



Análisis preliminar de la topología del genoma en la diferenciación de células endometriales humanas

Tesis de Licenciatura en Ciencias Biológicas

Luciana Ant

Directora:

Patricia Saragüeta

Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales
Laboratorio de Regulación Hormonal del Tracto Reproductivo Femenino, Instituto de
Biología y Medicina Experimental, IBYME-CONICET, Buenos Aires, Argentina.

Contenidos

1 Glosario y abreviaturas	3
2 Resumen	4
3 Abstract	6
4 Introducción	8
4.1 El endometrio y el ciclo menstrual	8
4.2 El rol de E2 y Pg en la decidualización	10
4.3 Marcadores de la decidualización	12
4.4 Decidualización a nivel celular	13
4.5 Los estímulos decidualizantes: P4 y R5020	15
4.6 Modelos estadísticos lineales y su aplicación al análisis morfológico	15
4.7 Exploración de la topología de la cromatina a través de Hi-C	19
5 Objetivos e hipótesis	25
6 Materiales y Métodos	26
6.1 Cultivo celular y tratamiento	26
6.2 Muestreo y cuantificación morfológica	26
6.3 Análisis de la circularidad	30
6.4 Hi-C	41
6.5 RNAseq	41
6.6 Asignación de los signos de PC1	42

6.7	PCA	49
7	Resultados	49
7.1	Verificación de la decidualización	49
7.2	Análisis exploratorio de los cambios morfológicos	51
7.3	Selección de Modelos estadísticos	54
7.4	Efecto de los tratamientos sobre la circularidad celular	61
7.5	Análisis preliminar de la topología génica durante la decidualización	63
7.6	Relación entre los cambios morfológicos y de compartmentalización	68
8	Discusión y Conclusiones	70
9	Anexo	77
10	Bibliografía	79

1 Glosario y abreviaturas

8-Br-AMPc: derivado bromado del monofosfato de adenosinacíclico

AMPc: Adenosín monofosfato cíclico

CTCF: CCCTC-binding factor

DSC: Decidual Stromal Cell (célula estromal decidual)

E2: estradiol

FC: Fold Change (veces de cambio)

log2FC: Log en base 2 del FC

FDR: False discovery rate (p-valor ajustado)

FOXO1: Forkhead box protein O 1

Hi-C: High-throughput chromosome conformation capture

IGFBP1/IGF: Insulin-Like Growth Factor Binding Protein 1 (Proteína de unión al factor de crecimiento insulínico tipo 1)

MPA: 17-acetato de medroxiprogesterona

P4: progesterona

PC1: Principal Component 1

PCA: Principal Component Analysis

PRL: prolactina

R5020: Promegestrona

TAD: Topological Associated Domain

t-HESC: transformed Human Endometrial Stromal Cells (células estromales endometriales humanas transformadas)

2 Resumen

La decidualización es un proceso biológico por el que transitan las células del estroma endometrial de los mamíferos euterios durante la preñez. Este proceso consiste en la transdiferenciación fenotípica de las células del estroma, que pasan de un tipo fibroblástico a uno epitelioide y secretor, culminando en la formación de la decidua. Dicha diferenciación se ve reflejada en cambios de la morfología celular, de la estructura cromatínica y de la expresión génica.

La cromatina presenta una disposición en el espacio determinada, no azarosa. Esta disposición u organización permite acercar o alejar ciertas regiones del genoma entre sí, permitiendo por ejemplo el encuentro de promotores y enhancers específicos. La plasticidad de la estructura de la cromatina permite a la célula responder y adaptarse a factores tanto intrínsecos como extrínsecos, pudiendo variar su organización frente a estos eventos. En nuestro laboratorio hemos reportado que en células epiteliales de endometrio estas estructuras son previas al estímulo hormonal, y le permiten responder a los cambios en el ambiente de manera más rápida. No se conoce cual es la relación entre los cambios en la topología de la cromatina y la diferenciación de las células estromales endometriales. Dado que este sistema podría constituir un modelo de estudio de la reestructuración de la cromatina como respuesta fisiológica a un estimulo externo es que se propuso caracterizarlo.

Los objetivos de este trabajo fueron cuantificar la morfología celular durante la decidualización y realizar un análisis global preliminar de los cambios transitorios y/o estables que ocurren en la organización de la cromatina de las células endometriales frente a los progestágenos implicados en la decidualización.

Con este fin se utilizó un modelo de células humanas estromales endometriales (t-HESCs) a las cuales se les aplicaron dos protocolos de decidualización: estradiol (E2), AMPc, y progesterona (P4) o la progestina R5020 ,un progestágeno sintético no metabolizable. Luego de los tratamientos se analizaron las respuestas a tiempos cortos (1h) o tiempos

largos (3 y 6 días) de tratamiento para la morfología y la organización de la cromatina. En primera instancia se cuantificaron los cambios morfométricos en el tiempo como medida del grado de avance de la diferenciación, para lo cual se aplicó un modelo estadístico que permitió explicar la variación de la circularidad de las células en función del tiempo y del tratamiento. Se observó que las células tratadas con ambos estímulos aumentaron significativamente su circularidad respecto de las células tratadas con el vehículo luego de 2 días de tratamiento.

Aunque las células para ambos tratamientos decidualizantes presentaron igual circularidad en el tiempo final (6 días), se observó que las células tratadas con R5020 presentaron un aumento anticipado respecto de las células tratadas con P4.

En segunda instancia, y con el objetivo de analizar los cambios iniciales y en los dos estadíos de decidualización descriptos, se realizaron protocolos de Hi-C a los tiempos T0, 1h, 3 y 6 días para ambos tratamientos y para el vehículo. Esta técnica permite predecir cómo se organiza la cromatina a partir de la probabilidad de contacto entre los distintos *loci*, luego de la cual es posible identificar las regiones abiertas (A) y cerradas (B) de la cromatina mediante análisis de correlación y de componentes principales. Ambos tratamientos decidualizantes produjeron una reestructuración global de la distribución de compartimientos A y B a los 3 y 6 días de ambos tratamientos respecto del T0. Además, la distribución A/B para los días 3 y 6 se diferenciaron claramente entre sí, lo que indica la presencia de al menos dos estadíos de la organización de la cromatina durante la decidualización. Por otro lado, a 1h de ambos tratamientos no se observaron cambios globales de la cromatina.

Tanto la morfología celular como la organización de la cromatina del estadio alcanzado al día 3 de cultivo con ambos tratamiento decidualizantes se diferenciaron del T0 y del tiempo final (6 días). Esto indica la presencia de al menos dos estadios durante la decidualización, un estadio temprano a los 3 días y un estadio tardío a los 6 días, que implican una organización de la cromatina y una circularidad característica y específica de ese estadio. Estos resultados muestran que la decidualización *in vitro* es buen modelo para analizar

cómo se desarrollan los cambios en la organización de la cromatina durante un proceso de reprogramación celular fisiológico, en el cual los factores de transcripción como el receptor de progesterona (PGR) tienen un papel fundamental.

3 Abstract

Genome topology during the trans-differentiation of human endometrial cells

Decidualization is a biological process that mammal endometrial stromal cells undergo during pregnancy. This process comprises a phenotypic transdifferentiation of the stromal cells, which transition from a fibroblastic phenotype to an epithelial and secretory type, culminating in the formation of the decidua. This differentiation impacts cellular morphology, chromatin structure and gene expression.

The spatial layout of the chromatin is non random and highly determined. This organization or layout can bring closer or insulate certain regions of the genome, allowing for example the encounter of specific enhancers and their promoters. The plasticity of the chromatin structure enables the cell to respond and adapt to external and internal factors. We have reported in previous studies that the chromatin structure in epithelial endometrial cells is already formed prior to hormone stimuli, allowing a quicker response. The relationship between chromatin topology and the differentiation of endometrial stromal cells is still unknown. Given that this biological system could constitute a suitable model to study the restructuring of the chromatin as a physiological response to an external stimuli, we aim to characterize it.

The goal of this work was to quantify the changes in cellular morphology that unfold during decidualization and to analyze the transient and stable changes that take place in the chromatin organization of the endometrial cells when treated with progestins involved in decidualization. To this end we employed a model using human endometrial stromal cells

(tHESCs) under two different decidualization treatments: estradiol (E2), cyclic AMP (cAMP) and either progesterone (P4) or the R5020, a non-metabolizable synthetic progestin. We analyzed the response after 1 hour and 3 and 6 days after treatment. On one hand we quantified the morphometric changes throughout the days in order to measure the progression of the decidualization process. We applied a statistical model that allowed us to explain the variation in the circularity of the cells depending on the time and treatment. Both decidualizing treatments produced an increment in cell circularity compared to the control treatment, only two days after induction. Although the cells for both treatments displayed the same level of circularity at final time (6 days after treatment), cells treated with R5020 presented an early increase of their circularity when compared to P4 treated cells.

On the other hand, in order to analyze the changes in the chromatin topology, we performed Hi-C protocols at different times (at time zero, 1 hour and 3 and 6 days) after treatment with the two decidualizing treatments and the control treatment. The analysis via Hi-C is able to predict the chromatin organization calculating the probability of interaction between any two given *loci*, from which is possible to identify open (A) and close (B) regions of the chromatin through correlation and principal component analysis. Both decidualizing treatments led to a global reorganization of A and B compartments, at 3 and 6 days. This restructuring was not found using control treatment.

For both cell morphology and chromatin topology, the state reached at day 3 are clearly distinguishable from the state at day 6 and at time zero. This indicates the presence of at least two clearly discernible states, an early state at day 3 and a late state at day 6, each one with a specific circularity and chromatin state profile.

This findings show how *in vitro* decidualization can serve as a model to analyze how the changes in chromatin organization unfold during a process of cellular reprogramming, in which transcription factors such as the progesterone receptor (PGR) play a crucial role.

4 Introducción

4.1 El endometrio y el ciclo menstrual

El endometrio es la mucosa que conforma la parte interior del útero. Está compuesto por un epitelio columnar simple, estroma de tejido conectivo y glándulas endometriales.

Experimenta etapas de proliferación, diferenciación y remodelamiento de manera cíclica modificando su composición celular, lo que permite la implantación en la fase de receptividad y el mantenimiento del embarazo una vez que el embrión ya se ha implantado. Estos cambios están dirigidos por las hormonas sexuales ováricas, progesterona (Pg) y estradiol (E2) (Fig.1).

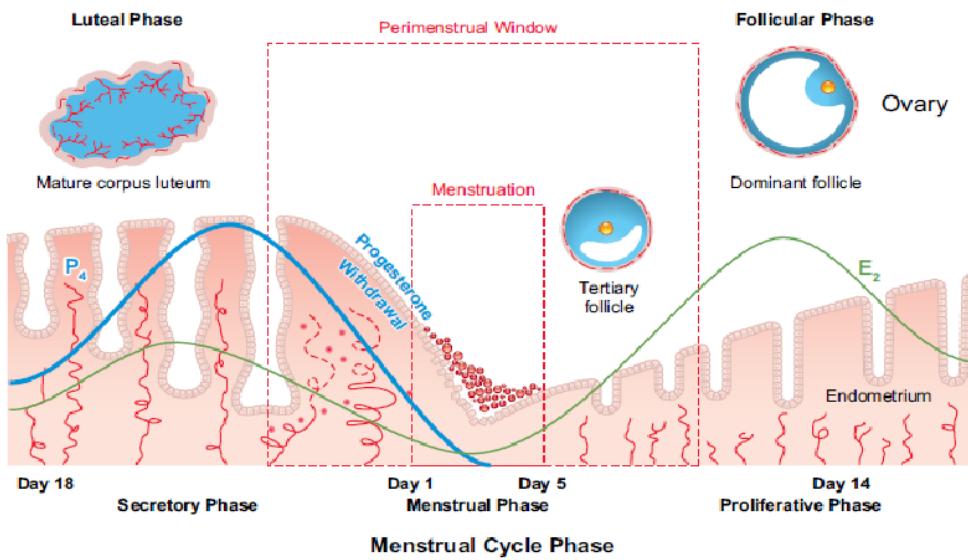


Figura 1: Etapas del ciclo menstrual en ovarios y endometrio (Critchley et al. 2020)

La fase proliferativa del endometrio ocurre en la primera mitad del ciclo, cuando los niveles de estrógeno se encuentran altos. Esta proliferación cesa a medida que disminuyen los niveles de estrógeno y aumentan los niveles de progesterona, ingresando así a la fase secretora. Es durante este pasaje que las células endometriales comienzan el proceso de decidualización, en respuesta al aumento en los niveles de progesterona.

La fase proliferativa se destaca por su estado pro-inflamatorio, un aumento en la angiogénesis y un gran influjo de células inmunitarias al espacio endometrial. En particular, las células inmunes Natural Killer (uNK) aumentan dramáticamente, llegando a componer hasta el 30-40% de las células en el compartimiento estromal (Gellersen & Brosens 2003). La comunicación con el sistema inmune durante el ciclo menstrual es crucial ya que en caso de que ocurra la implantación, será el encargado de otorgar la inmunotolerancia materna frente al feto antigénicamente distinto. Si ocurre la implantación, los niveles de progesterona continuarán aumentando y las células deciduales proseguirán el proceso de decidualización, culminando en la formación de la decidua. Puede así proporcionar nutrición, un micro ambiente inmunotolerante, un proceso de selección negativa para embriones no viables y protección contra factores estresantes fisiológicos así como contra la invasión excesiva del embrión en el compartimento uterino (Gellersen & Brosens 2014).

Por otra parte, si en cambio la implantación no ocurre, los niveles de progesterona caen y la pared externa del endometrio se desprende, dando lugar a la menstruación. Luego comienza nuevamente una fase proliferativa.

La decidualización constituye un paso esencial para el establecimiento y mantenimiento de la preñez en muchos mamíferos euterios. Su origen se encuentra en la rama basal de este grupo (Kin et al. 2014, Mess & Carter 2006), siendo la base de la serie de características que permiten la preñez prolongada en los euterios, incluyendo la implantación directa del blastocisto y trofoblasto en el endometrio materno, el reconocimiento materno exacerbado del embarazo, la comunicación materna-embrialaria, y la inmunotolerancia materna del embrión antigénicamente distinto (Lynch et al. 2015).

En la mayoría de los mamíferos la decidualización ocurre como respuesta a la presencia de un embrión. Se puede pensar, desde un punto de vista evolutivo, que la adquisición del control maternal de la decidualización debió haber sido esencial para el *fitness* reproductivo, de manera de justificar el gran gasto energético que conlleva el ciclo menstrual.

4.2 El rol de E2 y Pg en la decidualización

Ya sea durante la fase lútea (o secretora) del ciclo menstrual o con el inicio del embarazo, la mayoría de las respuestas fisiológicas a la progesterona derivada del ovario están mediadas por su receptor intracelular específico, el PGR. La caracterización estructural del PGR demostró que el receptor pertenece a la superfamilia de receptores nucleares de factores de transcripción, que incluye receptores afines para otras hormonas esteroides (incluido el estrógeno), así como para retinoides, prostanoïdes, hormona tiroidea y vitamina D3.

El PGR comprende dos isoformas del receptor: el PGR-B de longitud completa (~ 116 kDa) y el PGR-A truncado (~ 94 kDa), que carece de los primeros 164 residuos de aminoácidos (Fig.2). Se cree que la ausencia del dominio de activación 3 (AF3) en la isoforma PGR-A explica, en parte, las diferentes propiedades transactivacionales *in vitro* reportadas para estas dos isoformas de receptor. Las isoformas de PGR coexpresan con frecuencia en la mayoría de los tejidos blanco y están codificadas por un solo gen que utiliza un promotor distal PGR-B y un promotor proximal PGR-A en tandem en la expresión de las respectivas isoformas del receptor. Dada la coexistencia de ambas isoformas, se considera que la respuesta fisiológica de un tejido diana a la progesterona depende de la estequiometría neta de las isoformas PGR-B y PGR-A dentro de una población celular, en la que la formación de homodímeros y/o heterodímeros proyectan distintos controles transactivacionales sobre la expresión del conjuntos de genes asociados.

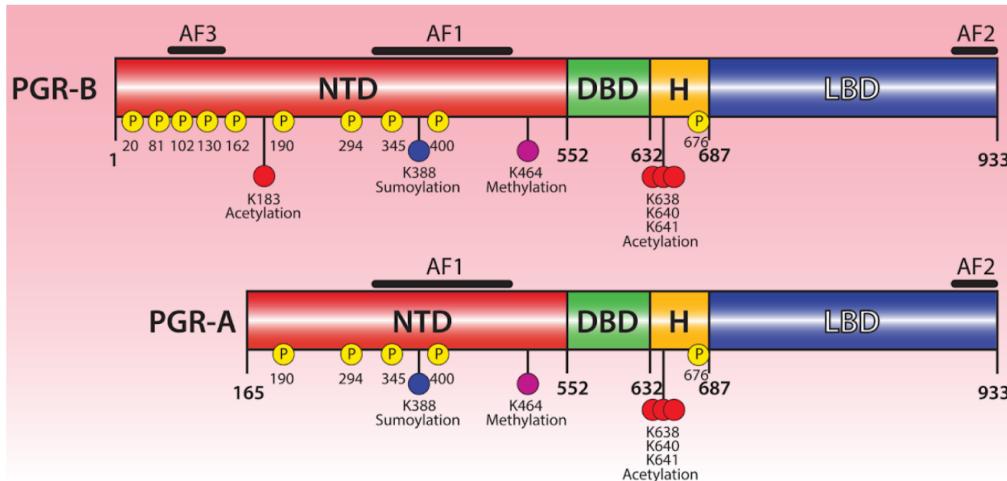


Figura 2: Organización funcional de dominios de las dos isoformas del receptor de progesterona humana. Se indican el dominio N-terminal (NTD), el dominio de unión al DNA (DBD), la región Hinge (H), y el dominio de unión al ligando (LBD). También se indican las posiciones de activación funcional (AF1, AF2, AF3). Los sitios de fosforilación se denotan con P. También se resaltan los sitios de sumoilación, acetilación y metilación (DeMayo & Lydon 2020, Grimm et al. 2016)

Lo que conocemos como el mecanismo clásico de la acción del PGR implica que la progesterona atraviese la bicapa lipídica de la membrana plasmática de la célula blanco, interactuando con el dominio de unión a ligando (LBD) y desencadenando un cambio conformacional del receptor, lo que da como resultado que el receptor unido a ligando se dirija al núcleo y se una directamente como un dímero a un elemento canónico de respuesta a progesterona (la secuencia PRE) dentro del promotor o de enhancers distantes de un gen blanco. Como parte de este proceso, los correguladores (es decir, coactivadores o correpresores) y cofactores asociados se reclutan en el complejo transcripcional para activar o silenciar la expresión génica, el resultado neto se manifiesta en la respuesta fisiológica a progesterona a nivel molecular. En muchos contextos fisiológicos, la expresión de PGR es inducida por el estrógeno a través de su receptor nuclear, el receptor de estrógeno- α (ESR1) (Grimm et al. 2016).

Los estudios realizados en líneas celulares de cáncer de mama humano revelaron que los niveles de expresión, la actividad transcripcional, el tránsito del PGR y la selección de genes blanco están modulados por modificaciones postraduccionales (PTM) (Hagan & Lange 2014), que incluyen fosforilación, ubiquitinación, sumoilación, acetilación y

metilación. El PGR también puede modular indirectamente la expresión génica a través de un modo de acción no clásico que prescinde de la necesidad de una unión directa al DNA. En este caso, el PGR se une a otros factores de transcripción unidos al DNA, como la proteína activadora 1 (AP1), SP1, NFkB y el transductor de señal activador de la transcripción 3 (Stat3) (Grimm et al. 2016). Además también se ha demostrado que el PGR se une a las quinasas activadas por los receptores del factor de crecimiento en la membrana plasmática, que desencadenan una activación rápida de las cascadas de señalización de las quinasas que con frecuencia fosforilan una miríada de factores de transcripción (incluido el mismo PGR) para influir en las fases posteriores de los programas de expresión genética (Vallejo et al. 2005, Hagan & Lange 2014, Grimm et al. 2016).

4.3 Marcadores de la decidualización

El gen de la prolactina (PRL) y el gen que codifica la proteína de unión al factor de crecimiento insulínico tipo 1 (IGFBP-1) constituyen dos de los principales genes marcadores de la decidualización y se usan muy frecuentemente para reconocer los distintos estados de la diferenciación.

PRL se expresa en células deciduales antes y durante el embarazo (Maslar & Riddick 1979) por la inducción con progesterona (Daly et al. 1983). Sus receptores (PRL-R), en endometrio donde no hay embarazo, se localizan en células epiteliales y tienen su mayor expresión en etapas más tardías del proceso de decidualización (Jabbour & Critchley 2001). Se le han adjudicado varias funciones a la PRL decidual. Se demostró que la PRL decidual, en ratas, inhibe la expresión de una enzima que cataboliza la progesterona, la 20α -hydroxysteroid dehydrogenase (20α -HSD); y, de la citoquina proinflamatoria IL-6, dos potenciales factores perjudiciales para el embarazo (Bao et al. 2007).

El segundo marcador de decidualización en las células estromales, IGFBP-1 (Rutanen et al. 1985), también se expresa en etapas tardías, como resultado de la inducción por progesterona (Rutanen et al. 1986). IGFBP-1 regula la disponibilidad de IGF-1 secuestrándolo, proceso que depende del estado de fosforilación del primero. En la decidua, se produce una variante de IGFBP-1 poco fosforilada que es menos efectiva en el secuestro de IGF-1 (Carter et al. 2004). Además, esta proteína estimula la invasión del trofoblasto (Gleeson et al. 2001). Una desregulación del sistema IGF/IGFBP-1 se vio implicado en complicaciones durante el embarazo. Se observa que valores elevados de IGFBP-1 son beneficiosos al inicio del embarazo, pero pueden ser dañinos en sus etapas más avanzadas (Carter et al. 2004).

