

# **DISECCIÓN MOLECULAR DE LA ETIOLOGÍA DE LA FALLA OVÁRICA PREMATURA POR SECUENCIACIÓN DIRECTA Y NGS: ESTUDIO DE LA IMPLICACIÓN DE LOS GENES *CITED2*, *CDKN1B*, *FOXO4*, *BMP15* y *ADAMTS19***

## **RESUMEN**

Entre las causas femeninas de infertilidad, la Falla Ovárica Prematura (FOP) es frecuente, ya que afecta al 1% de las mujeres menores de 40 años de edad (Coulam et al, 1986). La FOP se define clínicamente por amenorrea de más de 6 meses de duración, niveles plasmáticos elevados de gonadotrofinas y deficiencia de esteroides sexuales antes de los 40 años de edad. La FOP puede ocurrir por una variedad de mecanismos que conducen a la disminución del número de folículos, al incremento de la atresia folicular o a la falla en la maduración de los folículos (Beck-Peccoz et al, 2006).

La FOP se ha asociado a varias etiologías, pero en más del 80% de los casos, no se conoce la causa de la enfermedad. Esta alta prevalencia de casos idiopáticos sugiere la implicación de factores genéticos y ambientales (Maclaran et al, 2011). La complejidad de los procesos fisiológicos, celulares y genéticos subyacentes a la determinación sexual, la foliculogénesis y la ovulación sugiere la participación de numerosos genes regulados de una manera precisa en cada etapa del desarrollo (Laissue, 2015). En este contexto, es fundamental integrar el alcance de las distintas técnicas para la búsqueda de factores genéticos etiológicos de la FOP. En efecto, tanto la secuenciación directa como las aproximaciones multigénicas para el análisis genómico a gran escala son aproximaciones complementarias para este ejercicio.

La iniciativa principal de este trabajo doctoral consistió en realizar aproximaciones monogénicas (secuenciación de Sanger) y multigénicas *Next generation sequencing* (NGS) para la identificación de nuevos genes asociados a FOP. La identificación de nuevos genes y de variantes potencialmente deletéreas, así como su validación a través de estudios funcionales y análisis poblacionales, permitió aportar nuevo conocimiento en la etiología

molecular de la FOP. De manera más precisa, presentaré los cinco artículos científicos derivados de los resultados de este trabajo doctoral. Estos se encuentran relacionados en los anexos que se mencionan a continuación:

1. Anexo 5.1: Ojeda D, Lakhal B, **Fonseca DJ**, Braham R, Landolsi H, Mateus H, and Laissue P. (2011). Sequence analysis of the CDKN1B gene in patients with premature ovarian failure reveals a novel mutation potentially related to the phenotype. *Fertility and Sterility*, 9: 2658–60.
2. Anexo 5.2: **Fonseca DJ**, Garzón E, Lakhal B, Braham R, Ojeda D, Elhgezal H, Saad A, Restrepo CM, and Laissue P. (2012). Screening for mutations of the *FOXO4* gene in premature ovarian failure patients. *Reproductive Biomedicine Online*, 24: 339-41.
3. Anexo 5.3: **Fonseca DJ**, Ojeda D, Lakhal B, Braham R, Eggers S, Turbitt E, White S, Gorver S, Warne G, Zacharin M, Nevin A, Landolsi H, Elghezal H, Saad A, Restrepo CM, Fellous M, Sinclair A, Koopman and Laissue P. (2012). *CITED2* mutations potentially cause idiopathic premature ovarian failure. *Translational Research*, 160:384-88.
4. Anexo 5.5: **Fonseca DJ**, Ortega-Recalde O, Esteban-Perez C, Moreno-Ortiz H, Patiño LC, Bermúdez OM, Ortiz AM, Restrepo CM, Lucena E, and Laissue P.(2014). *BMP15* c.-9C>G promoter sequence variant may contribute to the cause of non-syndromic premature ovarian failure. *Reproductive Biomedicine Online*, 29:627-33.
5. Anexo 5.6: **Fonseca DJ**, Patiño LC, Suárez YC, de Jesús Rodríguez A, Mateus HE, Jiménez KM, Ortega-Recalde O, Díaz-Yamal I, and Laissue P.(2015).Next generation sequencing in women affected by nonsyndromic premature ovarian failure dispaly new potential causative genes and mutations. *Fertility and Sterility*, 104:154-62.

Adicionalmente, presento dos anexos que contienen resultados que hasta el momento no han sido publicados y que corresponden a las siguientes preguntas científicas:

1. Anexo 5.4. Son las variantes de secuencia del promotor y del ORF de *BMP15* una causa de FOP no-sindrómica e idiopática?
2. Anexo 5.7 ¿Es la mutación p.Thr943Ile de *ADAMTS19* causal de FOP no-sindrómica e idiopática?

He incluido las conclusiones y las perspectivas derivadas del presente trabajo doctoral en la parte final de este manuscrito.

Cabe anotar que los resultados presentados en esta tesis hacen parte del trabajo merecedor del premio a la investigación en ciencias básicas de la Academia Nacional de Medicina 2015, otorgado al Dr Paul Laissue Hormaza y a Dora Janeth Fonseca Mendoza.

Por último, estimo importante mencionar que durante el tiempo transcurrido en el desarrollo del doctorado, y bajo la dirección del Dr Paul Laissue, he participado en varias investigaciones adicionales cuyos resultados están publicados en **8 artículos científicos adicionales** (ver lista en el anexo 5.8). Estos estudios me permitieron abordar temáticas y técnicas diferentes que han complementado mi formación doctoral.

# **Molecular dissection of premature ovarian failure aetiology using direct sequencing and NGS approaches: *CITED2*, *CDKN1B*, *FOXO4*, *BMP15* and *ADAMTS19* genes**

## **Summary**

Premature ovarian failure (POF) is a frequent disease, affecting 1% of women under 40 years old (Coulam et al, 1986). Clinically, POF women are affected by primary or secondary amenorrhoea and increased plasma levels of gonadotrophins (e.g. FSH and LH). POF can be originated from a variety of mechanisms leading to a decrease in the number of follicles, to the increase of follicular atresia or to follicle maturation failure (Beck-Peccoz et al, 2006).

POF has been associated with different etiologies, but in the majority of the cases the cause remains unknown. The prevalence of idiopathic cases suggests the existence of genetic and environmental aetiological factors (Maclaran et al, 2011). The inherent molecular complexity of sex determination, folliculogenesis and ovulation underlines that hundreds of genes might be related to POF phenotype (Laissue, 2015). In this context, direct sequencing of single genes and large scale genomic approaches are complementary tools for identifying new POF molecular actors. The present PhD thesis focused on the use of direct sequencing, NGS and functional *in vitro* tests allowing the identification and validation of new POF causative sequence variants.

Five papers describing these approaches are presented in the following manuscript as the basis of the experimental work performed during my PhD training:

- Annex 5.1: Ojeda D, Lakhali B, **Fonseca DJ**, Braham R, Landolsi H, Mateus H, and Laissue P. (2011). Sequence analysis of the CDKN1B gene in patients with premature ovarian failure reveals a novel mutation potentially related to the phenotype. *Fertility and Sterility*, 9: 2658–60.

- Annex 5.2: **Fonseca DJ**, Garzón E, Lakhali B, Braham R, Ojeda D, Elhgezal H, Saad A, Restrepo CM, and Laissue P. (2012). Screening for mutations of the *FOXO4* gene in premature ovarian failure patients. *Reproductive Biomedicine Online*, 24: 339-41.
- Annex 5.3: **Fonseca DJ**, Ojeda D, Lakhali B, Braham R, Eggers S, Turbitt E, White S, Gorver S, Warne G, Zacharin M, Nevin A, Landolsi H, Elghezal H, Saad A, Restrepo CM, Fellous M, Sinclair A, Koopman and Laissue P. (2012). *CITED2* mutations potentially cause idiopathic premature ovarian failure. *Translational Research*, 160:384-88.
- Annex 5.5: **Fonseca DJ**, Ortega-Recalde O, Esteban-Perez C, Moreno-Ortiz H, Patiño LC, Bermúdez OM, Ortiz AM, Restrepo CM, Lucena E, and Laissue P.(2014). *BMP15* c.-9C>G promoter sequence variant may contribute to the cause of non-syndromic premature ovarian failure. *Reproductive Biomedicine Online*, 29:627-33.
- Annex 5.6: **Fonseca DJ**, Patiño LC, Suárez YC, de Jesús Rodríguez A, Mateus HE, Jiménez KM, Ortega-Recalde O, Díaz-Yamal I, and Laissue P.(2015). Next generation sequencing in women affected by nonsyndromic premature ovarian failure new potential causative genes and mutations. *Fertility and Sterility*, 104:154-62.

Two further annexes have been included which describe unpublished results. This material responds to the following scientific questions:

1. Annex 5.4. Are *BMRP15* promotor and ORF variants related to non-syndromic POF?
2. Annex 5.7. Is ADAMTS19 p.Thr943Ile mutation causative of non-syndromic POF?

The final section of the present manuscript includes conclusions and perspectives issued from the present work. Taken together, the results presented here received the National Academy of Sciences Award (2015). Finally, it is important to notice that during my PhD training program I worked in further parallel projects, which led us to publish 8 additional

articles (annex 5.8). This work described different topics but help me to understand other scientific features.

# **DISECCIÓN MOLECULAR DE LA ETIOLOGÍA DE LA FALLA OVÁRICA PREMATURA POR SECUENCIACIÓN DIRECTA Y NGS: ESTUDIO DE LA IMPLICACIÓN DE LOS GENES *CITED2*, *CDKN1B*, *FOXO4*, *BMP15* y *ADAMTS19***

## **1. INTRODUCCIÓN**

La infertilidad humana ha sido reconocida por la Organización Mundial de la Salud (OMS) como un problema mundial de salud pública. Esta patología se define como la incapacidad de conseguir un embarazo después de por lo menos un año de relaciones sexuales sin medidas anticonceptivas. La estimación de la prevalencia de la infertilidad realizada en varias poblaciones en el mundo ha indicado rangos entre el 3.5% y el 6.7% en naciones desarrolladas y entre un 6.9% y un 9.3% en naciones menos desarrolladas (Boivin et al, 2007). En los últimos años, la incidencia de la infertilidad humana ha aumentado sustancialmente. Se estima que cerca de 186 millones de personas son infértiles, lo que afecta cerca del 8% al 12% de las parejas en edad reproductiva (Inhorn et al, 2015).

Entre las causas femeninas de infertilidad, la Falla Ovárica Prematura (FOP) es frecuente ya que afecta al 1% de las mujeres menores de 40 años de edad (Coulam et al, 1986). La FOP se define clínicamente por amenorrea de más de 6 meses de duración, niveles plasmáticos elevados de gonadotrofinas y deficiencia de esteroides sexuales antes de los 40 años de edad. La FOP implica un amplio espectro de consecuencias físicas y psicológicas debido a los síntomas y a los efectos de largo plazo de la deficiencia de los esteroides sexuales (Maclaran et al, 2010). La FOP puede ocurrir por una variedad de mecanismos que conducen a la disminución del número de folículos, al incremento de la atresia folicular o a la falla en la maduración de los folículos (Beck-Peccoz et al, 2006).

La FOP se ha asociado a varias etiologías y el fenotipo puede ser consecuencia, entre otras, de alteraciones autoinmunes, infecciosas, metabólicas, genéticas o iatrogénicas. Sin embargo, en más del 80% de los casos no se conoce la causa de la enfermedad. Esta alta prevalencia de casos idiopáticos sugiere la implicación de factores genéticos y ambientales (Maclaran et al, 2011).

Existen numerosas evidencias de los mecanismos genéticos en la patogénesis de la FOP idiopática. Los defectos ováricos en pacientes afectadas para el síndrome de Turner o con anormalidades cromosómicas del X indican el rol de este cromosoma en la función ovárica. La alta heredabilidad estimada para la edad de inicio de la menopausia, la alta incidencia de FOP en formas familiares (4-31%) y la existencia de árboles genealógicos que indican modos de herencia dominante o recesiva permiten considerar que la FOP es una enfermedad compleja multifactorial en la que el fenotipo es el reflejo de la acción aditiva de múltiples alelos (Toniolo, 2006; Laissue, 2015). Los genes que participan en procesos complejos como la determinación y la diferenciación sexual, la foliculogénesis y la ovulación deben estar regulados de manera precisa para que cada etapa se desarrolle de manera adecuada. Desde la perspectiva de la identificación de genes involucrados en el fenotipo de la FOP los modelos animales KO (*Knock out*), KI (*Knock in*) o las mutaciones naturales en especies como la oveja (Galloway et al, 2000) han permitido candidatizar algunos de ellos como responsables de la etiología molecular de la enfermedad. Las mutaciones puntuales en algunos genes autosómicos y ligados al X se han relacionado con la etiopatología (Laissue et al. 2006, 2008). Sin embargo, la complejidad de los procesos fisiológicos, celulares y genéticos subyacentes a la determinación sexual, la foliculogénesis y la ovulación sugiere la participación de numerosos genes regulados de una manera precisa en cada etapa del desarrollo. De manera similar a lo que ocurre en otras enfermedades complejas, la FOP puede ser el resultado de efectos sinérgicos/compensatorios de los polimorfismos de nucleótido simple (SNP). De esta manera, el fenotipo resultante de la FOP puede ser considerado un rasgo cuantitativo que resulta de la interacción de numerosas regiones del genoma (Laissue, 2015).

Esta situación se ha documentado extensamente en los últimos años puesto que una cantidad importante de genes candidatos se han propuesto (Matzuk et al, 2008; Fonseca et al, 2015; Fonseca et al, 2012; Caburet et al, 2012; Ojeda et al, 2011; Lakhali et al, 2008; Laissue et al, 2008; Laissue et al, 2007; Laissue et al, 2006; De Baere et al, 2005). Por consiguiente, se han desarrollado varios proyectos por aproximaciones monogénicas en genes candidatos de la FOP mediante secuenciación directa. Los hallazgos derivados de estos análisis han descrito pocas mutaciones, que explican la causa molecular de un bajo porcentaje de casos de FOP. De manera particular, las mutaciones en los genes *FSHR*,

*BMPR2, LHCGR, NR5A1, NOBOX, FOXL2, FIGLA, BMP15, NANOS3* y *STAG3*, han sido validados funcionalmente como causales de FOP no sindrómica (Laissue, 2015; Patiño et al, 2016; Patiño et al, 2017; Fonseca et al, 2014; Laissue et al, 2009).

Es posible intuir que con esta aproximación metodológica se cuenten con pocos marcadores moleculares de esta patología. Este escenario puede ser el resultado de varios factores: en primer lugar la poca eficiencia de la técnica de Sanger para el análisis de regiones genómicas extensas. En segundo lugar, la baja frecuencia de FOP familiar que no permite realizar aproximaciones cartográficas como el análisis de ligamiento (conducentes a identificar regiones cromosómicas candidatas). Tercero, la complejidad molecular inherente a la reproducción de los mamíferos que debe ser considerada como un rasgo resultante de la interacción de cientos de genes.

Las aproximaciones de siguiente generación (NGS) para secuenciar todas las regiones codificantes (exoma) han sido ampliamente descritas y aplicadas para la identificación de nuevos genes, principalmente en enfermedades recesivas monogénicas. Esta metodología complementa estrategias de identificación como el análisis de ligamiento, el clonaje posicional a partir de modelos de roedores, e incluso la secuenciación de primera generación de genes candidato. Además, ha permitido abordar enfermedades frecuentes de transmisión mendeliana a partir de un número relativamente bajo de pacientes (Biesecker, 2010; Ellard et al, 2014; Frebourg, 2014).

Durante los últimos años, se ha incrementado de manera exponencial las publicaciones que identifican genes causales de diferentes síndromes por técnicas de NGS. Recientemente, nuestro grupo de investigación ha publicado la aplicación de NGS para el diagnóstico molecular de cuatro enfermedades monogénicas (paquidermoperiostosis, xeroderma pigmentoso, lipofuscinosis y síndrome de sobrecrecimiento ) (Diggle et al, 2012; Ortega-Recalde et al, 2013 ; Patiño et al, 2014; Ortega-Recalde et al, 2015; Fonseca et al, 2015; Laissue et al, 2015). Sin embargo, los avances en los métodos de secuenciación no solo han logrado identificar variaciones en los genes candidato relacionados con patologías monogénicas. De esta forma, se ha iniciado el análisis simultáneo de múltiples variaciones que responden al carácter cuantitativo de los fenotipos asociados a enfermedades complejas.

En lo concerniente a la aplicación de NGS para el estudio de enfermedades complejas, se ha propuesto que los ensayos de captura dirigidos a regiones específicas del genoma y la secuenciación de DNA en “pools” de un gran número de pacientes, pueden usarse para identificar variantes raras asociadas al fenotipo (Casals et al, 2012). Nejentsev et al. por ejemplo, utilizaron la plataforma 454 para analizar mediante *resequencing* la región codificante de 10 genes candidatos de diabetes tipo 1 en 480 pacientes y en 480 controles. Identificaron cuatro SNP de baja frecuencia (MAF<3%) asociados a la enfermedad (Nejentsev et al, 2009). En otro estudio, el *resequencing* de 103 genes candidatos realizado en 103 casos y en 43 controles permitió identificar mutaciones raras, que a través de estudios funcionales se validaron como causales de la deficiencia del complejo I (Calvo et al, 2010). Se ha sugerido que una alternativa en el análisis de enfermedades complejas mediante NGS es el *resequencing* de genes candidatos en familias con varios afectados o en pacientes con fenotipos extremos (Cirulli et al, 2010). En algunas enfermedades complejas como el autismo, se han identificado mediante NGS variantes de alta penetrancia relacionadas con el desarrollo del fenotipo (Pinto et al, 2010). Estos hallazgos han permitido plantear una hipótesis genética que sugiere que en las enfermedades complejas el fenotipo puede ser atribuido al efecto de algunas mutaciones con alta penetrancia (Kilpinen et al, 2013). Kryukov et al. han propuesto que el acúmulo de variantes no-sinónimas con un efecto deletéreo leve es la base genética de las enfermedades complejas (Kryukov et al, 2007). Se ha descrito que en el humano cerca del 20% de las mutaciones *missense* nuevas conducen a una pérdida de la función, que alrededor del 27% son neutrales y que el 53% restante tienen efectos deletéreos poco drásticos (Kryukov et al, 2007). En este contexto, la metodología de NGS permitiría identificar este tipo de variantes de baja frecuencia. Varios estudios han determinado que en los fenotipos de rasgos complejos se presenta un exceso de variantes no-sinónimas raras en genes candidatos causales de la enfermedad (Kryukov et al, 2007).

Alternativamente a la secuenciación del exoma, la aplicación de NGS a la selección genómica basada en microarreglos (MGS), permite el enriquecimiento de secuencias predefinidas. La MGS permite analizar paneles de genes específicos relacionados con enfermedades complejas (Okou et al, 2007). Esta aproximación fue utilizada en este trabajo de tesis para la identificación de nuevas variantes en nuevos genes FOP (ver más adelante).

En este contexto, es fundamental integrar la utilidad de las distintas técnicas para la búsqueda de factores genéticos etiológicos de la FOP. En efecto, tanto la secuenciación directa como las aproximaciones multigénicas de análisis genómico a gran escala son complementarias. Puede afirmarse que las técnicas multigénicas permiten efectuar un tamizaje a gran escala que luego debe ser validado por secuenciación directa en un número más importante de pacientes de grupos étnicos específicos.

La iniciativa de este trabajo doctoral consistió en realizar aproximaciones monogénicas (secuenciación de Sanger) y multigénicas (*NGS*) para la identificación de nuevos genes FOP. Además, esta aproximación constituyó una etapa necesaria para la comprensión de los mecanismos moleculares causales de la enfermedad. La identificación de nuevos genes y de variantes potencialmente deletéreas, así como su validación a través de estudios funcionales y análisis poblacionales, permitió aportar nuevo conocimiento en la etiología molecular de la FOP. Los resultados fueron publicados en cinco artículos científicos, que se anexan al presente documento (ver anexos 5.1, 5.2, 5.3, 5.5 y 5.6).

## **2. JUSTIFICACIÓN**

La infertilidad ha sido reconocida por la OMS como un problema de salud pública que impacta a millones de personas alrededor del mundo. Cerca del 20% de las parejas en edad reproductiva son incapaces de lograr una concepción exitosa lo que genera consecuencias físicas y psicológicas impactantes (World Health Organization, Gender and Genetics, 2004). Cerca del 30% de todos los casos de infertilidad son considerados como idiopáticos, lo que subraya la importancia de factores genéticos potencialmente asociados así como de factores ambientales subyacentes.

