



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**Maestría y Doctorado en Ciencias Bioquímicas**

“Análisis bioinformático de lncRNAs diferencialmente expresados durante la diferenciación dopaminérgica de células troncales pluripotenciales de humano”

TESIS

QUE PARA OPTAR POR EL GRADO DE:  
Maestro en Ciencias

PRESENTA:  
Ismael Portillo Pantoja

J. Iván Velasco Velázquez  
[Instituto de Fisiología Celular, UNAM](#)

Ernesto Soto Reyes Solís  
[UAM Cuajimalpa](#)

Leonor Pérez Martínez  
[IBt, UNAM](#)

Ciudad de México. Junio, 2022



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**Portillo Pantoja Ismael**  
**Estudiante de Maestría en Ciencias Bioquímicas**  
**Presente**

Los miembros del Subcomité Académico en reunión ordinaria del 14 de marzo del presente año, conocieron su solicitud de asignación de **JURADO DE EXAMEN** para optar por el grado de **Maestro en Ciencias**, con la réplica de la tesis **"Análisis bioinformático de lncRNAs diferencialmente expresados durante la diferenciación dopaminérgica de células troncales pluripotenciales de humano"**, dirigida por el/la Dr(a). **Velasco Velázquez Jaime Iván (IFC)**.

De su análisis se acordó nombrar el siguiente jurado integrado por los doctores:

PRESIDENTE	Recillas Targa Félix (IFC)
VOCAL	Chávez González Antonieta (IMSS)
VOCAL	Vázquez Laslop Matha Verónica (IBT)
VOCAL	Guerra Crespo Magdalena (IFC)
SECRETARIO	Aguilar Arnal Lorena (IIB)

#### **A LOS MIEMBROS DEL JURADO:**

Es obligación de los tutores de este programa participar en éstas y otras actividades académicas encomendadas por nuestro Comité Académico. Sin embargo, en caso de que tenga un impedimento académico o de salud para cumplir con esta encomienda, es muy importante contar con su respuesta (Formato anexo) en un plazo no mayor a una semana.

Tome en cuenta que usted tiene **20(Maestría)** días hábiles para emitir su voto con las rondas de revisión que considere necesarias.

Sin otro particular por el momento, aprovecho la ocasión para enviarle un cordial saludo.

**Atentamente**  
**"POR MI RAZA, HABLARÁ EL ESPÍRITU"**  
**Cd. Universitaria, Cd. Mx., a 14 de marzo de 2022**  
**COORDINADORA**



**Dra. Claudia Lydia Treviño Santa Cruz**

## **AGRADECIMIENTOS**

Agradezco infinitamente a:

Mis padres por su apoyo incondicional.

Mis hermanos por impulsarme a ser mejor en todos los sentidos.

Mis amigos Beto y Gaby por siempre estar a mi lado.

Mi pareja, Vannia por ser mi mayor motivación.

El doctor Iván Velasco por permitirme ser parte de su equipo de trabajo.

Mi comité tutorial por su tiempo y sus aportes al proyecto.

La Doctora Raquel Cuevas por su ayuda invaluable con los análisis bioinformáticos y por su gran disposición.

Mis compañeros de laboratorio y a los doctores Adolfo y Enrique por apoyarme con el proyecto.

El apoyo financiero de CONACyT (No. CVU: 1042621).

Los donativos CONACyT 272815, PAPIIT-UNAM IN213719 y PAPIIT-UNAM IN219122.

# ÍNDICE GENERAL

<b>ÍNDICE DE TABLAS</b> .....	1
<b>ÍNDICE DE FIGURAS</b> .....	1
<b>ABREVIATURAS</b> .....	2
<b>RESUMEN</b> .....	4
<b>ABSTRACT</b> .....	5
<b>INTRODUCCIÓN</b> .....	6
Las células troncales como modelo del desarrollo .....	6
Desarrollo de neuronas dopaminérgicas en vertebrados .....	7
Función de la dopamina en el cerebro .....	10
La Enfermedad de Parkinson (EP) .....	11
Tratamientos para la EP .....	13
Terapia celular de la EP con nDAs diferenciadas de hPSCs .....	13
Cambios epigenéticos en la diferenciación de células troncales.....	15
Clasificación y función de los lncRNAs .....	16
Regulación de la diferenciación celular por lncRNAs .....	20
<b>ANTECEDENTES</b> .....	22
Identificación de lncRNAs diferencialmente expresados (DE) durante el proceso de diferenciación dopaminérgica a partir de hESCs .....	22
Predicción de la función de lncRNAs por análisis bioinformáticos .....	24
<b>JUSTIFICACIÓN</b> .....	26
<b>HIPÓTESIS</b> .....	26
<b>OBJETIVOS</b> .....	26
Objetivo general .....	26
Objetivos particulares .....	26
<b>DISEÑO EXPERIMENTAL</b> .....	27
<b>METODOLOGÍA</b> .....	28
Extracción de RNA total.....	28
Tratamiento del RNA total con DNasa I .....	29
Síntesis de DNA complementario (cDNA) y qPCR.....	29
Análisis de sobrerrepresentación de conjunto de genes y matriz de asociación .....	31
Construcción de redes de coexpresión de lncRNAs y mRNAs de EFs .....	32
Construcción de redes de coexpresión entre lncRNAs y PCGs asociados.....	32
Análisis estadístico .....	32

<b>RESULTADOS</b> .....	33
Identificación de lncRNAs diferencialmente expresados en la diferenciación dopaminérgica implicados en estado pluripotencial de hESCs y diferenciación neural .....	33
Validación de lncRNAs seleccionados durante la diferenciación dopaminérgica a partir de hESC por RT-qPCR .....	34
Inferencia de la función de lncRNAs por análisis bioinformáticos .....	39
EFs identificados DE en la diferenciación dopaminérgica .....	43
Regulación <i>in cis</i> de lncRNAs .....	45
<b>DISCUSIÓN</b> .....	49
<b>CONCLUSIÓN</b> .....	54
<b>PERSPECTIVAS</b> .....	54
<b>ANEXOS</b> .....	55
Gel de electroforesis de muestras de RNA del D28 (NDA1 y NDA2) tratadas con DNasa I .....	55
Gel de electroforesis de muestras de cDNA del D28 .....	55
<b>BIBLIOGRAFÍA</b> .....	56

## ÍNDICE DE TABLAS

Tabla I. Lista de cebadores diseñados para la validación por RT-qPCR. ....	30
Tabla II. Los 30 lncRNAs más DE a la baja en el intervalo 0-28 D analizados por LNCpedia. ....	35
Tabla III. Los 30 lncRNAs más DE al alza en el intervalo 0-28 D analizados por LNCpedia.....	36

## ÍNDICE DE FIGURAS

Figura 1. Establecimiento de la identidad regional del tubo neural.....	9
Figura 2. Desarrollo de nDAs del mesencéfalo en vertebrados. ....	9
Figura 3. Principales vías dopaminérgicas.....	11
Figura 4. Características histológicas de la EP. ....	12
Figura 5. Protocolo de diferenciación dopaminérgica.....	15
Figura 6. Clasificación de lncRNAs por su relación con genes codificantes vecinos. ....	18
Figura 7. Regulación en <i>cis</i> y <i>trans</i> de lncRNAs. ....	19
Figura 8. Mecanismos de la función de lncRNAs. ....	20
Figura 9. Participación de lncRNAs en la pluripotencia y en la diferenciación neural.....	21
Figura 10. Genes DE durante la diferenciación dopaminérgica.....	23
Figura 11. Los lncRNAs tienen cambios dinámicos en su expresión. ....	23
Figura 12. Predicción de la función de lncRNAs en SCI en rata. ....	25
Figura 13. Productos de PCR punto final de lncRNAs seleccionados para validación.....	37
Figura 14. Expresión de lncRNAs por RT-qPCR. ....	38
Figura 15. Productos de PCR punto final para <i>ACTB</i> con muestras de RNA de D28 (NDA1 y NDA2) tratado con DNasa I. ....	39
Figura 16. Participación de lncRNAs en procesos biológicos.....	41
Figura 17. Correlación de expresión entre lncRNAs candidatos y marcadores de pluripotencia y dopaminérgicos. ....	42
Figura 18. Los lncRNAs DE a la baja tienen alto nivel de coexpresión con EFs.....	44
Figura 19. Los lncRNAs DE al alza tienen alto nivel de coexpresión con PCGs cercanos. ....	48
Figura 20. Participación <i>in cis</i> de lncRNAs en el desarrollo del SN. ....	48

## ABREVIATURAS

°C	Grados Celsius
cDNA	DNA complementario
D	Día
DE	Diferencialmente expresado
EFs	Factores epigenéticos (“Epigenetic Factors”)
ESCs	Células troncales embrionarias (“Embryonic Stem Cells”)
EP	Enfermedad de Parkinson
FDR	Tasa de descubrimientos falsos (“False Discovery Rate”)
FPKM	Fragmentos por kilobase por millón de lecturas mapeadas
GS	Conjunto de genes (“Gene Sets”)
GSEA	Análisis de sobrerrepresentación de genes (“Gene Set Enrichment Analysis”)
hESCs	Células troncales embrionarias humanas (“Human Embryonic Stem Cells”)
IGV	Visualizador genómico integrativo (“Integrative Genomics Viewer”)
iPSC	Célula troncal pluripotencial inducida (“Induced Pluripotent Stem Cells”)
Kb	Kilobase
lncRNA	RNA no codificante largo (“long non-coding RNA”)
log <sub>2</sub> FC	Veces de cambio en logaritmo base dos
nDAs	Neuronas dopaminérgicas
NES	Puntuación normalizada de enriquecimiento (“Normalized Enrichment Score”)
NSCs	Células troncales neurales (“Neural Stem Cells”)

nM	Nanomolar
μg	Microgramo
miRNA	MicroRNA
μL	Microlitro
mM	Milimolar
μM	Micromolar
NAT	Transcrito antisentido natural (“Natural Antisense Transcript”)
PCC	Coeficiente de correlación de Pearson (“Pearson Correlation Coefficient”)
PCGs	Genes que codifican proteínas (“Protein Coding Genes”)
PRC2	“Polycomb Repressive Complex 2”
RA	Ácido retinoico (“Retinoic Acid”)
RNA-seq	Secuenciación de RNA (“RNA-Sequencing”)
SCI	Lesión de médula espinal (Spinal Cord Injury)
SCs	Células troncales (“Stem Cells”)
SN	Sistema Nervioso
SNC	Sistema Nervioso Central
<i>SNpc</i>	<i>Substantia nigra pars compacta</i>
TF	Factor transcripcional

