



UNIVERSIDAD DE JAÉN
Facultad de Ciencias Experimentales

Trabajo Fin de Grado

APLICACIÓN DE LA BIOLOGÍA DE SISTEMAS PARA LA CARACTERIZACIÓN DE ALTERACIONES EN REPRODUCCIÓN HUMANA

Alumno: Elena Sánchez Escabias

Junio, 2019



UNIVERSIDAD DE JAÉN
Facultad de Ciencias Experimentales

Grado en Biología

Trabajo Fin de Grado

APLICACIÓN DE LA BIOLOGÍA DE SISTEMAS PARA LA CARACTERIZACIÓN DE ALTERACIONES EN REPRODUCCIÓN HUMANA

Alumno: Elena Sánchez Escabias

Junio, 2019

ÍNDICE

ABSTRACT	5
RESUMEN.....	5
1. INTRODUCCIÓN	6
1.1. Anovulación	6
1.1.1. Etiología	8
1.1.2. Tratamiento	9
1.1.3. Relación con el Síndrome del Ovario Poliquístico	10
1.2. Biología de Sistemas	11
2. OBJETIVOS.....	15
3. MATERIAL Y MÉTODOS.....	16
3.1. Obtención de datos	16
3.2. Análisis de redes.....	16
3.2.1. Análisis de interacción proteína-proteína.....	16
3.2.2. Análisis de parámetros de centralidad	17
3.3. Análisis de enriquecimiento funcional	18
3.4. Análisis de la intersección entre la anovulación y el PCOS.....	18
3.5. Esquema general	19
4. RESULTADOS.....	20
4.1. Anovulación	20
4.1.1. Obtención de datos	20
4.1.2. Análisis de interacción proteína-proteína.....	21
4.1.3. Análisis de parámetros de centralidad	22
4.1.4. Análisis de enriquecimiento funcional	24
4.2. Síndrome del Ovario Poliquístico (PCOS).....	25
4.2.1. Obtención de datos	25
4.2.2. Análisis de interacción proteína-proteína.....	27
4.2.3. Análisis de parámetros de centralidad	28
4.2.4. Análisis de enriquecimiento funcional	29
4.3. Análisis de la intersección entre la anovulación y el PCOS.....	31
5. DISCUSIÓN	33
6. CONCLUSIONES	36

7. BIBLIOGRAFÍA.....	37
----------------------	----

ABSTRACT

Anovulation is defined as the absence of ovulatory cycles for a long period of time and it is one of the most common causes of infertility in women of reproductive age. This work has been carried out with the aim of analysing the genes potentially related to this disease using a Systems Biology approach. Several conclusions have been drawn from the study of genetic interactions, the connectivity degree between genes and functional analysis of anovulation and Polycystic Ovary Syndrome (PCOS). In addition, the relationship between these two alterations was also studied. We have seen the importance of *INS* gene in anovulation and the presence of two genes that are not described yet in PCOS, but in anovulation.

Keywords: Anovulation – Functional analysis – Genetic interactions – PCOS – Systems Biology

RESUMEN

La anovulación se define como la ausencia de ciclos ovulatorios durante un tiempo prolongado y es una de las causas más comunes de infertilidad en las mujeres en edad reproductiva. En este Trabajo de Fin de Grado se han analizado los genes potencialmente relacionados con esta enfermedad a través del empleo de la Biología de Sistemas. Se han extraído varias conclusiones a partir del estudio de las interacciones génicas, del grado de conectividad entre genes y del análisis funcional de la anovulación y el Síndrome del Ovario Poliquístico (PCOS). Además, se ha analizado la relación existente entre estas dos alteraciones, de manera que se ha podido poner de manifiesto la importancia del gen *INS* en la anovulación, así como la existencia de dos genes que no se han descrito en el PCOS, pero sí en la anovulación.

Palabras clave: Análisis funcional – Anovulación – Biología de Sistemas – Interacciones génicas – PCOS

1. INTRODUCCIÓN

1.1. Anovulación

Según la Organización Mundial de la Salud (OMS), una de cada cuatro parejas en los países desarrollados tiene problemas para concebir. Entre las principales causas de infertilidad femenina se encuentra la anovulación, entendida como aquellos desórdenes en los que la ovulación ocurre de forma irregular o con baja frecuencia (Hakimi y Cameron, 2017). Suele estar asociada con la amenorrea (ausencia de menstruación) o la oligomenorrea (disminución de la frecuencia del número de menstruaciones y/o cantidad de pérdida menstrual), aunque también se han encontrado casos de mujeres con anovulación que presentaban menstruación (Li y Ng, 2012).

De un modo más concreto, la anovulación se define como la ausencia de ciclos ovulatorios durante un tiempo prolongado, y supone alrededor del 30% de los casos de infertilidad femenina, de ahí la importancia de su estudio (Hakimi y Cameron, 2017). Su origen más plausible puede ser un funcionamiento hormonal anormal del eje hipotálamo-hipófisis-gónada femenina. Como se observa en la Figura 1, en condiciones normales, la hormona liberadora de gonadotropinas (GnRH) es liberada de forma pulsátil por el hipotálamo. La GnRH viaja por el torrente sanguíneo hasta llegar a la hipófisis (o glándula pituitaria), donde induce la secreción de las hormonas folículo estimulante (FSH) y luteinizante (LH) hacia la sangre (Patel, 2018; Stamou *et al.*, 2015). Durante la fase folicular del ciclo ovárico, la FSH estimula el crecimiento de los folículos ováricos, uno de los cuales progresará hacia su maduración. En cuanto a la LH, ésta será responsable de la liberación del oocito del folículo ya maduro (ovulación) a mediados del ciclo (Schiller *et al.*, 2016).

A nivel fisiológico, la LH controla la esteroidogénesis. Se une a receptores de las células de la teca durante la fase folicular e induce la síntesis de androstenediona, la cual se dirige a las células de la granulosa para dar lugar a estrona, gracias a la acción de la aromatasa, y, finalmente, estradiol, por acción de otra enzima. La aromatasa y la enzima responsable de catalizar la conversión a estradiol son reguladas por la unión de la FSH a receptores de las células de la granulosa. En la mitad del ciclo aumentan los receptores de la LH en las células de

la granulosa, permitiendo que la LH se una y se libere mayor cantidad de progesterona durante la fase lútea (Oyola y Handa, 2017). Si no ha habido fecundación, el cuerpo lúteo se retrae y, por tanto, los niveles de estradiol y progesterona disminuyen, teniendo como consecuencia la menstruación. Si hay fecundación, el cuerpo lúteo se mantiene gracias a la gonadotropina coriónica humana (Schiller *et al.*, 2016).

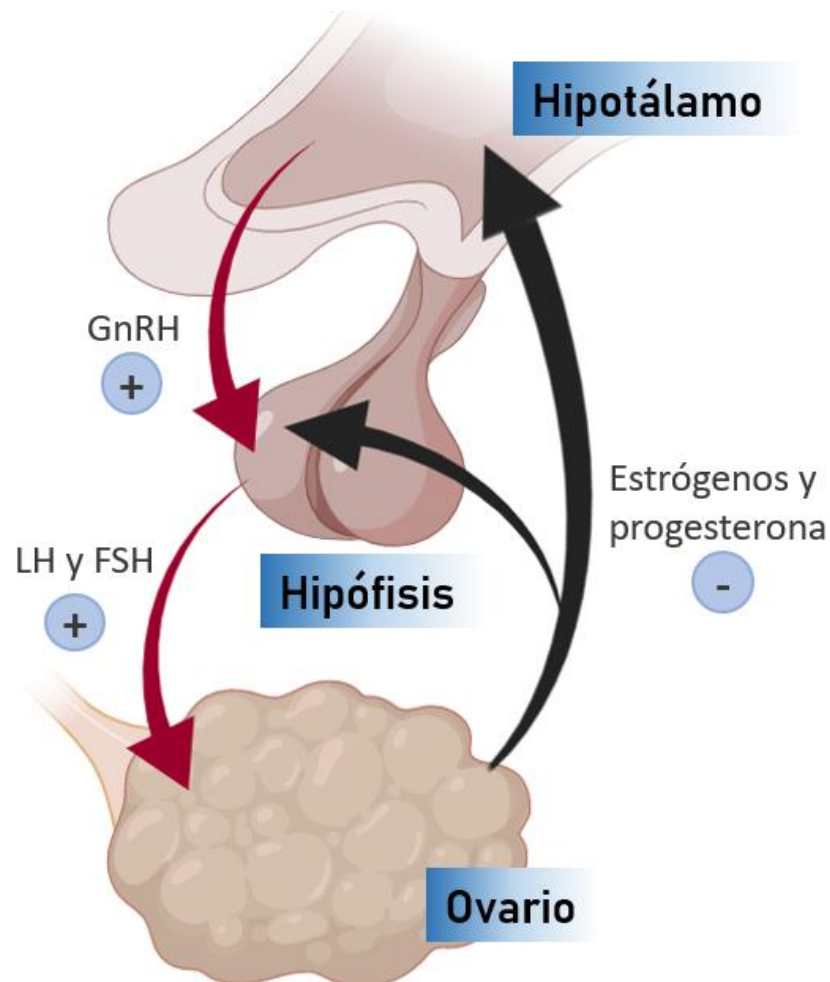


Figura 1. Representación esquemática del sistema de regulación del eje hipotálamo-hipófisis-gónada femenina. La hormona liberadora de gonadotropina (GnRH) hipotalámica favorece la secreción a la sangre de la hormona luteinizante (LH) y la hormona folículo estimulante (FSH) en la hipófisis o pituitaria. En el ovario, la FSH y la LH inducen la liberación de estrógenos y progesterona, que, en función del momento del ciclo ovárico, llevarán a cabo una retroinhibición o una retroalimentación sobre el hipotálamo y la hipófisis (adaptado de Schiller *et al.*, 2016). Imagen creada con [BioRender](#).

Cabe destacar que la OMS ha determinado tres tipos de anovulación en función de la concentración de gonadotropinas y estradiol en sangre (Lie *et al.*, 2015):

- a) Grupo I: Anovulación con niveles bajos de gonadotropinas y estradiol.

- b) Grupo II: Anovulación con niveles normales de gonadotropinas y estradiol.
- c) Grupo III: Anovulación con niveles altos de gonadotropinas y bajos de estradiol.

1.1.1. Etiología

La alteración hormonal entre el hipotálamo, la hipófisis y el ovario que conduce a la aparición de la anovulación puede tener su origen en un hipogonadismo hipogonadotrópico, un hipogonadismo hipergonadotrópico o una hiperprolactinemia. Asimismo, la anovulación también puede darse en condiciones normogonadotrópicas (Li y Ng, 2012; Lie *et al.*, 2015).

El hipogonadismo hipogonadotrópico produce la anovulación tipo I según la OMS (Li y Ng, 2012). La disminución de los niveles de gonadotropinas (FSH, principalmente) y estradiol se debe a un desorden hipotálamo-hipófisis, y suele estar asociado al ejercicio extremo, al estrés físico, a la anorexia nerviosa (Lie *et al.*, 2015) o a un tumor intracraneal (Li y Ng, 2012). Por otro lado, hay estudios que han concluido que el ejercicio moderado puede reducir el riesgo de anovulación, pero también señalan que la actividad física excesiva, como en el caso de mujeres atletas, puede aumentar la frecuencia de infertilidad anovulatoria debido a alteraciones metabólicas energéticas que repercuten en el ciclo ovulatorio (Hakimi y Cameron, 2017).