4.4 Decidualización a nivel celular

La transdiferenciación implica la reprogramación celular de una célula somática ya diferenciada, donde se transforma en otro tipo celular diferenciado sin pasar por un estadio pluripotente. La decidualización entraría dentro de esta categoría, ya que células somáticas fibroblasticas se diferencian hacia celulas somáticas deciduales.

La regulación espacio temporal de la decidualización se encuentra en estrecha relación con la expresión paracrína de morfógenos, citoquinas y factores de crecimiento, producidos en los distintos tipos celulares que componen el endometrio, incluyendo las células estromales, epiteliales, las células inmunes locales y la vasculatura.

Una vez iniciado el proceso de decidualización, las células atraviesan un cambio constante, pasando primero por un estado pro inflamatorio, que coincide con la ventana de implantación, donde se establece la comunicación con el embrión. Luego pasan hacia un estado secretorio, que encapsula al embrión una vez que atraviesa el epitelio luminal (Gellersen & Brosens 2014) (Fig.3).

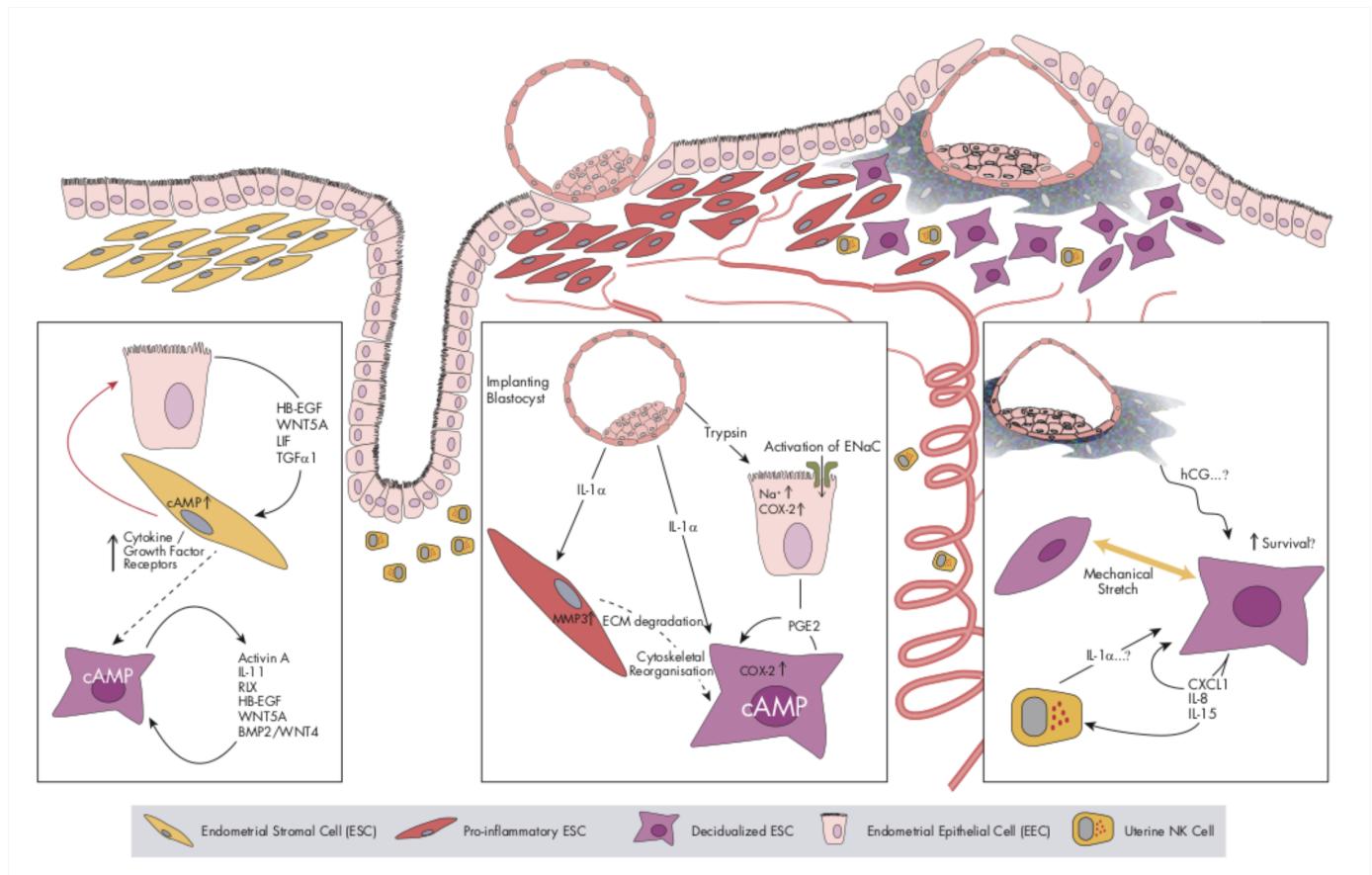


Figura 3: Esquema de los estadios por los que pasan células de endometrio durante la decidualización, así como las señales parácrinas y los tipos celulares que intervienen.

En línea con este esquema de una transdiferenciación de dos estados, un estudio reciente *in vitro* (Rytkonen et al. 2019) demostró que a los 3 días de inducir la decidualización en células tHESC, la vía inflamatoria STAT se regula positivamente. A los 8 días luego de la inducción, en lo que los autores denominan la fase tardía, la vía de STAT se encuentra reprimida, y en cambio aumenta la expresión de la vía NFkB, también inflamatoria. Estudios realizados con la tecnología de single cell RNAseq en cultivos primarios de células de estroma endometrial humano, reportan heterogeneidad de los tipos celulares endometriales luego de 6 días de ser tratados con los estímulos decidualizantes MPA y AMPc, hipotetizando que distintas poblaciones celulares puedan coexistir durante la decidualización, cumpliendo distintas funciones biológicas (Lucas et al. 2020).

4.5 Los estímulos decidualizantes: P4 y R5020

La progesterona, una vez dentro de las células del estroma endometrial, es rápidamente metabolizada a metabolitos que pueden presentar un efecto progestágeno menor. In vitro las células endometriales tratadas con progesterona tardan tan solo 8 hs en metabolizar el 70% de la hormona, mientras que luego de 24 hs., este porcentaje llega al 90% (Arici et al. 1999).

Teniendo esto en cuenta, es común utilizar progestinas sintéticas no metabolizables para el estudio de la decidualización in vitro, para así independizarse de los efectos que pudieran deberse a sus metabolitos. La progestina sintética más utilizada en los protocolos de decidualización es el 17-acetato de medroxiprogesterona (MPA), que a pesar de inducir efectos progestágenos, presenta afinidad también por el receptor de glucocorticoides, y puede inducir efectos de este tipo (Kuhl 2011), lo que hace difícil discernir la respuesta inflamatoria propia del proceso y la que el MPA pudiera disparar por el GR y no por el PGR. Por otra parte, la Promegestona (R5020), no presenta estos efectos glucocorticoides, por lo que podría ser un compuesto más adecuado para analizar la decidualización *in vitro* (Kuhl 2011) mediada por el PGR.

4.6 Modelos estadísticos lineales y su aplicación al análisis morfológico

Modelar los datos de un experimento permite contrastar las predicciones del investigador con las observaciones (Perelman & Garibaldi n.d.).

Para todo modelo estadístico, se pueden definir al menos tres componentes: la variable respuesta, las variables explicativas, y el error experimental. La primera es la variable

mediante la cual se mide el efecto del tratamiento. Las variables explicativas se miden para tener en cuenta y conocer su efecto sobre la variable respuesta. Finalmente, el error experimental es la suma de las variables no controladas, que también aportan variación a la variable respuesta, pero esta variación, a diferencia de las variables explicativas, no puede atribuirse a ninguna causa particular. Si se cuenta con más de una variable explicatoria, se puede evaluar también el efecto de su interacción; es decir, si el efecto de una variable cambia según cambie la otra.

Una manera de modelar estos datos estadísticamente es a partir de modelos lineales. Los modelos lineales son una formulación matemática de la distribución de probabilidad de la variable respuesta, de manera que la esperanza (valor esperado o media poblacional) de la variable respuesta es función matemática de las variables explicativas (los tratamientos o factores de interés). Luego, los desvíos observados respecto a esas esperanzas son los errores experimentales, donde dichos errores son independientes entre sí.

Este tipo de modelos asume que todas las observaciones, y por lo tanto todos los errores de esas observaciones respecto a la esperanza, son independientes. Sin embargo, muchas veces este no sucede.

En este trabajo por ejemplo, medimos la circularidad celular de múltiples fotos que provenían de una misma placa de cultivo celular. Estas fotos no son independientes entre sí, ya que todas se ven afectadas por las mismas condiciones de la placa.

Hay dos formas de incluir esta dependencia en el modelo lineal, que serán explicadas a continuación.

La primera añade una variable aleatoria en lo que llamaremos un modelo condicional o mixto. Este nombre lo toma por combinar los dos tipos de variables, las de efectos fijos y las de efectos aleatorios.

En general, una variable explicativa es considerada como de efectos fijos si sus niveles (categorías que puede tomar cierta variable categórica) fueron determinados explícitamente por el investigador y son de interés intrínseco para los objetivos del estudio. En nuestro caso, los diferentes niveles de la variable “tratamiento” (R5020, P4 y EtOH)

fueron seleccionados y controlados, y es de interés para el experimento cuantificar dichos efectos, por lo que son considerados efectos fijos.

En cambio, el objetivo de incluir una variable de efectos aleatorios es modelar la dependencia entre las muestras, ya que permite reducir el error no explicado y realizar comparaciones más potentes; es decir, que el modelo discernirá mejor la diferencia entre dos medias. Una variable explicativa se considera de efectos aleatorios si los niveles estudiados corresponden a una muestra aleatoria de una población de niveles para esa variable (Tognetti & Perez 2019). En este estudio, la población de niveles de la variable aleatoria “placa” estaría compuesta por todos los posibles efectos de una placa sobre la circularidad de las células, pudiendo disminuirla, aumentarla, o no modificarla (con efecto cero).

De toda esta población de efectos solo observamos 20, la cantidad de placas que utilizamos. El modelo entonces toma estos 20 efectos de las placas, reconstruye la distribución de probabilidades para toda la población (suponiendo que se distribuyen de manera normal con media cero) y de esta manera puede predecir qué circularidad tendrían las células de una placa promedio si el efecto de la placa fuera cero. No es de interés para el experimento cuantificar el efecto de cada placa, sino reducir el error no explicado y predecir cuál sería la circularidad si el efecto de la placa fuera nulo.

La segunda manera de incluir la dependencia en un modelo lineal es a partir de un modelo marginal. Aplicar un modelo marginal también reduce el error no explicado. En este caso, la falta de independencia no se incluye añadiendo una variable sino una matriz de covarianza para el error. Esta matriz contiene las covarianzas entre los niveles que presentan dependencia. En este estudio, la circularidad medida en determinada placa en determinado día, presenta una correlación con la circularidad medida en la misma placa en otro día. La matriz entonces explicita la correlación que presenta la circularidad entre los distintos días para las células de una misma placa.

Estas matrices de covarianza pueden presentar distintos patrones, las más utilizadas son la

simple, la compuesta, la regresiva de orden uno y la desestructurada (Fig.4).

$$\begin{array}{ll}
 \mathbf{A} & \begin{matrix} T1 & T2 & T3 & T4 \\ T1 & \sigma^2 & 0 & 0 & 0 \\ T2 & 0 & \sigma^2 & 0 & 0 \\ T3 & 0 & 0 & \sigma^2 & 0 \\ T4 & 0 & 0 & 0 & \sigma^2 \end{matrix} \\
 \mathbf{B} & \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix} \\
 \mathbf{C} & \begin{matrix} T1 & T2 & T3 & T4 \\ \sigma^2 & \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix} \end{matrix} \\
 \mathbf{D} & \begin{matrix} T1 & T2 & T3 & T4 \\ T1 & \sigma^2 & \sigma_{21} & \sigma_{31} & \sigma_{41} \\ T2 & \sigma_{12} & \sigma^2 & \sigma_{32} & \sigma_{42} \\ T3 & \sigma_{13} & \sigma_{23} & \sigma^2 & \sigma_{43} \\ T4 & \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma^2 \end{matrix}
 \end{array}$$

Figura 4: Matrices de covarianza más utilizadas, asumiendo que hay homocedasticidad. Sigma indica la varianza, de izquierda a derecha y de arriba hacia abajo: simple, compuesta, de regresión en grado 1 y desestructurada . Lambda al cuadrado indica la desviación estándar, distintos subíndices de lambda indican distintas desviaciones estándar. Rho es un parámetro calculado por el modelo.

En la matriz simple no hay covarianza (es equivalente a un modelo lineal donde las observaciones son independientes), mientras que en la compuesta la covarianza es la misma para todas las combinaciones. La matriz regresiva de orden uno es la que normalmente explica correlaciones entre días, ya que la correlación disminuye a medida que aumenta la distancia entre los tiempos. Por último la matriz desestructurada es la que presenta menos restricciones, pero es la que requiere un mayor número de parámetros a estimar.

4.7 Exploración de la topología de la cromatina a través de Hi-C

Las técnicas de High-throughput Chromosome Conformation Capture (Hi-C) permiten estudiar la organización en 3D de todo el genoma a una resolución que puede variar entre 1kb a 1Mb. Este protocolo consiste en fijar las células con formaldehído, realizando un crosslinking entre las regiones de la cromatina que se encuentran cerca en el espacio al momento de realizar el ensayo. Luego se fragmenta el DNA y se ligan los extremos sueltos. De esta manera quedan unidos en un mismo fragmento los *loci* que se encontraban interactuando en el genoma al momento de la fijación. Estos fragmentos son secuenciados y mapeados a las dos regiones correspondientes (Fig. 5a). Los resultados de la secuenciación indican cuántas veces se leyeron dos regiones en conjunto, pudiendo construir a partir de ellos mapas de frecuencias de interacción (Fig. 5b).

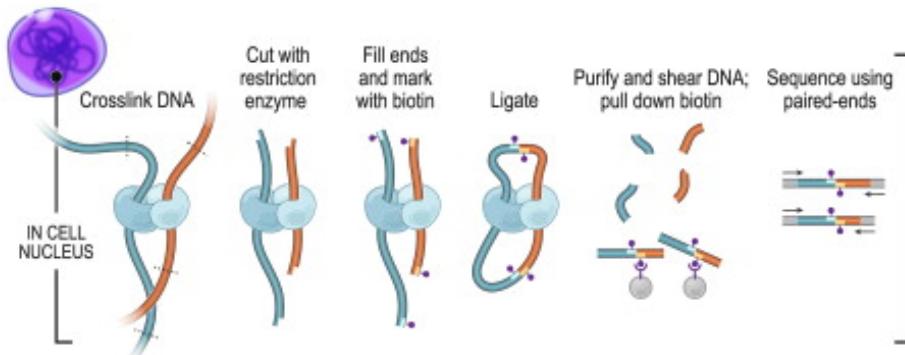
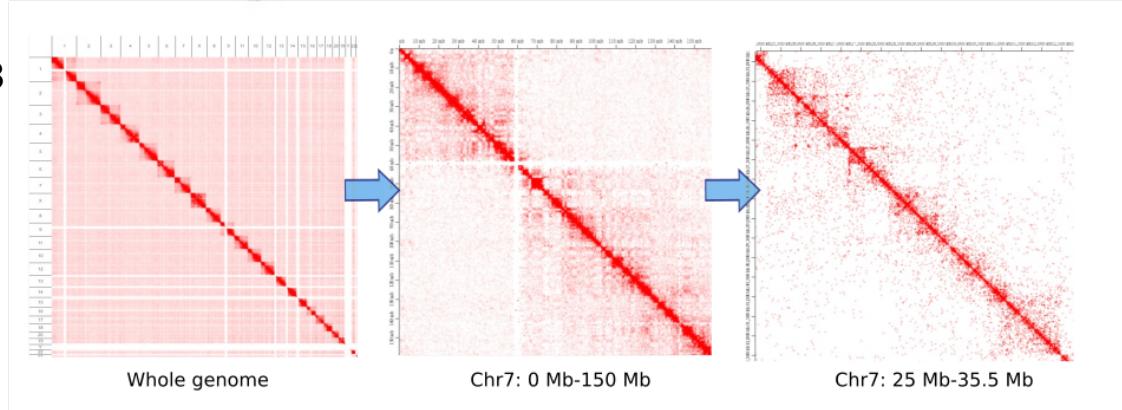
A**B**

Figura 5: a) Protocolo esquemático de Hi-C. En primer lugar, el DNA es crosslinkeado. Luego se fragmenta utilizando enzimas de restricción. Los extremos libres son biotinilados y luego religados. Luego el DNA es fragmentado nuevamente mediante enzimas de restricción, purificado mediante un pulldown y secuenciado (Lieberman-Aiden et al. 2009). b) Mapas de calor de la frecuencias de contactos para todo el genoma, para el cromosoma 7 y para una región de este. Estos mapas fueron obtenidos mediante el programa Juicer, y visualizados mediante JuiceBox (Durand et al. 2016).

El número de contactos correlaciona con la distancia a la que se encuentran espacialmente las regiones. Cuanto más lejos se encuentran dos *loci*, menos lecturas se obtendrán, y viceversa. Estas interacciones se pueden corregir para tener en cuenta la interacción esperada para cada *locus*. La probabilidad de interacción entre *loci* inmediatamente contiguos es muy alta, y decrece de manera logarítmica a medida que la distancia lineal aumenta (Fig.6). Sin embargo en la realidad experimental se encuentran *loci* que interaccionan más de lo esperado dada su distancia lineal en la secuencia genómica. Estas

regiones son las que conforman loops de interacción, por ejemplo entre promotores y enhancers (Fig.6).

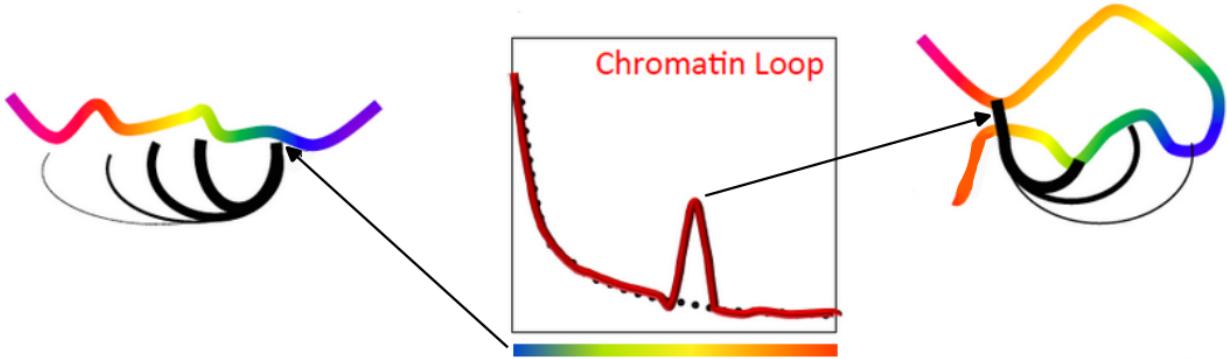


Figura 6: Esquema de la probabilidad de interacción esperada vs la observada. En la figura de la izquierda se muestra la probabilidad esperada de interacción para un loci cualquiera, donde el grosor de las líneas negras indica mayor probabilidad de interacción. En la figura central se muestra un gráfico de esta probabilidad. La línea punteada indica la interacción esperada mientras que la roja indica la observada. Ambas se comportan igual, cayendo logarítmicamente con la distancia, a diferencia del pico rojo, donde la interacción observada es mayor que la esperada. La figura derecha muestra el tipo de estructura de la cromatina que indica el pico de la figura central.

A partir de estos puntos donde la interacción es distinta a la esperada dada una estructura lineal, es posible reconstruir la estructura espacial que presentaba la cromatina en el núcleo. Los mapas de frecuencias de interacción resultantes entre diferentes regiones del genoma presentan dominios con zonas de alta interacción que pueden visualizarse como triángulos que surgen de los mapas de calor de dichas interacciones (Fig.7), denominados TADs (Topologically Associating Domains). Los TADs corresponden a dos tipos diferentes de dominios: CTCF loops, o loops dependientes de la proteína estructural CTCF (CCCTC-binding factor), y dominios compartimentales (Rao et al. 2017). Los compartimentos, o dominios compartimentales, pueden definirse como A o B, y se definen mediante análisis de componentes principales (PCA), a partir del componente que contiene el mayor porcentaje de la variación (PC1). Los dominios A presentan un estado más abierto, donde la cromatina se encuentra más laxa y activa transcripcionalmente. Los

dominos B presentan un estado de la cromatina más compactada, y es menos activa transcripcionalmente. Los análisis de PCA se realizan por lo general partiendo de mapas de contactos a una resolución de 1Mb, por lo que las zonas A y B se determinan a esta escala. Hay que tener en cuenta que si una región de 1Mb esta compuesta mayormente por cromatina abierta se la clasificará como A, a pesar de comprender también zonas más pequeñas en estado cerrado.

Los TADs formados por CTCF loops, a diferencia de los compartimentos, son definidos usando algoritmos para detectar cambios en la direccionalidad de las interacciones. Los lazos o loops de CTCF se forman a partir de la interrupción de la extrusión de la cohesina (complejo proteico involucrado en la estructuración de la cromatina) debido a la unión de CTCF. Los dominios generados por CTCF loops se pueden reconocer por una señal de interacción muy marcada en la cima del dominio, que indica la fuerte interacción entre los dos extremos del dominio a través de la unión de CTCF. Los dominios compartimentales, no poseen esta marca, y se clasifican como una zona A o B a partir del PC1, ya que las interacciones menos frecuentes entre dominios vecinos que pertenecen a compartimentos distintos crean límites compartimentales.

Es cada vez más evidente que esta compartimentalización es clave en la formación y mantenimiento de la identidad celular, ya que es capaz de generar microambientes en zonas específicas del genoma, facilitando el control sincronizado de la expresión de los genes que allí se encuentran. Así, la dinámica de las interacciones de la cromatina puede condicionar la posterior regulación de la transcripción.