## RESUMEN

Los RNAs largos no-codificantes (lncRNAs) pertenecen a un grupo muy diverso de transcritos que se caracterizan por tener una longitud mayor a los 200 nucleótidos y por carecer de marcos de lectura abierto (ORFs) por lo que no codifican para proteínas. En décadas pasadas se les consideraba como “RNA basura” o “ruido transcripcional” ya que se creía que el producto funcional de los genes eran las proteínas. Sin embargo, hoy en día se conocen lncRNAs que modulan procesos celulares muy importantes, entre ellos, el mantenimiento del estado indiferenciado de células troncales pluripotenciales (PSCs) y la diferenciación celular. Ambos procesos son de gran interés en la medicina regenerativa ya que su entendimiento permite obtener células de un linaje específico para su posible aplicación clínica. Se han establecido protocolos donde, a partir de PSCs, se obtienen neuronas dopaminérgicas (nDAs), las cuales mueren en la enfermedad de Parkinson. A pesar de que estos protocolos son eficientes, se generan otros tipos celulares por lo que es imperativo entender los mecanismos moleculares que regulan la diferenciación hacia un linaje en específico. Por esta razón, el objetivo del presente trabajo fue identificar lncRNAs que modulen el estado pluripotencial de PSCs y el proceso de diferenciación dopaminérgica. A partir del perfil transcriptómico del proceso de diferenciación dopaminérgica, se seleccionaron lncRNAs diferencialmente expresados (DE) para su validación por RT-qPCR y mediante análisis bioinformáticos se asociaron a procesos celulares importantes. Se encontró que los lncRNAs sobreexpresados en el estado pluripotencial (día 0) están enriquecidos en procesos relacionados a la proliferación celular, a la regulación del ciclo celular y de PSCs. Además, una fracción importante de ellos se correlacionaron con factores epigenéticos (EFs) y se encontró que los RNAs sobreexpresados en neuronas inmaduras (día 28) se asocian a diferenciación neuronal dopaminérgica y que algunos de ellos están físicamente cercanos a genes codificantes que participan en el desarrollo del sistema nervioso.

## **ABSTRACT**

Long non-coding RNAs (lncRNAs) are a very diverse group of transcripts that are longer than 200 nucleotides and for lacking open reading frames (ORFs) meaning that they do not code for proteins. In past decades they were considered "junk RNA" or "transcriptional noise" since it was believed that the only functional product of genes were proteins. Nowadays it is known that lncRNAs can modulate very important cellular processes, including the maintenance of pluripotent stem cells (PSCs) and cell differentiation. Both processes are of great interest in regenerative medicine since their understanding allows obtaining cells of a specific lineage for possible clinical application. There are protocols to obtain dopaminergic neurons (nDAs) from PSCs; this neuronal type degenerates in Parkinson's disease. Although these protocols are efficient, other cell types are generated, so it is imperative to understand the molecular mechanisms that regulate differentiation towards a specific lineage. For this reason, the goal of the present work was to identify lncRNAs that modulate the pluripotent state of PSCs and the dopaminergic differentiation process. From the transcriptomic profile of the dopaminergic differentiation process, differentially expressed (DE) lncRNAs were selected for validation by RT-qPCR and through bioinformatic analysis they were associated with important cellular processes where we found that the overexpressed lncRNAs in the pluripotent state (day 0) are enriched in processes related to cell proliferation, cell cycle and PSCs regulation. Furthermore, a significant fraction of them correlated with epigenetic factors (EFs) and the overexpressed RNAs in immature neurons (day 28) associates with dopaminergic neuronal differentiation and some of them are spatially close to coding genes that participate in the development of the nervous system.

## INTRODUCCIÓN

### Las células troncales como modelo del desarrollo

Las PSCs, como las células troncales embrionarias (ESCs) y las células troncales pluripotenciales inducidas (iPSCs), sirven como una fuente atractiva de células ya que ambas tienen la capacidad para autorrenovarse, lo que permite expansiones ilimitadas de estas células indiferenciadas y además, estas células pueden dar origen a cualquier tipo de célula diferenciada de adulto (Fan & Ng, 2020).

Las ESCs humanas (hESCs) se obtuvieron por primera vez de embriones congelados en etapa previa al blastocisto, producidos por fertilización *in vitro* para fines reproductivos. Estos embriones se donaron por consentimiento informado y se aprobó su uso por la junta de revisión institucional. Los embriones se cultivaron hasta la etapa de blastocisto de los cuales se aislaron 14 masas celulares internas y se derivaron cinco líneas celulares que se originaron de cinco embriones separados. Las células resultantes tuvieron una alta proporción de núcleo-citoplasma, nucleolos prominentes y una morfología de colonias como la de las ESCs de monos *Rhesus*. Tres líneas celulares (H1, H13 y H14) fueron de cariotipo masculino normal XY, y dos líneas celulares (H7 y H9) tuvieron un cariotipo femenino normal XX. Cada una de las líneas celulares se criopreservó y descongeló con éxito. Las hESCs tienen altos niveles de marcadores de superficie celular que caracterizan las ESCs, incluido el antígeno embrionario específico del estadio (SSEA)-3, SSEA-4, TRA-1-60, TRA-1-81, presentan actividad de la Telomerasa y mantienen el potencial para formar derivados de las tres capas germinales embrionarias (Thomson, 1998). Además, presentan altos niveles de los factores de pluripotencia OCT4, SOX2 y NANOG (Park et al., 2019).

Las iPSCs humanas se generan de células somáticas terminalmente diferenciadas, las cuales se reprograman al inducir la expresión constitutiva de los factores transcripcionales (TFs) *Oct4*, *Sox2*, *Klf4* y *c-Myc*, mismos que participan activamente en el mantenimiento de la

troncalidad. Estas células son muy similares a las ESCs ya que tienen un cariotipo normal, expresan marcadores asociados al estado pluripotencial y tienen una gran actividad de la telomerasa. Además, tienen la capacidad para formar teratomas (tumores formados por tejidos de las tres capas germinativas), lo cual evidencia su pluripotencia (Takahashi & Yamanaka, 2006).

La diferenciación dirigida de las hPSC hacia células especializadas en cultivo sigue los mismos principios que subyacen al desarrollo embrionario. Por lo tanto, los conocimientos obtenidos de la Biología del desarrollo han guiado las estrategias para usar factores de crecimiento o moduladores de las vías de señalización del desarrollo para mantener la pluripotencia o el destino de las células troncales hacia el linaje deseado y con alta eficiencia (revisado en Williams et al., 2012).

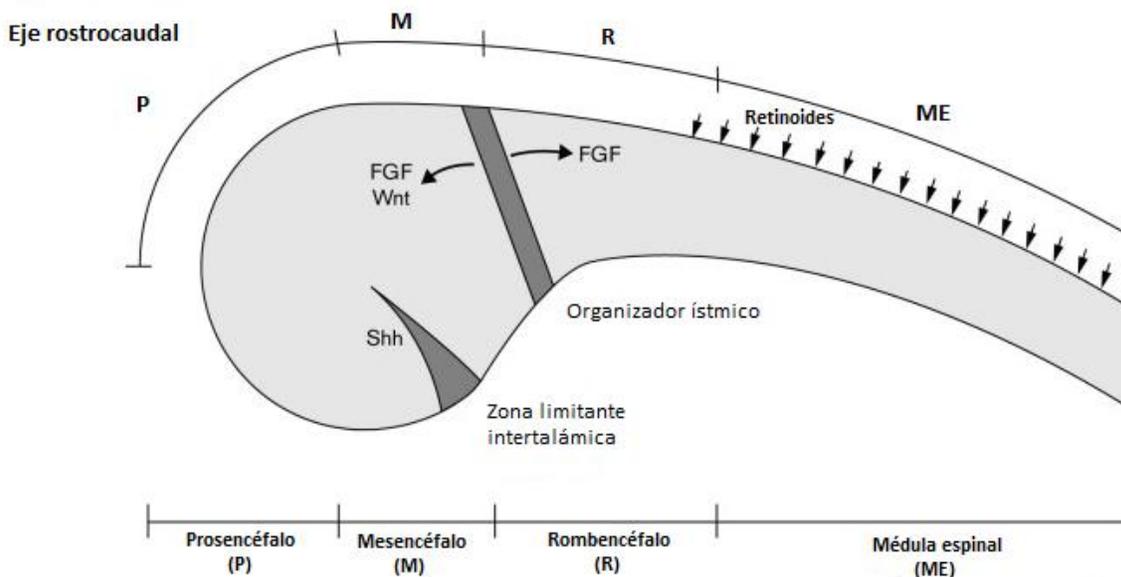
### **Desarrollo de neuronas dopaminérgicas en vertebrados**

La neurulación es un proceso dinámico que da lugar al tubo neural a partir del plegado de la placa neural, para culminar posteriormente en el desarrollo del SN. En el tubo neural existen grupos especializados de células que secretan factores que establecen la identidad regional de las vesículas que componen el sistema nervioso central (SNC) (Figura 1) así como el destino neuronal (revisado en Jessell & Sanes, 2000).

El ratón ha sido el modelo de estudio para entender cómo se especifican y diferencian las neuronas dopaminérgicas (nDAs) en el mesencéfalo en desarrollo y aunque existen diferencias en comparación con el humano (ej. tipos celulares intermedios específicos entre ambas especies), muchos tipos neuronales aparecen en tiempos homólogos (La Manno et al., 2016).

