



Viviendo al límite:

¿Importa la posición de los genes
dentro de los TADs?

2021-2022

Máster en Biología Molecular y Biomedicina

Universidad de Cantabria / Universidad del País Vasco

Estudiante: Thais Ealo Rodríguez

Instituto: Instituto de Biomedicina y Biotecnología de Cantabria

Laboratorio: Regulación transcripcional en enfermedades congénitas y de
desarrollo

Director: Dr. Álvaro Rada Iglesias

ÍNDICE

Abstract.....	3
1. Antecedentes del tema.....	4
1.1. Potenciadores.....	4
1.2. Organización 3D del genoma.....	4
1.3. Modelo de extrusión de bucle dependiente de cohesina.....	5
1.4. Reordenamientos cromosómicos asociados a patologías.....	7
1.5. Localización gen-potenciador dentro de los TAD.....	8
2. Objetivos.....	10
3. Metodología.....	10
3.1. Diseño de las secuencias guías.....	10
3.1.1 Vector pX330A.....	11
3.2. Clonaje del gRNA en el vector px330A.....	11
3.2.1 Digestión vector.....	11
3.2.2 Hibridación oligos.....	12
3.2.3 Ligación.....	12
3.3. Transformación celular.....	13
3.3.1 Colony PCR.....	13
3.3.2 Purificación del plásmido.....	14
3.4. Mantenimiento de las mESC.....	14
3.4.1 Líneas de mESC transgénicas.....	14
3.4.2 Transfección de las mESC.....	15
3.4.3 Genotipado de la población.....	15
3.4.4 Derivación líneas clonales mESC.....	17
3.5. Diferenciación de mESC a AntNPC.....	18
3.5.1 Extracción RNA.....	18
3.5.2 Síntesis de cDNA.....	18
3.5.3 RT-qPCR.....	19
4. Resultados.....	20
4.1. Comprobación de la diferencia en la expresión de Six3/Six2 entre las línea WT y la línea con la delección de la barrera Six3/Six2 tras la diferenciación de mESC a AntNPCs.....	20
4.2. Generación de la línea Six3 Inversión usando la tecnología CRISPR/Cas9.....	22
4.2.1 Genotipado Six3 INV.....	22

4.2.2 Análisis de la expresión por qPCR de Six3/Six2 tras la diferenciación de Inv mESC a AntNPC.....	25
4.3. Generación de las líneas Six3 ^{-/-} usando la tecnología CRISPR/Cas9.....	25
4.3.1 Genotipado Six3 KO	26
4.3.2 Análisis de la expresión por qPCR de Six3/Six2 tras la diferenciación de mESC a AntNPC.....	28
5. Discusión	29
6. Conclusiones	30
7. Bibliografía	30

Abstract

The genome is organized in large (Mb-scale) self-interacting domains that have termed as topological association domains (TADs). TADs are considered fundamental regulatory units that favor the interactions between regulatory elements (e.g. enhancers) and genes located within the same domain, while insulating genes from contacting ectopic regulatory elements located in neighboring domains. Previous analyses in the Rada-Iglesias laboratory revealed that developmental genes tend to be located in the borders of TADs close to CTCF binding sites (CBS), something that had not been reported to date. Since the modification of TAD organization can cause various diseases, we decided to investigate whether the positioning of genes near TAD boundaries was of any functional relevance. We performed CRISPR/Cas9-based genetic engineering of the *Six3/Six2 locus*. *Six3* and *Six2* are located at the border of two neighboring TADs separated by a large cluster of CBS. Moreover, *Six3* becomes activated upon differentiation of ESC into AntNPC, while *Six2* is inactive in both cell types. Deleting the CTCF boundary between both genes, we observed that, after differentiation to AntNPC, the expression of *Six2* increased significantly, confirming that the boundary prevents the ectopic activation of *Six2* by *Six3* enhancers. More interestingly, we found that the positioning of *Six3* close to the TAD boundary increases insulation, since by removing *Six3* plus the CTCF boundary, we showed that the expression of *Six2* increased significantly compared to the line where only the CTCF boundary was deleted. Finally, we showed that the gene itself does not benefit from being on the border of TAD, since after performing an inversion and placing *Six3* in the middle of the TAD, its expression was not affected. Overall, these results demonstrate that CBS and gene positioning at TAD borders synergistically control the insulation of developmental regulatory domains.

1. Antecedentes del tema

1.1. Potenciadores

El genoma no codificante está repleto de miles de elementos reguladores entre los que se encuentran los potenciadores. Estos elementos son regiones cortas de ADN *cis*-reguladores que permiten el control espacio-temporal de la expresión génica durante el desarrollo. Los potenciadores contienen regiones que actúan como sitios de unión para factores de transcripción (Smith et al., 2014). Cuando un potenciador se encuentra activo (en regiones de cromatina abierta) se caracteriza por la presencia de marcas de cromatina específicas como H3K4me1 y H3K27ac (Nathaniel et al., 2009).

Una de las características más importantes de los potenciadores es su capacidad de actuar sobre genes diana localizados a grandes distancias. Cabe señalar que un potenciador dado puede llegar a regular la expresión de un gen distal saltándose otros genes ubicados más próximos. Esto quiere decir que un gen no puede ser activado por cualquier potenciador ya que existe cierto grado de especificidad promotor-potenciador (Ghavi, 2019). Estudios realizados en los últimos años sugieren que mutaciones genéticas que alteran la función de los potenciadores o variantes estructurales que cambian la comunicación entre los potenciadores y sus genes días están asociado al desarrollo de enfermedades (Claringbould y Zaugg 2021) (Spielmann et al., 2018).

1.2. Organización 3D del genoma

El genoma de las células de mamífero puede llegar a medir hasta 2 metros, por lo que es un gran desafío para las células poder compactarlo en un espacio tan pequeño como es su núcleo de unas pocas micras de diámetro, asegurando, además, la funcionalidad adecuada del genoma en el momento concreto. Para ello, el genoma se encuentra plegado y empaquetado formando fibras de cromatina. Dicho plegamiento está altamente organizado y relacionado con la regulación transcripcional (Zheng y Xie, 2019).

Hoy en día es relativamente fácil estudiar la estructura tridimensional del ADN gracias a las mejoras en las técnicas de captura de la conformación cromosómica (3C) (Sati y Cavalli, 2017). En concreto, el Hi-C se convirtió en una herramienta imprescindible para poder comprender la organización del genoma en 3D, ya que proporciona de manera global todas las interacciones que existen dentro del genoma (Lieberman-Aiden et al., 2009).

Al analizar los mapas de calor de Hi-C (**Figura 1**) se demuestra la existencia de dominios estructurales con forma de `pirámides` llamados dominios de asociación topológica (TADs) (Dixon et al., 2016).

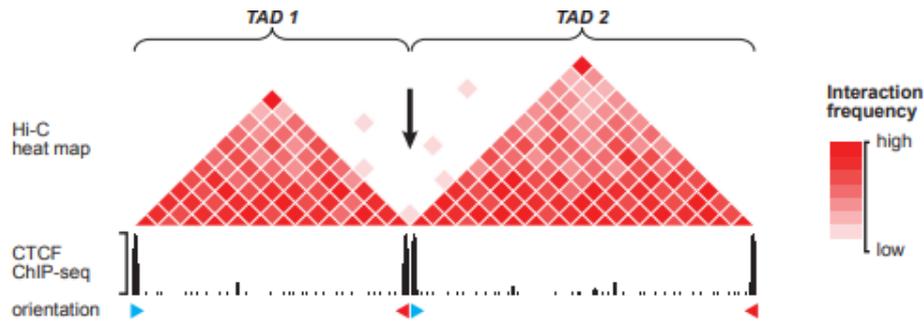


Figura 1. Esquema de un mapa de calor en el que se representan dos TADs obtenidos por Hi-C. La flecha señala la barrera CTCF entre TADs (Chang et al., 2019).

Estos dominios se caracterizan por un alto número de interacciones intradominio, entre las que se incluyen contactos potenciador-promotor, frente a un bajo número de interacciones interdominio. Por tanto, se considera que los TADs se encargan de producir aislamiento, impidiendo interacciones no deseadas entre genes y potenciadores ectópicos, y de esa manera contribuyen al control de la expresión génica (Sexton et al., 2012).

Entre distintos tipos celulares y de especies, los TAD están bastante conservados, teniendo, además, una configuración jerárquica en la que los TADs más pequeños pueden crear TADs de mayor tamaño. Los tamaños pueden variar de ~40 kb a ~3 MB, con un tamaño medio de ~185 kb (Hansen et al., 2018).

1.3. Modelo de extrusión de bucle dependiente de cohesina

Tras el descubrimiento de los TADs, gracias a los estudios de Hi-C, se realizaron experimentos por ChIP-Seq en lo que se observaron que los límites de los TADs estaban enriquecidos de la proteína aislante CTCF (86%) y las subunidades de cohesina RAD21 (86%) y SMC3 (87%). Por ello se sugirió que tanto CTCF como cohesina juegan un papel fundamental tanto en la formación como en el mantenimiento de los TAD (Rao et al., 2014).

CTCF se trata de una fosfoproteína con 11 dedos de zinc conservada evolutivamente, gracias a los cuales se puede unir a sitios diana. Los dedos de zinc se encuentran flanqueados por grandes dominios C- y N- cuya función aún no es conocida (Ohlsson et al., 2001). Clásicamente, CTCF se conoce como una proteína "aislante", jugando un papel muy importante en el plegamiento de la cromatina y la formación de los límites del bucle. Las interacciones mediadas por CTCF se facilitan cuando se une a sitios de unión con una orientación convergente (Lee et al., 2022).

Por su parte, la cohesina es un complejo multiprotéico con forma de anillo que está compuesto por Smc1, Smc3, Rad21 y SA1/2 (Hansen et al., 2018). Esta proteína, además de mediar la

cohesión entre las cromátidas hermanas, es capaz de regular la expresión y organización del genoma, gracias a la formación de bucles de cromatina (Richerova et al., 2017).

Uno de los modelos que postula la formación de los TAD es la extrusión del bucle (Fudenberg et al., 2016). Este modelo se basa en la extrusión de bucles de cromatina gracias a los anillos de cohesina, los cuales permiten la organización del genoma en dominios espaciales. Principalmente, el complejo de cohesina se asocia a la fibra de cromatina de manera aleatoria y comienza la formación progresiva de un bucle de ADN a través de extrusión bidireccional (**Figura 2**). La cohesina se detiene, y por tanto se bloquea la extrusión del bucle, cuando se encuentra una molécula de CTCF en una orientación determinada. Al encontrarse dos moléculas de CTCF orientados de manera convergente, en ambos bordes del TAD, el bucle se estabiliza. Esta hipótesis cambió la idea que se pensaba sobre la organización del genoma (Dekker y Mirny, 2016).

CTCF presenta una secuencia específica de unión al ADN implicada en la formación de bucles de cromatina, mientras que la cohesina se asocia con la cromatina de forma independiente a la secuencia de ADN. Gracias a este sistema, la cohesina y CTCF son capaces de aproximar secuencias que linealmente se encuentran muy lejanas (Richerova et al., 2017).