La FOP es una patología que afecta una proporción importante de mujeres en el mundo y es considerada una de las causas más frecuentes de infertilidad femenina. Sus consecuencias a largo plazo no solo se limitan a problemas reproductivos sino también se asocian con el desarrollo de patologías como desórdenes cardiovasculares, osteoporosis y efectos psicológicos (Maclaran et al, 2010). La mayoría de los casos se consideran idiopáticos, lo que sugiere causas genéticas subyacentes. El asesoramiento genético es actualmente recomendado cuando se sospecha o se identifica una paciente con FOP, de causa genética.

La enfermedad en los casos de insuficiencia ovárica no puede ser tratada si el diagnóstico se efectúa luego de que ha ocurrido la depleción folicular completa. En los casos en que se identifica una paciente con una mutación en un gen causal o de susceptibilidad de FOP se puede hacer una intervención preventiva. Es entonces plausible que la consejería genética se dirija hacia la necesidad de preservación de oocitos o hacia la búsqueda de una rápida concepción.

Por último, hasta la fecha del inicio de este trabajo doctoral no se habían realizado abordajes NGS de patologías de la reproducción, por lo cual se desconocía su eficiencia para el estudio de estas enfermedades. En la actualidad, las pacientes con FOP no cuentan con un método de diagnóstico molecular fiable. Estimo que los resultados del presente estudio son el fundamento de nuevas opciones diagnósticas, pronósticas y terapéuticas de la FOP.

### **3. OBJETIVOS GENERALES.**

1. Ofrecer a la comunidad médica y científica nuevos elementos para la comprensión de los mecanismos genéticos responsables del fenotipo FOP a partir de aproximaciones monogénicas (secuenciación directa) y multigénicas (secuenciación de siguiente generación-NGS).
2. Identificar y validar funcionalmente nuevas mutaciones y genes causales de FOP.
3. Determinar si la aproximación NGS puede ser utilizada como un método diagnóstico de la FOP.

### **4. OBJETIVOS ESPECÍFICOS.**

1. Secuenciar por técnica de Sanger la región codificante de los genes *CDKN1B*, *CITED2*, *FOXO4* y *BMP15* en pacientes FOP.
2. Determinar si las mutaciones de la región codificante de *CITED2*, *CDKN1B*, *FOXO4* y del promotor de *BMP15* son una causa frecuente de FOP.
3. Establecer funcionalmente si la variante *BMP15* c.-9C>G está asociada al fenotipo FOP.

4. Identificar por NGS nuevas variantes potencialmente causales del fenotipo FOP no sindrómica.
5. Validar mediante estudios funcionales si la mutación c.2828C>T, p.Thr943Ile de *ADAMTS19* es causal de FOP.

## 5. PREGUNTAS CIENTÍFICAS E HIPÓTESIS

### 1 ¿Son las mutaciones de *CDKN1B* una causa frecuente de FOP no-sindrómica e idiopática?

**Hipótesis:** las mutaciones de *CDKN1B* en humanos podrían generar un fenotipo similar al observado en ratones transgénicos. De esta forma mujeres FOP portadoras de mutaciones en la región codificante de *CDKN1B* podrían presentar una activación prematura y masiva del pool de folículos primordiales.

**Resultados publicados en:** Ojeda D, Lakhali B, Fonseca DJ, Braham R, Landolsi H, Mateus H, and Laissue P. (2011). Sequence analysis of the *CDKN1B* gene in patients with premature ovarian failure reveals a novel mutation potentially related to the phenotype. *Fertility and Sterility*, 9: 2658–60 (ver Anexo 5.1).

### 2. ¿Son las mutaciones de *FOXO4* una causa frecuente de FOP no sindrómica e idiopática?

**Hipótesis:** las mutaciones de *FOXO4* en humanos podrían generar un fenotipo similar al observado en ratones transgénicos. Las mujeres FOP portadoras de mutaciones en la región codificante de *FOXO4* podrían modificar la regulación de genes blanco como por ejemplo *CDKN1B*.

**Resultados publicados en:** Fonseca DJ, Garzón E, Lakhali B, Braham R, Ojeda D, Elhgezal H, Saad A, Restrepo CM, and Laissue P. (2012). Screening for mutations of the *FOXO4* gene in premature ovarian failure patients. *Reproductive Biomedicine Online*, 24: 339-41 (ver anexo 5.2).

### **3. ¿Son las mutaciones de *CITED2* una causa frecuente de FOP no-sindrómica e idiopática?**

**Hipótesis:** las mutaciones de *CITED2* en humanos podrían generar un fenotipo similar al observado en el modelo murino *Cited2*<sup>-/-</sup> hipomórfico para *Sfl*. Por consiguiente, mujeres FOP portadoras de mutaciones no sinónimas en *CITED2* podrían presentar una disminución transitoria de la expresión de genes pro-ováricos.

**Resultados publicados en: Fonseca DJ**, Ojeda D, Lakhali B, Braham R, Eggers S, Turbitt E, White S, Gorver S, Warne G, Zacharin M, Nevin A, Landolsi H, Elghezal H, Saad A, Restrepo CM, Fellous M, Sinclair A, Koopman and Laissue P. (2012). *CITED2* mutations potentially cause idiopathic premature ovarian failure. *Translational Research*, 160:384-88 (**ver anexo 5.3**).

### **4. ¿Son las variantes de secuencia del promotor y del ORF de *BMP15* una causa de FOP no-sindrómica e idiopática?**

**Hipótesis:** las variantes de secuencia del promotor de *BMP15* pueden modificar la unión de factores de transcripción afectando su transactivación lo que, *in fine*, podría traducirse en FOP. Esta hipótesis se fundamenta en que a pesar de que *BMP15* es un gen candidato relevante de FOP, se han identificado y validado pocas mutaciones causales. Además, podrían existir haplotipos de susceptibilidad compuestos por variantes localizadas sobre la región promotora y codificante del gen.

**Ver anexo 5.4.** En este anexo se describe la metodología y resultados de esta pregunta de investigación.

### **5. ¿Podría el polimorfismo del promotor de *BMP15* c.-9C>G estar asociado a la etiología de la FOP no-sindrómica e idiopática?**

**Hipótesis:** Esta variante podría modificar el sitio canónico de unión de un factor de transcripción modificando su capacidad de transactivación/transrepresión.

**Resultados publicados en:** Fonseca DJ, Ortega-Recalde O, Esteban-Perez C, Moreno-Ortiz H, Patiño LC, Bermúdez OM, Ortiz AM, Restrepo CM, Lucena E, and Laissue P. (2014). *BMP15* c.-9C>G promoter sequence variant may contribute to the cause of non-syndromic premature ovarian failure. *Reproductive Biomedicine Online*, 29:627-33 (ver anexo 5.5).

**6. ¿Es posible utilizar las técnicas de NGS para la identificación de nuevos marcadores moleculares de FOP?**

**Hipótesis:** Asumiendo que los procesos fisiológicos relacionados con la determinación sexual, la gametogénesis, la foliculogénesis y la ovulación están regulados por cientos de genes, proponemos que las técnicas de NGS pueden ser eficientes para la identificación de nuevos marcadores moleculares de esta patología.

**Resultados publicados en:** Fonseca DJ, Patiño LC, Suárez YC, de Jesús Rodríguez A, Mateus HE, Jiménez KM, Ortega-Recalde O, Díaz-Yamal I, and Laissue P. (2015). Next generation sequencing in women affected by nonsyndromic premature ovarian failure displays new potential causative genes and mutations. *Fertility and Sterility*, 104:154-62 (ver anexo 5.6).

**7. ¿Es la mutación p.Thr943Ile de ADAMTS19 causal de FOP no sindrómica e idiopática?**

**Hipótesis:** Esta mutación podría afectar la interacción proteína: proteína (sustratos?)

**Ver anexo 5.7.** En este anexo se describe la metodología y los resultados de esta pregunta de investigación.

Anexo 5.1. Ojeda D, Lakhali B, **Fonseca DJ**, Braham R, Landolsi H, Mateus H, and Laissue P. (2011). Sequence analysis of the CDKN1B gene in patients with premature ovarian failure reveals a novel mutation potentially related to the phenotype. *Fertility and Sterility*, 9: 2658–60.

Anexo 5.2: **Fonseca DJ**, Garzón E, Lakhal B, Braham R, Ojeda D, Elhgezal H, Saad A, Restrepo CM, and Laissue P. (2012). Screening for mutations of the *FOXO4* gene in premature ovarian failure patients. *Reproductive Biomedicine Online*, 24: 339-41.

Anexo 5.3: **Fonseca DJ**, Ojeda D, Lakhal B, Braham R, Eggers S, Turbitt E, White S, Gorver S, Warne G, Zacharin M, Nevin A, Landolsi H, Elghezal H, Saad A, Restrepo CM, FellousM, Sinclair A, Koopman and Laissue P. (2012). *CITED2* mutations potentially cause idiopathic premature ovarian failure. *Translational Research*, 160:384-88.

#### **Anexo 5.4. ¿Son las variantes de secuencia del promotor y del ORF de *BMP15* una causa de FOP no-sindrómica e idiopática?**

**Hipótesis:** Las variantes de secuencia del promotor de *BMP15* pueden modificar la unión de factores de transcripción afectando su transactivación lo que, *in fine*, podría traducirse en FOP. Esta hipótesis se fundamenta en que a pesar de que *BMP15* es un gen candidato relevante de FOP, se han identificado y validado pocas mutaciones causales. Además, podrían existir haplotipos de susceptibilidad formados por variantes localizadas sobre la región promotora y codificante del gen.

**Anexo 5.4. Este anexo describe la metodología y resultados que responden la pregunta de investigación numero 4 descrita en la sección 5 (preguntas científicas e hipótesis).**

**Pregunta Científica:**

**¿Son las variantes de secuencia del promotor y del ORF de *BMP15* una causa de FOP no-sindrómica e idiopática?**

**Hipótesis:** las variantes de secuencia del promotor de *BMP15* pueden modificar la unión de factores de transcripción afectando su transactivación lo que, *in fine*, podría traducirse se traduce en FOP. Esta hipótesis se fundamenta en que a pesar de que *BMP15* es un gen candidato relevante de FOP, se han identificado y validado pocas mutaciones causales. Además, podrían existir haplotipos de susceptibilidad formados por variantes localizadas sobre la región promotora y codificante del gen.

**5.4.1 Metodología para la secuenciación de la región promotora de *BMP15***

**5.4.1.1 Casos y controles**

El grupo de pacientes para el análisis de este gen estuvo constituido por 93 mujeres con diagnóstico de FOP procedentes de Túnez. El grupo de controles estuvo constituido por 71 mujeres procedentes de Túnez.

**5.4.1.2 Amplificación y secuenciación de la región promotora de *BMP15***

Todas las muestras de ADN fueron cuantificadas mediante espectrofotometría usando un Nanodrop 2000c (Thermo Scientific).

Las muestras fueron amplificadas usando primers específicos para analizar hasta -1952 pb de la región promotora (la adenina del primer ATG se considera como la posición 0). Estos primers fueron diseñados con el software primer 3 (Tabla 1). Se realizó una sola amplificación para todo el fragmento correspondiente al promotor de *BMP15*. Se mezclaron en un volumen final de 25  $\mu$ l: 100 ng de ADN, 12.5  $\mu$ l Master mix (Promega), 0.8  $\mu$ M de cada primer. Las condiciones de PCR fueron: desnaturalización inicial: 95°C x 10 min; 34

ciclos de: desnaturalización a 95°C x 40 segundos; alineamiento a 63°C x 30 segundos; extensión a 72°C x 2 min; y elongación final a 72°C x 10 minutos.

<b>Primer</b>	<b>Secuencia</b>
<b>hBMP15pr-1838F</b>	5'GTGCTATCGATCCATTTAATGTGCTATC 3'
<b>hBMP15+114R</b>	5'CTCAGCCAGAAGGGCAATAGAGGACT 3'

Tabla 1. Primers usados para la amplificación de la región promotora de *BMP15*

La detección del producto amplificado se realizó en geles de agarosa al 1% teñidos con bromuro de etidio. Los productos de PCR fueron purificados mediante fosfatasa alcalina y exonucleasa I antes de la secuenciación. Las reacciones de secuencia fueron realizadas con primers internos, diseñados con el software primer 3 (Tabla 2). El análisis de las secuencias de nucleótidos fue realizado a través de alineamientos múltiples usando el programa clustalW, comparando la secuencia de cada paciente con la WT para la región promotora de *BMP15*. Los cromatogramas además fueron analizados manualmente.

<b>Primer</b>	<b>Secuencia</b>
<b>hBMP15pr-1318SeqR</b>	5'TTCCAAGACAAATTTCTTCACTTCCCTAC3'
<b>hBMP15pr-1725SeqF</b>	5'CTAAATTTTGCCTATTGTGAATCATCTG3'
<b>hBMP15pr-1472SeqF</b>	5'CAGAGGAAGAGCACAAATAAATTTGAAC 3'
<b>hBMP15pr-883SeqF</b>	5'GGATCTCAACACCTTCCAAGGTAAAC 3'
<b>hBMP15pr-452SeqF</b>	5'GAAGTAGGATGGGAACAGGAGGCCTCTG3'

Tabla 2. Primers usados para la secuenciación de la región promotora de *BMP15*

Las variantes halladas en casos y en controles fueron tamizadas en dbSNP para corroborar si habían sido previamente descritas. Se realizó la determinación de las frecuencias alélicas y genóticas de las variantes de secuencia halladas en casos y en controles usando el programa estadístico SNPStats.

#### 5.4.2 Metodología para la secuenciación de la región codificante de *BMP15*

##### 5.4.2.1 Casos y controles

El grupo de pacientes para el análisis de este gen está constituido por 93 mujeres con diagnóstico de FOP procedentes de Túnez. El grupo de controles estuvo constituido por 71 mujeres procedentes de Túnez. Este grupo de pacientes y de controles son las mismas a las cuales se les analizó la región promotora de *BMP15*.

##### 5.4.2.2 Amplificación y secuenciación de la región codificante de *BMP15*

Las muestras fueron amplificadas usando primers específicos para analizar los dos exones codificantes de *BMP15*. Dichos primers fueron diseñados por el software primer 3 (Tabla 3). La PCR para los exones 1 y 2 fue realizada en volúmenes finales de 25  $\mu$ l: 100 ng de ADN, 12.5  $\mu$ l Master mix (Promega), 0.8  $\mu$ M de cada primer. Las condiciones de PCR fueron: desnaturalización inicial: 95°C x 10 min; 34 ciclos de: desnaturalización parcial a 95°C x 40 segundos; annealing a 63°C x 40 segundos; extensión a 72°C x 40; y elongación final a 72°C x 10 minutos.

<b>hBMP15Ex1-F1</b>	GATCCTCTTAGAGAAAGCAACATAGGG
<b>hBMP15Ex1-R1</b>	AAGTTGATGCCCTAAAGGGAAGTAGG
<b>hBMP15Ex2-F1</b>	CATTATCTTGAGTTAATTTAGGGCTGA
<b>hBMP15Ex2-R1</b>	AGCATTAAAGGGCTCTGATTAGTGGTC

Tabla 3. Primers usados para la amplificación de la región codificante de *BMP15*

La detección de los productos amplificados se realizó en geles de agarosa al 1%.

Los productos de PCR fueron purificados mediante fosfatasa alcalina y exonucleasa I antes de la secuenciación. Las reacciones de secuencia fueron realizadas con primers internos, diseñados con el software primer 3 (Tabla 4). El análisis de las secuencias de nucleótidos

fue realizado a través de alineamientos múltiples usando el programa clustalW, comparando la secuencia de cada paciente con la WT de *BMP15*. Los cromatogramas fueron analizados manualmente.

<b>hBMP15Ex1Sq-F1</b>	<b>CTTTTCTTGCCCCATCCTTGGTTG</b>
<b>hBMP15Ex1Sq-R1</b>	<b>CCAAGCCTAAAGAAACAACCTGCC</b>
<b>hBMP15Ex2Sq-F1</b>	<b>GCTGATTATAGCTATCAGTCTATATC</b>
<b>hBMP15Ex2Sq-R1</b>	<b>CTGGTATCAGAACTGTTTCTATAAA</b>

Tabla 4. Primers usados para la secuenciación de la región codificante de *BMP15*

Las variantes de secuencia halladas en los casos y en los controles fueron tamizadas en dbSNP para corroborar si habían sido previamente descritas. El alineamiento de las secuencias proteicas mutantes de *BMP15* humanas con las de diferentes especies para verificar la conservación de los aminoácidos implicados en el cambio, fue realizado con ClustalW. Las especies con las que se compararon fueron: *R. norvegicus*, *M. musculus*, *B. taurus*, *S. labiatus*, *G. gallus*, *X. leaveis*, *X. silurana tropicalis*, *D. rerio* y *C. idella*. El análisis del potencial efecto deletéreo de los cambios no-sinónimos se realizó usando el programa Polyphen-2, el cual predice el posible impacto de la sustitución de un aminoácido en la estructura y en la función de la proteína. Algunos de los parámetros usados por Polyphen2 son la comparación de secuencias de proteínas de diferentes especies ortólogas y parálogas y las características fisicoquímicas del cambio del residuo. Se realizó la determinación de las frecuencias alélicas y genotípicas de las variantes de secuencia halladas en casos y en controles usando el programa estadístico SNPStats.

#### **5.4.3 Metodología para el análisis de haplotipos de las variantes identificadas en las regiones promotora y codificante de *BMP15***

Con el fin de identificar potenciales haplotipos asociados a FOP se analizaron en conjunto las variantes localizadas en las regiones promotora y codificante del gen *BMP15*. Mediante los programas SNPStats y Haploview se identificaron los haplotipos en la población de estudio y se determinó si existían diferencias significativas entre la frecuencia identificada en casos y en controles, mediante el test de Chi-cuadrado.

## **5.5 Resultados**

### **5.5.1. Resultados de la secuenciación directa de la región promotora de *BMP15***

En la región promotora del gen *BMP15* se identificaron, tanto en casos como en controles, dos variantes de secuencia: -673T>C y c.-9C>G previamente reportadas en las bases de datos de SNP. Las frecuencias alélicas y genotípicas para estas variantes no presentaron diferencias significativas entre los casos y los controles ( $p>0.05$ ). En ambas poblaciones las frecuencias genotípicas estuvieron en equilibrio de Hardy Weinberg ( $p>0.05$ ). A través de haploview se demostró que las dos variantes estaban en desequilibrio de ligamiento ( $D' = 0.9044$ ). Estos hallazgos representan el primer reporte de secuenciación del promotor de *BMP15* en pacientes FOP no sindrómica e idiopática. Las variantes identificadas (c.-9C>G y -673C>T) han sido previamente reportadas en las bases de datos en frecuencias muy variables en diferentes poblaciones de individuos sanos. Algunos autores han reportado la variante c.-9C>G en pacientes FOP y en controles (Dixit et al, 2006; Tiotiu et al, 2010).

### **5.5.2 Resultados de la secuenciación de la región codificante de *BMP15***

La secuenciación directa de la región codificante de *BMP15* permitió identificar 6 variantes (Tabla 5). Entre ellas la única que se encontró exclusivamente en casos fue c.433T>C, p.Leu148Pro. Esta mutación había sido previamente reportada como causal del fenotipo FOP (Laissue et al, 2006; Rossetti et al, 2009; Tiotiu et al, 2010). Las variantes no sinónimas p.Ser5Arg, p.Asn103Ser y p.ins263Leu han sido previamente descritas como polimorfismos no funcionales (Wang et al, 2010; Di Pasquale et al, 2006; Dixit et al, 2006; Laissue et al, 2006; Laissue et al, 2010).

Exón	Cambio de Secuencia	dbSNP	Proteína
<b>1</b>	<b>c.13A&gt;C</b>	<b>rs113099187</b>	<b>p.Ser5Arg</b>
<b>1</b>	<b>c.308A&gt;G</b>	<b>rs41308602</b>	<b>p.Asn103Ser</b>
<b>2</b>	<b>c.788insTCT</b>	<b>rs78208991</b>	<b>p.ins263Leu</b>
<b>2</b>	<b>c.819A&gt;C</b>	<b>rs73488038</b>	<b>p.Ser273Ser</b>
<b>2</b>	<b>c.852C&gt;T</b>	<b>rs17003221</b>	<b>p.Ser284Ser</b>
<b>2</b>	<b>c.433T&gt;C</b>		<b>p.Leu148Pro</b>

Tabla 5. Variantes identificadas en la región codificante de *BMP15*

Las frecuencias alélicas y genotípicas para las variantes de secuencia halladas en la población de estudio no indicaron diferencias significativas entre los casos y los controles (SNPStats  $p > 0.05$ ) (Tabla 6). Todas las variantes identificadas en la población estudiada se hallaron en equilibrio de Hardy-Weinberg.