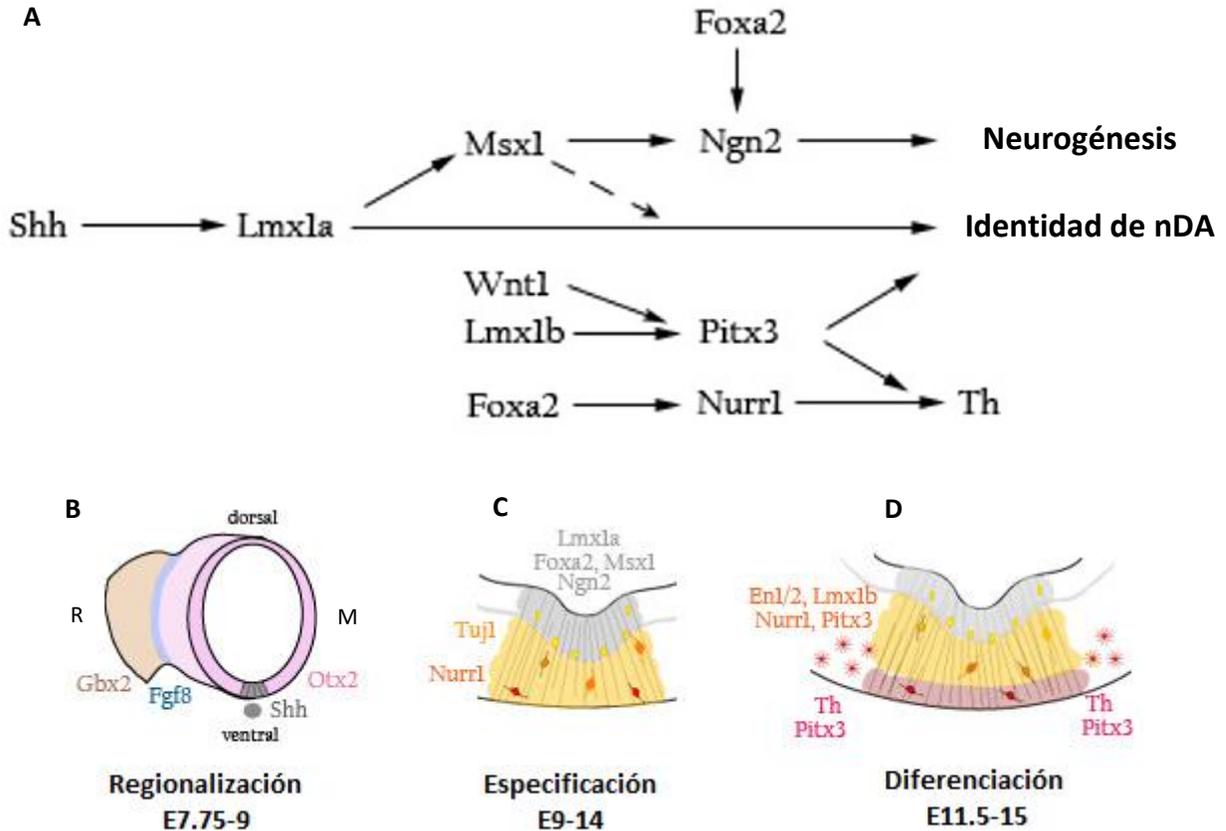
El desarrollo de las nDAs del mesencéfalo es un proceso altamente coordinado por morfógenos, los cuales establecen la información posicional de las células al activar vías de señalización específicas y TFs que encienden programas genéticos de manera orquestada (Figura 2A). Este proceso de diferenciación consta de las siguientes etapas:

1. Regionalización: el límite mesencéfalo-rombencéfalo se establece gracias a la expresión de *Otx2* en la porción mesencefálica y *Gbx2* en el rombencéfalo para dar origen al organizador ístmico (OI), el cual secreta FGFs y Wnts en la porción mesencefálica (Figura 1) y junto con Sonic hedgehog (SHH), el cual es secretado por la notocorda y posteriormente por la placa de piso del mesencéfalo, generan el patrón anteroposterior y dorsoventral, respectivamente (Figura 2B).
2. Especificación de la placa de piso: los precursores de nDA son células parecidas a la glia radial que se encuentran en la placa de piso del mesencéfalo en donde se activa la vía de SHH, esta regula la expresión de *Lmx1a* el cual a su vez activa la expresión de *Msx1*, este último junto con *Foxa2* inducen la expresión de *Ngn2* la cual permite la proliferación de precursores *Tuj1+/Nurr1+* para su posterior migración a la zona intermedia del mesencéfalo (Figura 2C).
3. Diferenciación: los precursores dopaminérgicos de la *substantia nigra pars compacta* (*SNpc*) y del área tegmental ventral (VTA) son indistinguibles hasta el día embrionario (E)13.5 donde, en la *SNpc*, se expresan *Nurr1*, *Pitx3*, *En1*, *En2* y *Lmx1b* los cuales actúan coordinadamente para generar el fenotipo dopaminérgico al inducir la expresión de la tirosina hidroxilasa (*Th*), gen que codifica para la enzima limitante en la síntesis de dopamina y también regulan la maduración y el mantenimiento de este fenotipo (Figura 2D) (revisado en Volpicelli et al., 2020).



**Figura 1. Establecimiento de la identidad regional del tubo neural.**

Tanto la zona limitante intertalámica (ZLI) como el organizador ístmico (OI) secretan factores que permiten la regionalización del tubo neural. El OI secreta FGFs y Wnts, ambos son necesarios para la diferenciación y el patrón del mesencéfalo y el rombencéfalo mientras que la ZLI secreta SHH en forma de gradiente importante para la ventralización del mesencéfalo (modificado de Jessell & Sanes, 2000).



**Figura 2. Desarrollo de nDAs del mesencéfalo en vertebrados.**

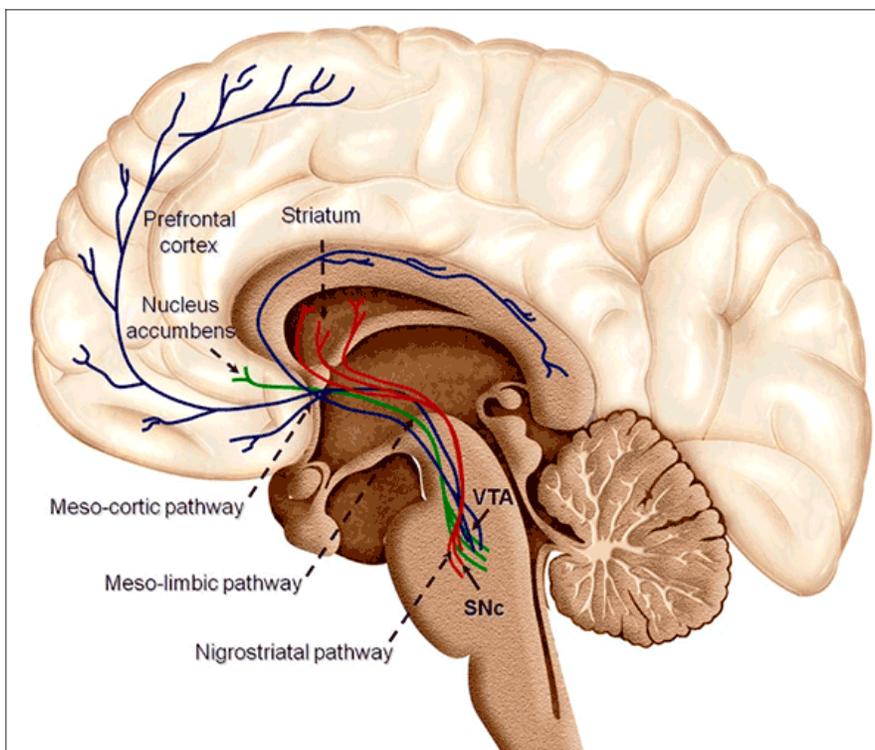
**(A)** Morfógenos y TFs involucrados en la neurogénesis e identidad de nDAs. **(B)** Regionalización del mesencéfalo (M) por morfógenos y TFs. R; Rombencéfalo. **(C)** En la zona proliferativa (gris) de la línea media ventral del mesencéfalo se expresan TFs que permiten la especificación de las células progenitoras dopaminérgicas que entran en la zona intermedia (amarillo). **(D)** Las células *Pitx3*<sup>+</sup> de la *SNpc* migran ventralmente hasta la zona periférica (rojo), eventualmente estas células migran lateralmente y expresan *Th* (modificado de Gale & Li, 2008).

## **Función de la dopamina en el cerebro**

La dopamina se sintetiza a partir del aminoácido L-tirosina, donde la Tirosina hidroxilasa (TH) es la enzima limitante en la velocidad de síntesis del neurotransmisor. Existen mecanismos de homeostasis que controlan con precisión tanto la síntesis, el almacenamiento, la liberación, la recaptura y el metabolismo de la dopamina. La desregulación de estos mecanismos se asocia a varios trastornos, uno de ellos es la pérdida de dopamina en el cuerpo estriado lo que causa la pérdida del control motor observado en pacientes con la EP.

La dopamina es una catecolamina que actúa como neurotransmisor tanto en el SNC, como en el SN periférico. Los núcleos importantes que contienen a las nDAs incluyen la *SNpc* que envía proyecciones al neostriado (vía nigroestriatal) y el área tegmental ventral (VTA), el núcleo que envía proyecciones hacia la corteza prefrontal (vía mesocortical) y a estructuras límbicas como el núcleo accumbens y la amígdala (vía mesolímbica) (Figura 2). Por esta razón, la dopamina desempeña un papel importante en muchas funciones cerebrales tales como el ciclo de sueño-vigilia, toma de decisiones, aprendizaje de recompensa, control del movimiento voluntario y procesamiento cognitivo, entre otras (Best et al., 2009).

Existen cinco subtipos de receptores dopaminérgicos, miembros de la superfamilia de receptores acoplados a proteínas G. Estos subtipos se dividen en dos familias principales, D<sub>1</sub> y D<sub>2</sub>. Los receptores del subtipo D<sub>1</sub> (D<sub>1</sub> y D<sub>5</sub>) estimulan la formación de AMP cíclico como el principal mecanismo de transducción. Estos receptores son esenciales reguladores de la recompensa, actividad motora y cognición del SNC. Por otro lado, al activar diferentes proteínas G (G<sub>ai</sub> y G<sub>ao</sub>), los receptores subtipo D<sub>2</sub> (D<sub>2</sub>, D<sub>3</sub> y D<sub>4</sub>), inhiben la formación de AMP cíclico, reducen la entrada de Ca<sup>2+</sup> a través de canales activados por voltaje y activan los canales de K<sup>+</sup>; en conjunto esto tiene un efecto inhibitorio en las neuronas (Beaulieu et al., 2011).