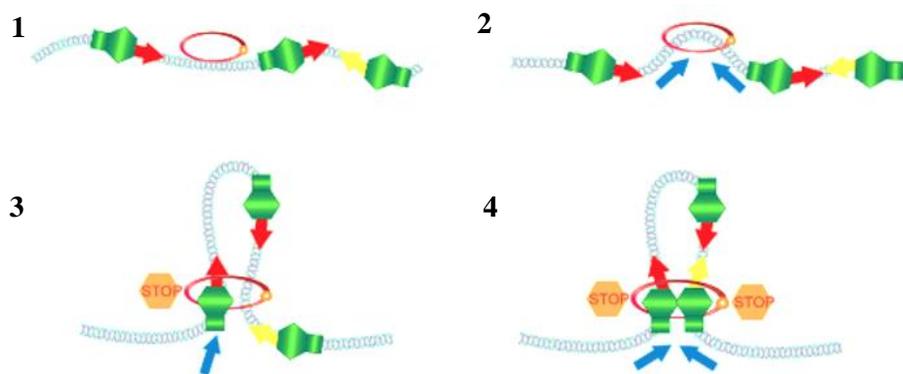


Figura 2. Formación de TADs mediante el modelo de extrusión de bucle. **1.** El anillo de cohesina se une a las fibras de cromatina. **2.** El anillo comienza a crear el bucle gracias a su desplazamiento a través de la cromatina. **3.** La presencia de CTCF en la correcta orientación termina la extrusión de un lado. **4.** El bucle se forma cuando la cohesina se une a un dímero de CTCF en una orientación convergente (Li et al., 2020).

La extrusión del bucle es un proceso activo que requiere energía, lo cual es posible gracias a que la propia cohesina tiene actividad ATPasa (Vian et al., 2018). Se ha informado que la transcripción mueve a la cohesina a lo largo del ADN hasta los sitios ocupados por CTCF, por lo que plantea la posibilidad de que las ARN polimerasas puedan ser una fuerza impulsora para la extrusión del bucle. El primer indicio de que la transcripción puede afectar la distribución de la cohesina dentro del cromosoma vino por la observación de que muchos sitios genómicos

enriquecidos con cohesina están flanqueados por genes transcritos (Richterova et al., 2017). En un estudio realizado por Busslinger et al., se observó la localización de la cohesina en fibroblastos embrionarios de ratón primario (MEF). En el tipo MEFs salvaje, la mayor parte de la cohesina se encuentra co-localizada con CTCF, sin embargo, en ausencia de CTCF, la mayor parte de la cohesina se comienza a acumular aguas abajo de genes activamente transcritos (Busslinger et al., 2017). Sin embargo, estudios más recientes indican que la ARN polimerasa puede representar una barrera para la cohesina y bloquear su progresión (Edward et al., 2022).

1.4. Reordenamientos cromosómicos asociados a patologías

Reordenamientos en la organización 3D de los TAD pueden llegar a provocar diversas enfermedades, como cáncer y trastornos del desarrollo. En una situación normal, los potenciadores contactan con sus genes diana gracias a su co-localización en el interior de un mismo TAD. Sin embargo, variantes estructurales (deleciones, duplicaciones, inversiones o translocaciones) que alteran la organización de los TAD pueden llegar a modificar la posición de los elementos reguladores y/o su comunicación con los genes. Estas modificaciones pueden presentar efectos perjudiciales, provocando alteraciones o creando nuevos contactos entre potenciadores y promotores, resultando en pérdida o ganancia de expresión génica con consecuencias patológicas (Lupiañez et al., 2016- Spielmann y Mundlos, 2013).

Uno de los ejemplos más estudiados es el caso del gen *Sony hedgehog* (*Shh*). Se trata de un gen encargado de controlar la organogénesis y la organización de sistema nervioso central, dedos y extremidades. La expresión de este gen en las extremidades está regulada por un potenciador localizado en su mismo TAD, aproximadamente a una megabase de distancia (Jeong et al., 2006). Symmons et al., 2016 y Williamson et al., 2019, realizaron estudios en los que generaban alteraciones en el TAD de *Shh*, con el fin de estudiar sus efectos en la expresión génica y la organización de la cromatina. Uno de los experimentos que generaron fue realizar deleciones en los sitios de unión de CTCF en el TAD de *Shh*. Al realizar una deleción en un sitio de unión a CTCF, la expresión de *Shh* no se vio afectada. Sin embargo, al deleccionar todos los sitios de unión de CTCF, se comenzaron a formar contactos ectópicos. De esta manera el potenciador comenzó a contactar con genes que se encontraban en el TAD vecino como es el caso del gen *Mnx1*, sin embargo, otros potenciadores localizados en la misma región no influyen en la expresión de *Mnx1*, lo que quiere decir que la interacción entre el potenciador-promotor es específica. Así mismo, crearon inversiones de tal manera que colocan una barrera de CTCFs entre *Shh* y el potenciador, en este caso la expresión de *Shh* se vió drásticamente afectada, con la consecuente malformación en el desarrollo de la pata, confirmando que las barreras del TAD actúan como aislantes.

Otro de los locus más estudiados sobre alteraciones cromosómicas es el caso de *Epha4*. Uno de los ejemplos más conocidos es el de una inversión que causa malformaciones graves en el desarrollo de la extremidad debido a nuevos contactos entre el potenciador de *Epha4* y el gen vecino *Wnt6*. Además, tras la inversión, la expresión de *Epha4* se ve drásticamente afectada al perder el contacto con su potenciador (Lupiáñez *et al.*, 2015). Este tipo de mecanismo patológicos que ocurre entre el potenciador y su nuevo gen diana se denomina adopción de potenciadores y se ha propuesto para explicar el 12% de las deleciones presentes en una base de datos de variantes genómicas raras (Lettice *et al.*, 2011).

1.5. Localización gen-potenciador dentro de los TAD.

Una de las cuestiones aún sin resolver es determinar qué factores permiten que los bordes de los TADs actúen como insuladores de la expresión génica. Los bordes de los TADs se encuentran enriquecidos, además de en los sitios de unión de CTCF ya mencionado previamente, en genes *Housekeeping* (Eisenberg *et al.*, 2013). Por tanto, estudios previos apuntan a que, además de CTCF, estos genes también podrían tener esa propiedad (Dixon *et al.*, 2016).

Debido a todas las incógnitas sobre la organización de los genes y potenciadores dentro de los TAD, un miembro de nuestro laboratorio realizó un análisis bioinformático en el que estudió la distribución de los sitios de comienzo de la transcripción (TSS), observando que la mayoría de ellos se localizaban cerca del borde de los TADs (**Figura 3.A**). Sin embargo, muchos de los TSS analizados correspondían a genes *housekeeping* lo cual ya estaba reportado, por lo que decidió realizar el mismo análisis teniendo en cuenta sólo genes de desarrollo. Además, cabe destacar que es precisamente la expresión de los genes de desarrollo, pero no la de los genes *housekeeping*, la que está regulada principalmente por potenciadores distales (Bergman *et al.*, 2022). Sorprendentemente, los genes de desarrollo mantenían la misma distribución al borde de los TADs (**Figura 3.A**). Del mismo modo, analizó la distribución de los potenciadores y observó que se distribuían de manera equitativa a lo largo del TAD (**Figura 3.B**).

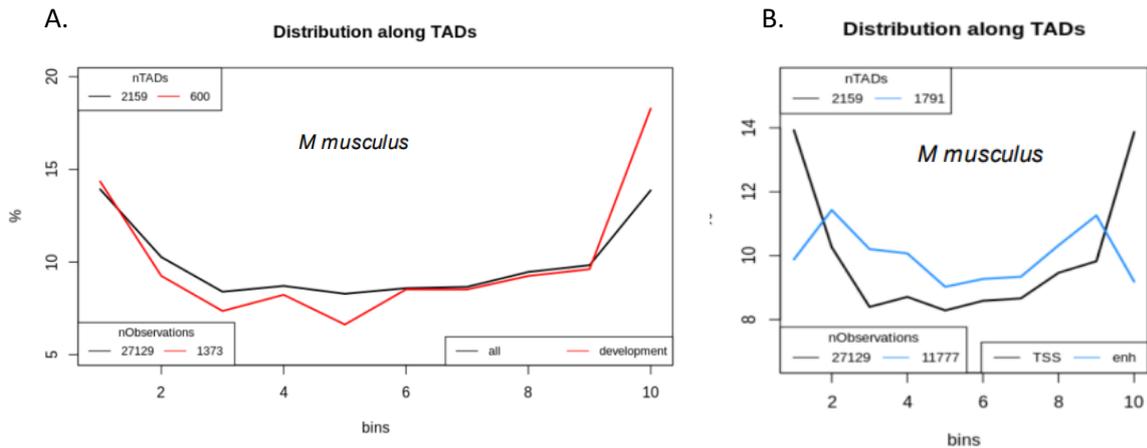


Figura 3. Representación gráfica de la distribución de genes y potenciadores dentro de los TAD en el genoma de ratón (*Mus musculus*). **A.** Acumulación de todos los genes (línea negra) y de los genes de desarrollo (línea roja) en el borde del TAD. **B.** Distribución de los potenciadores (línea azul) y de todos los genes (línea negra) a lo largo de los TAD. La posición de los genes se determinó teniendo en cuenta sus sitios de inicio de transcripción (TSS).

Tras estos interesantes resultados, nos comenzamos a preguntar el porqué de esta distribución. Para ello introdujimos modificaciones genéticas por CRISPR/Cas9 en células madre embrionarias de ratón (mESC) en el TAD de *Six3*, gen que se encuentra activo en progenitores neurales, para analizar tanto su expresión como la expresión de *Six2*, un gen que se encuentra en el TAD vecino y está inactivo en progenitores neurales. Usando el locus *Six3/Six2* como modelos, nos planteamos las siguientes preguntas:

1. ¿Es importante la posición de los genes cerca del borde de los TAD para su propia expresión? ¿Es capaz el potenciador de *Six3* de activar *Six2* una vez que se elimina la barrera? Para responder estas preguntas durante este proyecto se utilizó una línea ya generada previamente por miembros del laboratorio de Rada-Iglesias en la cual se eliminó los CBS que forman la barrera que separa los TAD de *Six3* y *Six2* (*Del* mESC). Además durante este proyecto se creó una inversión dentro del TAD de *Six3* para posicionar este gen en el medio de su TAD (*Inv* mESC).
2. ¿Es importante la posición de los genes cerca del borde de los TAD para aumentar la insulación de sus propios dominios reguladores? ¿Está el gen en el borde del TAD contribuyendo a la función insuladora? Para responder estas preguntas, durante este proyecto se eliminó *Six3* tanto en WT mESC (*KO* mESC) como en mESC donde ya se habían eliminado los CBS que forman la barrera que separa los TAD de *Six3* y *Six2* (*KO+Del* mESC).

En todas las líneas celulares anteriores se analizó la expresión tanto de *Six3* como de *Six2* tras la diferenciación de las mESC a progenitores neurales (AntNPC) siguiendo un protocolo ya establecido en el laboratorio.

2. Objetivos

1. Análisis de la expresión de *Six3/Six2* en la línea celular que contiene la barrera delecionada (*Del* mESC) tras su diferenciación a AntNPCs.
2. Generación de las líneas células en las que se elimina (*KO* mESC) o invierte (*Inv* mESC) *Six3* utilizando la tecnología CRISPR/Cas9.
3. Análisis de la expresión por qPCR de *Six3/Six2* en las líneas celulares generadas tras la diferenciación de mESC a AntNPCs.

3. Metodología

3.1. Diseño de las secuencias guías

En primer lugar, se llevó a cabo el diseño de las secuencias de las guías CRISPR/Cas9 (gRNA) utilizando la herramienta informática Benchling CRISPR (<https://www.benchling.com/crispr/>). Para ello se insertó la región de interés en la que queremos incorporar el gRNA. Dicha región se obtuvo utilizando UCSC genome browser (<https://genome.ucsc.edu/>). Una vez introducida la región de interés, se obtuvieron varias opciones de gRNA con diferentes secuencias PAM ordenadas en función de su eficiencia. Se seleccionaron aquellas con mejor eficiencia, es decir mayores valores ``off target`` y ``on target``. Con el objetivo de aumentar la eficiencia del gRNA al cortar, se añadió un nucleótido de guanina en la primera posición de la secuencia. Por último, se añadió al comienzo del gRNA la enzima de restricción *BbsI* (R0539L, NEB), realizando, además, el reverso complementario de dicha guía (**Tabla 1 y 2**).