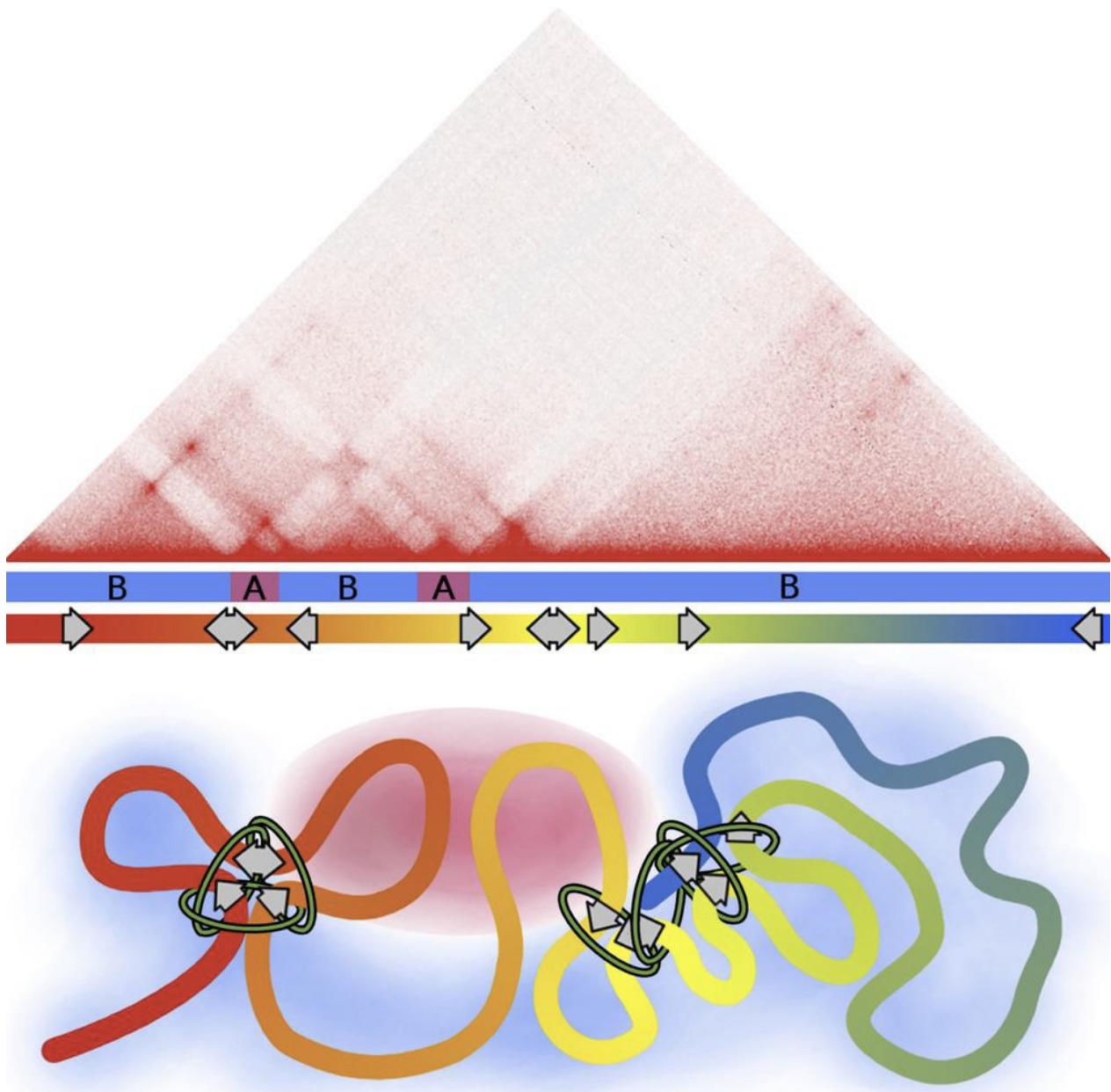


Figura 7: Esquema indicando cómo se relacionan los heatmaps de contactos con la compartimentalización y los loops de CTCF. El dibujo inferior indica la estructura de un fragmento de la cromatina, donde las flechas indican los sitios de unión de CTCF, las áreas rosadas zonas abiertas de la cromatina (A), activas transcripcionalmente, y las áreas azules las cerradas (B). Por encima se esquematiza el mismo fragmento de cromatina pero de manera lineal, sobre el cual se indica la compartimentalización. En la parte superior se encuentra el heatmap del que se puede deducir todos los datos esquematizados en la parte inferior. (Nichols & Corces 2021).

Como se mencionó antes, la manera en que los análisis de Hi-C permiten conocer en qué estado (A o B) se encuentra cada región de la cromatina es realizando análisis de

componentes principales.

Partiendo de los perfiles de contacto para cada *locus*, que describe las distintas probabilidades de interacción para dicho *locus* con el resto del genoma, se realizan mapas de correlación de Pearson. Estos mapas entonces muestran el grado de correlación entre los perfiles de contacto; si dos *loci* interactúan de manera similar con el resto del genoma, entonces su correlación será alta, y viceversa. Es a partir de estas mapas o matrices de correlación que se realizan los análisis de componentes principales (PCA), de tal manera de agrupar todos los sitios del genoma según cuán parecidos son sus perfiles de interacción. El PCA es un tipo de análisis estadístico que permite describir un conjunto de datos en términos de nuevas variables no correlacionadas, reduciendo su dimensionalidad. Estas nuevas variables, llamadas componentes principales, son combinaciones lineales de las variables originales, de manera tal que la dirección resultante maximice la varianza. De esta forma, cada componente principal capturará cierto grado de variación observada de los datos. El primer componente principal o eigenvector es el que captura la mayor varianza. Bajo la premisa de que dominios compartimentales que poseen la misma actividad transcripcional poseen perfiles de interacción similares entre sí, el eigenvector se divide clásicamente en dos sets de valores, positivos y negativos, que coinciden bien con las marcas epigenéticas de las regiones activas (A) e inactivas (B) (Lieberman-Aiden et al. 2009).

Cabe aclarar que a pesar de que esta categorización funciona bien en tipos celulares de mamíferos, la información que arrojan análisis de microscopía y epigenética, en conjunto con la utilización de otros métodos para agrupar los estados compartimentales (en lugar de el estándar PCA), sugiere que el estado transcripcional del genoma es más complejo, y su clasificación binaria podría ser una sobresimplificación (Nichols & Corces 2021). Por ejemplo el uso de una matriz oculta de Markov resolvió la cromatina en dos compartimentos activos A y 4 inactivos B en células GM12878, que difieren en distintas marcas cromatínicas como las modificaciones de histonas y la asociación al nucléolo o laminina (Rao et al. 2014).

5 Objetivos e hipótesis

Basándose en el claro cambio cualitativo en la morfología celular que se da durante la decidualización, este trabajo se propone explorar una metodología, simple, rápida y económica para cuantificar la decidualización midiendo la morfología celular.

Por otra parte, este trabajo se propone evaluar, en este mismo modelo, la reorganización de la compartmentalización A/B de la cromatina durante la decidualización. Esto último se debe a la hipótesis de que la organización de la cromatina juega un rol activo en la transdiferenciación de las células fibroblásticas.

Este trabajo se encuentra enmarcado dentro de un proyecto más amplio que busca dar luz sobre cómo se desarrollan e inter-relacionan causalmente los cambios en la expresión génica, la reorganización cromatínica y el posicionamiento de factores de transcripción durante la decidualización, a partir de análisis de RNA-seq, Hi-C y ChIP-seq.

6 Materiales y Métodos

6.1 Cultivo celular y tratamiento

El experimento analizado en este trabajo fue realizado previamente en nuestro laboratorio. Se utilizaron células estromales endometriales humanas transformadas (t-HESC, ATCC CRL-4003) mediante infección con sobrenadante de la línea celular empaquetadora pA317-hTERT (Geron Corp.; Menlo Park, CA) que expresaba hTERT y genes resistentes a puromicina (Krikun et al. 2004) compradas al LGC y cedidas por Ana Mestre-Citrinovitz (LGC Standards GmbH, Wesel, Alemania). Estas fueron cultivadas por 48 horas en medio DMEM-F12 (D2906, Sigma-Aldrich) suplementado con 10% de suero fetal bovino (FBS) charcoalizado y deplegado de hormonas (DCC-FBS) (Mestre Citrinovitz et al. 2019), , de acuerdo a las instrucciones recomendadas por ATCC. Las t-HESC fueron tripsinizadas, cuantificadas y plaqueadas en placas de 100mm con una densidad de 1×10^6 células/placa. Luego de 24hs, se removió el medio y se lavaron las células con PBS. A continuación, se les realizaron los siguientes tratamientos en medio DMEM-F12 suplementado con 2%-DCC-FBS: Progesterona (P4) 10^{-6} M, Estradiol (E2) 10nM, y 8-Bromo-AMPc (8-Br-cAMP, Sigma-Aldrich, St. Louis, MO, USA) 0,5mM; o, Progestina (R5020) 10^{-8} M, Estradiol (E2) 10nM y 8-Bromo-AMPc (8-Br-cAMP, Sigma-Aldrich, St. Louis, MO, USA) 0,5 ; o, vehículo Etanol (EtOH) 0,1%.

Para la cuantificación morfológica se tomaron fotos de cada placa (5 placas por tratamiento) durante todos los días, y para los ensayos de Hi-C y RNAseq se tomaron muestras de los cultivos a la hora, tres días y seis días, luego de la inducción.

6.2 Muestreo y cuantificación morfológica

El set de datos constó de la suma de 2 imágenes por placa para cada día para cada tratamiento (siendo 5 placas por tratamiento), a excepción de la placa 4 para el día 4 del

tratamiento R5020 que no contaba con imágenes, con lo cual fueron analizadas 188 fotos en total (10 fotos para el T0, 30 fotos para cada uno de los días 1, 2, 3, 5 y 6, y 28 para el día 4).

Todos los macros utilizados son originales de este trabajo. Fueron escritos en el entorno de ImageJ Macro Language, que permite automatizar series de comandos en este programa.

La circularidad se midió mediante el programa ImageJ de la siguiente manera.

En primer lugar se tomaron todas las fotos y se trazaron dos líneas verticales de manera automática utilizando ImageJ. Como consecuencia las imagen quedó dividida en 3 partes iguales. Las líneas verticales se trazaron agrupando todas las imágenes en un stack(agrupamiento de imágenes) y luego corriendo el siguiente macro para trazar las dos líneas sobre todas las imágenes al mismo tiempo:

```
setLineWidth(3); // determina el ancho de la linea
setColor("red"); // definir color de la linea
//Loop que toma cada slice del stack (cada imagen)
for (i=1; i<nSlices+1;i++) {
    Stack.setSlice(i); //toma el slice
    width = getWidth(); //mide el ancho
    height = getHeight(); //mide el alto
    //se dibujan las 3 lineas
    for (j=1; j<4; j++) {
        x = (width/4)*j; // ubicación de cada linea
        y = height; // largo de la linea, que es igual al
                    //largo de la foto
        drawLine(x,0,x, y); // dibuja la linea desde el punto (x,0)
                    //hasta el (x,y)
    }
    run("Select None"); //deselecciona el slice
```

}

Luego, se obtuvo el contorno o perímetro de las células en contacto con estas líneas guía. Esto se realizó mediante una segmentación de la imagen, donde se delimita o dibuja un área cerrada de la misma, definiéndola como un “área de interés” (ROI). ImageJ permite seleccionar estas áreas de interés de manera automática o de manera manual. En este caso, ante la imposibilidad de detectar las células de manera automática, la segmentación se realizó de manera manual, dibujando el contorno de las células que estaban en contacto con las líneas guía, como se observa en la figura 8.

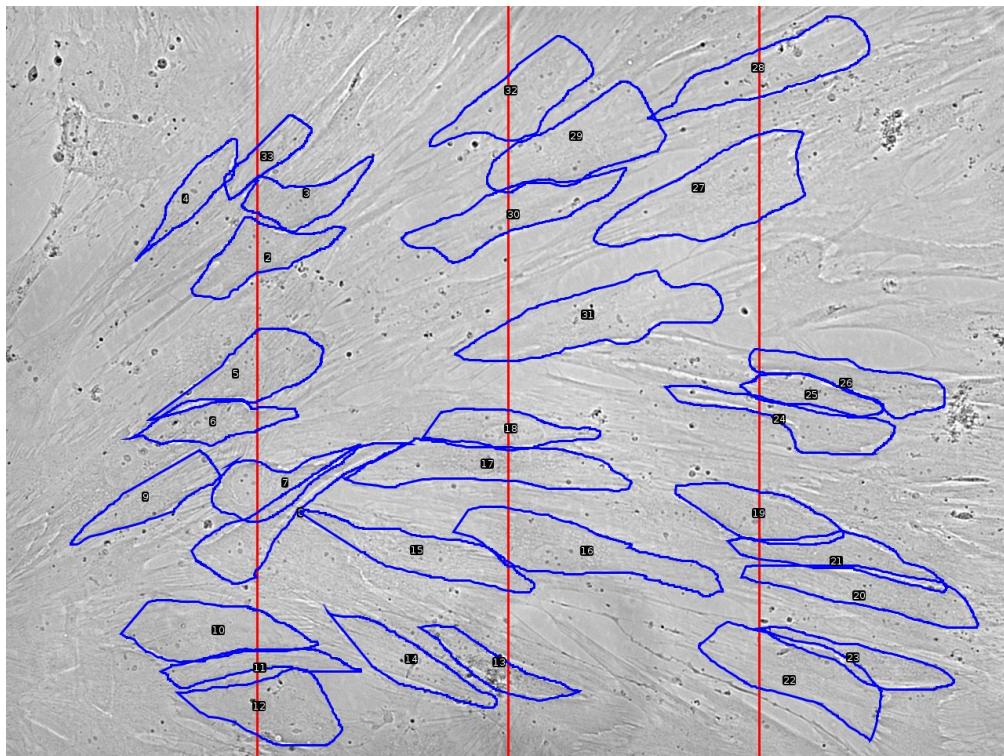


Figura 8: Imagen representativa de la segmentación celular realizada en ImageJ. En este caso la foto corresponde al día 3 de la placa nº5 tratada con R5020. Las líneas rojas son las líneas guía y las zonas azules son los contornos de las células, o ROIs, dibujadas a mano alzada.

Finalmente, se midió para cada foto la circularidad de las células segmentadas (ROIs), siguiendo la ecuación: $circularity = 4\pi \frac{area}{perímetro^2}$

Se exportaron los datos obtenidos (índice de circularidad por célula por foto) a una tabla mediante otro macro:

```
dir = getDir("image"); //toma la ubicación de la imagen abierta
name = getInfo("image.title"); //toma el nombre del archivo
Savename = File.getNameWithoutExtension(dir+name); //dónde guardar
//el archivo
dir=".~/Documents/my_path/" // Dónde guardar la tabla de mediciones

roiManager("measure"); // medir los ROIs dibujados
selectWindow("Results"); // seleccionar la ventana de resultados
//de la medición

// guardar los resultados con el mismo nombre que la imagen
saveAs("Measurements", dir + Savename +".csv");
// Guardar los ROIs en una carpeta .zip con el mismo
//nombre que la imagen
roiManager("save", dir + Savename +".zip");

selectWindow(name); // seleccionar la ventana de imagen
saveAs("tiff",dir+name); //guardar la imagen con los ROIs dibujados

roiManager("delete"); // Limpiar la ventana de ROIs
close(); // cerrar la ventana de ROIs
selectWindow("Results"); // cerrar la ventana de resultados
run("Close"); // cerrar la imagen
// El imageJ queda vacío para abrir la siguiente foto
```

Nota: los comentarios se denotan mediante “//”

Luego se unificaron todos los archivos csv obtenidos y se generó una tabla con la circularidad para cada célula con la metadata correspondiente, indicando placa, n° de imagen, tratamiento y día, como se observa en la imagen (Fig. 9).

cell	file	Area	Perim.	type	treatment	time	placa	Circularity	plateOK
15	15 SD P4 20	6302	598	SD P4	P4	5	4	0.4999449	P4 4
16	16 SD P4 26	11999	470	SD P4	P4	5	4	0.6825889	P4 4
17	17 SD P4 26	15833	559	SD P4	P4	5	4	0.6367214	P4 4
18	18 SD P4 26	17210	589	SD P4	P4	5	4	0.6233904	P4 4
19	19 SD P4 26	7044	419	SD P4	P4	5	4	0.5041981	P4 4
20	20 SD P4 26	4957	469	SD P4	P4	5	4	0.2831934	P4 4
21	21 SD P4 26	12362	557	SD P4	P4	5	4	0.5007122	P4 4
22	22 SD P4 26	6984	502	SD P4	P4	5	4	0.3482625	P4 4
23	23 SD P4 26	9071	442	SD P4	P4	5	4	0.5834726	P4 4
24	24 SD P4 26	14403	720	SD P4	P4	5	4	0.3491386	P4 4
25	25 SD P4 26	37054	758	SD P4	P4	5	4	0.8104133	P4 4
26	26 SD P4 26	16896	558	SD P4	P4	5	4	0.6819073	P4 4
27	27 SD P4 26	15497	569	SD P4	P4	5	4	0.6014963	P4 4
28	28 SD P4 26	20250	634	SD P4	P4	5	4	0.6330768	P4 4
29	29 SD P4 26	30897	820	SD P4	P4	5	4	0.5774288	P4 4
30	30 SD P4 26	15728	710	SD P4	P4	5	4	0.3920728	P4 4
31	31 SD P4 26	10785	561	SD P4	P4	5	4	0.4306300	P4 4
32	32 SD P4 26	9784	432	SD P4	P4	5	4	0.6588079	P4 4

Showing 15 to 33 of 4,425 entries, 10 total columns

Figura 9: Fragmento de la base de datos final obtenida a partir de imageJ, utilizada para el análisis morfológico. La columna “placaOK” indica un identificador único para cada placa.

6.3 Análisis de la circularidad

El análisis de la circularidad se realizó utilizando el lenguaje de programación R, en el entorno de desarrollo integrado (IDE) Rstudio.

Se realizó en primer lugar un análisis exploratorio de los datos. Luego se implementaron distintos modelos lineales y se midió su rendimiento. Se eligió el mejor modelo y se corroboró que dicho modelo cumpliera con los supuestos. Finalmente, se realizó un

análisis de la varianza (anova) y se comparó estadísticamente la circularidad predicha para los distintos tratamientos.

Para los modelos lineales implementados se tomó como variable respuesta la circularidad media por foto, y no la circularidad por célula. Esto es debido a que la distribución de la circularidad por célula no se ajusta a una distribución normal, supuesto necesario para aplicar un modelo lineal.

La distribución normal teóricamente abarca de menos infinito a más infinito, mientras que el índice utilizado para la medida de circularidad varía entre 0 y 1. Debido a esto la circularidad por célula pertenecía a una distribución del tipo beta, una familia de distribuciones continuas de probabilidad definidas en el intervalo [0 , 1], es decir, que pueden tomar valores del cero al uno .

Sin embargo, generalmente una distribución que presenta límites superior e inferior como en este caso, donde los límites son 0 y 1 y la distribución es estrictamente beta, puede ajustarse también a una distribución normal, mientras que la distribución se acerque a una campana de Gauss. A pesar de esto, la prueba estadística de Shapiro-Wilks realizada para la circularidad por célula arrojó un p-valor significativo (<0,05), indicando que la circularidad por célula no puede ajustarse a una distribución normal, por lo que no se puede aplicar un modelo lineal tomando la circularidad por célula como variable respuesta. Esto dificultaría el trabajo posterior ya que actualmente no se encuentran disponibles paquetes en R para modelar este tipo de distribución con las características que este diseño experimental presenta (un modelo marginal para distribución beta).

El Teorema Central del Límite postula que la distribución de la media de una muestra aleatoria, cuando el tamaño de la muestra es lo suficientemente grande, sigue aproximadamente una distribución normal. Teniendo esto en cuenta, tomamos como variable respuesta la circularidad media por foto. Posteriormente se corroboró gráficamente (mediante QQplot y gráfico de residuos vs valores predichos por el modelo) y mediante pruebas estadísticas (test Shapiro-Wilks) que esta variable se ajusta correctamente a una distribución normal (Fig. 12a).

La cantidad de células por foto fue tomada en cuenta para todos los modelos mediante el parámetro *weights*, de modo que se le dio más peso a las medias de las fotos con mayor cantidad de células.

Con el objetivo de mostrar claramente los pasos seguidos durante el análisis y permitir su reproducibilidad en las siguientes secciones se muestra el código utilizado para el análisis morfológico tal como figuran en el archivo Rmarkdown con el que se creó este documento. Todos los modelos planteados se implementaron mediante el paquete nlme.

