de expresión de los genes a lo largo del ciclo menstrual (figura 14) se obtuvieron con un p -valor $< 0,2$. Esta falta de significatividad menor que 0,05 pudo ser debida al escaso número de genes por grupo. No obstante, estos análisis aportaron otra perspectiva y permitieron obtener grupos de comportamiento descriptivos que

acompañaron al resto de resultados. Los estudios revelaron un grupo de 34 genes, que tenían su máxima expresión en receptividad (grupo 7), y otro de 47 genes que tenían su mínima expresión en receptividad (grupo 3) (figura 14). Con este análisis también se confirmó que existían grupos de genes incrementados (grupo 7) y disminuidos (grupo 3) en receptividad, los cuales podían ser usados como biomarcadores de receptividad endometrial y se justificaba su presencia en el *ERA*.

Por otra parte, el análisis funcional de los dos grupos de genes que se comportan según el perfil 7 y 3 de la figura 14 (tabla 4 y 5 respectivamente) apoyó los BP significativos al realizar el estudio funcional del *ERA* (tabla 3), así como las MF de ligando de unión y los CC relacionados con los microtúbulos y el huso mitótico. También se pudo encontrar una relación entre el análisis funcional de los grupos de genes 7 y 3 con lo descrito por CR Murphy (Murphy, 2004) en la transformación de la membrana plasmática (figura 3).

En el análisis funcional del grupo 7, los genes que tuvieron su máxima expresión en receptividad estuvieron relacionados con MF de unión a iones cadmio y cobre, y con receptores de unión; y el CC fue el espacio extracelular (tabla 4). En el análisis funcional del grupo 3 (tabla 5), los genes que disminuían su expresión en receptividad fueron genes relacionados con el citoesqueleto. Según CR Murphy en fase receptiva, aumentan las uniones estrechas sellándose las células, y la membrana lateral y basal se vuelve más sinuosa. Esto podría estar relacionado con que haya un grupo de genes relacionados con MF de receptores de unión y con CC en el espacio extracelular que aumente su expresión (grupo 7). Además, el cinturón de actina se desorganiza, las microvellosidades desaparecen, los desmosomas disminuyen y no existen hemidesmosomas, con ello parece plausible que un

conjunto de genes relacionados con el citoesqueleto tenga su mínima expresión en este momento del ciclo (grupo 3).

- **ERS**

El modo con el que se definió la *ERS* permitió que se aislaran los genes exclusivamente expresados diferencialmente durante la WOI de una manera muy consistente. El fenotipo transcriptómico de receptividad (*ERS*) se definió de una manera muy rigurosa porque, por una parte, se obtuvo a partir de la expresión de 238 genes seleccionados ya con anterioridad por estar diferencialmente expresados en fase pre-receptiva vs. fase receptiva. Y por otra parte, por ser obtenida tras utilizar la triple comparación R vs. P; R vs. PR y R vs. PS de estos 238 genes (figura 15).

También se relacionaron los estudios del comportamiento de expresión de los genes a lo largo del ciclo menstrual (figura 14), con los genes de la *ERS* (tabla 2). Los estudios de comportamiento revelaron un grupo de 34 genes que tenían su máxima expresión en receptividad (grupo 7) y 47 que tenían su mínima expresión en receptividad (grupo 3) (figura 14). De estos 81 genes (genes grupo 3 + genes grupo 7), 25 eran compartidos con la *ERS*. El hecho de que sólo fueran 25, pudo ser debido a que parte de esos 81 genes tuvieran también una expresión similar en fase post-receptiva con respecto a la expresión en fase receptiva, y por ello no fueron seleccionados en la *ERS*.

- **Análisis exploratorios**

Tanto en el análisis de agrupamiento como en el PCA, no se observó que hubiera presentes variables externas o internas que pudieran estar afectando en los datos y que fueran responsables del agrupamiento de las muestras.

Así mismo, estos métodos exploratorios también sirvieron como análisis de la capacidad del transcriptoma seleccionado (genes) para separar las muestras. El agrupamiento jerárquico del grupo control sano de ciclo natural, reveló que las muestras se agrupaban en función de la fase del ciclo (figura 17), al igual que con el PCA (figura 19). Esto estuvo en acuerdo con los trabajos previos de Ponnampalam y Talbi (Ponnampalam y cols., 2004; Talbi y cols., 2006) en los que las biopsias endometriales se separaban según su transcriptoma por PCA según la fase del ciclo menstrual, y a su vez nos confirmó la capacidad del transcriptoma *ERA* para separar las muestras control normales. Por otra parte, se vio que el transcriptoma seleccionado (genes *ERA*) no tenía capacidad para separar a las muestras R-FI de las muestras R (figura 18 y 20), sin embargo, si que tenía capacidad para separar las muestras R-FI entre ellas (figura 21).

- **Análisis de expresión génica diferencial**

Los análisis de expresión génica diferencial permitieron corroborar las diferencias génicas entre los distintos tipos de muestras control según parecían indicar los análisis exploratorios. Se corroboró que las muestras PF, PR y PS se diferenciaban de las R, y que las PR eran las más diferentes a las R, por tener más genes diferencialmente expresados de los 238 genes del *ERA* (vid. punto 6.5). Del mismo modo, también se comprobó que las muestras R-FI no tenían genes del *ERA* diferencialmente expresados con respecto a las R (vid. punto 6.8) y que eran heterogéneas entre ellas (figura 22). Si se tiene en cuenta que los genes del *ERA* fueron seleccionados a partir de la comparación entre muestras pre-receptivas vs. receptivas control normales, el que estos genes del *ERA* distinguieran mejor las muestras PR de las R era lo que cabía esperar. En cuanto a la dificultad en separar a las R-FI de las R se podía

justificar con que en la selección de genes no se incluyeron los genes diferencialmente expresados entre muestras receptivas vs. receptivas con FI.

No obstante, el seleccionar los genes importantes en la receptividad endometrial hizo que entre ellos pudiera haber genes responsables del FI que nos permitieran separar estas muestras R-FI. De hecho, el *ERA* contenía 61 de los 126 genes propuestos por Tseng como genes relacionados con alteraciones endometriales (Tseng y cols, 2009). Unido a esto, hay que recalcar el que el perfil del *ERA* si que fuera capaz de distinguir grupos de muestras dentro de las R-FI. Todo esto hizo pensar que la dificultad para clasificar las muestras R-FI pudiera deberse, más que a que los genes del *ERA* no fueran capaces de distinguir las muestras R-FI, a la heterogeneidad dentro de este grupo (figura 21 y figura 22).

En la actualidad, se está viendo que más que buscar los genes relativos a un proceso de forma lineal, es decir, centrarse sólo en los genes diferencialmente expresados, hay que buscar relaciones funcionales entre genes y encontrar las vías de señalización implicadas (Wang y cols., 2010), ya que diferentes firmas transcriptómicas pueden responder a genes que afectan a los mismas vías en distintos puntos y que causan el mismo fenotipo clínico (Shi y cols., 2010b; Vidal y cols., 2011). Debido a esto se pensó que la heterogeneidad de las muestras R-FI pudo ser debida al origen multifactorial del FI, habiendo distintos perfiles de expresión que se correspondían con el fenotipo de esta patología, bien por estar afectando a la misma vía en distintos genes o bien por haber distintas causas que la originen. Además, el hecho de que las tres muestras R-FI estuvieran más próximas a las pre-receptivas en el PCA (figura 20) (vid. punto 6.8), unido a que los 44 genes diferencialmente expresados fueran más coincidentes con la fase proliferativa y sobretudo con la pre-receptiva, nos podría hacer pensar que estas muestras R-FI estaban

relacionadas con una falta de “madurez” en la receptividad endometrial. Sin embargo, el hecho de que el mayor número de genes diferencialmente expresados lo posea la comparación R vs. PR (200 genes), seguido por R vs. PF (176 genes) y por R vs. PS (94 genes), hace que esto pueda ser debido a la mayor probabilidad por la cantidad de genes de partida.

Por otra parte, los estudios de expresión génica diferencial del *ERA* entre las distintas fases del ciclo menstrual, tanto al realizar un *T-Test* comparando todos los grupos con la expresión en receptivo como al realizar el ANOVA, mostraron que prácticamente todos los genes seleccionados para el *ERA* con la plataforma de Affymetrix estaban diferencialmente expresados con la plataforma de *Agilent*. Estos resultados están en acuerdo con lo publicado en el consorcio MAQCI (*Microarray Quality Control*) (Shi y cols., 2006, 2008a y 2008b) donde se establecía la consistencia de estos estudios de expresión génica diferencial entre distintas plataformas, así como el elevado nivel de concordancia dentro de una misma plataforma en los genes identificados como diferencialmente expresados.

Sin embargo, las pequeñas diferencias que se produjeron, puesto que había un grupo reducido de genes que no mostraban expresión génica diferencial en la plataforma de *Agilent* (figura 16 y tabla 6), pueden quedar justificadas con que la cohorte de muestras empleadas en cada plataforma fue distinta.

- **Predicción**

Los algoritmos de clasificación empleados para definir los modelos de predicción fueron el SVM, el KNN, y el RF. Estos algoritmos fueron los elegidos por ser aquellos que mejor funcionaban con las matrices de datos obtenidas a

partir de los experimentos de *microarrays* según el consorcio MAQC [MAQCII, (Shi y cols., 2010a)].

La predicción computacional se realizó a partir de los 238 genes del *ERA*, y a partir de los 56 genes obtenidos para la *ERS*. Los modelos de predicción generados con la *ERS* obtenían una clasificación muy parecida a los mismos modelos entrenados con las mismas muestras pero para la totalidad de los genes del *ERA* (tabla 7). En el caso del dataje endometrial, el modelo que mejor predecía fue el SVM y de manera muy similar con los genes *ERA* que con la *ERS*. Sin embargo, en el modelo de clasificación patológica se mejoraba el poder diagnóstico del modelo SVM con los 56 genes de la *ERS*. La explicación de por qué estos 56 genes de la *ERS* podían llegar a clasificar mejor las anomalías de receptividad endometrial que el *ERA*, se cree que recae más en el algoritmo matemático que en una explicación biológica.

El algoritmo SVM utiliza la información de las variables de una manera selectiva, con lo que introducir esos genes específicos de la *ERS* nos elimina variables que, al no ser tan específicas, pueden afectar en la clasificación. A diferencia del SVM, los algoritmos KNN y el RF mantenían su poder predictivo con la *ERS* con respecto a la predicción con el *ERA* (tabla 7C y tabla 7D). Del mismo modo, es importante recalcar que el generar el modelo predictivo con la *ERS* fue con el objetivo de testar el poder diagnóstico y reforzar la representación fenotípica de receptividad endometrial definida en la presente tesis doctoral (*ERS*). Sin embargo, utilizar los 56 genes de la *ERS* para generar el modelo de predicción no hubiera sido adecuado. Hay que recordar que los genes de la *ERS* fueron seleccionados a partir del mismo set de entrenamiento con el que se generó el modelo de predicción, a diferencia de los genes del *ERA*, y se podría estar usando una firma muy específica de este grupo de muestras. Una firma muy específica podría perder poder de generalización del

predictor cuando se usara en una población más amplia. Además, con todos los genes del *ERA* el predictor contenía más variables con las que hacer frente a nuevas variantes fenóticas. Es por estas razones que no se utilizaron los genes de la *ERS* y se definió el predictor con la totalidad de genes del *ERA*.

Otro de los principales objetivos en esta tesis fue conocer con profundidad el transcriptoma del endometrio en fase receptiva, pero sobretodo crear una herramienta molecular de uso clínico para un diagnóstico objetivo del estado de receptividad endometrial. Luego la finalidad era clasificar el endometrio en LH+7 (fase receptiva) en términos transcriptómicos. De este modo, aunque los resultados del entrenamiento mostraron el poder diagnóstico de esta herramienta molecular en todas las fases del ciclo menstrual, es de la fase receptiva de la se introdujeron más muestras (n=39), y es de esta fase de la que se van a tomar las biopsias a testar. De manera que la muestra se datará como receptiva o no receptiva.

En el modelo de dataje endometrial, si se consideraban todas las muestras no receptivas juntas, el set de entrenamiento se encontraba balanceado (R= 39, No receptivas= 40). El modelo con el algoritmo SVM poseyó una sensibilidad de 1, eso quiere decir que no hubo ninguna muestra que fuera receptiva que la clasificara mal, y cuando clasificaba una muestra mal de cualquier fase del ciclo, lo hacía como receptiva. Es por ello, que este modelo no tenía falsos negativos, esto implicaría que si una muestra fuera clasificada como no receptiva sería bastante fiable.

Las muestras R-FI fueron lanzadas con el modelo de dataje endometrial, clasificándose como receptivas sin distinciones con las muestras R, como era de esperar por los análisis preliminares en el agrupamiento jerárquico (figura 18), en el PCA (figura 20) y en análisis de expresión génica diferencial entre

estas muestras R-FI y las R (vid. punto 6.8). Es por ello, que se generó el modelo de clasificación patológica para intentar distinguir entre las R-FI y las R.

El modelo estaba desbalanceado, es decir, en el set de entrenamiento el 22% de las muestras eran R-FI y el 78% eran R (R-FI n=11; R n=39), por ello, la probabilidad de acertar por azar no era la misma de partida. Luego, del 90% de aciertos del predictor, el 76% de aciertos fueron debidos a las muestras R, que se sabía con anterioridad que se clasificaban bien, y el 14 % de los aciertos a las muestras R-FI (tabla 7C). Por la matriz de confusión se pudo observar que, en el mejor de los casos, el algoritmo SVM clasificaba bien el 63% de las muestras R-FI (7 muestras), frente a un 97% de las muestras R (38 muestras) (tabla 7C). Luego, la buena predicción de este segundo modelo fue debida a las muestras R. No obstante, aunque el modelo de dataje endometrial sólo conseguía clasificar algunas muestras R-FI permitió detectar con una sensibilidad de 0,98 cuando una muestra poseía FI, mientras que la especificidad era menor que en el modelo de dataje endometrial. De modo que muestras con FI se clasificaban como receptivas normales por la falta de especificidad, sin embargo, por la buena sensibilidad, cuando diagnosticaba una muestra como FI lo hacía de una manera muy eficaz (tabla 7C).

- **Herramienta molecular diagnóstica de la receptividad endometrial**

La herramienta molecular establece el diagnóstico de receptividad mediante el algoritmo SVM, que tiene una sensibilidad y especificidad muy buena (tabla 7A). Sin embargo, el diagnóstico para la clasificación patológica se realiza a partir del consenso entre dos algoritmos (SVM y KNN), optándose por el diagnóstico menos optimista. Esta forma de diagnosticar aumenta la capacidad de detectar un FI, ya que se prefiere alarmar y descartar el diagnóstico mediante otras pruebas, a que se quede sin detectar una muestra

FI. Así mismo, la probabilidad asociada a la predicción confiere un valor de referencia que aporta información sobre la fiabilidad del diagnóstico de receptividad.

El estudio endometrial por histología vs. la predicción del modelo de dataje utilizado en la herramienta molecular corroboró los trabajos realizados por Murray en cuanto a la variabilidad interobservador que supone el dataje histológico (Murray y cols., 2004), ya que se hubiera esperado una concordancia mucho mayor entre patólogos (tabla 8). Por otra parte, también se observó que la mayor variabilidad en el dataje entre patólogos se producía en la fase receptiva y mucho más en la post-receptiva (figura 23), lo cual también estuvo en acuerdo con lo reportado por Murray (figura 7). El dataje predictivo mostró una elevada capacidad diagnóstica en fase receptiva (tabla 8) superando el problema de la subjetividad.

Las muestras PF, PR y PS se diferenciaban de las R por 176, 200 y 94 genes diferencialmente expresados respectivamente (vid. punto 6.5). Como era de esperar por los genes diferencialmente expresados, fueron las muestras PR y R las que mejor fueron clasificadas por el predictor (matriz de confusión figura 7A y figura 23). Por el contrario, el predictor mostró menor número de aciertos en la fase proliferativa y aún menos en la post-receptiva, que fueron las muestras que menos genes diferencialmente expresados tenían con respecto a las R (matriz de confusión figura 7A y figura 23). Estos resultados eran de esperar al haberse realizado la selección de genes del *ERA* a partir de la comparación pre-receptivas vs. receptivas.

Sin embargo, no hay que olvidar que la herramienta molecular diagnóstica de la receptividad endometrial tenía como objetivo clasificar las muestras receptivas, ya que de esta fase del ciclo se va a tomar la biopsia. Es

por ello, que se seleccionaron los genes que cambian de la situación pre-receptiva a la receptiva, que es donde se produce la apertura de la WOI, y debido a esto el predictor clasificaba mejor la fase pre-receptiva y receptiva y peor la proliferativa y post-receptiva. Así mismo la clase peor clasificada (post-receptiva) fue la fase que más se parecía en expresión a la receptiva (sólo 94 genes diferencialmente expresados), y esto puede ser debido a que muchos genes que cambian del estado pre-receptivo al receptivo se mantengan o cambien muy poco en el post-receptivo hasta que se produzca la menstruación (figura 14).

- **Perspectivas en investigación y en clínica**

La predicción de la herramienta, y más concretamente la especificidad del modelo de clasificación patológica, podría ser mejorada con un número de muestras de FI mayor que contemplara los distintos perfiles que parece haber dentro del grupo R-FI y en un tamaño muestral suficiente. Pero el set de entrenamiento para el FI es muy complicado de establecer debido a la dificultad de definir y obtener las muestras. No obstante, el programa informático diseñado para la herramienta diagnóstica tiene la posibilidad de re-entrenar los modelos de predicción establecidos con nuevas muestras (anexo III).

Por otra parte, el programa informático diseñado también tiene la posibilidad de generar nuevos modelos predictivos para otros desordenes endometriales de los que se vaya obteniendo un set de entrenamiento adecuado. Es por ello que este trabajo de investigación, con el análisis de nuevas muestras de las que se tenga una clasificación clínica adecuada, ampliará sus aplicaciones diagnósticas a más desordenes endometriales y mejorará su poder diagnóstico.

Debido a la variabilidad de tratamientos clínicos que existen y a la gran cantidad de factores que influyen en la salud reproductiva, todavía no ha sido posible obtener resultados estadísticos para validar a nivel clínico el diagnóstico de la herramienta. No obstante, la validación clínica a gran escala está en marcha en este momento y supondrá poder correlacionar el diagnóstico de la herramienta molecular con las tasas de implantación y por consiguiente el diagnóstico endometrial con la probabilidad de implantar el embrión.

Del mismo modo, el *ERA*, de forma independiente a la herramienta molecular, podrá ser usado como *microarray* específico de receptividad endometrial para realizar otro tipo de investigaciones en las que se quiera estudiar sólo los genes relacionados con dicho proceso.

En definitiva, la herramienta molecular ofrece un nuevo nivel de análisis objetivo, que confiere un parámetro multivariante de uso clínico, el cual le permite al ginecólogo obtener más información sobre el estado endometrial de la mujer. Esta es la primera vez que se utiliza una herramienta molecular, basada en la tecnología de *microarrays*, que pueda ser usada clínicamente en medicina reproductiva para evaluar el endometrio (Díaz-Gimeno y cols., 2011). La herramienta molecular (*ERA + predictor*) es un método innovador y objetivo que permite la evaluación clínica y diagnóstico endometrial. Por otra parte, es una herramienta que incluye un número reducido y finito de genes relacionados con la receptividad endometrial, evitando usar *microarrays* de genoma completo, reduciendo el coste y simplificando el análisis de datos. A su vez, la metodología desarrollada para confeccionar esta herramienta de diagnóstico podrá ser el origen de nuevas investigaciones para diagnosticar y evaluar, no sólo a nivel endometrial, sino en otros aspectos de la medicina reproductiva.