Tabla 1. Diseño de las secuencias gRNA utilizadas para invertir *Six3*.

gRNA	Forward	Reverse
Six3 Inv A2	caccGACCCGTGAAATACCTGAC	aaacGTCAGGTATTTACGGGTC
Six3 Inv B2	caccGGCAGTTCTGAAAGTTGGG	aaacCCCAACTTTCAGAACTGCC

Tabla 2. Diseño de las secuencias gRNA utilizadas para eliminar *Six3*.

Nombre	Secuencia gRNA
F <i>Six3</i> KOA1	caccGTAATGTGCCCAACACCAT
R <i>Six3</i> KOA1	aaacATGGTGTTGGGCACATTAC
F B2 <i>Six3</i> KOA2	caccGGCAGTTCTGAAAGTTGGG
R B2 <i>Six3</i> KOA2	aaacCCCAACTTTCAGAACTGCC

3.1.1 Vector pX330A

El plásmido en el que se introdujeron las secuencias guías para su posterior expresión fue el px330A (**Figura 4**). Además de cas9, contiene los genes de resistencia a puromicina y ampicilina. El clonado del gRNA en el vector se llevó a cabo utilizando la enzima de restricción *BbsI*.

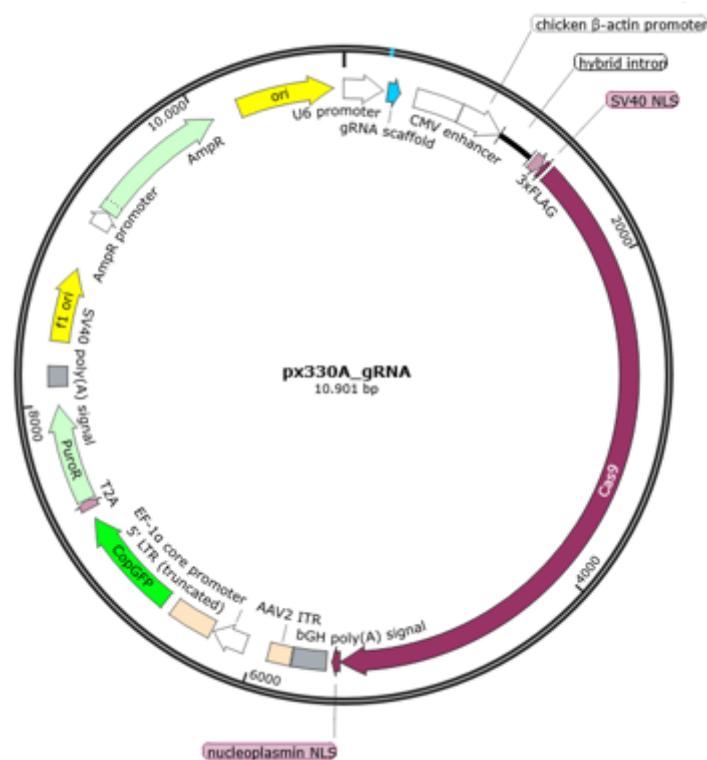


Figura 4. Esquema del vector px330A_hCas9_gRNA usado para la transfección de células embrionarias de ratón.

3.2. Clonaje del gRNA en el vector px330A

3.2.1 Digestión vector

En primer lugar, se llevó a cabo la digestión de 1 µg del vector pX330A empleando la enzima de restricción *BbsI*. La digestión se realizó preparando la mezcla de la **Tabla 3**, en un volumen total de 20 µl. La mezcla se incubó a 37°C durante 30 minutos.

Tabla 3. Mezclas de reacción para la digestión del vector px330A. La digestión del vector se llevó a cabo con el enzima de restricción Bbs I. Volumen total de 20 μ l

Compuesto	Cantidad
Vector px330A	1 μ g
BbsI	1 μ l
10 X BbsI Digestion Buffer	2 μ l
ddH ₂ O	X μ l

Total: 20 μ l

Se realizó a continuación, una electroforesis en gel de agarosa 1% (p/v) en TAE durante 20 min a 150 voltios. Bajo la luz UV se localizó y cortó la banda correspondiente al vector pX330A digerido. Posteriormente, se purificó el vector digerido siguiendo el protocolo PCR clean-up, gel extraction utilizando el kit Biotools speedtools pcr clean up (740609.50.205571, Biotools)

3.2.2 Hibridación oligos

La hibridación de los pares de oligos complementarios se realizó preparando la mezcla que se muestra en la **Tabla 4**, en un volumen total de 10 μ l.

Tabla 4. Mezcla de reacción para la hibridación de los pares de oligos. Volumen total de 10 μ l.

Compuesto	Cantidad
10 X t4 Ligation Buffer (NEB)	1 μ g
Oligo 1 (100 μ M)	1 μ l
Oligo 2 (100 μ M)	1 μ l
ddH ₂ O	7 μ l

Total: 10 μ l

A continuación, se introdujo la muestra en el termociclador estableciendo los siguientes parámetros: 37 °C durante 30 min; 95 °C durante 5 min; se descendió la temperatura a 25 °C at 5 °C min⁻¹. Se cuantificó la cantidad de ADN mediante espectrofotometría a 260 nm en un nanoDrop.

3.2.3 Ligación

Se realizó la clonación de los oligos previamente hibridados con 50 ng del vector digerido con *Bbs I*. Se preparó la mezcla recogida en la **Tabla 5**, en un volumen total de 20 μ l. Tras ello, la reacción de ligación se incubó durante 1 h a temperatura ambiente (EL0013, Thermo Fisher).

Tabla 5. Mezcla de reacción para la ligación de los oligos con el vector digerido por *Bbs I*. Volumen total de 10 μ l

Compuesto	Cantidad
Oligos hibridados (1:200)	1 μ l
Vector digerido	50 ng
10X T4 Ligasa Buffer	1 μ l
T4 Ligasa	1 ul
ddH₂O	X ul

Total: 10 μ l

3.3. Transformación celular

Se llevó a cabo la transformación de las células competentes TOP10 de la cepa *E. coli* con el plásmido recombinante. CEn primer lugar, se descongelaron en hielo las células TOP10 y se añadió el producto de la ligación obtenido en el apartado 2.3. La mezcla se incubó en hielo durante 1 min, y se introdujo en un termobloque a 37°C durante 1 min e inmediatamente después se volvió a colocar en hielo durante 1 min. Dicha mezcla se añadió a un tubo que contenía 1 ml de LB (sin antibiótico). Posteriormente, se sembró en placas de LB agar + ampicilina a 37°C durante toda la noche.

3.3.1 Colony PCR

Para comprobar la inserción correcta del gRNA se realizó una colony PCR siguiendo el protocolo de NZY Tech Green Master Mix (MB35803, Nzytech). Para ello se seleccionaron 5 colonias y se resuspendieron en 50 μ l de H₂O. De éstos 50 μ l se cogieron 5 μ l para realizar la colony PCR (Tabla 6 y 7), para ello se usó como primer Forward el respectivo "Oligo 1" y como primer Reverse "GGAAAGTCCCTATTGGCGTT", que es una secuencia que se encuentra en el vector pX330. La banda esperada en el caso de las colonias positivas es de 300 bp.

Tabla 6. Mezcla de reacción para la identificación de las colonias positivas por colony PCR. Volumen total de 25 μ l.

Compuesto	Cantidad
Green Master Mix	12,5 μ l
Primer F	0,5 μ l
Primer R	0,5 μ l
DNA	5 ul
H₂O	6,5 ul

Total: 25 μ l

Tabla 7. Programa PCR.

Paso	Temperatura	Tiempo	Ciclos
Desnaturalización inicial	95°C	3 min	1
Desnaturalización	94°C	30 sec	37
Hibridación	65 °C	30 sec	
Extensión	72°C	1 min	
Extensión final	72°C	5 min	1

3.3.2 Purificación del plásmido

Una vez obtenidas las colonias positivas (banda de 300 bp) se cogieron los 45 µl restantes de las colonias resuspendidas en agua y se añadieron a diferentes tubos con 4 ml de LB + ampicilina. Se dejaron creciendo durante la noche a 37°C en movimiento.

Al siguiente día se llevó a cabo la purificación del plásmido usando el kit comercial NZY Tech Miniprep Kit (MB01001, Nzytech). Tras realizar la purificación se cuantificó la concentración de la muestra usando Nanodrop y se envió a secuenciar mediante el método Sanger, utilizando el mismo primer reverse que se utilizó para realizar la colony PCR. Una vez confirmado que el inserto del gRNA es correcto se prosiguió con la transfección.

3.4. Mantenimiento de las mESC

En el sistema CRISPR/Cas9 se utilizaron células madre embrionarias de ratón (mESC) cultivadas en placas recubiertas de gelatina. Dichas células adherentes, se cultivaron en medio KnockOut™ DMEM (10829018, ThermoFisher) suplementado con 15% de suero bovino fetal (10500064, ThermoFisher Scientific) y factor inhibitorio de la leucemia (LIF). El LIF fue añadido con el objetivo de mantener las células en un estado de pluripotencia. En el medio de cultivo también se añadieron antimicóticos y antibióticos (A5955-100ML, Sigma-Aldrich) con el objetivo de evitar contaminaciones, β-mercaptoetanol (21985023, ThermoFisher Scientific), Glutamax (35050038, ThermoFisher Scientific) y MEM NEAA (11140035, ThermoFisher Scientific). Las células fueron cultivadas a 37°C con un 5% de CO₂ atmosférico.

3.4.1 Líneas de mESC transgénicas

Para este proyecto se usaron mESC E14 a partir de las cuáles se generaron varias líneas:

Tabla 8. Líneas celulares utilizadas en el proyecto

Línea celular	Referencia
WT (E14) mESC	Wysocka Laboratory
Delección Barrera Six3 mESC (<i>Del</i>)	Laboratorio Rada-Iglesias (Ya disponible)
Six3 -/- mESC (<i>KO</i>)	Generada durante este proyecto
Six3 -/- + Delección barrera (<i>KO+Del</i>)	Generada durante este proyecto
Inversión Six3/Enhancer (<i>Inv</i>)	Generada durante este proyecto

3.4.2 Transfección de las mESC

La transfección de las mESC con los vectores de expresión de los gRNAs (**Tabla 1 y 2**) se llevó a cabo siguiendo las recomendaciones del kit Lipofectamine 3000 (L3000001, Thermo Fisher). Se usaron 300 ng de cada vector. Tras 24 horas se comprobó si el plásmido se había incorporado en las células, para ello se analizó la señal de GFP. Al observar colonias verdes significa que la transfección ha funcionado por lo que se continuó con la selección de puromicina (1 ul 1000x puromicina/1 ml S+L). Una vez que la selección con puromicina terminó (las células del control están muertas) se obtuvo una población heterogénea con diferentes alteraciones genéticas. Se dejaron a las células recuperarse tras la transfección y se procedió a la extracción de DNA.

La extracción se realizó utilizando Lysis Buffer, compuesto de 25 mM KCl (27810.295, Sigma-Aldrich), 5mM TRIS (0497-5KG, Sigma-Aldrich) pH8.3, 1.25mM MgCl₂ (BDH7899-1, VWR), 0.225% IGEPAL (I8896-50ML, Sigma-Aldrich) y 0.225% Tween20 (VWR). Antes de usarlo se añadió proteinasa K (EO0492, ThermoFisher Scientific) a una concentración final de 0,4 ug/ul.