1. Librerías utilizadas

```
library(ggplot2) # Gráficos  
library(plyr) # Manipulación y operaciones con data frames  
library(gridExtra) # Para el correcto display de las figuras  
library(cowplot) # Para el correcto display de las figuras  
library("reshape2") # Para modificar dataframes  
library(GGally) # Para hacer el gráfico de correlaciones  
library(nlme) # Para modelar los datos  
library(knitr) # Para crear archivo pdf  
library(xtable) # Para insertar las tablas en pdf  
library(emmeans) # Para comparación de efectos simples  
library(ggeffects) # Para graficar las comparaciones  
library(car) # Para el correcto display de las figuras  
library(dplyr) # Manipulación y operaciones con data frames  
# library(kableExtra)
```

2. Preparación de los datos

```

#creo mi dataframe con mis datos

path<- "/home/luciana/Documents/LABO/morfología_DF_FINAL_5.4.csv"

#path a la base de datos

df <- read.table(path, header = T, sep = ",")


# Creo un identificador único para cada placa

library(dplyr)

df<- df %>%
  mutate(placaOK = paste(treatment,placa))



# Creo nuevo df con circularidad promedio por foto, df2

df2 <- df %>%
  group_by(file) %>%
  summarise(circularidadMedia = mean(Circularity, na.rm = TRUE),
            tiempo= mean(time, na.rm =TRUE),
            tratamiento=unique(treatment), Nrocelulas=n(),
            placa=unique(placaOK), Foto=unique(file))




# Cambio los datos del día de tipo numérico a factor

df2$tiempo <- as.factor(df2$tiempo)


# Creo la base de datos solo con T0 y día 1, dfT0

df_T0 <- df2[df2$tiempo %in% c("0", "1"), ]

df_T0 <- as.data.frame(df_T0) # Defino df_T0 como una base de datos,
#para que ggeffects lo tome correctamente


# Creo la base de datos con el resto de los tiempos menos T0,
#para el modelo de medidas repetidas

```

```
df_wo_T0 <- df2 %>% filter(tiempo !=0)
```

3. Gráficos exploratorios

```
#### Gráfico de densidad  
dens<-ggplot(data = df2, aes(circularidadMedia)) +  
  geom_density(aes(fill=factor(tratamiento), alpha=0.1)) +  
  xlab("Circularidad media por foto")+ylab("Densidad") +  
  scale_x_continuous(guide = guide_axis(check.overlap = TRUE,  
                                         angle = 45))+  
  theme(legend.position="right",  
        legend.text=element_text(size = 8),  
        legend.title = element_text(size=8, face="bold")) +  
  facet_grid(. ~tiempo, labeller=label_both )  
  
#### Boxplot  
box <- ggplot(df2, aes(x=tiempo, y=circularidadMedia,  
                        colour=tratamiento)) +  
  geom_boxplot(aes(color=tratamiento), coef=TRUE)+  
  theme_bw() +  
  geom_jitter(alpha=0.2, size=0.2,aes(color=tratamiento),  
              position = position_jitter(width = .2))+  
  theme(legend.position="right",  
        legend.text=element_text(size = 8),  
        legend.title = element_text(size=8, face="bold")) +  
  ylab("Circularidad media por foto") + xlab("Día")  
  
#### Spaguetti plot
```

```

# Creo una base de datos con el promedio por placa para poder
#graficar solo el promedio de cada placa

cdata <- ddply(df_wo_T0, c("tratamiento", "tiempo", "placa"),
               summarise,
               N      = sum(!is.na(circularidadMedia)),
               circularidad = mean(circularidadMedia),
               sd     = sd(circularidadMedia),
               se     = sd / sqrt(N))

#plot

spaghetti<-ggplot(cdata, aes(x =tiempo, y = circularidad,
                               group = placa, colour=placa)) +
  geom_line() +
  geom_point()+theme(legend.title = element_text(size = 8),
                     legend.text = element_text(size = 8),
                     legend.position="right")+
  facet_grid (. ~ tratamiento)+ labs(y="Circularidad por placa")

```

4. Análisis de correlación

```

# Genero una columna por día para la tabla df_wo_T0

datos_wide <- dcast(df_wo_T0,
                      formula = placa+tratamiento ~ tiempo,
                      value.var = "circularidadMedia",
                      fun.aggregate=mean)

tiempos <-datos_wide[3:8] # Selecciono las columnas de los días,
#si quiero hacer una plot de correlación sólo para determinado

```

```

#tratamiento seleccionar solo las filas de ese tratamiento

tiempos[tiempos == 0] <- NA # Reemplazo los ceros por NA para
#que no los tenga en cuenta en la correlación

# Gráfico de correlación
corrPlot <- ggpairs(tiempos, axisLabels="none" ,columnLabels=c(
  "Day 1", "Day 2", "Day 3", "Day 4", "Day 5", "Day 6"))

```

5. Selección de modelos

```

##### Modelo medidas repetidas #####
#Modelo 1: modelo condicional
m1 <- lme(circularidadMedia ~ tratamiento*tiempo, random= ~1|placa,
           weights = ~(1/Nrocelulas), data = df_wo_T0)

#Modelo 2: Simetría compuesta.
m2<-gls(circularidadMedia ~ tratamiento*tiempo, correlation =
          corCompSymm(form = ~ 1 | placa), weights = ~(1/Nrocelulas)
          , df_wo_T0)

#Modelo 3: matriz autoregresiva de orden 1
m3<-gls(circularidadMedia ~ tratamiento*tiempo, correlation =
          corAR1(form = ~ 1 | placa), weights = ~(1/Nrocelulas),
          df_wo_T0)

#Modelo 5: matriz desestructurada NO CONVERGE, no es tenido en cuenta
#m5<-gls(circularidadMedia ~ tratamiento*tiempo,

```

```

#correlation = corSymm(form = ~ 1 | placa), df2) no
#pudimos obtener un resultado, no convergía el modelo

#TABLA de comparación de modelos
AIC<-AIC(m1,m2,m3) # Lista con los AIC de cada modelo
# Lista con los cuadrados medios del error de cada modelo
CMe<-c(summary(m1)$sigma^2,summary(m2)$sigma^2,summary(m3)$sigma^2)
modelo <- c("condicional","compuesto","AR1") # Nombres de los modelos

# Creo una tabla con las listas como columnas, llamada comp
comp <- cbind(modelo, CMe, AIC)
colnames(comp)<-c("modelo", "CMe", "gl","AIC") # Nombro las columnas

##### Modelo T0 vs D1 #####
mT0 <- lme(circularidadMedia ~ tratamiento,
            random= ~1|placa, weights = ~(1/Nrocelulas),
            data = df_T0)

```

6. Cumplimiento de supuestos

```

##### MODELO 1 #####
e_m1 <- resid(m1, type="pearson") # Residuos de Pearson
pre_m1 <- predict(m1) # Predichos
# Coeficientes aleatorios estimados
Bi_m1 <- m1[["coefficients"]][["random"]][["placa"]]

res_vs_pred_m1 <- {plot(pre_m1, e_m1, xlab="Pred.", ylab =
                         "Pearson Res.", main="RE vs PRED")
                     abline(0,0)} # Plot de residuos vs predichos

```

```

# QQplot de los residuos
qqplot_e_m1 <- qqPlot(e_m1, main = "QQPlot standarized res.")

# QQplot de los efectos aleatorios
qqplot_B_m1 <- qqPlot(Bi_m1, main="QQPlot Random effects")

# Prueba de Shapiro para e y Bi
shapiro.test(Bi_m1)
shapiro.test(e_m1)

##### MODELO T0 #####
e_mT0 <- resid(mT0, type="pearson") # Residuos de Pearson
pre_mT0 <- predict(mT0) # Predichos
# Coeficientes aleatorios estimados
Bi_mT0 <- mT0[["coefficients"]][["random"]][["placa"]]

res_vs_pred_mT0 <- {plot(pre_mT0, e_mT0, xlab="Pred.",
                           ylab = "Pearson Res.", main="RE vs PRED")
abline(0,0)} # Plot de residuos vs predichos

# QQplot de los residuos
qqplot_e_mT0 <- qqPlot(e_mT0, main = "QQPlot standarized res.")

# QQplot de los efectos aleatorios
qqplot_B_mT0 <- qqPlot(Bi_mT0, main="QQPlot Random effects")

# Prueba de Shapiro para e y Bi
shapiro.test(Bi_mT0)
shapiro.test(e_mT0)

```

7. Anova

```

anovaM1 <- anova(m1) # Anova del modelo
#seleccionado para medidas repetidas
anovaT0 <- anova(mT0) # Anova del modelo para T0

```

8. Comparación de efectos simples

```

# Seteo de emmeans
options(emmeans= list(emmeans = list(infer = c(TRUE, TRUE)),
                      contrast = list(infer = c(TRUE, TRUE)))))

# Comparaciones para el modelo 1,
#me fijo las diferencias significativas para
#asignar las letras en el gráfico final
comp1<-emmeans(m1, pairwise ~ tratamiento|tiempo )

# Comparaciones para el modelo T0, me fijo las diferencias
#significativas para asignar las letras en el gráfico final
compT0<-emmeans(mT0, pairwise ~ tratamiento )

# Gráficos finales, asigno las letras a mano
estim <- ggemmeans(m1, terms = c( "tiempo","tratamiento"),
                     ci.level=0.95)
estimT0 <- ggemmeans(mT0, terms = c( "tratamiento"), ci.level=0.95)
p1 <- plot(estim)+
  labs(y = "circularidad", title =
    "Circularidad media marginal por día por tratamiento.",
    size = 5, main=2)+annotate("text",
    x = c(0.91,0.98,1.08,

```

```

1.92, 2, 2.1,
2.92, 3, 3.1,
3.92, 3.99, 4.075,
4.918, 5, 5.075,
5.917, 6.0, 6.1),
y = c(0.372, 0.408, 0.45,
0.42, 0.5, 0.507,
0.401, 0.57, 0.55,
0.487, 0.59, 0.67,
0.468, 0.64, 0.72,
0.52, 0.71, 0.724),
label = c("A", "AB", "B",
"A", "B", "B",
"A", "B", "B",
"A", "B", "C",
"A", "B", "C",
"A", "B", "B"), size=2)+

theme(text = element_text(size=12))

p2 <- plot(estimT0)+

labs(y = "circularidad", title = "Circularidad media marginal",
size = 5, main=1)+annotate("text",
x = c(1,2,3,4),
y = c(0.43, 0.36, 0.39, 0.43),
label = c("A", "B", "AB", "AAB"), size=3)+

theme(text = element_text(size=8))

```

6.4 Hi-C

El protocolo de Hi-C se realizó siguiendo el protocolo *in-nuclei* descrito a continuación (Rao et al. 2014). El pellet de células se cross-linkeo con formaldehído y luego fue tratado con buffer de lisis a 0 C°. Luego se resuspendió en NEBuffer2 (New England Biolabs) conteniendo 400U MboI de enzima de restricción MboI. Para la reparación de los extremos y marcado de los fragmentos se utilizaron biotin-11-dATP y DNA Polymerase I Large (Klenow) fragment. Para ligar y revertir el cross-linking se utilizó T4 DNA ligasa y proteinasa K, respectivamente, dejándolas actuar toda una noche. La purificación de DNA fue realizada mediante el kit de purificación de DNAeasy (Qiagen) y se cuantificó usando un fluorómetro Qbit. Las librerías se prepararon de acuerdo a las especificaciones de Ilumina, antes de cortar el DNA utilizando un Bioruptor apto para DNA (Diagenode BioruptorPico) y precipitación de biotina usando Streptavidin T1 beads magnéticas. La amplificación de las librerías sobre las beads fueron realizadas de acuerdo a las especificaciones de Ilumina para una secuenciación Hi-Seq 2000 paired-end, con 4-12 ciclos de PCR y luego se purificó el DNA mediante AMPure XP beads. Las lecturas resultantes (200 millones de reads por muestra) se mapearon al genoma de referencia (hg38) usando BWA (Li & Durbin 2009). Los datos obtenidos a partir del Hi-C, incluidos los mapeos, generación de matrices de interacción, normalización ICE (Imakaev et al. 2012), generación de matrices de correlación y la obtención de los componentes principales que se desprenden de estos últimos, fueron obtenidos mediante la herramienta TADbit (Serra et al. 2017).

6.5 RNAseq

A las muestras tomadas a los distintos tiempos, de los distintos tratamientos, se les extrajo el RNA total, usando un RNeasy Plus Mini Kit (QIAGEN), se armaron las librerías y se secuenciaron mediante secuenciador Illumina HiSeq 2000 and 2500. La secuenciación se

hizo en las instalaciones de la Unidad de Genómica del Centro de Regulación Genómica (CRG), en Barcelona. Las lecturas fueron alineadas con el genoma humano de referencia versión hg38, conservando todas aquellas que se alineaban con menos de dos bases nucelotídicas sin alinear. Los resultados obtenidos de la secuenciación fueron procesados con un script para análisis de RNA utilizando el paquete DESeq2 del entorno de software libre y lenguaje de programación R y la herramienta DRIMSeq de R/Bioconductor. De esta forma, se obtuvo una tabla de datos de las lecturas por gen normalizadas.

6.6 Asignación de los signos de PC1

Se obtuvo a partir de los análisis realizados en Tadbit los resultados del PCA de las matrices de correlación. Este PCA se realizó a una escala de 100kb, por lo que a cada región de 100kb (en adelante referido como bin) le correspondía un valor de PC1.

Para cada cromosoma de cada muestra todas las zonas abiertas presentaban el mismo signo de PC1, positivo o negativo de manera azarosa, mientras que las zonas cerradas presentaban el signo contrario (negativo o positivo según correspondiera). Se corrigió la asignación de signo de los eigenvalues (PC1) mediante un script de Python original, invirtiendo los signos de PC1 de ser necesario de manera de que las zonas abiertas siempre presenten un PC1 positivo, y las zonas B negativo. Para saber qué signo presentaban las zonas A y B (y así saber si era necesario invertir los signos o no) se calculó para cada cromosoma de cada muestra la correlación entre los PC1 y la cantidad de reads normalizados de RNAseq en esa región (100kb). Teniendo en cuenta que las zonas abiertas (A) son más activas transcripcionalmente que las cerradas (B), Se consideró como zonas A los valores positivos del PC1 en los casos en que las lecturas de RNAseq aumentaban conforme aumentaba el PC1 (correlación positiva), mientras que cuando esta correlación era negativa, se consideraron zonas A los valores negativos del PC1. Los pasos realizados para realizar este análisis se enumeran a continuación, junto a los bloques de código que les corresponden.

Para analizar el número de lecturas de RNAseq por bin se buscaron primero dentro de que bin se encontraba cada gen (paso 3). Luego de modificar los nombres de las muestras en la base de datos de PC1 y RNAseq, de manera que las muestras se nombren de igual modo (paso 4) se agruparon las lecturas por bin, para saber la cantidad de reads totales para cada bin (independientemente de los genes a los que correspondían) (paso 5).

Luego se calculó el índice de correlación de Pearson entre los valores de PC1 y las lecturas de RNAseq para cada cromosoma de cada muestra. Los valores de PC1 no se modificaron en los casos en los que la correlación era positiva, y se invirtieron los valores de PC1 en caso de ser negativa (paso 6).

1. Paquetes utilizados

```
import pandas as pd
import seaborn as sns
from scipy import stats
import matplotlib.pyplot as plt
import numpy as np
from tqdm import tqdm
```

2. Bloque de código para acceder a las bases de datos de los PC1 y las lecturas normalizadas de RNAseq desde Rstudio

```
# PC1 = abro el tsv con los PC1
PC1 = pd.read_csv(
    "Bases de datos/Resigned Eigen tHESC whole genome.txt",
    sep="\t", decimal=",")
# RNA = abro el tsv de los reads de RNA-seq normalizados
RNA=pd.read_csv(
```

```
"Bases de datos/2021-02-19_FLeDily_RNAseq_tHESC_FLD_hg38_allsamples_normalized.txt", sep="\t")
```

3. Bloque de código para iterar sobre cada gen y asignar su bin correspondiente del PC1 (que tiene una resolución de 100kb).

```
#función para redondear
def roundint(value, base):
    return int(value) - int(value) % int(base)

#iterar por los genes asignarles un bin
#start indica el cominezo del gen y end el final,
#resolution es la resoslución del PC1
def assignBins(start, end, resolution):

    # en qué bin caerían el principio y final del gen
    roundStart = roundint(start, resolution)
    roundEnd = roundint(end, resolution)

    #si cae entre dos bins defino el bin dependiendo de que lado
    #hay más bases
    if roundStart != roundEnd:
        # n° bases por debajo del límite del bin
        Below = roundEnd-start
        # n° bases por arriba del límite del bin
        Above = end-roundEnd

        # si hay más gen en el bin inferior, asigno ese
```

```

if Below>Above:
    bin=roundStart

    # y asigno el superior si sucede lo contrario
else:
    bin=roundEnd

#Si el gen cae en un solo bin es indistinto
else:
    bin=roundStart

return bin

# corro la función
RNA["bin"] = RNA.apply(lambda x: assignBins(x.start ,x.end, 100000),
                        axis = 1)
#output: rna con un bin asignado para cada gen

```

4. Bloque de código para compatibilizar los nombres de las columnas de ambas tablas (PC1 y RNA). Se modificaron los nombres de las muestras en las dos bases de datos para unirlas.

```

# igualo la manera qen que se indican los cromosomas
PC1["chr"] = "chr"+PC1.chromosome
# Agrupo las lecturas del RNA por bin
RNA = RNA.groupby(["chr","bin"], as_index = False).sum()

# Igualo la manera de nombrar los samples

```

```

meta_data = pd.read_excel("Metadata_Hi-C_tHESC.xlsx", header=None)
namesRNA = RNA.filter(like='PDR', axis=1).columns.tolist()
namesRNA.sort(key = lambda x: int(x.strip('PDR_')) )
namesPC1 = [" ".join(i.split(" ",2)[:2])
            for i in PC1.columns.tolist()]
            if i.startswith('D')]==True]#list PC1 columns

# Promedio las lecturas de las tres réplicas de RNA
for n in range(10):
    RNA[namesPC1[n]] = (RNA[namesRNA[n]]+
                           RNA[namesRNA[n+10]]+
                           RNA[namesRNA[n+20]])/3
RNA = RNA.drop(namesRNA, axis=1)

```

5. Bloque de código para el cálculo de las correlaciones

```

#input: rna con un bin asignado para cada gen y df pc1

#Uno el df de PC1 y rna
merged = PC1.merge(RNA, how='inner', on=["chr","bin"])
# Creo una lista de todos los cromosomas
chromosomes= merged.chr.unique().tolist()
PC1cols = [i for i in PC1.columns.tolist()
            if i.startswith('D')]==True] #todas los samples
dfSign = pd.DataFrame() #tabla donde poner los resultados

for col in PC1cols: #itero por los samples
    #itero por los cromosomas de cada sample

```

```

for chromosome in chromosomes:

    compartment=merged.loc[merged.chr==chromosome][col]

    expression=merged.loc[merged.chr==chromosome][" ".join(col.split(" ")):2

    sign=0

    # Si la correlación es positiva asigno un 1
    # en la tabla de resultados

    if stats.pearsonr(compartment, expression)[0]>0:
        sign=1

    else: #si es negativa asigno -1
        sign=-1

    dfSign.loc[chromosome, col]=sign

#graphics, lo descomento si quiero obtener los gráficos de correlacion

    # print(col, chromosome,stats.pearsonr(compartment,
    #expression)[0])

    # print(sign)

    # plt.figure()

    # sns.regplot(x=compartment, y=expression, data=merged,
    #             line_kws={"color":"red"})

    # plt.show()

#output: tabla con el sentido de la correlacion

```

6. Bloque de código para el cambio de los PC1 según el signo de la correlación entre RNAseq y PC1

```

#input=lista con el sentido de la correlacion

# itero sobre el df de PC1 y modiflico los valores dependiendo
# si la correlación es negativa. Si no lo dejo como está.

tqdm.pandas()

def changeSign(row, col):

    return row[col]*dfSign.loc[row.chr,col]

# EL resultado lo guardo en nuevas columnas como
# "changed_" + el nombre de la columna original

for col in tqdm(PC1cols):

    PC1["changed_"+col]=PC1.apply(changeSign, args=(col,),axis=1)

#output: df modificado

```

7. Bloque de código para la generación de la tabla final

```

#Selecciono solo las columnas que me quiero quedar

PC1ok=PC1[['chr','bin', 'changed_D0 T0 Rep1', 'changed_D0 EtOH 60 Rep1',
            'changed_D0 P60 Rep1', 'changed_D0 R60 Rep1',
            'changed_D3 EtOH 60 Rep1', 'changed_D3 P60 Rep1',
            'changed_D3 R60 Rep1', 'changed_D6 EtOH 60 Rep1',
            'changed_D6 P60 Rep1', 'changed_D6 R60 Rep1',
            'changed_D0 T0 Rep2', 'changed_D0 EtOH 60 Rep2',
            'changed_D0 P60 Rep2', 'changed_D0 R60 Rep2',
            'changed_D3 EtOH 60 Rep2', 'changed_D3 P60 Rep2',
            'changed_D3 R60 Rep2', 'changed_D6 EtOH 60 Rep2',
            'changed_D6 P60 Rep2', 'changed_D6 R60 Rep2']]]

#renombro las columnas (le quito el "changed")

```

```

renameDict = {name:name.split("_")[1] for name in
              PC1ok.columns.tolist() if
              name.startswith('changed')}

PC1ok=PC1ok.rename(renameDict, axis=1)
#Guardo la tabla con los signos del PC1 reasignados
PC1ok.to_csv("PC1_compartments_signed_changed.csv", index=False)

PC1ok

```

6.7 PCA

Los análisis de componentes principales fueron realizados mediante el paquete PCAtools, tomando como input los PC1 de todo el genoma con su signo ya modificado.

7 Resultados

7.1 Verificación de la decidualización

Con el objetivo de analizar la morfología y la organización cromatínica durante la decidualización, se indujo este proceso en células tHESC mediante el tratamiento con estrógeno, AMP cíclico y progesterona o R5020 (Fig. 10a). Se realizaron ensayos de Hi-C y RNAseq inmediatamente antes de la inducción y 60 minutos, 3 y 6 días luego del tratamiento. También se tomaron fotos por microscopía por luz transmitida de las placas, en el tiempo 0 y a todos los días posteriores hasta el día 6 inclusive.

Con el objetivo de corroborar la decidualización se realizaron ensayos de QRT-PCR de los marcadores de decidualización PRL e IGFBP1 y los factores de transcripción PGR y FOXO1 (Fig. 10b,c).

Se observó que los dos marcadores de decidualización (PRL e IGFBP1) aumentaron su expresión a partir del día 3. PGR también aumentó su expresión al día 3, mientras que

FOXO1 presentó un aumento a los 60 minutos después de la inducción.

A partir de estos resultados se puede afirmar que las células tratadas experimentaron un proceso de decidualización, por lo que procedimos a analizar y cuantificar la morfología y la organización cromatínica durante la decidualización a partir de los datos obtenidos mediante las imágenes de luz transmitida y los ensayos de Hi-C, respectivamente.