*“Las ciencias no tratan de explicar, incluso apenas tratan de interpretar,
construyen modelos principalmente.*

*Por modelo, se entiende una construcción matemática que,
con la adición de ciertas interpretaciones verbales, describe los fenómenos observados.*

*La justificación de tal construcción matemática es
sólo y precisamente que se espera que funcione”*

John Vonn Newman

Conclusiones

8. CONCLUSIONES

1. Se ha seleccionado un conjunto de genes estadísticamente significativos que posee relevancia funcional en el proceso de la receptividad endometrial.
2. Se ha construido el *microarray* personalizado *ERA* que analiza la expresión de dicho conjunto de genes.
3. Se ha definido y caracterizado la *ERS* que permite distinguir y clasificar el fenotipo receptivo.
4. Se ha desarrollado un modelo de predicción que es capaz de diagnosticar el fenotipo endometrial receptivo.
5. La herramienta molecular diagnostica la receptividad endometrial normal y facilita el diagnóstico de endometrios en fase receptiva de mujeres con fallo de implantación, mediante la utilización de un segundo modelo de predicción.
6. La herramienta molecular diagnóstica de la receptividad endometrial mejora el dataje de los métodos histológicos clásicos en fase receptiva e incorpora un sistema de detección objetivo del estatus de receptividad endometrial.

PUBLICACIONES DERIVADAS DE LA PRESENTE TESIS DOCTORAL:

Patente

Inventores: C Simón; JA Horcajadas; **P Díaz-Gimeno**; A Pellicer.

Título: Perfil De Expresión Génica Como Marcador De La Receptividad Endometrial

Nº De Solicitud: Ax080129WO/ PCT2008/000513

Fecha de prioridad: 22jul 2008 Ampliación 22jul 2009

Entidad Titular: Equipo IVI Investigación S.L

Países: País de prioridad España, Extendido a Europa y USA

Empresas que lo explotan iGenomix S.L.

Revista Internacional

P Díaz-Gimeno, JA Horcajadas, JA Martínez-Conejero; FJ Esteban; P Alama; A Pellicer, C Simón. A Genomic Diagnostic Tool For Human Endometrial Receptivity Based On The Transcriptomic Signature. Fertility and Sterility 2011; 95: 50-60.

Congresos Internacionales

P Díaz-Gimeno, JA Martínez-Conejero, N Garrido, A Pellicer, C Simón., JA Horcajadas. The Endometrial Receptivity Array (ERA) Improves The Histological Dating And Provides An Objective Method For Clinical Evaluation Of Endometrial Receptivity.

Tipo de participación Póster

Congreso 65th Annual Meeting ASRM (American Society For Reproductive Medicine)

Denver Colorado.23-27 Octubre 2010.

P Díaz-Gimeno, JA Horcajadas, JA Martínez-Conejero, P. Alamá, A Pellicer, C Simón. Defining The Gene Expression Profile Of Human Endometrial Receptivity. 57th Annual Meeting SGI. Journal Of The Society For Gynecologic Investigation 2010: 3: 335a-336a.

Tipo de participación Póster

Orlando Florida. 23-27 Marzo 2010.

P Díaz-Gimeno, JA Horcajadas, JA Martínez-Conejero, FJ Esteban, A Pellicer, C Simón. A New Genomic Method for Human Endometrial Receptivity Diagnosis. 3rd Internacional IVI Congress: P-15: 69.

Tipo de participación Póster
Madrid Spain 14-16 Mayo 2009.

P Díaz-Gimeno, JA Horcajadas, JA Martínez-Conejero, I. Medina, A Pellicer, C Simón. Development Of A Customized Array For The Molecular Diagnosis Of Human Endometrial Receptivity. American Society For Reproductive Medicine'S 64th Annual Meeting 2008: P-164: 59.

Tipo De Participación Póster
San Francisco Ca EEUU 8-12 Noviembre 2008.

P Díaz-Gimeno, JA Horcajadas, JA Martínez-Conejero, FJ Esteban, A Pellicer, C Simón. Development Of A Customized Array For Endometrial Receptivity (ERA). 2nd SGI Internacional Summit On Reproductive Medicine. Scientific Program And Abstracts 2007: P.8: 67.

Tipo De Participación Póster
Valencia, Spain. 8-10 Noviembre 2007.

Congreso Nacional

P Díaz-Gimeno, JA Horcajadas, FJ Esteban, I Medina, JA Martínez-Conejero, C Simón. Desarrollo de una herramienta Molecular basada en el perfil de expresión génica para evaluar la receptividad endometrial humana. XXXVII Congreso De La Sociedad Española De Genética: Ps8.

Tipo de participación Póster
Torremolinos Málaga España 29 Septiembre-02 Octubre 2009.

OTRAS PUBLICACIONES RELACIONADAS CON LA PRESENTE TESIS:

Revista Internacional

JA Horcajadas, **P Díaz-Gimeno**, A Pellicer, C Simón. Uterine Receptivity and the Ramifications of Ovarian Stimulation on Endometrial Function. Seminar Reproductive Medicine 2007; 25: 454-460.

Capítulos de Libro Internacional

P Díaz-Gimeno, JA Horcajadas. "El Paso De La Medicina Anatómica A La Medicina Molecular: ¿Qué es un microarray?. En: C. Simón; JA. Horcajadas; J. Garcia Velasco; A. Pellicer. El Endometrio Humano: De La Investigación A La Clínica. Médica Panamericana. Buenos Aires, Madrid, 2009.

P Díaz-Gimeno; F Loro; D Valbuena; JA Horcajadas. "Estudio Genómico Del Endometrio Humano En Ciclos Naturales Y Estimulados". En C. Simón; JA. Horcajadas; J. Garcia Velasco; A. Pellicer. El Endometrio Humano: De La Investigación A La Clínica. Médica Panamericana. Buenos Aires, Madrid, 2009.



*“Si he conseguido ver más lejos,
es porque me he aupado en hombros de gigantes”*

Isaac Newton

Bibliografía

9. BIBLIOGRAFÍA

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

Acosta AA, Elberger L, Borghi M y cols. Endometrial dating and determination of the window of implantation in healthy fertile women. *Fertil Steril* 2000; 73: 788-98.

Achache H, Revel A. Endometrial receptivity markers, the journey to successful embryo implantation. *Hum Reprod Update* 2006; 12: 731-46.

Aghajanova L, Simon C, Horcajadas J. Are favourite molecules of endometrial receptivity still in favour? *Expert Rev Obstet Gynecol* 2008a; 3: 487-501.

Aghajanova L, Hamilton AE, Giudice LC. Uterine Receptivity to Human Embryonic Implantation: Histology, Biomarkers, and Transcriptomics. *Semin Cell Dev Biol*. 2008b; 19: 204-11.

Al-Shahrour F, Diaz-Uriarte R, Dopazo J. FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. *Bioinformatics* 2004; 20: 578-80.

Al-Shahrour F, Minguez P, Vaquerizas JM y cols. BABELOMICS: a suite of web tools for functional annotation and analysis of groups of genes in high-throughput experiments. *Nucleic Acids Res* 2005; 33 (Web Server issue): W460-4.

Azuaje F, Dopazo J, editors. *Data analysis and visualization in Genomics and Proteomics*. John Willey and sons ltd. England; 2005.

Balash J, Fabregues F, Creus M y cols. The usefulness of endometrial biopsy for luteal phase evaluation in infertility. *Hum Reprod* 1992; 7: 973-7.

Barrett JC, Kawasaki ES. Microarrays: the use of oligonucleotides and cDNA for the analysis of gene expression. *Drug Discov Today* 2003; 8: 134-41.

Batista MC, Cartledge TP, Merino MJ y cols. Midluteal phase endometrial biopsy does not accurately predict luteal function. *Fertil Steril* 1993; 59: 294-300.

Bellver J, Melo MA, Bosch E y cols. Obesity and poor reproductive outcome: the potential role of the endometrium. *Fertil Steril* 2007; 88: 446-51.

Benjamini YH, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc B MET* 1995; 57: 289–300.

Bentin-Ley U, Sjogren A, Nilsson L y cols. Presence of uterine pinopodes at the embryo-endometrial interface during human implantation in vitro. *Hum Reprod* 1999; 14: 515-20.

Bergh PA, Navot D. The impact of embryonic development and endometrial maturity on the timing of implantation. *Fertil Steril* 1992; 58: 537-42.

Borthwick JM, Charnock-Jones DS, Tom BD y cols. Determination of the transcript profile of human endometrium. *Mol Hum Reprod* 2003; 9: 19-33.

Bourgain C, Santonja N. Estudio histológico del endometrio. En Simón C y cols. *El endometrio humano: de la investigación a la clínica* Ed. Médica Panamericana. Buenos Aires, Madrid.; 2009; p. 40-5.

Brar AK, Handwerger S, Kessler CA y cols. Gene induction and categorical reprogramming during in vitro human endometrial fibroblast decidualization. *Physiol Genomics* 2001; 7: 135-48.

Brazma A, Hingamp P, Quackenbush J y cols. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 2001; 29: 365-71.

Breiman L. Random Forests. *Machine Learning* 2001; 45: 5–32.

Bulmer JN, Johnson PM. Immunohistological characterization of the decidual leucocytic infiltrate related to endometrial gland epithelium in early human pregnancy. *Immunology* 1985; 55: 35-44.

Buyse M, Loi S, van't Veer L y cols. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst* 2006; 98: 1183-92.

Carson DD, Lagow E, Thathiah A y cols. Changes in gene expression during the early to mid-luteal (receptive phase) transition in human endometrium detected by high-density microarray screening. *Mol Hum Reprod* 2002; 8: 871-9.

Conesa A, Nueda MJ, Ferrer A y cols. maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments. *Bioinformatics* 2006; 22: 1096-102.

- Cooke ID. Implantation and the endometrium. *Curr Opin Obstet Gynecol* 1995; 7: 233-8.
- Costa J. Reacción en cadena de la polimerasa (PCR) a tiempo real. *Enferm Infecc Microbiol Clin* 2004; 5: 299-305.
- Coutifaris C, Myers ER, Guzick DS y cols. Histological dating of timed endometrial biopsy tissue is not related to fertility status. *Fertil Steril* 2004; 82: 1264-72.
- Creus M, Ordi J, Fabregues F y cols. α 3 integrin expression and pinopod formation in normal and out-of-phase endometria of fertile and infertile women. *Hum Reprod* 2002; 17: 2279-86.
- Critchey H. Endometrial steroid receptor expression throughout the menstrual cycle. En: O'Brien P y cols. *Disorders of the Menstrual Cycle*. Ed. RCOG Press. London; 2000.
- Cheon YP, Li Q, Xu X y cols. A genomic approach to identify novel progesterone receptor regulated pathways in the uterus during implantation. *Mol Endocrinol* 2002; 16: 2853-71.
- Churchill GA. Fundamentals of experimental design for cDNA microarrays. *Nat Genet* 2002; 32 Suppl: 490-5.
- Díaz-Gimeno P, Horcajadas JA. El paso de la medicina anatómica a la medicina molecular: ¿qué es un microarray?. En: Simón C; JA Horcajadas; J Garcia Velasco y cols. *El endometrio humano: de la investigación a la clínica*. Ed. Médica Panamericana. Buenos Aires, Madrid; 2009; p. 99-106.
- Díaz-Gimeno P, Horcajadas JA, Martínez-Conejero JA y cols. A genomic diagnostic tool for human endometrial receptivity based on the transcriptomic signature. *Fertil Steril* 2011; 95: 50-60.
- Dickey RP, Olar TT, Curome DN y cols. Endometrial pattern and thickness associated with pregnancy outcome after assisted reproduction technologies. *Hum Reprod* 1992; 7: 418-21.
- Domenech J. Unidad 14: Medida del cambio: análisis de diseños con medidas intrasujeto. Estudio de la concordancia. *Fundamentos de Diseño y Estadística*. Ed. Signo. Barcelona; 2006.
- Domínguez F, Loro F, Simón C. Funcionalidad del endometrio. En Simón C; JA Horcajadas; J Garcia Velasco y cols. *El endometrio humano: de la investigación a la clínica*. Ed. Médica Panamericana. Buenos Aires, Madrid.; 2009; p. 19-37.

Domínguez F, Yanez-Mo M, Sanchez-Madrid F y cols. Embryonic implantation and leukocyte transendothelial migration: different processes with similar players?. *Faseb J* 2005; 19: 1056-60.

Evans D. Genetics of gynaecological cancer. *Curr Obstet Gynaecol* 1995; 5: 201-5.

Garcia-Velasco JA, Isaza V, Caligara C y cols. Factors that determine discordant outcome from shared oocytes. *Fertil Steril* 2003; 80: 54-60.

Giudice LC, Saleh W. Growth factors in reproduction. *Trends Endocrinol Metab* 1995; 6: 60-9.

Glas AM, Floore A, Delahaye LJ y cols. Converting a breast cancer microarray signature into a high-throughput diagnostic test. *BMC Genomics* 2006; 7: 278.

Hannan NJ, Salamonsen LA. Role of chemokines in the endometrium and in embryo implantation. *Curr Opin Obstet Gynecol* 2007; 19: 266-72.

Haouzi D, Mahmoud K, Fourar M y cols. Identification of new biomarkers of human endometrial receptivity in the natural cycle. *Hum Reprod* 2009; 24: 198-205.

Hastie T TR, Friedman J. The elements of statistical learning. Ed. Springer. New York; 2001.

Hewitt SC, Deroo BJ, Hansen K y cols. Estrogen receptor-dependent genomic responses in the uterus mirror the biphasic physiological response to estrogen. *Mol Endocrinol* 2003; 17: 2070-83.

Hong SH, Nah HY, Lee JY y cols. Analysis of estrogen-regulated genes in mouse uterus using cDNA microarray and laser capture microdissection. *J Endocrinol* 2004; 181: 157-67.

Horcajadas JA, Minguez P, Dopazo J y cols. Controlled ovarian stimulation induces a functional genomic delay of the endometrium with potential clinical implications. *J Clin Endocrinol Metab* 2008; 93: 4500-10.

Horcajadas JA, Pellicer A, Simon C. Wide genomic analysis of human endometrial receptivity: new times, new opportunities. *Hum Reprod Update* 2007; 13: 77-86.

Horcajadas JA, Riesewijk A, Martin J y cols. Global gene expression profiling of human endometrial receptivity. *J Reprod Immunol* 2004; 63: 41-9.

Ishiwata I, Tamagawa T, Tokieda Y y cols. Organogenesis of heart-vascular system derived from mouse 2 cell stage embryos and from early embryonic stem cells in vitro. *Hum Cell* 2003; 16: 15-22.

Jabbour HN, Kelly RW, Fraser HM y cols. Endocrine regulation of menstruation. *Endocr Rev* 2006; 27: 17-46.

Jolivet A, Gautray JP. Clinical investigation of the menstrual cycle. I. Diagram of the normal menstrual cycle. *Fertil Steril*. 1978; 29: 40-2.

Kabir-Salmani M, Nikzad H, Shiokawa S y cols. Secretory role for human uterodomes (pinopods): secretion of LIF. *Mol Hum Reprod* 2005; 11: 553-9.

Kao LC, Germeyer A, Tulac S y cols. Expression profiling of endometrium from women with endometriosis reveals candidate genes for disease-based implantation failure and infertility. *Endocrinology* 2003; 144: 2870-81.

Khan J, Wei JS, Ringner M y cols. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nat Med* 2001; 7: 673-9.

Kliman HJ, Honig S, Walls D y cols. Optimization of endometrial preparation results in a normal endometrial function test (EFT) and good reproductive outcome in donor ovum recipients. *J Assist Reprod Genet* 2006; 23: 299-303.

Lessey BA, Castelbaum AJ, Sawin SW y cols. Integrins as markers of uterine receptivity in women with primary unexplained infertility. *Fertil Steril* 1995; 63: 535-42.

Li TC, Dockery P, Rogers AW y cols. How precise is histologic dating of endometrium using the standard dating criteria? *Fertil Steril* 1989; 51: 759-63.

Martel D, Frydman R, Glissant M y cols. Scanning electron microscopy of postovulatory human endometrium in spontaneous cycles and cycles stimulated by hormone treatment. *J Endocrinol* 1987; 114: 319-24.

Martin J, Dominguez F, Avila S y cols. Human endometrial receptivity: gene regulation. *J Reprod Immunol* 2002; 55: 131-9.

Mata J, Marguerat S, Bahler J. Post-transcriptional control of gene expression: a genome-wide perspective. *Trends Biochem Sci* 2005; 30: 506-14.

McWilliams GD, Frattarelli JL. Changes in measured endometrial thickness predict in vitro fertilization success. *Fertil Steril* 2007; 88: 74-81.

Medina I, Carbonell J, Pulido L y cols. Babelomics: an integrative platform for the analysis of transcriptomics, proteomics and genomic data with advanced functional profiling. *Nucleic Acids Res* 2010; 38 Suppl: W210-3.

Medina I, Montaner D, Tarraga J y cols. Prophet, a web-based tool for class prediction using microarray data. *Bioinformatics* 2007; 23: 390-1.

Mirkin S, Arslan M, Churikov D y cols. In search of candidate genes critically expressed in the human endometrium during the window of implantation. *Hum Reprod* 2005; 20: 2104-17.

Montaner D, Tarraga J, Huerta-Cepas J y cols. Next station in microarray data analysis: GEPAS. *Nucleic Acids Res* 2006; 34 (Web Server issue): W486-91.

Murphy CR. The cytoskeleton of uterine epithelial cells: a new player in uterine receptivity and the plasma membrane transformation. *Hum Reprod Update* 1995; 1: 567-80.

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

Murphy CR, Shaw TJ. Plasma membrane transformation: a common response of uterine epithelial cells during the peri-implantation period. *Cell Biol Int* 1994; 18: 1115-28.

Murray MJ, Meyer WR, Zaino RJ y cols. A critical analysis of the accuracy, reproducibility, and clinical utility of histologic endometrial dating in fertile women. *Fertil Steril* 2004; 81: 1333-43.

Naciff JM, Jump ML, Torontali SM y cols. Gene expression profile induced by 17alpha-ethynyl estradiol, bisphenol A, and genistein in the developing female reproductive system of the rat. *Toxicol Sci* 2002; 68: 184-99.

Nevins JR, Potti A. Mining gene expression profiles: expression signatures as cancer phenotypes. *Nat Rev Genet* 2007; 8: 601-9.

Nikas G. Cell-surface morphological events relevant to human implantation. *Hum Reprod* 1999; 14 Suppl 2: 37-44.

Niklaus AL, Pollard JW. Mining the mouse transcriptome of receptive endometrium reveals distinct molecular signatures for the luminal and glandular epithelium. *Endocrinology* 2006; 147: 3375-90.

Noyes N, Liu HC, Sultan K y cols. Endometrial thickness appears to be a significant factor in embryo implantation in in-vitro fertilization. *Hum Reprod* 1995; 10: 919-22.