SNP	ALELOS Y GENOTIPOS	FRECUENCIA ALÉLICA Y GENOTÍPICA DE CONTROLES	FRECUENCIA ALÉLICA Y GENOTÍPICA DE CASOS
c.13A>C	A	0.98	0.98
	C	0.02	0.02
	AA	0.96	0.96
	AC	0.04	0.04
	CC	0	0
c.308A>G	A	0.98	0.97
	G	0.02	0.03
	AA	0.96	0.94
	AG	0.04	0.06
	GG	0	0
c.788ins TCT	D	0.89	0.89
	I	0.11	0.11
	DD	0.77	0.79
	DI	0.23	0.2
	II	0	0.01
c.819A>C	A	0.99	0.99
	C	0.01	0.01
	AA	0.99	0.98
	AC	0.01	0.02
	CC	0	0
c.852C>T	C	0.9	0.9
	T	0.1	0.1
	CC	0.8	0.8
	CT	0.2	0.2
	TT	0	0
c. 433T>C	T	1	0.984
	C	0	0.016
	TT	1	0.97
	TC	0	0.03
	CC	0	0

Tabla 6. Frecuencias alélicas y genotípicas de las variantes de secuencia identificadas en la región codificante de *BMP15*.

### 5.5.3. Resultados del análisis de haplotipos de la región promotora y codificante de *BMP15*

Con las variantes descritas en la región promotora (n=2) y codificante (n=5) y utilizando el programa SNPStats y Haploview, se identificaron 19 haplotipos diferentes para la

población de estudio (Tabla 7). Solo uno de ellos fue encontrado de manera exclusiva en casos y corresponde a TGAAIAT (Tabla 7, haplotipo número 10).

Haplotype frequencies estimation (n=161)											
	X.673	X.9	X13	X308	insTCT	X819	X852	Total	group.0.Control	group.1.case	Cumulative frequency
1	C	C	A	A	D	A	C	0.6845	0.6629	0.7036	0.6845
2	T	G	A	A	D	A	C	0.1559	0.1668	0.1488	0.8404
3	C	C	A	A	I	A	T	0.0588	0.0715	0.0457	0.8993
4	T	G	A	G	D	A	C	0.0204	0.0141	0.0238	0.9196
5	C	C	A	A	I	C	T	0.0093	0.007	0.0111	0.9289
6	C	C	C	A	I	A	T	0.0089	NA	0.0167	0.9379
7	C	G	A	A	I	A	T	0.0088	0.0201	NA	0.9467
8	C	G	A	A	D	A	C	0.0074	0.0154	NA	0.9541
9	T	C	A	A	D	A	C	0.007	NA	0.0127	0.9611
10	T	G	A	A	I	A	T	0.0063	0	0.0123	0.9674
11	T	G	A	A	I	A	C	0.0058	0.0047	0.0056	0.9733
12	T	C	A	A	I	A	T	0.0058	NA	0.0102	0.9791
13	C	C	C	A	D	A	C	0.0047	0.0097	NA	0.9838
14	C	C	A	A	I	A	C	0.0035	0.0094	NA	0.9873
15	C	G	A	G	D	A	C	0.0032	0.007	NA	0.9905
16	C	G	C	A	D	A	C	0.0031	0.0068	NA	0.9936
17	C	C	C	A	I	A	C	0.0031	NA	0.0037	0.9967
18	T	G	C	A	D	A	C	0.002	0.0046	NA	0.9987
19	T	G	A	G	I	A	T	0.0013	NA	0.004	1

Tabla 7. Haplotipos para *BMP15* entre la región promotora y codificante. Tomado del programa Haploview.

Con los SNP analizados se evidenció la formación de un bloque haplotípico entre los marcadores c.852C>T y c.788\_799 insTCT (Figura 1). Estos marcadores mostraron un fuerte desequilibrio de ligamiento ( $D'0.9992$ ) (Tabla 8).

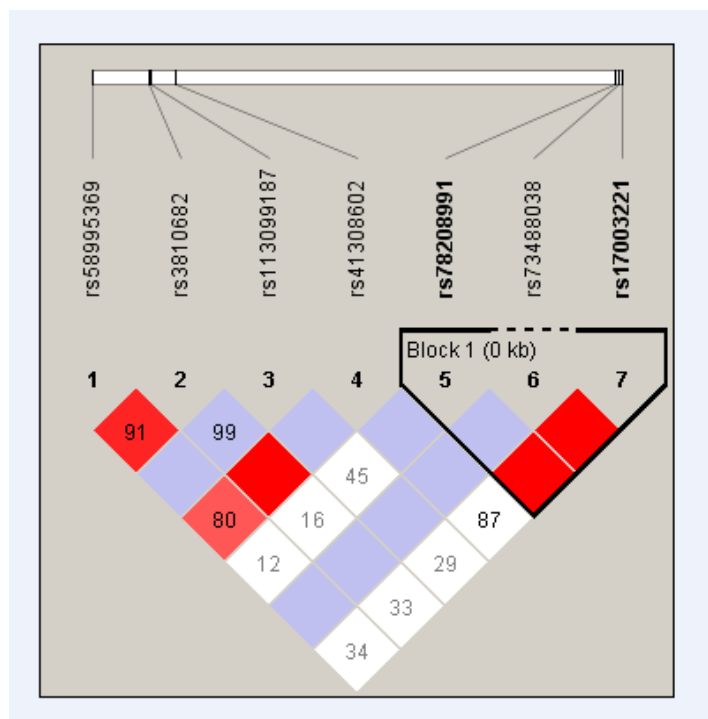


Figura 1. Bloques Haplotipicos *BMP15*. Tomado del programa Haploview

**D' statistic**

	X.673	X.9	X13	X308	insTCT	X819	X852
X.673	.	0.9251	0.9845	0.8062	0.123	0.9593	0.3467
X.9	.	.	0.9806	0.996	0.1606	0.9621	0.3309
X13	.	.	.	0.8758	0.4515	0.6283	0.2892
X308	.	.	.	.	0.9724	0.6748	0.8767
insTCT	.	.	.	.	.	0.9908	0.9992
X819	.	.	.	.	.	.	0.991
X852	.	.	.	.	.	.	.

Tabla 8. Análisis de desequilibrio de ligamiento de variantes de *BMP15*. Tomado del programa Haploview.

Teniendo en cuenta los hallazgos positivos previos de asociación haplotípica de variantes del gen *BMP15* con FOP (Dixit et al, 2006), realizamos este análisis en la población de estudio para identificar potenciales combinaciones alélicas relacionadas con el fenotipo

FOP. Desde esta perspectiva identificamos un haplotipo exclusivo de los casos (TGAAIAT). Sin embargo no fue posible demostrar diferencias significativas entre casos y controles. Esto puede ser atribuido al insuficiente tamaño de la muestra. Se puede apreciar que uno de los SNP del haplotipo (cuya base no corresponde al WT) es la variante del promotor de *BMP15* c.-9C>G. Este cambio podría estar implicado con la función ovárica, lo que sustentaría los hallazgos previos de asociación con infertilidad, con FOP y como modulador de la respuesta a la rFSH (Dixit et al, 2006; González et al, 2008; Hanevik et al, 2011).

Anexo 5.5: **Fonseca DJ**, Ortega-Recalde O, Esteban-Perez C, Moreno-Ortiz H, Patiño LC, Bermúdez OM, Ortiz AM, Restrepo CM, Lucena E, and Laissue P. (2014). *BMP15* c.-9C>G promoter sequence variant may contribute to the cause of non-syndromic premature ovarian failure. *Reproductive Biomedicine Online*, 29:627-33.

Anexo 5.6: **Fonseca DJ**, Patiño LC, Suárez YC, de Jesús Rodríguez A, Mateus HE, Jiménez KM, Ortega-Recalde O, Díaz-Yamal I, and Laissue P. (2015). Next generation sequencing in women affected by nonsyndromic premature ovarian failure displays new potential causative genes and mutations. *Fertility and Sterility*, 104:154-62.

**Anexo 5.7 ¿Es la mutación p.Thr943Ile de *ADAMTS19* causal de FOP no sindrómica e idiopática?**

**Hipótesis:** Esta mutación podría modificar la interacción proteína:proteína y/o alterar su actividad enzimática.

**Anexo 5.7.** Este anexo describe la metodología y los resultados que buscan responder a la pregunta de investigación número 7 descrita en la sección 5 (preguntas científicas e hipótesis).

**Pregunta científica central:**

**¿Es la mutación p.Thr943Ile de ADAMTS19 causal de FOP no sindrónica?**

**Hipótesis:** Esta mutación podría modificar la interacción proteína:proteína y/o alterar su actividad enzimática.

### **5.7.1 Antecedentes**

#### **5.7.1.1 ADAMTS19**

Las proteínas ADAMTS (*a disintegrin and metalloproteinase with thrombospondin motifs*) fueron descritas por primera vez por Kuno et al. en 1997 en ratones. Desde entonces se han identificado en otros mamíferos y en *Caenorhabditis elegans* (Kuno et al, 1997). Según la base de datos de MEROPS, las proteínas ADAMTS forman parte de la subfamilia B (*adamalysin subfamily*), familia 12 y clan MA de metalopeptidasas (Barrett et al, 2001). Están estructural y evolutivamente relacionadas con ADAMT (*a disintegrin and metalloproteinase*) y de manera más distante con proteínas MMP (*matrix metalloproteinase*).

En humanos, se han identificado 19 factores ADAMTS. Mediante el análisis de dendogramas construidos con los alineamientos de la secuencia de la proteína, se ha determinado que ADAMTS puede ser clasificada en 5 grupos según su homología estructural y funcional. El primer grupo contiene las ADAMTS 1, 4, 5,8 y 15. El segundo involucra a ADAMTS 2,3 y 14. Desde el tercer hasta el quinto grupo, las proteínas están agrupadas en pares: ADAMTS 19/17, ADAMTS 18/16, ADAMTS 12/7, ADAMTS 10/6 y ADAMTS 9/20. Finalmente, ADAMTS 13 no está asociada a ningún otro miembro de la familia (Kelwick et al, 2015). Desde el punto de vista estructural, todos los miembros de la familia ADAMTS están conformados por un dominio proteinasa y un dominio auxiliar. La mayor variabilidad se presenta en la región C-terminal de la proteína (dominio auxiliar) (Jones et al, 2005; Kelwick et al, 2015). Las proteínas ADAMTS son sintetizadas

inicialmente como pre-proenzimas inactivas. La estructura proteica de ADAMTS19 está ilustrada en la figura 2.

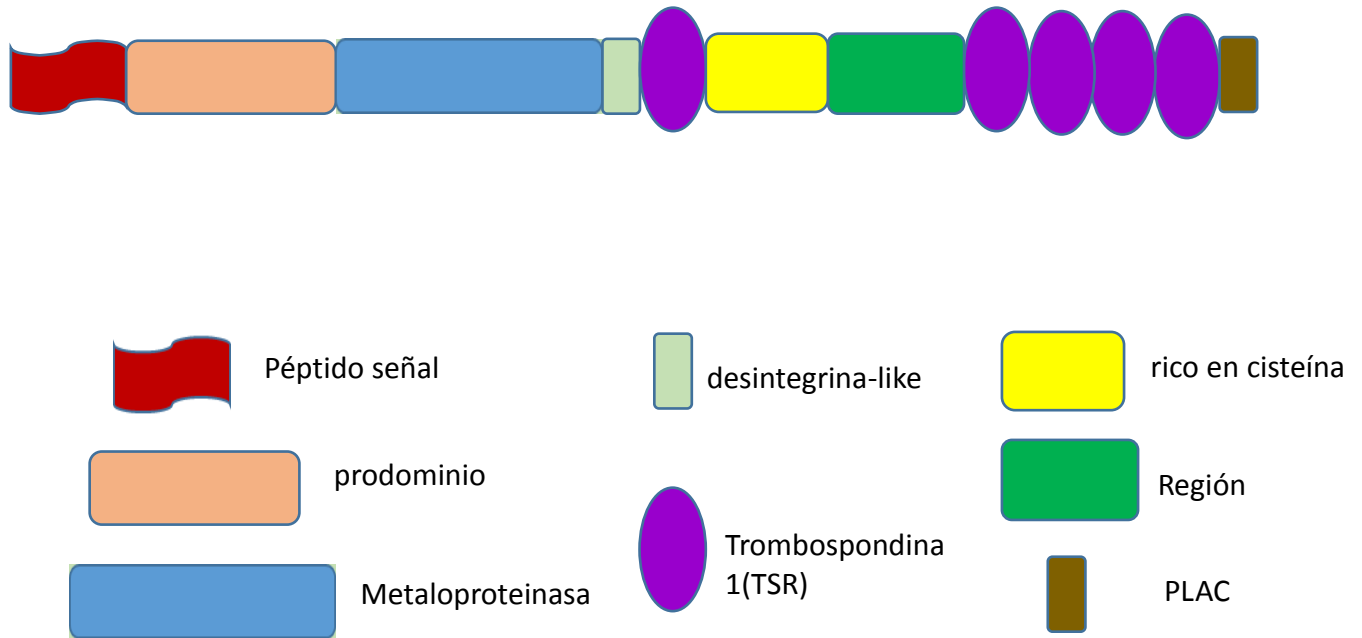


Figura 2. Estructura de la proteína ADAMTS19.

A continuación se describen los dominios que conforman las proteínas ADAMTS: (Jones et al, 2005; Tortorella et al, 2009).

- Péptido de señalización
- Pro-dominio: mantiene a la proteasa en un estado latente que previene el acceso a los sustratos. La longitud del propeptido es de 41 a 295 aminoácidos. Este prodominio debe ser removido para permitir la activación de la función catalítica.
- Dominio catalítico metaloproteínica: es la región que contiene el sitio activo de la enzima responsable de la hidrólisis. De manera semejante a todas las metaloproteasas los miembros de la familia, ADAMTS utilizan zinc unido a agua como nucleófilo. Este átomo de zinc está asociado con la proteína a través de la unión en un dominio conservado HEXXHXXGXXH.

- Dominio *disintegrin-like*: es una región que exhibe similaridad de secuencia con el motivo *disintegrin* encontrado en miembros de la familia ADAMT. Este dominio se une a varios receptores de integrina tales como  $\alpha 3$ ,  $\alpha 6$  y  $\beta 1$  (Huang, 1998). El dominio tiene entre 57 y 96 aminoácidos de longitud.
- Dominio *thrombospondin type I (TS)*: contiene una secuencia de 53 a 56 aminoácidos característica de la *thrombospondin I* y *II*. Este motivo ha sido implicado en la unión con glicosaminoglicanos (GAG) y con el receptor CD36 (Zheng et al, 2001).
- Dominio *cysteine-rich*: este dominio tiene la función de reconocimiento y de localización en la matriz extracelular. Es probable que este dominio determine la movilidad, la localización y la especificidad del sustrato de los miembros ADAMTS (Tortorella et al, 2009).
- Dominio TS: presenta un número variable de repeticiones de TS. Algunas proteínas tienen hasta 14 repeticiones, como es el caso de ADAMTS-9 mientras que otras no tienen ninguna repetición (ADAMTS-4) (Huang, 1998).
- PLAC (*Protein that localizes at cell-cell contacts*): este dominio se encuentra en la mitad de los miembros ADAMTS (2, 3, 10, 12, 14, 16, 17,18 y 19). Está localizado en el extremo C-terminal y se encuentra también en varias pro-proteínas convertasas que corresponden a un grupo de proteasas responsables de la activación de ADAMTS (Zheng et al, 2001).

Adicionalmente, ADAMTS 9 y 20 contienen un dominio de cerca de 200 aminoácidos denominado GON-1 también presente en *C. elegans* relacionado con el desarrollo gonadal (Somerville et al, 2003). ADAMTS-13 es el único miembro de la familia que tiene dos dominios CUB (*complement subcomponent Clr/Clf/embryonic sea urchin protein Uegf/bone morphogenic protein 1*). Este dominio es funcionalmente independiente y es encontrado en un rango amplio de proteínas extracelulares. Algunas proteínas con dominios CUB se han relacionado con procesos del desarrollo como la embriogénesis y la organogénesis (Bork et al, 1993).

De manera similar a otras metaloproteasas, las proteínas ADAMTS son sintetizadas como enzimas inactivas que incluyen propéptidos que enmascaran el dominio catalítico, lo que previene el acceso a los sustratos. Las ADAMTS tienen un motivo que consiste en 4 aminoácidos Arg-X-Lys/Arg-Arg localizados entre los dominios pro y catalítico. Este motivo es reconocido por las proproteínas convertasas que remueven el propéptido de la enzima (Seidah et al, 2002). La activación del extremo N-terminal puede ocurrir dentro o fuera de la célula sobre la superficie celular o en la matriz extracelular (Tortorella et al, 2009). Por ejemplo, pro-ADAMTS-1 y pro-ADAMTS-4 son activadas intracelularmente por la furina en una variedad de tipos celulares y ADAMTS-5 por PACE4 en la matriz extracelular (Malfait et al, 2008).

Adicionalmente a la eliminación del prodominio, la activación de ADAMTS puede producirse por otros mecanismos como el clivaje del dominio *cysteine-rich* (Tortorella et al, 2009). Las funciones biológicas de las ADAMTS son variadas. Por ejemplo, ADAMTS 2,3, y 14 actúan en el procesamiento del propéptido amino-terminal de las procollagenasas I, II y III. Las mutaciones en *ADAMTS-2* causan el síndrome de Ehlers-Danlos tipo VIIC. En el modelo murino KO para *Adamts-2* los animales son normales al nacer pero luego desarrollan fragilidad de la piel. Los machos son infértiles (Li et al, 2001). ADAMTS-3 por su parte presenta una expresión mayor en el cartílago que en la piel y actúa como una N-propeptidasa pro colágeno tipo II (Fernandes et al., 2001). ADAMTS13 actúa sobre un precursor del factor de Von Willebrand (vWF) para producir un factor activo implicado en la anticoagulación. En humanos las mutaciones en *ADAMTS-13* están relacionadas con púrpura trombótica trombocitopénica, una enfermedad que presenta gran cantidad de multímeros vWF (Scheiflinger et al, 2003). *Adamts-20* es el ortólogo murino del gen *gon-1* de *C.elegans* y es requerido para la migración y la supervivencia de los melanoblastos durante la embriogénesis. Las mutaciones homocigotas en *Adamts-20* generan ratones con manchas blancas denominados *Belted* (Rao et al, 2003). ADAMTS-1,4,5 y 9 (*aggrecanases*) catabolizan componentes importantes del sistema vascular y nervioso (Somerville et al, 2003).

Algunos miembros de la familia ADAMTS se han relacionado con el desarrollo folicular y con procesos ovulatorios. El modelo murino KO para *Adamts1* por ejemplo,

demostró que las hembras homocigotas eran infértiles o presentaban pequeñas camadas (Shindo et al, 2000). En las hembras mutantes, los oocitos maduros quedaban atrapados en los folículos ováricos lo que permitió sugerir que *Adamts-1* es requerido para la ovulación normal (Mittaz et al, 2004).

Con el objetivo de estudiar detenidamente el rol de *Adamts-1* en la función del ovario Shozu et al. generaron un modelo murino mediante cruces de hembras KO heterocigotas (+/-) y machos homocigotos (-/-) de *background* genético 129/Sv y C57BL/6. Cuando las hembras inmaduras fueron tratadas con PMSG (*pregnant mare serum gonadotropin*)/hCG (*human chorionic gonadotropin*) el número de oocitos ovulados fue mucho menor en las hembras mutantes (Shozu et al, 2005). En las hembras con la deficiencia de *Adamts-1*<sup>-/-</sup>, la proporción de folículos anovulados en relación al total de folículos maduros fue mucho más alto y el número de folículos preantrales tardíos fue muy reducido. Estos datos demostraron que la alteración en la función ovárica secundaria a la ablación de *Adamts-1* se presentaba en dos niveles: en el desarrollo de los folículos en crecimiento y en el proceso ovulatorio (Shozu et al, 2005). Adicionalmente, Shozu et al. observaron que en las hembras mutantes, los ovarios contenían una proporción elevada de folículos atrésicos con pérdida de capas de células de la granulosa. Este hallazgo indicó que *Adamts1* es importante en el desarrollo folicular y en el mantenimiento de las capas de células de la granulosa en los folículos (Shozu et al, 2005).