La anovulación normogonadotrópica está asociada con la anovulación tipo II según la OMS (Li y Ng, 2012). Un gran número de mujeres con Síndrome del Ovario Poliquístico (PCOS, por las siglas en inglés de *Polycystic Ovary Syndrome*) presentan este tipo de anovulación. Aunque no se conoce con claridad su causa, parece estar asociada a alteraciones en la hipófisis, concretamente en la liberación de FSH (Lie *et al.*, 2015). También se puede deber a un aumento de los niveles de andrógenos en la mujer (Li y Ng, 2012).

El hipogonadismo hipergonadotrópico está relacionado con la anovulación tipo III según la OMS (Li y Ng, 2012), y se debe al mal funcionamiento del propio ovario, como es el caso de la insuficiencia ovárica (Lie *et al.*, 2015).

Por último, la hiperprolactinemia puede interferir en la liberación pulsátil de la GnRH en el hipotálamo, provocando la aparición de irregularidades hormonales y, por tanto, de la anovulación. Las causas más comunes que producen un aumento de

la secreción de prolactina son el prolactinoma, el hipotiroidismo primario o la insuficiencia renal crónica (Li y Ng, 2012).

1.1.2. Tratamiento

La anovulación es un problema para la concepción, por lo que es necesario un abordaje terapéutico para intentar revertir este desorden ovulatorio. Los principales tratamientos se centran en la administración de medicamentos o en la cirugía, aunque también se han observado resultados positivos del ejercicio moderado y de la dieta en la restauración de la ovulación, principalmente en mujeres anovulatorias con sobrepeso u obesidad, lo cual es frecuente en pacientes con PCOS (Hakimi y Cameron, 2017).

El citrato de clomifeno es el principal fármaco utilizado para inducir la ovulación en mujeres anovulatorias. Es un compuesto no esteroideo semejante estructuralmente a los estrógenos y que actúa como agonista y antagonista de estos, ya que se unen competitivamente a los receptores de estrógenos (Von Hofe y Bates, 2015). Otro medicamento similar es el tamoxifeno, pero su uso no está tan extendido (Li y Ng, 2012). El citrato de clomifeno es capaz de unirse a los receptores de estrógenos nucleares de modo competitivo durante más tiempo que los propios estrógenos, lo que resulta en el envío de una señal al hipotálamo, que reconoce una falsa bajada de estrógenos. En respuesta, el hipotálamo aumenta la secreción de GnRH por mecanismos de retroalimentación. El aumento de GnRH hace que la hipófisis libere más gonadotropinas (FSH y LH), las cuales se dirigen al ovario, favoreciendo la actividad folicular (Von Hofe y Bates, 2015).

Se ha observado que el citrato de clomifeno tiene una gran tasa de éxito en mujeres con anovulación tipo II, mientras que en los casos de hipogonadismo hipogonadotrópico (anovulación tipo I) e hipergonadotrópico (anovulación tipo III) no se han detectado resultados tan satisfactorios (Von Hofe y Bates, 2015). Si la paciente no responde al citrato de clomifeno, se puede sugerir la administración de inhibidores de aromatasas, como el letrozol. Las aromatasas catalizan el paso de androstenediona a estrona y de testosterona a estradiol. Como consecuencia, los inhibidores de aromatasas reducen los niveles de estrógenos, lo que desencadena la activación de los mecanismos de regulación hipotálamo-hipófisis-ovario, haciendo

que aumente la liberación de la FSH endógena, necesaria para la maduración del folículo ovárico (Li y Ng, 2012).

La cirugía es otra opción indicada en caso de resistencia al citrato de clomifeno en mujeres con PCOS, concretamente la perforación ovárica laparoscópica, ya que parece destruir tejido productor de andrógenos en el ovario (Li y Ng, 2012).

En el caso de que haya una afección hipotalámica que cause hipogonadismo hipogonadotrópico (anovulación tipo I), se puede administrar GnRH exógena de forma pulsátil y transitoria a la paciente si quiere quedar embarazada. Si lo que existe es un fallo en la liberación de gonadotropinas por parte de la hipófisis, se pueden suministrar de forma exógena tanto la FSH como la LH para la maduración del folículo y la preparación del endometrio para la implantación, respectivamente (Li y Ng, 2012).

Los agonistas de la dopamina tales como la cabergolina o la bromocriptina son utilizados para combatir la hiperprolactinemia, llegando a restaurar la ovulación en un 90% de los casos. La bromocriptina es menos utilizada debido a que presenta una menor efectividad y a los posibles efectos secundarios que pueden aparecer, como náuseas, dolores de cabeza, dolor abdominal, etc. Sin embargo, la cabergolina también puede producir fiebres y valvulopatías cardíacas si se usa en altas dosis. No obstante, los agonistas de la dopamina no suponen un riesgo para la fertilidad o el desarrollo fetal, aunque se aconseja el cese del tratamiento una vez la mujer queda embarazada, a no ser que la hiperprolactinemia se deba a un adenoma hipofisario, ya que su función es inhibir la secreción de prolactina (Li y Ng, 2012).

1.1.3. Relación con el Síndrome del Ovario Poliquístico

La anovulación puede ser consecuencia del PCOS, el cual supone el principal problema de infertilidad femenina de origen endocrino, pues afecta al 5-10% de las mujeres en edad reproductiva (Liu *et al.*, 2019). Se estima que más del 12% de los desórdenes endocrinos en mujeres se deben al PCOS (Meier, 2018), de ahí la importancia de su estudio.

El PCOS se debe a una alteración de los niveles de hormonas sexuales femeninas, lo cual conduce a la formación de quistes en los folículos antrales del ovario. El oocito queda encerrado en este quiste llegando a provocar la anovulación

y, como consecuencia, se interrumpe el ciclo menstrual y se produce la amenorrea. Asimismo, el quiste secreta andrógenos, lo que hace que tenga lugar un hiperandrogenismo que, en última instancia, da lugar a la aparición de acné, hirsutismo o hiperinsulinemia, entre otros (Patel, 2018).

Puede existir una predisposición genética a padecer PCOS, de manera que ciertas mutaciones (SNPs [por las siglas en inglés de *Single Nucleotide Polymorphism*], inserciones, deleciones, etc.) en los genes que median en la aparición de PCOS pueden asociarse a su incidencia, en combinación con factores ambientales y estilo de vida. Asimismo, se está estudiando la importancia de la epigenética en el riesgo de desarrollar esta enfermedad, y se ha descubierto que la modificación del patrón de metilación del ADN podría aumentar el riesgo de padecerla (Patel, 2018).

Otro síntoma común en mujeres con PCOS es la resistencia a la insulina, que conduce a la aparición de hiperinsulinemia como mecanismo de compensación. Aunque no se conoce bien el mecanismo, se cree que este aumento de insulina podría estar asociado a un aumento de la hormona antimülleriana (AMH), producida por las células de la granulosa. La AMH está implicada en el proceso de foliculogénesis (fase de maduración folicular), inhibiendo a la FSH y, por ende, la expresión de la aromatasa (esteroidogénesis). En definitiva, el aumento de la AMH evita la maduración completa y ovulación de los folículos antrales, produciendo la anovulación (Liu *et al.*, 2019).

1.2. Biología de Sistemas

En la actualidad, las ciencias ómicas, tales como la genómica, transcriptómica, proteómica o epigenómica están cobrando una gran importancia en relación al estudio de las interacciones, estructuras y procesos que ocurren en funciones biológicas complejas, lo que ayuda a generar un gran número de datos (Djordjevic *et al.*, 2019; Vargas *et al.*, 2018). Debido a la enorme cantidad de información que se está generando gracias a la investigación con estas tecnologías ómicas de alta resolución, puede resultar complicado la inclusión y análisis de todos los resultados obtenidos de forma aislada pues, en realidad, los distintos sistemas, ya sea a nivel de ambiente o de organismo, suelen trabajar de forma conjunta e interrelacionada

(Vargas *et al.*, 2019). La Biología de Sistemas es la encargada de estudiar los sistemas biológicos empleando como herramientas la computación y los modelos matemáticos desde un punto de vista holístico (Nielsen, 2017). Gracias a esta disciplina se consiguen extraer resultados cualitativos y cuantitativos (Delles y Husi, 2017; Nielsen, 2017), lo que permite la integración de todos estos datos de manera interdisciplinar (Vargas *et al.*, 2019). Con este fin, se han desarrollado bases de datos como *Database of Annotation, Visualization and Integrated Discovery* ([DAVID](#)) que buscan unificar la biología molecular con la de sistemas (Eddy *et al.*, 2015; Vargas *et al.*, 2019). Gracias a las investigaciones con enfoque reduccionista se ha podido conseguir una gran cantidad de datos, como si de piezas de un puzzle se tratara, y es la Biología de Sistemas la que busca recomponer y organizar todos los componentes biológicos que lo conforman (Eddy *et al.*, 2015).

Por nombrar un ejemplo de la evolución que han sufrido estas técnicas a lo largo de los últimos años, hasta hace poco la secuenciación de unos pocos genes se realizaba con la técnica de Sanger, pero gracias al desarrollo de las técnicas de secuenciación de nueva generación se ha conseguido reducir el tiempo empleado y abaratar los costes, logrando obtener información de la secuencia de numerosos genes que podrían estar implicados en enfermedades concretas (Stamou *et al.*, 2015).

En definitiva, los estudios ómicos pueden proporcionar información sobre distintos biomarcadores para predecir y prevenir enfermedades, a la vez que se pueden utilizar con fines biomédicos como herramienta para determinar el tratamiento más adecuado para el paciente (Guzzi *et al.*, 2016). Se habla de Medicina de Sistemas cuando el objetivo último de la Biología de Sistemas es hallar y aplicar en la práctica clínica las conclusiones obtenidas (Vargas *et al.*, 2018). Muchas de estas investigaciones emplean *microarrays* en ensayos de expresión génica (Guzzi *et al.*, 2016) para el estudio de cáncer de mama, para determinar el metabolismo de un fármaco o para genotipar el tipo de virus del papiloma humano que presenta el paciente. Esto ha permitido desarrollar y crear productos con aplicaciones médicas, como *MammaPrint*, *AmpliChip* o *PailloCheck*, respectivamente (Li, 2016).

En relación a la reproducción humana, la Biología de Sistemas supone una gran ventaja a la hora de predecir el resultado de un tratamiento, ya que se pueden estudiar distintos marcadores que ayuden a determinar el éxito, o no, de un fármaco

o procedimiento. Asimismo, la Medicina de Sistemas está cobrando importancia en las técnicas de reproducción asistida, permitiendo estudiar nuevos métodos para conseguir transferir embriones óptimos en la madre (Vargas *et al.*, 2018). Sin embargo, tal magnitud de resultados ómicos, muchos obtenidos por la tecnología de *microarrays*, se deben unificar y es aquí donde también juega un papel importante la Biología de Sistemas (Nielsen, 2017; Vargas *et al.*, 2019).