Tras tripsinizar las células con el objetivo de despegarlas de la placa, se pasaron a un tubo de 1,5 ml y se centrifugaron a 1000 rpm durante 3 minutos. Posteriormente, se eliminó el sobrenadante y el pellet se resuspendió en Lysis Buffer suplementado con proteinasa k. La mezcla se incubó a 65°C durante 6 minutos y para inactivar la proteinasa k se incubó a 98°C durante 2 minutos. La concentración de la muestra fue medida en Nanodrop y se utilizaron 100ng de DNA genómico para cada reacción de PCR.

3.4.3 Genotipado de la población

Con el objetivo de verificar si la población contenía la modificación deseada, primeramente, se extrajo DNA genómico como se describe anteriormente. La secuencia de la región de *Six3* se obtuvo utilizando UCSC genoma browser y seleccionando el genoma de ratón *mm9*. Del mismo modo se seleccionó la región de la delección o inversión deseada. Basándonos en dichas secuencias

se diseñaron los correspondientes primers (Ver Tablas 9 y 10). El master mix de la PCR y el programa utilizado se describe en las siguientes tablas (Tabla 11 y 12 respectivamente). Los productos de PCR fueron cargados en un gel de agarosa 1,5% en el cual se añadió GreenSafe Premium (NZYTech) para detectar la doble hebra de DNA debajo de la luz UV. Para estimar los tamaños de los productos se usó el Ladder VII de NZYTech.

Tabla 9. Secuencia de los primers utilizados para genotipar la población transfectada con los gRNA *Six3* Inversión, así como los tamaños correspondientes.

Primer	Forward 5'3'	Reverse 5'3'	Tamaño (bp)
WT (1+2)	TCTTGGTGCCATGGATGACT	CTGGAAGGGAAGGAGCAGAA	620
WT (3+4)	TGACATCAGGAACAGGGACC	CATCCCATGCGACTTGACTG	963
INV (1+5)	TCTTGGTGCCATGGATGACT	GCAGCCCACATTTCTGACAA	729
INV (2+6)	CTGGAAGGGAAGGAGCAGAA	AGGCCAAGCATACAGACTGT	684
DEL (1+6)	TCTTGGTGCCATGGATGACT	AGGCCAAGCATACAGACTGT	251

Tabla 10. Secuencia de los primers utilizados para genotipar las poblaciones transfectadas con los gRNA *Six3* KO, así como los tamaños correspondientes.

Primer	Forward	Reverse	Tamaño
Delección (1+2)	AAACACAACACAGCGACTCC	AGGCCAAGCATACAGACTGT	852 bp
Control WT (3+4)	CAGGGTCAGGAAGGAGAAGG	CTTCCACCTCCACCTCCAT	203 bp

Tabla 11. Mezcla de reacción para la identificación de las colonias positivas por colony PCR. Volumen total de 25 μ l.

Compuesto	Cantidad
Green Master Mix	12,5 μ l
Primer F	0,5 μ l
Primer R	0,5 μ l
DNA	100 ng
H ₂ O	X

Total: 25 μ l

Tabla 12. Programa PCR

Paso	Temperatura	Tiempo	Ciclos
Desnaturalización inicial	95°C	3 min	1
Desnaturalización	94°C	30 sec	37
Hibridación	65 °C	30 sec	
Extensión	72°C	1 min	
Extensión final	72°C	5 min	1

3.4.4 Derivación líneas clonales mESC

Una vez confirmado que las modificaciones deseada estaba presente en la población heterogénea, se procedió a obtener clones homocigotos para dichas modificaciones. Para ello se preparó una placa de 96 pocillos cubierta con gelatina durante toda la noche. Las células de la población se tripsinizaron 5 minutos a 37°C y se añadió el doble de cantidad de medio S+L para inactivar la tripsina. Se resuspendieron varias veces para asegurarnos de que las células se encontraban individuales y no formando grupos.

A continuación, se hicieron una serie de diluciones con el objetivo de tener 1 única célula en cada pocillo. Las células se dejaron crecer durante 5 días sin cambiar el medio. Después se marcaron aquellos pocillos que contenían una célula y se genotiparon siguiendo las mismas condiciones que las **tablas 9 y 10**. Una vez caracterizados los clones se expandieron y se congelaron las células seleccionadas.

3.5. Diferenciación de mESC a AntNPC

Para llevar a cabo la diferenciación a progenitores neurales se plaqueron 8000 células/ cm² en placas recubiertas con geltrex durante 1 hora en medio N2B27 suplementado. Para la suplementación del medio de diferenciación N2B27 se siguieron las recomendaciones establecidas previamente en el protocolo (Cruz-Molina et al., 2017) con algunas modificaciones. Cada día se suplementó el medio con los siguientes componentes: β -FGF (PHG0368, ThermoFisher Scientific) 10 ng/ml de día 0 a día 2, Xav939 (X3004-5MG, Sigma-Aldrich) 5 μ M de día 2 a día 6, BSA (MB04602, ThermoFisher Scientific) 1mg/ml a día 0 y 40 μ g/mL el resto de días de la diferenciación (**Tabla 13**).

Tabla 13. Suplementos añadidos cada día de la diferenciación a progenitor neural.

D0	D1	D2	D3	D4	D5	D6
bFGF (1X) BSA (1mg/ml)	bFGF(1X) BSA (40ug/ml)	bFGF (1X) Xav (1X) BSA (40ug/ml)	Xav (1X)	Xav (1X)	Xav (1X)	Collect RNA

La activación de la señal β -FGF durante los primeros días de diferenciación nos permitieron obtener células progenitoras neurales (Gouti et al., 2014) y al añadir Xav939, un inhibidor de la señal de Wnt, se favoreció la homogeneidad de la diferenciación enriqueciendo destinos ectodérmicos (Matsuda et al., 2014).

3.5.1 Extracción RNA

La extracción del RNA se llevó a cabo usando el kit NZY Total RNA Isolation (MB13402, NzyTech) siguiendo las instrucciones del fabricante.

3.5.2 Síntesis de cDNA

Para realizar análisis de expresión génica, el ARN se transcribió inversamente en cDNA utilizando el kit NZY First-Strand cDNA Synthesis Kit (MB12502, NZYTech). Por cada reacción, 1 ug de RNA se incubó con 10 μ L de NZYRT 2x Master Mix, 2 μ L de NZYRT Enzyme Mix y nuclease-free water hasta un volumen total de 20 μ L a 25°C durante 10 minutos, seguido

de 30 minutos a 50 °C. La enzima se inactivó por calor a 85°C durante 5 minutos. Para digerir el resto de RNA se añadió 1µL of NZY RnaseH a la reacción y se incubó a 37°C durante 20 minutos.

3.5.3 RT-qPCR

Para analizar los niveles de expresión de los genes de interés, se extrajo el RNA y sintetizó el cDNA como se describe previamente. En primer lugar, las muestras de cDNA se diluyeron 1:3 en dH₂O y 0,4 µL de esa dilución se usaron para cada 10 µL de reacción, siguiendo el kit NZYSpeedy qPCR Green Master Mix (2x), ROX plus (NZYTech) (**Tabla 14**). En la **tabla 15** se describe el programa de qPCR utilizado.

Tabla 14. Reacción RT-qPCR. 1X (10µl)

Compuesto	Cantidad
NZYSpeedy qPCR Green Master Mix (2X)	5µl
ROX plus (NZY Tech)	
10µM F, R primer Mix	0.125µl
cDNA dilution	0,4µl
H ₂ O	4.475µl

Tabla 15. Programa de amplificación usado en las reacciones de RT-qPCR

Paso	Temperatura	Tiempo
Desnaturalización inicial	50.0	2 min
	95.0	2 min
40 Ciclos	95.0	15 sec
	60.0	15 sec
	72.0	1 min
	95.0	15 sec
Curvas melting	60.0	1 min
	95.0	15 sec

Todas las reacciones se llevaron a cabo realizando triplicados técnicos, y la desviación estándar se representó como barras de error. Los cebadores fueron diseñados con el fin de amplificar regiones que cubrían la unión exón-exón (**Tabla 16**). Gracias a este diseño, se asegura la amplificación de cDNA desde ARN mensajero maduro. Los genes *Eef1a* y *Hprt1* se utilizaron

como controles de cargado. Para llevar a cabo el cálculo de la expresión de cada gen se normalizaron sus valores Ct frente a los genes controles (WT D6) utilizando la fórmula $2^{-\Delta\Delta CT}$.

Tabla 16. Cebadores usados para analizar la expresión génica por RT-qPCR

Cebadores	Secuencia
Hprt1 F	CAAGGGCATATCCAACAACA
Hprt1 R	GCCCCAAAATGGTTAAGGTT
Eef1a F	TAGACGAGGCAATGTTGCTG
Eef1a R	AGCGTAGCCAGCACTGATTT
Six3 F	CCTCACCCCCACACAAGTAG
Six3 R	CTGATGCTGGAGCCTGTTCT
Six2 F	GCAAGTCAGCAACTGGTTCA
Six2 R	CTTCTCATCCTCGGAACTGC

4. Resultados

4.1. Comprobación de la diferencia en la expresión de Six3/Six2 entre las línea WT y la línea con la delección de la barrera Six3/Six2 tras la diferenciación de mESC a AntNPCs.

Six3 y *Six2* son dos genes vecinos que se encuentran separados por una barrera de TAD conservada y con varios CBS. Durante la embriogénesis ambos genes muestran patrones de expresión diferentes, *Six3* se expresa en el cerebro y se requiere para la formación del telencéfalo y del diencefalo, actuando además en la proliferación y diferenciación de células progenitoras neurales a través de la activación de varios genes, mientras que *Six2* es un factor de transcripción necesario el desarrollo de distintos órganos como es el caso de riñones, cráneo y estómago.

Por otro lado, *Six3* contiene un potenciador a una distancia de 146 kb que controla su inducción en AntNPCs sin ningún efecto en *Six2* (situado a 218 kb del potenciador de *Six3*). Es por ello que se llevó a cabo una delección de 36-kb que incluye varios CBS para eliminar la barrera que separa los dos TADs en los que se encuentran *Six3* y *Six2* (línea generada previamente por los miembros del laboratorio de Rada-Iglesias denominada más adelante como Del) con el objetivo de ver si tras la delección, el potenciador es capaz de inducir *Six2* y además, observar qué ocurre con la inducción de *Six3* al dejar de estar lado de la barrera (Pachano et., 2021) (**Figura 5**). El protocolo de diferenciación gracias al cual las células se diferencian de mESC a AntNPC se basa en dirigir las células a las primeras etapas del desarrollo del cerebro. Como resultado, tras 6 días de diferenciación, las células son principalmente progenitores neurales anteriores y dorsales.

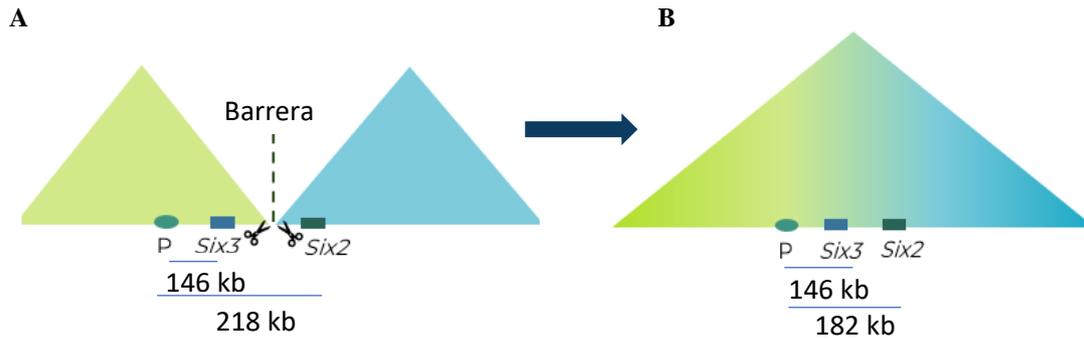


Figura 5. Representación esquemática de los TADs de *Six3/Six2* antes (A) y después de eliminar la barrera (B) (36kb) con los respectivos gRNA representados con unas tijeras.