Además de *Adamts1*, otros miembros de la familia se han relacionado con funciones a nivel ovárico. Antes de la ovulación, *Adamts-4* y *Adamts-5* son coexpresados en las células de la granulosa. La expresión de *Adamts-5* se ha observado también en células de la granulosa de folículos atrésicos. Richards et al, sugirieron que *Adamts1* y *4* actúan en la formación, la estabilización y la función de la matriz que rodea y protege al oocito. La elevada expresión de estos genes en el sitio de ruptura de la superficie del ovario sugiere que esas proteasas tiene una importante función en la angiogénesis (Richards et al, 2005).

ADAMTS19 está incluida junto con ADAMTS 6, 10, 16,17, y 18 dentro del grupo de las proteínas ADAMTS huérfanas, denominadas de esta manera porque no se han descrito sus sustratos específicos. A pesar de esto, se ha evidenciado que ADAMTS-7,10 y 12 son proteolíticamente activas. ADAMTS -7 tiene una isoforma (7B) con un dominio mucina. Esta isoforma sufre varias modificaciones post-traduccionales de O-glicosilación. Esta característica sugiere que ADAMTS-7B puede funcionar como proteoglicano en algunas células y tejidos (Somerville et al, 2004). Otros miembros de este grupo como ADAMTS-6 y 18 son expresados diferencialmente en el cáncer de seno (Porter et al., 2004) mientras que *ADAMTS-12* y *ADAMTS-16* están desreguladas en la osteoartritis (Davidson et al, 2006). *ADAMTS-16* se expresa en el riñón y en el ovario. La FSH induce su expresión en las células de la granulosa luteinizante. *ADAMTS-16* es capaz de catalizar la  $\alpha$ -2 macroglobulina, un sustrato común de las proteasas (Gao et al, 2007). *ADAMTS-10* se ha relacionado con el crecimiento y con el desarrollo cardíaco en humanos. Las mutaciones de *ADAMTS-10* se han relacionado con síndrome Weill-Marchesani, una condición caracterizada por baja talla, anormalidades oculares, miopía y glaucoma (Dagoneau et al, 2004).

*ADAMTS-19* no tiene una función caracterizada hasta la fecha. Sin embargo, en un estudio de GWAS realizado en 99 pacientes caucásicas con FOP idiopática, se encontró una evidencia de asociación con *ADAMTS19*. El SNP asociado (rs246246; P-value  $<2 \times 10^{-7}$ ) se localizó en una región intrónica (intrón 21) en un bloque de desequilibrio de ligamiento de 200 Kb en el cromosoma 5q31 (Knauff et al, 2009). Este hallazgo no pudo ser replicado en una cohorte independiente de origen alemán (60 pacientes FOP) (Knauff et al, 2009).

Más recientemente, Pyun et al. plantearon que *IGF2R* (*Insuline-like growth factor receptor*) podría ser un gen candidato para el desarrollo de FOP ya que su síntesis regula la concentración de IGF (*insuline-like growth factor*). Se ha demostrado que IGF actúa sinérgicamente con la FSH para inducir la esteroidogénesis en las células de la granulosa en bovinos ejerciendo una regulación positiva sobre el *FSHR* (Pyun et al, 2013). Esta observación y el resultado previo de Knauff et al. sobre una posible causalidad de variantes en *ADAMTS19* en la FOP, fue la base conceptual para que Pyun

et al. realizaron un estudio de asociación de estos dos genes en pacientes con esta enfermedad. Se realizó un estudio de casos y de controles mediante el análisis de SNP's intra y extragénicos de los genes *IGF2* y *ADAMTS19*. Se determinó que existe una asociación estadísticamente significativa ( $p < 0.0001$ ) entre diplotipos de estos genes y el fenotipo FOP. El SNP de *ADAMTS19* (rs246246) que se encontró asociado a la FOP fue monomórfico en la población analizada por Pyun et al. (Pyun et al, 2013). Se ha descrito que IGF2 incrementa la producción de estradiol y de progesterona en las células de la granulosa. La expresión de algunos miembros de la familia ADAMTS es regulada por la progesterona y por su receptor. Pyun et al. hipotetizaron que *ADAMTS19* puede ser regulado por la progesterona y controlado por IGFR2 y que la acción sinérgica entre estos genes puede contribuir al desarrollo de FOP (Pyun et al, 2013).

En ratones, la expresión de *Adamts19* es mucho mayor en el ovario embrionario que en los testículos durante la diferenciación sexual (Menke et al, 2002). En otros tejidos como los riñones, el corazón, el músculo esquelético y el pulmón, la expresión del gen es menor (Menke et al, 2002). Estos hallazgos sugieren que *ADAMTS19* puede constituir un gen FOP.

Los experimentos de NGS realizados en este trabajo doctoral, identificaron en una paciente FOP la mutación c.2828C>T (p.Thr943Ile) en *ADAMTS19*. Esta mutación no había sido previamente descrita en las bases de datos de SNP y no fue encontrada en ninguno de los 521 controles analizados (ver anexo número 5.6). Es coherente sospechar un efecto de esta variante en la etiología de la FOP por varias razones: en primer lugar fue encontrada exclusivamente en un caso FOP. En segundo lugar la posición 943 está conservada en varias especies de vertebrados. Tercero, el análisis *in silico* mediante Polyphen indicó un potencial efecto deletéreo. Además, la mutación p.Thr943Ile está localizada en el dominio auxiliar de *ADAMTS19*, en la primera repetición trombospondina (TSR), el cual se ha asociado funcionalmente con la especificidad de su unión a los sustratos, con la regulación de la actividad proteolítica y con la interacción con la matriz extracelular (Kelwick et al, 2015; Apte SS, 2009). Por consiguiente, nosotros hipotetizamos que la mutación p.Thr943Ile en *ADAMTS19*

podría alterar la interacción con otra proteína y/o afectar su localización subcelular. En este marco, iniciamos ensayos de validación funcional de las hipótesis mencionadas. Los detalles metodológicos que permiten responder a la pregunta central acerca de la potencial implicación en la etiología molecular de esta mutación serán expuestos más adelante.

## **5.8 ¿Es la mutación p.Thr943Ile de ADAMTS19 causal de FOP no sindrónica?**

A la fecha la proteína ADAMTS19 es considerada como huérfana, ya que su sustrato no se conoce. Por consiguiente, para validar la potencial patogenicidad de la mutación ADAMTS29 c.2828C>T fue necesario determinar primero los *partners* proteicos de la metaloproteinasa en estudio. En el siguiente aparte se describen los métodos y resultados para la identificación de la interacción proteína:proteína de ADAMTS19.

### **5.8.1. ¿Cuál(es) son lo (s) *partner(s)* proteicos de interacción con ADAMTS19?**

#### **5.8.1.1 Metodología para la identificación de las proteínas *partners* de ADAMTS19.**

##### **Sistema de doble híbrido en levaduras.**

En 1989, las primeras descripciones del sistema de doble híbrido en levaduras iniciaron una nueva era en la investigación de la interacción proteína:proteína. Esta metodología está basada en la reconstrucción de un factor de transcripción funcional, cuando dos proteínas o polipéptidos de interés interactúan (Fields, 1989). El proceso de interacción ocurre entre cepas de levaduras genéticamente modificadas, en las que la transcripción de un gen reportero lleva a la expresión de un fenotipo específico. De manera frecuente, el fenotipo permite el crecimiento de las levaduras sobre un medio selectivo o confiere cambio en el color de las colonias de las levaduras ([www.hybrigenics-services.com](http://www.hybrigenics-services.com)). Respecto a los genes reporteros, dos son comúnmente usados: el primero corresponde a *His3*, un gen que permite el crecimiento en un medio carente de histidina. El segundo es el gen *Lacz* que permite realizar el tamizaje de las levaduras mediante ensayos de colorimetría (Fields, 1989).

Para la realización del ensayo, se generan dos proteínas fusión: la primera (llamada *prey*) corresponde a una de las proteínas de interés que está unida al dominio de unión al ADN del factor de transcripción. La segunda (llamada *bait*) es el híbrido proteico entre el dominio de activación del factor de transcripción y la proteína que se asume ejercerá una interacción con la primera. La reconstitución del factor de transcripción funcional permite

que se active el gen reportero. De manera interesante, el sistema de doble híbrido puede evaluar la interacción entre dos proteínas específicas, o usar una proteína de interés como un *bait* sobre la cual se hace un tamizaje con el producto de una librería de ADNc preparada de tejidos específicos (Fields,1989). Con esta última estrategia, la identificación de los *partners* (*preys*) unidos al *bait* de interés se obtiene mediante secuenciación de los plásmidos en las colonias de levaduras seleccionadas.

El sistema de doble híbrido en levadura que sustenta los resultados de esta tesis doctoral, fue realizado por el laboratorio *Hybrigenics* en París, Francia ([www.hybrigenics.com](http://www.hybrigenics.com)). La proteína “*bait*” fusionada al dominio de unión de ADN del factor de transcripción, fue generada de clones de ADNc de *Adamts19-Mus musculus* (*Gene ID*: 240322) e introducido en los vectores p27 y pB66. La proteína fusión traducida contenía un fragmento de Adamts19 de 535 aminoácidos (aa 637-1172). Se encontraban presentes los dominios de Adamts19 correspondientes a: *thrombospondin type 1 repeat* (SM0029 o PF00090), ADAM-TS Spacer 1 (PF05896) y PLAC (PF08686). Este fragmento correspondía a la región homóloga humana en la que se encontró la mutación p.Thre973Ile en una de las pacientes FOP estudiada por NGS. Las proteínas “*prey*” fusionadas al dominio de activación del factor de transcripción, fueron generadas de clones de una librería de ADNc de ovario de *Mus musculus* en los vectores p27 y pB66. Un total de 148.2 millones de interacciones fueron evaluadas. Después de la selección sobre el medio deficiente de histidina, los clones positivos fueron elegidos y los fragmentos correspondientes al “*prey*” se amplificaron y secuenciaron en sus extremos 3’ y 5’.

Para cada interacción, los resultados fueron expresados en un Puntaje Biológico Predicho (PBS) que evalúa la fiabilidad de la interacción y que se fundamenta en la comparación entre el número de fragmentos de las proteínas sustrato que interactúan con la proteína de interés y la posibilidad de que estos hallazgos se deban al azar. Los resultados de la interacción se clasifican en categorías discriminadas desde la A hasta la F (Fig. 3).

La identificación del(os) “*prey*” que establecen interacción proteína-proteína con el “*bait*” permite la definición del SID (Selection Interaction Domain). Este dominio corresponde a la secuencia de aminoácidos específica del “*prey*” necesaria para su interacción con el “*bait*”.

<b>A</b>	Very high confidence in the interaction
<b>B</b>	High confidence in the interaction
<b>C</b>	Good confidence in the interaction
<b>D</b>	Moderate confidence in the interaction This category is the most difficult to interpret because it mixes two classes of interactions : - False-positive interactions - Interactions hardly detectable by the Y2H technique (low representation of the mRNA in the library, prey folding, prey toxicity in yeast)
<b>E</b>	Interactions involving highly connected (or relatively highly connected) prey domains, warning of non-specific interaction. The total number of screens performed on each organism is taken into account to set this connectivity threshold: 20 interactions to different bait proteins in our entire database for Human, 10 for Mouse, Drosophila and Arabidopsis and 6 for all other organisms. They can be classified in different categories: - Prey proteins that are known to be highly connected due to their biological function - Proteins with a prey interacting domain that contains a known protein interaction motif or a biochemically promiscuous motif
<b>F</b>	Experimentally proven technical artifacts

Figura 3. Categorías de la interacción entre proteínas, según el sistema de Hybrigenics. Tomada de Hybrigenics, Global PBS.

### 5.8.1.2 Resultados del ensayo de doble híbrido en levaduras.

A continuación se presentan los resultados del sistema de doble híbrido en levadura que identificó posibles interacciones de Adamts19.

Clone Name	Type Seq	Gene Name (Best Match)	Start..Stop (nt)	Frame	Sense	%Id 5p	%Id 3p	PBS
pB66_A-17	5p/3p	Mus musculus - Col6a1	990..2011	IF		78.1	87.0	<b>B</b>
pB66_B-87	5p/3p	Mus musculus - Col6a1	1217..2131	OOF2		99.7	96.9	<b>B</b>
pB66_B-114	5p/3p	Mus musculus - Col6a1	1524..2042	IF		100.0	97.2	<b>B</b>
pB66_A-1	5p/3p	Mus musculus - Col6a1	1587..2038	IF		100.0	100.0	<b>B</b>
pB66_A-28	5p/3p	Mus musculus - Col6a2	963..1992	IF		97.7	97.2	<b>A</b>
pB66_A-16	5p/3p	Mus musculus - Col6a2	1053..2008	IF		95.1	97.0	<b>A</b>
pB66_A-39	5p/3p	Mus musculus - Col6a2	1206..1981	IF		99.6	97.1	<b>A</b>
pB66_A-23	5p/3p	Mus musculus - Col6a2	1236..1970	IF		97.4	96.1	<b>A</b>
pB66_A-15	5p/3p	Mus musculus - Col6a2	1284..1992	IF		98.7	97.4	<b>A</b>
pB66_A-32	5p/3p	Mus musculus - Col6a2	1329..2169	IF		96.9	97.1	<b>A</b>
pB66_B-26	5p/3p	Mus musculus - Col6a2	1338..2010	IF		98.2	98.2	<b>A</b>
pB66_B-40	5p/3p	Mus musculus - Col6a2	1338..2010	IF		99.4	96.8	<b>A</b>
pB66_B-70	5p/3p	Mus musculus - Col6a2	1452..1991	IF		100.0	94.6	<b>A</b>
pB66_B-28	5p/3p	Mus musculus - Col6a2	1524..2014	IF		100.0	97.0	<b>A</b>
pB66_A-33	5p/3p	Mus musculus - Col6a2	1698..1992	IF		99.0	100.0	<b>A</b>

Tabla 9. Resultados del ensayo de doble híbrido en levaduras

Los estudios de doble híbrido en levaduras permitieron la identificación de una única proteína (Col6a2) en la categoría “A”. Se determinó que el SID de Col6a2 implicado en la interacción entre las dos proteínas correspondía a los aminoácidos 567 a 656. También se identificó una potencial interacción (categoría B) de Adamts19 con Col6a1.

Teniendo en cuenta los resultados descritos anteriormente, se procedió a validar por otro método la interacción entre las proteínas humanas ADAMTS19:COL6A2. Para ello, se realizaron diferentes clonajes de ADNc humano con secuencias de regiones homólogas a las del ADNc de *Mus musculus* usadas en el tamizaje de doble híbrido de levaduras (ver los detalles metodológicos más adelante).

En primer lugar, se realizó un análisis similar de doble híbrido en células eucariotas (*CheckMate™ Mammalian Two-Hybrid System*), cuyos detalles metodológicos y resultados se describen a continuación. Posteriormente, se realizaron ensayos de coimmunoprecipitación.

#### **5.8.1.3 Sistema de doble híbrido en células eucariotas (*CheckMate™ Mammalian Two-Hybrid System*).**

El sistema de doble híbrido en levadura realizado en *Hybrigenics* utilizó clones de ADNc de una librería de ovario de *Adamts19-Mus musculus* (*Gene ID*: 240322). Por consiguiente, se quiso realizar la verificación de la interacción proteína: proteína utilizando construcciones plasmídicas de ADNc de *ADAMTS19* y *COL6A2* humano. Para este propósito, se utilizó el sistema de doble híbrido en células eucariotas *CheckMate Mammalian Two-Hybrid System* (Promega). En esta aproximación experimental, cada uno de los ADNc que codifican para las proteínas de interés fue clonado dentro de los vectores pACT o pBIND para generar proteínas fusión, que al interactuar logran reconstituir el factor de transcripción que se une al promotor *GAL4*. El gen reportero que es transactivado por el promotor *GAL4* unido a este factor de transcripción es la luciferasa (figura 4).

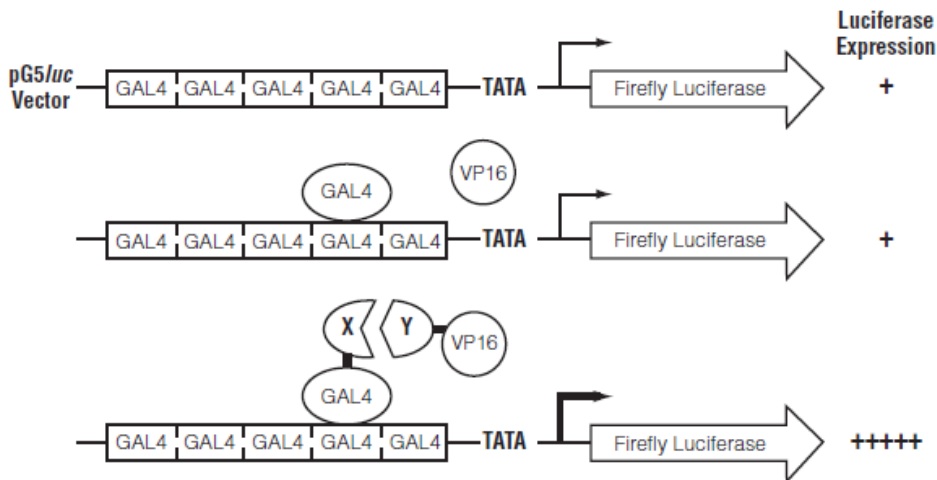


Figura 4. Sistema *CheckMate™ Mammalian Two-Hybrid System*

#### 5.8.1.3.1 Construcciones plasmídicas de ADNc de *ADAMTS19* y *COL6A2* (secuencias humanas).

Los ADNc de *ADAMTS19* (*GenBank accession number* NM\_133638.3) y *COL6A2* (*GenBank accession number* NM\_001849.3), correspondientes a los aminoácidos 634-1169 y 306-649, respectivamente, fueron amplificados usando DNA Taq Polimerasa (GoTaq® Long PCR Master Mix-Promega). El ADNc de *COL6A2* fue obtenido de una biopsia de piel mediante extracción de RNA por trizol y posterior síntesis usando el kit *SuperScrip III First-Strand* (Invitrogen) y Oligo dTs. Por otro lado, el ADNc de *ADAMTS19* se adquirió clonado en el vector pCMV6-AC-GPF (Origene). Los fragmentos de *ADAMTS19* y *COL6A2* fueron clonados en fase en los sitios *EcoRI/XbaI* de los vectores pACT y pBIND respectivamente. Con el objetivo de realizar la mutagénesis de *ADAMTS19* c.2828C>T (mutación encontrada en la paciente FOP), se utilizó el kit *Gene Art* (Invitrogen). Al final se obtuvieron las construcciones pACT-*ADAMTS19*WT, pACT-*ADAMTS19*Mut y pBIND-*COL6A2*.

#### 5.8.1.3.2 Co-transfección de las construcciones plasmídicas

Células CHO (Chinese Hámster Ovary) fueron cultivadas en DMEM-F12 suplementado con SFB al 10% y penicilina/estreptomicina al 1% (Invitrogen-Gibco). Las células fueron sembradas 24 horas antes de la transfección a una densidad de  $6 \times 10^6$  células (en cajas de

24 pozos). Se co-transfectaron 200ng de pACT-ADAMTS19WT ó pACT-ADAMTS19Mut, 200 ng de pBIND-COL6A2 y 200 ng de pG5luc. Cada experimento se realizó al menos tres veces en sextuplicado. Para efectuar el proceso de transfección, se utilizó *FUGENE 6 Transfection Reagent* (Promega). A las 48 horas post-transfección, se realizó un lavado con PBS1X y lisis celular con buffer1X. La actividad del gen reportero luciferasa se cuantificó utilizando el protocolo de *Dual Luciferase Reporter Assay Systems* (promega), mediante las lecturas de Renilla luciferasa y *firefly* luciferasa en un luminómetro GloMax 20/20. Los resultados fueron expresados en unidades relativas de luciferasa (URL) y la significancia estadística fue estimada con la prueba t-student (Figura 5).