Un *microarray* es un chip que contiene sondas complementarias a secuencias de ADN y que permite analizar la expresión de una gran cantidad de material genético de forma eficiente (Vargas *et al.*, 2019). Existen distintos tipos de *microarrays* en función de la molécula objeto de estudio: de ácidos nucleicos, de proteínas, de polisacáridos, de péptidos o de células (Li, 2016). En los *arrays* de expresión se puede utilizar ADN copia (ADNc) (no se utiliza el ARNm debido a su inestabilidad) u oligonucleótidos (secuencias cortas de ADN de unos 10-50 nucleótidos) como sondas que hibriden (Marzancola *et al.*, 2016; Vargas *et al.*, 2019). Asimismo, se puede hacer uso de uno o dos fluoróforos. En el primer caso se consiguen valores absolutos de expresión génica, mientras que para el estudio de la expresión diferencial se utilizan dos fluoróforos (cianinas Cy3 [verde] y Cy5 [rojo]) (Vargas *et al.*, 2019). En referencia al material utilizado, los chips de *microarray* suelen fabricarse sobre soportes de cristal, aunque también se emplean de silicona o polímeros (Marzancola *et al.*, 2016).

Todos los datos que se obtienen son recogidos en distintas bases de datos para, posteriormente, ser tratados para su utilización en la investigación. Algunos científicos centran su estudio en la detección de genes con una expresión correlacionada dentro de un conjunto, es decir, realizan un análisis exploratorio. El *clustering*, o análisis de conglomerados, es un método de análisis exploratorio que permite agrupar en distintas clases (*clusters*) los diferentes datos, ya sean genes, muestras, o ambos, en función de unos parámetros y algoritmos que establece el investigador. El análisis de componentes principales (PCA, por las siglas en inglés de *Principal Component Analysis*) es otra técnica de análisis exploratorio en la que se realiza un análisis estadístico multivariante con el fin de reducir la complejidad del sistema (Vargas *et al.*, 2019).

Una vez obtenidos los datos de los genes con expresión diferencial, se puede llevar a cabo un análisis de enriquecimiento funcional de los mismos para analizar los procesos en los que está implicado ese conjunto concreto de genes y, además,

construir una red biológica. Para ello, se emplearán distintos tipos de software que comentaremos más adelante (Vargas *et al.*, 2019).

Otra opción que ofrece la Biología de Sistemas es la realización de los llamados meta-análisis (Vargas *et al.*, 2018). La información que se extrae de genes con expresión diferencial en los distintos experimentos con *microarrays* es muy heterogénea. La realización de meta-análisis puede ser una herramienta útil para corroborar o no la veracidad de los resultados mediante el empleo de métodos estadísticos (Ghojavand *et al.*, 2019).

2. OBJETIVOS

El principal objetivo que ha llevado a la realización de este Trabajo de Fin de Grado es conocer y aplicar algunas de las múltiples herramientas que ofrece la Biología de Sistemas en el estudio genético de enfermedades que afectan a la reproducción humana, concretamente la anovulación.

Del planteamiento de este objetivo general se derivan una serie de objetivos específicos:

1. Aprender a utilizar las herramientas computacionales propias de la Biología de Sistemas.
2. Obtener la lista de genes potencialmente implicados en la anovulación y PCOS.
3. Analizar las redes biológicas, incluyendo la interacción proteína-proteína y parámetros de centralidad, a través del empleo de programas computacionales.
4. Extraer información sobre los principales procesos biológicos, funciones moleculares y componentes celulares relacionados con la anovulación y el PCOS mediante el uso de programas bioinformáticos.
5. Estudiar la posible relación existente entre la anovulación y el PCOS.
6. Integrar conocimientos adquiridos en diferentes asignaturas del Grado.

3. MATERIAL Y MÉTODOS

3.1. Obtención de datos

La obtención de datos es uno de los pilares de este Trabajo de Fin de Grado y supone el acceso a la lista de genes cuyas variaciones podrían estar implicadas en la anovulación, así como en el PCOS. Como herramienta para conseguir esta información, se ha empleado [Phenopedia](#) (Yu *et al.*, 2010), una base de datos que recoge asociaciones genéticas humanas mediante una continua búsqueda y revisión de los artículos científicos presentes en [PubMed](#) desde 2001. Para utilizar esta plataforma, se escribe el nombre de la enfermedad que se desea buscar en inglés — en nuestro caso, *Anovulation* y *Polycystic Ovary Syndrome*. [Phenopedia](#) proporciona varios resultados dentro de los términos [MeSH](#) (de las siglas en inglés *Medical Subject Headings*), además de incluir el número de publicaciones, el número de genes estudiados, el número de investigadores y el lugar y fecha de publicación. Asimismo, genera una tabla con la lista de genes en orden descendente en función del número de veces que se haya asociado cada uno con la enfermedad buscada (Yu *et al.*, 2010).

3.2. Análisis de redes

Una vez obtenido el listado de genes que se han asociado con la anovulación y el PCOS, se procedió a determinar las relaciones existentes entre dichos genes para cada enfermedad.

3.2.1. Análisis de interacción proteína-proteína

Las proteínas pueden estar relacionadas unas con otras funcionalmente, como ocurre en las rutas metabólicas o en interacciones reguladoras, ya sea directa o indirectamente. Estas conexiones se pueden visualizar y representar mediante redes biológicas, lo que puede favorecer la predicción del efecto de un medicamento, ayudar a entender mejor las enzimas que actúan en una ruta metabólica, etc. Por

ello, su análisis puede resultar muy interesante desde la perspectiva de la Biología de Sistemas (Szklarczyk *et al.*, 2011).

Para la construcción de la red biológica se utilizó la plataforma web [STRING](#) (Szklarczyk *et al.*, 2015), donde los diferentes nodos (elementos que componen la red) están unidos por aristas (conexiones que existen entre los nodos) (Vargas *et al.*, 2019). [STRING](#) es un programa de libre acceso que recoge información de interacciones proteína-proteína y que incluye en su base de datos las relaciones tanto físicas como funcionales que se establecen entre un conjunto cualquiera de genes (Vargas *et al.*, 2018).

3.2.2. Análisis de parámetros de centralidad

Para el estudio de la estructura de la red biológica, es decir, de su grado de interacción, conexión, importancia de cada gen en la red, etc., se utilizó [Cytoscape 3.7.0](#) (Shannon *et al.*, 2003), una herramienta muy útil y gratuita. Este software de código abierto permite visualizar redes complejas e integrarlas con propiedades que nos interese analizar (Delles y Husi, 2017); en nuestro caso, son la centralidad de grado o *degree* y la intermediación o *betweenness*. La centralidad de grado (*degree*) es el número de interacciones directas que establece un nodo con otros nodos, o aplicado a nuestro caso, las conexiones que establece un gen con otros genes. Cuanto mayor sea el valor de *degree* de un gen, mayor importancia tendrá dicho gen dentro de la red. Por su parte, la intermediación (*betweenness*) se define como la medida de la relevancia de un nodo dado contando el número de geodésicas que lo atraviesan. El *betweenness* mide, por tanto, la importancia de un gen como puente entre dos genes que están alejados entre sí, siendo su valor máximo 1 (Boccaletti *et al.*, 2006).

Como criterio, se seleccionaron los genes que presentaban un *degree* por encima de la suma de la media y la desviación estándar de los *degrees* de todos los genes. El mismo procedimiento se aplicó para el *betweenness*.

3.3. Análisis de enriquecimiento funcional

El análisis de enriquecimiento funcional de un conjunto de genes se realiza para estudiar los procesos biológicos en los que estos están implicados. Para abordar este análisis se empleó el complemento *GeneMANIA* (Warde-Farley *et al.*, 2010) de [Cytoscape 3.7.0](#) para la anovulación, mientras que el estudio de los procesos biológicos, la función molecular y los componentes celulares de la lista de genes del PCOS se realizó con *Database of Annotation, Visualization and Integrated Discovery* ([DAVID](#)) (Huang *et al.*, 2009).

GeneMANIA es una herramienta de [Cytoscape 3.7.0](#) capaz de predecir interacciones gen-gen a partir de estudios previos y nos permite estudiar las funciones biológicas en las que participan un conjunto de genes, dando un q-valor para cada proceso. Como convenio, solo consideramos como funciones biológicas significativas aquellas que poseían un q-valor por encima de 0,05. Asimismo, con este complemento también somos capaces de generar una red de interacciones entre genes, unos pertenecientes a la lista de genes que introducimos en el programa, y otros predichos por *GeneMANIA* (Vargas *et al.*, 2018).

En cuanto a [DAVID](#), es un software especializado que permite estudiar varios grupos de genes a la vez, facilitando, además, valores de significación estadística para los diferentes resultados (Vargas *et al.*, 2019). Para cada una de estas categorías se detectaron los genes del PCOS implicados en una función biológica, molecular o componente celular, acompañados de su correspondiente p-valor de significación estadística y el FDR (de sus siglas en inglés *False Discovery Rate*); establecimos como requisito un valor de FDR superior a 0,05 para considerar un proceso, función o componente como estadísticamente significativo. El FDR se calcula debido a que se está trabajando con una gran cantidad de datos procedentes de múltiples experimentos, por lo que es necesario determinar la probabilidad de que aparezcan falsos positivos (van den Oord, 2008).

3.4. Análisis de la intersección entre la anovulación y el PCOS

Debido a la estrecha relación existente entre la anovulación y el PCOS, se decidió realizar un análisis más profundo para determinar qué genes eran comunes a ambas

alteraciones y cuáles eran propios de la anovulación y, por otra parte, cuáles del PCOS. Con este motivo, se empleó la herramienta [VENNY 2.1.0](#) (Oliveros, 2007), una plataforma informática de libre acceso que permite generar un diagrama de Venn para detectar la intersección entre varios conjuntos de datos (en nuestro caso, para ver los genes en común de alteraciones distintas: anovulación y PCOS).

3.5. Esquema general

El procedimiento secuencial que seguimos para la obtención de resultados se muestra en la Figura 2.

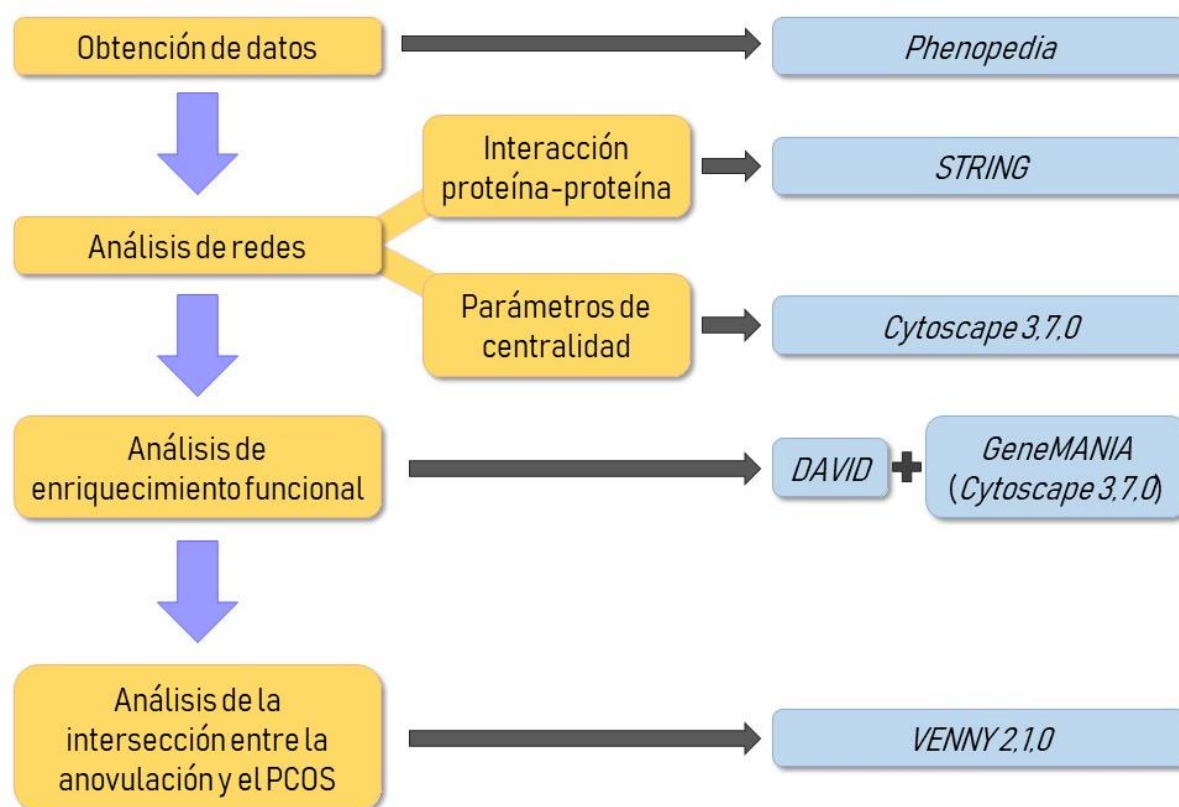


Figura 2. Esquema general de trabajo para la realización de un análisis génico mediante la Biología de Sistemas. A la izquierda y en amarillo se muestran las etapas que hemos seguido en este Trabajo de Fin de Grado, y a la derecha y en azul se muestran las plataformas empleadas para cada tipo de análisis. PCOS: Polycystic Ovary Syndrome (Síndrome del Ovario Poliquístico).