Tras la diferenciación a AntNPCs, *Six3* se ve afectado y su expresión se redujo su expresión en 2,5 fold vs. WT (**Figura 6.A**), mientras que *Six2* se indujo en la línea DEL (6-fold vs. WT) (**Figura 6.2**). Estos resultados confirman la función de aislamiento de los CTCF que forman la barrera, ya que bloquean el mecanismo activo por el cual la cohesina forma los bucles (Chang et al., 2019). Por su parte la bajada de la expresión de *Six3* podría venir por el hecho de que existen reguladores negativos en el TAD de *Six2* que tras la eliminación de la barrera son capaces de actuar sobre *Six3* (Batut et al., 2022) o bien porque el alejar a *Six3* de la barrera del TAD afecta negativamente a su expresión.

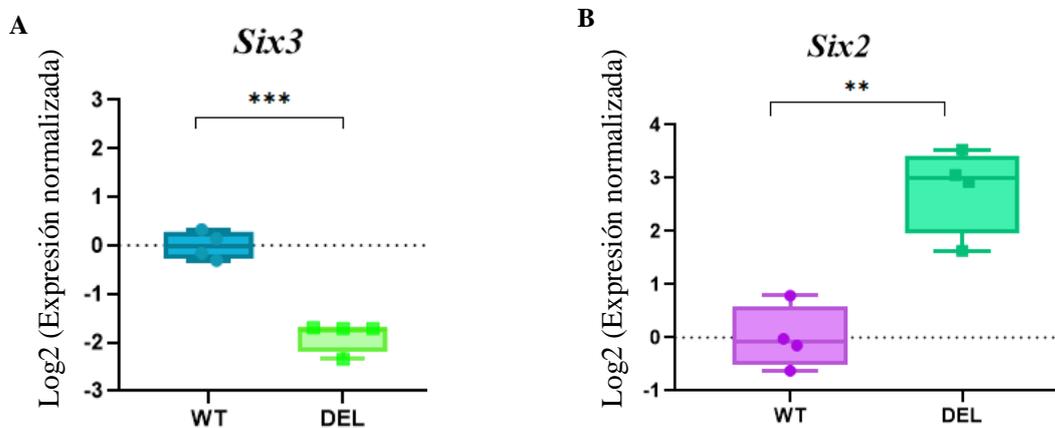


Figura 6. Expresión de *Six3* y *Six2* medida por RT-qPCR tras la diferenciación a AntNPCs. En ambos casos los valores están normalizados frente al WT (Día 6). A) Tras la diferenciación, en la línea DEL (verde) la expresión de *Six3* baja significativamente (p-value: 0,0001) al normalizarlo frente al WT (azul). B) Con respecto a *Six2*, se observa un aumento de su expresión significativo tras eliminar la barrera (p-value: 0,0014). Análisis estadísticos realizados usando un análisis t-test. *** representa fold change ≥ 2 y p-value $< 0,0005$. **representa fold change ≥ 2 y p-value $< 0,0005$.

4.2. Generación de la línea Six3 Inversión usando la tecnología CRISPR/Cas9

Tras observar la bajada de expresión de *Six3* en la línea en la que se encontraba deletionada la barrera, surgió la duda de si esta bajada es debida a que hay elementos reguladores negativos (i.e. silenciadores) en el TAD de *Six2* actuando sobre *Six3* o es por el hecho de que la expresión de *Six3* se ve favorecida por la proximidad a la barrera. Es por ello por lo que decidimos hacer una inversión en el TAD de *Six3* dejando la barrera intacta y mediante la cual alejamos a *Six3* de la barrera mientras que la posición relativa del potenciador se mantuvo invariable (**Figura 7**). En las condiciones WT, *Six3* se encuentra a 8 kb de la barrera, sin embargo, después de la inversión, éste se encuentra en el medio del TAD a 231 kb de la barrera. Por su parte, el potenciador que originalmente se encontraba a 164 kb de la barrera, ahora se encuentra a 106 kb por lo que apenas se modificó la distancia. De esta manera ni *Six3* ni el potenciador estarán cerca de la barrera, la cual se encuentra intacta.



Figura 7. Representación esquemática de la inversión de *Six3* dentro de su propio TAD. Dicha modificación se realizó mediante dos gRNA (mostrados como tijeras) en la línea WT ESC.

4.2.1 Genotipado Six3 INV

Después de la transfección con las guías A2 y B2 (**Tabla 1**) se diseñaron diferentes primers (**Tabla 9**) para poder amplificar tanto las inversiones, como el WT y la delección (**Figura 8**), se comprobó por PCR que la población contenía la inversión de *Six3* (**Figura 9**).

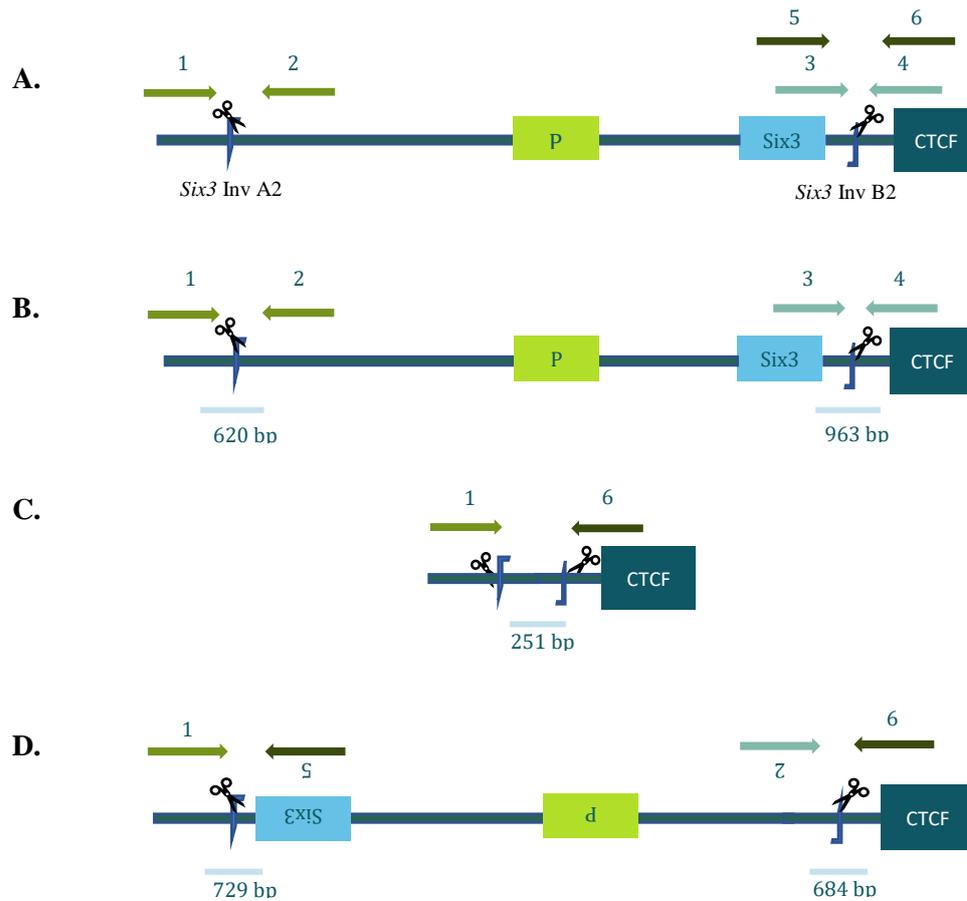


Figura 8. A. Representación esquemática de los primers utilizados para genotipar la población que contiene la inversión de *Six3*. **B.** Primers utilizados para detectar el alelo WT. Una pareja de primers se diseñaron a cada de lado de los gRNA. La pareja 1-2 situada a cada lado del gRNA “*Six3* Inv A2” amplifica en caso de existir un alelo WT, dando una banda de 620 bp, por su parte, la pareja 3-4 situada a cada lado del gRNA “*Six3* Inv B2” amplifica también el alelo WT dando una banda de 963 bp **C.** Con el objetivo de detectar una posible delección, se utilizaron los primers 1-6, los cuales dan una banda de 251 bp **D.** Las parejas de primers 1-5 y 2-6 se usaron para identificar la inversión deseada, dando productos de 729 y 684 bp respectivamente.

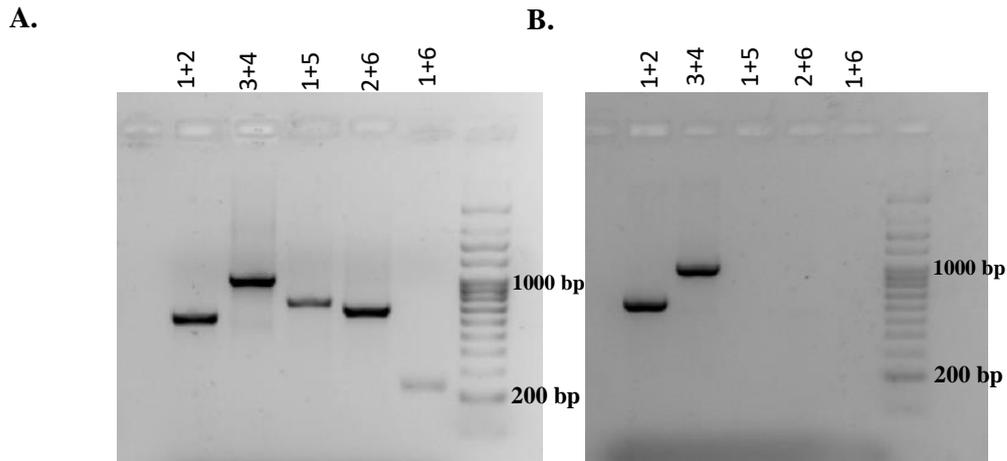


Figura 9.A. Genotipado de la población transfectada para generar la línea *Six3* Inversión (Inv). Tras realizar la transfección con el objetivo de obtener la inversión de *Six3*, se genotipó la población. Como se muestra en la **Figura 10** las combinaciones 1+2 y 3+4 se refieren a la banda WT (620 y 963 bp respectivamente), las combinaciones 1+5 y 2+6 a las bandas de la inversión (729 y 684 bp respectivamente), y, por último, la delección se detectó utilizando la combinación de primer 1+6 (251 bp). **B.** Las mismas combinaciones de primers se utilizaron en el WT, obteniendo banda solo en las parejas de primers diseñadas para amplificar el alelo WT.

Una vez comprobado por PCR que la población contenía la inversión (**Figura 10**), se plaqueron las células en una placa de 96 pocillos con el objetivo de derivar líneas clonales. Una vez obtenidos los clones, se genotiparon obteniendo 2 clones homocigotos.