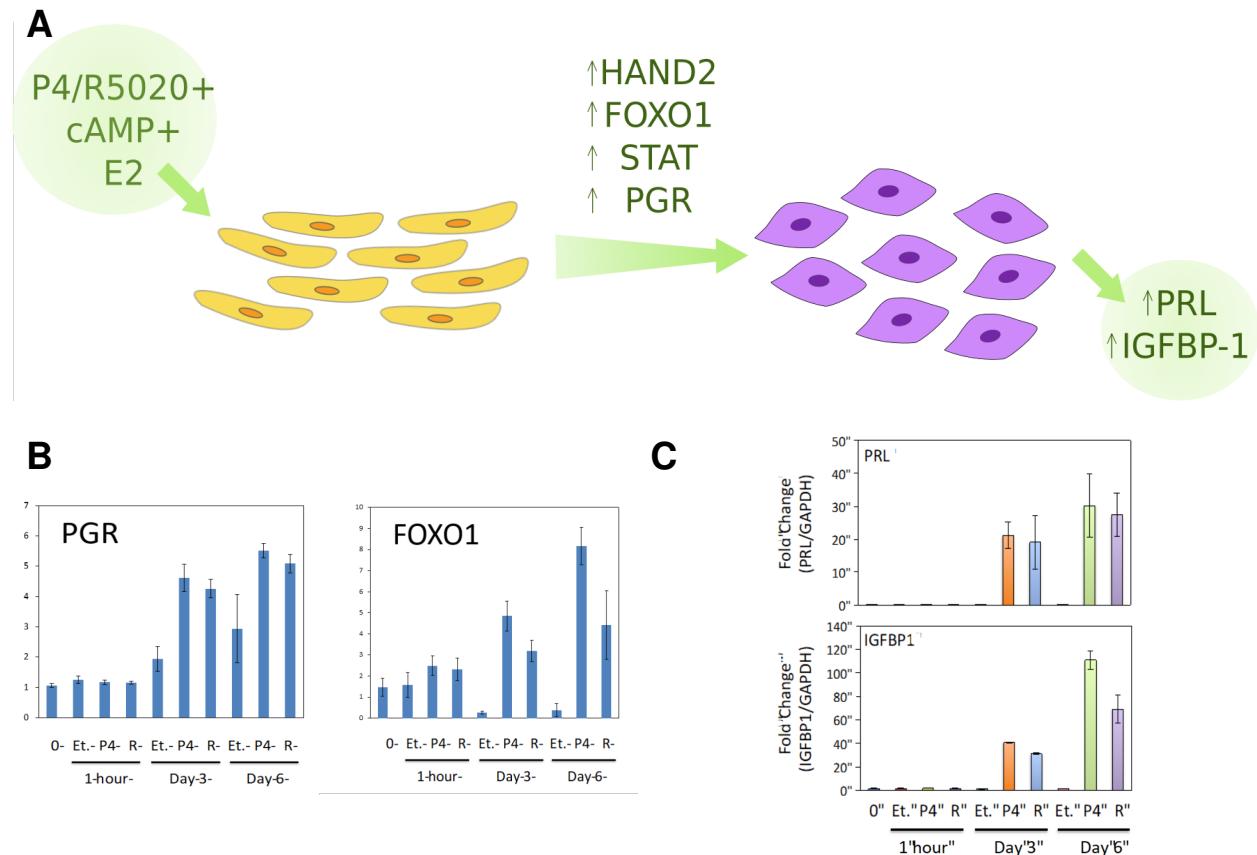


Figura 10: a) Esquema general del sistema de decidualización utilizado y los distintos ensayos realizados. c) Cuantificación de la expresión de los factores de transcripción PGR y FOXO1 para los tratamientos con progesterona (P4), R5020 (R) y vehículo(Et), para el tiempo cero (0), los 60 minutos y los días 3 y 6. d) Cuantificación de la expresión de los marcadores de decidualización PRL y IGFBP1, para los tratamientos con progesterona (P4), R5020 (R) y control (Et), para el tiempo cero (0), los 60 minutos y los días 3 y 6.

7.2 Análisis exploratorio de los cambios morfológicos

A partir de las fotos tomadas durante el experimento, se observó de manera cualitativa un aumento de circularidad en la mayoría de las células tratadas con estímulos decidualizantes al cabo de 6 días, en comparación con las células control (Fig. 11a). Con el objetivo de estudiar estos cambios morfológicos de manera cuantitativa, se decidió aplicar un modelo estadístico del tipo lineal para comparar los cambios en la circularidad que inducen los tratamientos decidualizantes. Los distintos pasos llevados a cabo en este análisis se desarrollan a continuación y se encuentran esquematizados en la figura 11b.

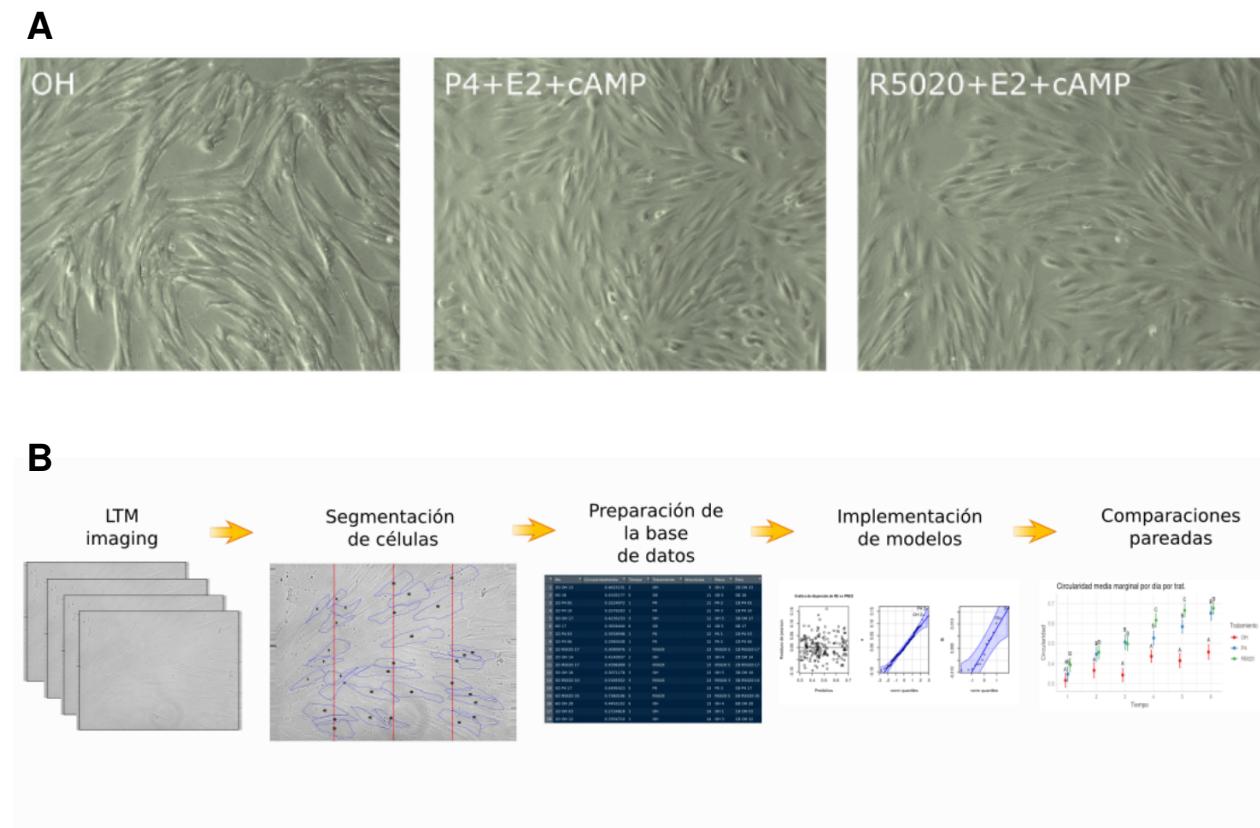


Figura 11: a) Imágenes representativas de las células al día 6, para los distintos tratamientos utilizados b) Esquema de la metodología de trabajo utilizada en esta tesis para analizar cuantitativamente el cambio en la circularidad de las células durante la decidualización. En primer lugar se tomaron fotos de luz transmitida de las células para cada día y tratamiento. Luego se segmentaron las células a través del programa ImageJ y se midió su índice de circularidad. Se preparó la base de datos obtenida de las circularidades y se implementaron y compararon diferentes modelos lineales. Finalmente, a partir del modelo elegido, se compararon las circularidades promedio para los distintos tratamientos.

Luego de la obtención de las medidas de circularidad, se procedió a realizar un análisis exploratorio, para averiguar cómo se comporta esa variable en la población de datos.

En la figura 12a se observa un gráfico de densidades, que indica cuán frecuentes son los distintos valores de circularidad para cada variable. Las células bajo tratamiento decidualizante mostraron un corrimiento del pico de densidades hacia circularidades mayores a medida que transcurren los días; lo que no sucede en la población de células tratadas con el vehículo como control. En la misma figura 12a se observa que las curvas se comportan de manera similar a una normal. Aunque esta apreciación es meramente cualitativa, este dato nos dio pie a utilizar modelos para distribuciones normales, a pesar de que la variable no es estrictamente normal.

Comparando los cuartiles mediante un gráfico de cajas (Fig. 12b) también se observa un claro aumento de la circularidad para las células tratadas. El hecho de que la circularidad para las células tratadas con etanol quede prácticamente fija nos condujo a trabajar con un modelo de comparación de medias, ya que no parecería haber una relación lineal entre la variable respuesta (circularidad media por foto) y las explicativas (tiempo y tratamiento). Además a partir del gráfico de cajas se observa que el efecto del tratamiento sobre la circularidad varía según el día y viceversa, indicando una interacción entre las variables. Por ejemplo, el efecto de R5020 sobre la circularidad no parecería ser igual para todos los días, así como el efecto del día 4 sobre la circularidad varía según si las células fueron tratadas con P4, EtOH o R5020.

Por último realizamos un Spaghetti plot para corroborar si alguna placa presentaba una respuesta atípica frente al tratamiento, ya que la presencia de datos atípicos podría influenciar los resultados. En la figura 12c se observa que ninguna placa presenta un comportamiento muy diferente a las otras bajo el mismo día y tratamiento.

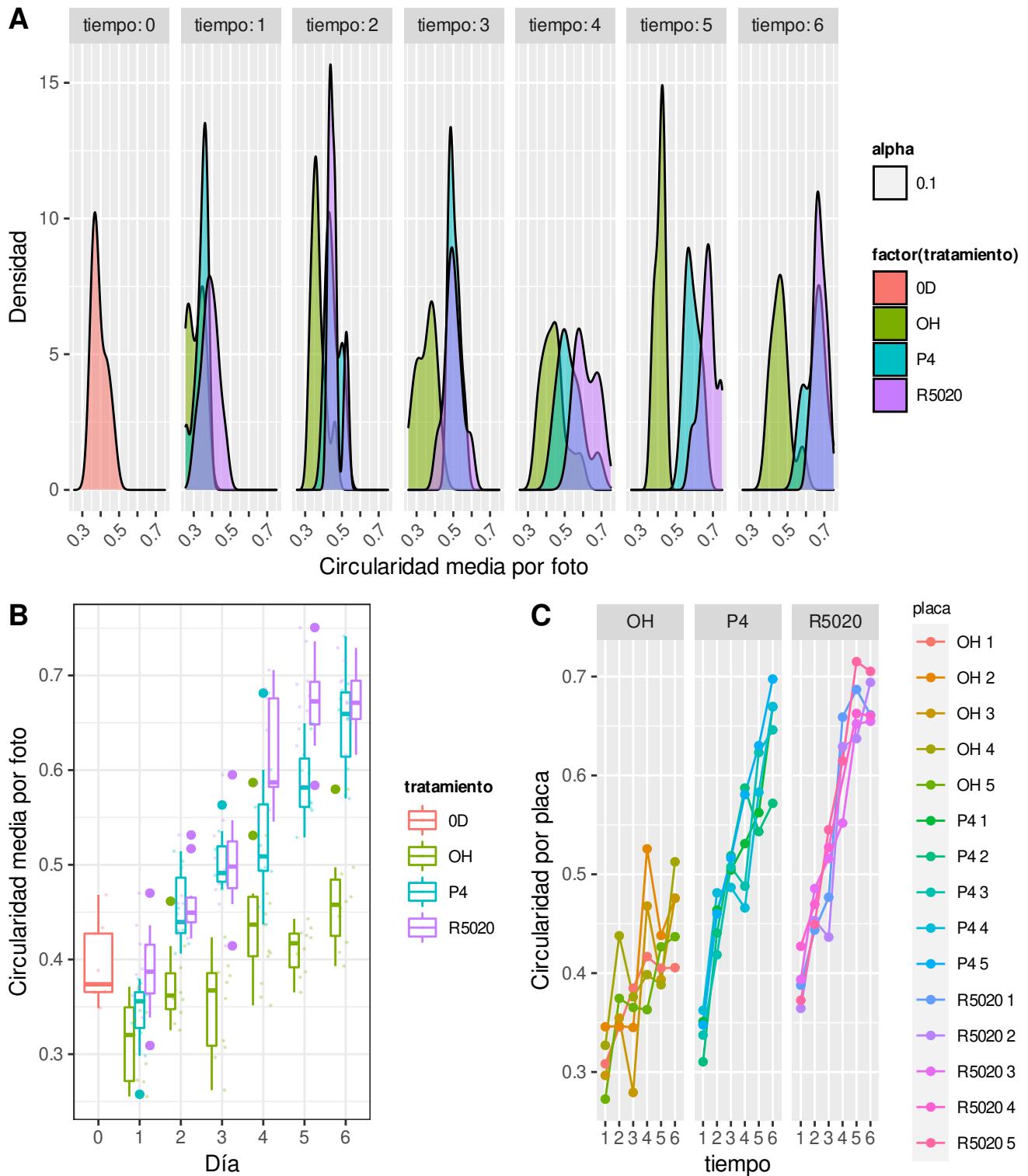


Figura 12: a) Gráfico de densidad de la circularidad indicando cuán frecuentes son las distintas circularidades para cada día y tratamiento, donde el pico de cada distribución indica la circularidad por foto más frecuente . b) Boxplot indicando los cuartiles para la circularidad por tratamiento y día. c) Spaghetti plot indicando el promedio de la circularidad por placa según día y tratamiento.

7.3 Selección de Modelos estadísticos

Debido a que el análisis exploratorio mostró un claro aumento de la circularidad celular durante la decidualización, se procedió a determinar qué modelo lineal sería el más adecuado para modelar estos datos, y definir si estas diferencias en la circularidad efectivamente eran significativas.

Debido a que se midieron las mismas placas desde el día 1 al día 6, se planteó aplicar un modelo de medidas repetidas, que tenga en cuenta el efecto del tratamiento y del día (variables explicativas) y la interacción entre ellas. Adicionalmente, debido a que la circularidad de las células de una misma placa en distintos días no son independientes entre sí, el modelo debía incluir esta falta de independencia. Esto puede hacerse de manera implícita, añadiendo una variable aleatoria al modelo que corresponda a la placa, en cuyo caso se denomina modelo condicional, o bien explícitamente añadiendo una matriz de covarianza, pudiendo elegir distintos patrones para esta, siendo entonces un modelo marginal. Para analizar si la correlación de la circularidad para la misma placa entre los distintos días presentaba un patrón particular, realizamos una matriz de correlación para los distintos días (Fig. 13).

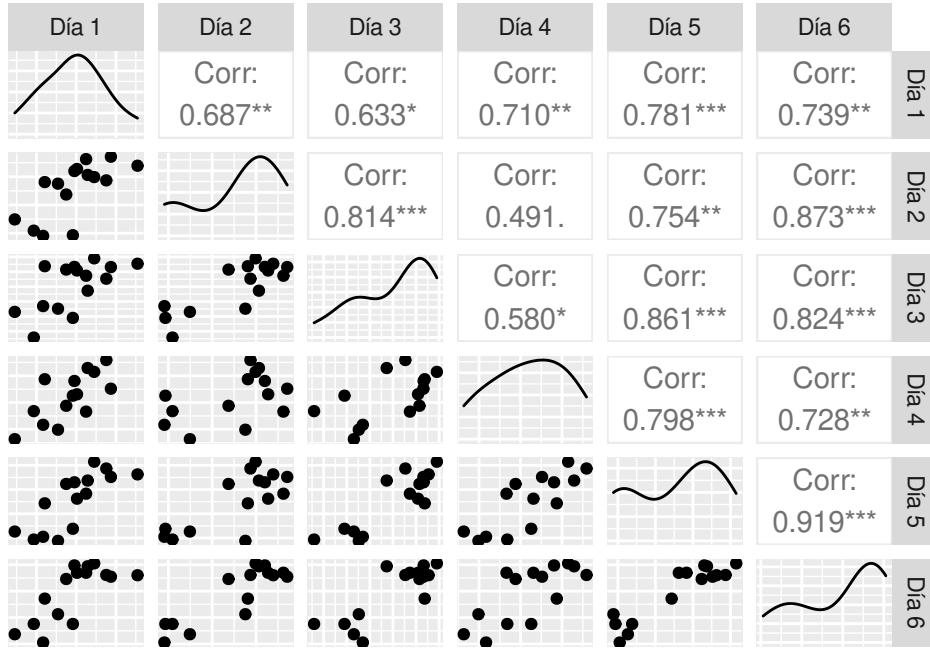


Figura 13: Matriz de correlación entre las medias de circularidad por placa para los diferentes días. La esquina superior indica el coeficiente de correlación, mientras que la esquina inferior muestra el gráfico de puntos correspondiente. En la diagonal se encuentra la distribución de la circularidad media por foto.

En la figura 13 no se reconoce de manera clara ningún patrón de correlación de los explicados anteriormente (Fig.4), donde por ejemplo todas las correlaciones sean iguales (matriz de correlación compuestas) o decaigan a mayor distancia entre los días (matriz de correlación regresiva de orden 1).

Debido a que no era evidente el patrón de correlación entre los días, se probaron modelos marginales con todas las matrices de covarianza posibles. También se probó un modelo condicional con el efecto de la placa como una variable aleatoria.

Se midió y comparó el desempeño de los modelos a partir del cuadrado medio del error (CMe) y su score en el criterio de información de Akaike (AIC) (Tabla 1). El cuadrado medio del error es un parámetro que varía según el promedio de los errores, es un indicador de cuánto se alejan los valores observados a la media predicha por el modelo. Un modelo con mejor desempeño presentará un menor CMe. El AIC por su parte, es una medida de la calidad relativa de un modelo estadístico, para un conjunto dado de datos.

Sigue la fórmula $AIC = 2k - 2 \ln(L)$, donde k es el número de parámetros del modelo y

L es el máximo valor de la función de verosimilitud para el modelo. El índice de AIC es menor cuanto mejor se ajustan los datos al modelo (bondad de ajuste) pero penalizando a su vez la complejidad del modelo (aumenta cuantos más parámetros tenga el modelo). Cuanto menor sea el AIC del modelo mejor será el balance entre bondad de ajuste y complejidad del modelo.

Tabla 1: Comparación de modelos

	modelo	CMe	gl	AIC
m1	condicional	0.0479673	20	-453.3368
m2	compuesto	0.0517463	20	-452.2906
m3	AR1	0.0522599	20	-453.4576

* CMe: cuadrado medio del error, gl: grados de

libertad, AIC:criterio de información de Akaike

Todos los modelos mostraron un desempeño similar, sin embargo, se decidió utilizar el modelo con menor AIC y CMe; el modelo condicional con variable aleatoria. El modelo final se puede expresar mediante la siguiente fórmula:

$$E(Circularidad) = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + B_{k(i)}$$

i=[1:3]

j=[1:6]

k=[1:5]

y suponiendo:

$$E_{ijk} \sim N(0, \sigma^2)$$

$$B_k \sim N(0, \sigma_{plate}^2)$$

Donde α representa el efecto por tratamiento, β el efecto del día, $\alpha\beta$ la interacción entre ellos y B el efecto aleatorio de la placa.

En cuanto al tiempo cero, se tomaron fotos de 5 placas previo al tratamiento. Estas 5 placas fueron inmediatamente utilizadas para ensayos de Hi-C y RNAseq correspondientes a ese tiempo, por lo que del día 1 al 6 se fotografiaron otras 5 placas. Por esto no fue posible incluir las placas correspondientes al tiempo cero al modelo de medidas repetidas, ya que no compartían tratamiento, día, ni placa con las otras muestras.

Esto nos obligó a aplicar otro modelo adicional, para analizar si hubo un cambio en la circularidad entre el día 1 de cada tratamiento y el tiempo cero (en adelante modelo T0).

Este modelo de comparación de medias solo incluye una variable explicativa que llamamos tratamiento, que consta de los niveles: tiempo cero (T0), día 1 R5020 (R5020), día 1 Progesterona (P4) y día 1 control (EtOH).

El modelo entonces sigue la ecuación:

$$E(Circualridad) = \mu + \alpha_i + B_{k(i)}$$

$$i=[1:4]$$

$$k=[1:5]$$

y suponiendo:

$$E_i \sim N(0, \sigma^2)$$

$$B_k \sim N(0, \sigma_{plate}^2)$$

Donde α es el efecto del día y/o tratamiento (T0; Día 1 R5020; Día 1 EtOH; Día 1 P4) y B el efecto aleatorio dado por la placa.

Se comprobaron los supuestos para todos los modelos. En la figura Fig. 14 se muestran

únicamente los gráficos que comprueban los supuestos de los modelos elegidos (Fig. 14). Se grafican, de izquierda a derecha, los residuos vs los valores predichos, los cuantiles de los residuos estandarizados vs los cuantiles esperados para una distribución normal (QQplot), y los cuantiles de los efectos aleatorios estimados (el efecto de cada placa) vs los cuantiles esperados para una distribución normal (QQplot).

El gráfico de residuos (diferencia entre los valores predichos por el modelo y los observados) vs predichos (los valores de circularidad predichos por el modelo) no presenta ningún patrón aparente, por lo que no hay indicios de falta de ajuste a una distribución normal o de la presencia de outliers. Los QQPlots indican que tanto la circularidad media por foto como los efectos aleatorios estimados se ajustan a una distribución normal.