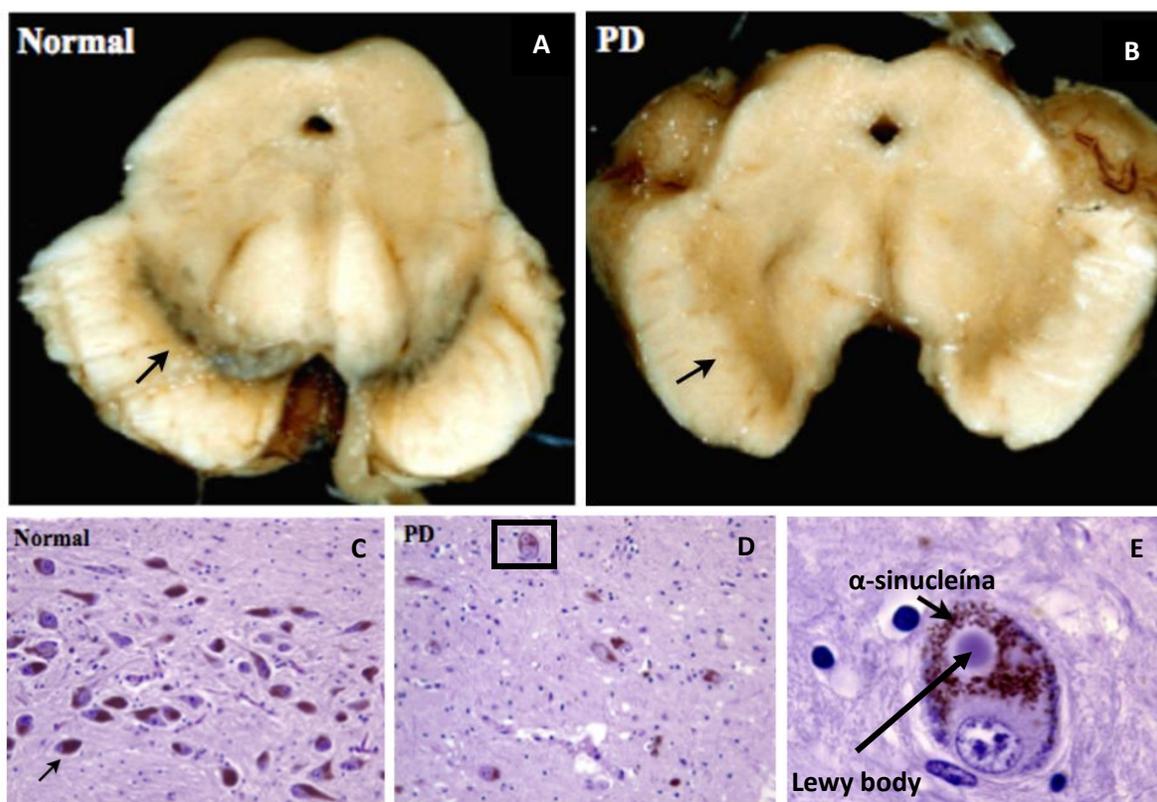
**Figura 3. Principales vías dopaminérgicas.**

Representación del cerebro humano en un plano sagital que muestra los núcleos de nDAs (*SNpc* y *VTA*), las tres principales vías dopaminérgicas: vía nigroestriatal (rojo), vía mesolímbica (verde) y vía mesocortical (azul) y las estructuras cerebrales hacia donde proyectan los axones (Best et al., 2009).

### **La Enfermedad de Parkinson (EP)**

Las enfermedades neurodegenerativas son un grupo heterogéneo de trastornos que se caracterizan por la pérdida progresiva de la estructura y función del sistema nervioso (SN). Las enfermedades neurodegenerativas más comunes incluyen la enfermedad de Alzheimer y la EP (Wang et al., 2019). La EP se describió inicialmente en el siglo XIX por James Parkinson como "la parálisis temblorosa" y hoy en día, es el segundo trastorno neurodegenerativo más común. Se estima que la prevalencia de la EP es de aproximadamente 1% en personas mayores de 60 años y aumenta del 3-4% a la edad de 80 años. La EP está caracterizada por agregaciones citoplasmáticas y axonales de  $\alpha$ -sinucleína (proteína que forma parte de la maquinaria de anclaje de las vesículas de almacenamiento de los neurotransmisores), conocidas como cuerpos de Lewy y la pérdida progresiva de nDAs en la *SNpc* del

mesencéfalo (Figura 3) que conlleva a la liberación disminuida de dopamina en los tractos extrapiramidales encargados de controlar el movimiento (Vallerga et al., 2020). Los signos clínicos son evidentes al perderse más del 80% de las nDAs de la *SNpc*. Entre los síntomas más característicos de la EP se encuentran la bradicinesia (lentitud en la ejecución de movimientos), rigidez, temblor, alteración de la postura y la marcha, trastornos del habla y deterioro cognitivo (Tarsy, 2012).



**Figura 4. Características histológicas de la EP.**

Región mesencefálica de una persona sin la EP (**A**) en donde se señala con flecha la pigmentación típica de las nDAs de la *SNpc* (neuromelanina) en comparación con el mesencéfalo de un paciente con la EP (**B**) en el cual no existe esta pigmentación por la pérdida de esas neuronas. Fotomicrografía del área de la *SNpc* donde se observan neuronas pigmentadas con neuromelanina en un cerebro sano (**C**), mientras en el cerebro de pacientes con EP hay una gran pérdida de estas neuronas (**D**). Al hacer una magnificación del área limitada por un recuadro negro de la imagen **D**, se observan agregaciones

citoplásmicas de  $\alpha$ -sinucleína y ubiquitina conocidos como cuerpos de Lewy (**E**). PD, Parkinson's disease; Lewy body, Cuerpo de Lewy (Aron, 2010).

### **Tratamientos para la EP**

La EP se puede tratar con terapias farmacológicas dopaminérgicas donde se usan precursores de dopamina como levodopa lo que mejora significativamente a muchos de los síntomas motores de la EP, particularmente en las primeras etapas de la enfermedad. Sin embargo, con el tiempo, la enfermedad avanza y, además, el uso prolongado de las terapias farmacológicas causa complicaciones motoras inducidas por fármacos, como discinesias inducidas por levodopa. Existen formas de administrar una terapia dopaminérgica más continua, aunque estas son más invasivas, como las bombas de apomorfina (vía subcutánea) y DuoDopa® (levodopa/carbidopa vía duodenal), así como intervenciones neuroquirúrgicas como la estimulación cerebral profunda. Estas terapias solo tratan los síntomas sin ningún intento de regenerar la población celular perdida durante la enfermedad. Por esta razón, se intentan encontrar tratamientos que puedan tanto aliviar los síntomas como reemplazar la inervación neuronal perdida. (revisado en Wijeyekoon & Barker, 2009).

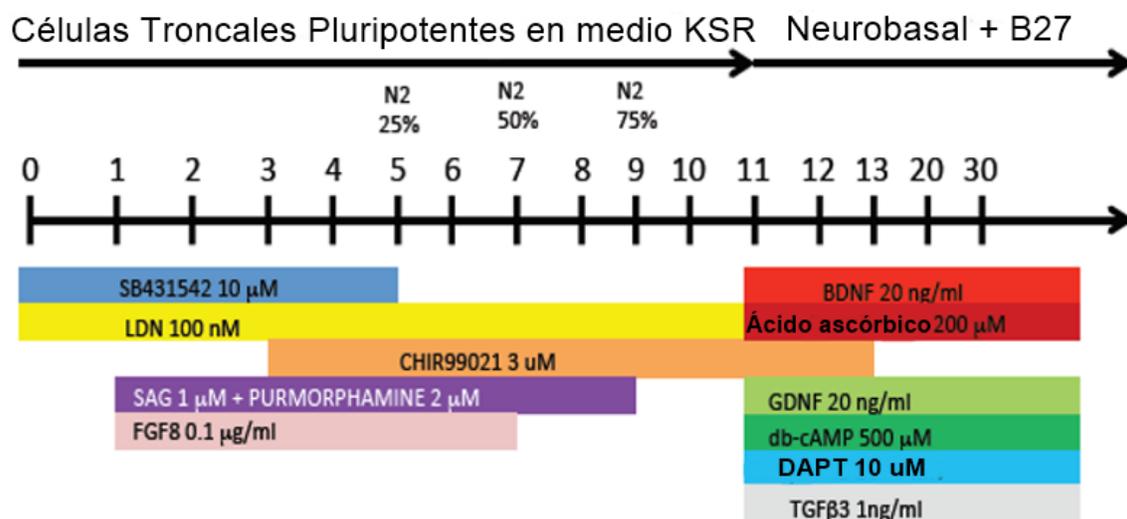
### **Terapia celular de la EP con nDAs diferenciadas de hPSCs**

Una alternativa prometedora para tratar la EP es la terapia celular en la cual se trasplantan nDAs con el objetivo de reemplazar las neuronas que han muerto en esta patología y así, reestablecer los niveles de dopamina en el estriado (revisado en Wijeyekoon & Barker, 2009).

Para obtener células dopaminérgicas del mesencéfalo a partir de hESCs se utilizan protocolos de diferenciación que manipulan vías de señalización con el objetivo de comprometer a las células a un linaje específico. La inhibición de la vía TGF- $\beta$ /BMP inhibe a las SMAD, mejorando así la diferenciación neural usando la combinación de SB431542 (inhibidor de ALK) y Nogina (inhibidor de BMP). Estos inhibidores desestabilizan la red de pluripotencia mediada por TGF- $\beta$  / Activina y NANOG, suprimen los destinos mesodérmicos

al inhibir a la Activina endógena y a las señales Nodal, lo que promueve la neuralización del ectodermo primitivo, a través de la inhibición de las SMADs (Chambers et al., 2009).

La vía de señalización canónica Wnt- $\beta$ -Catenina es un sistema de comunicación importante para la renovación de células troncales, la proliferación y la diferenciación, durante la embriogénesis y la homeostasis del tejido adulto. La señalización de Wnt- $\beta$ -Catenina es crucial para definir los ejes antero-posterior y dorso-ventral del cuerpo en múltiples especies animales. En concordancia con su papel en el desarrollo embrionario, la diferenciación de ESCs en diferentes capas germinales requiere la activación espaciotemporal de la señalización de Wnt- $\beta$ -Catenina (Steinhart & Angers, 2018). Para la diferenciación de hPSC a nDAs se utiliza CHIR99021, el cual es un inhibidor de GSK3 $\beta$ , que es responsable de la fosforilación y degradación continua de  $\beta$ -Catenina, lo que activa la vía de señalización canónica de Wnt y promueve la diferenciación. La señalización de Wnt- $\beta$ -Catenina y la activación de la vía de SHH inducen a la coexpresión de *FOXA2* y *LMX1A*, TFs presentes en el mesencéfalo en desarrollo. La activación de estas vías de señalización por moléculas pequeñas (Figura 5) enriquece la cantidad de nDAs mesencefálicas derivadas de hPSC. Sin embargo, a pesar de la alta eficiencia se observa heterogeneidad en el cultivo, es decir, no sólo se encuentran nDAs, sino también neuronas GABAérgicas (5%) y serotoninérgicas (1%) (Kriks et al., 2011).



**Figura 5. Protocolo de diferenciación dopaminérgica.**

Representación de las condiciones de cultivo que ilustra el tiempo en días y la concentración de moléculas pequeñas y factores neurotróficos usados para la obtención de nDAs a partir de hPSCs (modificado de Kriks et al., 2011).

**Cambios epigenéticos en la diferenciación de células troncales**

Las redes genéticas que controlan la identidad de las células troncales son de gran interés, debido a su potencial terapéutico y su gran relevancia para los modelos de desarrollo temprano (Mohamed et al., 2010).