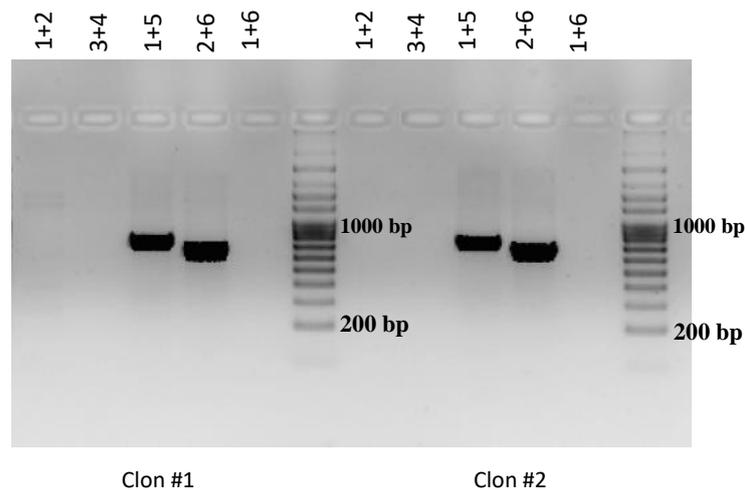


Figura 10. Genotipado clones homocigotos para la línea *Six3* Inversión. Se confirmaron ya que sólo presentan banda en las combinaciones de primers 1+5 y 2+6 las cuales corresponden a las inversiones.

4.2.2 Análisis de la expresión por qPCR de *Six3*/*Six2* tras la diferenciación de Inv mESC a AntNPC.

Tras la diferenciación a progenitor neural se analizó la expresión tanto de *Six3* como *Six2* en la línea de la inversión y en WT. Tras normalizar los valores frente a WT (D6) no se observó ningún cambio significativo para ninguno de los dos genes (**Figura 11**). Estos resultados sugieren el hecho que la posición de *Six3* cerca del borde del TAD no favorece de forma significativa su propia expresión. Con respecto a *Six2*, tras la inversión su expresión se mantiene igual que en la línea WT ya que la barrera de CTCF se sigue manteniendo intacta.

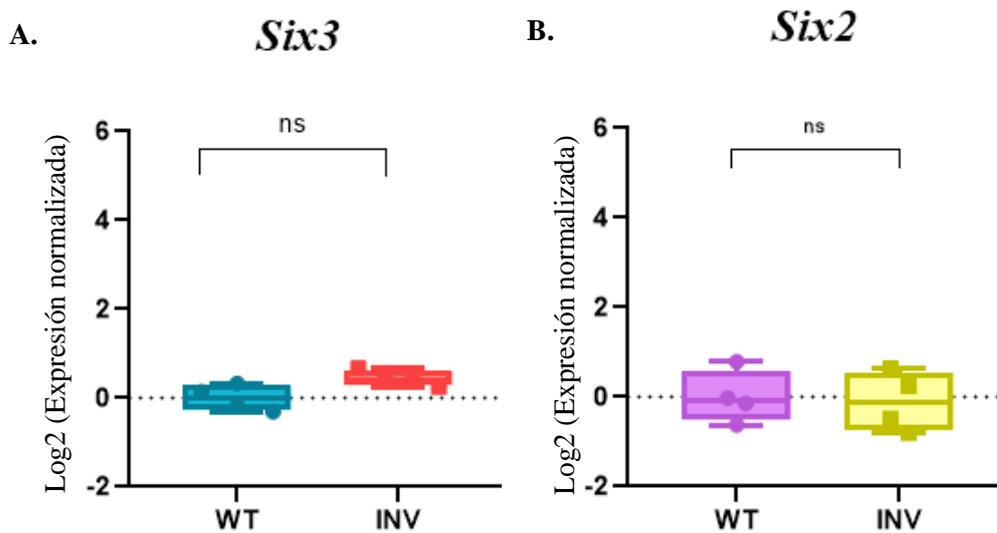


Figura 11. Expresión de *Six3* y *Six2* medida por RT-qPCR tras la diferenciación a AntNPCs. Valores normalizados frente al WT (Día 6). Tras la inversión, tanto la expresión de *Six3* como de *Six2* es similar a la de las células WT. Análisis estadísticos realizados usando un análisis t-test. n/s representa fold change <2 y p-value >0,1.

4.3. Generación de las líneas *Six3*^{-/-} usando la tecnología CRISPR/Cas9

Posteriormente, con el objetivo de determinar si la presencia de *Six3* al lado de la barrera contribuye a la función aisladora para prevenir activaciones ectópicas en el TAD vecino (i.e. *Six2*), se eliminó este gen en la línea celular mESC WT y en la línea que contenía la barrera ya delecionada (**Figura 12**).

Para conseguir eliminar este gen se utilizó la tecnología CRISPR/Cas9 basada en el diseño de dos gRNAs (**Tabla 2**) encargados de realizar un corte de doble hebra a cada lado de *Six3* (22 kb).

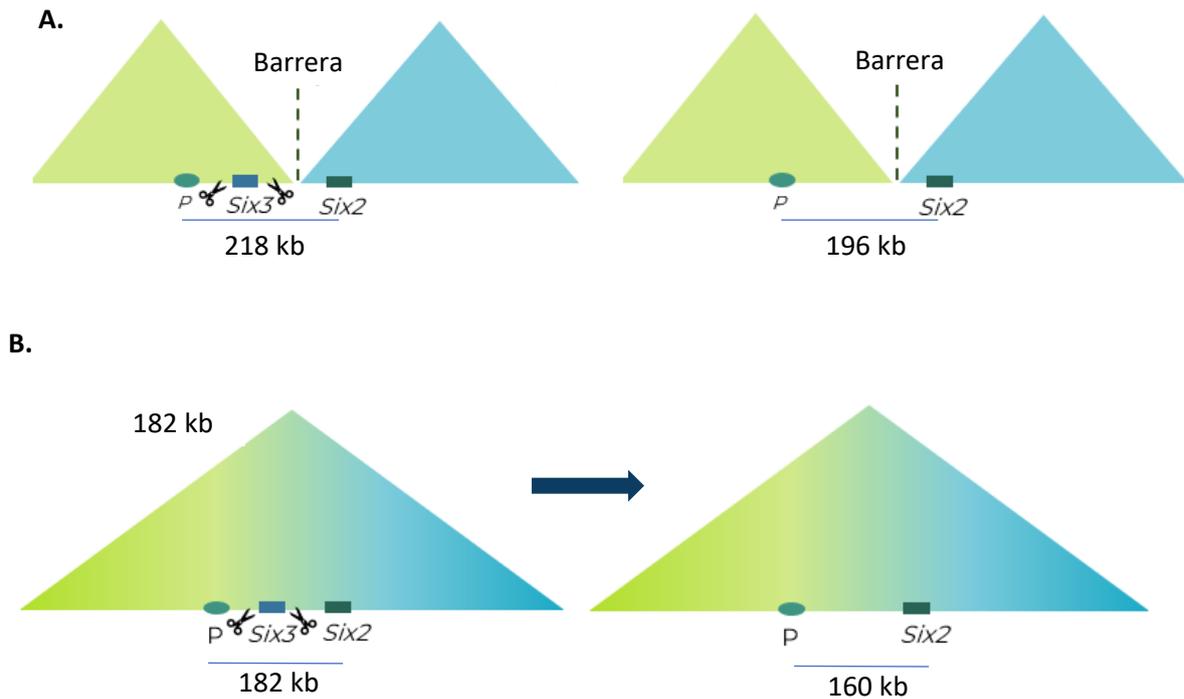


Figura 12. Representación esquemática de la eliminación de *Six3* mediante dos gRNA (mostrados como tijeras) tanto en la línea WT (A) como en la línea DEL (B).

4.3.1 Genotipado *Six3* KO

Después de la transfección, se diseñaron una pareja de primers fuera de *Six3*, que sólo en el caso de que el gen esté deletionado conseguirían amplificar (852 bp), así como unos primers internos que se encuentran dentro de *Six3* que solo amplifican en ausencia de la delección (203 bp) (**Tabla 10**) (**Figura 13**). Se comprobó por PCR que tanto la población de mESC WT como Del contenían la delección de *Six3* (**Figura 14**).

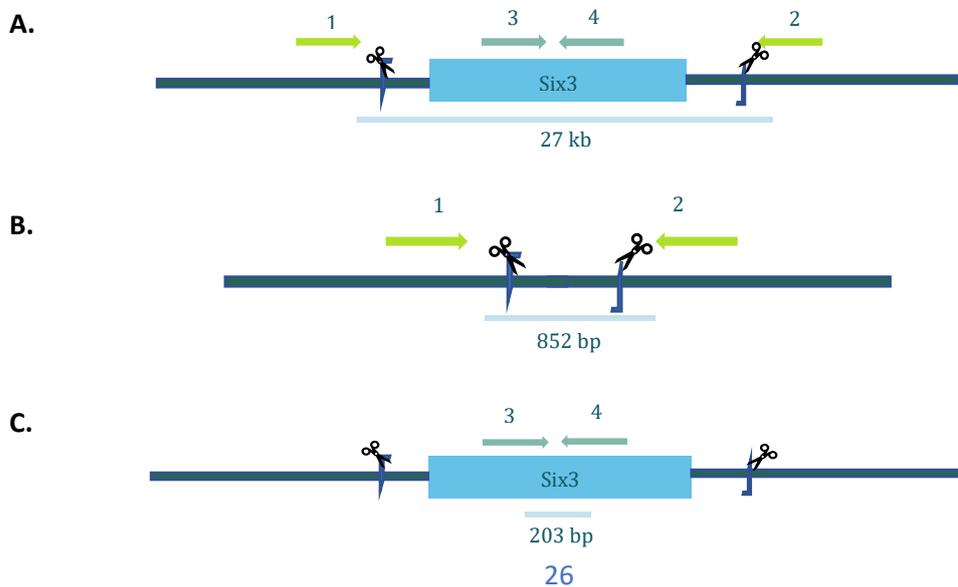


Figura 13. A. Representación esquemática de los primers utilizados para genotipar las dos poblaciones en las que se ha eliminado *Six3*. **B.** Pareja de primers utilizados para confirmar la delección de *Six3* (852 bp). **C.** Pareja de primers utilizados para la detección de WT (203 bp).

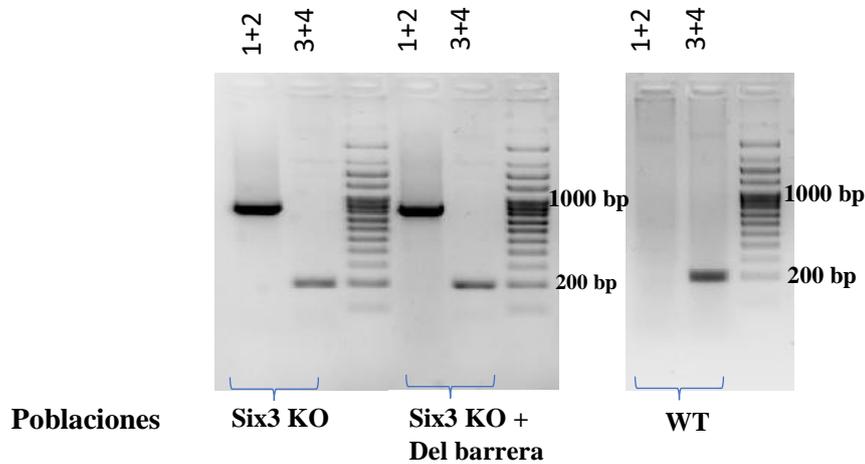


Figura 14. Para la identificación de la delección de *Six3* en ambas poblaciones, se utilizaron los primers 1+2 y como control se utilizaron los primers 3+4. Los resultados de la PCR obtenidos de las 2 líneas y del WT se muestran en la figura.

Una vez comprobado por PCR que ambas poblaciones tenían la delección deseada (**Figura 14**), se plaquero las células en una placa de 96 pocillos con el objetivo de derivar líneas clonales. Una vez obtenidos los clones, se genotiparon obteniendo 3 clones homocigotos para cada población (KO *Six3* y KO *Six3* + Delección barrera) (**Figura 15**).