- Noyes R, Hertig A, Rock J. Dating the endometrial biopsy. *Fertil Steril* 1950; 1: 3-25.
- Noyes RW, Haman JO. Accuracy of endometrial dating; correlation of endometrial dating with basal body temperature and menses. *Fertil Steril* 1953; 4: 504-17.
- Noyes RW, Hertig AT, Rock J. Dating the endometrial biopsy. *Am J Obstet Gynecol* 1975; 122: 262-3.
- Okada H, Nakajima T, Yoshimura T y cols. Microarray analysis of genes controlled by progesterone in human endometrial stromal cells in vitro. *Gynecol Endocrinol* 2003; 17: 271-80.
- Oliveira JB, Baruffi RL, Mauri AL y cols. Endometrial ultrasonography as a predictor of pregnancy in an in-vitro fertilization programme after ovarian stimulation and gonadotrophin-releasing hormone and gonadotrophins. *Hum Reprod* 1997; 12: 2515-8.
- Oliveros JC. VENNY (2007). An interactive tool for comparing lists with Venn Diagrams. <http://bioinfogp.cnb.csic.es/tools/venny/index.html>
- Ordi J, Creus M, Quinto L y cols. Within-subject between-cycle variability of histological dating, alpha v beta 3 integrin expression, and pinopod formation in the human endometrium. *J Clin Endocrinol Metab* 2003; 88: 2119-25.
- Papanikolaou EG, Tournaye H, Verpoest W y cols. Early and late ovarian hyperstimulation syndrome: early pregnancy outcome and profile. *Hum Reprod* 2005; 20: 636-41.
- Park SJ, Goldsmith LT, Skurnick JH y cols. Characteristics of the urinary luteinizing hormone surge in young ovulatory women. *Fertil Steril* 2007; 88: 684-90.
- Ponnampalam AP, Weston GC, Trajstman AC y cols. Molecular classification of human endometrial cycle stages by transcriptional profiling. *Mol Hum Reprod* 2004; 10: 879-93.
- Popovici RM, Kao LC, Giudice LC. Discovery of new inducible genes in in vitro decidualized human endometrial stromal cells using microarray technology. *Endocrinology* 2000; 141: 3510-3.
- Psychoyos A. From Lataste to the "window of implantation": 100 years of fascinating discoveries. *Contracept Fertil Sex* 1993; 21: 333-8.
- Puente J, Serna J, Cerrillo M y cols. Ultrasonografía en el estudio del endometrio. En: Simón C y cols. *El endometrio humano: de la investigación a la clínica*. Ed. Médica Panamericana. Buenos Aires, Madrid; 2009; p. 48-65.

Puerto B, Creus M, Carmona F y cols. Ultrasonography as a predictor of embryo implantation after in vitro fertilization: a controlled study. *Fertil Steril* 2003; 79: 1015-22.

Punyadeera C, Dassen H, Klomp J y cols. Oestrogen-modulated gene expression in the human endometrium. *Cell Mol Life Sci* 2005; 62: 239-50.

Quackenbush J. Computational analysis of microarray data. *Nat Rev Genet* 2001; 2: 418-27.

Quackenbush J. Microarray data normalization and transformation. *Nat Genet* 2002; 32 Suppl: 496-501.

Quackenbush J. Computational approaches to analysis of DNA microarray data. *Yearb Med Inform* 2006a: 91-103.

Quackenbush J. Microarray analysis and tumor classification. *N Engl J Med* 2006b; 354: 2463-72.

Quinn CE, Casper RF. Pinopodes: a questionable role in endometrial receptivity. *Hum Reprod Update* 2009; 15: 229-36.

Raga F, Bonilla-Musoles F, Casan EM y cols. Assessment of endometrial volume by three-dimensional ultrasound prior to embryo transfer: clues to endometrial receptivity. *Hum Reprod* 1999; 14: 2851-4.

Reese J, Das SK, Paria BC y cols. Global gene expression analysis to identify molecular markers of uterine receptivity and embryo implantation. *J Biol Chem* 2001; 276: 44137-45.

Remohi J, Ardiles G, Garcia-Velasco JA y cols. Endometrial thickness and serum oestradiol concentrations as predictors of outcome in oocyte donation. *Hum Reprod* 1997; 12: 2271-6.

Riesewijk A, Martin J, van Os R y cols. Gene expression profiling of human endometrial receptivity on days LH+2 versus LH+7 by microarray technology. *Mol Hum Reprod* 2003; 9: 253-64.

Ripley B. Pattern recognition and neural networks. Ed. Cambridge University Press. England; 1996.

Rockett JC, Hellmann GM. Confirming microarray data--is it really necessary? *Genomics* 2004; 83: 541-9.

Saeed AI, Bhagabati NK, Braisted JC y cols. TM4 microarray software suite. *Methods Enzymol* 2006; 411: 134-93.

Saeed AI, Sharov V, White J y cols. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* 2003; 34: 374-8.

Salamonsen LA, Hannan NJ, Dimitriadis E. Cytokines and chemokines during human embryo implantation: roles in implantation and early placentation. *Semin Reprod Med* 2007; 25: 437-44.

Salamonsen LA, Nie G, Hannan NJ y cols. Society for Reproductive Biology Founders' Lecture 2009. Preparing fertile soil: the importance of endometrial receptivity. *Reprod Fertil Dev* 2009; 21: 923-34.

Schena M. Genome analysis with gene expression microarrays. *Bioessays* 1996; 18: 427-31.

Schena M, Shalon D, Davis RW y cols. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995; 270: 467-70.

Shi L, Campbell G, Jones WD y cols. The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. *Nat Biotechnol* 2010a; 28: 827-38.

Shi L, Jones WD, Jensen RV y cols. The balance of reproducibility, sensitivity, and specificity of lists of differentially expressed genes in microarray studies. *BMC Bioinformatics* 2008a; Suppl 9: S10.

Shi L, Perkins RG, Fang H y cols. Reproducible and reliable microarray results through quality control: good laboratory proficiency and appropriate data analysis practices are essential. *Curr Opin Biotechnol* 2008b; 19: 10-8.

Shi L, Reid LH, Jones WD y cols. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 2006; 24: 1151-61.

Shi W, Bessarabova M, Dosymbekov D y cols. Functional analysis of multiple genomic signatures demonstrates that classification algorithms choose phenotype-related genes. *Pharmacogenomics J* 2010b; 10: 310-23.

Shoupe D, Mishell DR, Jr., Laccarra M y cols. Correlation of endometrial maturation with four methods of estimating day of ovulation. *Obstet Gynecol* 1989; 73: 88-92.

Simón C. El endometrio humano. En C Simón; JA Horcajadas; J Garcia Velasco y cols. El endometrio humano: de la investigación a la clínica. Ed. Médica Panamericana. Buenos Aires, Madrid.; 2009; p. 1-10.

Soares SR, Troncoso C, Bosch E y cols. Age and uterine receptiveness: predicting the outcome of oocyte donation cycles. *J Clin Endocrinol Metab* 2005; 90: 4399-404.

Strauss III JF, Gurbide E. Endometrio: regulación y disfunción. En: Yen/Jaffe. *Endocrinología de la Reproducción*. 3ªedn. Vol. 9: Panamericana; 1993; p. 333-81.

Tackels-Horne D, Goodman MD, Williams AJ y cols. Identification of differentially expressed genes in hepatocellular carcinoma and metastatic liver tumors by oligonucleotide expression profiling. *Cancer* 2001; 92: 395-405.

Talbi S, Hamilton AE, Vo KC y cols. Molecular phenotyping of human endometrium distinguishes menstrual cycle phases and underlying biological processes in normo-ovulatory women. *Endocrinology* 2006; 147: 1097-121.

Tan BK, Tan LK, Yu K y cols. Clinical validation of a customized multiple signature microarray for breast cancer. *Clin Cancer Res* 2008; 14: 461-9.

Tan YF, Li FX, Piao YS y cols. Global gene profiling analysis of mouse uterus during the oestrous cycle. *Reproduction* 2003; 126: 171-82.

Thie M, Harrach-Ruprecht B, Sauer H y cols. Cell adhesion to the apical pole of epithelium: a function of cell polarity. *Eur J Cell Biol* 1995; 66: 180-91.

Tierney EP, Tulac S, Huang ST y cols. Activation of the protein kinase A pathway in human endometrial stromal cells reveals sequential categorical gene regulation. *Physiol Genomics* 2003; 16: 47-66.

Tseng LH, Chen I, Chen MY y cols. Genome-based expression profiling as a single standardized microarray platform for the diagnosis of endometrial disorder: an array of 126-gene model. *Fertil Steril* 2009; 94: 114-9.

Tynan S, Pacia E, Haynes-Johnson D y cols. The putative tumor suppressor deleted in malignant brain tumors 1 is an estrogen-regulated gene in rodent and primate endometrial epithelium. *Endocrinology* 2005; 146: 1066-73.

van 't Veer LJ, Dai H, van de Vijver MJ y cols. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 2002; 415: 530-6.

Van de Vijver MJ, He YD, van't Veer LJ y cols. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002; 347: 1999-2009.

- Van Mourik MS, Macklon NS, Heijnen CJ. Embryonic implantation: cytokines, adhesion molecules, and immune cells in establishing an implantation environment. *J Leukoc Biol* 2009; 85: 4-19.
- Vapnik VN. An overview of statistical learning theory. *IEEE Trans Neural Netw* 1999; 10: 988-99.
- Vaquerizas JM, Conde L, Yankilevich P y cols. GEPAS, an experiment-oriented pipeline for the analysis of microarray gene expression data. *Nucleic Acids Res* 2005; 33 (Web Server issue): W616-20.
- Vidal M, Cusick ME, Barabási AL. Interactome networks and human disease. *Cell*. 2011; 144: 986-98.
- Von Rokitsansky C. Ueber Uterusdrusen-Neubildung. Uterus and Ovarialsarcomen. *Zeitschrift Gesellschaft für Aerzte zu Wien* 1860; 37: 577-81.
- Wang J, Jia M, Zhu L y cols. Systematical detection of significant genes in microarray data by incorporating gene interaction relationship in biological systems. *PLoS One* 2010; 5: e13721.
- Watanabe H, Suzuki A, Kobayashi M y cols. Analysis of temporal changes in the expression of estrogen-regulated genes in the uterus. *J Mol Endocrinol* 2003; 30: 347-58.
- White CA, Salamonsen LA. A guide to issues in microarray analysis: application to endometrial biology. *Reproduction* 2005; 130: 1-13.
- Wilcox AJ, Baird DD, Weinberg CR. Time of implantation of the conceptus and loss of pregnancy *New Engl J Med* 1999; 340: 1796-9.
- Wu X, Pang ST, Sahlin L y cols. Gene expression profiling of the effects of castration and estrogen treatment in the rat uterus. *Biol Reprod* 2003; 69: 1308-17.
- Yanaihara A, Otsuka Y, Iwasaki S y cols. Differences in gene expression in the proliferative human endometrium. *Fertil Steril* 2005; 83 Suppl 1: 1206-15.
- Yoshioka K, Matsuda F, Takakura K y cols. Determination of genes involved in the process of implantation: application of GeneChip to scan 6500 genes. *Biochem Biophys Res Commun* 2000; 272: 531-8.
- Zhang S, Cao J. A close examination of double filtering with fold change and T test in microarray analysis. *BMC Bioinformatics* 2009; 10: 402.

Zhang X, Chen CH, Confino E y cols. Increased endometrial thickness is associated with improved treatment outcome for selected patients undergoing in vitro fertilization-embryo transfer. Fertil Steril 2005; 83: 336-40.

URLs

URL1: http://www.nlm.nih.gov/medlineplus/spanish/ency/esp_imagepages/17087.htm

URL2: [http://1.bp.blogspot.com/_d6R7JCGWJPM/R1GAWreXzcl/AAAAAAAAADQ/OKkTOFKYs-c/s1600-R/Ciclo_menstrual+\(wikipedia\).png](http://1.bp.blogspot.com/_d6R7JCGWJPM/R1GAWreXzcl/AAAAAAAAADQ/OKkTOFKYs-c/s1600-R/Ciclo_menstrual+(wikipedia).png)

URL3: <http://gepas.bioinfo.cipf.es/cgi-bin/loadtool.cgi?tool=prophet>

URL4: <http://bioinfogp.cnb.csic.es/tools/venny/index.html>

URL5: <https://earray.chem.agilent.com/earray/>

URL6: <http://www.bioconductor.org/>

URL7: <http://www.r-project.org/>

URL8: <http://bibiservs.techfak.uni-bielefeld.de/genefisher/old.html>

URL9: <http://www.ncbi.nlm.nih.gov/genome/seq/BlastGen>

URL10: http://www.gen-es.org/12_publicaciones/docs/pub_74_d.pdf

URL11: http://www.cnio.es/es/news/docs/oncochip_1.pdf



*“¡Triste época la nuestra!
¡Es más fácil desintegrar un átomo que un prejuicio!”*
Albert Einstein

Anexos

10. ANEXOS

Anexo I. Consentimiento informado.

INFORMACIÓN CLÍNICA A LA PACIENTE

UTILIZACIÓN DE TEJIDO ENDOMETRIAL EN EL PROYECTO DE DESARROLLO DE UN BIOCHIP DIAGNÓSTICO DE RECEPTIVIDAD ENDOMETRIAL ERA (Endometrial Receptivity Array)

Iniciales de la paciente: _____

Nº de paciente: _____

Nombre del Investigador Principal: _____

METODOLOGÍA: DESCRIPCIÓN Y PROPÓSITO DEL ESTUDIO

Se le ha propuesto que participe en un estudio piloto para testar un nuevo método diagnóstico que permitirá saber si el endometrio (la mucosa del interior de la matriz) esta bien en términos de receptividad endometrial. Este nuevo método de diagnóstico molecular, está basado en el perfil de expresión génica de las células endometriales en el momento del ciclo menstrual que se le toma la muestra. Supondrá conocer el correcto estado receptivo del endometrio como control de la salud del órgano reproductor femenino, de modo que podrá ser ofrecido de forma rutinaria en las consultas ginecológicas, así como para descartar problemas de fertilidad endometriales.

Para este proyecto necesitaremos de una biopsia de su endometrio, es decir, una muestra del mismo. No podemos usar una técnica menos invasiva debido a que necesitamos obtener cantidad suficiente de material endometrial. Usted notará molestias derivadas del procedimiento y puede sangrar un poco tras la biopsia, pero es un proceso habitual sin riesgo añadido.

Con el fin de que pueda decidir si desea participar en nuestra investigación, usted debe comprender las ventajas e inconvenientes de colaborar en la misma para que sea capaz de tomar una decisión informada al respecto. Este proceso es lo que se conoce como consentimiento informado.

Este consentimiento informado le da información detallada sobre el estudio y además, su médico comentará esta información con usted. Cuando haya comprendido en qué consiste el proyecto, se le solicitará que firme este consentimiento informado si desea participar en él. Se le dará una copia de este documento para que pueda guardarlo.

Su participación en nuestra investigación es voluntaria, Ud no cobrará nada ni pagará nada por colaborar con nosotros y puede abandonar dicha participación en cualquier momento si así lo decidiera.

RIESGO ASOCIADO AL ESTUDIO

Usted no tiene ningún riesgo adicional por estar incluida en este proyecto, ya que vamos a utilizar únicamente un fragmento de tejido de su útero. Dejando siempre una parte suficiente para su análisis anatómico-patológico del que se encargará la Dra. Nuria Bosch del IVI.

BENEFICIOS ASOCIADOS CON EL ESTUDIO

Los principales beneficios que se pueden generar de la consecución de esta investigación son:

En Ginecología, servir de alarma y diagnosticar de forma precoz, enfermedades que subyacen al patrón de expresión alterado, tales como la endometriosis. Evitando el sufrimiento de la paciente, así como el avance de la enfermedad.

En medicina reproductiva y enfermedades relacionadas con infertilidad, los beneficios también son obvios, ya que podremos dar sentido a tratamientos de reproducción asistida sólo si el endometrio es adecuado.

Sin embargo, es posible que Ud no se beneficie de los resultados obtenidos por razón del tiempo que se necesita para generarlos. Puede solicitar al investigador una mayor información acerca del estudio.

PERMISO PARA LA REVISIÓN DE SUS DATOS / CONFIDENCIALIDAD

Su identidad y todos los datos referentes a su información personal serán confidenciales, salvo que se solicite lo contrario por la ley. No se publicará ninguna información que lo identifique en los informes o publicaciones resultantes de este trabajo. Se permitirá el acceso, para la revisión de sus registros médicos, al Comité Ético, y a las Autoridades Sanitarias.

Firmando este documento usted está de acuerdo con el mismo.

Sus datos se encontrarán protegidos por la Ley Orgánica 15/1999 de Protección de Datos Personales.

DERECHO A PREGUNTAR SOBRE EL ESTUDIO Y A ABANDONARLO

Si necesita más información sobre el estudio o tiene cualquier pregunta relativa a sus derechos como paciente, contacte con la Dra Gemma Castellón si Ud es paciente de IVI Barcelona, el Dr. Juan García Velasco si Ud es paciente de IVI Madrid, el Dr Manuel Fernández si Ud es paciente de IVI Sevilla y el Dr. Carlos Simón si Ud es paciente de IVI Valencia.

Tiene el derecho a informarse más, así como a abandonar su colaboración en cualquier momento, sin que por ello se altere la relación

médico-enfermo ni se produzca perjuicio en su tratamiento. Le será comunicada cualquier nueva información que pueda afectar a su buena disposición para continuar su participación en este proyecto.

COSTES

Usted no cobrará cantidad económica alguna por participar en este estudio.

MODELO DE CONSENTIMIENTO POR ESCRITO

Nº de paciente _____

TÍTULO DEL ESTUDIO: Desarrollo de un biochip diagnóstico de receptividad endometrial (ERA).

Yo _____
(nombre y apellidos)

He leído la hoja de información que se me ha entregado.

He podido hacer preguntas sobre el estudio.

He recibido suficiente información sobre el estudio.

He hablado con _____
(nombre del investigador)

Comprendo que mi participación es voluntaria.

Comprendo que puedo retirarme del estudio:

1. Cuando quiera
2. Sin tener que dar explicaciones
3. Sin que esto repercuta en mis cuidados médicos

Presto libremente mi conformidad para participar en el estudio.

Firma del participante

Firma del Investigador

Fecha:

Fecha:

Anexo II. Secuencia de las sondas *Agilent* para cada uno de los genes del *microarray* de receptividad endometrial (Endometrial Receptivity Array, ERA).