Debido a que los cuantiles observados para la circularidad y los teóricos para una distribución normal no difieren significativamente, es que la regresión presenta una pendiente que no difiere significativamente de 1.

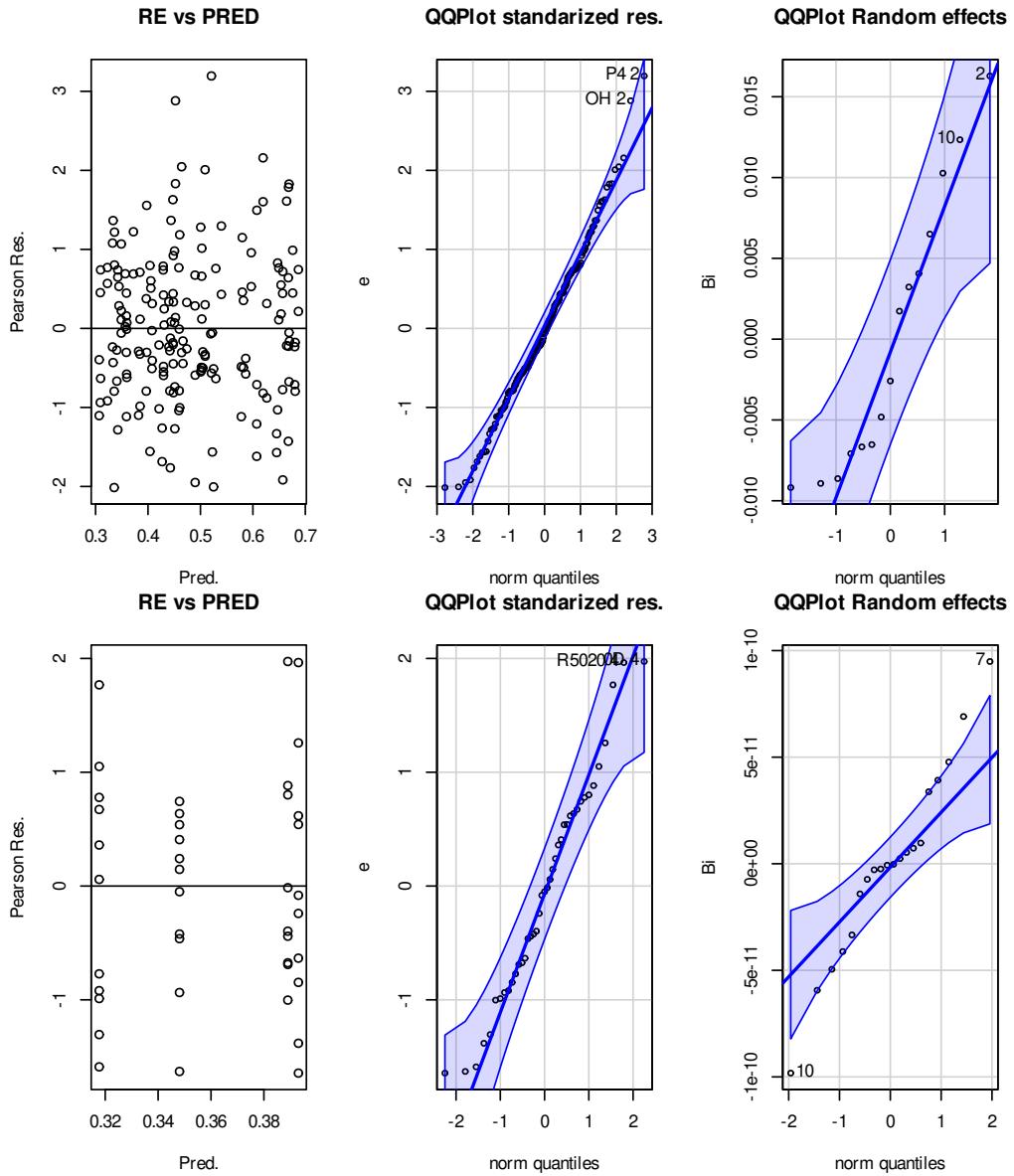


Figura 14: Verificación gráfica del cumplimiento de los supuestos del modelo para el modelo de medidas repetidas (superior) y el modelo que incluye el tiempo cero y el día uno para los distintos tratamientos (inferior). Se grafica, de izquierda a derecha, los residuos vs los predichos, el QQplot para los residuos estandarizados y el QQplot para los efectos aleatorios. El área azul indica el intervalo de confianza y la línea central la correlación. Los test de Shapiro-Wilks tampoco mostraron evidencia de falta de normalidad (p valor > 0.05) para ninguno de los dos modelos.

Luego, para corroborar si las distintas variables explicativas (tiempo y tratamiento) tienen un efecto sobre la circularidad celular, realizamos un Anova para ambos modelos.

El modelo de medidas repetidas (que incluye del día 1 al 6) arrojó un p valor menor a 0.05 tanto para las variables tiempo y tratamiento como para sus interacciones. De esta manera

se puede decir que la circularidad se ve afectada de manera diferencial dependiendo la combinación específica de tratamiento y día en la que se encuentra.

Para el modelo T0, el efecto de la variable llamada “tratamiento” (cuyos niveles son T0, 1 día P4, 1 día R5020 y 1 día EtOH) sobre la circularidad también es significativa, por lo que al menos una de las circularidades difiere del resto.

Tabla 2: Anova para el modelo condicional (m1)

	numDF	denDF	F-value	p-value
(Intercept)	1	153	10484.591018	0
tratamiento	2	12	112.311464	0
tiempo	5	153	124.075519	0
tratamiento:tiempo	10	153	7.831016	0

* Los parametros con un p-valor menor a 0,05 se

consideran significativos

Tabla 3: Anova para el modelo T0 y Día 1 (mT0)

	numDF	denDF	F-value	p-value
(Intercept)	1	21	3323.645461	0.0000000
tratamiento	3	16	8.343365	0.0014424

* Los parametros con un p-valor menor a 0,05 se

consideran significativos

7.4 Efecto de los tratamientos sobre la circularidad celular

Una vez determinado que el tratamiento y el día tienen un efecto sobre la circularidad, se procedió a realizar una comparación de efectos simples para comparar estadísticamente las medias de la circularidad. La comparación de efectos simples permite realizar una menor cantidad de comparaciones, ya que compara las medias para los diferentes tratamientos de una variable, dejando la otra variable fija.

En este caso se compara, dentro de cada día, las medias entre los distintos tratamientos. De esta manera, se realizaron 3 comparaciones para cada día, es decir 18 comparaciones en total, mientras que si se hubiera hecho una comparación de todos contra todos se hubieran realizado 153 comparaciones. Como la potencia de las comparaciones disminuye cuantas más comparaciones se realicen (aumenta la posibilidad de falso negativo), las comparaciones de efectos simples presentan una mayor potencia que si se realiza una comparación del tipo todos contra todos. Sin embargo, no permite comparar entre distintos niveles del factor que fue dejado fijo, en este caso no se podrían sacar conclusiones acerca de si la circularidad para un tratamiento aumenta de un día a otro.

En la figura 15 se observan los resultados de las comparaciones de entre tratamientos. Las circularidades para los tratamientos decidualizantes no difieren del día cero. En cuanto al tratamiento control EtOH, sí es significativamente menor su circularidad respecto al tiempo cero (Fig.15a). Es por esto que el efecto del “tratamiento” arrojaba un p-valor significativo en el análisis de Anova (tabla 3). A pesar de eso, la circularidad de EtOH al día 1 no difiere de los otros tratamientos al día 1 (P4 y R5020). A su vez, P4 y R5020 al día 1 no difieren del día cero. Además, a pesar de no ser significativa, se nota una disminución también en el tratamiento con P4, por lo que esta disminución bien podría deberse a un efecto del etanol, o consecuencia de una mayor adherencia de las células a la placa, que habían sido sembradas 24 hs previo al comienzo del experimento. Se podría decir entonces que no hay cambios biológicamente relevantes en la morfología entre el día 1 y el día 0 para los estímulos decidualizantes.

En cuanto a la comparación para los días 1 a 6, se pueden diferenciar dos etapas (Fig.15b). Hasta el día 3 ambos tratamientos actúan de igual manera, aumentando su circularidad significativamente en comparación al control a partir del día 2. Luego, en el día 4, el tratamiento con R5020 presenta aumento, de la circularidad celular respecto de las células tratadas con P4. Esta diferencia se mantiene en el día 5, y recién al día 6 ambos tratamientos decidualizantes vuelven a igualarse (Fig.15b).

Se puede concluir que, a pesar de que ambos tratamientos presentan una circularización de su morfología en respuesta a la transdiferenciación causada por el estímulo decidualizante, su dinámica es distinta.

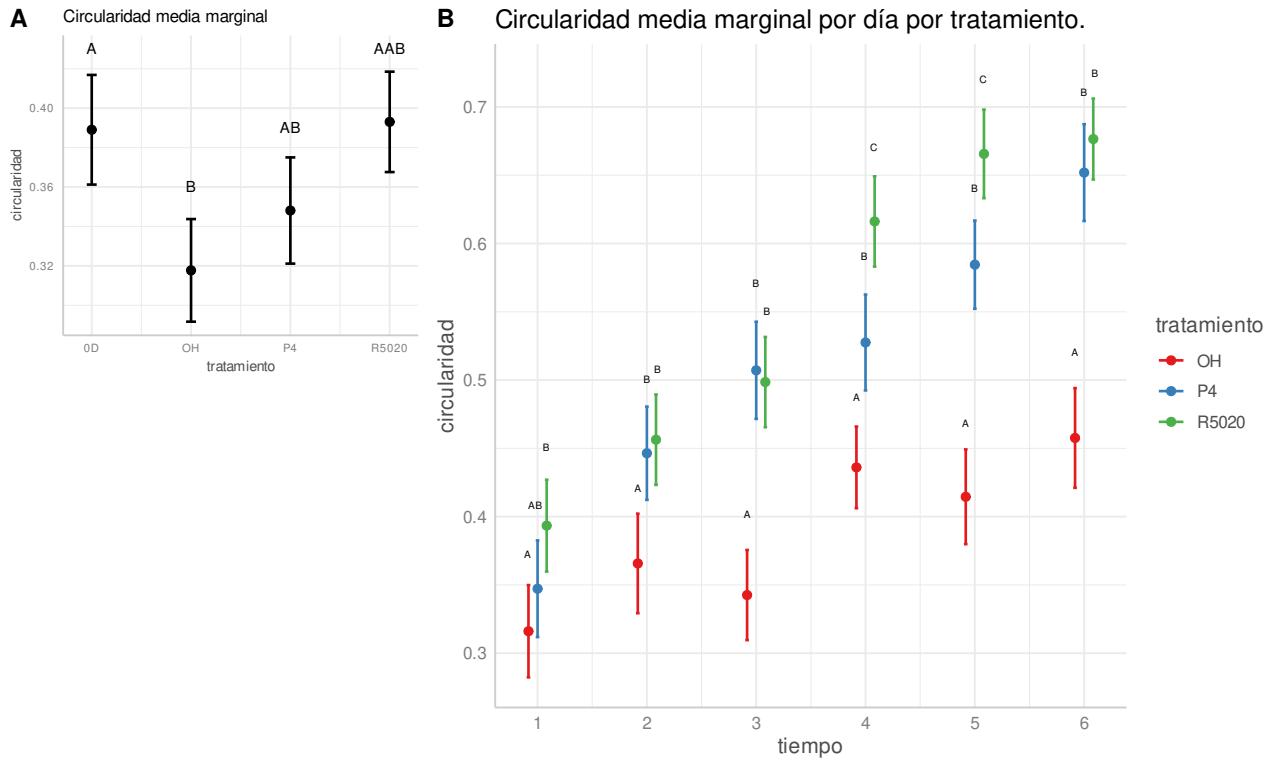


Figura 15: a) circularidad media para los distintos tratamientos para el día 1 y el tiempo cero. b) circularidad media marginal para los distintos tratamientos por día. Ambos gráficos fueron realizados mediante la función `ggemmeans` del paquete `ggeffects`, con un intervalo de confianza del 95 %. Letras diferentes indican diferencias significativas entre las medias de circularidad para los diferentes días dado un tratamiento.

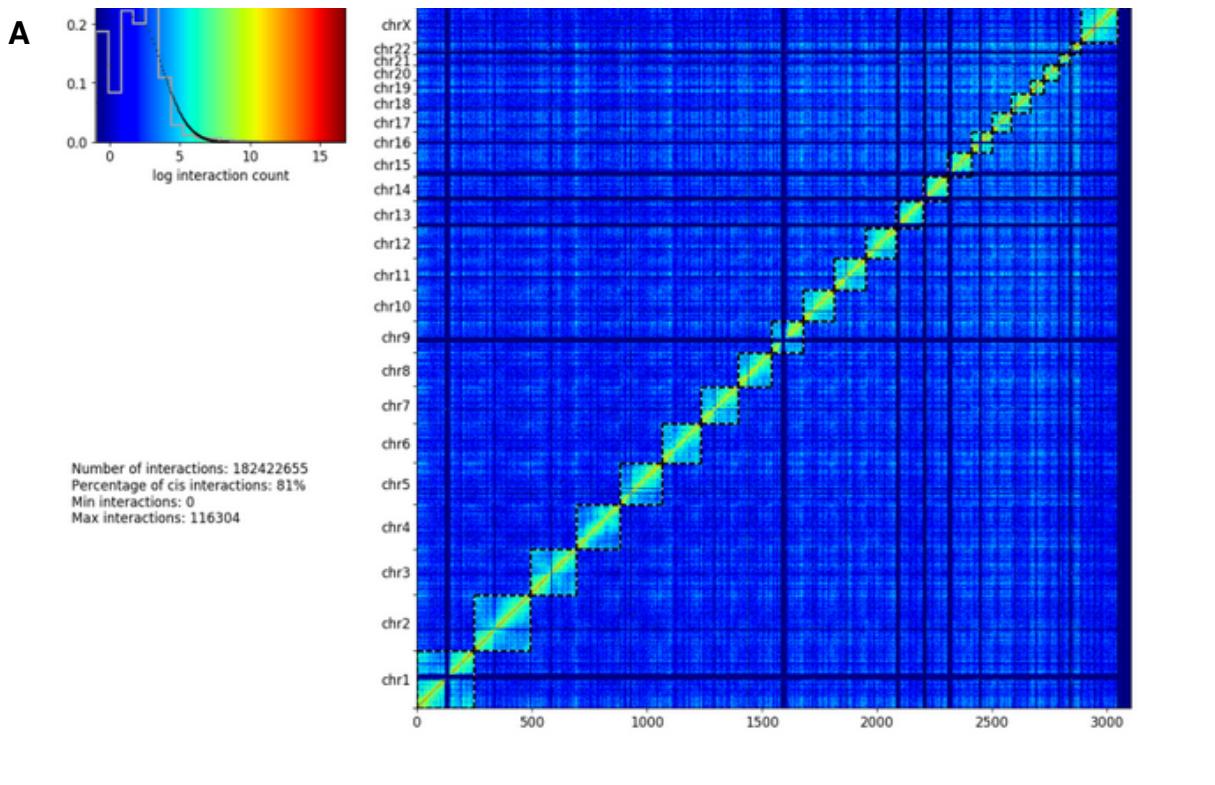
7.5 Análisis preliminar de la topología génica durante la decidualización

Paralelamente, con el objetivo de evaluar la organización de la cromatina, se realizaron ensayos de Hi-C obtenidos en el tiempo cero, a los 60 minutos, y a los 3 y 6 días de tratamiento con ambos tratamientos y con el tratamiento control.

Decidimos realizar estos ensayos a tiempos cortos y a tiempos largos ya que nos interesa saber si es necesaria una reorganización previa al cambio de fenotipo, para luego estudiar más específicamente dichos cambios. Dado el alto costo de estos experimentos, realizamos Hi-C a un tiempo muy corto (1 hora) que permitiría estudiar el disparo de la señal, y a los dos tiempos en los cuales el fenotipo decidual ya

se ha establecido (día 3 y 6), según lo visto mediante la cuantificación de la morfología y la expresión de genes marcadores (PRL e IGFBP1).

El paradigma actual ubica a cada cromosoma dentro del núcleo en una región relativamente fija, llamada territorio cromosómico. Es por esto que en las técnicas de Hi-C se espera que las interacciones intercromosómicas sean bajas (aunque no nulas) en relación a las intracromosómicas. Un análisis de calidad estándar para este tipo de ensayo incluye el análisis de la relación entre estos dos tipos de interacciones. En la figura 17a se muestra un ejemplo que presenta una correcta relación de sus interacciones inter e intra cromosómicas. Las interacciones de cada cromosoma se visualizan en los cuadrados celestes resultantes de las probabilidades más altas de su interacciones intra cromosómicas, contrastando con las regiones azules intercromosómicas, de menor frecuencia de contacto. Además, para cada cromosoma se pueden notar dos regiones, aisladas entre sí por los centrómeros. En la figura 17b se muestran los ratios entre el número de contactos intracromosómicos (Cis) e intercromosómicos (Trans) para cada muestra. La mayoría de las muestras presentan una relación entre los contactos Cis y Trans cromosómicos dentro del rango aceptado para este parámetro (celdas blancas). A pesar de que 4 muestras, coloreadas en rosado, presentan un ratio un poco más bajo, son consideradas aceptables por la facility que realizó el preprocesamiento de los datos. La réplica 1 para el día 6 tratado con P4 sin embargo, presenta un ratio de contactos cis/trans de tan solo 15. Esto podría indicar que hubo una falla durante el procedimiento, y que los resultados que se obtuvieron no son confiables. Es por esto que no tuvimos en cuenta esta muestra en los análisis posteriores.



B

Tratamiento	Cis/Trans Rep1	Cis/Trans Rep2
Day 0	81	82
Day 0 EtOH	83	85
Day 0 P4	84	80
Day 0 R5020	84	80
Day 3 EtOH	46	83
Day 3 P4	60	84
Day 3 R5020	81	82
Day 6 EtOH	48	83
Day 6 P4	15	75
Day 6 R5020	41	77

Figura 16: a) Mapa de calor de todo el genoma para la muestra correspondiente a la réplica 1 del T0. El panel izquierdo determina el color para cada número de contactos, siendo azules para interacciones bajas y rojas para interacciones altas. La curva gris indica la frecuencia de los números de contactos. Los gráficos superiores corresponden a los valores de los primeros 3 componentes principales: Eigenvalue 1 (E1), Eigenvalue 2 (E2) y Eigenvalue 3 (E3). Se indica también el número de interacciones, el porcentaje de interacciones en cis y el mínimo y máximo de interacciones que presentan los bins. b) Tabla indicando la relación entre el número de contactos intracromosómicos y intercromosómicos. Las muestras que presentan un valor moderadamente bajo se indican con rojo claro, mientras que las muestras con un valor críticamente bajo se indican en color rojo.

Luego de corroborar la calidad de los datos, se analizó la compartmentalización global partiendo de los PC1 obtenidos a partir de las matrices de correlación. El valor del PC1 indica el grado de apertura de la cromatina como se ha explicado anteriormente (ver

Materiales y Métodos), asignando el mismo signo de PC1 (positivo o negativo de manera azarosa) para todas las zonas abiertas de la cromatina, y asignando el signo contrario a las zonas cerradas.

Debido a que el PCA para obtener el PC1 se realiza para cada cromosoma por separado, el signo del PC1 para los compartimentos A y B podría variar para cada cromosoma. Se procedió primero a unificar el criterio, cambiando los signos de ser necesario para que todos los compartimentos A presentaran signo positivo y los B signo negativo. Teniendo en cuenta que las regiones abiertas presentan una expresión génica más activa, se calculó para cada cromosoma el coeficiente de correlación entre la transcripción y el PC1 de las distintas zonas (Fig.16a). Si este coeficiente resultaba negativo, se reescribían los PC1 manteniendo su valor pero cambiando su signo (Fig.16b).

Luego, fue posible clasificar de manera global los compartimentos que variaron de un estado al otro durante la transdiferenciación, ya que los compartimentos que pasaban de B a A cambiaban de un PC1 negativo a uno positivo, y lo contrario ocurría con los compartimentos que pasaban de A a B.

Se realizó un PCA de los PC1 modificados, con el objetivo de evaluar el cambio global en la compartmentalización. Observando los Componentes principales, se puede decir que la mayor parte de la variación está dada por el PC1 (Fig.16c), por lo que un gráfico de PC1 vs PC2 resulta una buena manera de observar gráficamente la variación global (Fig.16d).

Al observar el análisis de componentes principales, es claro que la compartmentalización global cambia durante la decidualización, ya que las réplicas se agrupan por tratamiento y por día. Este resultado indica que el perfil global de compartmentalización varía entre el día cero y el día 3, y nuevamente entre el día 3 y el 6. Se puede concluir entonces que la compartmentalización A/B, durante la decidualización, transita por al menos dos estados diferentes.

La compartmentalización a los 60 minutos, para los tratamientos decidualizantes y para el tratamiento control, se agrupan junto al tiempo cero, lo que indica que sus perfiles de compartmentalización son similares. Esto no implica que no ocurran cambios en el estado

de la cromatina a la hora, sino que estos no son suficientes como para generar un cambio a nivel global.