La epigenética estudia los mecanismos reversibles y heredables mitóticamente o meióticamente que regulan la expresión genética en respuesta a estímulos ambientales, sin modificar la secuencia del DNA. El principio de esta regulación se basa en la accesibilidad (nivel de compactación de la cromatina) a elementos regulatorios de los genes como “enhancers” los cuales son secuencias de DNA que aumentan los niveles de transcripción de su gen blanco o a las secuencias promotoras de los genes, donde se une la maquinaria transcripcional. Si la cromatina se encuentra accesible (laxa), se favorece la unión de la maquinaria y TFs a las secuencias regulatorias lo que hace posible llevar a cabo la transcripción. De lo contrario, si la cromatina está compacta en esa región, este proceso no se lleva a cabo (revisado en Corso-Díaz et al., 2018).

Entre los mecanismos epigenéticos que controlan el estado de la cromatina se encuentra la metilación de citosinas del DNA por las DNA metiltransferasas en secuencias promotoras de los genes en regiones ricas en guanina y citosina (islas CpG). Este mecanismo se asocia generalmente a la represión genética ya que existen proteínas que reconocen estas modificaciones, las cuales, reclutan a complejos remodeladores de la cromatina cuya acción provoca el silenciamiento del gen. Sin embargo, esta modificación es reversible, ya que las proteínas TET oxidan el grupo metilo para su posterior eliminación, permitiendo nuevamente la expresión del gen. Además de la metilación del DNA existen otros mecanismos epigenéticos como los orquestados por enzimas que depositan o retiran

grupos químicos a las colas de las Histonas (escritores y borradores, respectivamente) para que otras proteínas interpreten esta modificación (lectoras), remodelando el estado de la cromatina y así, regular la expresión genética. También existen remodeladores de la cromatina dependientes de ATP que alteran los contactos DNA-Histona promoviendo así la accesibilidad del DNA y el intercambio de Histonas por variantes de Histonas ( no canónicas) proporcionando otro mecanismo de regulación génica (revisado en Allis & Jenuwein, 2016).

Estos mecanismos epigenéticos mantienen el estado pluripotencial de las ESCs al silenciar genes importantes en la determinación del destino celular y el desarrollo al mismo tiempo que mantienen transcripcionalmente activos a genes de proliferación y programas de pluripotencia (Bernstein et al., 2006), mientras que en el proceso de diferenciación, se activan genes específicos de un linaje celular a la vez que se mantienen silenciados genes de otros linajes y se silencian los genes de pluripotencia (revisado en Godini et al., 2018, Schuettengruber et al., 2017).

En las últimas décadas se han identificado una variedad heterogénea de RNAs no codificantes, entre ellos, RNAs largos no codificantes (lncRNAs) involucrados en la regulación de la expresión genética, entre ellos, en la regulación epigenética al conferir especificidad de secuencia a las actividades modificadoras de la cromatina (revisado en Roberts et al., 2014) por mecanismos moleculares muy variados. Sin embargo, hasta la fecha solo se ha caracterizado una mínima fracción de ellos (revisado en Banta & Richards, 2018).

### **Clasificación y función de los lncRNAs**

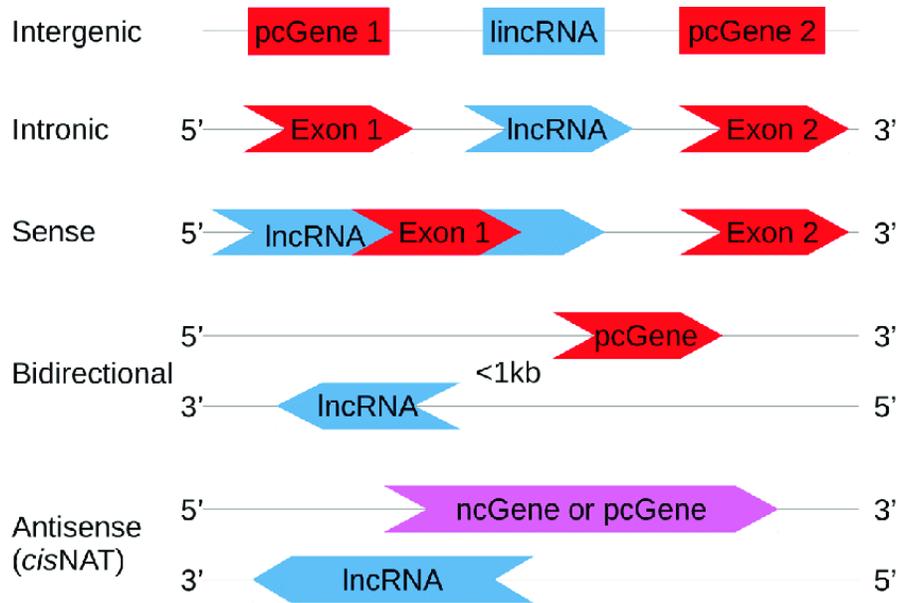
Los lncRNAs pertenecen a una clase de RNAs no codificantes que incluye miles de especies diferentes (> 60,000 en el genoma humano), son transcritos de RNA que tienen más de 200 nucleótidos de longitud, a menudo tienen “capping”, pueden o no estar poliadenilados, pueden experimentar “splicing” y carecen de ORFs evidentes. Además, los lncRNAs generalmente se expresan en un tejido y en una etapa de desarrollo muy específica y a órdenes de magnitud más bajos en comparación con los genes codificantes. Estos transcritos se pueden clasificar de muchas formas: por su tamaño, localización subcelular,

biogénesis, función o proceso biológico específico asociado. Sin embargo, la forma más común de clasificar a estos transcritos es de acuerdo con la localización relativa de su *locus* con respecto a genes codificantes (Figura 6), ya que, por una parte no se necesitan estudios experimentales exhaustivos previos para su clasificación y por otra, porque esta información es de gran utilidad para inferir la posible función de estos transcritos al estudiar los genes codificantes asociados (revisado en Jarroux et al., 2017).

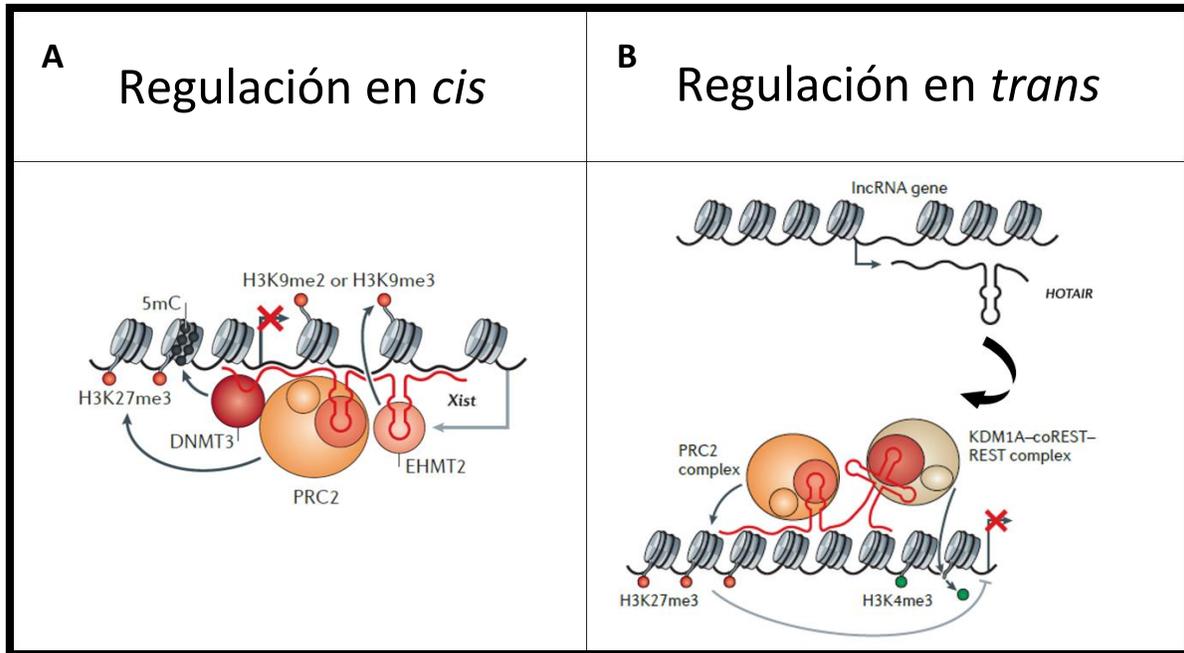
Los lncRNAs tienen dos formas de regular la expresión genética:

1. En *cis*, donde el transcrito regula el mismo *locus* que le dio origen, por ejemplo: *Xist* (“X-inactive specific transcript”), uno de los primeros lncRNAs reportados y mejor caracterizados, se expresa al inicio de la diferenciación en uno de los cromosomas X de células de hembras en mamíferos. Este RNA cubre físicamente al cromosoma X que le dió origen y recluta remodeladores de cromatina para silenciar este cromosoma, lo que permite balancear la diferencia de expresión derivada del cromosoma X entre ambos sexos (Figura 7A).
2. En *trans*, regula un *locus* distinto del que se originó, por ejemplo: *HOTAIR* (“*HOXA* transcript antisense RNA”) se origina del “cluster” *HOXC*. Este transcrito se traslada al *locus* de *HOXD* y en su extremo 5’ interactúa con PRC2 y en su extremo 3’ con la proteína lisina desmetilasa 1 (KDM1) así, PRC2 coloca la marca represiva H3K27me3 y KDM1 elimina la marca activadora H3K4me3, como resultado se silencia el “cluster” *HOXD* (Figura 7B). Esto permite coordinar de manera espacio-temporal precisa, la expresión de los genes de este “cluster” que regulan el plan corporal del embrión y los procesos de diferenciación (revisado en Fatica & Bozzini, 2014).

La naturaleza intrínseca de ácido nucleico confiere a los lncRNAs una enorme versatilidad, pudiendo interactuar con proteínas o mediar interacciones de apareamiento de bases en sitios específicos de RNA o DNA para ejercer su función. En la Figura 8 se ilustran los mecanismos moleculares por los que los lncRNA ejercen su función (revisado en Hu et al., 2012).