Gen	Agilent	Secuencia de la sonda
ABCC3	A_23_P207507	GTAAGTGTGCTGAATGCTTTAGATGAGGAAATGATCCCCAAGTGGTGAATGACACGCCTA
ABLIM3	A_01_P004704	GGCGAGATTGAAGGGCTTTTGTATTGTTGTTGGATATTTTTGTTGCCATAAAAGCACAA
ABLIM3	A_24_P123408	TTTGACCGCAGCATGCACAAGCTCCAAAGTGAATTGGCCGGCTGATTCTGAAGGAAGAA
ABLIM3	A_23_P256205	TTTCTGGAAGAGATTGCATCTGAGGAAATTCAGGAAGGATCTTTGTAGATTGGGGGGAGA
ABLIM3	A_23_P256204	TGTGTTCTCATGTAGGATGTCAGCCCTCCCTGCAACTTCTCTTTTGGCCAATGCTTTTT
ACADSB	A_24_P189516	AAGCCATTCATAAAAGAAGCGTCAATGGCCAAATACTATGCATCAGAGATTGCAGGACAAA
ACADSB	A_23_P158570	CAACATCCAGTTGAACACCATTGCAAAGCATATCGATGCAGAATACTGACGTCTATAGGA
ACADSB	A_32_P31945	AATTCAGAAATGGGTTTTGGTTCAGTGATTCTCAAGAAAAAGATCTCTTGCCCATTAAG
ACTA2	A_23_P150053	ATCATGAAGTGTGATATTGACATCAGGAAGGACCTCTATGCTAACAATGCCTATCAGGG
ACTA2	A_01_P006531	ACCGCAAATGCTTCTAAAACACTTTCCCGCTCTCTGCTCTAGCACACAACTGTGAA
ADAMTS1	A_23_P342275	ACCTCTGAGCAGTGATATAGCATAATAAAGCCCCGGGCATTATTATTATTATTTCTTTTG
ADAMTS1	A_23_P211039	TATGGGGGTAGATAGAAAAGGAGTTGAATCATCAGAGTAAACTGCCAGTTGCAAATTTGA
ADAMTS1	A_01_P013378	CTTTATAAGTATTGGTTTGGGTGTTCCCTTCCAAGAAGGACTATAGTTAGTAATAAATGCC
ADAMTS8	A_01_P016559	GAAGAGGAGAGGGGATGGGGCGACAGATCCTATCATCAACTGTCCAGTGGACTGGACCTT
ADAMTS8	A_23_P86956	CACCAATTATGGCTACAATGACATTGTCACCATCCAGCTGGTGCCACTAATATTGACGT
ADRA2A	A_23_P138706	ATCATGTCATTGATGAACTGCCAAAGTCAGGGGAGGAGGGCAGAGACTTTGTGTTTACAT
ADRA2A	A_01_P002991	ACTACAAGGACATGGCTCACAAAAGGTTAATGGATGGTTACCTAGAGGATTACCTAGCCCC
AGR2	A_23_P31407	CTCCTCAATCTGGTTTATGAAACAACACTGACAACACCTTCTCCTGATGGCCAGTATGTC
ALPL	A_01_P005774	TTTTTCTCTTTTGGTGGTGGTTAAAGGGAACACAAAACATTTAAATAAAACTTTCCA
ALPL	A_23_P97043	TCCTCGGAAGACACTCTGACCGTGGTCACTGCGGACCATTCCCAGTCTTACATTTGGT
ALPL	A_24_P353619	GCTACAAGTGGTGGGGGTGAACGAGAGAATGCTCCATGGTGGACTATGCTCACAAACA
ALPL	A_23_P97046	ACCTGGAAGAGCTTCAAACCGAGATACAAGCACTCCCACCTCATCTGGAACCGCACGGAA
AMIGO2	A_23_P14083	TCTTAGCTTTTAGCACTATTGGTAATTCAGAGTAGGCCAAAGGTGATGACTCCCAT
ANG	A_23_P428738	GGCCCAAAGAAAGAGCTACCTGGACCTTTTGTCTGTTGACAACATGTTTAATAAAT
ANG	A_01_P018201	AGAGCTACCTGGACCTTTTGTCTGCTGACAACATGTTTAATAAATAAAATGTCAC
ANK3	A_23_P202269	GTGTCATATTAGGTTAATAAGGCTGCTGTGTTTTAAAGGGCATTTTTATTGGGTTTTGG
ANK3	A_23_P301530	CTCGTAACAGCGAACGGTCAGTCAAGGGATCATAAGTTTTACTGCCAGTATTGAGAAAT
ANK3	A_24_P162173	ACTTTTATTCTTTCAAAGAAAATAGACTGCCATTTCCATCAAGATTAGAGACACCAGCC
ANK3	A_01_P007944	ATCAGCATTATTGACACATAGCTCTAATGACATATGTATGAAAAACCATACACTGGATG
ANK3	A_32_P33304	TGTTGGAATACCGCGGTGATCTGCTTTTATAAATCACTGATTTAAAGGAAAGATGA
ANK3	A_32_P33309	AGCACCCGTAATGGACTTTGGTCTCAATGCTTTGACTCTTTGCCGTGGTTTTGGATGTT
ANXA4	A_23_P16976	AATTA AAAACGCCTACAGCTGCCTCCTAGAATATAGACTGTCTGTATTATTACCTAT
AOX1	A_01_P009551	CTTCTGCTGTCACCTTCGTCTTGTCAGAATGAATATAGACTGTATCTAAGTGGGACC
AOX1	A_23_P154037	TTGGTTTCTCTAGGGTGATATTGTCATTACTCTGCTCTTCAATCCATCCAGCTAAAT
AQP3	A_23_P112481	CTGTCATCTGTGCATAAGGAGAGGAAAGTCCAGGGTGTGTATGTTTCAGGGGCTTCC

AQP3	A_23_P112482	AATTTGGGTCAATACATCCTTTTGTCTCCCAAGGGAAGAGAATGGGCAGCAGGTATGTGT
ARG2	A_23_P128728	TTTGGTCAGACAAGAGAAGGAGGGCATATTGTCTATGACCAACTCTACTCCCGATTCA
ARHGDIB	A_23_P151075	TCAGGTTCCAGTAGTTCATTCTAATGCCTAGATTCTTTTGGTTGTGTGCTGGCCCAATG
ARHGDIB	A_01_P018222	CTAGTCACATCCCCTGCCAGATGGAGTCTTCTTTTGTGAGAGACTGTAACAACAACA
ARID5B	A_23_P97871	AAGACTGGCAAAACAGATGGCAAGGGATGCCCTCTTTTTTCATAAAACTCTCCAAGTTCA
ARID5B	A_01_P007755	TGGTGCATGAATTCTCAAGTACTGTATTTCACTGTGTTGGTGTGTCTGATGGAAATTCG
ARID5B	A_32_P18440	GAGCACTTAAAGTCCAGTGTGGCTGTAGTGTATTGATATTCTGCCTGTCTCCTCATG
ARID5B	A_24_P170667	AGAACGAATACCCTATTTAGGTTTTAAACAGATTAACCTTTGGACTATGTTTCAAGCTGC
ARID5B	A_24_P944437	ATGTAGAGATCCAGTGTAAAGAGTTCATTGCTTCAATTAATTATTACCTCCTGTGG
ARID5B	A_01_P007754	ATCAGGGCATATACAAAAGGGTTTGTAAAACCTCGATGTTAACTTTACAACCTTCTGACC
ARNSE4	A_23_P205531	GGTTAGATGCCACCATGTAGGGATTATCGCGAGTGGTTGACCTTACACTACTCCTTAAA
ARNSE4	A_01_P017372	GCTTTTCTGTAATAAGCTCCTTTTATAATAGTCTCAGCTTAGCTCTCAGATCCTAT
ASPM	A_23_P52017	ATCACAAATCCCCTGCAAGCTATTCAAATGGTGATGGATACGCTTGGCATTCTTATTAG
ASPM	A_32_P231556	TCTTTGGAAGATAACAGTGTATGCACAGGGTTGGCTATGTTTATTCTGAATCGCCTACTT
ASPM	A_24_P911179	AGGCCGAGTAAGAGAGATTATGACATTTCTGTAAATGATGTTCTGAAACACCAAGCTAT
ASPM	A_01_P018684	ATCACAAATCCCCTGCAAGCCATTCAAATGGTGATGGATACACTTGGCATTCTTATTAG
ASS1	A_01_P018905	GTGGGGGGCAGCTGCGGTGGGGAGCTATAAAATGACAATTAAGATACACTAGTCTTTT
ASS1	A_23_P31922	ATTGTTGTGATAATTTGTAATTGTGACTTGTCTCCCGGCTGGCAGCGTAGTGGGGCTG
ASS1	A_23_P31921	TCCCTGGAGGATGCCTGAATTCTACAACCGGTTCAAGGGCCGCAATGACCTGATGGAGTA
ASS1	A_32_P13102	TGTGAAGGTGACCAACGTC AAGGATGGCACCACCACCAGACCTCCTTGGAGCTCTTCAT
ATP1B1	A_01_P014301	AATAATTGCCAGGAGTACAGTGTCTTGTGTGATCTTGTATTGATCAGTCAGGTTAAAACAATG
ATP1B1	A_23_P146943	AACTGTCATACGTATGGGACCTACACTTAATCTCTATGCTTTACACTAGTCTTGCATTT
ATP1B1	A_23_P62932	AATAATTGCCAGGAGTACAGTGTCTTGTGTGATCTTGTATTGATCAGTCAGGTTAAAACAACG
ATP1B1	A_01_P010458	TGGGTGGGGGTTGTCATGGGGAACTGCCCTTTAAATTTAAGTGACACTACAGAAAAA
ATP6V0E2L	A_23_P61960	TGAAGCAGAGAAAACATTACACACAAAAAGCAACATAGTCATGTGGGTCCAGATGGCCTC
ATP6V0E2L	A_23_P61956	ATGTGGGTCCAGATGGCCTCAGTCTAGATGTTGGCACCCCTTGTCTGTCTCCTCAGAG
ATP6V1A	A_24_P396994	TATGCTCAGATATTGATCGTAAGTCTCCCTCACCTGTTACAGAGTTTTCAGATCGGTCAC
ATP6V1A	A_23_P211965	TTTGTCTAGGTGTACCCTTTCCTCATCTCTATTAATTTGTAACAGGACTACTGCATGT
ATP6V1A	A_01_P010392	GTAGAGGAAAGGGTGTGTAACATGGTAACAATCTCAAATACCCAAATTTGATAGCAT
BARD1	A_23_P67771	TTGATGGATGCTACTTCTATTTGTGGGAACTTCAAACACCATCCAAAGGACAACCTTA
BARD1	A_01_P005782	GCACGGTTTGTGAGAGCCAGTCATTGTGCTGTTTTTAATTTTTCACATTTTACAATA
BCL6	A_23_P57856	CTGCGTTAAAGGCTCGATTTTGTATCTGCAGGCAGACACGGATCTGAGAATCTTTATTGA
BCL6	A_01_P001583	TTTGGCTGTGCTAAACTGCATCACCGGTTGTAATAATAGCTGTACAAATATCAGAAT
BIRC3	A_01_P001085	ATCAGCGTCTATACATTGAAGGTGTGCATATATGTTGAATGACATTTTAGGGACATGGT
BIRC3	A_23_P98350	TCCTATACATCGAAGGTGTGCATATATGTTGAATGACATTTTAGGGACATGGTGTTTTA
BUB1B	A_01_P000082	CTAAACAGGCTCATTACAATGGTTACCTTGTATTTAACCAATTTGTCTACTTTTCC
BUB1B	A_23_P163481	CACAGTGTCTGTTCTTGGGGAGCTTGCAGCAGAAAATGAATGGGGTTTTTGACACTACATT
C10orf10	A_24_P329795	TGTGTTTTCTGGAGATAGAATGTAAACCATATTAAGGAAAAAGTTTCAGACAAGCA
C10orf10	A_23_P35597	AAGACCACCGCAGGATGGGCAGCAGACTCTGACCTAAGAAGCTGGACTTTTGGGCAGTC

Desarrollo de una Herramienta Molecular Diagn3stica de la Receptividad Endometrial

C10orf10	A_23_P35595	CGCCCCAGCAGTGCCTCAGAACACTCTACTCGCACCTCCCGGTGATCCATGAACTCTGA
C10orf3	A_01_P003476	GTAAACCAAAAACTTTTAAATTTCTTCAGGTTTTCTAACATGCTTACCACCTGGGCTACTG
C11orf8	A_01_P012231	TCGCTAGCACAAAATATTGTCGCTAATAGTCATTCTGTTTTCCATTGTAATGCTGTT
C14orf161	A_23_P77041	ACACACGCTTATTGCCGTGGCAACAGCGGTAGTGTAGGGGATTAATTTTTATAGCATT
C14orf161	A_23_P77043	TCAACTTAAGCATAAAGGGCTCTGAACTTTTCCACTTTAGAGTGACCGTCATTTCAGGAG
C3	A_01_P000840	ATGCCAGGACCTCGGCACCTTCACTGAGAACATGGTTGCTTTGGGTGCCCAACTGACC
C3	A_23_P101407	GTGCCGCTGTGCTGAGGAGAATTGCTTCATACAAAAGTCGGATGACAAGTCACCTGGA
C3	A_23_P101400	ATGCCAAGACGAAGAGAACCAGAAACAATGCCAGGACCTCGGCGCCTTCCCGAGAGCAT
C4,4 ^a	A_01_P000020	ACCGTGCCTTATGCCTGTGTGATCAGTTTCTGGCACACAGATGCCTCAATAAGATTT
C4BPA	A_01_P005981	CCTCTTGCAAATCAATACAGATCAGTTTAGCAAATCTACTGTCAATTTGGCAGTGATATT
C4BPA	A_23_P97541	TTCTCAAAAGAAGGAGGAAAAGGTGCTTGTGGCTTGCCTCTTGCAATTCAATACAGAT
CALB2	A_23_P21092	GCTGGATGCCCTTTTGAAGGATCTGTACGAGAAAAACAAAAGGAAATGAATATTCAACA
CAPN6	A_23_P217570	GTAGTACTTAGCCTACCTAGACCAGCAAGCATTCTTTTTAGCTCGCTCATTTTTTACCA
CAPN6	A_01_P018103	CTTCTGCAGTGCCTTCACTACACTGCCTTACATAAACCAAAATCACAATAAAGTTCATATT
CBR3	A_23_P40453	GTGGTGAATATCAGTAGTTTGCAGTGTTTAAGGGCTTTTGAAAATGCAGTGAAGATCTG
CBR3	A_24_P934477	CAGGTCATGAGGTTTCATTATCGGCAGTAACTCGTTGCACATGTTTCTAGTGGCAAAAA
CBR3	A_23_P40445	TTGGTCCATGACAAAAGTTGTGCAAACTGGTAAACGTCTGCTTCGGAGCTGTGCTTAA
CCNB2	A_32_P72822	TGTGTAGTTAAGCACCAACAGCAGAGAGAACTTAGACACTACCACCAAGCCTTGTGA
CCNB2	A_23_P65757	TTGATTTTGTACATAGTCTCTGGTCTATCTCATGAAACCTTCTCTCAGACCAGTTTTCT
CCNB2	A_01_P000086	AGTCCTCTGGTCTATCTCATGAAACCTTCTCTCAGACCAATTTTCTAAATATATATTGAG
CD7	A_23_P118862	GGCAGAGGACGGGAGGAGACCAGTCCCCACCAGCCGTACCAGAAATAAAGGCTTCTGT
CD7	A_01_P019708	CATGTCCCGCAGCCGTGCAACACGCTGCTCCTCCCCAACCCAGTACCAGTGACCCAGTGG
CD7	A_23_P431815	GACAACCTGACTATCACCATGCACCGCTGCAGCTGTCGGACACTGGCACCTACACCTGC
CDA	A_01_P001621	ATCAGCCTCCCCAAGGTTCTGTCTGTTCCGAGCAACTTTTCTAATTATAAACATCACAG
CDA	A_23_P34597	AGAATGCCCACTGCCTGTAACAGCCACCTGGAGAACTTCATAAAGATGTCTCACAGCCCT
CDC2	A_23_P138507	CCCATGTCAAAAACCTTGGATGAAAATGGCTTGGATTGCTCTCGAAAATGTTAATCTATG
CDC2	A_01_P011602	TATCAAATTAAGCTGTACTTCATCTTCTAATTTCAAAGTATAACTTAAAAATGTAAT
CDC20	A_23_P149195	ATCCACCAAGGCATCCGCTGAAGACCAACCCATCACCTCAGTTGTTTTTATTTTTCTAA
CDC20	A_01_P005979	CCGCTGAAGACTAACCCATCACCTCAGTTGTTTTTATTTTTCTAATAAAGTCATGTCTC
CDC20	A_23_P149200	GGTAATGATAACTTGGTCAATGTGTGGCCTAGTGTCTCTGGAGAGGGTGGCTGGTTCTCT
CENPE	A_23_P253524	CTGAGTGCAAACTCAGTAGACTCCTCTTTGTCACTTCTCTGGAGATCCAGCATTCCTTA
CENPE	A_01_P006340	GGCTTTACTTTGCCACTAGAGTTGAAATATAAGGGAACAGGAAATGAATGCAATTGGTA
CES1	A_23_P206733	ACCTGAAGAGCTTCAAGCTGAAAGGAATTTCCACACTGTCCCTACATGGTGGGAATTA
CFD / DF	A_01_P001685	AGCAATAAAGTCATCCACTCCTGCATCTGGTTGGTCTTTATTGAGCACCTACTATATGCA
CFD/ DF	A_32_P540234	GAGCAATGAAGTCATCCACTCCTGCATCTGGTTGGTCTTTATTGAGCACCTACTATATGC
CKB	A_23_P25674	AACCTGGGCAAGCATGAGAAGTTCTCGGAGGTGCTTAAGCGGCTGCGACTTCAGAAGCGA
CKB	A_24_P61537	AGAAATGAAGCCCGGCCACACCCGACACCAGCCCTGCTGCTTCTAACTTATTGCTGG
CKB	A_01_P017811	CTAACTTATTGCTGGGAGTGCACCCATGCACCCCTGATGTTCCGCGCCTGGCGAGCC
CLDN10	A_01_P007324	GTAAGAACTGTCTGATATGAATCACAACATGGATGAATGTAGTATTTTCTGAAAGTGTG

CLDN10	A_23_P48350	AAAGTATGAATTAGGAGCCGCTGTTTATTGGATGGGCAGGAGCCTCACTGTGCATAAT
CLDN4	A_23_P19944	GGAAGTCCTGGGGTTTTCTCTTCTTCTTGTGGTTTCTGTTTTGTAATTAAGAAGA
CLDN4	A_24_P115183	AACATCATCCAAGACTTCTACAATCCGCTGGTGGCCTCCGGGCAGAAGCGGGAGATGGGT
CLDN4	A_01_P000670	GATGGCTCAACCGTGACTTTGGGCTGCTTGCATCGGTGTGGCCACTGCCCCATTTA
CLU	A_23_P215918	TCCAGGAAGAACCCTAAATTTATGGAGACCGTGGCGGAGAAAAGCGCTGCAGGAATACCGC
CLU	A_23_P215913	CGCTGAGAGGTTGACCAGGAAATACAACGAGCTGCTAAAGTCTACCAGTGGAGATGCT
COBL	A_23_P313623	GTCTACAGGCCAAAATGCGCACAGTTGATTTTCGGTGTGTTCTGTATAACGGCTTGAAA
COBL	A_32_P86173	GTATTTGGCCACCCAGCATTTTTGGGCCGAAGAAAAAATCAAACCTGTTGTCCAGAGAC
COL16A1	A_01_P003861	AAACCATGAAGGGCCTTTTGGCTGAAATCACACCTGCCTTTGGATGAAGACTCCGTT
COL16A1	A_23_P160318	TGGATGAAAGACTCCGTTGGGAATAAATGGCCAAAGCTTATAGGACTCTGTGACAGTTG
COMP	A_24_P264943	TTACCGCTGCAATGACACCATCCAGAGGACTATGAGACCCATCAGTCCGCGCAAGCCTA
COMP	A_23_P90436	GCTGCTGTGGAAGGACCCGCGAAACGTGGGTTGGAAGGACAAGAAGTCTATCGTTGGTT
CORO1A	A_23_P106761	TATCTCTCCATGTTTCAGTTCGAAGGAGTCCCAGCGGGGCATGGGCTACATGCCAAAACGT
COTL1	A_23_P83620	TTTCTTTGATCTCTTTTCTTTTCTCCCCCTCTTTTTTGTCTAAAGAAAAGTCATTTT
COTL1	A_24_P416131	GGATCGTTTTGTTTTGTTTTAAAGAAAGGTGAGATTGGCTTGGTCTTCATGAGCACAT
COTL1	A_32_P158385	AGAGCTATATCAAATGTGCTCATGAAGAACCAAGCCAATCTCACCTTTCTTAAAAACAA
COTL1	A_23_P3866	TCATGAGCACATTTGATATAGCTCTTTTTCTGTTTTCTTCTGCTCATTTCTGTTTTGGGA
COTL1	A_23_P83624	AGACCCTGGTGAAGGAGGTCGTACAGAATTCGCTAAGGAGTTTGATCAGTGATCGGA
CP	A_01_P014841	AACAATGTATGTGAAAGTGAATAAGAAATGTTACTTTAGAATGACTATAAACATTAATA
CP	A_23_P414793	TCACGGCCATAGCTTCCAATACAAGCACAGGGGAGTTTATAGTTCTGATGCTTTTGACAT
CRABP2	A_23_P405064	CCCAACTCTCTGGCAACTGGAATAATCATCCGATCGGAAAACCTCGAGGAATTGCTCAAA
CRABP2	A_01_P005286	TGAGGGCCTGAGCAGGAAAGACTGGCCCTCTGGCTTCTACACTTTGCTCCCTGAGCCTAT
CRABP2	A_23_P115064	CCCTACACCAACAAGAGGAATGGCTGCAAGAGCCAGATCACCCATTCCGGGTTCACTC
CREB3L1	A_01_P003175	CGTTCAGGAGAAAAGGCTCTACTTCCCAGCCTTCTTCCCTGACATCTGGACTCTT
CREB3L1	A_23_P150407	TTCGCCCCCTCCCTTGTTTTATATTTTATGAAGTTAGTGGGGCTTTGCTGCTCCCTGGCC
CRISP3	A_23_P419760	ATAACAACAACCTTTGGATTTTTATATATAAATTTGTGATTTAAATTTACTGAATTAAT
CSRP2	A_01_P010473	CTGTACTTGGATAGGCTGGCTAACTTGTAGGAAGAGAGCACTGTATCGTATCCTTTTGGCT
CSRP2	A_23_P44724	TGTAATGGATAGGCTGGCTAACTCGTAGGAAGAGAGCACTGTATGGTATCCTTTTGGCT
CTNNA2	A_24_P33477	CTTATCCAGGCAGCTAAAAACCTGATGAATGCTGTTGTCTCAGCGTAAAGCATCCTAT
CTNNA2	A_32_P14986	ATAGGTGTAGCTTGGAGTGTGGTATCTAATATACCATTGTATTTCACTAACTAACTCAAA
CTNNA2	A_23_P84736	GTGAAAAATCTGGAAGTGAATGGTAGAACATAAACTTGTATTGCTTCTGTTTCAGTGC
CTSW	A_24_P396167	GACTTCATCATGCTGCAGAACACGAGCACAGAATTGCGCAGTACCTGGCCACTTATGGC
CTSW	A_23_P13031	TGACCATCAACATGAAGCCCTTTCAGCTATACCGGAAAGGTGTGATCAAGGCCACACCCA
CTSW	A_23_P13027	TAGAGACCCTGTGGCGCATCAGTTTCTGGGATTTTGTGGACGCTCCGTGCATGAAGTGC
CXCL13	A_23_P121692	GGCGGGGCCGGGGGACTCTGGTATCTAATCTTTAATGATTCTATAAATCTAATGACA
CXCL13	A_01_P003023	AAATTTGCTGTCTAAGATTAATAGCATTCAAAGATCCCAGACTTCATAGAATACTCAGG
CXCL13	A_23_P121695	CAAGAGGCAAAGGAATCCATGTAGTAGATATCCTCTGCTTAAAAACTCACTACGGAGGAG
CXCL14	A_01_P019786	CGTACATGTCAGAAAACATTAGCATTGCATGCAGGTTTCATATTCTTTCTAAGATGGAAA
CXCL14	A_23_P213745	GTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAAGAATAGGGTAAAAACCTCA

Desarrollo de una Herramienta Molecular Diagnóstica de la Receptividad Endometrial

CYBRD1	A_24_P345451	CCCTCATTTTTGGATAGTACCAGACCGCAATGGAAACGCTCTAAGGAGCCAAATTCTA
CYBRD1	A_01_P005800	TTAGATAACTCCAGCAGAAAACTGTAAGTCTATGTCTTCAGGAAAAATGTAGAAGAAAGA
CYBRD1	A_23_P209564	CTTCCTGTGCTGCTTGTGTCAAATGGAACCTGCCCTCTAAAGCAGCTTTCTTTCTTTACT
CYP2J2	A_23_P103486	CACGTGTTCTGAAACCCTGGTGTCTGCTCAGATGTGTTGGGACAAAATGAAAGTGACTT
CYP3A5	A_24_P916916	ATGTGGCTCCCTCCTGCTGTCTACAGTACACACATGGAGTTTGTCTTTTTCTCTGACAG
CYP3A5	A_23_P8807	AGGATTTCTACTTTGGTCTTCAAGAAAGCTGTGCCCCAGAACACCAGAGATTTCAACTTA
CYP3A5	A_23_P8801	TTCAGAGATGGAACCCTAAGTGGAGAATGAGTTATTCTAAGGATTTCTACTTTGGTCTT
DDX52	A_32_P153423	AGTACACCTTGGAGTTTTTCGAAATATGGGTGGGTTTTTGGGCTCTGGTTGATTTAA
DDX52	A_23_P118660	GTTTATGCTAAAAATAACACCAAAATGTGGTGAAGTCTTAAGGACTTTCCCTCAAGTG
DDX52	A_24_P1929	TTCCTGAGGATGATAAGCATTATTAAGAAGCGTTGCTAATGTTATACAGCAGGCTGGG
DEFB1	A_01_P010027	ACTTTTTATAAGCATTCTTTAATAAAGGAAAATTGTTTTGAAGTATACCTCCTTTGGG
DEFB1	A_23_P71480	GCACCTGTTACAGAGGGAAGGCCAAGTGTGCAAGTGAGCTGAGAGTGACCAGAAGAAAT
DEPDC6	A_23_P60166	GGGGGTAGTTAGGGAGAGACTACATGAAATGGTGTGCCCTATTTCTTTCTGATCCTAA
DFNB31	A_01_P000551	CTCTGCACCCCGCAGCTATCCCAGAGGCTTGCAGGTGACCAGCAGTGTCTATTGATTTA
DFNB31	A_24_P376129	TGTGTAGTGTATGTTAATAACAGGCTAAGTGCACACAATAAAGAGCTTAAATACTGTG
DFNB31	A_23_P83351	AGATGGGGCCAACAAACCGCTGGACTTCTGGAGCCACGTCCTACTCTGGTCCGTGTGAA
DHRS3	A_23_P33759	AGGCAGGAAAACAGCCAGAAGCCACCTTGACACTTTTGAACATTCCAGTTCTGTAGAGT
DKK1	A_23_P24129	AATAACACCTTCCAAAACTGGAGTGAAGAGCTTTGTTCTTTATGGAACCTCCCTGT
DLG7	A_23_P88331	ATCCATTTACTCAGCTGGAGAGGAGACATCAAGAACATGCCAGACACATTTCTTTGGTG
DPP4	A_24_P97104	AGTCATGAGCAGAGCTGAAAAATTTAAACAAGTTGAGTACCTCCTTATTTCATGGAACAGC
DPP4	A_23_P39885	GTGCCAGTTATGGCTATAGGTGTACAAAAACACAGCAAGGGTGTGGGAAAGCATTGTA
DPP4	A_01_P008447	AAATCATTCCGCATCAGCTGTGAAACAACAAATAGGAATTGTTTTATGGAAGCTTTG
DUOX1	A_23_P54291	AGTCTAACACCACAATAATTTCAACCAAGTTTTAAGCACGTTCTTTTCATCAGACCCTG
DUOX1	A_24_P316586	AAAGCAAAGGAGTGAAGTCTAATGCTGTAATTTAGACTCACATGGTTGCGCACATGGA
ECM1	A_23_P160559	GATGAACAGGTCAACTGCTTCAACATCAATTATCTGAGGAACGTGGCTCTAGTGTCTGGA
ECM1	A_01_P019449	TGGAGGGTCAAAATGGGGGGAACCCACCTACCCACCCCTTGAACACTCATTACAGTA
ECM1	A_23_P149180	GAACAGAAGGAAGGAATGCCAGCTCCATTTGGGGACCAGAGCCATCCAGAACCTGAGTCC
EDN3	A_01_P007491	CATTGCGCAAACACTCAGAAAGTACTGCCAAAAGCCTAATAAAAAATCTAAAGTTTGCTC
EDN3	A_01_P017090	CATTGCGCAAACACTCAGAAAGTACTGCCAAAAGCCTAATAAAAAATCTAAAGTTTGCTC
EDN3	A_01_P013980	CATTGCGCAAACACTCAGAAAGTACTGCCAAAAGCCTAATAAAAAATCTAAAGTTTGCTC
EDN3	A_23_P17438	AAGTTATTACAGGTATAAAGTATGATGACCTATCATGAGGAAATGAAAGTGGCTGATTGC
EDNRB	A_01_P010992	CAGAAAGAAAGAGCAATAAATAAATTACATGCCATGTGGATTCTATTATAAATCAC
EDNRB	A_23_P2831	GCTTCACTGAATTCCTGCATTAACCAATTGCTCTGTATTTGGTGGGAAAAGATTCAAA
EDNRB	A_24_P330263	GTAAACCAAAACCCAACAATGTGGCCAGAAAGAAAGAGCAATAAATAAATTACACACAC
EDNRB	A_01_P015376	TGTGACAAGTGGACACTATTTATGTTAAATATACAATCATCAAGGAAGTATGAAGTATT
EFEMP1	A_23_P501007	AGGCAGCCATCATAACCATTGAATAGCATGCAAGGGTAAGAATGAGTTTTTAAGTCTTT
EFEMP1	A_01_P017336	GGAGGCAGCCATCATAACCATTGAATAGCATGCAAGGGTAAGAATGAGTTTTTAAGTCTTT
EFNA1	A_01_P016340	GGAACCACTCCCACCACAGGCACAAGCTGCACCTAGCAGCCTCAAACGGGTGAGTATT
EFNA1	A_23_P254512	TAAAGGGCAGGGCCACGTTATAGTATCTGTATATAAGTTGCTGTGTCTGCTGCTGAT

EFNA1	A_23_P113005	ATGGAGAAAGAAGTGGAGACAGTCCCTTCCACCATTCTGCCTTTAAGCCAAAGAAACA
EMCN	A_23_P382065	AGTTTTCCGGGCCAAGAATTTTATCCATGAAGACTTTCCTACTTTTCTCGGTGTTCTTA
ENPEP	A_23_P333605	TGTATGGATTAGCATCAGTGAAGAAGTACTCTTTTGTCAAGGTATTTGGATTGCTCA
ENPEP	A_23_P144596	TTGTGAATCTATTGTTTCTCCTCTGAAGCATTGGTGGCCTAATTTACAAGCAGATGGA
EPHB3	A_23_P259220	AGATTACACAACCTTACGACAGTGGTGGTGGCTGGATGCCATCAAGATGGGGCGGTA
EPHB3	A_23_P95060	AACTGGAGGAGGGACTCCAGGAATGGGAAATGTGACACCACCATCTGAAGCCAGCTT
EVC	A_23_P167051	CCAAACGAGCTGTTCAATAACAGCGATTTTGCCTAACATTATATCATAAGCGTGTCCAAT
EVC	A_24_P60441	TTGCTCCCTGATTAATAAGATGAGATATGGCTATTTGGAAGACACTGCATTTTAGCCAGTGA
FAM59A	A_24_P370042	TAAATAGGATAAGAGAACTCTTACTATGCAGATTACGTTTTTGAATGGTGAACAGGCT
FAM59A	A_23_P66948	AACACTGGAGGTGTGCCTTGTACGCTCAACAAACACTCAGCAGCTGCTAAAAGAAA
FGB	A_23_P136125	GAGTTGGTGTTCATAATTCAGTTCTAGTTGATTGCGAGAATTTTCAAATAAGGAAGAGG
FGB	A_23_P114011	GTGAATAGCAATATCCCACTAACCTTCGTGTGCTTCGTTCAATCCTGGAAAACCTGAGA
FLJ10719	A_01_P018037	TGAGCAAAGGAGCTTAATGCTAAGGTCAAAGGAGAGTGAAAGGTTGAGAACAATTGTC
FLJ20152	A_24_P266048	TGTGATCATCAGCAATAAAGATATAATAACTCTGTTTTCTTAGCCTGTATAGAGGAGAGG
FLJ20152	A_23_P167599	ACTGAAGACTTTGAACACTTGCTTTTTGTGATTGCTTATGTCATTAGTGCCTCATGACTG
FOSL2	A_01_P009941	ACTCAGCTCTCACAGGGTAATCATCTCAAGTGGTATTTGTAGCCAAGTGGGAGCTATTT
FOSL2	A_23_P218555	TCCCTCATTCTTGCAGCAAAATCCCATTTCTTGAAGCCTTGGAGAAGCTCGGTTTG
FOSL2	A_23_P348121	TGGGACTGGAAGTGACCTGTACAAGTATGCAGAAAGGAGGGTTCAAAGAAAAGGATT
FOSL2	A_23_P218553	TATTGGAAGACTTGGGTTGATCTCTTAGAAGCCATGGGACCTCCTCCCTCATTCTCTG
FOSL2	A_23_P345581	TGCTGTCACTCCGGGCACCTCGAACCTCGTCTTACCTATCCTAGCGCTCCTGGAGCAGGA
FXD2	A_24_P196562	TGACTATGAGACCGTTCGCAATGGGGCCTGATCTTCGCTGGACTGGCCTTCATCGTGGG
FXD2	A_23_P161769	CTTTAGACCTTTGCCCCGCTACTGCCAGCGCTTGGGCTGAAGGAAGCTCCAGACTCAAT
G0S2	A_23_P74609	CCAACACTGTGTAATTATCTAAATGCGTCTACCATTTTGCAGTGGGAGGAAGGATAAA
G0S2	A_01_P004760	GACGCGTCCACTACTGTGCAATTATCTAAATACGTCTACCATTTTGCCTAGGGAGGAA
GABARAPL1	A_23_P162640	GATCTCTTACCTTTGGAAAATAGGGGTTAGGCATGAAGGTGGTGTGATTAAGAAGATGG
GABARAPL1	A_24_P4816	GGATTGGCTTTGATAGAGGAATGGGGATGATGTAAGTTTACAGTATTCCTGGGGTTAAT
GABARAPL1	A_23_P65817	TGTGTGCAGTTCTGTGCTGCAAGTCTTGTGTAATGAAGTCAATGCTGTCAGGCCAA
GADD45A	A_23_P23221	CCTGCCTTAAGTCAACTTATTTGTTTTGCCGGGAAAGTCCGCTACATGGATCAATGGGTT
GALNT12	A_01_P005123	TTGTTCCACTCTTACGAGACTGCACCAACTCGGATCATCAGAAATGGTTCTTCAAGGAGC
GALNT12	A_23_P257731	GATTTGTGAATGATCCAGACCAACCTGAGATTTTGTCAACCTGATTAAGTCAATATGA
GALNT12	A_23_P415652	GAGAATCAGAAGTTCATCTTGCAGGAGGATGGATCTTTATTTACGAAACAGTCCAAGAAA
GALNT4	A_24_P942370	TCCCTATAACTCAAATAACTGTTTGTAAAAGAAAATTTGTTTACTTACCATTAGTAA
GALNT4	A_01_P010093	ACTTTTCTAGGATGAAGACAGCTTATTTTAAAGTGTATGGTCTTAGTTGGTTAGGGTC
GALNT4	A_23_P413576	GGTATTTGAAAACGTTTTTCTAATTTACATGTTCCAGAGGATAGACCAGGCTGGCATG
GALNT4	A_23_P116922	AAAGCTGGGGTTCATTTTGGTATATCACACTGAAACTGGGTACCAGAGTCTGCTGTTTT
GAS1	A_01_P005076	CTGGGGACAGATAGAGGGGATGGTGGGGATACTCCAAAACCTTTTCAAGTCAACTT
GAS1	A_23_P83134	GGATTAAGGGTCCAAAATGCTGATCTAAGGGGTTGCCATGGTGTGAACAATGCAACTT
GAST	A_23_P159190	ATGGATGGATGGACTTCGGCCCGCAGTCTGAGGATGAGAACTAACAATCCTAGAACC
GAST	A_23_P159191	AGCCTATGGATGGATGGACTTCGGCCCGCAGTCTGAGGATGAGAACTAACAATCCTA

Desarrollo de una Herramienta Molecular Diagn3stica de la Receptividad Endometrial

GBP2	A_01_P012911	GCTGAAGTGGTGAATATGATGAAAACTGATGAGACCAACAAGCCATGGGGCACAGT
GBP2	A_23_P85693	ATACAGGCCAAAGAGGTGCTGAAAAAATATTTGGAGTCCAAGGAGGATGTGGCTGATGCA
GDF15	A_23_P16523	TATTGGGGTGACCTCTTGGGGACTCGGGGCTGGTCTGATGGAAGTGTATTATTTA
GDF15	A_01_P016695	TTGGGGTGACCTCTCGGAACTGGGGGCTGGTCTGATGGAAGTGTATTATTTAAA
GLIPR1	A_24_P390096	AAGGGATAAACCTAAATATTTACTTGTATCATTAGAGAGGGAACATCAAATGCTGGGAC
GLIPR1	A_23_P364024	CAGCTCAAGTACCCTAATTTAGTCTTTTGGACTAATACAATTCAGGAAAGAAAAACCC
GNLY	A_23_P209954	GTGTGTAGGACGGGGAGGTACAGATGGCGCGACGTCTGCAGAAATTCATGAGGAGGTAT
GPR64	A_24_P932547	TGTGGAATTCATCCATCCAATTAATTCAGCTTATAATATTTCTGGTCCACATTTGATGC
GPR64	A_01_P016209	CCATAGTAATAATACACATTTCTGTGAGTGTGACTTGTCTTTGCAATATTTCTTCTCTG
GPR64	A_23_P253692	ACCCCTTTTAGTTAAGTCTTTCCTAAGTTCTCTTGCATATATTTCAAGTGAATGTTGG
GPR64	A_23_P253695	AGGGAAGTCTCTACACTGCTATTGTTGCTACATGTATCGAGCCTTGATTGCTCCTAGTT
GPRC5C	A_32_P109029	ATGAGACCATCTGAAAGAGCAGAAGGGTCAGAGCATGTTCTGGGAGAACAAGGCCCTTTT
GPRC5C	A_32_P544510	CCTGAAAGAGCAGAAGGGTCAGAGCATGTTCTGGGAGAACAAGGCCCTTTTCCATGGATGA
GPRC5C	A_23_P38167	TCCAGCCAAATAGTGTCTCGGGGTGGTGGCTGGGCAGCGCCTATGTTTCTCTGGAGATT
GPRC5C	A_23_P346673	GCAGGAACAGGTGTCATGACCCTGATTACCTGGATCTCACAGTTCTGGGAATGCAATA
GPRC5C	A_23_P346670	GGGCTGCTGTGGCTATTCCAACATCTCTGGTCAGGGCCTGGAAGTACCTTAGGACACAT
GPRC5C	A_01_P006986	AGGCCTGGATCTTGTCTCTCTGTGAGGAACAAGGGTGCCTAATAAAAAACATTTCTGCTTT
GPX3	A_23_P133474	TGGGTGTGGGTGCATGTGGGTGTTACACACATGCCTACAGGTATGCGTGATTGTGTGTG
GPX3	A_23_P133475	CACTCAAGGCCCCAGCCTGGCACAATGGATGCATACAGTTCTGTGTACTGCCAGGCATG
GREM2	A_23_P329821	TCGTTGGGAAATTGATGACCGGAACTAGAAAGCCCTATAATGTGGAATATGCATATGT
GREM2	A_23_P329822	CTGACGTTACTAAGCATTCGAGCACAATGTAGAAATGGCTTGGGATGGATAGGTATAGG
GREM2	A_24_P40626	GCAAACACGAGCACACTCTTTCGAACCCAATTGTTGGGTGTAGCAATGAAAGCAATATGA
GREM2	A_23_P97181	AAGTCCACTGTTAGTCTCTCCGTTAGCACCAGGGACACACTTGTCTGAGTTTGTTCAT
GREM2	A_01_P017852	CTCCTTCTGGGGTATCGTATGTACAAAGTTACCTTATAATGGCTCAAATTGTATTTA
GZMA	A_23_P133445	GAATGAATATGGTTTGTGCTGGAAAGCCTCCGAGGTGGAAAAAAGCTCGTCAATGGAGATT
HABP2	A_01_P014820	AAATTAGGCGCCTGTTTGTAGCTGCATTTACACTTCTTTAGAGCTAGCTGACCTTTGGC
HABP2	A_23_P1173	AACCCAAGGAATGATGGAATCAACACAACATAGTATGTTTCTTCCCTTACCAATTGTA
HAL	A_23_P61643	TTCTGCACAAGAAATCCACCAAAATCCCGGAGTCTGAGGACCTTAATGGGCTTTGTCAT
HAL	A_23_P61637	AGATCAATCCATTGTATCATTAGTCTTCTAAAGCCTACGTTGGTTAGGCTGATGGCAG
HAL	A_01_P003831	CCAGGTATTTTCAGATTGTAGGAGTTTTCTTCTTAACAATTTCAACAGGCCACTCACTC
HAL	A_32_P502420	GCAACAGATAATCCTATGGTCTTTGCCAATAGGGGAGAGACAATTTCTGGAGGAAACTTC
HAND2	A_23_P7154	TTGGGGTCCGATCTAGTGTATCCATTATCTGTGGTCCCTCTTTGAAGATGTTT
HAND2	A_23_P373521	TCTGAGGACTCCTTGCAATTTGGAATCATCCGGTTTATTTATGTGCAATTTCCCTCCCTC
HEY1	A_23_P157659	ATAGGTTTGTACTGTGCCTAATTTACTTTGTAACCAGAATGATCCGTTTTTCCCTCAC
HEY1	A_01_P001077	AAACGACCTCCACTGAACTGGGTTGACCTCTGTTGTACTGATGTGCTGACTAAATAA
HEY1	A_32_P83845	GGGCTATAAAATATCATTTTTAGGTTTATTTCTTTAGCAGGTGTAGTTAAACGACCTCC
HEY2	A_01_P012352	TAAAAGATAAGAGACACAACATGATTATGCACTTCATTTCTACTGTGTGGAGAAAGC
HEY2	A_23_P168351	AAAGGAGAGGGATTATAGAGAAAAGGCCTCGGGATCGGATAAAATAACAGTTTATCTGAGT
HEY2	A_24_P363408	AAGCTCTTTGTCAAGTTAGTATTGATTTGATCCAAAACAAGATGAATGTATGCAATG