#### 5.8.1.4 Resultados del sistema de doble híbrido en células eucariotas *Check Mate Mammalian Two-Hybrid System*

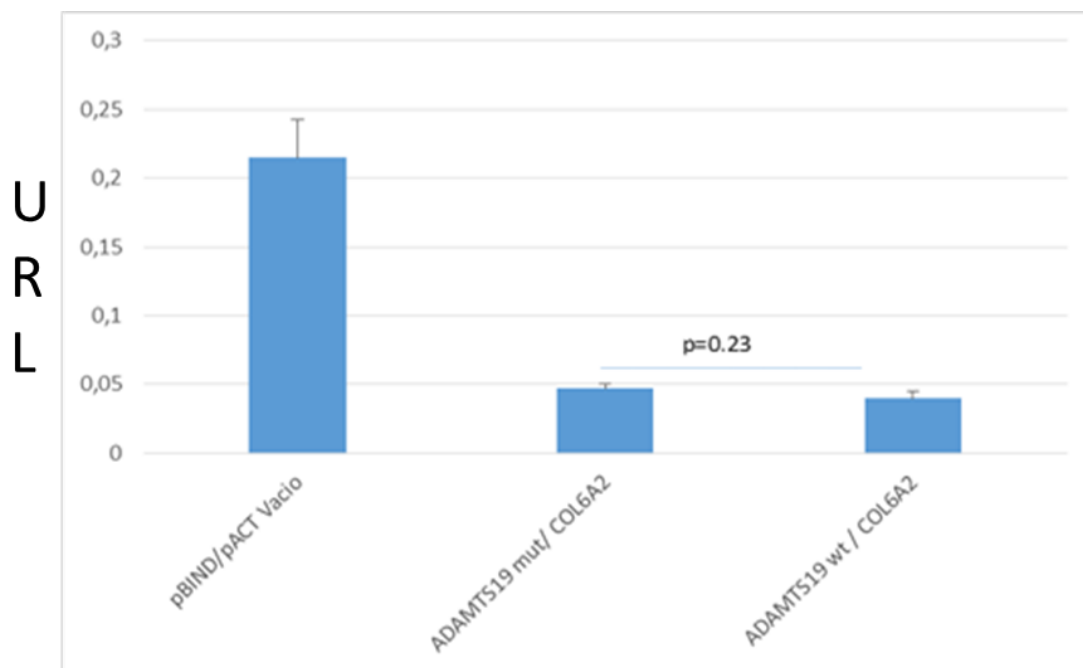


Figura 5. Resultados expresados en URL (unidades relativas de luciferasa) del ensayo doble híbrido en eucariotas.

El sistema de doble híbrido en células eucariotas *CheckMate Mammalian Two-Hybrid System* (Promega), utilizó ADNc de *ADAMTS19WT* ó *ADAMTS19* mutante y *COL6A2* clonados en los vectores pACT y pBIND respectivamente. Los resultados, evaluados en

términos de la cuantificación de las unidades relativas de luciferasa (URL), indicaron que no se produjo la interacción entre las regiones proteicas traducidas a partir de las construcciones transfectadas.

Teniendo en cuenta este resultado, que no permitió establecer la unión entre ADAMTS19 y COL6A2, se procedió a determinar si estos fragmentos proteicos estaban presentes en las células luego de su co-transfección. Para este fin, se efectuaron ensayos de Western Blot (WB) e inmunofluorescencia indirecta (IFI) post-transfección. Las células CHO usadas en los experimentos de WB e IFI no expresan de manera basal las proteínas ADAMTS19 y COL6A2. Esta afirmación es consecuente con los resultados de los ensayos previos realizados de RT-PCR y WB en las células CHO no transfectadas. En RT-PCR no se evidenció la amplificación de *ADAMTS19* ni de *COL6A2* a partir de ADNc obtenido del RNA de las células no transfectadas. El extracto proteico utilizado para WB no indicó la presencia de las proteínas ADAMTS19 y COL6A2.

A continuación se detallan la metodología y los resultados que permiten abordar la pregunta acerca de la producción potencial de fragmentos de proteínas traducidas a partir de las construcciones *pACT-ADAMTS19WT*, *pACT-ADAMTS19Mut* y *pBIND-COL6A2* transfectadas.

### **5.8.1.5 ¿Las proteínas COL6A2 y ADAMTS19 (WT y mutante) son producidas a partir de la transfección de las construcciones pACT-ADAMTS19WT, pACT-ADAMTS19Mut y pBIND-COL6A2?**

#### **5.8.1.5.1 Construcciones plasmídicas**

Las construcciones plasmídicas usadas corresponden a *pACT-ADAMTS19WT*, *pACT-ADAMTS19Mut* y *pBIND-COL6A2*, generadas según lo descrito en el aparte 5.8.1.3.1.

#### **5.8.1.5.2 Cultivo celular y transfecciones**

Células CHO fueron cultivadas en DMEM-F12 suplementado con SFB al 10% y penicilina/estreptomicina al 1% (Invitrogen-Gibco). Las células fueron sembradas 24 horas antes de la transfección a una densidad de  $6 \times 10^6$  células (en cajas de 24 pozos). Se transfectó en las cajas de 24 pozos 300ng de: *pACT-ADAMTS19WT*, *pACT-*

ADAMTS19Mut y pBIND-COL6A2 por independiente para experimentos posteriores de Western blot. Para los experimentos adicionales de colocación mediante inmunofluorescencia indirecta se realizaron también cotransfecciones de 300 ng de las construcciones pACT-ADAMTS19WT ó pACT-ADAMTS19Mut y pBIND-COL6A2. Para el proceso de transfección se utilizó *FUGENE 6 Transfection Reagent* (Promega).

#### **5.8.1.5.3 Western blot y anticuerpos**

Las células CHO fueron lisadas 48 horas después de la transfección en el buffer de lisis RIPA (Na-Ortovanadato 1mM, PMSF 2mM y Coctel inhibidor de proteasas 20uL/mL, Santa Cruz). Los extractos proteicos fueron cuantificados usando el kit Pierce *BCA Protein Assay* (Thermo Scientific). Aproximadamente 30 µg del extracto fueron analizados mediante SDS-PAGE usando geles de acrilamida-bisacrilamida al 5-10%. Las proteínas fueron transferidas a membranas PVDF (Hybond-P GE Healthcare). Se revelaron usando los anticuerpos *ADAMTS-19 goat polyclonal* (sc-21513) o *COL6A2 mouse monoclonal* (sc 374566), mediante el sistema de quimioluminiscencia ECL (*SuperSignal™ West Pico Chemiluminescent substrate – Thermo Scientific*). Los epítopes para los anticuerpos *ADAMTS-19 goat polyclonal* y *COL6A2 mouse monoclonal* corresponden a la región C-terminal y a una región entre los aminoácidos 241 y 540, respectivamente.

#### **5.8.1.5.4 Inmunofluorescencia indirecta**

Las células CHO fueron lavadas con PBS 1X 48 horas después de la transfección, fijadas durante 25 minutos con PFA 4%, permeabilizadas con Tritón 1% y bloqueadas con BSA 1%. Las células fueron incubadas toda la noche a 4°C con los anticuerpos primarios anti ADAMTS-19:1/300 (sc 21513) o anti- COL6A2: 1/300 (sc 374566). Se realizaron tres lavados en PBS1X y se incubó durante 2 horas con los anticuerpos secundarios *Donkey anti-goat IgG CFL594:1/250* (sc 362275) y *goat anti-mouse IgG CFL 48: 1/250* (sc 362257) respectivamente. El montaje se realizó con *Ultra Cruz® Mounting Medium* (sc-24941) que contiene DAPI para teñir los núcleos. Las células fueron visualizadas en un microscopio de fluorescencia Nikon Eclipse NiE y las imágenes fueron capturadas con una cámara fotométrica digital (Coolnap EZ) usando el *NIS-Elements Advance Research software*.

## 5.8.1.6 Resultados

### 5.8.1.6.1 Western blot

A continuación se presentan los resultados del Western blot realizado para verificar la producción de las proteínas fusión luego de la transfección de las construcciones pACT-ADAMTS19WT, pACT-ADAMTS19Mut y pBIND-COL6A2.



Figura 6. Western blot. Extracto de proteínas de células CHO transfectadas con pACT-ADAMTS19WT, pACT-ADAMTS19Mut y pBIND-COL6A2. Control:  $\beta$ -Actina.

Los resultados del western blot indican que se produjo la traducción de las proteínas luego de la transfección en células CHO de las construcciones pACT-ADAMTS19WT, pACT-ADAMTS19Mut y pBIND-COL6A2. El peso molecular de cada una de las proteínas identificadas correspondió al esperado. Para COL6A2 fue de 39.7 kDa y para ADAMTS19 de 75 kDa. Fue utilizado como control positivo, la  $\beta$ -actina.

Esta verificación indica que las construcciones plasmídicas utilizadas en el sistema de doble híbrido en células eucariotas *CheckMate Mammalian Two-Hybrid*, producen las proteínas esperadas.

### 5.8.1.6.2 Inmunofluorescencia indirecta

A continuación se presentan los resultados de la inmunofluorescencia indirecta realizada para la verificación de la producción y la potencial colocalización de las proteínas producidas luego de la transfección de las construcciones pACT-ADAMTS19WT, pACT-ADAMTS19Mut y pBIND-COL6A2.

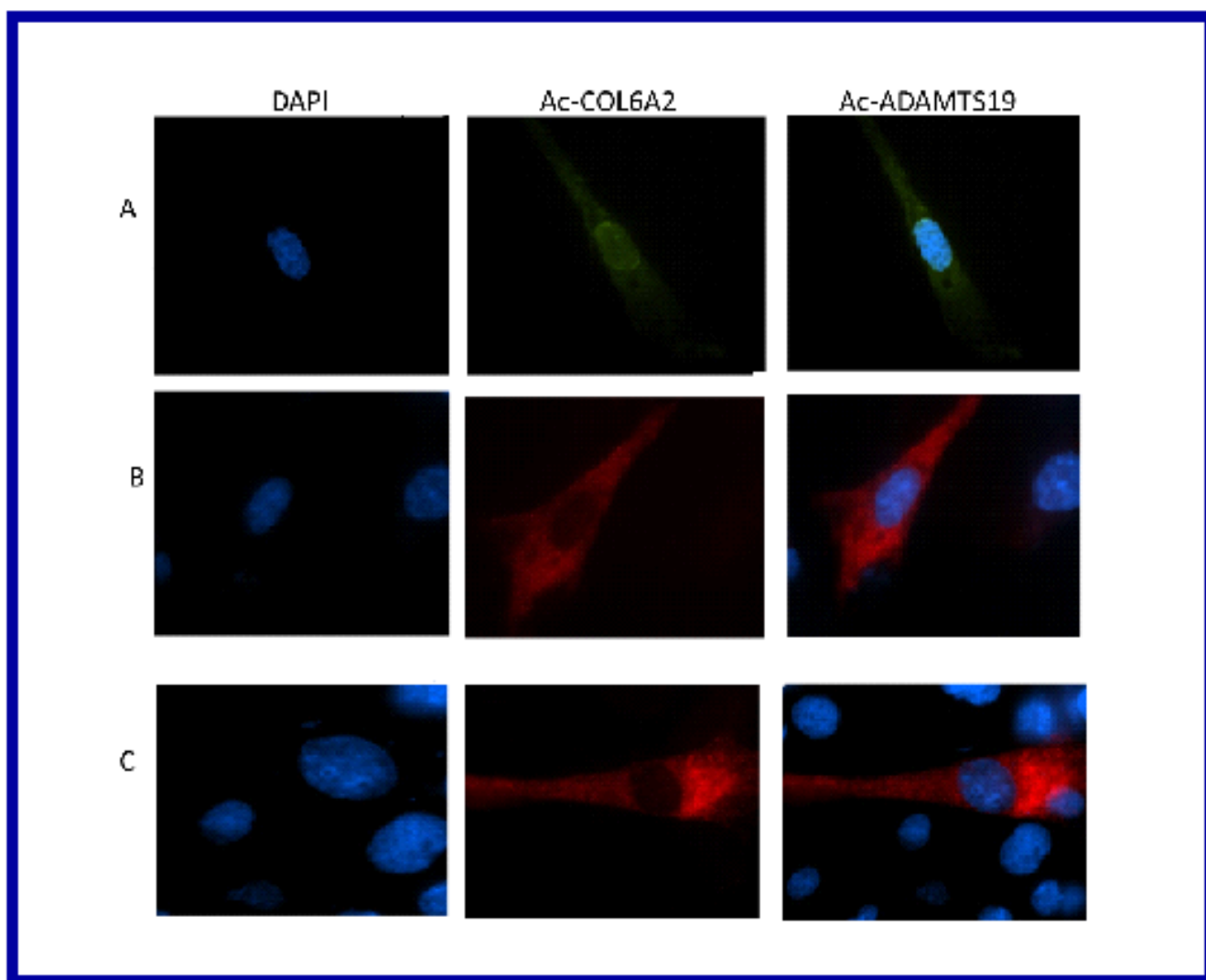


Figura 7. Resultados de la inmunofluorescencia, luego de la transfección en células CHO. Se observa localización citoplasmática de las dos proteínas. No hay diferencia entre ADAMTS19WT y Mut. A. pBIND-COL6A2; B. pACT-ADAMTS19WT y C. pACT-ADAMTS19Mut. Fotos con Microscopio de fluorescencia Nikon Eclipse NiE. 40X.

Los resultados de la inmunofluorescencia indirecta presentados en la figura 7, indican que se realizó la traducción de las proteínas luego de la transfección en células CHO de las construcciones pACT-ADAMTS19WT, pACT-ADAMTS19Mut y pBIND-COL6A2. Se evidencia que la localización de las dos proteínas es citoplasmática, con distribución homogénea. No se observó diferencia entre las proteínas ADAMTS19 WT y mutantes.

Esta verificación es coherente con los resultados obtenidos con el WB (descrito en la sección anterior) e indica que las construcciones plasmídicas utilizadas en el sistema de doble híbrido en células eucariotas *CheckMate Mammalian Two-Hybrid*, producen las proteínas esperadas.

A continuación se presentan los resultados de la inmunofluorescencia indirecta realizada para el estudio de la potencial colocalización de las proteínas luego de la transfección de las construcciones pACT-ADAMTS19WT ó pACT-ADAMTS19Mut y pBIND-COL6A2.

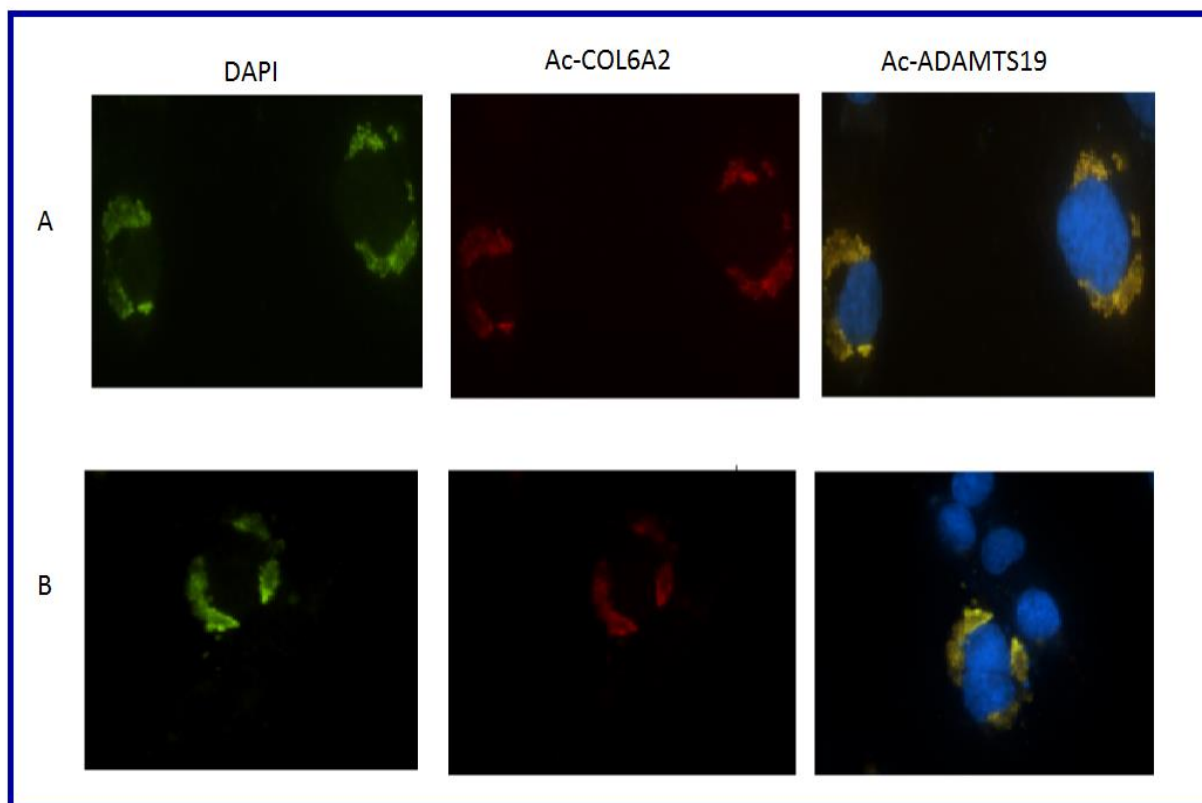


Figura 8. COL6A2 se colocaliza y forma agregados citoplasmáticos con ADAMTS19 wild type y con ADAMTS19 mutante en células CHO co-transfectadas con las construcciones plasmídicas **A.** pBIND-COL6A2 y pACT-ADAMTS19WT y **B.** pBIND-COL6A2 y pACT-ADAMTS19Mut. Fotos obtenidas con Microscopio de fluorescencia Nikon Eclipse NiE. 40X.

Los resultados de la inmunofluorescencia indirecta presentados en la figura 8, indican que se realizó la traducción de las proteínas luego de la co-transfección en células CHO de las construcciones pACT-ADAMTS19WT o pACT-ADAMTS19Mut con pBIND-COL6A2. Se observa la colocalización de las dos proteínas. Sin embargo, se producen imágenes similares a las de agregados citoplasmáticos lo que implica una clara diferencia con la distribución homogénea de las proteínas luego de transfectarse de manera individual (Figura 7). Con este hallazgo podemos hipotetizar que las proteínas producidas tienden a agregarse lo que eventualmente podría afectar su adecuada interacción. Esto podría explicar los resultados negativos del experimento del sistema de doble híbrido en células eucariotas *CheckMate Mammalian Two-Hybrid*.

Con el objetivo de realizar una segunda verificación de la interacción entre ADAMTS19 y COL6A2, se realizaron ensayos de coinmunoprecipitación utilizando fragmentos de estas proteínas. Para este fin, se crearon construcciones plasmídicas con ADNc de *ADAMTS19* y *COL6A2* de origen humano clonado en el vector pCDNA3.0. El detalle metodológico y la presentación de los resultados se describen a continuación.

## **5.8.2 ¿Es posible identificar la interacción proteína:proteína y la potencial colocalización de fragmentos humanos de ADAMTS19WT o ADAMTS19Mut y COL6A2 ?**

### **5.8.2.1 Construcciones plasmídicas**

Los ADNc de *ADAMTS19* (*GenBank accession number* NM\_133638.3) y *COL6A2* (*GenBank accession number* NM\_001849.3), correspondientes a los aminoácidos 634-1169 y 306-649 respectivamente, fueron amplificados usando DNA Taq Polimerasa (GoTaq® Long PCR Master Mix-Promega). En el diseño de los primers se tuvo en cuenta la inclusión del ATG de inicio en una secuencia kozac, para garantizar la adecuada traducción de los fragmentos de ADNc de *ADAMTS19*WT y mutante así como de *COL6A2*. Los fragmentos de *ADAMTS19* y *COL6A2* fueron clonados en fase en los sitios *EcoRI/XbaI* del vector pCDNA3.0. Se obtuvieron las construcciones *ADAMTS19*-WT-pCDNA3.0, *ADAMTS19*-T943I-pCDNA3.0 y *COL6A2*-pCDNA3.0.

### 5.8.2.2 Cultivo celular y transfecciones

Células CHO fueron cultivadas en DMEM-F12 suplementado con SFB al 10% y penicilina/estreptomicina al 1% (Invitrogen-Gibco). Las células fueron sembradas 24 horas antes de la transfección a una densidad de  $6 \times 10^6$  células (en cajas de 24 pozos). Se transfectaron 300ng de: *ADAMTS19*-WT-pCDNA3.0, *ADAMTS19*-T943I-pCDNA3.0 y *COL6A2*-pCDNA3.0 por independiente. Además se realizaron co-transfecciones de las construcciones: *ADAMTS19*-WT-pCDNA3.0 /*COL6A2*-pCDNA3.0 y *ADAMTS19*-T943I-pCDNA3.0 /*COL6A2*-pCDNA3.0. Para el proceso de transfección se utilizó *FUGENE 6 Transfection Reagent* (Promega).