4. RESULTADOS

4.1. Anovulación

4.1.1. Obtención de datos

Mediante el uso de [Phenopedia](#) se han detectado un total de 39 genes que han sido relacionados con la anovulación en humanos en distintos artículos. En la Tabla 1 se muestra la lista de genes por orden decreciente de frecuencia con la que se han asociado a la enfermedad.

Genes potencialmente relacionados con la anovulación	
Símbolo	Nombre completo del gen
CYP19A1	Cytochrome P450 family 19 subfamily A member 1
INS	Insulin
INSR	Insulin receptor
LHCGR	Luteinizing hormone/choriogonadotropin receptor
CYP17A1	Cytochrome P450 family 17 subfamily A member 1
CYP11A1	Cytochrome P450 family 11 subfamily A member 1
AKR1C3	Aldo-keto reductase family 1 member C3
PPARG	Peroxisome proliferator activated receptor gamma
DENND1A	DENN domain containing 1A
THADA	THADA armadillo repeat containing
FSHR	Follicle stimulating hormone receptor
AMH	Anti-Mullerian hormone
AMHR2	Anti-Mullerian hormone type 2 receptor
NR3C1	Nuclear receptor subfamily 3 group C member 1
HLA-DQB1	Major histocompatibility complex, class II, DQ beta 1
HLA-DRB1	Major histocompatibility complex, class II, DR beta 1
HSD11B1	Hydroxysteroid 11-beta dehydrogenase 1
FST	Follistatin
CAPN10	Calpain 10
COMT	Catechol-O-methyltransferase
NOBOX	NOBOX oogenesis homeobox
ACE	Angiotensin I converting enzyme
DRD3	Dopamine receptor D3
ESR1	Estrogen receptor 1
ESR2	Estrogen receptor 2
F2	Coagulation factor II, thrombin
F5	Coagulation factor V
FSHB	Follicle stimulating hormone beta
TCF7L2	Transcription factor 7 like 2
FTO	FTO alpha-ketoglutarate dependent dioxygenase
C9orf3	Chromosome 9 open reading frame 3
HSD17B6	Hydroxysteroid 17-beta dehydrogenase 6
BMP15	Bone morphogenetic protein 15
H6PD	Hexose-6-phosphate dehydrogenase/glucose 1-dehydrogenase
MC4R	Melanocortin 4 receptor
MTHFR	Methylenetetrahydrofolate reductase
NEUROG3	Neurogenin 3

INSIG2	Insulin induced gene 2
PLIN	Perilipin

Tabla 1. Lista de los 39 genes que podrían estar implicados en el proceso de anovulación obtenida a partir de [Phenopedia](#). A la izquierda se muestra el símbolo del gen, y a la derecha, el nombre completo.

4.1.2. Análisis de interacción proteína-proteína

A partir de los 39 genes relacionados con la anovulación se construyó una red biológica donde se establecieron las interacciones entre las proteínas mediante el uso de la plataforma [STRING](#) (Figura 3). Cada uno de los 39 nodos representa todas las proteínas producidas por un locus en el ADN. Dentro de algunos de ellos se puede distinguir la estructura de una proteína, lo que indica que se conoce su estructura 3D. Asimismo, si se selecciona un nodo de la red, se puede ver información general sobre ese gen.

Se han establecido 175 conexiones entre los distintos nodos, pero esto no conlleva que estén cerca físicamente unos de otros. [STRING](#) ofrece información sobre el tipo de conexión a través de un código de colores para las aristas:

- Interacciones conocidas:
 - Azul: interacciones conocidas a partir de bases de datos (ej.: entre INS e INSR).
 - Morado: interacciones conocidas a partir de experimentos (ej.: entre INS e INSR).
- Interacciones predichas
 - Verde: interacciones predichas a partir de su cercanía física (ej.: entre H6PD y HSD17B6).
 - Rojo: interacciones predichas a partir de fusiones génicas (no se han encontrado conexiones de este tipo en la red biológica).
 - Azul oscuro: interacciones predichas a partir de coocurrencia génica (ej.: entre CYP17A1 y CYP11A1).
- Otros:
 - Amarillo: interacciones determinadas por minería de texto (ej.: entre MTHFR y F5).
 - Negro: interacciones determinadas por coexpresión (ej.: entre NOBOX y BMP15).

- Gris: interacciones determinadas por homología en proteínas (ej.: entre NR3C1 y ESR1).

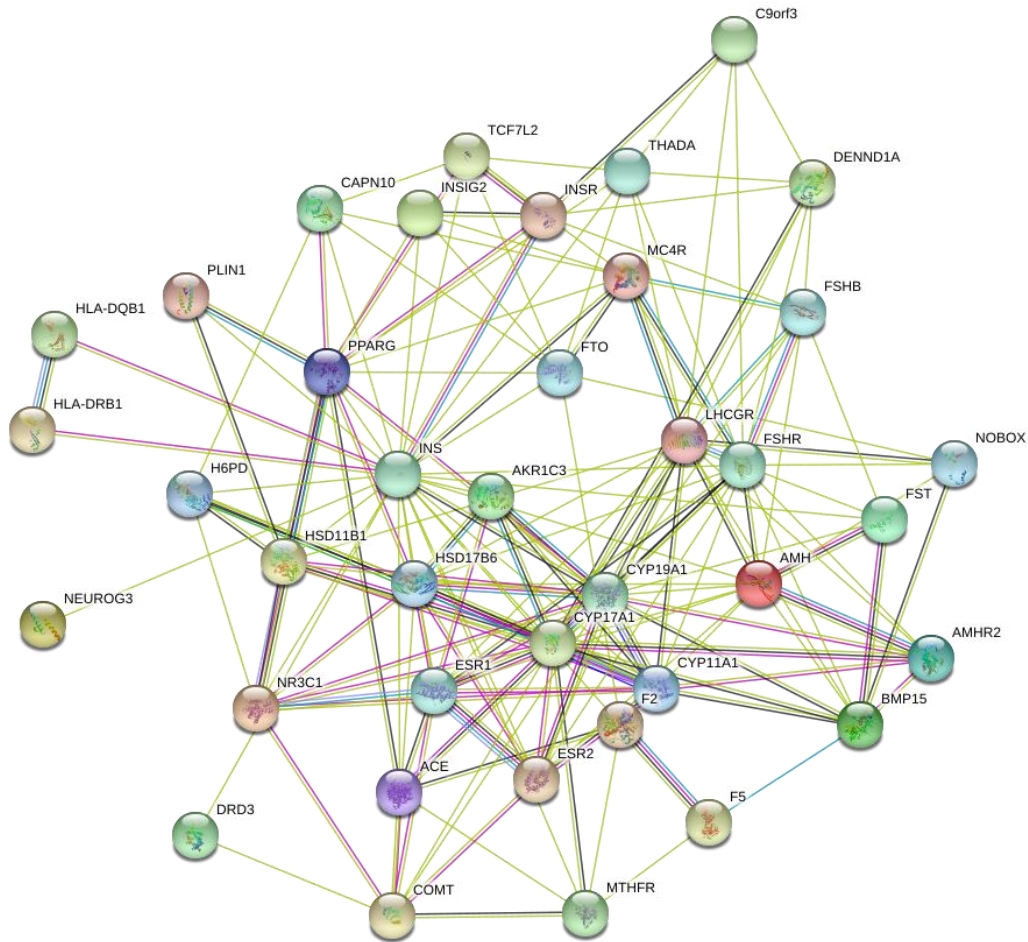


Figura 3. Red biológica generada con el programa [STRING](#), donde se muestran las distintas relaciones entre los 39 genes que están relacionados con la anovulación según [Phenopedia](#).

4.1.3. Análisis de parámetros de centralidad

Otra forma de representar la red biológica es mediante el programa [Cytoscape 3.7.0](#), como se muestra en la Figura 4. Esta red se basa en los valores de centralidad (*degree* y *betweenness*) que se han establecido para los 39 genes de la anovulación. Visualmente, se puede observar que el gen *INS* (gen de la insulina) tiene los valores más elevados de *degree* y *betweenness*, lo que se corrobora con los datos de la Tabla 2.