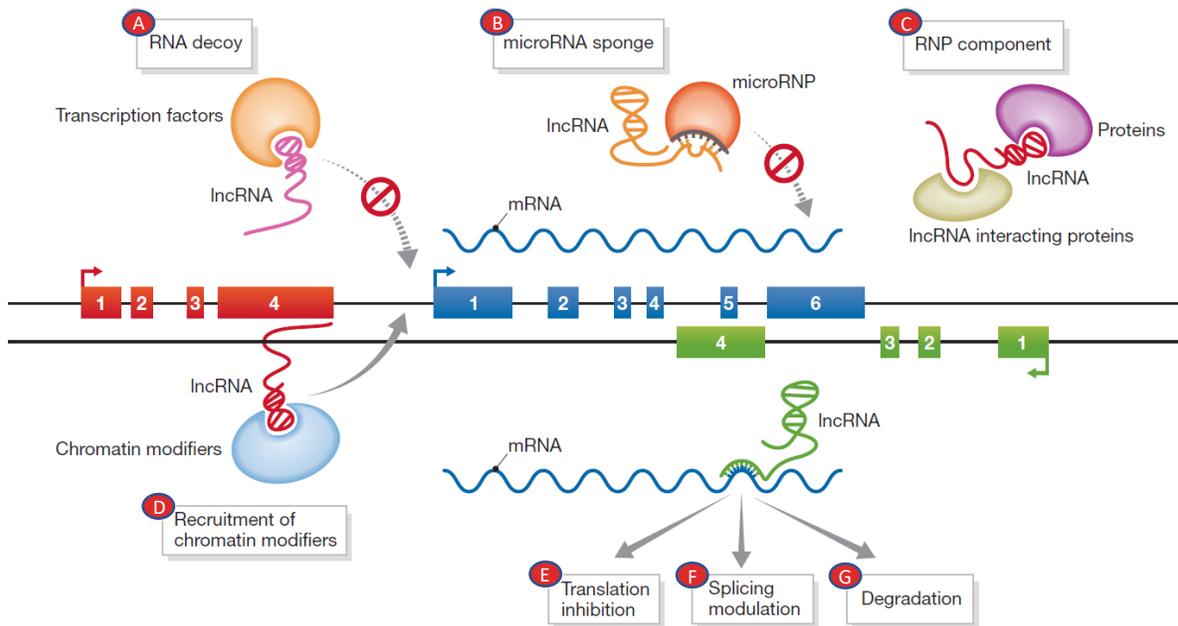


**Figura 6. Clasificación de lincRNAs por su relación con genes codificantes vecinos.** Las flechas rojas indican genes codificantes, las azules lincRNAs y la morada puede ser un gen codificante o no codificante (la forma de las figuras indica la dirección de transcripción). Los lincRNAs se dividen en cinco subtipos: los intergénicos (lincRNA), localizados entre dos genes codificantes (no importa la orientación), los intrónicos se encuentran en los intrones de los genes codificantes, mientras que los sentido se sobrelapan con exones, los bidireccionales se transcriben en direcciones opuestas (su sitio de inicio de la transcripción se encuentra a menos de 1 kb de distancia) y los antisentido o “Natural Antisense Transcript” (NAT) son secuencias de RNA que se transcriben de las dos cadenas del mismo *locus* en direcciones opuestas (tomado de Latg et al., 2018).



**Figura 7. Regulación en *cis* y *trans* de lncRNAs.**

**(A)** *Xist* induce la formación de cromatina represiva a través del reclutamiento de la DNA metiltransferasa 3 (DNMT3), que induce la metilación del DNA; PRC2, que trimetila la lisina 27 de la Histona H3 (H3K27me3) y EHMT2, que di- y trimetila la lisina 9 de la Histona H3 (H3K9me2 y H3K9me3). **(B)** *HOTAIR* es un lncRNA que tiene su *locus* en el “cluster” *HOXC*, actúa en *trans* en los genes del “cluster” *HOXD*. Sirve como andamio para dos complejos represivos distintos, PRC2 y el complejo desmetilante H3K4 KDM1A-coREST-REST en la misma región genómica dando como resultado la represión genética del *locus* (modificado de Fatica & Bozzini, 2014).



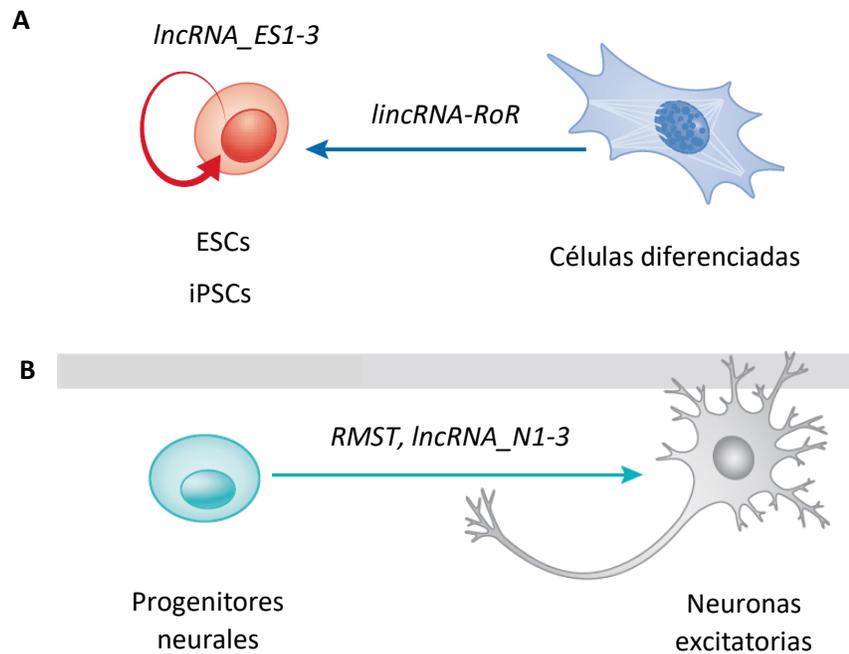
**Figura 8. Mecanismos de la función de lncRNAs.**

Los mecanismos por los cuales los lncRNAs regulan sus sitios blanco dependen de la secuencia primaria, estructura secundaria que adoptan y su *locus*. **(A)** Pueden actuar como señuelos donde secuestran TFs y evitan que se unan a su secuencia blanco. **(B)** Pueden tener secuencias complementarias con el complejo microRNA-ribonucleoproteína (microRNP) evitando su unión a la región 3'UTR del RNA mensajero (mRNA) y su posterior degradación. **(C)** En otros casos los lncRNAs sirven de andamio para unir proteínas reguladoras. **(D)** Otros lncRNAs reclutan complejos remodeladores de la cromatina a sus sitios diana del DNA. **(E)** Algunos otros lncRNAs (NAT) pueden modular el procesamiento de su mRNA blanco como la traducción, **(F)** el "splicing" alternativo y **(G)** la degradación (modificado de Hu et al., 2012).

### Regulación de la diferenciación celular por lncRNAs

En las últimas décadas se han identificado lncRNAs que participan en modular la diferenciación de células especializadas de mamífero a partir de sus progenitores, unos son requeridos para el mantenimiento del estado pluripotencial de ESCs o para la desdiferenciación (Figura 9A), otros para el mantenimiento de células progenitoras somáticas como *ANCR* (en queratinocitos) y *PINC* (en células epiteliales mamarias), otros participan en la regulación de genes que participan en el establecimiento del plan corporal (*HOTAIR*, *HOTTIP*, y *HOTAIRM1*), algunos otros asociados a la hematopoyesis (*HOTAIRM1*, *lincRNA-EPS* y *EGO*) o al desarrollo vascular (*Tie-1AS*).

También se han identificado lncRNAs que participan en la diferenciación de células musculares a partir de mioblastos (*linc-MD1*) y, con respecto a la diferenciación neural, se ha observado que muchos lncRNA están DE y específicamente se localizan en tejidos neurales durante el desarrollo mostrados en la Figura 9B (Hu et al., 2012).



**Figura 9. Participación de lncRNAs en la pluripotencia y en la diferenciación neural.** **(A)** Los lncRNAs *lncRNA\_ES1*, *lncRNA\_ES2* y *lncRNA\_ES3* modulan el mantenimiento del estado pluripotencial de ESCs mientras que *lincRNA-RoR* está involucrado en la reprogramación de células diferenciadas hacia iPSCs. **(B)** *RMST* y los lncRNAs *lncRNA\_N1*, *lncRNA\_N2* y *lncRNA\_N3* participan en la diferenciación de progenitores a neuronas excitatorias (modificado de Hu et al., 2012).

En mamíferos, aproximadamente el 40% de los lncRNAs se expresa exclusivamente en el cerebro (Statello et al., 2020). Además, existe una fuerte correlación entre la expansión del tamaño del cerebro a través de la evolución, con la expansión en el número de lncRNAs individuales, y ya que estos se expresan en linajes neuronales específicos, se puede especular que estos RNAs son de gran relevancia en el SN (Jarroux et al., 2017).

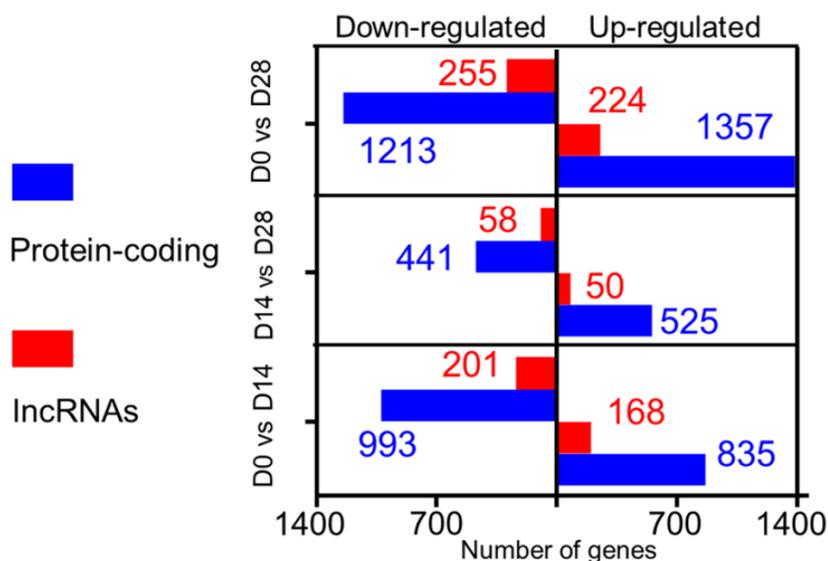
Actualmente, se sabe que muchos de los lncRNAs enriquecidos en el cerebro tienen papeles importantes en el neurodesarrollo y neuropatologías (Zhang et al., 2019).

A pesar de que se han identificado lncRNAs implicados en la pluripotencia de PSCs y en el proceso de diferenciación neural/neuronal, se conoce muy poco acerca de su función específica y el papel que juegan otros lncRNAs no caracterizados aún. Sin embargo, ya que los niveles de expresión de genes que codifican proteínas (PCGs) pueden ser regulados por lncRNAs en *cis*, es posible predecir la función potencial de estos lncRNA mediante análisis bioinformáticos donde se generen redes de coexpresión lncRNA-PCGs para posteriormente analizar en que procesos biológicos están asociados estos genes (Li et al., 2020). En este sentido, es posible que el estudio y manipulación de lncRNA permita el desarrollo de protocolos de diferenciación celular más eficientes (Perry & Ulitsky, 2016).