HEY2	A_23_P168354	TACTCGGGGCAAAGTACTAGCTCTGTGATTAGATTGAATTCCTCAACAACAACATCTCAG
HLA-DOB	A_32_P477383	TTCCAGTCCCAATCATTACTTTTCTGTGGTCCAGCCCTACTCTATAAGTCATGATCT
HLA-DOB	A_24_P911788	GGTAGGGAGGCAGGAAGAGGGAAACATTGTGTCTGTTTAGGATCCTTATTGTGTGATC
HLA-DOB	A_23_P30736	CCAAATCATTACTTTTCTGTGGTCCAGCCCTACTCTATAAGTCATGATCTCCAAAGCT
HLA-DOB	A_32_P343332	CCAAATCATTACTTTTCTGTGGTCCAGCCCTACTCTATAAGTCATGATCTCCAAAGCT
HLA-DOB	A_01_P015398	TGTCTCCAAGTGCAGTCTCCAGCTTTCAGAAGACAAATGCTCAGGTAGTCACTGTTT
HMHA1	A_23_P374053	GGTCTCTCGGAAGCCACCGTGTGGTCTTTTCACAGGCACGTTTATTTTGTGAAATAAAA
HPGD	A_24_P71904	AGACAAAGTAGTGAACATCAATGAACATCTGATAGAGATAAACTGTAATCAGGCATAAGC
HPGD	A_23_P213050	CATCATTATCAATATGTCATCTTAGCAGGACTCATGCCGTTGCACAGCAGCGGTTTA
HPGD	A_01_P012816	CCTCTGCTCACTTTAAGAACTTAACTGACTCCAAAAATCTCAGGAATTAAGTGTAAAC
HPSE	A_23_P256107	AGTGGATGAAAACCTCGATCCTTACCTGATTATTGGCTATCTCTTGTTCAGAAATTT
HRASLS3	A_23_P116414	CATCAAGGGGGTCTGTTTGTCTAGAGAGTTGGGGTTGGTTTGTGGATTCATTGTGA
HSD11B2	A_23_P14986	TGCAGATAGTGCCCTCTGCAAACTAAGGAGTGACTAGGTGGGTTGGGGACCCCTCAGGAT
ID4	A_23_P59375	AAAGCCACCGGAGGAAAGGAAAAAACATCGGGCCAACCTAGAAAAGTTTTTCATTCGTCATT
IDH1	A_32_P45009	GTTCTGGTGCATAGATGTCCATTTTGTGAGGTAGAGCTGTGCATAAACTTGCACATG
IDH1	A_23_P5376	GCACAGGGAAGTTCTGGTGCATAGATATCCCGTTTTGTGAGGTAGAGCTGTGCATAAA
IDH1	A_01_P006229	GCACAGGGAAGTTCTGGTGCATAGATATCCCGTTTTGTGAGGTAGAGCTGTGCATAAA
IER3	A_01_P015343	CGACTGTCGAGATTGCCAGTATGTTCTGTGAACACAAATAAACTTGATTTACTGTCTGC
IER3	A_23_P42257	TGTGAGATGTTCCCGCTGTGAAATGCAGGCTCTTGGTATTTATTGAGCTTTGTGGGA
IER3	A_01_P000583	CGACTGTCGAGATTGCCAGTATGTTCTGTGAACACAAATAAACTTGATTTACTGTCTGC
IER3	A_32_P419552	GTACGTAATATTTATTTAACTTATGAAGGGTGTGAGATGTTCCCGCTGTGAAATGC
IGF2	A_23_P203458	GCCACAACAACCTCTTAAACTAATTGGCTTTTTAGAAACACCCACAAAAGCTCAGAA
IGF2	A_23_P421379	TGCTTCCGGACAACCTCCCGAGATACCCCGTGGGCAAGTTCTTCCAATATGACACCTGGA
IGF2	A_23_P150609	CTCAACTCAGCTCCTTAAACGCTAATATTTCCGGCAAAATCCCATGCTTGGGTTTTGTCT
IGFBP1	A_23_P42869	CATGAAACACTTCTCATCATATTGTATGTAAGTAATTGCATTCTGCTCTTCCAAAGCTC
IGFBP1	A_23_P42868	TTTCTGCTCTTCCAAAGCTCCTGCGTCTGTTTTTAAAGAGCATGGAAAAATACTGCCTAG
IGFBP1	A_01_P002590	AATCAAAGCTACCTGTGGTGTGTTGCCACCGGTTAAAATGTACTGATATGTTGTTA
IL15	A_23_P29953	CTGTGTGGAACCACTGACTACTGGCTCTCATTGACTTCCTTACTAAGCATAGCAAACAGA
IL2RB	A_24_P203000	TTGAGGTTGTCTGAGTCTTGGGCTATGCCTTGAAAAAAGCTGAATTATTGGACAGTCTC
IL2RB	A_01_P020028	CCACAAAGTGGACGCCTGCTGTATCTTCCAACAGTGGCTTACAGACCCACAAGAGAAG
IL2RB	A_23_P255430	CTCCCGGATGCCTTGGAGATAGAGCCCTGCCAGGTGACTTTACTTACGACCCCTACTCA
IMPA2	A_01_P011439	TCCCCCAGAATAATAATCTCAGGTAATAAGGCTTTAGAACTGCTGATAAAGCGGATCGT
IMPA2	A_23_P50081	TAGCTGTTCTCTCTTAACTCAGTACGCTTTTTCAGGTTAGTACGTTCTTCTGTC
INDO	A_23_P112026	ATCTGCAAAATCGTACTAAGTACATCCTGATTCTGCAAGCCAGCAGCCAAAGGAGAATA
KCNG1	A_01_P006886	ACCCCGAGTTCCGGGCTCCTGCTCTTCTCTGCGTGGCCATCGCCCTCTCGCGCCC
KCNG1	A_23_P210581	CCCTATGTGTGTTCCCTCAATAAGGAGATGCCTTGTCTTTTCCACATGCAAATAACAT
KCNG1	A_24_P237117	CCTGTCCGTTGCTCTTCCGTGACCGTCAACCTCTCCGTCAGCACCTTGGCCAG
KCNG1	A_23_P210580	CCAGCCCTATGTGTTTCCCTCAATAAGGAGATGCCTTGTCTTTTCCACATGCAAAT
KCNG1	A_01_P002764	CGCCGATGCACCCGAGTTCCGGCTCCTGCTGCTCTTCTCTGCGTGGCCATCGCCCT

Desarrollo de una Herramienta Molecular Diagn3stica de la Receptividad Endometrial

KCNJ2	A_23_P329261	AAGTTGAACAATCCTAGCCATTGACAATCGTGATAGTTATTATTTCCCATTTGCTGTCT
KCNJ2	A_23_P55219	TCGTTTTGTGTTTTCCCAAACCTGAACTGCAGGCAAGCCTTGTTGGGTATTTGATT
KCNK7	A_23_P86874	CCTGCGCCATTGCCTGCTGCTCAGCCGCCACGTCCTGGGTAGCGGTCCACTG
KCNK7	A_01_P003027	GCTTGCTGAAGCGTCAGGTGACCGAGTTCAGCTCCCATAGGTGGCGGCACCTAAGGAGG
KHDRBS3	A_01_P016715	CTTTTGAACCTCAAGTCTTGAATGTTCTGTAGTGTAAAGCAAAGTCTCCTCTTGCTTGA
KHDRBS3	A_23_P257335	ATGATGAACAGAGTTATGATTCTATGATAACAGCTATAGCACCCAGCCCAAAGTGGTG
KIAA0802	A_01_P004436	AAACACACAGCTGCTGAATGTTCAACCTGTGAAACTGAGATGTTTCTAGAATGAAACAGT
KIAA0802	A_32_P38763	AGGATACCAAATGAAACACATGATGATGCCTCTGGGTCTGTATGAGACCGTATGAAGT
KIAA0802	A_23_P360605	CTCAAAGTTTTCTAAGGGAAACACTACAAAAGTCACAAGGATACCAAATGGAAACAC
KIAA0888	A_23_P423074	AAGGGCTTCAATCATATTGGAGGCTTCAATCATCCTTTGATTGTACTGAGTACTGGTTG
KIAA0888	A_32_P84242	AGAAAGGCTTACCTTCTGTCATCAAGTGATTGTATCATCCTGGATCGTCAATTCGAAGGA
KIAA0888	A_32_P84241	TGGATCGTCATTTCAAGGAAGTACGCTTTCTTTTCTAAGCGTCTGTATGTGTTCTAAA
KIAA0888	A_01_P007808	CACATAGTCTTTGTGAACTGTTTGTGAAGGAAGTTCACCTTTTGTGTACATACGTTGA
KIAA0888	A_01_P007810	TTGCAAGGCATTAGAAAAAATTCACAATTACAGGGGACTGAAAATGTGATTTCAACCCAG
KIF11	A_24_P227091	ACACTCCAAGACCTGTGCCTTTTAGAGAAGCTCACAATGATTTAAGGACTGTTTGAACCT
KIF11	A_23_P52278	AACACACTGGAGAGGTCTAAAGTGAAGAACTACAGAGCACTTGGTTACAAAGAGCAGA
KIF20A	A_23_P256956	TCAAGCCTTGACCCTTGTGATGACATCTTAATCAACAGGACCAGACTCTGGCTGAACCT
KIF4A	A_23_P148475	GCTTGGGAGATGCTTTCAGGTTGCAGCCAGAAGGGGTTTTTAAATGACTTCTCTGGATT
KLRC1	A_23_P151046	GGTTGAAATAGGAGATGACCTCTAAGTATGATAGAACGTTACTTTGTGTCGTGATGAAACT
KLRC1	A_23_P139654	ATTGTTGGGATCCTGGGAATTATCTGTCTTATCTTAATGGCCTCTGTGGTAACGATAGTT
KMO	A_24_P387926	AGTGGAAATTAAGTAAAAGTTGGGCACTAATCTGGATTAACATTGAGGAAATCAGTTG
KMO	A_23_P200838	ATAGGACAGATCCCATCTCCTCCACCAATACATTATTAGACTGAACTGTGACCTGAAAT
KMO	A_24_P77082	TCATGCGATTATGCCATCGACCTTTATCCCTCTCTATACAATGGTCACCTTTTCCGAAT
KNTC2	A_23_P50108	AAAGTGGGAAATAACTTGCAACGCTGTTAGAGATGGTTGCTACACATGTTGGTCTGTA
KNTC2	A_01_P010921	AGAATGAAGATAAAATGTTGATCATGTATATATATCCATAGTGAATAAAATGTCTCAGT
KNTC2	A_24_P14156	AGTGTATTCGACAACCTCTGTGAGTTTCTTACAGAAAATGGTTATGCACATAATGTGCCA
KRT7	A_23_P313591	GAAGATGAAATTAACCGCCGCACAGCTGCTGAGAATGAGTTTGGGTGCTGAAGAAGGAT
KRT7	A_23_P381945	GAAGATGAAATTAACCGCCGCACAGCTGCTGAGAATGAGTTTGGGTGCTGAAGAAGGAT
LAMB3	A_23_P86012	GCCTTAGTTCTCCACTGGGGAGGAATCCTGGACCAAGCACAAAACTTAACAAAAGTGA
LEPREL1	A_01_P018078	ATGGCCAGGTTGCCTACGAGTGGGGTCCACTGATGAAAAGAGGTTTTTGTACTTACATA
LEPREL1	A_23_P69179	CCAGGTTGCCTAGGAGTGGGGTCCACTGATGAAAAGAGGTTTTTGTACTTACATAAGAA
LIF	A_24_P122137	CATTTCCCTGCAGATGGTACAGATGTTCTGCTTAGAGTCACTCTAGTTCCCCACCTC
LIF	A_24_P233488	GTCTTCCAGAAGAAGAAGCTGGGCTGTCAACTCCTGGGGAAGTATAAGCAGATCATCGCC
LIF	A_23_P155017	TCTGGCACCCGGGAGGTAGCATTCCCTGCAGATGGTACAGATGTTCTGCTTAGAGT
LMCD1	A_01_P019499	AGGAGCAGGCTGGGATCCCAACTATCGCTTGTGCCTTTTTCAAGTGGAAATTTGAATTT
LMCD1	A_23_P6771	AGGAGCAGGCTGGGATCCCAACTATCGCTTGTGCCTTTTTCAAGTGGAAATTTGAATTT
LMOD1	A_01_P000575	ACACACTGGCTCCTAGACCTAAAGGGTATGAGCTGGAGCTAAGGCCAGCTAGAGCTTCCA
LMOD1	A_23_P201940	ACTTGAAGTTTGTGATGCAAAAGCAGACATTTTCTCACACCCACCTGCTGCTGATGAAT
LRFN4	A_23_P63980	GGTGTGATGGACGGGCAGCTTCTGTGTGCTCCAAGGATGAGCCTCGTGGGCAGAGGG

LRFN4	A_01_P007807	CTTTCCTCGGTTCTGGCCTCCAGACCAGAGTAAGGGGCAGGTCCCTCCAACAGGTGCTCA
LRP4	A_24_P403561	AGCACTGATTTACTCTGTAAAAAGCAAAATCTCTGTGCTAAACTAATGGAAGCGATT
LRP4	A_23_P159349	TCTGGCTCCCTGGATGACACAGAGACGGAGCAGCTGTACAGGAAGAGCAGTCTGAGTGT
LRRC1	A_23_P215024	TTTTTCTGAATAAGTCTCTCATAATGAGTGCAGTGCAGACTGTGCCTACTCTGATGGT
LRRC1	A_01_P020198	TTAAGGGATATTGGGAAAAGTTTTGGTGTGTTTCTGTTGACTCTTTTTGTATGCTGTG
LRRC1	A_01_P016431	AGGAAGCTTATAGTGGGTCAAACAAGGAGGTGTTAGTGTGTTTAAAAAGAAGGCTC
LRRC17	A_23_P253958	GCACCACTACAATGTCATTTTAATGGCCTGGAATGCAAACGCCTGAAGAATACAAAGG
MAOA	A_23_P83857	CTGCTGAAGATCATTGGATTTTCCACATCAGTAACCTGCCCTGGGGTTTGTGCTGTACAAA
MAOA	A_23_P96410	CCCCAGTGTACACGTTGGAGTATACCTATGTGTGTGCTTTGCCACTGAAGTAAGATTTT
MAP2K6	A_01_P016743	TTTTCTTTCTGGCCTCAAGTTCAATATGGAGAGGGATTGCTTCCCTGAATCCTCTCTTC
MAP2K6	A_23_P207445	ACAGCATCAATAGAAAAGTCATCTTTGAGATAATTTAACCCCTGCCTCAGAGGGTTTTCT
MFAP2	A_23_P1021	GTCTGCTGGATTTCCTACAGCCCCGTGGGCATGGACCACCTTTATTTTATACAAAATT
MFAP2	A_23_P1029	CCGCCTCTACTCCATACACAGGCCCTTGAACACAGTGTCTCAACGAGGTCTGCTTCTACAG
MFAP5	A_23_P87700	TGAGAGGAAGGAGGATCCTCTTCTTCCAACCCATTGACAGCTAACCCCTTAGACAGTATT
MFAP5	A_01_P003020	TTGACCCTCTACTCCACATTGCAACATTTGCATCAGACAGCATTCAATCCAGTATTA
MGC11242	A_01_P013777	AGATGACTGTTTCTCATGCCTTTATCTTCTTCTATGTAAGTAAAGTGGACCTTTGTGCTC
MMP26	A_23_P2052	TATTC AACCTGTCCTTT CAGGGAGTTT ATTGGAGGATCAAAGAACTGAAAGCACTAGAGC
MRPS2	A_23_P135104	TTCTCAATCTAAATGCCTTT CAGGTGGGCCGCTTCTTGGCTACCTGGTCCAGGGGGCT
MSX1	A_23_P110430	AAGTGGCTGGAAGAGTCCCTTAGTACTCTTCTAGCATTTAGATCTACACTCTCGAGTTAA
MSX1	A_01_P001372	AACATTTGCTCTGGGGGGGCAGGGAATACACAGATGCGTTGCAAAGGTAGGTGAAGGGA
MSX1	A_24_P345837	ACCGCCGCGCCAAGGCAAGAGACTACAAGAGGCAGAGCTGGAGAAGCTGAAGATGGCCG
MT1G	A_23_P60933	AGCTGCTGTGCCTGATGTGCGGGACAGCCCTGCTCCCAAGTACAAATAGAGTGACCCGTAA
MT1G	A_23_P206707	TTGGGAACCTAGTCTCGCCTCGGGTGC AATGGACCCCACTGCTCCTGTGCCGCTGGT
MT1G	A_23_P206701	GCATCGGAGAAGTGCAGTGTGCTGCGCCTGATGTGCGGACAGCCCTGCTCCCAAGTACAAA
MT1H	A_23_P414343	TGTGCCAAGTGTGCCACGGCTGCATCTGCAAAGGGACGTCGGAGAAGTGCAGTGTGCTGT
MT2A	A_23_P106844	CAACCCTGACCGTGACCGTTTGCTATATTCCTTTTTCTATGAAATAATGTGAATGATAAT
MT2A	A_24_P361896	CGTCCAGATGTAAAGAACGGACTTCCACAAACCTGGATTTTTATGTACAACCTGA
MUC16	A_32_P306874	CACAATGATTCCAACCTCAACTCTTCCCTTGGTTACTAGAGACTACAGGCTTACTGGC
MUC16	A_32_P889536	GAGCACACGCTAGTTCAGAAAAGTCCAAGCACCATCAAACCTACCATGGCTTCAAGTAGTA
MUC16	A_32_P307960	TGGTGTCTTTATCTGCAGTAACCTCACCTTCGCCACTTTATTCCACACCATCTGAGAGTA
MUC16	A_32_P393950	CCAAGACCTCCTTGGTTGAAACAACCTGATGGAACGCTAGTGACCACCATAAAGATGTCAA
MUC16	A_32_P309031	CAGAATCAGCAGAAATGATGATCAAGACACAACAGATCCTCCTGGGTCTACACCAGAGA
MUC16	A_32_P424560	CAGGGATCCACTCAGCTATGACTCATGGATTTTACAATTGGATGTGACCCTTATGA
MUC16	A_32_P378729	TATCCCTTTTTCCATGACACTAAGCAATGCAGAAACAAGTCCGAAAGGGTCAGAAGCAC
MUC16	A_23_P5211	AGCATCAAGATTATTTTTCTGACTGTCAAGTTTCAACATTCAGTCTGTCCCCAACAGG
MYL9	A_23_P210428	TGAGTCTGGGATCAGACACCCCTTACGTGTATCCCACACAAATGCAAGCTACCAAG
MYL9	A_23_P210425	TGAGGAAGTGGACGAGATGTACCGGGAGGCACCCATTGATAAAGAAAGGCAACTCAACTA
MYL9	A_01_P014392	TGTGCTCAGGAGTTGCGGGCAGCATGGACATCTGTCCAGAGGAGGCAGAAATCTCCAATA
NDRG2	A_24_P244800	TTTTCAGCCAGGAAGACTCTCTGAAATCTGAGTTGATACAAAAGTACAGAAATATCA