### 5.8.2.3 Inmunoprecipitación y anticuerpos

Las células CHO fueron lisadas 48 horas después de la transfección en el buffer de lisis RIPA (Na-Ortovanadato 1mM, PMSF 2mM y Coctel inhibidor de proteasas 20uL/mL, Santa Cruz). Los extractos proteicos fueron cuantificados usando el kit Pierce *BCA Protein Assay* (Thermo Scientific). Aproximadamente 400 µg de proteínas fueron inmunoprecipitadas mediante incubación a 4°C por toda la noche con los complejos *Immunocruz IP/WB Optima E system (sc-45042)*/anticuerpo primario COL6A2 (2ug) (sc 374566) e *Immunocruz IP/WB Optima D system (sc-45041)*/anticuerpo primario ADAMTS-19 (2ug) (sc 21513). Luego de realizar cuatro lavados con buffer RIPA, las proteínas unidas fueron sometidas a ebullición en 50 µl de buffer Laemmli-SDS 2X. Los eluidos fueron separados mediante SDS-PAGE usando geles de Acrilamida-bisacrilamida 5-10%. Las proteínas fueron transferidas a membranas PVDF (Hybond-P GE Halthcare). Se revelaron usando los anticuerpos *ADAMTS-19 goat polyclonal (sc-21513)* o *COL6A2 mouse monoclonal (sc 374566)* mediante el sistema de quimioluminiscencia ECL (*SuperSignal<sup>TM</sup> West Pico Chemiluminescent substrate-Thermo Scientific*). El mismo procedimiento se realizó para el lisado de proteínas obtenido a partir de líquido folicular de una paciente sana donante de óvulos.

#### 5.8.2.4. Resultados de la coimmunoprecipitación

A continuación se presentan los resultados de la coimmunoprecipitación de las proteínas ADAMTS19Wt y ADAMTS19Mut con COL6A2. Las proteínas fueron obtenidas luego de la cotransfección de los plásmidos *ADAMTS19*-WT-pCDNA3.0, *ADAMTS19*-T943I-pCDNA3.0 y *COL6A2*-pCDNA3.0 en células CHO.

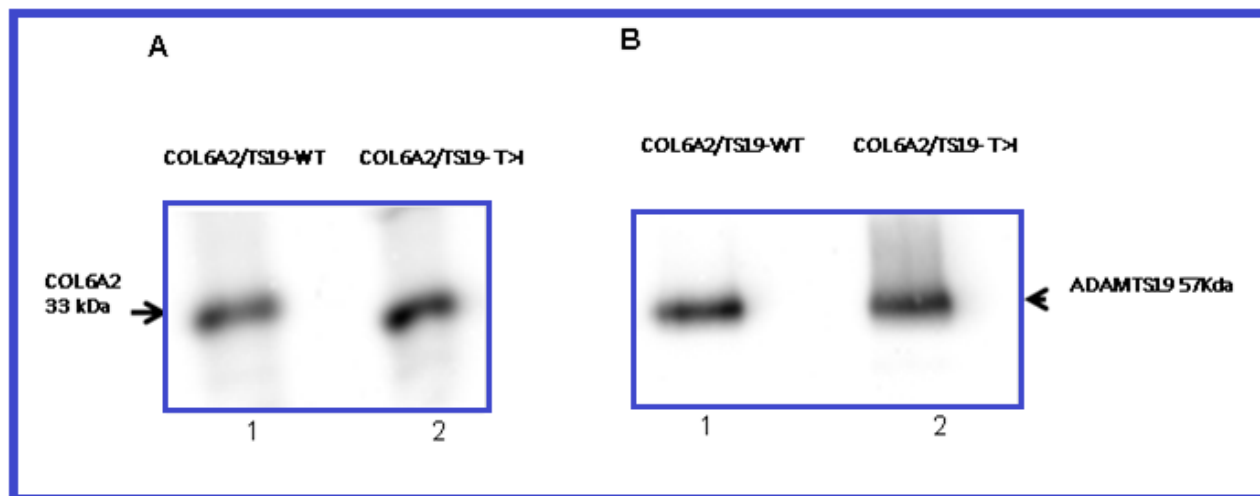


Figura 9. COL6A2 interacciona con ADAMTS19 wild type (TS19-WT) y con ADAMTS19 p.Thr943Ile (TS19-T>I) en células CHO transfectadas **A.** El complejo de inmunoprecipitación fue incubado con un anticuerpo Anti-ADAMTS19 y en el análisis de Western blot se detectó con un anticuerpo anti-COL6A2. **B.** El complejo de inmunoprecipitación fue incubado con un anticuerpo anti-COL6A2 y en el análisis de Western blot se detectó con un anticuerpo anti-ADAMTS19

Los resultados del estudio de la interacción de los fragmentos humanos de ADAMTS19 WT o mutante con COL6A2 mediante coimmunoprecipitación fueron positivos. Esta interacción no había sido previamente reportada, ya que hasta la fecha la proteína ADAMTS19 había sido considerada como huérfana. El análisis de coimmunoprecipitación no indicó diferencias de interacción con COL6A2 de los fragmentos de proteínas generadas a partir de las construcciones *ADAMTS19*-WT-pCDNA3.0 o *ADAMTS19*-T943I-pCDNA3.0. De esta manera, estimamos que la mutación p.Thr973Ile no afecta la interacción entre la proteína ADAMTS y su *partner* COL6A2.

Con el objetivo de estudiar una colocalización de ADAMTS19WT o mutante con COL6A2, se realizaron ensayos de inmunofluorescencia indirecta post-transfección de las mismas regiones utilizadas en el ensayo de coimmunoprecipitación. El detalle metodológico y la presentación de los resultados se describen a continuación.

#### **5.8.2.5 Inmunofluorescencia indirecta**

Células CHO fueron lavadas con PBS 1X 48 horas después de la transfección, fijadas durante 25 minutos con PFA 4%, permeabilizadas con Tritón 1% y bloqueadas con BSA 1%. Las células fueron incubadas toda la noche a 4°C con los anticuerpos primarios anti ADAMTS-19:1/300 (sc 21513) o anti- COL6A2: 1/300 (sc 374566). Se realizaron tres lavados en PBS1X y se incubó durante 2 horas con los anticuerpos secundarios *Donkey anti-goat IgG CFL594*:1/250 (sc 362275) y *goat anti-mouse IgG CFL 48*: 1/250 (sc 362257) respectivamente. El montaje se realizó con *UltraCruz*® *Mounting Medium* (sc-24941) que contiene DAPI para teñir los núcleos. Las células fueron visualizadas en un microscopio de fluorescencia Nikon Eclipse NiE y las imágenes fueron capturadas con una cámara fotométrica digital (Coolnap EZ) usando el *NIS-Elements Advance Research software*.

#### **5.8.2.6 Resultados de la inmunofluorescencia indirecta**

A continuación se presentan los resultados de la inmunofluorescencia indirecta de las proteínas ADAMTS19WT y ADAMTS19Mut con COL6A2. Las proteínas fueron obtenidas luego de la cotransfección de los plásmidos *ADAMTS19-WT-pCDNA3.0*, *ADAMTS19-T943I-pCDNA3.0* y *COL6A2-pCDNA3.0* en células CHO.

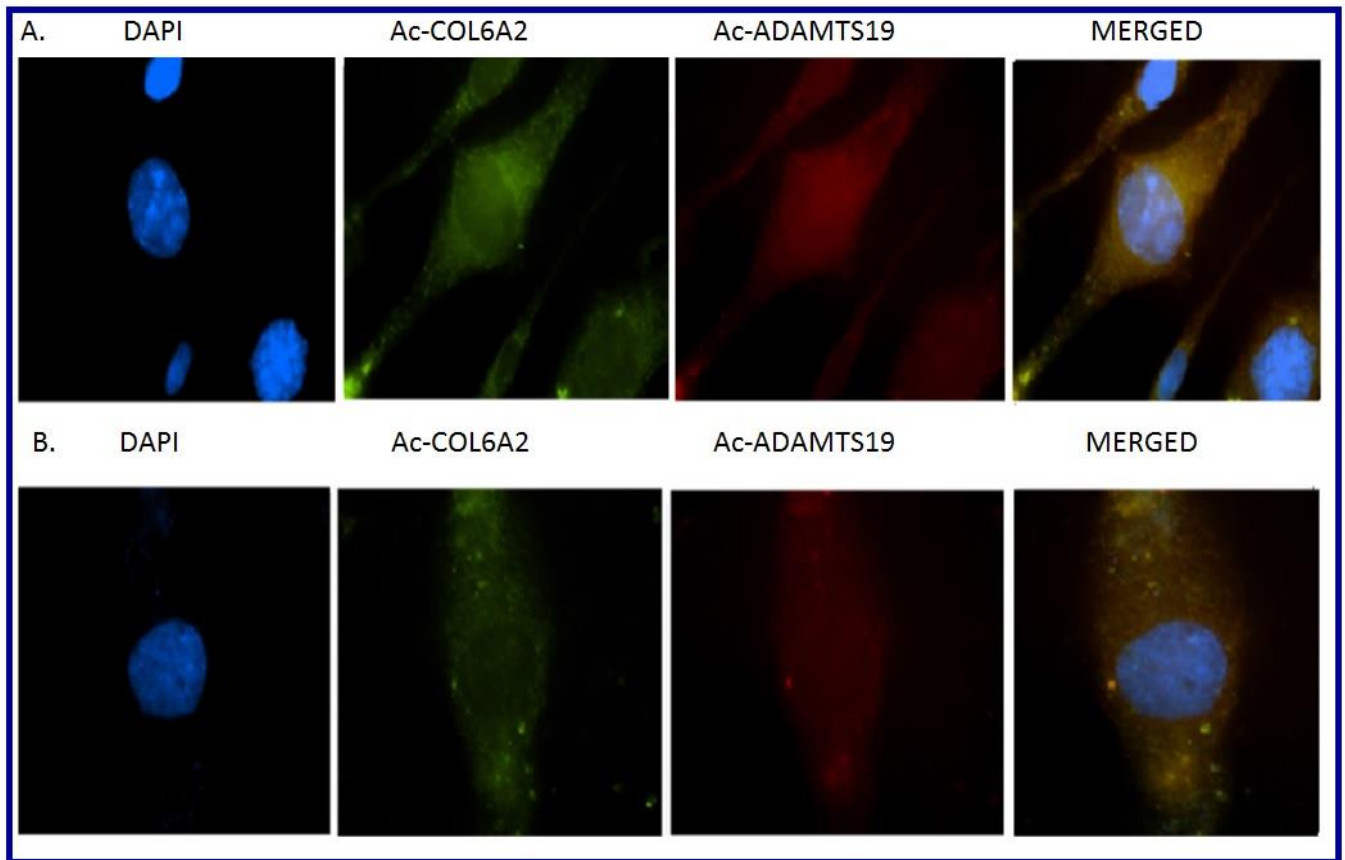


Figura 10. COL6A2 se colocaliza con ADAMTS19 wild type (TS19-WT) y con ADAMTS19 p.Thr943Ile (TS19-T>I) en células CHO co-transfectadas con las construcciones plasmídicas **A.** *ADAMTS19*-WT-pCDNA3.0 y *COL6A2*-pCDNA3.0 y **B.** *ADAMTS19*-T943I-pCDNA3.0 y *COL6A2*-pCDNA3.0. Fotos con Microscopio de fluorescencia Nikon Eclipse NiE. 40X.

Los experimentos de inmunofluorescencia indirecta indicaron una colocalización de los fragmentos de proteínas generados de las construcciones plasmídicas co-transfectadas *ADAMTS19*-WT-pCDNA3.0, *ADAMTS19*-T943I-pCDNA3.0 con *COL6A2*-pCDNA3.0. La distribución fue citoplasmática con un patrón uniforme. Los resultados no indicaron diferencias entre los fragmentos de proteínas generadas de las construcciones *ADAMTS19*-

WT-pCDNA3.0 y *ADAMTS19*-T943I-pCDNA3.0. De esta manera, estimamos que la mutación p.Thr973Ile no afecta la colocalización de *ADAMTS19* y *COL6A2*.

### **5.8.3 ¿Los genes *ADAMTS19* y *COL6A2* se coexpresan en una muestra de líquido folicular humano? ¿Se observa el mismo efecto en una muestra de ovario de ratón?**

Teniendo en cuenta que el gen *ADAMTS19* fue propuesto como un candidato nuevo, cuyas mutaciones pudieran ser responsables de la etiología de la falla ovárica prematura, fue necesario determinar si este gen se co-expresa con *COL6A2* en un contexto ovárico. El detalle metodológico y la presentación de los resultados se describen a continuación.

#### **5.8.3.1 Extracción de RNA, síntesis de ADNc y PCR para identificación de *Adamts19* y *Col6a2* en ovarios de ratón y humano**

RNA total de ovario de ratón Balb/c de 8 semanas y de líquido folicular humano (obtenido de una donante de óvulo- Centro Colombiano de Fertilidad y Esterilidad CECOLFES) fue extraído usando TRIzol (Invitrogen), de acuerdo a las recomendaciones de la casa comercial. La síntesis del ADNc total fue realizada con el kit *SuperScript III First-Strand* (Invitrogen) y oligo dTs. El ADNc fue amplificado con primers específicos de ratón y de humano, que fueron previamente diseñados con el software primer 3 (tabla 10 y tabla 11 respectivamente). La detección del producto amplificado se realizó en geles de agarosa al 1% teñidos con bromuro de etidio. Los productos de PCR fueron purificados mediante fosfatasa alcalina y exonucleasa I antes de la secuenciación. El análisis de las secuencias de nucleótidos fue realizado a través de alineamientos múltiples usando el programa clustalW. Los cromatogramas fueron analizados manualmente.

<i>mAdamts19F</i> 5'GCTAGGTTCTGAAGCAAGGGA 3'
<i>mAdamts19R</i> 5'GTGCATGCTTTGGGTAC 3'
<i>mCol6a2F</i> 5'TTGGATCGGATGGTCGGAAG 3'
<i>mCol6a2R</i> 5'CTCCAGGGTATCCATCGGGG 3'

Tabla 10. Primers usados en la amplificación del ADNc de ovario de ratón Balb/c.

<i>hADAMTS19F</i> 5'ACATGACTTCAGTGGATGCAGC 3'
<i>hADAMTS19R</i> 5'CTGCTTGCTACAAACAAATCCGC 3'
<i>hCOL6A2F</i> 5'CCGTGGACAGAAGGGTGCCAAG 3'
<i>hCOL6A2R</i> 5'CAACAGTGGAGCCCCAGGAAGTC 3'

Tabla 11. Primers usados en la amplificación del ADNc de líquido folicular humano.

### 5.8.3.2 Resultados

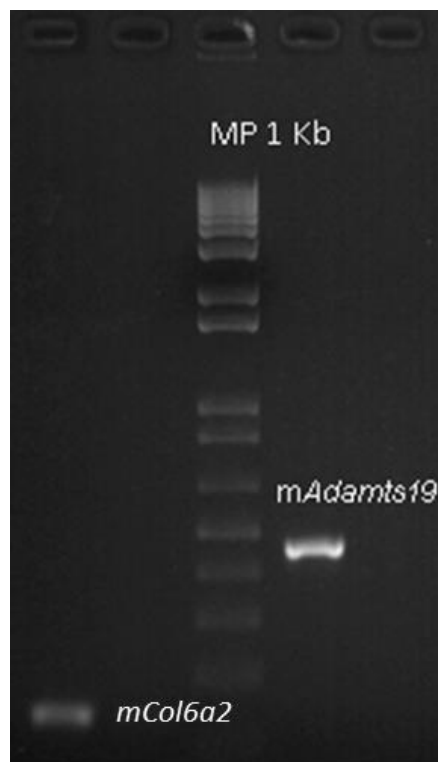


Figura 11. Los genes *mCol6a2* y *mAdamts19* se coexpresan en ovario de ratón.

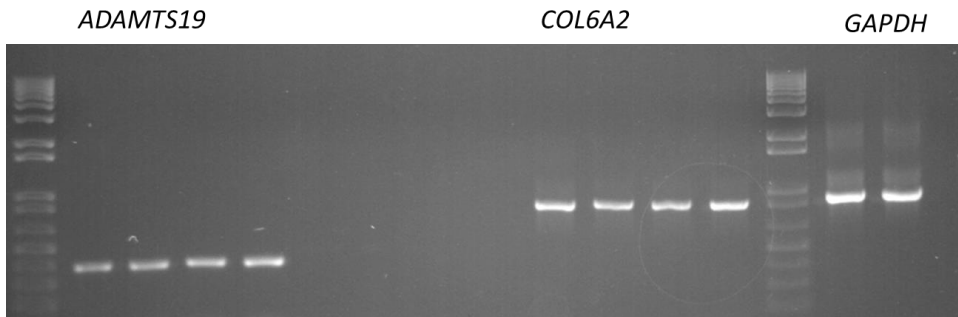


Figura 12. Los genes *COL6A2* y *ADAMTS19* se coexpresan en una muestra de líquido folicular. Se observa la amplificación a varias cantidades de ADNc.

#### **5.8.4 ¿Las proteínas COL6A2 y ADAMTS19 están presentes en el líquido folicular humano? ¿Si es el caso, existe una interacción entre ellas?**

Con la evidencia de que los transcritos de *ADAMTS19* y *COL6A2* se coexpresaban en una muestra de líquido folicular humano, se estudió mediante western blot la presencia a nivel proteico de las proteínas correspondientes. Además, se realizó una coimmunoprecipitación de las dos proteínas obtenidas del líquido folicular. El detalle metodológico y la presentación de los resultados se describen a continuación.

##### **5.8.4.1 Inmunoprecipitación y anticuerpos**

El líquido folicular fue extraído en una paciente sana, donante de óvulos en el Centro Colombiano de Fertilidad y Esterilidad (CECOLFES). Posteriormente, 1 ml de la muestra fue lisada en el Buffer de lisis RIPA (Na-Ortovanadato 1mM, PMSF 2mM y Coctail inhibidor de proteasas 20µL/mL, Santa Cruz). Los extractos proteicos fueron cuantificados usando el kit Pierce *BCA Protein Assay* (Thermo Scientific). Aproximadamente 400 µg de proteínas fueron inmunoprecipitadas mediante incubación a 4°C por toda la noche con los complejos *Immunocruz IP/WB Optima E system (sc-45042)*/anticuerpo primario COL6A2 (2µg) (sc 374566) e *Immunocruz IP/WB Optima D system (sc-45041)*/anticuerpo primario ADAMTS-19 (2µg) (sc 21513). Luego de realizar cuatro lavados con buffer RIPA, las proteínas unidas fueron sometidas a ebullición en 50 µl de buffer Laemmli-SDS 2X. Los eluidos fueron separados mediante SDS-PAGE usando geles de acrilamida-bisacrilamida 5-10%. Las proteínas fueron transferidas a membranas PVDF (Hybond-P GE Healthcare). Se

revelaron usando los anticuerpos *ADAMTS-19 goat polyclonal* (sc-21513) o *COL6A2 mouse monoclonal* (sc 374566) mediante el sistema de quimioluminiscencia ECL (*SuperSignal<sup>TM</sup> West Pico Chemiluminescent substrate-Thermo Scientific*).

#### 5.8.4.2. Resultados western Blot

A continuación se presentan los resultados del Western blot realizado de extractos proteicos obtenidos de líquido folicular.



Figura 13. Western blot. Extracto de proteínas totales de líquido folicular humano.

El extracto de proteínas fue inmunoprecipitado con *Immunocruz IP/WB Optima D system* (sc-45041)/anticuerpo primario ADAMTS-19 (2µg) (sc 21513) o *Immunocruz IP/WB Optima E system* (sc-45042)/anticuerpo primario COL6A2 (2µg) (sc 374566). El inmunoblot fue revelado con los anticuerpos *ADAMTS-19 goat polyclonal* (sc-21513) o *COL6A2 mouse monoclonal* (sc 374566) respectivamente.

Los resultados obtenidos indicaron que las proteínas ADAMTS19 y COL6A2 se expresaban en muestras de líquido folicular humano. Este hallazgo no se había documentado previamente. En este marco, los resultados permiten hipotetizar que el complejo ADAMTS19/COL6A2 participa en el proceso de recambio de la matriz extracelular de los folículos maduros en fase de ovulación (ruptura del folículo de Graaf?).

### 5.8.4.3. Resultados de la co-inmunoprecipitación en líquido folicular humano

A continuación se presentan los resultados de coimmunoprecipitación realizado de extractos proteicos obtenidos de líquido folicular.