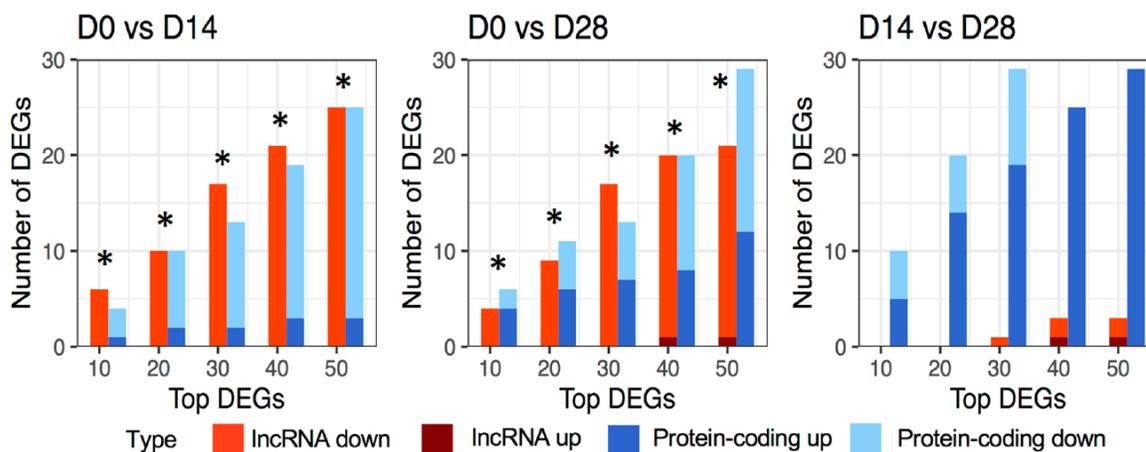
## **ANTECEDENTES**

### **Identificación de lncRNAs diferencialmente expresados (DE) durante el proceso de diferenciación dopaminérgica a partir de hESCs**

En un trabajo desarrollado por Meléndez-Ramírez y colaboradores donde se obtuvo el perfil transcripcional (RNA-seq) de hESC (H9-GFP) al día 0 (D0), día 14 (D14) y día 28 (D28) de diferenciación dopaminérgica, se encontraron diferencias significativas en los niveles de expresión de 4009 genes (3377 genes codificantes y 632 lncRNAs) en los intervalos 0-14 días (0-14 D), 14-28 días (14-28 D) y 0-28 días (0-28 D), donde se observa que los genes que mayormente cambian en todos los intervalos son genes codificantes de proteínas (PCGs) (Figura 10) (Meléndez-Ramírez et al., 2021). Sin embargo, al analizar los genes DE que tienen los cambios más drásticos en los intervalos 0-14 D y 0-28 D estos son lncRNAs (Figura 11).



**Figura 10. Genes DE durante la diferenciación dopaminérgica.** Número de genes DE tanto a la baja como al alta en los intervalos 0-14 D, 14-28 D y 0-28 D del proceso de diferenciación dopaminérgica a partir de hESCs (tomado de Meléndez-Ramírez et al., 2021).

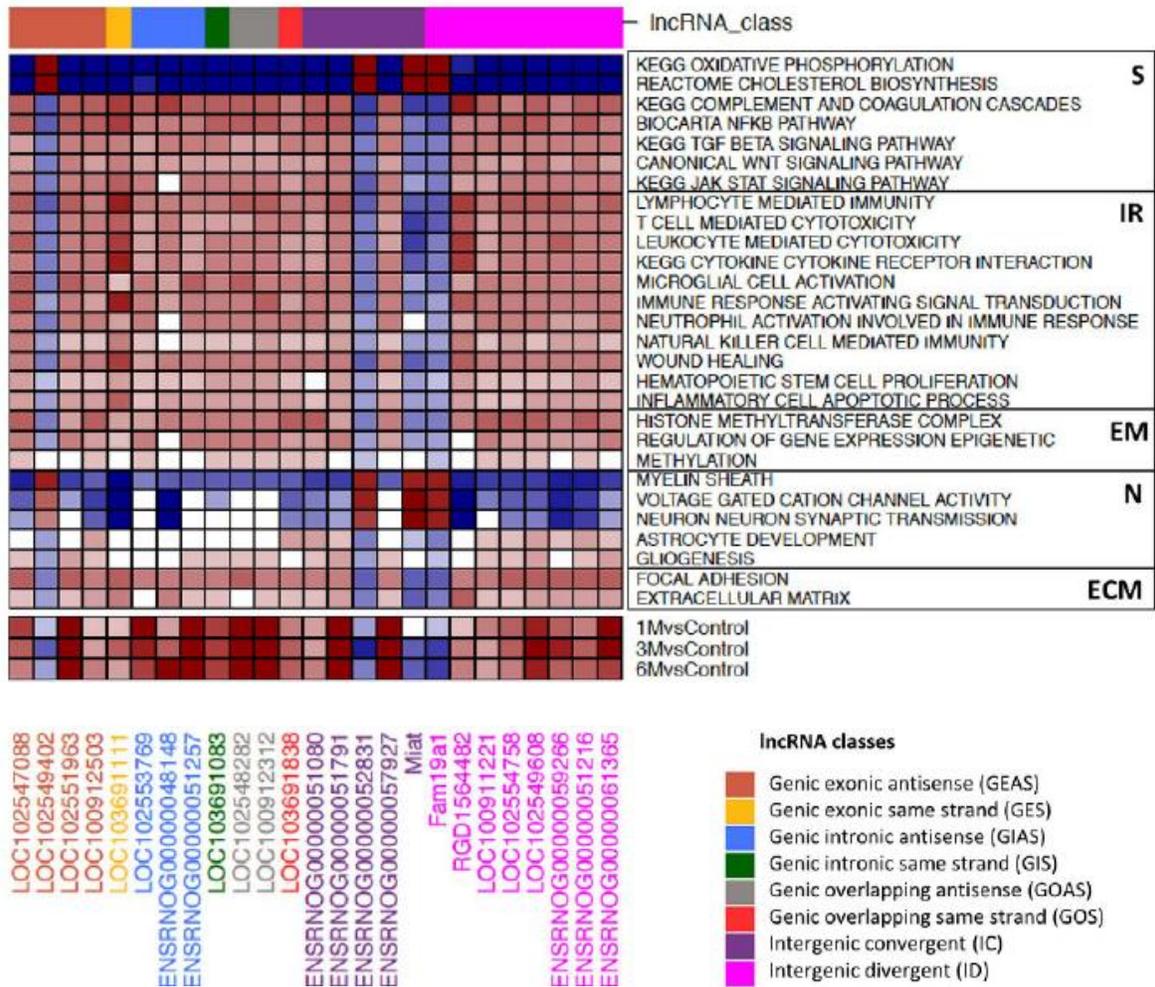


**Figura 11. Los lncRNAs tienen cambios dinámicos en su expresión.** Se muestran los 10, 20, 30, 40 y 50 genes DE que tienen la mayor tasa de cambio en los intervalos 0-14 D, 0-28 D y 14-28 D de la diferenciación dopaminérgica, así como la proporción de los genes que se expresan a la baja (“down”) como los que se sobreexpresan (“up”) en estos tres intervalos. El asterisco indica que el valor  $p$  de los lncRNAs es más bajo que el valor  $p$  de los PCGs (Meléndez-Ramírez et al., 2021).

## **Predicción de la función de lncRNAs por análisis bioinformáticos**

La función biológica de los lncRNAs se puede inferir usando herramientas bioinformáticas. De esta forma se han identificado lncRNAs que participan en la progresión de la lesión de médula espinal (SCI) subcrónica y crónica en modelos de rata examinando los cambios en la expresión de PCGs y lncRNAs a 1 mes (1 M), 3 meses (3 M) y 6 meses (6 M) después de la lesión mediante RNA-seq. Los datos de secuenciación se normalizaron a FPKMs (fragmentos por kilobase por millón de lecturas mapeadas), se obtuvieron genes DE tomando tres criterios principales: 1) que al menos una de las muestras tenga un FPKM > 1; 2) que las veces de cambio (FC) sean > 2; 3) que tenga un “False Discovery Rate” (FDR) < 0.01 (Duran et al., 2017).

Para inferir la función potencial de lncRNAs en la progresión de SCI, se realizó una matriz de asociación de los lncRNAs seleccionados y las vías enriquecidas usando ontología de genes (Figura 17). Se encontraron cinco categorías en las que están implicados estos genes: 1) Vías de señalización (S); 2) Respuesta inmune (IR); 3) Modificación epigenética (EM); 4) SN (N); 5) Matriz Extracelular (EMC). Estos análisis bioinformáticos permiten la inferencia de la posible función de lncRNAs en un determinado contexto biológico (Duran et al., 2017).



**Figura 12. Predicción de la función de lncRNAs en SCI en rata.**

Matriz de asociación donde se muestran los lncRNAs DE en las columnas y en las filas se encuentran las vías enriquecidas por ontología de genes y agrupadas en cinco términos funcionales, donde la intensidad de color indica la fuerza de asociación entre el lncRNA y la vía enriquecida (modificado de Duran et al., 2017).

## **JUSTIFICACIÓN**

El estudio de los lncRNAs permitirá ampliar el conocimiento básico sobre los factores y mecanismos que modulan el mantenimiento de PSCs y el proceso de diferenciación neuronal dopaminérgica.

## **HIPÓTESIS**

La decisión entre el mantenimiento de la troncalidad o la diferenciación neuronal dopaminérgica de hESCs está parcialmente modulada por lncRNAs específicos al regular la expresión de genes importantes en estos procesos celulares.

## **OBJETIVOS**

### **Objetivo general**

Identificar posibles lncRNAs que participan en la regulación de la expresión de genes importantes en el estado pluripotencial de hESCs y durante la diferenciación dopaminérgica mediante análisis bioinformáticos.

### **Objetivos particulares**

1. Validar por RT-qPCR lncRNAs seleccionados al alza y a la baja que fueron identificados por RNA-seq al diferenciar hESCs a nDAs.
2. Evaluar los procesos biológicos en los que participan los lncRNAs mediante análisis bioinformáticos.
3. Identificar redes potenciales de regulación entre lncRNAs y PCGs DE en la diferenciación dopaminérgica.