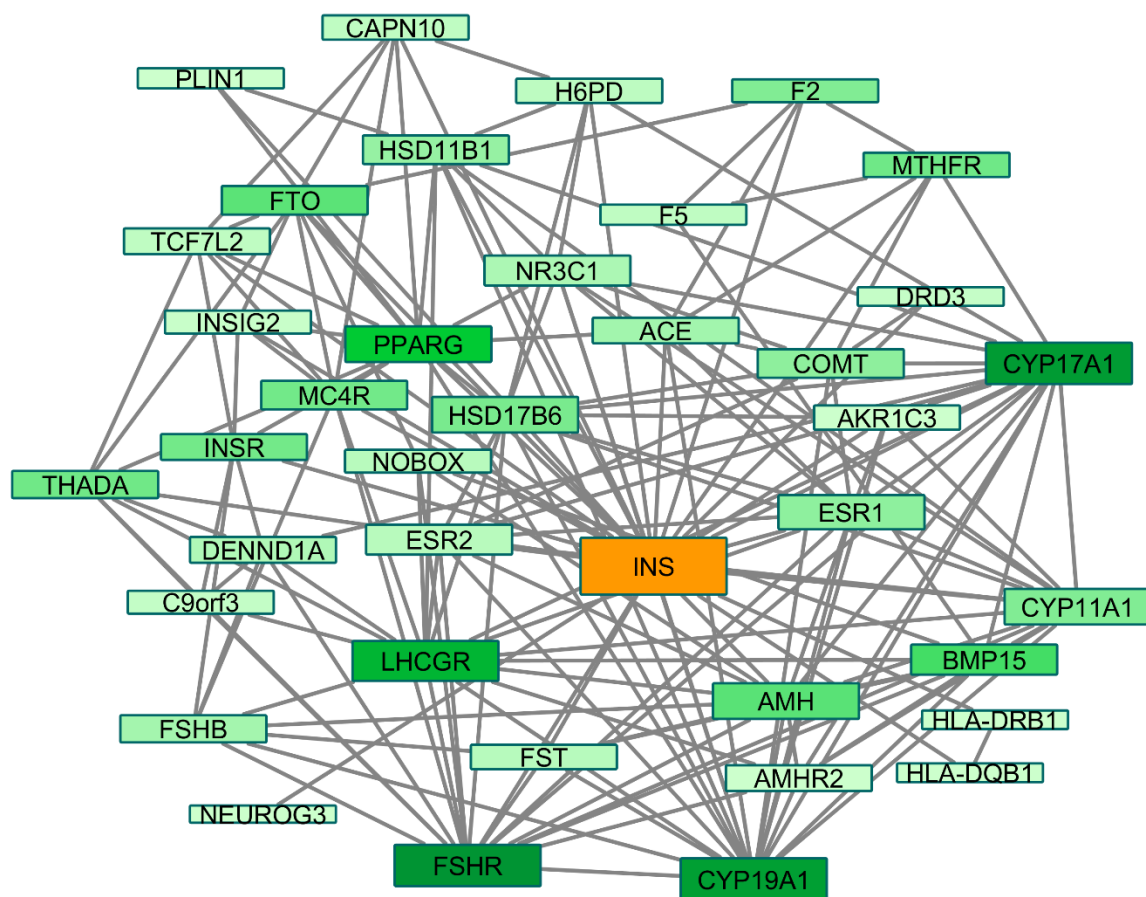


Figura 4. Red biológica generada a partir de [Cytoscape 3.7.0](#) en la que se muestran los genes implicados en la anovulación. Por orden decreciente de *betweenness*, los genes aparecen en color naranja y de verde oscuro a verde claro. El *degree* se ha representado en función del tamaño, de manera que el tamaño de los nodos incrementa a medida que lo hace el *degree*.

Gen	Degree	Betweenness Centrality
INS	29	0,399408
CYP19A1	19	NS
FSHR	18	NS
CYP17A1	18	NS
LHCGR	17	NS

Tabla 2. Grado de conectividad (*Degree*) e interrelación (*Betweenness Centrality*) de los genes de la anovulación en los que se han encontrado diferencias significativas. NS = no significativo.

Tanto para el *degree* como para el *betweenness*, se han seleccionado los genes que presentan un *degree/betweenness* por encima de la suma de la media y la desviación estándar (Tabla 2). Los genes que cumplen con estos requisitos son *INS*, *CYP19A1*, *FSHR*, *CYP17A1* y *LHCGR* para el *degree*, mientras que solo el gen *INS* tiene un *betweenness* más cercano a 1.

4.1.4. Análisis de enriquecimiento funcional

Una vez obtenidas las redes biológicas, se determinaron los procesos biológicos en los que están implicados los genes relacionados con la anovulación. Para ello, utilizamos *GeneMANIA*, un complemento de [Cytoscape 3.7.0](#), ya que con [DAVID](#) solo se obtuvo un resultado significativo. Con *GeneMANIA* se detectaron hasta 150 funciones biológicas significativas, muchas de ellas relacionadas con el metabolismo y la síntesis de hormonas (Tabla 3), lo que muestra la importancia de la función hormonal en la anovulación.

Nº de genes	Proceso biológico	q-valor
19	Regulación del nivel de hormonas	1,7246E-17
15	Metabolismo hormonal	4,5619E-17
10	Biosíntesis de hormonas	7,3889E-14
9	Metabolismo hormonal a nivel celular	1,8793E-09
10	Biosíntesis de esteroides	2,3217E-09
6	Metabolismo de pregnenolona	1,9583E-07
10	Metabolismo de esteroides	6,3930E-07
6	Procesamiento de hormonas peptídicas	6,3930E-07
8	Diferenciación sexual	1,1953E-06
9	Proceso de desarrollo involucrado en la reproducción	3,0983E-06

Tabla 3. Primeros diez procesos biológicos estadísticamente significativos extraídos de *GeneMANIA* — complemento de [Cytoscape 3.7.0](#) — en los que podrían estar implicados los distintos genes de la anovulación.

GeneMANIA también nos ofrece la posibilidad de generar una red con las distintas interacciones que se establecen entre el conjunto de genes obtenidos a partir de [Phenopedia](#) (nodos en negro) y otros que son predichos por el programa a partir de estudios previos (nodos en gris) (Figura 5).

Genes potencialmente relacionados con el PCOS						
INS	PIK3R1	GNRH1	FMN2	PLIN1	DNAJB1	ALDH1B1
PPARG	GSK3B	HLA-DRB1	RPS6KA1	PON2	APOA1	FMO1
INSR	SOD2	APOB	CAT	POU1F1	HTR5A	DRD1
IRS1	POMC	CDKN2B	CHRNA6	FAM46A	IGFBP4	DRD5
AR	ADRB2	TMEM18	MFHAS1	PRKCI	IGFBP5	SSTR2
ADIPOQ	C9orf3	PPP1R3A	TLR2	PRLR	SIM1	trnD
FSHR	BMP15	IGF2R	ARRB2	IGFBP1	SLC6A3	TGFB3
FTO	GHRL	FAS	SMAD5	CYP2B6	PRKCZ	TNFRSF1B
LHCGR	STK11	UGT1A1	MEF2A	SMAD6	ADAMTS9	TPH1
CAPN10	PCSK1	GDF3	MEP1A	MAF	CHRNA10	TPO
SHBG	HSD3B2	CHFR	OPRK1	MC3R	PTH1R	TSC2
CYP11A1	IGF1	PTPN11	PCM1	GREB1	ZNF71	UCP3
SERPINE1	TP53	ROCK1	ADIPOR1	USP34	NEK4	TRIB3
CYP17A1	SMAD3	EBF2	PDYN	UGT2B4	VIP	BCAT1
CYP19A1	GNB3	BMP6	CDK12	POF1B	SSTR4	RBP4
TCF7L2	MTCH2	BRCA1	PEPD	INHBE	ACVR2B	BRD2
DENND1A	NEGR1	BRCA2	POLG	CASR	CBR3	LIPE
MTHFR	TOX3	SST	PRKCA	CBL	DOT1L	SMAD2
VDR	AGT	TGFBR1	CYP2C9	CCK	GABBR2	MMP1
THADA	AKT2	TLR4	CYP3A4	PLCZ1	INPPL1	OPRD1
PON1	YAP1	CEP70	APOC1	SMC3	IRF1	PIK3R2
TNF	CYP1B1	IKBKG	DBH	BMPR1B	KCNQ1	PLIN
IL6	EPHX1	EIF2B2	AKR1C2	trnK	IGFBP2	B3GNT5
IRS2	ACVR2A	EIF2B5	DRD2	STXBP1	IL6R	KISS1R
HSD11B1	VCAM1	KIR3DS1	COG2	TGFBR2	HMGCR	SUMO1P1
ACE	WFS1	LDLR	CDK5	FEM1C	HSD3B1	GDF6
ESR1	LIF	LPA	ADAMTS13	RPS6KB1	MTR	SCARB1
LHB	PPARD	ATM	CHRNA2	CCL2	NODAL	CD40LG
AKR1C3	SREBF2	NFKB1	CHRNA2	CCL5	PKD1	CYP3A7
CYP21A2	SORBS1	NUCB2	CHRNA2	INHA	PLCB3	OCTN3
ESR2	CRP	OGG1	NLRP5	BMP2	SH2B2	FRAT1
CYP1A1	HLA-A	TM7SF3	ANKK1	UGT2B7	GNA13	FLOT1
H6PD	INHBB	PIK3CA	FGFR1OP2	UGT2B7	SLC2A6	CHRNA5
IL1B	HSD17B2	SERPINA12	GATA6	SSTR1	CHRNA4	CHRNA5
FBN3	GYS2	CYP2A6	MSTN	trnR	CHRNA4	CHRNA5
FSHB	GDF9	ETV5	GH1	MAP3K7	CHRNA7	CHRNA5
HSD17B6	GDF9	ETV5	GNAS	TGFBR3	CHRNA7	CHRNA5
SRD5A1	FEM1B	LEFTY1	GSK3A	BMP4	CHRNA7	CHRNA5
AMH	PGR	AKAP11	GYS1	BMP4	CHRNA7	CHRNA5
IGF2	MTNR1B	CHRNA3	SEC16B	BMPR1A	CHRNA7	CHRNA5
AMHR2	GNRHR	CHRNA3	ACVRL1	SLC22A1	CHRNA7	CHRNA5
KCNJ11	HTR1A	CHRNA3	SSTR5	CASC1	CHRNA7	CHRNA5
IL1RN	INHBA	GJA4	SULT1A1	CHRNA9	CHRNA7	CHRNA5
IL10	ACVR1	BMP10	LEFTY2	PRKAA2	CHRNA7	CHRNA5
NR3C1	FABP4	SULT1A1	KIR2DL4	SLC2A4RG	CHRNA7	CHRNA5
ENPP1	AKT1	MCF2L2	KIR2DS4	NR0B2	CHRNA7	CHRNA5
HHEX	INSIG2	RAD54B	LPL	EIF2B4	CHRNA7	CHRNA5
FST	STAR	GATA4	MIR146A	NOG	CHRNA7	CHRNA5
MC4R	RAB5B	GDF2	PTPN1	MDM2	CHRNA7	CHRNA5
APOE	BDNF	HTR1B	PROK2	MIF	CHRNA7	CHRNA5
CDKAL1	CYP11B2	AQP8	BMP5	PDPK1	CHRNA7	CHRNA5
SRD5A2	FMR1	KCNJ1	BMP7	PIK3CB	CHRNA7	CHRNA5
SLC30A8	KISS1	MAOA	BMP8B	LMNA	CHRNA7	CHRNA5
F2	STS	MAOB	FOXL2	LPP	CHRNA7	CHRNA5
FEM1A	SGTA	NBN	SPAST	LRP5	CHRNA7	CHRNA5
TGFB1	SLC2A4	PPIF	MAPK14	LTA	CHRNA7	CHRNA5
LEP	GHR	KHDRBS3	GDF7	SMAD1	CHRNA7	CHRNA5
	SLC6A4	SLC12A7	CHRNA1	KRAS	CHRNA7	CHRNA5
	HMGA2	DNASE1L3		CGA	CHRNA7	CHRNA5

VEGFA	UCP2	AGTR1	CITED4	PROKR2	HTR2A	trnC
ICAM1	CGB5	FAIM2	DRD3	CYP2D6	HTR2C	TNFRSF11A
F5	IDE	LPIN1	ABCA1	INHBC	IGFALS	TAX1BP1
CNR1	NOS3	RHOQ	GLP1R	ITGB3	IGFBP3	PAPSS2
IL1A	IGF2BP2	ABCC8	GPX1	KAL1	IGFBP6	ACVR1B
NR1H3	GNPDA2	BMP3	ANXA5	LEPROTL1	INSL3	BUB3
IGF1R	SREBF1	SLC18A2	HIVEP1	GALT	INSRR	PPP1CA
ADRB3	TH	BMPR2	HSD17B4	OR7A15P	HP	DOK5
PPARGC1A	KCTD15	trnE	AFF2	GDF1	EME1	PTEN
SULT2A1	PLA2G7	PENK	VTLG	DDC	SLC2A12	TACC1
PTH	COMT	PLAU	FYN	EML1	CYP1A2	TCF4
LEPR	SMAD4	BCMO1	GDF10	FAAH	ADRBK2	TG
IL18	SUOX	PRKAG3	OPRM1	FGFR1	CHRNA4	TGFB2
RETN	SH2B1	PPARA	PDE8A	HLA-B	CCR3	NKX2-1
CAPN5	GNAQ	PRL	PLCB2	AIRE	FABP1	PKD2
SLC22A2						

Tabla 4. Lista de los 505 genes que podrían estar implicados en el Síndrome del Ovario Poliquístico (PCOS) obtenida a partir de [Phenopedia](#).