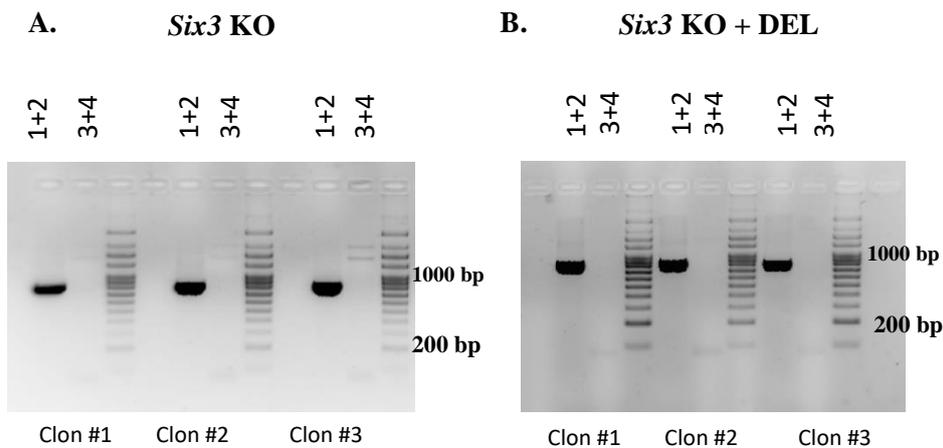


Figura 15. Líneas clonales en las que se ha eliminado *Six3*. A) Clones homocigotos para la delección de *Six3*. B) Clones homocigotos para la delección de *Six3* en la línea DEL. Identificados usando los primers 1+2 que flanquean *Six3* (banda de 852 bp) así como los primers internos 3+4.

4.3.2 Análisis de la expresión por qPCR de *Six3*/*Six2* tras la diferenciación de mESC a AntNPC.

Tras la diferenciación a progenitores neurales se analizó la expresión tanto de *Six3* como *Six2* en células WT, KO y KO+Del (**Figura 16**). Como era de esperar, en la línea KO hay ausencia de expresión de *Six3*. (**Figura 16.A**). Como se describe previamente, tras eliminar la barrera de CTCF que actuaba como aislante (Línea DEL) la expresión de *Six2* aumentó (p-value: 0,0014) en comparación con el WT. Tras deleccionar *Six3*, la expresión de *Six2* no aumenta significativamente (ns), por lo que, aunque se elimine el gen del borde, mientras siga estando la barrera intacta, el potenciador no es capaz de activar *Six2*. Sin embargo, sorprendentemente, tras eliminar *Six3* a la vez que la barrera (KO+DEL) la expresión de *Six2* incrementa de forma significativa y sinérgica en comparación a la línea en la que está eliminada la barrera (DEL) (p-value: <0,0001) Estos resultados sugieren que *Six3* se encuentra en el borde del TAD contribuyendo a la barrera de insulación junto con los CBS. Además, la función insuladora de *Six3* parece ser epistática, ya que solo se manifiesta en ausencia de los CBS.

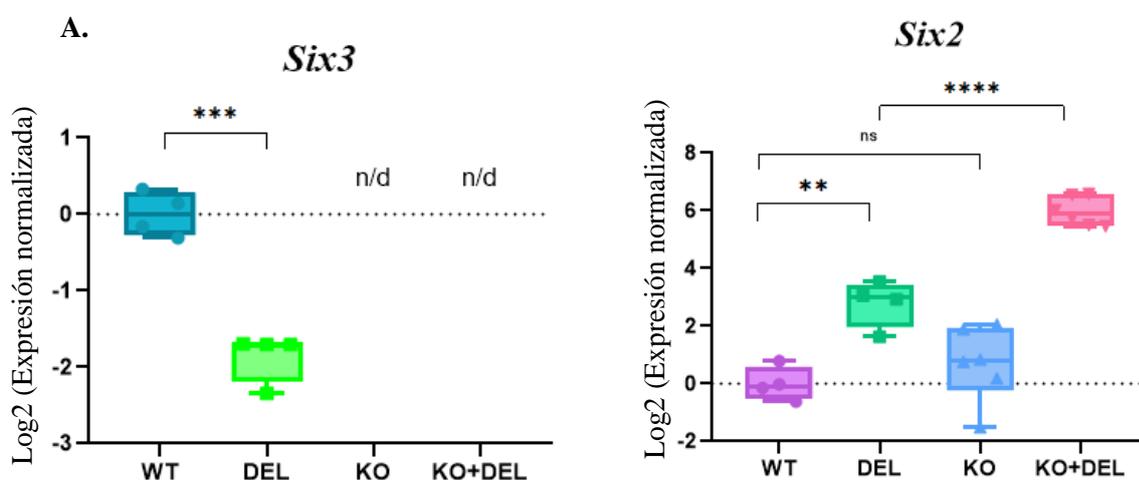


Figura 16. Expresión de *Six3* y *Six2* medida por RT-qPCR tras la diferenciación a AntNPCs. Valores normalizados frente al WT (Día 6). Como se muestra en la **Figura 6**, tras la diferenciación, en la línea de la DEL (verde) la expresión de *Six3* baja significativamente (p-value: 0,0001) al normalizarlo frente al WT (azul). En el caso de KO y KO+DEL, la expresión de *Six3* es no detectable. B) Con respecto a *Six2*, se observa un aumento de su expresión significativo tras eliminar la barrera (p-value: 0,0014). Al eliminar sólo *Six3* su incremento no es significativo, sin embargo al combinar las dos cosas a la vez (DEL + KO) la expresión de *Six2* aumenta significativamente en comparación a la DEL (p-value: <0,0001). Análisis estadísticos realizados usando un análisis t-test. *** representa fold change ≥ 2 y p-value <0,0002. **representa fold change ≥ 2 y p-value <0,0021, **** representa fold change ≥ 2 y p-value <0,0001.

5. Discusión

Entender como cambios en la comunicación entre los potenciadores y los genes de desarrollo dentro de los TADs puede afectar a los patrones de expresión génica puede tener importantes repercusiones médicas, ya que cambios en dicha comunicación debido a variantes estructurales es una causa frecuente de enfermedades congénitas (Lupiañez et al., 2016 - Spielmann y Mundlos, 2013). Siguiendo los modelos actuales, las proteínas aislantes (CTCF) delimitan los límites de los TADs restringiendo la actividad de los potenciadores para que actúen en genes ubicados dentro de sus mismos TADs (Kikuta et al., 2007). Durante este proyecto confirmamos este modelo ya que tras eliminar la barrera de CTCF que separa *Six3* (activo en AntNPCs) de *Six2* (inactivo en AntNPCs), el potenciador de *Six3* es capaz de activar de manera significativa a *Six2* en AntNPC (Figura 6).

Fudenberg et al., 2016 describieron la formación de los TADs como un mecanismo de extrusión de bucles de cromatina en el que un anillo de cohesina se une de manera aleatoria a la cromatina y comienza a crear un bucle hasta que se encuentra con dos dominios de CTCF ubicados de manera convergente. En relación con este modelo un estudio reciente de Edward et al., 2022 demuestran que la cohesina también se acumula en sitios de inicio de transcripción activos gracias a la presencia de la RNA polimerasa la cual actúa como barrera. Este modelo concuerda con la organización de los genes Housekeeping en los bordes de los TADs, los cuales se encuentran siempre activos (Evonne y John 2021). Sin embargo, gracias al análisis bioinformático de un miembro del laboratorio Rada-Iglesias descubrimos que los genes de desarrollo mantienen una distribución similar dentro de los TAD que los genes *housekeeping*, es decir, se localizan cerca los bordes de los TADs. Sorprendentemente, tras la realización de este proyecto podemos confirmar que *Six3* se encuentra en el borde del TAD contribuyendo de manera epistática junto a los CBS a la insulación de su propio dominio regulador, ya que tras eliminar *Six3* en la línea que ya tenía delecionada la barrera, la expresión de *Six2* aumentó significativamente en comparación a la línea DEL. Sin embargo, tras eliminar *Six3* en la línea WT, *Six2* se mantiene inactivo. Estos resultados sugieren la existencia de una relación de sinergia entre el gen y los sitios de CTCF para crear una barrera de insulación robusta. En este proyecto también descubrimos que tras realizar una inversión y localizar el gen en el medio del TAD su expresión no se ve afectada lo cual indica que al gen no le beneficia de manera significativa estar al lado de la barrera del TAD. Esto apoya también que el principal motivo del posicionamiento de los genes de desarrollo cerca de los bordes de los TADs es reforzar el efecto barrera/insulador de los CBS. A nivel mecanístico, la presencia de RNA polimerasa una vez que *Six3* está activo y/o los complejos proteicos unidos a las numerosas islas de CpG que rodean a este gene pueden bloquear el avance de la cohesina, si bien esta posibilidad debe aún explorarse.

Tras estos interesantes resultados determinaremos si existen contactos físicos entre el potenciador de *Six3* y *Six2* tanto en la línea DEL, como DEL+ *Six3* KO utilizando la técnica de Micro-Capture-C (MCC), un nuevo método 3C de alta resolución (Hua et al., 2021). Por otra parte, sería interesante realizar experimentos de ChIP-seq para observar la distribución de la cohesina ya que si *Six3* está actuando como barrera se observaría cierta acumulación de cohesina en el gene y tras eliminar tanto *Six3* como la barrera de CTCF dicha cohesina debería comenzar a acumularse en el TAD de *Six2*. Por último, también sería interesante realizar los mismos experimentos en otros loci diferente para comprobar el posicionamiento de los genes de desarrollo cerca de los bordes de los TAD contribuyen a la insulación de sus dominios reguladores de forma genérica. Si esto se corrobora, nuestros hallazgos pueden ayudar entender y predecir los posibles efectos funcionales y patológicos de variantes estructurales que ocurren dentro de los TAD y cuyos mecanismos de acción aún se desconocen en gran medida.

6. Conclusiones

En resumen, mi proyecto fin de máster representa un hallazgo importante para el entendimiento de la organización tridimensional de los genes de desarrollo dentro de los TADs. En este proyecto se ha confirmado 1) la pérdida de la barrera de CTCF provoca la activación ectópica de *Six2*. 2) *Six3* se encuentra en el borde del TAD actuando de manera epistática junto a CTCF en la insulación de su dominio regulador. 3) La expresión de *Six3* no se ve afectada tras posicionar el gen lejos del borde de su TAD.

7. Bibliografía

Batut PJ, Bing XY, Sisco Z, Raimundo J, Levo M, Levine MS. Genome organization controls transcriptional dynamics during development. *Science*. 2022 Feb 4;375(6580):566-570. doi: 10.1126/science.abi7178. Epub 2022 Feb 3. PMID: 35113722.

Bergman DT, Jones TR, Liu V, Ray J, Jagoda E, Siraj L, Kang HY, Nasser J, Kane M, Rios A, Nguyen TH, Grossman SR, Fulco CP, Lander ES, Engreitz JM. Compatibility rules of human enhancer and promoter sequences. *Nature*. 2022 May 20. doi: 10.1038/s41586-022-04877-w. Epub ahead of print. PMID: 35594906.

Busslinger, G., Stocsits, R., van der Lelij, P. *et al.* Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* 544, 503–507 (2017).

Chang LH, Ghosh S, Noordermeer D. TADs and Their Borders: Free Movement or Building a Wall? *J Mol Biol.* 2020 Feb 7;432(3):643-652. doi: 10.1016/j.jmb.2019.11.025. Epub 2019 Dec 27. PMID: 31887284.