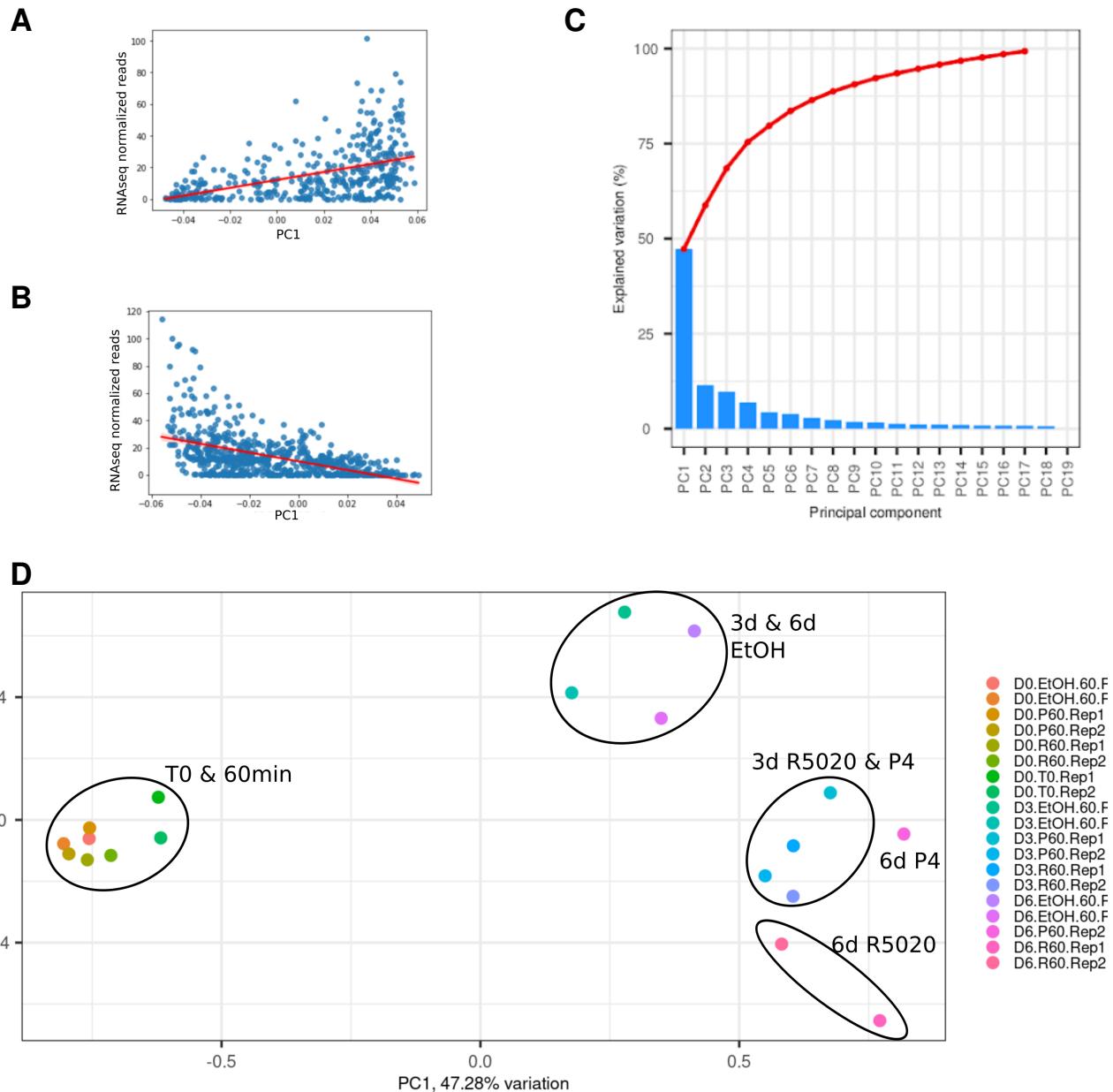


Figura 17: a) Scatter plot de los valores de PC1 vs las lecturas de RNAseq normalizadas, para cada región (100kb) del cromosoma 16 de la réplica 1 del tratamiento con R5020 al día 3. b) Scatter plot de PC1 vs lecturas normalizadas, para cada región (100kb) del cromosoma 16 de la réplica 1 del tratamiento con EtOH a los 60 minutos. c) Scree plot detallando cuánta variación es explicada por los distintos componentes principales del PCA de compartmentalización. Las barras indican el porcentaje de la varianza de los perfiles de compartmentalización que explica cada componente principal. La curva roja indica varianza explicada acumulada con cada componente principal. d) PCA de los perfiles de compartmentalización.

7.6 Relación entre los cambios morfológicos y de compartmentalización

Con el objetivo de analizar si los cambios morfológicos se dan de manera concertada respecto a los de compartmentalización, se realizó un scatter plot (Fig.18) indicando los correspondientes PC1 para cada muestra provenientes del PCA (Fig.16d) y sus respectivas circularidades predichas por el modelo (fig. 15). La figura (Fig.18) muestra la correlación entre estas dos features. Esta regresión sin embargo es a modo cualitativo, ya que consideramos que no son suficientes los puntos para analizar la relación de manera cuantitativa.

Se puede observar que para los dos tratamientos decidualizantes la relación entre circularidad y PC1 es directa. Para el EtOH el coeficiente de correlación es de sólo 0,52, por lo que es menos probable una asociación.

Para los tratamientos decidualizantes se observa una diferencia tanto para la circularidad como para la compartmentalización entre el tiempo cero y los 3 días. Entre el día 3 y el 6, el cambio en el PC1 es menor respecto al cambio de PC1 entre el día 0 y 3, mientras que en la circularidad la diferencia entre el día 3 y 6 es similar a la que ocurre entre el día 0 y 3. Esto indica que el cambio más grande en la compartmentalización a nivel global se da durante los primeros 3 días de la decidualización, mientras que los cambios en la circularidad parecerían estar distribuidos más equitativamente a lo largo de los días. Sin embargo, los cambios en la compartmentalización entre el día 3 y 6 son suficientes para agruparlos de manera separada en el análisis de PCA.

Se puede concluir que las células t-HESC durante la decidualización bajo este modelo presentan un perfil distintivo para los diferentes días (tiempo cero, día 3 y día 6) tanto en su circularidad como en su perfil de compartmentalización, que permite diferenciarlos claramente entre sí.

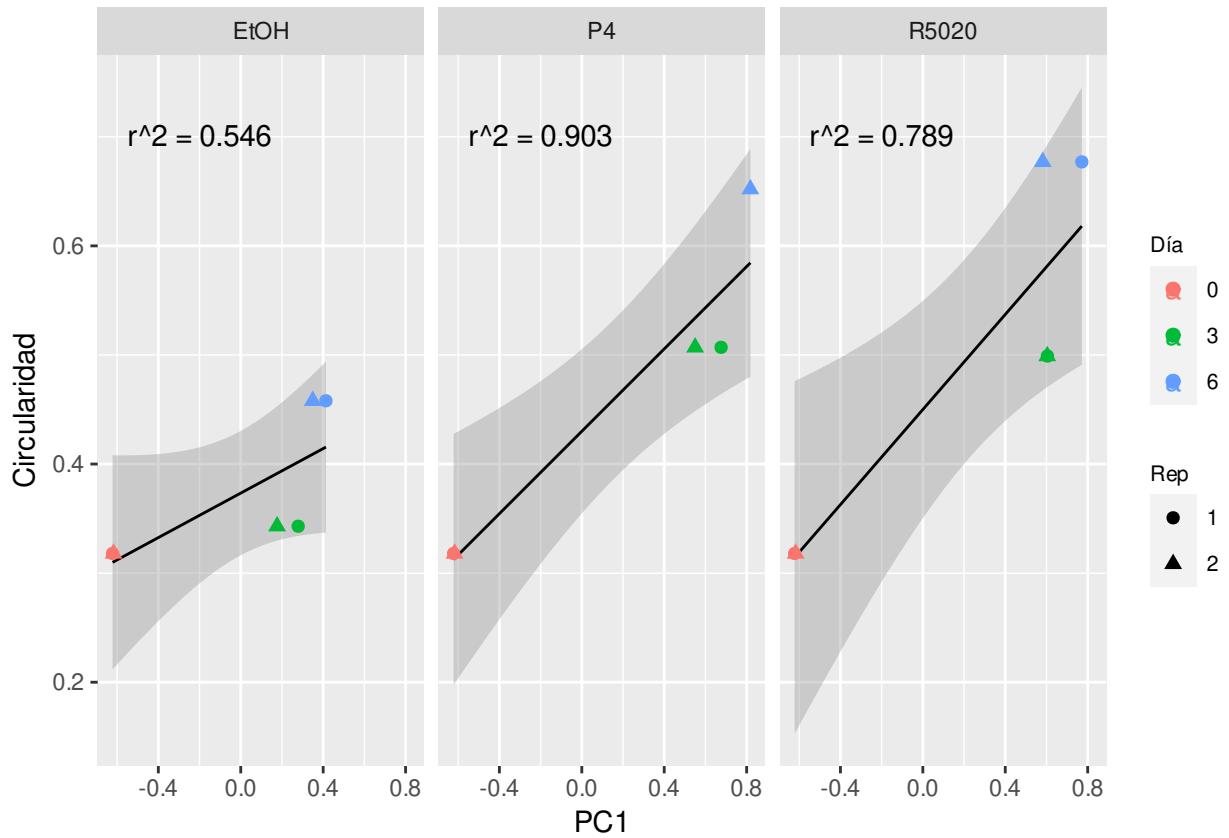


Figura 18: Gráfico de dispersión para los diferentes tratamientos. A ambas réplicas para cada día y tratamiento del protocolo de Hi-C se le asignó la circularidad media predicha por el modelo lineal. En los 3 gráficos se utilizaron los mismos datos para el tiempo cero. La línea negra denota la regresión lineal y el área gris el intervalo de confianza del 95%. En la parte superior se encuentra el coeficiente de correlación de Pearson.

8 Discusión y Conclusiones

Este trabajo presenta una metodología para cuantificar los cambios morfológicos que ocurren durante la decidualización *in vitro* y su asociación con la reorganización global de la cromatina.

En primer lugar se caracterizaron las condiciones del modelo experimental corroborando la inducción de la decidualización por los estímulos hormonales P4 y la progestina sintética R5020 en presencia de AMPc y E2. La diferencia entre ambos radica en que R5020 no se metaboliza dentro de la célula, a diferencia de P4. Utilizar los dos tipos de estímulo nos permitió diferenciar el efecto de las progestinas en sí de los efectos que pudieran tener sus metabolitos. Se comprobó que ambas progestinas son capaces de producir la decidualización de las células t-HESC. Sin embargo, se observaron diferencias en la dinámica de esta decidualización, que podrían ser atribuidas a los productos de la degradación de P4. Estas diferencias también se podrían adjudicar a una disminución de la progesterona con el correr del tiempo de cultivo, mientras que R5020, al no degradarse, mantendría su concentración estable. Justamente, previniendo esto, es que se utilizaron mayores concentraciones de P4.

El avance de los cambios morfométricos durante la decidualización mostró que las células presentaron un aumento significativo de la circularidad a tan sólo 2 días luego de la inducción, tanto para las células tratadas con P4 como R5020. Al día 4, la morfología para ambos tratamientos diverge, ya que la circularidad de las células tratadas con R5020 aumenta significativamente en comparación con P4. Al día 5 esta diferencia se mantiene y recién al día 6 la circularidad de ambos tratamientos vuelve a igualarse.

Estos resultados nos muestran que el tratamiento con R5020 presenta un cambio de manera anticipada respecto al tratamiento con P4.

In vivo, los cambios morfológicos asociados a la decidualización suceden 9 días después de la ovulación, momento en el cual comienzan a aumentar los niveles de progesterona circulantes. Esta diferencia temporal entre el aumento de los niveles de progesterona y los

cambios morfológicos indica que serían otras las señales que se requieren para desencadenar la decidualización (Gellersen & Brosens 2003). En este sentido, los cultivos primarios de células endometriales requieren de 7 a 10 días de tratamiento sólo con progesterona para lograr la inducción de la expresión de PRL. Por otro lado, la producción de PRL en explantes de endometrio en la fase media secretora se detiene luego de 3 a 4 días, pero puede ser restaurada mediante el tratamiento con progesterona (Maslar & Ansbacher 1986). La progesterona entonces cumpliría un rol en el mantenimiento del fenotipo decidualizado, pero es poco probable que sea el estímulo desencadenante por sí solo. Es cada vez más evidente que la señal capaz de desencadenar la decidualización debe ser aquella que aumente los niveles de AMPc intracelulares (Gellersen & Brosens 2014).

In vitro esta señal se saltea agregando AMPc capaz de permear dentro de la célula. Se podría pensar entonces que la discrepancia en las dinámicas de cambio de la morfología se observan recién al día 4 debido a que el desencadenamiento y primeras etapas de la decidualización estarían gobernadas más bien por AMPc. Las progestinas mantendrían el fenotipo decidualizado una vez que la decidualización ha sido desencadenada. Se podría esperar que unos días después del tratamiento, cuando las progestinas son las encargadas de mantener el fenotipo decidual, la progesterona presente niveles más bajos que R5020, ya que durante los primeros días pudo haber sido metabolizada. Es por esta razón que en el protocolo utilizado en este trabajo se colocaron concentraciones mayores de P4 (10-6M) que de R5020 (10-8M). Sin embargo, no podemos descartar que el cambio más lento de morfología que se observa con la progesterona pueda deberse a una menor concentración de esta hormona, en comparación con R5020.

Aún considerando una disminución de la concentración de progesterona en el tiempo, el efecto no se asemejaría a lo observado cuando se retira la progesterona del medio (Yu et al. 2016). Yu *et al* (Yu et al. 2016) registraron los cambios al tratar cultivos primarios de HESC con P4, E2 y AMPc, y luego de retirar las hormonas del medio por dos semanas (Yu et al. 2016). En dicho trabajo observaron una regresión del fenotipo decidual al fibroblástico al retirar el tratamiento. Esta regresión se evidenció tanto de manera

morfológica como en la expresión de marcadores de deciduallización , indicando una plasticidad fenotípica en las células ya deciduallizadas in vitro. En nuestro modelo no se observa tal dinámica, ya que ambas morfologías al día 6 no difieren entre sí, y sí difieren del control al día 6. La diferencia observada en progesterona entonces no se correspondería a un retroceso del fenotipo, sino a una ralentización en la transdiferenciación.

En cuanto a la metodología utilizada para el análisis morfológico, esta se podría mejorar en futuros ensayos, tomando fotos de las mismas placas también en tiempo cero, para así tener en cuenta esta variable aleatoria durante toda la secuencia temporal. Además, se podría agilizar el proceso de medición de la circularidad segmentando las células de manera automática por algoritmos de deep learning utilizando en conjunto los programas Cellpose (Stringer et al. 2021) y LabelsToRois (Waisman et al. 2021). Actualmente CellPose no reconoce las células de esta línea celular mediante imágenes de luz transmitida, posiblemente por ser estas muy planas. Sería necesario marcar las células de manera fluorescente, con un marcador citoplasmático por ejemplo, para que de esta manera sean reconocidas por el programa. Luego mediante LabelsToRois se podría convertir los resultados del CellPose de manera que puedan ser tomados por el programa ImageJ, y luego continuar con la medición de la circularidad como se describió en este trabajo.

En los análisis de la organización de la cromatina, se corroboró la calidad de los datos provenientes de los ensayos de Hi-C. Se obtuvieron los perfiles de interacción, mediante los cuales es posible conocer la frecuencia de interacción entre loci distantes. Además, a partir de estos perfiles se pudo determinar los perfiles de compartmentalización, mediante el cual se infiere el estado de la cromatina en los distintos tiempos. A partir de estos resultados se realizarán numerosos análisis para tratar de elucidar cómo es que la organización de la cromatina en el espacio acompaña y/o causa el proceso de deciduallización. Estos análisis incluirán la cuantificación de la cantidad de compartimientos que presenten un cambio en el estado de la cromatina y la identificación de TADs que presenten un cambio en sus interacciones intraTAD y contengan (o no) genes que cambien su expresión. También se realizará la integración de los datos provenientes

del Hi-C y RNAseq con los obtenidos a partir de los ensayos de ATACseq en el mismo experimento, que actualmente se encuentran en etapa de preprocesamiento y análisis de calidad en la facility del CRG en Barcelona.

Los análisis de componentes principales (PCA) de los perfiles de compartimentalización mostraron una reestructuración general luego de 3 días del estímulo decidualizante. Esta estructura cambió nuevamente a los 6 días. La decidualización comprendería una remodelación de la cromatina en al menos dos etapas. Esta transición en dos estadios se condice con estudios de transcripción realizados previamente (Rytkonen et al. 2019), donde se identifican dos perfiles de expresión distintos por los que pasan las células durante la decidualización, uno temprano (3 días) y uno tardío (8 días). En dicho estudio, Rytkonen et al. utilizaron la misma línea celular que en este trabajo (tHESC), e indujeron su decidualización mediante 8-br-cAMP y la progestina sintética acetato de medroxiprogesterona (MPA), renovando el estímulo cada dos días, seguidos de ensayos de RNAseq al tiempo cero y a los 3 y 8 días. Pudieron diferenciar claramente distintos perfiles de expresión para los distintos estadios. Rytkonen et al. (Rytkonen et al. 2019) caracterizaron el estadio temprano alcanzado a los 3 días por una aumento en la vía inflamatoria STAT, mientras que el estadio tardío alcanzado a los 8 días se lo caracterizó por una represión de STAT y un aumento en la vía de NFkB.

En un trabajo reciente de nuestro laboratorio (Tesis de licenciatura de Julieta Erramouspe, 2022) se analizaron los perfiles de expresión mediante RNAseq en el mismo modelo utilizado en este trabajo, al día 3 y 6 tras la inducción. Se observaron dos perfiles de expresión diferentes, al igual que en el trabajo de Rytkonen et al. Sin embargo, dichos perfiles no eran exactamente iguales en cuánto a las ontologías sobrerepresentadas, ya que a pesar de observarse un aumento de la vía NfkB a los 6 días no observamos un aumento de la senescencia. A pesar de diferir en los análisis ontológicos, ambos trabajos coinciden en que se pueden distinguir claramente dos perfiles de expresión, uno para un estadio temprano y otro para un estadio tardío. Según nuestros resultados los cambios en la estructura de la cromatina acompañarían estos cambios en el transcriptoma.

Por otra parte, Stadhouders *et al.* (Stadhouders et al. 2018), con el objetivo de estudiar la relación causal entre cambios en la dinámica de la transcripción y organización de la cromatina, estudiaron dichos cambios en un modelo de desdiferenciación transgénica, partiendo de células b de ratones en las que se induce la expresión de los genes OCT4, SOX2, KLF4 y c-MYC, que llevan a la desdiferenciación celular. Al analizar las regiones del genoma donde hay un mayor cambio en la transcripción, observaron que un número significativo de estas regiones presenta un cambio de compartimentalización de manera previa al cambio en la transcripción. De esta manera establecieron una asociación entre estos dos fenómenos en el tiempo. En nuestro caso análisis futuros en nuestro modelo de diferenciación nos permitirán establecer si los cambios sincrónicos de expresión y topología presentan una relación similar a lo descripto para el modelo de la desdiferenciación en células B de ratón, donde los cambios de compartimiento están ligados a un cambio en la transcripción, o si bien son ambos consecuencias de la decidualización, con la posibilidad de que sean independientes entre sí.

Nuestros resultados muestran que a la resolución (100 kb) a la que se analizaron los datos de HiC en nuestro modelo no hay un cambio global de la topología génica tras una hora de tratamiento. Esto es contrario a lo que sucede por ejemplo en células de cáncer de mama T47D en respuesta a un estrés osmótico (Amat et al. 2019), donde la cromatina se reorganiza como respuesta al ambiente. En dicho trabajo (Amat et al. 2019). se aplicó un estrés osmótico (110 mM NaCl, 488 mOsm) durante una hora y se pudo observar un cambio global en la organización A/B de la cromatina, así como en el número de TADs y el posicionamiento de proteínas estructurales de la cromatina, como CTCF. Además, estos cambios fueron reversibles como así lo demuestra el modelo experimental en el que se retira a las células del estrés osmótico (Amat et al. 2019).

A pesar de que en nuestro caso los cambios provocados a la hora no se dan de manera generalizada en el genoma, es posible que la respuesta a los estímulos decidualizantes a corto plazo provoque cambios de manera puntual, por lo que no sería necesaria una respuesta general como en el caso del estrés osmótico, donde se da una represión

generalizada de la transcripción. En el trabajo de nuestro laboratorio antes mencionado (Tesis de licenciatura de Julieta Erramouspe, 2022) se analizaron también los perfiles de expresión a una hora luego de la inducción, y se observaron alrededor de 100 genes que aumentan su expresión. Actualmente ya comenzamos a investigar si ocurren cambios en la organización de la cromatina alrededor de estos 100 genes. Nuestra hipótesis es que cambios en la organización de la cromatina de manera localizada en estas regiones podrían estar permitiendo el disparo de la expresión de genes clave, ya sea para la respuesta transitoria a progesterona o para la transición a una respuesta estable.

Si en cambio no se encuentran cambios en el estado de la cromatina a la hora, podría estar ocurriendo un mecanismo similar al propuesto por LeDily *et al.* (Le Dily et al. 2019) para la línea celular de cáncer de mama T47D. En dicho trabajo se observó que los dominios enriquecidos en genes que se activan como parte de la respuesta frente a progesterona y estradiol, se encuentran en contacto entre sí. Esta interacción entre dominios que responden a P4 y E2 está presente incluso antes del tratamiento, mediante los sitios de unión al receptor de progesterona (PGR), y presentan marcas correspondientes a un estado abierto de la cromatina. La cromatina se encuentra organizada en forma previa a los estímulos hormonales, de manera tal de responder rápidamente a ellos, por lo que una hora luego del tratamiento se observa un cambio de expresión y no en la topología de la cromatina. En otro trabajo realizado en nuestro laboratorio en la línea celular epitelial endometria Ishikawa (La Greca et al. 2022) tampoco se observó un cambio en la compartimentalización A/B luego de un tratamiento de 12 horas con E2 o R5020. A pesar de no presentar un cambio en la compartimentalización, sí mostraban una respuesta en la expresión génica, gobernadas por zonas de regulación donde se unen los factores de transcripción PGR y PAX2. Estas regiones regulatorias interactúan con los promotores de los genes regulados por progestinas que se encuentran en el mismo TAD. Esto indica que la cromatina en las células Ishikawa se encuentra organizada para responder a la regulación hormonal incluso antes del tratamiento.