4.2.2. Análisis de interacción proteína-proteína

De los 505 genes obtenidos de [Phenopedia](#), se ha generado una red biológica a través de la plataforma [STRING](#) (Figura 6), donde se observan 492 nodos y 8718 conexiones (interacciones conocidas a partir de bases de datos y experimentos, interacciones predichas a partir de su cercanía física, de fusiones génicas y de coocurrencia génica, interacciones determinadas por minería de texto, por coexpresión o por homología en proteínas).

Por otra parte, no se ha conseguido establecer la conexión de los genes *R3HCC1*, *EML1*, *ZNF683*, *ZNF71*, *SLC2A6*, *CITED4*, *AQP8*, *NEK4* y *CHRNA10* con el resto de elementos que conforman la red.

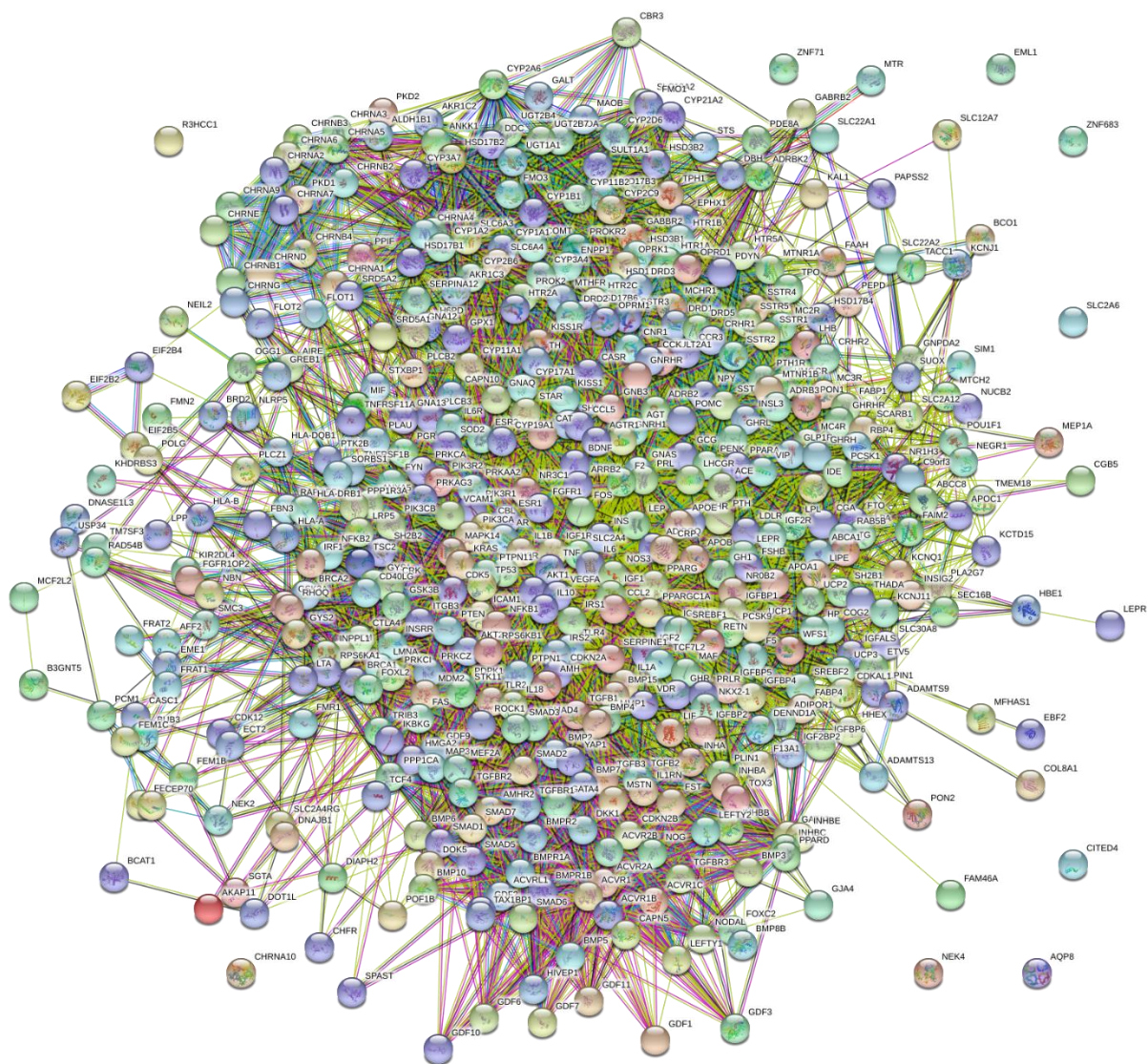


Figura 6. Red biológica generada con el programa [STRING](#), donde se muestran las distintas relaciones entre los 492 genes relacionados con el Síndrome del Ovario Poliquístico (PCOS).

4.2.3. Análisis de parámetros de centralidad

Como criterio para el análisis de los parámetros de centralidad *degree* y *betweenness*, se han seleccionado los genes que presentan un *degree/betweenness* por encima de la suma de la media y la desviación estándar. Los genes que cumplen con esta condición aparecen en la Tabla 5. Tanto para el *degree* como para el *betweenness*, el gen *INS* tiene los valores más altos, es decir, es el que establece más conexiones con otros genes y, además, conecta un gran número de genes separados (actúa como puente entre genes aparentemente no relacionados).

Gen	Degree	Gen	Degree	Gen	Betweenness Centrality
INS	247	TGFB1	79	INS	0,133220
AKT1	180	GNB3	79	AKT1	0,052019
IL6	159	SMAD3	77	TP53	0,046976
IGF1	156	CYP19A1	77	IGF1	0,030427
TP53	132	CCL2	76	IL6	0,026914
VEGFA	132	APOB	75	BDNF	0,026138
LEP	128	CAT	75	GSK3B	0,022436
TNF	125	CRP	75	ESR1	0,020944
POMC	117	NR3C1	75	VEGFA	0,019392
ESR1	112	KRAS	73	FOS	0,015855
FOS	112	SERPINE1	72	POMC	0,015447
PPARG	105	IL10	72	LEP	0,015270
BDNF	103	IGF2	71	AR	0,014856
IRS1	98	MAPK14	71	INSR	0,013784
ADIPOQ	96	BMP4	69	TNF	0,013026
AGT	90	TLR4	69	PIK3R1	0,011855
APOE	88	GSK3B	68	APOE	0,011295
SMAD2	86	SST	68	PIK3CA	0,011258
PRL	85	SLC2A4	68	SMAD2	0,010987
AR	85	TGFBR1	66	ATM	0,010957
PIK3CA	83	AGTR1	66		
SMAD4	83	F2	65		
IL1B	82	DRD2	65		
IGF1R	81	ADRB2	65		
GNRH1	80	GHRL	65		
PIK3R1	79	CCK	65		
NOS3	79	INSR	64		

Tabla 5. Grado de conectividad (*Degree*) e intermediación (*Betweenness Centrality*) de los genes altamente conectados en el Síndrome del Ovario Poliquístico (PCOS).

4.2.4. Análisis de enriquecimiento funcional

Para el análisis de los procesos biológicos, funciones moleculares y componentes celulares en los que estarían implicados los genes del PCOS se ha utilizado [DAVID](#). Teniendo en cuenta un valor de FDR de 0,05, se han encontrado 176 procesos biológicos significativos, 48 funciones moleculares significativas y 23 componentes celulares significativos. En la Tabla 6 y en la Figura 7 se muestran los diez primeros resultados significativos para cada uno de los análisis.

Nº de genes	Procesos biológicos	p-valor	FDR
37	Regulación positiva de la vía restringida de fosforilación de la proteína SMAD	3,8832E-46	7,1599E-43
35	Transducción de señal de la proteína SMAD	1,7775E-36	3,2774E-33
30	Vía de señalización BMP	1,7709E-25	3,2652E-22
23	Regulación de la cascada MAPK	1,0200E-23	1,8807E-20
91	Regulación positiva de la transcripción desde el promotor de la ARN polimerasa II	1,5830E-23	2,9187E-20
30	Vía de señalización del receptor del factor de crecimiento transformante beta	9,8266E-23	1,8118E-19
22	Desarrollo celular	1,9932E-22	3,6751E-19
49	Respuesta a drogas	2,2869E-22	4,2166E-19
27	Respuesta celular al estímulo de insulina	2,2289E-21	4,1096E-18
36	Envejecimiento	6,9457E-21	1,2806E-17
Nº de genes	Funciones moleculares	p-valor	FDR
32	Unión al receptor del factor de crecimiento transformante beta	1,4243E-39	2,2336E-36
42	Actividad del factor de crecimiento	1,3460E-27	2,1107E-24
42	Actividad de citoquinas	4,4895E-26	7,0402E-23
28	Actividad hormonal	3,8220E-20	5,9934E-17
16	Actividad del canal de catión selectivo activado por acetilcolina	2,0309E-18	3,1848E-15
16	Unión a la acetilcolina	2,9427E-17	4,6145E-14
15	Actividad del receptor de la acetilcolina	1,1907E-16	1,7764E-13
15	Actividad de los canales iónicos activados por ligando	4,0083E-13	6,2850E-10
14	Unión al receptor de insulina	8,7439E-13	1,3712E-09
9	Actividad serina/treonina quinasa del receptor proteico transmembrana	3,5425E-12	5,5552E-09
Nº de genes	Componentes celulares	p-valor	FDR
105	Espacio extracelular	8,6923E-24	1,2124E-20
107	Región extracelular	5,1385E-19	7,1672E-16
16	Complejo de canales controlados por acetilcolina	1,9870E-18	2,7715E-15
96	Componente integral de la membrana plasmática	1,4481E-17	2,0198E-14
192	Membrana plasmática	1,8928E-17	2,6401E-14
23	Complejo receptor	2,6215E-12	3,6564E-09
46	Superficie celular	1,4125E-11	1,9701E-08
18	Membrana de orgánulos	1,1343E-10	1,5821E-07
31	Cuerpo celular neuronal	1,6942E-09	2,3631E-06
25	Membrana postsináptica	2,2728E-09	3,1701E-06

Tabla 6. Primeros diez procesos biológicos, funciones moleculares y componentes celulares en los que podrían estar implicados los distintos genes del Síndrome del Ovario Poliquístico (PCOS). *FDR* = *False Discovery Rate*.