Desarrollo de una Herramienta Molecular Diagn3stica de la Receptividad Endometrial

NDRG2	A_23_P37205	CGTTTGGCTGCACTAACTTTGGTAGCTCAGTGTGCATCTAGAGTGGGACTGGGGAGGGAG
NKG7	A_23_P119042	GTTTTGAGCGTTGTATTCCAAGGCCTCATCTGGAGCCTCGGAAAGTCTGGTCCCACAT
NNMT	A_01_P015240	AGGAAGCTGAGCAGATCCCTGTGATGCCTGTGACCTCAATTAAGCAATTCCTTTGACCT
NNMT	A_23_P127584	CTGCTGTGAAAGAGGCTGGCTACACAATCGAATGGTTTGAGGTGATCTCGCAAAGTTATT
NP	A_23_P140256	CTACTAGCTCTTTGAGATAAATACATTCCGAGGGGCTCAGTTCTGCCTTATCTAAATCACC
NR4A2	A_01_P012780	AGCAGCTTTTGACTGTTCCAGAGTGCCTATAATATACATAACTCCCTGGAAATTAAGTGA
NR4A2	A_23_P131207	CTAAATGTTGCGTGGTGGCATGAGTTGAAGAAGGCAAAGGCTTGTAAATTTACCCAATG
NR4A2	A_23_P131208	AAGTTTTCTGCTGTAAGAAAGCTGTAATATATAGTAAACTAAATGTTGCGTGGGTGGC
NRG2	A_23_P213695	TATAACAAAATGTTTTATTTTCATTTTAGCAAAAATTGCTTATAACTAGCTAACGGC
NRG2	A_01_P018499	GCTAACGGCAAAGACGTTTTTATAGGGAAACTATTTATATGTAACATCCTGATTTACAGC
NRG2	A_01_P009356	CATCTGCGTGGCTCTGCTGGTCTGGGCATCGTCTGTGTGGTGGCTACTGCAAGACCAA
NRG2	A_23_P349857	GGACAGAGATGTTGGAGAACTGCCTTTGCGATTGTACATGCCAGATCCTAAGCAAAGT
NRG2	A_23_P213699	CTAGCTAACGGCAAAGGCGTTTTTATAGGGAAACTATTTATATGTAACATCCTGATTTAC
OFD1	A_24_P134653	CTGCAATCTAGTGACAAAGTCGAAAGTTTAAACAGGCTTTTCTCATGAAGAAGTACAGCAGC
OFD1	A_23_P254226	TAGTCCCTGATTCTGACCTTGAGTTGTAGCCAATACTAAGGCAAGGGTCAAAGAGCT
OFD1	A_01_P014581	TTTTCTCACGAAGAACCAGATGACTCTTGGTAACCATGTTTGTGCTCCAGCTTCTAACTT
OLFM1	A_01_P010547	AAGGGCCCTCCAAGCCTTAATGGCACCCCTAAGCCTCCATGCCAGGCCAAAAGATGCTT
OLFM1	A_23_P219161	GGGCTAACTTAAAGAGTTTTTCAATGCTGCAGTGAAGAAAGCAGTCCACTCCCAT
OLFM1	A_23_P304311	AGGGCTATGCCTGTGCTTATTGAGACACCTTGGCAAAGAGATGGCTGATTCTGGGTGGT
OLFM1	A_24_P406601	TCAAAGGAATTACTCTCTTCTGTTAAATTAGCTAAATCATGTAACCCGAGATAGGAAGG
OLFM4	A_24_P181254	TTTTCTTTGATGTCAAGTCCTAGTCTATAGGATTGGCAGTTAAATGCTTACTCCC
OLFM4	A_23_P2789	TTAGATATCTGCAGGGGTGCTAAAAGTGTGTTTATTTGTCAGCAATGTTTAGGTGCATA
OPRK1	A_23_P169061	TTGAAGGCAAAGATCATCAATATCTGCATCTGGCTGCTGCTCATCTGTTGGCATCTCT
OPRK1	A_32_P33576	CAGTCTTAAACGTGTATAACTATTGTCAGACAATTTATAGGTGTTTCATCTAGTCTGG
OPRK1	A_32_P33572	GAAAATCAGATCCAAAACGGTATAACCCAGCTTCTTAAGGCAATTTTCTCTCTGAAA
OPRK1	A_01_P003195	TAGTCGTGGAGATGCTTCTGTCAGTCTTTCAGGAAGAGAGGAGTTCATGATCTGGGTT
PAEP	A_23_P257129	TACGGTGGCGAACGAGGCCAGCTGCTCGATACTGACTACGACAATTTCTGTTTCTCTG
PAQR4	A_23_P397334	CAAGGCCCTCTGTCTTTTTCAGAACACATGGACTTGGAGGCAGATTTGAAATAAACTT
PAQR4	A_23_P397341	GCCATGTTGCCACATGAGCAAGCTTGGGTGCTCCCAAGGTTCAAATACTTTTATTAGAC
PAQR4	A_23_P66213	TTCCCTCCAAGCTCTATTTTACTGTGTCAGCTGGAAGGAAACCTTTCCCTCTGGGAC
PAQR4	A_23_P66211	TTACTGTGTCAGCTGGAAGGAAACCTTTCCCTCTGGGACCTTTTACCCTCTGTGACCT
PBK	A_32_P62997	AGCTCTGGAACAGATGTCTAGTGATCATCTCAGCTGAAGTGTGGCTGCGTAAATAACT
PBK	A_23_P82699	ATTGGCTATAAGCACTTGAATGTACTGGGTTTTCTGTAAGTTTTAGAACTAGCTAC
PECI	A_23_P156852	CTTCAGGGAAGATGGCTATCAGATGAATGCACAAATGCTGTGGTGAACCTTCTATCCAGA
PENK	A_23_P254584	TTCATTGTCTGGATAACTATACAACCTGAAAAGTGCATTTTCAGGTTCTGTGCTTTTTT
PENK	A_23_P417918	TTTGCTAGCCAAAAGGTATGGGGCTTCATGAAAAGGTATGGAGGCTTCATGAAGAAAAT
PENK	A_01_P015221	CATTTTCAGGTTCTGTGCTCTTTTGGAGTCTTAAGCTCAGTATTGGTCTGTTCAGCTA
PLA1A	A_23_P29816	GAGAAAAGCAAAGCTCTTTCTTATTTTCTCATAATCAGCTACCCTGGAGGGGAGGGAGA
PLA1A	A_24_P294408	TTACCTGAACCAGTGAACCTACAAGCAAGTGTGACTGTTTCTGTGACCTGAAGATAGCC

POLD4	A_23_P127367	CTGCTGCCTGTCTCCCTGACCCATGATCTGGCAAGTTAGGCACAGTCAGACATGGACAGT
POLD4	A_01_P004358	GGAGAGTGGTCACTGGGGAAAAGGACCTGGCCATCACCTTCCAGTACCTGCTGCCTGTCT
POLD4	A_23_P360215	CTATGAGGCACCACGTAAGACCTCTGCCCTTAGCTCTTGTCTCACCACCAAGAACCT
POSTN	A_24_P347411	AGAATCTGACATCATGACAACAAATGGTGAATTCATGTTGTAGATAAACTCCTCTATCC
POSTN	A_23_P205111	AGGAAGTTGCAAGCCAACAAAAAGTTCAAGGATCTAGAAGACGATTAAGGGAAGGTCGT
PPARGC1A	A_23_P18443	TTGAGCCCTTGCCGGGCCITTTTTCCACCTGCCAATTCTACATGTATTGTTGGTTTTA
PPARGC1A	A_24_P303052	TCTCCAGGCAGTAGATCCTCTTCAAGATCCTGCTATTACTATGAGTCAAGCCACTACAGA
PPARGC1A	A_23_P18447	AATCACTGTAGTCTAAGACCTGATCTATAGATGACCTAGAATAGCCATGTACTATAATGT
PPARGC1A	A_01_P016440	TTTTTAAATGTGCAGTACACATCAGCCTCACTGAGCTAATAAAGGGAACGAATGTTTC
PRC1	A_23_P206059	TCTCAATCACTACTCTTCTTGAAGCACTATTATTTATTCTCCGCTGTCTGCCTGCAGCA
PRKCQ	A_23_P1374	ACTGATCAACAGCATGGACCAGAATATGTTTCAGGAACITTTCTTCTATGAACCCGGGAT
PROM1	A_23_P258463	TGGGGTGTGTTGCCATTGGATGCATTCTATCAAACTCTATCAAATGTGATGGCTAG
PROM1	A_01_P010711	TCAAATGTGATGGCTAGATTCTAACATATTGCCATGTGTGGAGTGTGCTGAACACACACC
PROM1	A_23_P258462	TCAAATGTGATGGCTAGATTCTAACATATTGCCATGTGTGGAGTGTGCTGAACACACACC
PROS1	A_24_P119141	AGCAGCACAGGAATCTTACTTCTTGGCAGTGCAGTCTGTCAAGATGAGACATCAGATTA
PROS1	A_23_P73114	CCAGAACAATTTTAAACAAAAGGACAACCACAGAGGGATATAGTGAATATCGTATCATTG
PROS1	A_32_P500684	TATACGAAGCTGGAATTTGATGAAGCAAGGAGCTTCTGGAATAAAGGAAATTATTCAAGA
PROS1	A_23_P84510	TGGTACTGTGATGCATTCAAGTGGCAGTTTTATCACGTTGAATCTACCATTCTATAGCC
PROS1	A_24_P383480	ATTTGATTTCCGGACATATGATTGAGAAGGCGTGATACTGTACGCAGAATCTATCGATCA
PSMB10	A_23_P140805	CAGTGAAGCCACTAACCCCTGGAGCTAGTGGAGGAACTGTGCAGGCTATGGAGGTGGAGT
PSMB10	A_23_P140807	ACACTGAGCTCACCCACAGAGCCCGTGAAGAGGTCTGGCCGCTACCCTTTGTGCCTGGA
PTGER2	A_23_P151710	CCAGCTGCCTATTGATTTAAGCTTCTCTGTTGAATGACAAAGTATGGTTTTGTAATTT
PTGER2	A_01_P012436	GAACATGGTTGACTCATCTTATATGGGAAACCATGTAGCAGTGAGTCATATCTTAATAT
PTPRR	A_23_P47924	TCTTGGAATGGACTGTCAGCTGTTAACTGTTCTGTTTTGAAGTGCTATTACCTTTCT
PTPRR	A_32_P770346	TGATCATTTTTGGCAATTAATCAGAAGAAGAGTGGGAAGCCGGTATTTCATTTATAAGCA
PTPRR	A_32_P461885	AGACCAAAAATGTAACCGATTCAATTGAGCACCTACATTAATGCTAATTATATTAGGGGC
PTPRR	A_32_P480310	GTTTTAACTACTTGTCTCTTTTTGCTAAGAAGGGATTTTTGAATATGCTATCTACCTGG
RAD54B	A_01_P017426	ACTCAAGCTACTGGCACATAATGAAAGATTACTTCATGACATTCCATTGCTCTCTTTTG
RAD54B	A_23_P82738	CAAGCTACTGGCACATAGTGAAGATTACTTCTGACATTCCATTGCTCTCTTTTGAAAA
RAD54B	A_23_P94141	AAATTTGGATTGTATGTTTCAGGAGAAGAGGGATGGATTGAAAAGAAGGCAGCAGCTAGA
RANBP17	A_23_P58819	TCTGGGTCTCAGGACAGTGATGTTGGCTAGCCAGGGGAATGATTTTTCAAACATACA
RANBP17	A_01_P013828	TAAGACATGAAAGGTTGGCCTTACTGTTGAACAAGAAGTAAATCCACAGGCTCCTGTTAT
RARRES1	A_23_P18078	TGAGACTCATCTGGGATTTGGCTTTCCTTGAAGCTCTTACGTGATGTGGAAATGACAA
RARRES3	A_23_P1962	AGGCGTTCTCTAGATCCTTCTCTGTTCCCTCTCTCGCTGGCAAAAAGTATGATCTAAT
RASSF2	A_23_P166087	GATGGTACATGACTTGATTCAACGTTTGGTTCTGAACTTACACACTGATGCGTTTACTCA
RPRM	A_01_P017520	AAGACCCAGAATAGAACTCGTAGCTCGTACTGCACGGTTTACGCCACAAAAGTGCTCTT
RPRM	A_23_P5370	ACCTTTGATTGTAACGTGCCGTTTTAAGAGGTTTTGTTGTTGTTGCTTGAATACAAATG
RPRM	A_23_P5365	TGCACGGTTTACGCCACAAAAGTGCTCTTACATCCGTGACACCGTTTTGACTTTTTGTT
RRAS	A_01_P012092	GGGCTACTGCCACTGTGTGCTTCCGCCAACACCTCCTGTCCACCTAAGCCTGGTGGG

Desarrollo de una Herramienta Molecular Diagn3stica de la Receptividad Endometrial

RRAS	A_23_P39076	TAGCTGCCTTCGCACCTTGCTGTGTGACCTGAGGCCCTCACTGAGCCTCAATTCCTCAT
RRAS	A_23_P39074	GTGGGGGTGAGGGGCTCCGGGTCCTGCTGTATATAACTCCCTCCCCAGAAAAATAAA
S100A1	A_23_P383227	CTCACAGTGGCCTGTAACAATTTCTCTGGGAGAACAGTTGAGCAGACAGCCACATTGGG
S100A4	A_23_P94800	ATAAGCAGCCCAGGAAGAAATGAAAACTCCTCTGATGTGGTTGGGGGTCTGCCAGCTGG
S100A4	A_32_P900698	CTTCACCAAGTACTCGGGCAAAGAGGGTGACAAGTTCAAGCTCAACAAGTCAGAATAAA
S100P	A_23_P58266	CCAAAAGTGTGGTTGGCAATTATCCCTAGGCTGAGCCTGCTCATGTACCTCTGATTA
SCGB2A2	A_23_P161940	CTTAACCAAACGGATGAAACTCTGAGCAATGTTGAGGTGTTATGCAATTAATATATGAC
SCGB2A2	A_01_P016031	CTTAACCAAACGGATGAAACTCTGAGCAATGTTGAGGTGTTATGCAATTAATATATGAC
SCYE1	A_01_P002141	AACATAAAAGCTAGGAGATGTGGCATCTGAACATTTTTGCTTTGCTGCCAGAGTAACCCT
SCYE1	A_23_P121686	GGTCATGTGTGCTAGTTCACCAGAGAAAATTGAAATCTTGGCTCCTCAAATGGGTCTGT
SERPINA5	A_23_P205355	AACAACCTGTCAGTTCATCTGCTGATGGGAAAAATGTTGGAATGGGAGTCTGAAATGGGGCT
SERPINA5	A_24_P321766	GTATGCTGGTAGCTAGTGATTTACACAGTTTAGTTGACTAATGAGGCATTACAAATAAT
SERPING1	A_23_P139114	CATGCTCTCCAAACCCTTTTTGCGAGCTTCTCTAGTTCAAGTTCACCAGACTCTATAAA
SERPING1	A_23_P139123	GACAACATTTGATCCCAAGAAAACCAGAATGGAACCCCTTCACTTCAAAAACCTCAGTTAT
SFRP4	A_23_P215328	TGTGTGCATACTCTAGAAGAGTAGGGAAAAATAATGCTTGTGTACAATTCGACCTAATATGT
SLC15A1	A_24_P8165	TGACCACTGGAGAGTCTAAGCCATCCTTAGCATGGGCATCCATCTTCTAAACTGTTATT
SLC15A1	A_23_P128609	GTAGTGCTTATACCTATATAGTCCAAAGGAAGAATGACAGCTGCCCTGAAGTGAAGGTGT
SLC15A2	A_23_P92107	ATACCCTGCAGTGCCTGTAGAACAGAAGATAAGAACTTTTCTGAAATTTGGGCTTCT
SLC15A2	A_01_P004018	TATGTATGATGTGATCTGGTCCAGCCAGGGCTGGTTTGTGAGCTATCTAGGTTTGATAA
SLC16A6	A_23_P152791	GAGATATGTAGAAGACTCTTTGGTTACATCCGATATTAATAAGTACATGAACTGG
SLC16A6	A_01_P007822	AATAGTTTTGTAAACCTTTTGTAAAGTACCAAGGCTCCCATTAACAATTTACGGCCTC
SLC16A6	A_24_P731648	ACTACATTTTAAAGGAATGTGTATGTGAAGAGCACTACCAACATCGCTTTTGTTTAAG
SLC1A1	A_01_P003853	AGATAGCAGAAGAGTAATAAGTACTCAGTATTGACCACCTACATCTGAAATCTACAACA
SLC1A1	A_23_P216468	GAAATCCAAAATAGTCATGTTTCTGCAGTATTCTGTAGCCAACCTTAAACTGTGCTTTCA
SLC1A1	A_24_P232252	CATTGTGAATCCCTTTGGCTTGGAAATCCACAATCCTTGACAACGAAGACTCAGACACCAA
SLC38A1	A_23_P363399	TTGGGGATGATTCTTACCTTGGTAATTAATGAAGCTACACATTTGGGTAATCTAGCAA
SLC38A1	A_01_P001379	GTACCTAAGTAAGGATGATCTAGGATAAGTAACCTCTGTTTTATATTGAGTACTTTAGGG
SLC7A1	A_23_P117242	GTGAACATTTCAACCAGCCTTATAGCTGTTCTCATCATCACCTCTGCATTGTGACCGTG
SLC7A1	A_24_P253251	CAAGGCAAATGATTTTGTGTTTCTTGATGACAGACTATTAAGTTGGGACTTATTTCCC
SLPI	A_23_P91230	AGGTCCTTCCACCCTGAGACTTGGCTCCACCCTGATATCCTCCTTTGGGAAAGGCTT
SLPI	A_01_P005468	ACAGCAGGATTTCCAGGAAGTCCAGTTGATCAATGAATAAATAAATGAGCCTATTTCTCT
SLPI	A_24_P190472	AAGAGATGTTGCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC
SNX10	A_23_P8513	TTAAGAGCGATCCTCATCCCTCAGCAATATGATTTGAGTTCACACTATTTCTGTTTT
SNX10	A_24_P98109	AAGAGTGGCAGAGGCTACTACAAAAGCAACCTTTTCATTTTCACTAAGAGTTTAAAGC
SOD2	A_24_P935819	GAAGATAATCGATAGTCATGTTTTTAGACTCTCTGTATTGCTTGGTAAGCTACGTAGTA
SOD2	A_23_P134176	TTGATGTGTGGGAGCAGCCTTACTACCTCAGTATAAAAATGTCAGGCCTGATTATCTAA
SORD	A_32_P89691	CGTCTGGGATGAAGGCTTGTGAGCACTTCCAGTTTGAACGCAATGTTTCTAGAGACATA
SORD	A_32_P85311	CTCTGGAGAAAGCTCTGGAGGCCCTTGAACATTTAAAAGGGATTGGGGTTGAAAATCA
SORD	A_32_P127153	GTAGCCATGACATAGCTTGAGCTATAGCCTTAATTCCTTACTTTGGCTATGGGTGGAGG