Al comparar la dinámica de cambio en la morfología y la compartimentalización,

observamos que ambas correlacionan. Al día 3 y 6 la morfología de ambos tratamientos difiere del vehículo (EtOH), al igual que lo hace la compartimentalización de la cromatina. Tanto en la compartimentalización como en la morfología se observa un cambio marcado entre el día 3 y el tiempo cero y entre el día 3 y el 6. Sin embargo, el cambio para la compartimentalización es más abrupto entre el tiempo cero y el tercer día que entre este y el sexto día. El PC1 entre el día 3 y 6 es mucho más similar que entre el tiempo cero y el día 3. Esto podría indicar que los eventos que llevan a una remodelación general de la cromatina tienen lugar durante los primeros días de la deciduallización.

Resumiento, a partir de este abordaje cumplimos con los dos objetivos propuestos. Fue posible corroborar la correcta deciduallización del modelo elegido, cuantificando los cambios morfológicos que ocurren durante la transdiferenciación *in vitro* de las células del estroma endometrial, así como comparar las dinámicas entre los distintos tratamientos. También se logró analizar los cambios en la topología de la cromatina de manera global, observando una reestructuración general de la misma asociada a la diferenciación alcanzada.

A partir de estos resultados es posible comenzar los análisis para caracterizar en profundidad dichos cambios topológicos, y de qué manera podrían relacionarse con los cambios en la expresión génica.

9 Anexo

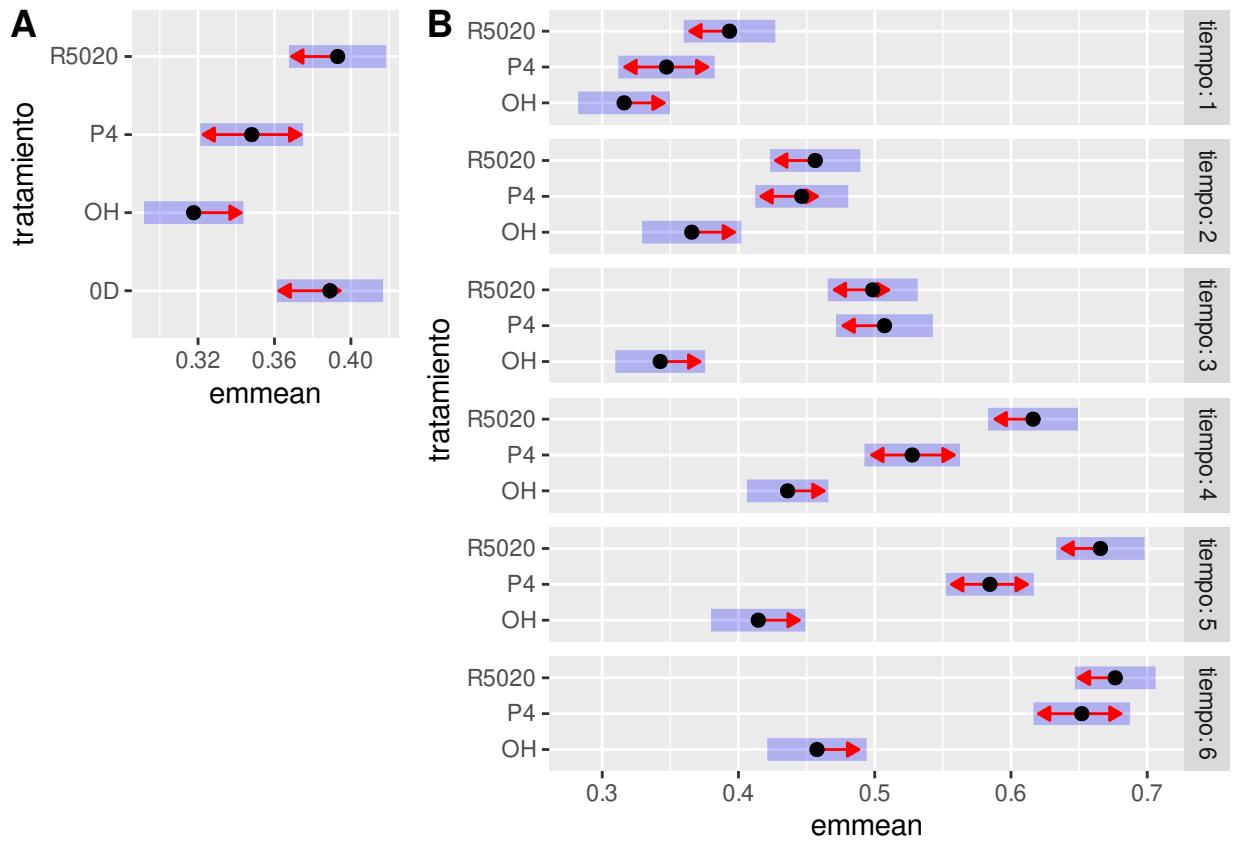


Figura S1: Resultado de las comparaciones de circularidad realizadas. Los puntos negros muestran para cada día y tratamiento el promedio estimado por el modelo resultante, las barras azules muestran el intervalo de confianza de dicho promedio con una confianza del 95%. Las flechas rojas denotan las comparaciones realizadas, si las flechas de dos tratamientos se superponen indica que su circularidad no difiere significativamente.

```

tiempo = 1:
contrast estimate SE df lower.CL upper.CL t.ratio p.value
OH - P4    -0.03112 0.0226 12  -0.0915   0.0292  -1.376  0.3834
OH - R5020 -0.07734 0.0220 12  -0.1362   -0.0185  -3.508  0.0111
P4 - R5020 -0.04621 0.0224 12  -0.1060   0.0135  -2.063  0.1397

tiempo = 2:
contrast estimate SE df lower.CL upper.CL t.ratio p.value
OH - P4    -0.08070 0.0231 12  -0.1424   -0.0190  -3.491  0.0115
OH - R5020 -0.09062 0.0228 12  -0.1514   -0.0298  -3.975  0.0048
P4 - R5020 -0.00991 0.0218 12  -0.0681   0.0483  -0.454  0.8935

tiempo = 3:
contrast estimate SE df lower.CL upper.CL t.ratio p.value
OH - P4    -0.16459 0.0224 12  -0.2244   -0.1047  -7.335 <.0001
OH - R5020 -0.15596 0.0216 12  -0.2136   -0.0983  -7.219 <.0001
P4 - R5020  0.00862 0.0223 12  -0.0508   0.0681  0.387  0.9213

tiempo = 4:
contrast estimate SE df lower.CL upper.CL t.ratio p.value
OH - P4    -0.09146 0.0213 12  -0.1483   -0.0346  -4.294  0.0028
OH - R5020 -0.18011 0.0206 12  -0.2351   -0.1252  -8.745 <.0001
P4 - R5020 -0.08866 0.0221 12  -0.1477   -0.0296  -4.008  0.0046

tiempo = 5:
contrast estimate SE df lower.CL upper.CL t.ratio p.value
OH - P4    -0.17000 0.0219 12  -0.2285   -0.1115  -7.751 <.0001
OH - R5020 -0.25114 0.0220 12  -0.3098   -0.1925 -11.427 <.0001
P4 - R5020 -0.08115 0.0210 12  -0.1372   -0.0251  -3.863  0.0059

tiempo = 6:
contrast estimate SE df lower.CL upper.CL t.ratio p.value
OH - P4    -0.19435 0.0235 12  -0.2571   -0.1316  -8.257 <.0001
OH - R5020 -0.21893 0.0218 12  -0.2770   -0.1608 -10.051 <.0001
P4 - R5020 -0.02458 0.0212 12  -0.0811   0.0320  -1.159  0.4981

Degrees-of-freedom method: containment
Confidence level used: 0.95
Conf-level adjustment: tukey method for comparing a family of 3 estimates
P value adjustment: tukey method for comparing a family of 3 estimates

```

Figura S2: Tabla de resultados de los contrastes o comparaciones realizadas. Se muestra para cada contraste la diferencia en la circularidad estimada (estimate), el error estándar (SE), los grados de libertad (df), el límite del intervalo de confianza de la diferencia inferior (lower.CL) y superior (upper.CL), indicando que la diferencia entre esos tratamientos se encuentra con un 95% de confianza en ese intervalo. Se indica también el t valor (t.ratio) y el p-valor correspondiente (p.value).

10 Bibliografía

Amat, R., Böttcher, R., Le Dily, F., Vidal, E., Quilez, J., Cuartero, Y., Beato, M., de Nadal, E. & Posas, F. (2019), ‘Rapid reversible changes in compartments and local chromatin organization revealed by hyperosmotic shock’, *Genome Research* **29**(1), 18–28.

URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6314167/>

Arici, A., Marshburn, P. B., MacDonald, P. C. & Dombrowski, R. A. (1999), ‘Progesterone Metabolism in Human Endometrial Stromal and Gland Cells in Culture’, *Steroids* **64**(8), 530–534.

URL: <https://linkinghub.elsevier.com/retrieve/pii/S0039128X9900029X>

Bao, L., Tessier, C., Prigent-Tessier, A., Li, F., Buzzio, O. L., Callegari, E. A., Horseman, N. D. & Gibori, G. (2007), ‘Decidual Prolactin Silences the Expression of Genes Detrimental to Pregnancy’, *Endocrinology* **148**(5), 2326–2334.

URL: <https://academic.oup.com/endo/article/148/5/2326/2502147>

Carter, A., Hills, F., O’Gorman, D., Roberts, C., Sooranna, S., Watson, C. & Westwood, M. (2004), ‘The Insulin-like Growth Factor System in Mammalian Pregnancy—A Workshop Report’, *Placenta* **25**, S53–S56.

URL: <https://linkinghub.elsevier.com/retrieve/pii/S0143400404000335>

Critchley, H. O. D., Maybin, J. A., Armstrong, G. M. & Williams, A. R. W. (2020), ‘Physiology of the Endometrium and Regulation of Menstruation’, *Physiological Reviews* **100**(3), 1149–1179.

URL: <https://journals.physiology.org/doi/10.1152/physrev.00031.2019>

Daly, D. C., Maslar, I. A. & Riddick, D. H. (1983), ‘Prolactin production during in vitro decidualization of proliferative endometrium’, *American Journal of Obstetrics and Gynecology* **145**(6), 672–678.

URL: <https://linkinghub.elsevier.com/retrieve/pii/0002937883905720>

DeMayo, F. J. & Lydon, J. P. (2020), '90 YEARS OF PROGESTERONE: New insights into progesterone receptor signaling in the endometrium required for embryo implantation', *Journal of Molecular Endocrinology* **65**(1), T1–T14.

URL: <https://jme.bioscientifica.com/view/journals/jme/65/1/JME-19-0212.xml>

Durand, N. C., Shamim, M. S., Machol, I., Rao, S. S., Huntley, M. H., Lander, E. S. & Aiden, E. L. (2016), 'Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments', *Cell Systems* **3**(1), 95–98.

URL: <https://linkinghub.elsevier.com/retrieve/pii/S2405471216302198>

Gellersen, B. & Brosens, J. (2003), 'Cyclic AMP and progesterone receptor cross-talk in human endometrium: a decidualizing affair', *Journal of Endocrinology* **178**(3), 357–372.

URL: <https://joe.bioscientifica.com/view/journals/joe/178/3/357.xml>

Gellersen, B. & Brosens, J. J. (2014), 'Cyclic Decidualization of the Human Endometrium in Reproductive Health and Failure', *Endocrine Reviews* **35**(6), 851–905.

URL: <https://academic.oup.com/edrv/article/35/6/851/2354669>

Gleeson, L. M., Chakraborty, C., McKinnon, T. & Lala, P. K. (2001), 'Insulin-Like Growth Factor-Binding Protein 1 Stimulates Human Trophoblast Migration by Signaling through a5b1 Integrin via Mitogen-Activated Protein Kinase Pathway¹', *The Journal of Clinical Endocrinology & Metabolism* **86**(6), 2484–2493.

URL: <https://academic.oup.com/jcem/article-lookup/doi/10.1210/jcem.86.6.7532>

Grimm, S. L., Hartig, S. M. & Edwards, D. P. (2016), 'Progesterone Receptor Signaling Mechanisms', *Journal of Molecular Biology* **428**(19), 3831–3849.

URL: <https://linkinghub.elsevier.com/retrieve/pii/S002228361630242X>

Hagan, C. R. & Lange, C. A. (2014), 'Molecular determinants of context-dependent progesterone receptor action in breast cancer', *BMC Medicine* **12**(1), 32.

URL: <http://bmcmedicine.biomedcentral.com/articles/10.1186/1741-7015-12-32>

Imakaev, M., Fudenberg, G., McCord, R. P., Naumova, N., Goloborodko, A., Lajoie, B. R., Dekker, J. & Mirny, L. A. (2012), ‘Iterative correction of Hi-C data reveals hallmarks of chromosome organization’, *Nature Methods* **9**(10), 999–1003.

Jabbour, H. & Critchley, H. (2001), ‘Potential roles of decidual prolactin in early pregnancy’, *Reproduction* pp. 197–205.

URL: <https://rep.bioscientifica.com/view/journals/rep/121/2/197.xml>

Kin, K., Maziarz, J. & Wagner, G. P. (2014), ‘Immunohistological Study of the Endometrial Stromal Fibroblasts in the Opossum, *Monodelphis domestica*: Evidence for Homology with Eutherian Stromal Fibroblasts1’, *Biology of Reproduction* **90**(5).

URL:

<https://academic.oup.com/biolreprod/article-lookup/doi/10.1095/biolreprod.113.115139>

Krikun, G., Mor, G., Alvero, A., Guller, S., Schatz, F., Sapi, E., Rahman, M., Caze, R., Qumsiyeh, M. & Lockwood, C. J. (2004), ‘A novel immortalized human endometrial stromal cell line with normal progestational response’, *Endocrinology* **145**(5), 2291–2296.

Kuhl, H. (2011), ‘Pharmacology of progestogens’, *Journal für Reproduktionsmedizin und Endokrinologie-Journal of Reproductive Medicine and Endocrinology* **8**(Special Issue 1), 157–76.

URL: <https://www.kup.at/kup/pdf/10168.pdf>

La Greca, A., Bellora, N., Le Dily, F., Jara, R., Nacht, A. S., Quilez Oliete, J., Villanueva, J. L., Vidal, E., Merino, G., Fresno, C., Tarifa Reischle, I., Vallejo, G., Vicent, G., Fernández, E., Beato, M. & Saragüeta, P. (2022), ‘Chromatin topology defines estradiol-primed progesterone receptor and PAX2 binding in endometrial cancer cells’, *eLife* **11**, e66034.

URL: <https://doi.org/10.7554/eLife.66034>

Le Dily, F., Vidal, E., Cuartero, Y., Quilez, J., Nacht, A. S., Vicent, G. P.,

- Carbonell-Caballero, J., Sharma, P., Villanueva-Canas, J. L., Ferrari, R., De Llobet, L. I., Verde, G., Wright, R. H. G. & Beato, M. (2019), ‘Hormone-control regions mediate steroid receptor-dependent genome organization’, *Genome Research* **29**(1), 29–39.
- Li, H. & Durbin, R. (2009), ‘Fast and accurate short read alignment with Burrows-Wheeler transform’, *Bioinformatics (Oxford, England)* **25**(14), 1754–1760.
- Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S. & Dekker, J. (2009), ‘Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome’, *Science* **326**(5950), 289–293.
URL: <https://www.science.org/doi/10.1126/science.1181369>
- Lucas, E. S., Vrljicak, P., Muter, J., Diniz-da Costa, M. M., Brighton, P. J., Kong, C.-S., Lipecki, J., Fishwick, K. J., Odendaal, J., Ewington, L. J., Quenby, S., Ott, S. & Brosens, J. J. (2020), ‘Recurrent pregnancy loss is associated with a pro-senescent decidual response during the peri-implantation window’, *Communications Biology* **3**(1), 1–14.
URL: <https://www.nature.com/articles/s42003-020-0763-1>
- Lynch, V., Nnamani, M., Kapusta, A., Brayer, K., Plaza, S., Mazur, E., Emera, D., Sheikh, S., Grützner, F., Bauersachs, S., Graf, A., Young, S., Lieb, J., DeMayo, F., Feschotte, C. & Wagner, G. (2015), ‘Ancient Transposable Elements Transformed the Uterine Regulatory Landscape and Transcriptome during the Evolution of Mammalian Pregnancy’, *Cell Reports* **10**(4), 551–561.
URL: <https://linkinghub.elsevier.com/retrieve/pii/S221112471401105X>
- Maslar, I. A. & Ansbacher, R. (1986), ‘Effects of progesterone on decidual prolactin production by organ cultures of human endometrium’, *Endocrinology* **118**(5), 2102–2108.
- Maslar, I. A. & Riddick, D. H. (1979), ‘Prolactin production by human endometrium

during the normal menstrual cycle', *American Journal of Obstetrics and Gynecology* **135**(6), 751–754.

URL: <https://linkinghub.elsevier.com/retrieve/pii/0002937879903867>

Mess, A. & Carter, A. M. (2006), 'Evolutionary transformations of fetal membrane characters in Eutheria with special reference to Afrotheria', *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **306B**(2), 140–163.

URL: <https://onlinelibrary.wiley.com/doi/10.1002/jez.b.21079>

Mestre Citrinovitz, A. C., Strowitzki, T. & Germeyer, A. (2019), 'Decreased Autophagy Impairs Decidualization of Human Endometrial Stromal Cells: A Role for ATG Proteins in Endometrial Physiology', *International Journal of Molecular Sciences* **20**(12), 3066.

URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6628477/>

Nichols, M. H. & Corces, V. G. (2021), 'Principles of 3D compartmentalization of the human genome', *Cell Reports* **35**(13), 109330.

URL: <https://linkinghub.elsevier.com/retrieve/pii/S2211124721007063>

Perelman, S. B. & Garibaldi, L. A. (n.d.), *INTRODUCCION A LOS MODELOS MIXTOS*.

Rao, S., Huntley, M., Durand, N., Stamenova, E., Bochkov, I., Robinson, J., Sanborn, A., Machol, I., Omer, A., Lander, E. & Aiden, E. (2014), 'A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping', *Cell* **159**(7), 1665–1680.

URL: <https://linkinghub.elsevier.com/retrieve/pii/S0092867414014974>

Rao, S. S., Huang, S.-C., Glenn St Hilaire, B., Engreitz, J. M., Perez, E. M., Kieffer-Kwon, K.-R., Sanborn, A. L., Johnstone, S. E., Bascom, G. D., Bochkov, I. D., Huang, X., Shamim, M. S., Shin, J., Turner, D., Ye, Z., Omer, A. D., Robinson, J. T., Schlick, T., Bernstein, B. E., Casellas, R., Lander, E. S. & Aiden, E. L. (2017), 'Cohesin Loss Eliminates All Loop Domains', *Cell* **171**(2), 305–320.e24.

URL: <https://linkinghub.elsevier.com/retrieve/pii/S0092867417311200>

Rutanen, E.-M., Koistinen, R., Sjöberg, J., Julkunen, M., Wahlström, T., Bohn, H. & Seppälä, M. (1986), ‘Synthesis of Placental Protein 12 by Human Endometrium*’, *Endocrinology* **118**(3), 1067–1071.

URL: <https://academic.oup.com/endo/article-lookup/doi/10.1210/endo-118-3-1067>

Rutanen, E.-M., Koistinen, R., Wahlstrom, T., Bohn, H., Ranta, T. & Seppala, M. (1985), ‘Synthesis of Placental Protein 12 by Human Decidua’, *Endocrinology* **116**(4), 1304–1309.

URL: <https://academic.oup.com/endo/article-lookup/doi/10.1210/endo-116-4-1304>

Rytkonen, K. T., Erkenbrack, E. M., Poutanen, M., Elo, L. L., Pavlicev, M. & Wagner, G. P. (2019), ‘Decidualization of Human Endometrial Stromal Fibroblasts is a Multiphasic Process Involving Distinct Transcriptional Programs’, *Reproductive Sciences* **26**(3), 323–336.

URL: <http://journals.sagepub.com/doi/10.1177/1933719118802056>

Serra, F., Baù, D., Goodstadt, M., Castillo, D., Filion, G. J. & Marti-Renom, M. A. (2017), ‘Automatic analysis and 3D-modelling of Hi-C data using TADbit reveals structural features of the fly chromatin colors’, *PLOS Computational Biology* **13**(7), e1005665.

URL: <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005665>

Stadhouders, R., Vidal, E., Serra, F., Di Stefano, B., Le Dily, F., Quilez, J., Gomez, A., Collombet, S., Berenguer, C., Cuartero, Y., Hecht, J., Filion, G. J., Beato, M., Marti-Renom, M. A. & Graf, T. (2018), ‘Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming’, *Nature genetics* **50**(2), 238–249.

URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5810905/>

Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. (2021), ‘Cellpose: a generalist algorithm for cellular segmentation’, *Nature Methods* **18**(1), 100–106.

URL: <https://www.nature.com/articles/s41592-020-01018-x>

Tognetti, P. M. & Perez, A. (2019), *Introduccion a los modelos mixtos*, Universidad de Buenos Aires. Facultad de Agronomia.

URL: <https://ri.conicet.gov.ar/handle/11336/138828>

Vallejo, G., Ballare, C., Lino Baranao, J., Beato, M. & Saragueta, P. (2005), ‘Progestin Activation of Nongenomic Pathways via Cross Talk of Progesterone Receptor with Estrogen Receptor b Induces Proliferation of Endometrial Stromal Cells’, *Molecular Endocrinology* **19**(12), 3023–3037.

URL: <https://academic.oup.com/mend/article/19/12/3023/2738001>

Waisman, A., Norris, A. M., Elías Costa, M. & Kopinke, D. (2021), ‘Automatic and unbiased segmentation and quantification of myofibers in skeletal muscle’, *Scientific Reports* **11**(1), 11793.

URL: <https://www.nature.com/articles/s41598-021-91191-6>

Yu, J., Berga, S. L., Johnston-MacAnanny, E. B., Sidell, N., Bagchi, I. C., Bagchi, M. K. & Taylor, R. N. (2016), ‘Endometrial Stromal Decidualization Responds Reversibly to Hormone Stimulation and Withdrawal’, *Endocrinology* **157**(6), 2432–2446.