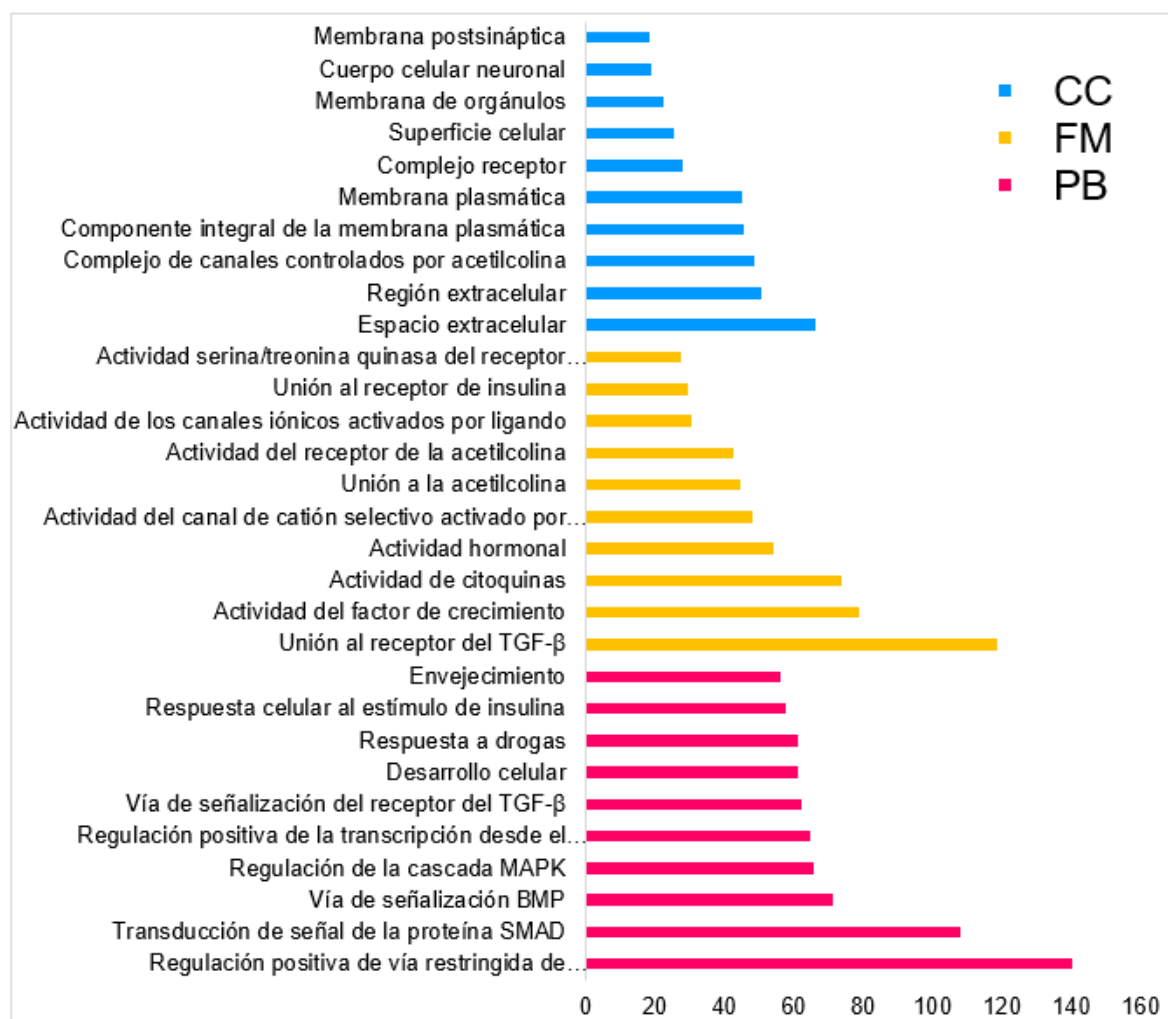


Figura 7. Principales procesos biológicos, funciones moleculares y componentes celulares relacionados con los genes implicados en el PCOS. Para cada elemento, se muestra el valor de Z-score, calculado a partir del $-\log_2$ del FDR.

4.3. Análisis de la intersección entre la anovulación y el PCOS

A partir de los resultados de [Phenopedia](#), de los 39 genes que se ha visto que pueden tener alguna relación con la anovulación, 37 de ellos también estarían implicados en el Síndrome del Ovario Poliquístico (había un total de 505 genes relacionados con esta alteración) (Figura 8). En total, un 94,9% de los genes de la anovulación son compartidos con el PCOS. El porcentaje restante corresponde a dos genes: *NOBOX* y *NEUROG3*.

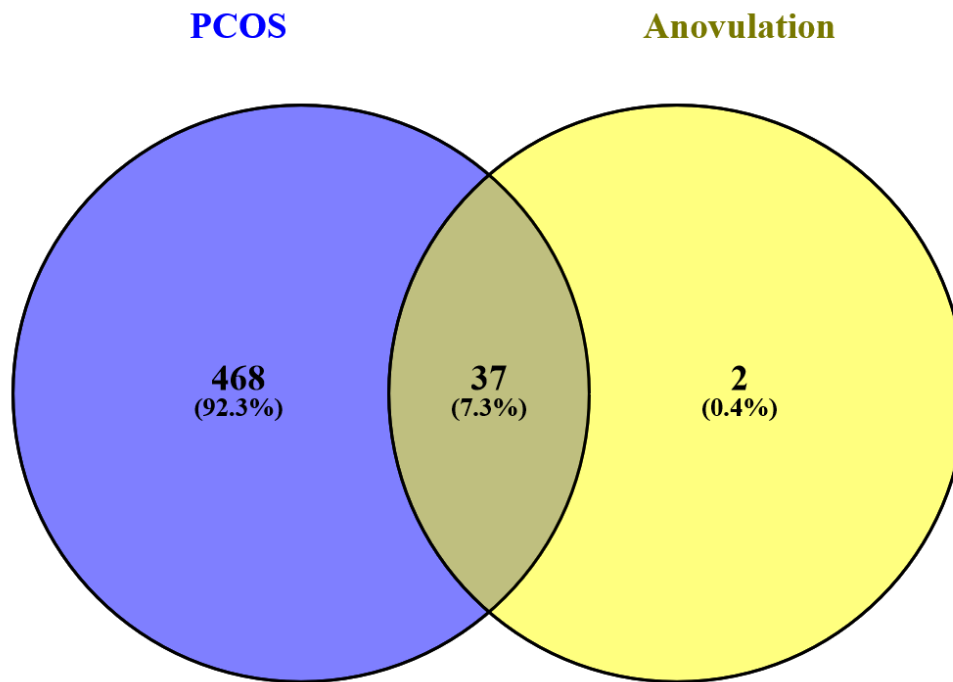


Figura 8. Intersección entre los conjuntos de genes implicados en el Síndrome del Ovario Poliquístico (PCOS) y la anovulación (*Anovulation*) obtenida con [VENNY 2.1.0](#) a partir de la lista de genes generada por [Phenopedia](#).

5. DISCUSIÓN

La anovulación es una de las causas más comunes de infertilidad en mujeres en edad reproductiva, por lo que es importante el estudio de esta alteración desde un punto de vista genético, no de forma aislada, sino del conjunto de los 39 genes encontrados que están potencialmente relacionados con la anovulación. Para ello, la Biología de Sistemas es una disciplina que nos ha permitido entender y explicar algunas de las conexiones que existen entre genes endocrinos, inmunológicos o incluso de regulación de otros genes. Así pues, hemos podido determinar que el gen de la insulina (*INS*), supone una parte muy importante de la red biológica de los genes de la anovulación, con un grado de conectividad diferencialmente significativo. Establece múltiples conexiones con numerosos genes, de manera que mutaciones en su secuencia podrían provocar la desconexión entre genes que, de no ser por *INS*, no estarían relacionados.

Algunos estudios de meta-análisis, como el de Song *et al.* (2014), han planteado que el polimorfismo en un minisatélite del gen *INS* involucrado en la resistencia a la insulina podría conducir a la anovulación. La insulina es una hormona secretada por las células beta del páncreas, y alteraciones en su concentración se ha visto que provocan cambios en la liberación pulsátil de GnRH y LH. En concreto, la hiperinsulinemia provoca una mayor amplitud en la liberación de la LH, es decir, aumenta el tiempo de su secreción entre pulsos. Aún no se conoce si esto es debido a la disminución de la GnRH y/o a una menor sensibilidad de la hipófisis a esta hormona. Lo que sí parece claro es que la insulina actuaría directamente en el hipotálamo, uniéndose a los receptores de la insulina de sus células (Sliwowska *et al.*, 2014).

Además, existen otros estudios realizados en humanos que también plantean una correlación positiva entre la anovulación en el PCOS y la disminución de la sensibilidad hipofisaria a la GnRH mediada por la hiperinsulinemia/resistencia a la insulina. Otra prueba de que los genes relacionados con la insulina estarían implicados en la anovulación sería la utilización de metformina — aumenta la sensibilidad a la insulina — como tratamiento para estas pacientes (Song *et al.*, 2014). Todo ello parece indicar que el gen *INS* tiene una gran importancia en el desarrollo de ciclos sin ovulación, de manera que la alteración del eje hipotálamo-hipófisis-gónada provocada por cambios en el gen *INS* podrían afectar a la

capacidad reproductiva de la mujer y conducir a la aparición del proceso anovulatorio.

También hemos detectado evidencias de la importancia del gen *CYP19A1*. Este gen codifica la enzima aromatasa, necesaria para la biosíntesis de estrógenos (cataliza la transformación de testosterona a estradiol) (Kaur *et al.*, 2018), mientras que la FSH es la encargada de inducir la actividad de esta enzima (Laven *et al.*, 2003). La disminución de la expresión de este gen (debido a un polimorfismo) provocaría el PCOS y, por tanto, la anovulación. Sin embargo, en el estudio de Kaur *et al.* (2018) no encontraron relación entre los polimorfismos de este gen y el PCOS. Esto seguramente se deba a que el PCOS es una enfermedad que depende de múltiples factores, por lo que el análisis y estudio de uno solo de forma aislada no sería suficiente para encontrar diferencias significativas. En ese mismo trabajo también se encontraron evidencias de que el aumento de la actividad del gen *CYP17A1*, el cual también se ha observado que presenta una elevada conectividad en la red biológica de la anovulación, puede conducir a la aparición del PCOS. Este gen se expresa en la glándula adrenal, las células de Leydig (en hombre) y las células de la teca del ovario, y codifica enzimas implicadas en la biosíntesis de andrógenos (Kaur *et al.*, 2018), lo cual parece estar relacionado con el hiperandrogenismo característico del PCOS y que conduciría a la anovulación.

Por otro lado, se han encontrado SNPs en el gen *FSHR* — gen del receptor de la hormona FSH — que podrían estar relacionados con niveles altos de FSH en sangre, debido a que disminuye su actividad de unión a la FSH (Laven *et al.*, 2003; Singhasena *et al.*, 2014). Se cree que estos SNPs también podrían interactuar con polimorfismos de otros genes implicados en la esteroidogénesis y foliculogénesis, pudiendo producirse la anovulación, aunque existe cierta controversia al respecto (Singhasena *et al.*, 2014). La LH y la hormona coriónica humana también intervienen en la síntesis de esteroides, las cuales ejercen su acción una vez se han unido al receptor de la LH. Esta proteína transmembrana está codificada por el gen *LHCGR*, demostrándose la implicación de este gen en la maduración folicular y la ovulación, por lo que la presencia de polimorfismos estaría relacionada con desórdenes ovulatorios (Cui *et al.*, 2015; Thathapudi *et al.*, 2015).