Claringbould A, Zaugg JB. Enhancers in disease: molecular basis and emerging treatment strategies. *Trends Mol Med.* 2021 Nov;27(11):1060-1073. doi: 10.1016/j.molmed.2021.07.012. Epub 2021 Aug 20. PMID: 34420874.

Cruz-Molina S, Respuela P, Tebartz C, et al. PRC2 Facilitates the Regulatory Topology Required for Poised Enhancer Function during Pluripotent Stem Cell Differentiation. *Cell Stem Cell.* 2017;20(5):689-705.e9. doi:10.1016/j.stem.2017.02.004

Dekker J, Mirny L. The 3D Genome as Moderator of Chromosomal Communication. *Cell.* 2016 Mar 10;164(6):1110-1121. doi: 10.1016/j.cell.2016.02.007. PMID: 26967279; PMCID: PMC4788811.

Dixon JR, Gorkin DU, Ren B. Chromatin Domains: The Unit of Chromosome Organization. *Mol Cell.* 2016;62(5):668–680. doi: 10.1016/J.MOLCEL.2016.05.018

Edward J. Banigan, Wen Tang, Aafke A. van den Berg, Roman R. Stocsits, Gordana Wutz, Hugo B. Brandão, Georg A. Busslinger, Jan-Michael Peters, Leonid A. Mirny. Transcription shapes 3D chromatin organization by interacting with loop-extruding cohesin complexes. *bioRxiv* 2022.01.07.475367; doi: <https://doi.org/10.1101/2022.01.07.475367>
Eisenberg E, Levanon EY. Human housekeeping genes, revisited. *Trends Genet.* 2013 Oct;29(10):569-74. doi: 10.1016/j.tig.2013.05.010. Epub 2013 Jun 27. Erratum in: *Trends Genet.* 2014 Mar;30(3):119-20. PMID: 23810203.

Edward J. Banigan, Wen Tang, Aafke A. van den Berg, Roman R. Stocsits, Gordana Wutz, Hugo B. Brandão, Georg A. Busslinger, Jan-Michael Peters, Leonid A. Mirny. Transcription shapes 3D chromatin organization by interacting with loop-extruding cohesin complexes. *bioRxiv* 2022.01.07.475367; doi: <https://doi.org/10.1101/2022.01.07.475367>

Evonne McArthur, John A. Capra. Topologically associating domain boundaries that are stable across diverse cell types are evolutionarily constrained and enriched for heritability. *The American Journal of Human Genetics*, Volume 108, Issue 2, 2021, Pages 269-283, ISSN 0002-9297, <https://doi.org/10.1016/j.ajhg.2021.01.001>.

Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* 2016 May 31;15(9):2038-49. doi: 10.1016/j.celrep.2016.04.085. Epub 2016 May 19. PMID: 27210764; PMCID: PMC4889513.

Ghavi-Helm Y. Functional Consequences of Chromosomal Rearrangements on Gene Expression: Not So Deleterious After All? *J Mol Biol.* 2020 Feb 7;432(3):665-675. doi: 10.1016/j.jmb.2019.09.010. Epub 2019 Oct 15. PMID: 31626801.

Gouti, M.; Tsakiridis, A.; Wymeersch, F. J.; Huang, Y.; Kleinjung, J.; Wilson, V.; Briscoe, J., 2014: In vitro generation of neuromesodermal progenitors reveals distinct roles for wnt signalling in the specification of spinal cord and paraxial mesoderm identity. *PLoS Biology.*, 12

Hansen AS, Cattoglio C, Darzacq X, Tjian R. Recent evidence that TADs and chromatin loops are dynamic structures. *Nucleus.* 2018 Jan 1;9(1):20-32. doi: 10.1080/19491034.2017.1389365. Epub 2017 Dec 14. PMID: 29077530; PMCID: PMC5990973.

Hua, P., Badat, M., Hanssen, L.L.P. *et al.* Defining genome architecture at base-pair resolution. *Nature* **595**, 125–129 (2021). <https://doi.org/10.1038/s41586-021-03639-4>

Jeong Y, El-Jaick K, Roessler E, Muenke M, Epstein DJ. A functional screen for sonic hedgehog regulatory elements across a 1 Mb interval identifies long-range ventral forebrain enhancers. *Development.* 2006 Feb;133(4):761-72. doi: 10.1242/dev.02239. Epub 2006 Jan 11. PMID: 16407397.

Kikuta H, Laplante M, Navratilova P, Komisarczuk AZ, Engström PG, Fredman D, Akalin A, Caccamo M, Sealy I, Howe K, Ghislain J, Pezeron G, Mourrain P, Ellingsen S, Oates AC, Thisse C, Thisse B, Foucher I, Adolf B, Geling A, Lenhard B, Becker TS. Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates. *Genome Res.* 2007 May;17(5):545-55. doi: 10.1101/gr.6086307. Epub 2007 Mar 26. PMID: 17387144; PMCID: PMC1855176.

Lee R, Kang MK, Kim YJ, Yang B, Shim H, Kim S, Kim K, Yang CM, Min BG, Jung WJ, Lee EC, Joo JS, Park G, Cho WK, Kim HP. CTCF-mediated chromatin looping provides a topological framework for the formation of phase-separated transcriptional condensates. *Nucleic Acids Res.* 2022 Jan 11;50(1):207-226. doi: 10.1093/nar/gkab1242. PMID: 34931241; PMCID: PMC8855298.

Lettice LA, Daniels S, Sweeney E, Venkataraman S, Devenney PS, Gautier P, Morrison H, Fantès J, Hill RE, FitzPatrick DR. Enhancer-adoption as a mechanism of human developmental disease. *Hum Mutat.* 2011 Dec;32(12):1492-9. doi: 10.1002/humu.21615. Epub 2011 Oct 20. PMID: 21948517.

Li, M., Gan, J., Sun, Y. *et al.* Architectural proteins for the formation and maintenance of the 3D genome. *Sci. China Life Sci.* **63**, 795–810 (2020). <https://doi.org/10.1007/s11427-019-1613-3>

Lupianez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*. 161 (2015) 1012-1025.

Lupiáñez, D. G., Spielmann, M. & Mundlos, S. Breaking TADs: How Alterations of Chromatin Domains Result in Disease. *Trends in Genetics* 32, 225–237 (2016).

Matsuda, K.; Kondoh, H., 2014: Dkk1-dependent inhibition of Wnt signaling activates *Hesx1* expression through its 5' enhancer and directs forebrain precursor development. *Genes to Cells*, 19, 374–385.

Nathaniel D Heintzman, Gary C Hon, R David Hawkins, Pouya Kheradpour, Alexander Stark, Lindsey F Harp, Zhen Ye, Leonard K Lee, Rhona K Stuart, Christina W Ching, Keith A Ching, Jessica E Antosiewicz-Bourget, Hui Liu, Xinmin Zhang, Roland D Green, Victor V Lobanekov, Ron Stewart, James A Thomson, Gregory E Crawford, Manolis Kellis, Bing Ren. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 459, 108–112 (2009)

Ohlsson R, Renkawitz R, Lobanekov V. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet.* 2001;17:520–7. PMID:11525835 [https://doi.org/10.1016/S0168-9525\(01\)02366-6](https://doi.org/10.1016/S0168-9525(01)02366-6).

Pachano T, Sánchez-Gaya V, Ealo T, Mariner-Faulí M, Bleckwehl T, Asenjo HG, Respuela P, Cruz-Molina S, Muñoz-San Martín M, Haro E, van IJcken WFJ, Landeira D, Rada-Iglesias A. Orphan CpG islands amplify poised enhancer regulatory activity and determine target gene responsiveness. *Nat Genet.* 2021 Jul;53(7):1036-1049. doi: 10.1038/s41588-021-00888-x. Epub 2021 Jun 28. PMID: 34183853; PMCID: PMC7611182.

Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014 Dec 18;159(7):1665-80. doi: 10.1016/j.cell.2014.11.021. Epub 2014 Dec 11. Erratum in: *Cell*. 2015 Jul 30;162(3):687-8. PMID: 25497547; PMCID: PMC5635824.

Richterova J, Huraiova B, Gregan J. Genome Organization: Cohesin on the Move. *Mol Cell*. 2017 May 18;66(4):444-445. doi: 10.1016/j.molcel.2017.05.002. PMID: 28525739

Sati, S., Cavalli, G. Chromosome conformation capture technologies and their impact in understanding genome function. *Chromosoma* **126**, 33–44 (2017). <https://doi.org/10.1007/s00412-016-0593-6>

Smith E, Shilatifard A. Enhancer biology and enhanceropathies. *Nat Struct Mol Biol.* 2014 Mar;21(3):210-9. doi: 10.1038/nsmb.2784. PMID: 24599251.

Spielmann M, Lupiáñez DG, Mundlos S. Structural variation in the 3D genome. *Nat Rev Genet.* 2018 Jul;19(7):453-467. doi: 10.1038/s41576-018-0007-0. PMID: 29692413.

Symmons O, Pan L, Remeseiro S, Aktas T, Klein F, Huber W, Spitz F. The Shh Topological Domain Facilitates the Action of Remote Enhancers by Reducing the Effects of Genomic Distances. *Dev Cell.* 2016 Dec 5;39(5):529-543. doi: 10.1016/j.devcel.2016.10.015. Epub 2016 Nov 17. PMID: 27867070; PMCID: PMC5142843.

T. Sexton, E. Yaffe, E. Kenigsberg, F. Bantignies, B. Leblanc, M. Hoichman, H. Parrinello, A. Tanay, G. Cavalli. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell*, 148 (2012), pp. 458-472

Uhlmann F. SMC complexes: from DNA to chromosomes. *Nat Rev Mol Cell Biol.* 2016 Jul;17(7):399-412. doi: 10.1038/nrm.2016.30. Epub 2016 Apr 14. PMID: 27075410.

van Arensbergen J, van Steensel B, Bussemaker HJ. In search of the determinants of enhancer-promoter interaction specificity. *Trends Cell Biol.* 2014 Nov;24(11):695-702. doi: 10.1016/j.tcb.2014.07.004. Epub 2014 Aug 24. PMID: 25160912; PMCID: PMC4252644.

Vian L, Pękowska A, Rao SSP, Kieffer-Kwon KR, Jung S, Baranello L, Huang SC, El Khattabi L, Dose M, Pruett N, Sanborn AL, Canela A, Maman Y, Oksanen A, Resch W, Li X, Lee B, Kovalchuk AL, Tang Z, Nelson S, Di Pierro M, Cheng RR, Machol I, St Hilaire BG, Durand NC, Shamim MS, Stamenova EK, Onuchic JN, Ruan Y, Nussenzweig A, Levens D, Aiden EL, Casellas R. The Energetics and Physiological Impact of Cohesin Extrusion. *Cell.* 2018 May 17;173(5):1165-1178.e20. doi: 10.1016/j.cell.2018.03.072. Epub 2018 Apr 26. Erratum in: *Cell.* 2018 Sep 20;175(1):292-294. PMID: 29706548; PMCID: PMC6065110.

Williamson I, Kane L, Devenney PS, Flyamer IM, Anderson E, Kilanowski F, Hill RE, Bickmore WA, Lettice LA. Developmentally regulated *Shh* expression is robust to TAD perturbations. *Development.* 2019 Sep 30;146(19):dev179523. doi: 10.1242/dev.179523. PMID: 31511252; PMCID: PMC7212092.

Zheng H, Xie W. The role of 3D genome organization in development and cell differentiation. *Nat Rev Mol Cell Biol.* 2019 Sep;20(9):535-550. doi: 10.1038/s41580-019-0132-4. PMID: 31197269.