SORD	A_23_P77103	CAGCACTTCCAGTTTAGAACGCAATGTTTCTAGAGACATATTGGCTGTTTGTGTTGATGA
SORD	A_01_P000174	ATGAGGGCTGAGTTATGAAAAGATAACTTCTGAAGACTTAACCTGGCCCGAAGCTGATTT
SOX17	A_23_P82775	ACCCTATTTCCAAGTTCAAGTTAACTAGCTTTGAATGTGCCAAAACAGCTTCTCCAT
SOX17	A_01_P001801	GTCCAAAACAGCTTCTCCATTTCTGAAAGTTTATTGATCAAAGAAATGTTGTCGCCG
SOX17	A_24_P46946	CTGGGCGCGCTCCGCTCGAGGGCTACCCGTTGCCACGCCCAGACAGTCCCGCTGGAC
SPDEF	A_23_P111194	CCGCCATGAACTACGACAAGCTGAGCCGCTCCATCCGCCAGTATTACAAGAAGGGCATCA
SPP1	A_23_P7313	TTCCACAGCCATGAATTTACAGCCATGAAGATATGCTGGTTGTAGACCCCAAAGTAAG
SPP1	A_01_P017618	GAATGTAATAAGAATCTGGTGGTGTCAATTGCTTACTTGTTCACAGTTGTCCAGCA
STAR	A_32_P527817	GAAATATAGTCCAAGCTTCTCTGTGAAAAAGACAAAACCTATTAGTAGACATGTTTCC
STAR	A_23_P8981	TCATTAGTAGACATGTTTCCCTATTGCTTTCATAGGCACCAGTCAGAATAAAGAATCATA
STEAP4	A_24_P351906	ATATCCCCTTGGATTTCACTTGCATTGTGCAATAAGCAAAGAAGGGTTGATAAAAGTTCT
STEAP4	A_23_P255231	CTGGAAGAGAACCACATTTTATCTCAGGTTAGTGAAGAATCAGTGCAGGTCCTGACTCT
SYNE2	A_32_P109727	AATTAAGAGGTACCCCTTTGAGACACTAGCCCAGTGGAACTAGAACACGCTACTTTAAA
SYNE2	A_23_P205553	AGGGAATAATGCAACTGGAAGGAGCTGGTAGAACCTCAACCTCAAGTGGACATGTTA
SYNE2	A_23_P128887	AACCTCAAGTGGACATGTTACAGGAGATTTCAAACAGCCTTCTCATTAAAGGGACATGGAG
SYNE2	A_23_P140277	AATCTGGGAGGCAGAAGCCAAATCTGTTTTGGATCAAGATGATGTGGACCTCAATGGA
SYNE2	A_24_P68311	CAATGCGCCACCTATGAGTCTGCAATGTGACCGATTTTAAAGTCAAGTTGGAGAAATGGG
SYNE2	A_24_P307759	AAGAAGCTAAAGAGAAAGTCCAGATCAATGTGGTAAAACCTATTGCAGCGTTGAAGAACT
TACC3	A_23_P212844	AGCTCCACGGGAAGAACCTGGAAGTGGGAAGATCATGGACAGGTTCAAGAGGTTGTGT
TAGLN	A_01_P002824	TTGAAGAAGAACCAGCCAGCCTGCCTCCTATCTTTTCCGGAATATTTTGGGGTTGGA
TAGLN	A_23_P87011	ACCCGTGTGGTACCTTCAGCCCTGGCCAAGCTTTGAGGCTCTGTCAGTACGCAATGGTAA
TAGLN	A_23_P87013	CCGTGGAGATCCCAACTGGTTTATGAAGAAAGCGCAGGAGCATAAGAGGGAATTCACAGA
TBC1D2	A_23_P123662	CAAGCAAGAGCACTGCCTCTATAGGGTAACTGGAACATTCTCTAGGTTATATCAATATA
TBC1D2	A_23_P123666	GCCGGAAGCTGATGAACATCGCCTCAATGACATGAACCCCTTCCGCATGAACAGCTGC
TCN1	A_23_P64372	CCAGAAAATGAATGATACTATATTTGGTTTCACAATGGAGGAGCGCTCATGGGGCCCTA
TH	A_01_P009817	TCCGCCCTTCCCTGACACTGTCTGTGCCCAATCGCCGTCACAATAAAAAGAAAGTGTGG
TH	A_24_P924602	TCACAGGCTCTCCGTGGCCTGGAAGTGCAGCCCAGCTGCATCCTACACCCCACCCCAA
TH	A_23_P258633	CTGTCCACGCTGACTGTTTACGGTGGAGTTCGGGCTGTGTAAGCAGAACGGGGAGGTG
TH	A_23_P95891	GGATGAGAAATTGAGAAGCTGTCCACGCTGCATGTTTACGGTGGAGTTCGGGCTGTG
THBD	A_01_P013083	TTCCAGACTGCTTCCAATTTTCTGGAACACATTAATATGGATCAGTTATAAGTAGCAG
THBD	A_23_P91390	ACAGGGTCTGCAAGGCTTTGGTTTACAGTAAGCTAGGAATGAAATCCTGCTTCAAGTGTAT
THBD	A_32_P192334	TCTTTCTGTAACCATAACAACCTCATATATGAGGACTTGTGTCTCTGTGCTTTAAATGC
THBS2	A_23_P253652	ACTTTTTACTTTTTTGGCGTGGAGCTGATTCCCGAGACCAACGAAGCGTTGGGATACT
THBS2	A_23_P62021	TCGTTGTGTTGTTTGTGCTGCACTTTTTACTTTTTTGGCGTGGAGCTGATTCCCGAGAC
THBS2	A_24_P605612	TGTGGAGTAAAATTGGGCATGCCATTACATTGCTTTTTTCTGGTGGTAAAAGAATGA
TIMP3	A_23_P399078	CAAGATGGCACCCAAGTGTGGCTTCTGGCTACCTAAGGTTAACATGCTACTAGAGTAT
TIMP3	A_01_P005541	CTGTCTAGAAGGAATGATTTGTTGCTAAATTTGTAGCACTGTTTACAGTTTCTCCCA
TMEM16A	A_01_P014722	CCTTTTATGGAACCGTTTTTAAAAAAGTGAATGTACACAAATCCACAGAAGACTGTGG
TMEM16A	A_24_P87036	ATGAGTGATCTAAATTTGAGCAATGATACTAAACAACCTCTGAAATTTCTCAAGCACC

Desarrollo de una Herramienta Molecular Diagn3stica de la Receptividad Endometrial

TMEM16A	A_24_P110831	CTTATCTGGTGTTCATATAGAATCACCTAGAAGGATAAAGTCGCTGTAGAGTTAATGA
TMEM16A	A_23_P98304	TCAGGGTTTCAAGAAGTCTTAGGGCTCCAGGGGTCCCCTGGAAGCTTTAGAATATTTAT
TMEPAI	A_01_P015348	GAAGAACTGCTTGTGTGTATCAGTAATCATTAGTGGAATGATGACATTCTGAAAAGC
TMEPAI	A_24_P413126	AAGAACTGCTTGTGTGTATCAGTAATCATTAGTGGAATGATGACATTCTGAAAAGCT
TMEPAI	A_23_P57089	GCCGGCTGGGGCTGCGTAGGTGAAAAGGCAGAACACTCCGCGCTTCTAGAAGAGGAGT
TMSL8	A_23_P137173	TGTAGTCTTGCATCCTATACTGGATAAGCCTGTTTTAACCTGCTATGATGGGTGCTTCC
TMSL8	A_23_P137179	TAGAGATTTTAGGCGTCTTCGGATATCTTCTCACCTATGTTCCCTGGCTAAGAAGTCAGA
TOP2A	A_23_P118834	TGTGAGGCGATTATTTTAAAGTAATTATCTTACCAAGCCCAAGACTGGTTTTAAAGTTACC
TRA@	A_23_P106061	CTCACCTTTGGGACAGGCCTCAGCTAGAAGTGGGACTCAATATCCAGAACCCTGACCCT
TRA@	A_23_P258504	ATCTTCAGTGGTTCTCTGGGCTCTAGGCTCCTGGAGAATGTTGTGAGGGGTTATTTTT
TRA@	A_32_P806170	TTTCTTTAACCATTTTTGAAACCCCTTCAAAGGCAGAGACTTGTCCAGCCTAACCTGCCTG
TRA@	A_23_P44857	ATCATTTCATGGTCTAAAAACAATGAAACCAATGAAATGGCCTCTCTGATCATCACAGAA
TRA@	A_23_P258798	TAGTGTTCATAAAGAAATACATAGTATTCTTCTTCTCAAGACGTGGGGGAAATTATCTC
TRH	A_23_P132760	TCTGGCTGTGGGATAAATGTGTGTTGGAATATTGAAACATCGCCTAGGAATTGTGGTTG
TRH	A_01_P011590	AGCCCTTTCCCTTGTCAATGACAGTCATCCTAATGATAATAAAACCTGCATCCAGATAA
TSPAN8	A_23_P36531	GCTGATTGGGAAATAATTTCAACACTATCCTGAATTATGTGCCTGTCTAGATAAGCAG
TSPAN8	A_23_P36528	TTGAGATACTGGGTTTGGTGTCTTCTATGGTCCTGATTGCCAGATCGGGAACAAATGAA
TTC21B	A_23_P5422	TGATGACTCTGCTTCTATGAGGTCACCAGCAGTAGAACCATATCTTGCTGGCATACTT
TTC21B	A_24_P111511	ATCCAACCTATCCAAAAATCAGAAAGGATATACTTGATAAGGCCCGTGCCTTTAAGAC
VCAM1	A_23_P34345	AAAGCCAACATGAAGGGGTCATATAGTCTTGTAGAAGCACAGAAATCAAAGTGTAGCTA
XCL1	A_23_P200752	CTTTACTGCGAATAAGCTTTTAAATGCTCCAAATGCTGACCCATGCAATATTTCTCATGT
XCL1	A_24_P45476	AGTTTATACTCACCTTTTATGAAAGCACTGCATGAATAAAATTTATCTTTGTATTTTA
XCL1	A_01_P019136	CACCTTAGAGTCCAGCTTGTAAATGGTCTTTACACATGAGTCACAAGTTACAGCTGTGA
XCL2	A_23_P321431	TCACTGCATACATTGTGGAAGGTGTAGGGAGTGAAGTCTCACATAGGAGGACCTGTGTGA
XCL2	A_01_P017396	CATGAATAAAGTTATTCCTTTGTATTTTACTTTTAAATGTCTTCTGCATTCACCTTATAT
XCL2	A_23_P51534	CCAGTTAGCAGAATCAAGACCTACACCATCACGGAAGGCTCCTTGAGAGCAGTAATTTTT
XCL2	A_23_P51532	CTCATTTCACCTTACACCTCATGGACTGAGATTATACTCACCTTTATGAAAGCACTGC

Anexo III. Detalle del *software* de la herramienta molecular diagnóstica de la receptividad endometrial.

El programa consta de cuatro pestañas: upload data, data management, analysis y diagnose (figura IIIA).

La función **upload data** es indispensable para introducir cualquier fichero con el que se vaya a trabajar en el programa. La función **diagnose** es la que sirve para diagnosticar cualquier muestra de la que deseemos testar su fenotipo, y las otras dos funciones (**data management** y **analysis**) sirven para investigar, innovar y reentrenar la herramienta predictora.

La pestaña **upload data** sirve para cargar el fichero de los *microarrays* a analizar (muestra) y definir sus características. Se especifica el tipo de fichero en la opción **data file/ select your data file** pudiendo seleccionar el tipo de fichero que sale del escáner del *microarray* bruto (**gpr array**) o introducir un tipo de fichero ya normalizado (**normalized data**). También se puede introducir información acerca de la muestra en el apartado **some additional data**. Una vez definido se introduce el fichero en **data file/ examinar**, se etiqueta la muestra (**label your data**) y se carga todo en el programa (**upload**).

Después de introducido el fichero de la muestra que queremos testar, se va a la función **diagnose** donde se seleccionan los dos modelos de entrenamiento establecidos en la presente tesis doctoral (figura IIIA).

El primer modelo está compuesto por el set de entrenamiento de las muestras control sanas PF, PR, R y PS; y el segundo modelo por las muestras control R y las muestras control R-FI. De ambos modelos se introduce un fichero con la matriz de expresión de cada muestra (gpr_files_recep_vs_prerecep_vs_prolif_vs_postrecep.txt

/gpr_file_recep_vs_fi.txt) y otro fichero con la clase a la que pertenece cada muestra (classes.txt) (figura IIIA).

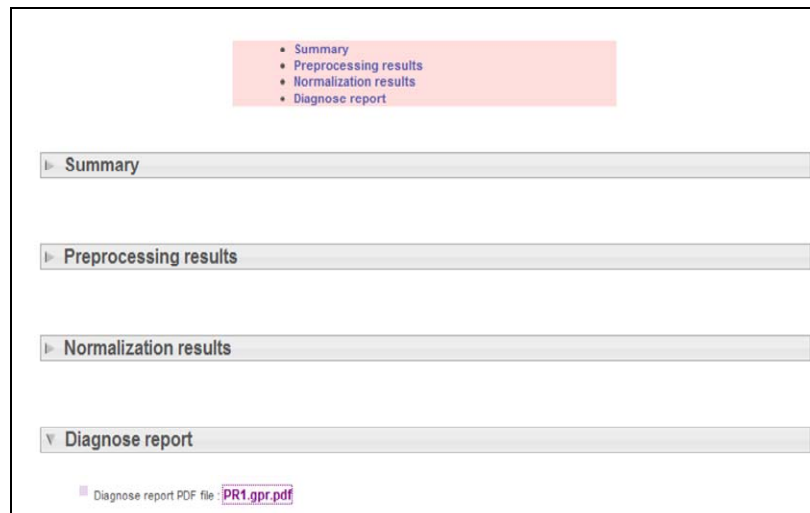
The screenshot shows the 'Diagnose' interface with the following sections:

- Navigation:** Upload Data, Data management, Analysis, Diagnose (active).
- User/Project Info:** anonymous working on project default 2.70 Mb of 1.00 Gb (0.26)
- Diagnose Section:**
 - Select the data training set for the first model:**
 - Choose a filter for the first model: gpr_files_recep_vs_prerecep_vs_prolif_vs_postrecep.txt
 - Choose a class file for the first model: classes.txt
 - Choose a filter for the second model: gpr_files_recep_vs_fi.txt
 - Choose a class file for the second model: classes.txt
 - Select the data prediction set:**
 - browse server no data selected.
 - Or go to Upload Data form: Upload [gpr]
 - Report parameters:**
 - Report number: [input field]
 - Clinic history number: [input field]
 - Doctor name: [input field]
 - Patient name: [input field]
 - Job:**
 - Job name: sample name
 - Job description: [input field]

Figura IIIA. Ventana del diagnóstico del software. Se muestran las opciones para elegir el set de entrenamiento (*training set*) con el que se quiere predecir, y la muestra que se quiere diagnosticar (*prediction set*).

Una vez elegidas las muestras de entrenamiento, se carga el fichero previamente introducido con la función upload (del que se quiere saber su diagnóstico) en **select the data prediction set** (figura IIIA). Se introducen los datos clínicos en **report parameters** y se da un nombre al trabajo diagnóstico que va a realizar el programa (**job**).

En los resultados del predictor obtenemos un sumario de las muestras utilizadas como set de entrenamiento para cada modelo (**summary**), el preprocesado de la muestra introducida (**preprocessing results**), la normalización de la muestra con respecto al resto de muestras del entrenamiento (**normalization results**) y un informe diagnóstico (**diagnose report**) (figura IIIB).



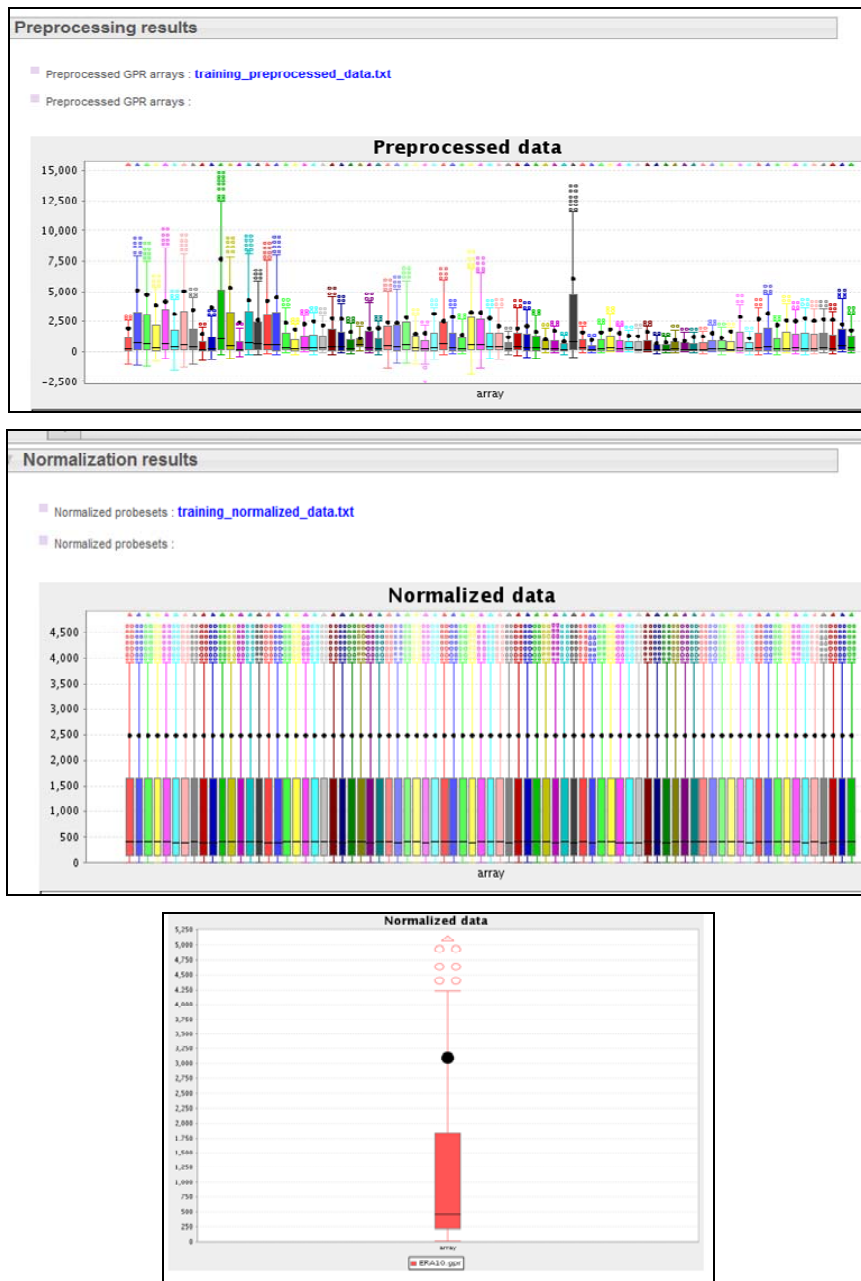


Figura IIIB. Ventana de resultados del diagnóstico. Se muestran cada una de las pestañas de los resultados (**summary**, **preprocessing results**, **normalization results**, **diagnose report**). Los resultados de preprocesado y normalización aparecen en forma de diagrama de intensidades. La muestra introducida se normaliza en función del set

de entrenamiento. A su vez también aparece el diagrama de cajas de la muestra de forma aislada.

El informe diagnóstico se genera con una probabilidad de diagnóstico asociada.

En resumen, se carga en **upload** un *microarray* en formato gpr de una biopsia endometrial en LH+7, se selecciona en la función de **diagnose** ese archivo y se obtiene el informe diagnóstico.

La función de **data management** sirve para preprocesar muestras nuevas con las que se vaya a ampliar el set de entrenamiento o para crear un set de entrenamiento nuevo para otro modelo de predicción.

La función de **analysis** sirve para reentrenar y crear modelos de predicción nuevos. Esta aplicación contiene las funciones necesarias para normalizar (**normalization**), entrenar (**training**) y predecir (**prediction**).