En cuanto al PCOS, hemos detectado la participación de un elevado número de genes en los distintos procesos biológicos, funciones moleculares y componentes celulares, lo que pone de manifiesto el hecho de que las distintas proteínas,

secuencias de regulación, hormonas, etc. trabajan de manera conjunta en el organismo, dependiendo el resultado de la acción de todos ellos. Esto se corrobora con nuestros resultados de interacción proteína-proteína, donde se puede observar una red biológica compleja con múltiples conexiones entre genes (Figura 6).

En el PCOS también se observa la gran importancia del gen *INS* en las interacciones entre genes, lo que ya nos empieza a indicar que, efectivamente, existe una estrecha relación entre esta enfermedad y la anovulación. Concretamente, en el análisis de los genes compartidos por ambas alteraciones, alrededor del 95% de los genes de la anovulación participan también en el PCOS. Solo los genes *NEUROG3* y *NOBOX* no se han incluido en la intersección entre las dos patologías, lo cual puede ser debido a que hasta la fecha no se han realizado estudios sobre la posible relación de estos genes con el PCOS, lo que podría resultar interesante para futuras investigaciones.

NEUROG3 es un gen proendocrino relacionado con la maduración de las células beta pancreáticas (Zhu *et al.*, 2017). Fallos en este gen supondrían la presencia de células beta alteradas y, por tanto, una formación y secreción anormal de la insulina. Esta relación existente entre los genes *NEUROG3* e *INS* se puede observar también en la Figura 4, donde el gen *NEUROG3* solo establece conexión con el gen *INS*.

Por su parte, el gen *NOBOX* codifica un factor de transcripción que tiene un papel esencial en la foliculogénesis. Mutaciones encontradas en este gen parecen estar asociadas a la insuficiencia ovárica prematura (Bouilly *et al.*, 2015), la cual es una de las causas de la anovulación, en concreto, de la anovulación tipo III. Esto explicaría por qué este gen no está relacionado con el PCOS, pero sí con la anovulación. Sin embargo, al tener un papel crucial en la maduración folicular, sería conveniente la realización futura de investigaciones que pudiesen comprobar su posible rol en el PCOS.

Por último, es importante destacar la necesidad de llevar a cabo más estudios experimentales y bibliográficos sobre la anovulación en los próximos años, debido a que es una de las principales causas de infertilidad femenina, suponiendo alrededor de un 30% de los casos diagnosticados en las clínicas. Esto permitiría que se pudiese extraer una mayor cantidad de resultados significativos mediante el uso de la Biología de Sistemas, lo que puede tener aplicaciones en el diagnóstico, seguimiento y desarrollo de posibles nuevos tratamientos.

6. CONCLUSIONES

- Se ha constatado la gran utilidad de las herramientas que ofrece la Biología de Sistemas para determinar qué genes son compartidos por la anovulación y el PCOS, tales como [Phenopedia](#) para la obtención de la lista de genes potencialmente relacionados con la anovulación y el PCOS, [STRING](#) y [Cytoscape](#) para la construcción de redes biológicas y *GeneMANIA* y [DAVID](#) para el análisis de enriquecimiento funcional.
- Se ha detectado la importancia del gen de la insulina (*INS*) en la anovulación, así como en el PCOS, de manera que mutaciones en este gen podrían conducir a alteraciones del ciclo ovulatorio.
- Se ha comprobado el papel crucial de la función hormonal en la anovulación.
- Se ha observado la existencia de un elevado grado de conectividad entre los genes del PCOS, utilizando procedimientos estadísticos y computacionales, y se ha puesto de manifiesto la implicación de numerosos genes en una misma función.
- Además, se destaca la estrecha relación existente entre la anovulación y el PCOS. Es más, proponemos que los genes *NOBOX* y *NEUROG3*, los cuales fueron hallados en la anovulación pero no en el PCOS, podrían estar también implicados en este síndrome y ser objeto de futuras investigaciones.
- El abordaje multidisciplinar de este Trabajo de Fin de Grado ha permitido integrar conocimientos de diferentes asignaturas estudiadas durante el Grado.

7. BIBLIOGRAFÍA

- Boccaletti S, Latora V, Moreno Y, Chavez M y Hwang DU. 2006. Complex networks: Structure and dynamics. *Physics Reports*, 424:175-308
- Bouilly J, Roucher-Boulez F, Gompel A, Bry-Gauillard H, Azibi K, Beldjord C, Dodé C, Bouligand J, Mantel AG, Hécart AC, Delemer B, Young J y Binart N. 2015. New NOBOX mutations identified in a large cohort of women with primary ovarian insufficiency decrease KIT-L expression. *The Journal of Clinical Endocrinology & Metabolism*, 100(3):994-1001
- Cui L, Li G, Zhong W, Bian Y, Su S, Sheng Y, Shi Y, Wei D, Zhang W, Zhao H y Chen ZJ. 2015. Polycystic ovary syndrome susceptibility single nucleotide polymorphisms in women with a single PCOS clinical feature. *Human Reproduction*, 30(3):732-6
- Delles C y Husi H. 2017. Systems Biology Approach in Hypertension Research. *Methods in Molecular Biology*, 1527:69-79
- Djordjevic M, Rodic A y Graovac S. 2019. From biophysics to 'omics and systems biology. *European Biophysics Journal*, in press.
- Eddy JA, Funk CC y Price ND. 2015. Fostering synergy between cell biology and systems biology. *Trends in Cell Biology*, 25(8):440-5
- Ghojavand S, Bagheri F y Mesrian Tanha H. 2019. Integrative meta-analysis of publically available microarray datasets of several epithelial cell lines identifies biological processes affected by silver nanoparticles exposure. *Comparative Biochemistry and Physiology - Part C: Toxicology & Pharmacology*, 216:67-74
- Guzzi PH, Agapito G, Milano M y Cannataro M. 2016. Methodologies and experimental platforms for generating and analysing microarray and mass spectrometry-based omics data to support P4 medicine. *Briefings in Bioinformatics*, 17(4):553-61
- Hakimi O y Cameron LC. 2017. Effect of Exercise on Ovulation: A Systematic Review. *Sports Medicine*, 47(8):1555-1567
- Huang da W, Sherman BT y Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4(1):44-57

- Kaur R, Kaur T y Kaur A. 2018. Genetic association study from North India to analyze association of CYP19A1 and CYP17A1 with polycystic ovary syndrome. *Journal of Assisted Reproduction and Genetics*, 35(6):1123-1129
- Laven JS, Mulders AG, Suryandari DA, Gromoll J, Nieschlag E, Fauser BC y Simoni M. 2003. Follicle-stimulating hormone receptor polymorphisms in women with normogonadotropic anovulatory infertility. *Fertility and Sterility*, 80(4):986-92
- Li PC. 2016. Overview of Microarray Technology. *Methods in Molecular Biology*, 1368:3-4
- Li RH y Ng EH. 2012. Management of anovulatory infertility. *Best Practice & Research: Clinical Obstetrics & Gynaecology*, 26(6):757-68
- Lie Fong S, Schipper I, Valkenburg O, de Jong FH, Visser JA y Laven JS. 2015. The role of anti-Müllerian hormone in the classification of anovulatory infertility. *European Journal of Obstetrics & Gynecology & Reproductive Biology*, 186:75-9
- Liu XY, Yang YJ, Tang CL, Wang K, Chen JJ, Teng XM, Ruan YC y Yang JZ. 2019. Elevation of antimüllerian hormone in women with polycystic ovary syndrome undergoing assisted reproduction: effect of insulin. *Fertility and Sterility*, 111(1):157-167
- Marzancola MG, Sedighi A y Li PC. 2016. DNA Microarray-Based Diagnostics. *Methods in Molecular Biology*, 1368:161-78
- Meier RK. 2018. Polycystic Ovary Syndrome. *Nursing Clinics of North America*, 53(3):407-420
- Nielsen J. 2017. Systems Biology of Metabolism. *Annual Review of Biochemistry*, 86:245-275
- Oliveros JC. 2007. VENNY. An interactive tool for comparing lists with Venn Diagrams. <http://bioinfogp.cnb.csic.es/tools/venny/index.html>
- Oyola MG y Handa RJ. 2017. Hypothalamic-pituitary-adrenal and hypothalamic-pituitary-gonadal axes: sex differences in regulation of stress responsivity. *Stress*, 20(5):476-494
- Patel S. 2018. Polycystic ovary syndrome (PCOS), an inflammatory, systemic, lifestyle endocrinopathy. *The Journal of Steroid Biochemistry and Molecular Biology*, 182:27-36

- Schiller CE, Johnson SL, Abate AC, Schmidt PJ y Rubinow DR. 2016. Reproductive Steroid Regulation of Mood and Behavior. *Comprehensive Physiology*, 6(3):1135-60
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B y Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11):2498-504
- Singhasena W, Pantasri T, Piromlertamorn W, Samchimchom S y Vutyavanich T. 2014. Follicle-stimulating hormone receptor gene polymorphism in chronic anovulatory women, with or without polycystic ovary syndrome: a cross-sectional study. *Reproductive Biology and Endocrinology*, 12:86
- Sliwowska JH, Fergani C, Gawalek M, Skowronska B, Fichna P y Lehman MN. 2014. Insulin: its role in the central control of reproduction. *Physiology & Behavior*, 133:197-206
- Song LY, Luo JR, Peng QL, Wang J, Xie L, He Y, Li S y Qin X. 2014. Lack of association of INS VNTR polymorphism with polycystic ovary syndrome: a meta-analysis. *Journal of Assisted Reproduction and Genetics*, 31(6):675-81
- Stamou MI, Cox KH y Crowley WF Jr. 2015. Discovering Genes Essential to the Hypothalamic Regulation of Human Reproduction Using a Human Disease Model: Adjusting to Life in the "-Omics" Era. *Endocrine Reviews*, 36(6):603-21
- Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguéz P, Doerks T, Stark M, Müller J, Bork P, Jensen LJ y von Mering C. 2011. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Research*, 39(Database issue):D561–D568
- Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ y von Mering C. 2015. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Research*, 43(Database issue):D447-52
- Thathapudi S, Kodati V, Erukkambattu J, Addepally U y Qurratulain H. 2015. Association of luteinizing hormone chorionic gonadotropin receptor gene polymorphism (rs2293275) with polycystic ovarian syndrome. *Genetic Testing and Molecular Biomarkers*, 19(3):128-32

- van den Oord EJ. 2008. Controlling false discoveries in genetic studies. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, Part B*, 147B(5):637-44
- Vargas E, Díaz L y Esteban FJ. 2019. Integración y análisis de datos mediante Biología de Sistemas. *Biología Molecular y celular. Vol 1. Técnicas y Fundamentos*. Universitaria Española
- Vargas E, Esteban FJ y Altmäe S. 2018. Computational Approaches in Reproductomics. Chapter 18. *Academic Press*, 347-383
- Von Hofe J y Bates GW. 2015. Ovulation induction. *Obstetrics and Gynecology Clinics of North America*, 42(1):27-37
- Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes CT, Maitland A, Mostafavi S, Montojo J, Shao Q, Wright G, Bader GD y Morris Q. 2010. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Research*, 38(Web Server issue):W214-20
- Yu W, Clyne M, Khoury MJ y Gwinn M. 2010. Phenopedia and Genopedia: disease-centered and gene-centered views of the evolving knowledge of human genetic associations. *Bioinformatics*, 26(1):145-6
- Zhu Y, Liu Q, Zhou Z y Ikeda Y. 2017. PDX1, Neurogenin-3, and MAFA: critical transcription regulators for beta cell development and regeneration. *Stem Cell Research & Therapy*, 8(1):240