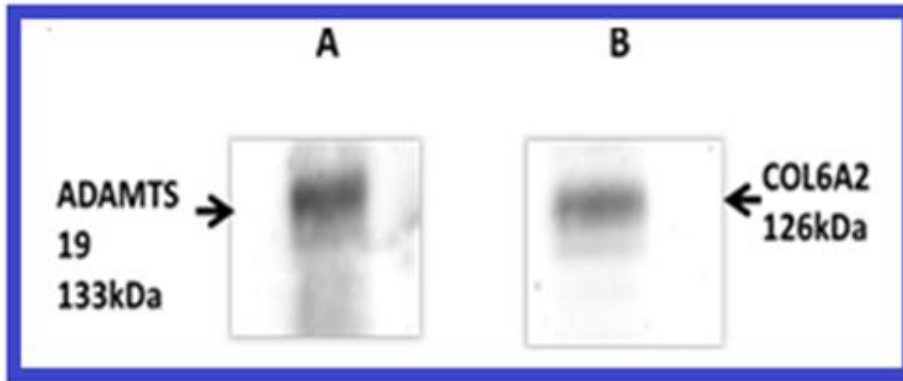


Figura 14. Coimmunoprecipitación de ADAMTS19 y COL6A2 en una muestra de líquido folicular humano.

- A. El complejo de inmunoprecipitación fue incubado con *Immunocruz IP/WB Optima E system (sc-45042)*/anticuerpo primario COL6A2 (sc 374566) y el inmunoblot fue revelado con el anticuerpo *ADAMTS-19 goat polyclonal (sc-21513)*.
- B. El complejo de inmunoprecipitación fue incubado con *Immunocruz IP/WB Optima D system (sc-45041)*/anticuerpo primario ADAMTS-19 (sc 21513) y el inmunoblot fue revelado con el anticuerpo *COL6A2 mouse monoclonal (sc 374566)*.

Los resultados obtenidos de la coimmunoprecipitación de ADAMTS19 y COL6A2 indicaron que en un contexto ovárico como el líquido folicular, estas dos proteínas interactúan. En concordancia con los hallazgos de co-expresión tanto a nivel transcriptómico como proteómico, la validación funcional de la interacción entre ADAMTS19 y COL6A2 permite inferir acerca de su potencial función en esta fase del desarrollo folicular.

#### **5.8.4.4 Discusión de los resultados obtenidos de la validación funcional de la mutación *ADAMTS19* c.2828C>T**

Las proteínas ADAMTS han sido relacionadas con la degradación de la matriz extracelular actuando sobre sustratos específicos (Kelwick et al, 2015). De manera similar, se ha descrito que miembros de la familia de colágenos están presentes en la matriz extracelular cumpliendo funciones de proliferación, esteroidogénesis y luteinización, entre otros. Teniendo en cuenta que el ovario es un tejido en donde distintas estructuras, como los folículos y el cuerpo lúteo, están en constante recambio, el adecuado funcionamiento de los componentes de la matriz extracelular es fundamental (Rodgers, et al 2003). Desde esta perspectiva, las variaciones en genes como *ADAMTS19* pueden afectar los procesos normales de recambio de la matriz. En este contexto, se han descrito varios miembros de la familia ADAMTS con funciones críticas en la expansión y en la formación del complejo del COC, como la degradación de sustratos requeridos en el remodelamiento de la matriz extracelular (como el versican), en la ovulación y en la angiogénesis (Demircan et al, 2014). La expresión de las proteínas ADAMTS 1, 5, 4, 9, 16 y 17 se ha identificado en muestras de ovario sugiriendo una función clave de esta familia proteica en el contexto de la reproducción. En pacientes afectadas con síndrome de ovario poliquístico (PCOS) la expresión de *ADAMTS1* está significativamente sub-regulada. Además, en estas pacientes la expresión disminuida de *ADAMTS9* se ha asociado con anomalías meióticas de los oocitos (Russell et al, 2015). En cuanto a la etiología de la FOP, *ADAMTS19* ha sido propuesto como un candidato a través de estudios de GWAS y de asociaciones diplóticas con los genes *IGF2* y *ACVR2B* (Pyun et al, 2013; Pyun et al, 2015). Estimamos que *ADAMTS19* puede estar involucrada en la función de recambio de la matriz extracelular posiblemente durante procesos de maduración folicular/ovulación. El hallazgo descrito en este trabajo de la co-expresión transcriptómica/proteómica de estas moléculas en el líquido folicular puede ser un indicativo de su función en este contexto. Es posible, como algunos estudios lo sugieren, que la formación del líquido folicular requiere de un proceso activo de remodelamiento celular que permita el transporte de fluidos dentro del folículo (Rodgers et al. 2010). *ADAMTS19* podría participar de este evento fisiológico y su disfunción podría

originar fenotipos ováricos. Por último no descartamos que ADAMTS19 pueda actuar en otras etapas del desarrollo folicular.

**Anexo 5.8.** Lista de artículos científicos adicionales publicados durante la permanencia en el programa doctoral.

- **Fonseca DJ**, Prada CF, Siza LM, Angel D, Gomez YM, Restrepo CM, Douben H, Rivadeneira F, de Klein A, Laissue P. (2012). A de novo 14q12q13.3 interstitial deletion in a patient affected by a severe neurodevelopmental disorder of unknown origin. *American Journal of Medical Genetics*,158A: 689-93.
- Diggle CP, Parry DA, Logan CV, Laissue P, Rivera C, Restrepo CM, **Fonseca DJ**, Morgan JE, Allanore Y, Fontenay M, Wipff J, Varret M, Gibault L, Dalantaeva N, Korbonits M, Zhou B, Yuan G, Harifi G, Cefle K, Palanduz S, Akoglu H, Zwijnenburg PJ, Lichtenbelt KD, Aubry-Rozier B, Superti-Furga A, Dallapiccola B, Accadia M, Brancati F, Sheridan EG, Taylor GR, Carr IM, Johnson CA, Markham AF, Bonthron DT. (2012). Prostaglandin transporter mutations cause pachydermoperiostosis with myelofibrosis. *Human Mutation*, 33:1175-81
- Ortega-Recalde O, **Fonseca DJ**, Patiño LC, Atuesta JJ, Rivera-Nieto C, Restrepo CM, Mateus HE, van der Knaap MS, Laissue P. (2013). A novel familial case of diffuse leukodystrophy related to NDUFV1 compound heterozygous mutations. *Mitochondrion*,13:749-54.
- Castro T, Mateus HE, **Fonseca DJ**, Forero D, Restrepo CM, Talero C, Vélez A, Laissue P. (2013). Sequence analysis of the ADRA2A coding region in children affected by attention deficit hyperactivity disorder. *Neurological Sciences*,34: 2219-22.
- Ortega-Recalde O, Vergara JI, **Fonseca DJ**, Ríos X, Mosquera H, Bermúdez OM, Medina CL, Vargas CI, Pallares AE, Restrepo CM, Laissue P. (2013). Whole-exome sequencing enables rapid determination of xeroderma pigmentosum molecular etiology. *PLoS One*,8: e64692.

- Niño MY, Mateus HE, **Fonseca DJ**, Kroos MA, Ospina SY, Mejía JF, Uribe JA, Reuser AJ, Laissue P.(2013). Identification and Functional Characterization of GAA Mutations in Colombian Patients Affected by Pompe Disease. *Journal of Inherited Metabolic Disease Reports*, 7:39-48.
- **Fonseca DJ**, Rojas RF, Vergara JI, Ríos X, Uribe C, Chávez L, Velandia F, Vargas CI, Restrepo CM, Laissue P.(2013). A severe familial phenotype of Ichthyosis Curth-Macklin caused by a novel mutation in the KRT1 gene. *British Journal of Dermatology*, 168:456-8.
- Ortega-Recalde O, Moreno MB, Vergara JI, **Fonseca DJ**, Rojas RF, Mosquera H, Medina CL, Restrepo CM, Laissue P.( 2015). A novel TGM1 mutation, leading to multiple splicing rearrangements, is associated with autosomal recessive congenital ichthyosis. *Clinical and Experimental Dermatology*,40:757-60.

## 6. CONCLUSIONES

1. En lo referente a *CDKN1B*, una de las variantes identificadas fue exclusiva de los casos y no había sido previamente reportada en las bases de datos de SNP en ninguna población. El cambio corresponde a una mutación no-sinónima en estado heterocigoto: c.356T>C, p.Ile119Thr. Los alineamientos de las secuencias proteicas en esta región indicaron que la isoleucina en la posición 119 está conservada de manera estricta en algunas especies de mamíferos (*Homo sapiens*, *Sus scrofa*, *Mus musculus*, *Rattus norvegicus*, *Felis catus*). La mutación identificada sugiere un efecto deletéreo. La sustitución de una isoleucina por una treonina implica un cambio de hidrofobicidad, ya que la cadena lateral de los residuos es diferente en su polaridad. En esta posición (119), la proteína CDKN1B establece una conformación de hoja beta que podría verse afectada por la diferencias de hidrofobicidad de los aminoácidos involucrados. Además, la conservación de la posición 119 a lo largo de la evolución, puede indicar un rol funcional en la etiología de la FOP.
2. En cuanto a *CITED2*, se identificó la mutación c.604C>A (*p.Pro202Thr*) heterocigota de manera exclusiva en una paciente afectada. El análisis comparativo de la proteína entre diferentes especies de vertebrados demostró una conservación estricta del residuo Pro en la posición 202. El software Polyphen2 clasificó esta mutación como probablemente patogénica. La sustitución de prolina por treonina implica un cambio importante en términos de estructura e hidrofobicidad, ya que el primer aminoácido es no-polar y el segundo es polar. Esta sustitución puede estar relacionada con un inadecuado plegamiento de la proteína, lo que podría generar una alteración de su función. La mutación descrita se encuentra próxima a la región *serine-glycine rich junction (SGJ)* de la proteína (aa 162-199), la cual ha sido descrita como un hot-spot mutacional en pacientes afectados por cardiopatías congénitas (Sperling et al., 2005). En el estudio de Sperling et al. el análisis funcional indicó que los mutantes reducían la capacidad de trans-represión de *CITED2* sobre HIF1A. Este resultado permitió proponer que las mutaciones de la región SGJ pueden alterar la capacidad de unión de *CITED2* con cofactores esenciales (Sperling et al., 2005). A pesar de que a la fecha las mutaciones de *CITED2* solo se han asociado con anomalías cardíacas, es posible que

mutaciones localizadas en la región SGJ puedan afectar también el contexto ovárico. La haploinsuficiencia de *CITED2* podría afectar la expresión de factores que participan en el programa de diferenciación gonadal femenina como *SFI*, *FOXL2*, *RSPO1* y *WNT4*.

3. Al referirnos a la secuenciación de la región codificante de *FOXO4*, concluimos que en la población de Túnez (analizada en el presente trabajo doctoral) *FOXO4* no está relacionada frecuentemente con la etiología de la FOP, lo cual no descarta que variantes potencialmente patogénicas afecten mujeres de características étnicas diferentes a las estudiadas y que las mutaciones en este gen puedan causar la enfermedad.
4. En lo referente al abordaje molecular mediante secuenciación por Sanger de la región promotora del gen *BMP15*, se identificaron, tanto en casos como en controles, dos variantes de secuencia: c.-673T>C y c.-9C>G previamente reportadas en las bases de datos de SNP. Las frecuencias alélicas y genotípicas para estas variantes no presentaron diferencias significativas entre los casos y los controles ( $P > 0.05$ ). En ambas poblaciones, las frecuencias genotípicas estuvieron en equilibrio de Hardy Weinberg ( $p > 0.05$ ). A través de Haploview se demostró que las dos variantes están en desequilibrio de ligamiento ( $D' = 0.9044$ ). Estos hallazgos representan el primer reporte de secuenciación del promotor de *BMP15* en pacientes FOP no sindrómica e idiopática. Las variantes identificadas (c.-9C>G y c.-673C>T) habían sido previamente reportadas en las bases de datos en frecuencias muy variables en diferentes poblaciones de individuos sanos. Algunos autores han reportado la variante c.-9C>G en pacientes FOP y en controles (Dixit et al, 2006; Tiotiu et al, 2010)
5. En lo concerniente a la secuenciación directa de la región codificante de *BMP15*, se identificaron 6 variantes. Entre ellas, la única que se encontró exclusivamente en casos fue la variante no sinónima, c.433T>C, p.Leu148Pro. Esta mutación había sido previamente reportada como causal del fenotipo FOP (Laissue et al, 2006; Rossetti et al, 2009; Tiotiu et al, 2010). Con las variantes descritas en la región promotora (n=2) y codificante (n=5) y utilizando los programas SNPStats y

Haploview, se identificaron 19 haplotipos diferentes en la población de estudio. Teniendo en cuenta hallazgos positivos previos de asociación haplotípica de variantes del gen *BMP15* con FOP (Dixit et al, 2006), realizamos este análisis en la población de estudio para identificar potenciales combinaciones alélicas relacionadas con el fenotipo FOP. Desde esta perspectiva, identificamos un haplotipo exclusivo de los casos (TGAAIAT). Sin embargo, no fue posible demostrar diferencias significativas entre casos y controles. Esto puede ser atribuido al insuficiente tamaño de la muestra. Se puede apreciar que uno de los SNP del haplotipo (cuya base no corresponde al WT) es la variante del promotor c.-9C>G. Este cambio podría estar implicado con la función ovárica, lo que explicaría la asociación sugerida con infertilidad, con FOP y como modulador de la respuesta a la rFSH (Dixit et al., 2006; González et al, 2008; Hanevik et al, 2011).

6. En cuanto a la validación funcional del polimorfismo c.-9C>G del promotor de *BMP15*, se identificó por primera vez el factor de transcripción PITX1 como regulador directo de este gen. Los ensayos de expresión por RT-PCR indicaron que *PITX1* y *BMP15* están coexpresados en los oocitos humanos adultos así como en los ovarios de ratones adultos. Los experimentos de luciferasa demostraron que PITX1 es capaz de activar las dos versiones alélicas del promotor de *BMP15* (-9G y c.-9C>G). La actividad transcripcional del promotor *BMP15-G*, por la transactivación de PITX1 fue 1.6 mayor en relación a la observada para el promotor *BMP15-C* ( $p=1.4 \times 10^{-4}$ ). Estos resultados demuestran que la variante *BMP15* -9G>C modifica la transcripción del gen *BMP15*, por una modificación de la transactivación de su promotor vía la fijación de PITX-1. Estos resultados son coherentes con los hallazgos haplotípicos descritos por Dixit et al. De esta forma concluimos que esta variante debe contribuir con el fenotipo FOP. Sin embargo, genes modificadores y causas ambientales deben intervenir en la aparición del fenotipo ya que mujeres control también presentan el cambio.
7. Teniendo en cuenta los hallazgos descritos en este trabajo doctoral y el de otros grupos de investigación, es posible indicar que las aproximaciones “monogénicas” y la secuenciación directa de genes han explicado la etiología molecular de la FOP en

muy pocos casos (Laissue, 2015). En este contexto, se puede asumir que las aproximaciones multigénicas (e.g. NGS) de análisis genómicos son pertinentes. Por lo tanto, de manera innovadora se realizó la aproximación NGS de un panel de genes específicos, para la identificación de nuevos mecanismos moleculares causales de la enfermedad. Desde esta perspectiva, en las 12 pacientes FOP, se encontraron 4 variantes con un potencial efecto. Los cambios no-sinónimos se identificaron en tres pacientes con amenorrea secundaria. Ninguna de las variantes había sido previamente reportada en las bases de datos de SNP y los alineamientos de las secuencias proteicas entre varias especies de vertebrados indicó una total conservación de los residuos implicados en el cambio. Este resultado determinó que el NGS de regiones codificantes de genes candidatos para FOP tiene una alta sensibilidad de detección de variantes posiblemente causales. Los experimentos identificaron variantes potencialmente patogénicas en 3 de los genes analizados en el microarreglo (*LHCGR*, *ADAMTS19* y *BMPR2*).

8. Los experimentos para *ADAMTS19*, permitieron aportar nuevo conocimiento en los aspectos moleculares de este gen y su producto proteico. Identificamos por primera vez la interacción proteína: proteína entre COL6A2 y ADAMTS19.
9. En cuanto a la consecuencia de la mutación en *ADAMTS19* descrita en la paciente FOP analizada, concluimos que no afecta la interacción con COL6A2. Sin embargo, es probable que esté relacionada con la modificación de la actividad enzimática u otro proceso molecular.
10. El presente trabajo doctoral abordó el análisis genético y la validación funcional de genes relacionados con una patología compleja frecuente en la población, la falla ovárica prematura (FOP). De esta manera, este estudio aporta al nuevo conocimiento en los aspectos moleculares de la FOP y así integra a la práctica médica nuevas herramientas de diagnóstico innovadoras.
11. Con el desarrollo del presente trabajo, se logró la identificación de variantes en genes potencialmente causales de falla ovárica prematura (FOP) idiopática a través de las metodologías de gen candidato (evaluado mediante secuenciación por Sanger) y NGS (evaluado mediante microarreglos de genes específicos).

12. Posponemos nuevos marcadores moleculares de la patología, lo que genera una herramienta útil en el diagnóstico.

## 7. PERSPECTIVAS

El desarrollo de este trabajo doctoral permitió la realización de la primera aproximación NGS en una patología compleja y frecuente asociada a la reproducción. Los resultados obtenidos demostraron una alta eficiencia en la identificación de nuevas variantes y nuevos genes potencialmente causales de FOP. Por consiguiente, proponemos la aplicación de esta metodología para el análisis de un mayor número de pacientes. En este contexto, está en curso un proyecto financiado por Colciencias (bajo la dirección del Dr. Paul Laissue) que busca analizar por NGS (exoma) a 75 pacientes afectadas con FOP no-sindrómica. Los resultados preliminares muestran variantes nuevas potencialmente patogénicas, que una vez sean validadas en estudios funcionales, permitirán proponer nuevos marcadores moleculares asociados a la enfermedad. De esta forma, se generará un método innovador de diagnóstico molecular de enfermedades complejas que supere el alcance de las aproximaciones de gen candidato mediante secuenciación de Sanger. La descripción de nuevos marcadores permitirá ofrecer diagnósticos moleculares más exactos y realizar asesoramientos genéticos adecuados conducentes a la identificación de mujeres en riesgo. Por otro lado, las aproximaciones metodológicas desarrolladas en este trabajo doctoral, son la base para su aplicación en otras investigaciones. La identificación y validación funcional de nuevos *partners* mediante ensayos de doble híbrido en levaduras, coinmunoprecipitación e IFI indirecta están aplicándose en el estudio de nuevas proteínas asociadas a FOP. Antes de la realización de este trabajo doctoral no se conocía el sustrato de ADAMTS19. En este contexto, la validación funcional de la interacción entre ADAMTS19 y COL6A2, abre la posibilidad de generar nuevos proyectos que indaguen acerca de los mecanismos moleculares subyacentes a su interacción.

Por otro lado, en colaboración de los institutos Sanger (UK) e INRA (Instituto Nacional de Investigación Agronómica de Francia) hemos generado un modelo murino KO *Adamts19*<sup>-/-</sup>. A la fecha, se han estudiado algunos fenotipos relacionados con la reproducción. Este modelo servirá para estudios de la proteína Col6a2 en distintos estadios del desarrollo folicular. Experimentos sobre tejidos humanos (e.g. cortes de tejido ovárico) permitirán igualmente precisar el perfil de expresión espacio-temporal de ADAMTS19 y COL6A2.

## 8. REFERENCIAS

- Apte, S.S. (2009). A disintegrin-like and metalloprotease (reprolysin-type) with thrombospondin type 1 motif (ADAMTS) superfamily: functions and mechanisms. *Journal of Biological Chemistry*, 284(46):31493-7.
- Barrett, A. J., Rawlings, N. D., O'Brien, E. A (2001). The MEROPS database as a protease information system. *Journal of Structural Biology*, 134(2-3), 95–102.
- Beck-Peccoz P1., Persani L. (2006). Premature ovarian failure. *Orphanet Journal of Rare Disorders*, 6;1:9.
- Biesecker, L. G. (2010). Exome sequencing makes medical genomics a reality. *Nature Genetics*, 42(1), 13–4.
- Boivin, J., Bunting, L., Collins, J. A., Nygren, K. G. (2007). International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. *Human Reproduction*, 22(6), 1506–12.
- Bork, P., Beckmann, G. (1993). The CUB domain. A widespread module in developmentally regulated proteins. *Journal of Molecular Biology*, 231(2), 539–45.
- Caburet S, Zavadakova P, Ben-Neriah Z, Bouhali K, Dipietromaria A, Charon C, Besse C, Laissue P, Chalifa-Caspi V, Christin-Maitre S, Vaiman D, Levi G, Veitia RA, Fellous M. (2012) Genome-wide linkage in a highly consanguineous pedigree reveals two novel loci on chromosome 7 for non-syndromic familial Premature Ovarian Failure. *PLoS One*;7(3):e33412
- Calvo, S. E., Tucker, E. J., Compton, A. G., Kirby, D. M., Crawford, G., Burt, N. P., Mootha, V. K. (2010). High-throughput, pooled sequencing identifies mutations in NUBPL and FOXRED1 in human complex I deficiency. *Nature Genetics*, 42(10), 851–8.
- Casals, F., Idaghdour, Y., Hussin, J., Awadalla, P. (2012). Next-generation sequencing approaches for genetic mapping of complex diseases. *Journal of Neuroimmunology*, 248(1-2), 10–22.
- Cirulli, E. T., Goldstein, D. B. (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nature Reviews. Genetics*, 11(6), 415–25.

- Coulam, C. B., Adamson, S. C., Annegers, J. F. (1986). Incidence of premature ovarian failure. *Obstetrics and Gynecology*, 67(4), 604–6.
- Dagoneau, N., Benoist-Lasselín, C., Huber, C., Faivre, L., Mégarbané, A., Alswaid, A., Cormier-Daire, V. (2004). ADAMTS10 mutations in autosomal recessive Weill-Marchesani syndrome. *American Journal of Human Genetics*, 75(5), 801–6.
- Davidson, R. K., Waters, J. G., Kevorkian, L., Darrah, C., Cooper, A., Donell, S. T., Clark, I. M. (2006). Expression profiling of metalloproteinases and their inhibitors in synovium and cartilage. *Arthritis Research Therapy*, 8(4), R124.
- De Baere E, Copelli S, Caburet S, Laissue P, Beysen D, Christin-Maitre S, Bouchard P, Veitia R, Fellous M. (2005) Premature ovarian failure and forkhead transcription factor FOXL2: blepharophimosis-ptosis-epicanthus inversus syndrome and ovarian dysfunction. *Pediatr Endocrinol Rev.* 2(4):653-60.
- Demircan. K., Cömertoğlu, İ., Akyol, S., Yiğitoğlu, B.N., Sarıkaya, E. (2014). A new biological marker candidate in female reproductive system diseases: Matrix metalloproteinase with thrombospondin motifs (ADAMTS). *Journal of the Turkish German Gynecology Association*, 15(4):250-5
- Di Pasquale, E., Rossetti, R., Marozzi, A., Bodega, B., Borgato, S., Cavallo, L., Persani, L. (2006). Identification of new variants of human BMP15 gene in a large cohort of women with premature ovarian failure. *The Journal of Clinical Endocrinology and Metabolism*, 91(5), 1976–9.
- Diggle, C. P., Parry, D. A., Logan, C. V., Laissue, P., Rivera, C., Restrepo, C. M., Bonthron, D. T. (2012). Prostaglandin transporter mutations cause pachydermoperiostosis with myelofibrosis. *Human Mutation*, 33(8), 1175–81.
- Dixit, H., Rao, L. K., Padmalatha, V. V., Kanakavalli, M., Deenadayal, M., Gupta, N., Singh, L. (2006). Missense mutations in the BMP15 gene are associated with ovarian Hum Mutat.;34(11):1583-7.
- Ellard, S.1., Patrinos, G.P., Oetting, W.S. (2014). Clinical applications of next-generation sequencing: the 2013 human genome variation society scientific meeting. *Plos One*.
- Fonseca DJ, Patiño LC, Suárez YC, de Jesús Rodríguez A, Mateus HE, Jiménez KM, Ortega-Recalde O, Díaz-Yamal I, Laissue P. (2015) Next generation sequencing in women affected by nonsyndromic premature ovarian failure displays new potential causative genes and mutations. *Fertil Steril.* 104(1):154-62.

- Fonseca DJ, Ortega-Recalde O, Esteban-Perez C, Moreno-Ortiz H, Patiño LC, Bermúdez OM, Ortiz AM, Restrepo CM, Lucena E, Laissue P. (2014) BMP15 c.-9C>G promoter sequence variant may contribute to the cause of non-syndromic premature ovarian failure. *Reprod Biomed Online*. 29(5):627-33.
- Fonseca DJ, Ojeda D, Lakhali B, Braham R, Eggers S, Turbitt E, White S, Grover S, Warne G, Zacharin M, Nevin Lam A, Landolsi H, Elghezal H, Saâd A, Restrepo CM, Fellous M, Sinclair A, Koopman P, Laissue P. (2012) CITED2 mutations potentially cause idiopathic premature ovarian failure. *Transl Res*; 160(5):384-8.
- Fonseca DJ, Garzón E, Lakhali B, Braham R, Ojeda D, Elghezal H, Saâd A, Restrepo CM, Laissue P. (2012) Screening for mutations of the FOXO4 gene in premature ovarian failure patients. *Reprod Biomed Online*. 24(3):339-41
- Fernandes, R. J., Hirohata, S., Engle, J. M., Colige, A., Cohn, D. H., Eyre, D. R., Apte, S. S. (2001). Procollagen II amino propeptide processing by ADAMTS-3. Insights on dermatosparaxis. *The Journal of Biological Chemistry*, 276(34), 31502–9.
- Fields, S., Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature*, 340(6230):245-6.
- Galloway, S. M., McNatty, K. P., Cambridge, L. M., Laitinen, M. P., Juengel, J. L., Jokiranta, T. S., Ritvos, O. (2000). Mutations in an oocyte-derived growth factor gene (BMP15) cause increased ovulation rate and infertility in a dosage-sensitive manner. *Nature Genetics*, 25(3), 279–83.
- Frebourg T (2014). The challenge for the next generation of medical geneticists. *Human Mutation*, 35(8):909-11.
- Gao, S., De Geyter, C., Kossowska, K., Zhang, H. (2007). FSH stimulates the expression of the ADAMTS-16 protease in mature human ovarian follicles. *Molecular Human Reproduction*, 13(7), 465–71
- González, A. I., Ramírez-Lorca R, Calatayud, C., Mendoza, N., Ruiz, A., Sáez, M. E., Morón, F. J. (2008). Association of genetic markers within the BMP15 gene with anovulation and infertility in women with polycystic ovary syndrome. *Fertility and Sterility*, 90(2):447-9.
- Hanevik, H. I., Hilmarsen, H. T., Skjelbred, C. F., Tanbo, T., Kahn, J. A. (2011). A single nucleotide polymorphism in BMP15 is associated with high response to ovarian stimulation. *Reproductive Biomedicine Online*, 23(1), 97–104.

- Huang, T. F. (1998). What have snakes taught us about integrins? *Cellular and Molecular Life Sciences*, 54(6), 527–40.
- Inhorn, M.C., Patrizio, P. (2015). Infertility around the globe: new thinking on gender, reproductive technologies and global movements in the 21st century. *Human Reproduction Update*, 21(4):411-26.
- Jones, G. C., Riley, G. P. (2005). ADAMTS proteinases: a multi-domain, multi-functional family with roles in extracellular matrix turnover and arthritis. *Arthritis Research Therapy*, 7(4), 160–9.
- Kelwick, R., Desanlis, I., Wheeler, G.N., Edwards, D.R. (2015). The ADAMTS (A Disintegrin and Metalloproteinase with Thrombospondin motifs) family. *Genome Biol*, 16:113.
- Kilpinen, H., Barrett, J. C. (2013). How next-generation sequencing is transforming complex disease genetics. *Trends in Genetics* : 29(1), 23–30.
- Knauff, E. a H., Franke, L., van Es, M. a, van den Berg, L. H., van der Schouw, Y. T., Laven, J. S. E., ... Fauser, B. C. J. M. (2009). Genome-wide association study in premature ovarian failure patients suggests ADAMTS19 as a possible candidate gene. *Human Reproduction*, 24(9), 2372–8.
- Kryukov, G. V, Pennacchio, L. A., Sunyaev, S. R. (2007). Most rare missense alleles are deleterious in humans: implications for complex disease and association studies. *American Journal of Human Genetics*, 80(4), 727–39.
- Kuno ,K., Kanada, N., Nakashima, E., Fujiki, F., Ichimura, F., Matsushima, K.(1997). Molecular cloning of a gene encoding a new type of metalloproteinase-disintegrin family protein with thrombospondin motifs as an inflammation associated gene. *Journal of Biological Chemistry*, 272(1):556-62.
- Lakhal B, Laissue P, Braham R, Elghezal H, Saâd A, Fellous M, Veitia RA. (2010) BMP15 and premature ovarian failure: causal mutations, variants, polymorphisms? *Clin Endocrinol (Oxf)*. 72(3):425-6.
- Lakhal B, Laissue P, Braham R, Elghezal H, Saâd A, Fellous M, Veitia RA. (2009) A novel BMP15 variant, potentially affecting the signal peptide, in a familial case of premature ovarian failure. *Clin Endocrinol (Oxf)*. 71(5):752-3.

- Lakhal B, Laissue P, Elghèzal H, Fellous M. (2008) Genetic analysis of premature ovarian failure: role of forkhead and TGF-beta genes]. *Gynecol Obstet Fertil*. Sep;36(9):862-71.
- Laissue, P., Christin-Maitre, S., Touraine, P., Kuttenn, F., Ritvos, O., Aittomaki, K., ... Veitia, R. A. (2006). Mutations and sequence variants in GDF9 and BMP15 in patients with premature ovarian failure. *European Journal of Endocrinology*, 154(5), 739–44.
- Laissue, P., Vinci, G., Veitia, R. a, Fellous, M. (2008). Recent advances in the study of genes involved in non-syndromic premature ovarian failure. *Molecular and Cellular Endocrinology*, 282(1-2), 101–11.
- Laissue P.( 2015). Aetiological coding sequence variants in non-syndromic premature ovarian failure: From genetic linkage analysis to next generation sequencing. *Molecular and Cellular Endocrinology*, 411:243-57.
- Laissue P, Vinci G, Veitia RA, Fellous M. (2008) Recent advances in the study of genes involved in non-syndromic premature ovarian failure. *Mol Cell Endocrinol*.30;282(1-2):101-11.
- Laissue P, Christin-Maitre S, Bouchard P, Fellous M, Veitia RA. (2007) Mutations in the NOG gene are not a common cause of nonsyndromic premature ovarian failure. *Clin Endocrinol (Oxf)*.66(6):900.
- : Laissue P, Christin-Maitre S, Touraine P, Kuttenn F, Ritvos O, Aittomaki K, Bourcigaux N, Jacquesson L, Bouchard P, Frydman R, Dewailly D, Reyss AC, Jeffery L, Bachelot A, Massin N, Fellous M, Veitia RA. (2006) Mutations and sequence variants in GDF9 and BMP15 in patients with premature ovarian failure. *Eur J Endocrinol*. 154(5):739-44.
- Li, S.W., Arita, M., Fertala, A., Bao, Y., Kopen, G.C., Långsjö, T.K., Hyttinen, M.M., Helminen, H.J., Prockop, D.J.( 2001). Transgenic mice with inactive alleles for procollagen N-proteinase (ADAMTS-2) develop fragile skin and male sterility. *Biochemical Journal*, 355(Pt 2):271-8.
- Maclaran, K., Horner, E., Panay, N. (2010). Premature ovarian failure: long-term sequelae. *Menopause International*, 16(1), 38–41.
- Maclaran, K., Panay, N. (2011). Premature ovarian failure. *The Journal of Family Planning and Reproductive Health Care*, 37(1), 35–42.

- Malfait, A.-M., Arner, E. C., Song, R.-H., Alston, J. T., Markosyan, S., Staten, N., ... Tortorella, M. D. (2008). Proprotein convertase activation of aggrecanases in cartilage in situ. *Archives of Biochemistry and Biophysics*, 478(1), 43–51.
- Matzuk, M. M., Lamb, D. J. (2008). The biology of infertility: research advances and clinical challenges. *Nature Medicine*, 14(11), 1197–213.
- Menke, D. B., Page, D. C. (2002). Sexually dimorphic gene expression in the developing mouse gonad. *Gene Expression Patterns : GEP*, 2(3-4), 359–67.
- Mittaz, L., Russell, D. L., Wilson, T., Brasted, M., Tkalcevic, J., Salamonsen, L. A., Pritchard, M. A. (2004). Adamts-1 is essential for the development and function of the urogenital system. *Biology of Reproduction*, 70(4), 1096–105.
- Nejentsev, S., Walker, N., Riches, D., Egholm, M., Todd, J. A. (2009). Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science*, 324(5925), 387–9.
- Ojeda D, Lakhal B, Fonseca DJ, Braham R, Landolsi H, Mateus HE, Restrepo CM, Elghezal H, Saâd A, Laissue P. (2011) Sequence analysis of the CDKN1B gene in patients with premature ovarian failure reveals a novel mutation potentially related to the phenotype. *Fertil Steril*. Jun 30;95(8):2658-60.
- Lakhal B, Laissue P, Braham R, Elghezal H, Saâd A, Fellous M, Veitia RA. (2010) BMP15 and premature ovarian failure: causal mutations, variants, polymorphisms? *Clin Endocrinol (Oxf)*. 72(3):425-6.
- Okou, D. T., Steinberg, K. M., Middle, C., Cutler, D. J., Albert, T. J., Zwick, M. E. (2007). Microarray-based genomic selection for high-throughput resequencing. *Nature Methods*, 4(11), 907–9.
- Ortega-Recalde, O., Vergara, J. I., Fonseca, D. J., Ríos, X., Mosquera, H., Bermúdez, O. M., Laissue, P. (2013). Whole-exome sequencing enables rapid determination of xeroderma pigmentosum molecular etiology. *PloS One*, 8(6), e64692.
- Ortega-Recalde, O., Beltrán, O., Gálvez, J.M., Palma-Montero, A., Restrepo, C.M., Mateus, H.E., Laissue, P. (2015) Biallelic HERC1 mutations in a syndromic form of overgrowth and intellectual disability. *Clinical Genetics*, 88(4):e1-3

- Patiño, L.C., Battu, R., Ortega-Recalde, O., Nallathambi, J., Anandula, V.R., Renukaradhya, U., Laissue, P. (2014). Exome sequencing is an efficient tool for variant late-infantile neuronal ceroid lipofuscinosis molecular diagnosis. *PLoS One*, 9(10):e109576.
- Patiño LC., Silgado D., Laissue P.(2017). A potential functional association between mutant BMPR2 and primary ovarian insufficiency. *Syst Biol Reprod Med*, 17,1-5.
- Liliana C. Patiño., Kelly L. Walton ., Thomas D. Mueller., Katharine E. Johnson., William Stocker., Dulama Richani., David Agapiou., Robert B. Gilchrist., Paul Laissue., Craig A. Harrison. (2016) BMP15 Mutations Associated With Primary Ovarian Insufficiency Reduce Expression, Activity, or Synergy With GDF9. *J Clin Endocrinol Metab*, 102 (3), 1009-1019.
- Pinto, D., Pagnamenta, A. T., Klei, L., Anney, R., Merico, D., Regan, R., Betancur, C. (2010). Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*, 466(7304), 368–72.
- Porter, S., Scott, S. D., Sassoon, E. M., Williams, M. R., Jones, J. L., Girling, A. C., Edwards, D. R. (2004). Dysregulated expression of adamalysin-thrombospondin genes in human breast carcinoma. *Clinical Cancer Research*, 10(7), 2429–40.
- Pyun, J.A., Kim, S., Cha, D.H., Kwack, K.( 2013). Epistasis between IGF2R and ADAMTS19 polymorphisms associates with premature ovarian failure. *Human Reproduction*, 28(11):3146-54.
- Pyun, J.A., Kim, S., Kwack, K.(2015). Epistasis between polymorphisms in ACVR2B and ADAMTS19 is associated with premature ovarian failure. *Menopause*, 2(2):212-6.
- Rao, C., Foerzler, D., Loftus, S. K., Liu, S., McPherson, J. D., Jungers, K. A., Beier, D. R. (2003). A defect in a novel ADAMTS family member is the cause of the belted white-spotting mutation. *Development*, 130(19), 4665–72.
- Richards, J. S., Hernandez-Gonzalez, I., Gonzalez-Robayna, I., Teuling, E., Lo, Y., Boerboom, D., Sandy, J. D. (2005). Regulated expression of ADAMTS family members in follicles and cumulus oocyte complexes: evidence for specific and redundant patterns during ovulation. *Biology of Reproduction*, 72(5), 1241–55.

- Rodgers, R.J., Irving-Rodgers, H.F., Russell, D.L.(2003).Extracellular matrix of the developing ovarian follicle. *Reproduction*, 126(4):415-24.
- Rodgers, R.J., Irving-Rodgers, H.F.(2010).Formation of the ovarian follicular antrum and follicular fluid. *Biol Reprod*, 82(6):1021-9.
- Rossetti, R., Di Pasquale, E., Marozzi, A., Bione, S., Toniolo, D., Grammatico, P., Persani, L. (2009). BMP15 mutations associated with primary ovarian insufficiency cause a defective production of bioactive protein. *Human Mutation*, 30(5), 804–10.
- Russell, D.L., Brown, H.M., Dunning, K.R. (2015).ADAMTS proteases in fertility. *Matrix Biology*, 44-46:54-63.
- Scheiflinger, F., Knöbl, P., Trattner, B., Plaimauer, B., Mohr, G., Dockal, M.,Rieger, M. (2003). Nonneutralizing IgM and IgG antibodies to von Willebrand factor-cleaving protease (ADAMTS-13) in a patient with thrombotic thrombocytopenic purpura. *Blood*, 102(9), 3241–3.
- Seidah, N. G., Prat, A. (2002). Precursor convertases in the secretory pathway, cytosol and extracellular milieu. *Essays in Biochemistry*, 38, 79–94.
- Shindo, T., Kurihara, H., Kuno, K., Yokoyama, H., Wada, T., Kurihara, Y.,Matsushima, K. (2000). ADAMTS-1: a metalloproteinase-disintegrin essential for normal growth, fertility, and organ morphology and function. *The Journal of Clinical Investigation*, 105(10), 1345–52.
- Shozu, M., Minami, N., Yokoyama, H., Inoue, M., Kurihara, H., Matsushima, K., Kuno, K. (2005). ADAMTS-1 is involved in normal follicular development, ovulatory process and organization of the medullary vascular network in the ovary. *Journal of Molecular Endocrinology*, 35(2), 343–55.
- Somerville, R. P. T., Longpré, J.-M., Apel, E. D., Lewis, R. M., Wang, L. W., Sanes, J. R., Apte, S. S. (2004). ADAMTS7B, the full-length product of the ADAMTS7 gene, is a chondroitin sulfate proteoglycan containing a mucin domain. *The Journal of Biological Chemistry*, 279(34), 35159–75.
- Somerville, R. P. T., Longpre, J.-M., Jungers, K. A., Engle, J. M., Ross, M., Evanko, S., Apte, S. S. (2003). Characterization of ADAMTS-9 and ADAMTS-20 as a distinct ADAMTS subfamily related to *Caenorhabditis elegans* GON-1. *The Journal of Biological Chemistry*, 278(11), 9503–13.

- Sperling, S., Grimm, C. H., Dunkel, I., Mebus, S., Sperling, H.-P., Ebner, A., Hammer, S. (2005). Identification and functional analysis of CITED2 mutations in patients with congenital heart defects. *Human Mutation*, 26(6), 575–82.
- Tiotiu, D., Alvaro Mercadal, B., Imbert, R., Verbist, J., Demeestere, I., De Leener, a, Delbaere, a. (2010b). Variants of the BMP15 gene in a cohort of patients with premature ovarian failure. *Human Reproduction*, 25(6), 1581–7.
- Toniolo, D. (2006). X-linked premature ovarian failure: a complex disease. *Current Opinion in Genetics Development*, 16(3), 293–300.
- Tortorella, M. D., Malfait, F., Barve, R. A., Shieh, H.-S., Malfait, A.-M. (2009). A review of the ADAMTS family, pharmaceutical targets of the future. *Current Pharmaceutical Design*, 15(20), 2359–74.
- Wang, B., Wen, Q., Ni, F., Zhou, S., Wang, J., Cao, Y., Ma, X. (2010). Analyses of growth differentiation factor 9 (GDF9) and bone morphogenetic protein 15 (BMP15) mutation in Chinese women with premature ovarian failure. *Clinical Endocrinology*, 72(1), 135–6.
- Zheng, X., Chung, D., Takayama, T. K., Majerus, E. M., Sadler, J. E., Fujikawa, K. (2001). Structure of von Willebrand factor-cleaving protease (ADAMTS13), a metalloprotease involved in thrombotic thrombocytopenic purpura. *The Journal of Biological Chemistry*, 276(44), 41059–63.