## DISEÑO EXPERIMENTAL

A partir de los datos obtenidos del perfil transcripcional (RNA-seq) de hESCs a los 0, 14 y 28 días de diferenciación dopaminérgica, se tomaron los 30 lncRNAs más diferencialmente expresados tanto al alza como a la baja en el intervalo 0-28 días (de acuerdo con sus veces de cambio en logaritmo base dos [ $\log_2FC$ ]) y se analizaron con la base de datos LNCipedia (Volders et al., 2013). Esta base de datos contiene información básica de 21 488 lncRNA humano anotadas y además la predicción de la estructura secundaria, el potencial para codificar proteínas, sitios de unión de microRNAs y bibliografía disponible de lncRNAs específicos. De los lncRNAs evaluados se seleccionaron como controles para la validación por RT-qPCR a *ESRG*, el cual, está involucrado en la pluripotencia de hPSCs (Zhao et al., 2007) y que tuvo niveles de expresión altos solo en el D0 y a *MIR124-2HG*, sobreexpresado en el D28, el cual se expresa exclusivamente en cerebro y es precursor de la familia de *miR-124*, estos *miRs* están involucrados en la diferenciación neuronal del SNC (Makeyev et al., 2007) y en el desarrollo neural y crecimiento dendrítico (Yoo et al., 2009). Para la evaluación de la expresión de lncRNAs se seleccionaron cuatro transcritos que cumplieron los siguientes criterios: 1) tuvieron un FPKM > 1, al menos en uno de los tiempos en los que se comparó sus niveles de expresión (al tomar este corte se aseguró que el lncRNA se está expresando); 2) que mayormente cambiaron de acuerdo con los datos del RNA-seq (dos a la baja y dos al alta); 3) que el diseño de los cebadores fuera posible. Los lncRNAs que cumplieron estos criterios fueron: *AC019155.2* Y *LINC00678* (a la baja) y, *DIO3OS* y *AC018563.1* (al alza).

Para validar los niveles de expresión de lncRNAs diferencialmente expresados (DE), se partió de RNA total de hESCs (D0) y de neuronas dopaminérgicas (D28). El RNA total se trató con DNasa I para eliminar el posible DNA genómico. A continuación, se sintetizó el cDNA con transcriptasa reversa y oligonucleótidos dT como cebadores y se cuantificó la concentración de cDNA con un espectrofotómetro Nanodrop. Posteriormente, con el cDNA obtenido y los cebadores previamente diseñados para cada lncRNA seleccionado, se realizaron qPCRs utilizando SYBR® Green Master con el objetivo de cuantificar los niveles de expresión de estos transcritos. Los niveles relativos de la expresión génica se

normalizaron con *ACTB* (gen que codifica para la proteína beta actina) por el método de Livak (Livak & Schmittgen, 2001).

Posteriormente, a partir de la información del RNA-seq se realizó un análisis de enriquecimiento de conjunto de genes (GSEA) de los lncRNAs DE donde se asociaron a procesos biológicos. A continuación, se buscaron lncRNAs que, aunque no mostraron cambios muy drásticos en su expresión durante el proceso de diferenciación, estuvieron fuertemente asociados a procesos celulares relacionados al estado pluripotencial, neurogénesis y diferenciación neuronal dopaminérgica. Con la finalidad de inferir su participación en la regulación epigenética de la expresión genética, se identificaron EFs DE y se construyó una red de coexpresión entre lncRNAs seleccionados y EFs. También se realizó una red de coexpresión entre lncRNAs y PCGs que se encontraron cercanos genómicamente con la finalidad de establecer una posible regulación *in cis*.

Al realizar los experimentos anteriores se identificaron posibles lncRNAs que participan en la regulación tanto del estado pluripotencial de hESCs como en el proceso de diferenciación dopaminérgica.

## **METODOLOGÍA**

### **Extracción de RNA total**

Para la lisis celular de hESC H9-GFP (D0) y neuronas dopaminérgicas (D28) se añadieron 0.75 mL de TRIzol™ LS Reagent por cada 0.25 mL de muestra, se homogeneizó y se incubó por 5 min. Posteriormente se añadieron 0.2 mL de cloroformo, se incubó por 3 min y se llevó a centrifugación por 15 min a 1310 g a 4°C para la separación de fases. Después la fase acuosa se transfirió a un tubo nuevo y se incubó con 0.5 mL de isopropanol para la precipitación del RNA. Se realizaron 3 lavados del RNA a concentraciones crecientes de etanol y se dejó secar la pastilla por 10 min. Finalmente, el RNA se resuspendió con agua libre de RNasas.

### **Tratamiento del RNA total con DNasa I**

Se eliminó el DNA genómico (gDNA) de las muestras de RNA total con el kit RQ1 RNase-Free DNase (Promega). Primero se tomaron 8  $\mu$ L de RNA total y se añadió 1  $\mu$ L del Buffer 10X RQ1 y 1U de DNasa libre de RNasa para un volumen total de 10  $\mu$ L. Después, se homogeneizó la mezcla, se incubó a 37°C por 30 min y posteriormente se añadió 1  $\mu$ L del reactivo Stop solution y se incubó nuevamente 10 min a 65°C para detener la reacción. Se observó su integridad con gel de agarosa al 1% (Anexo I), y se almacenó a -20°C para su uso posterior.

### **Síntesis de DNA complementario (cDNA) y qPCR**

A partir del RNA total y con el kit High Capacity cDNA Reverse Transcription (Applied Biosystems) y oligonucleótidos (dT) (ya que los datos de secuenciación se obtuvieron por RNA-seq de PolyA), se preparó un “master mix” para cuatro reacciones (tres de muestras y una como control negativo). El “master mix” para la retrotranscripción se colocó en un microtubo estéril el cuál consistió de: 8  $\mu$ L de buffer 10X de transcripción reversa, 3.2  $\mu$ L del “mix” dNTP 25X (100mM), 4  $\mu$ L de oligonucleótidos (dT) (50  $\mu$ M), 4  $\mu$ L de la enzima retrotranscriptasa, 4  $\mu$ L de inhibidor de RNasa y se agregaron 16.8  $\mu$ L de agua libre de RNasa para llevarlo a un volumen final de 40  $\mu$ L. Posteriormente, para cada muestra se transfirieron 10  $\mu$ L del “master mix” a un microtubo, se agregó 1  $\mu$ g de RNA y se llevó a un volumen final de 20  $\mu$ L con agua estéril (en el caso del control negativo se agregaron 10  $\mu$ L de agua estéril). Enseguida se llevaron los microtubos al termociclador donde se programaron 4 pasos: 1) 10 minutos a 25°C; 2) 120 minutos a 37°C; 3) 5 minutos a 85°C; 4) mantener a 4°C. El cDNA sintetizado se cuantificó nuevamente mediante NanoDrop™ y se observó en gel de agarosa al 2% (anexo II), posteriormente las muestras se congelaron a -20°C.

Para realizar la cuantificación de los lncRNAs seleccionados se realizó el diseño de cebadores en la plataforma Primer-BLAST del NCBI (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi>), posteriormente se evaluó la capacidad que tiene cada primer de generar horquillas, homodímeros y/o heterodímeros

con OligoAnalyzer de IDT (<https://www.idtdna.com/calc/analyzer>), y finalmente se realizó una PCR *in silico* para corroborar que los cebadores fueran específicos, hibridaran en zonas de interés y generaran el amplicón esperado usando la plataforma UCSC *In-Silico* PCR (<https://genome.ucsc.edu/cgi-bin/hgPcr>). Los pares de cebadores diseñados y sintetizados, así como el tamaño de su amplicón esperado en pb se muestran en la Tabla I.

En cada reacción de RT-qPCR se adicionó 5 µL de iTaq™ Universal SYBR® Green Supermix (Bio-Rad), 1 µL del cebador directo (500 nM), 1 µL del cebador reverso (500 nM), 50 ng de la muestra y se llevó a un volumen final de 10 µL con agua libre de DNAsa, se mezcló gentilmente la solución en cada pozo y para cada reacción se realizaron dos réplicas técnicas. La placa se tapó y se colocó en el termociclador de tiempo real CFX96 Touch de Bio-Rad (USA). Usando el software CFX maestro se hizo el diseño de la placa y se programó el equipo con los siguientes pasos: 1) activación de la polimerasa a 95°C por 30 segundos; 2) 5 segundos a 95°C para desnaturalizar el DNA; 3) hibridación de los cebadores, la extensión y la lectura de la placa a 60°C por 30 segundos (el paso 2 y 3 se repiten 39 veces); 4) análisis de la curva de fusión 5 segundos a 65°C e incrementos de 0.5°C por 5 segundos hasta llegar a 95°C. Una vez que se corrió el programa se obtuvo la información con la cual se realizaron los cálculos y los análisis estadísticos.

**Tabla I. Lista de cebadores diseñados para la validación por RT-qPCR.**

<b>DE A LA BAJA</b>				
<b>Gen</b>	<b>Cadena</b>	<b>Cebadores</b>	<b>Tm</b>	<b>Tamaño Producto</b>
<b><i>ESRG</i></b>	F	ACAGCCTTGTACCCTGGTCT	59.69	83
	R	CAATGGTGCGAAGCTGTGTT	60.41	
<b><i>AC019155.2</i></b>	F	AGACTTGGGCAAATCCAAC	58.04	80
	R	ACTGGCATTGATTGAAGTGGT	58.77	
<b><i>LINC00678</i></b>	F	ACCAGTCTGCCCTGCTCATA	60.62	103
	R	TCACACTCCCCTCATTCTGC	60.04	
	R	GCTTCTCGGCTGAATCCCTC	60.53	

<b>DE AL ALZA</b>				
<b>Gen</b>	<b>Cadena</b>	<b>Cebadores</b>	<b>Tm</b>	<b>Tamaño Producto</b>
<b>DIO3OS</b>	F	CATCCTCTTCAGCGGTCCTC	59.90	161
	R	CCTGGTGCTGGTTTGGGTTA	60.18	
<b>AC018563.1</b>	F	TAGCTGGGTTTGATGCGTGT	59.96	226
	R	GGATAGGATCGGGTCAGGGA	59.89	
<b>MIR124-2HG</b>	F	GCCGTGAGAGTCCTTTGACA	59.97	264
	R	TCCTCCGCTTATTGCCACTG	60.11	
	R	GCAAGCAGGTCCTCAAGTCA	60.25	
<b>Control</b>				
<b>Gen</b>	<b>Cadena</b>	<b>Cebadores</b>	<b>Tm</b>	<b>Tamaño Producto</b>
<b>ACTB</b>	F	GCTATCCAGGCTGTGCTATC	57.5	166
	R	TGAGGTAGTCAGTCAGGTCC	55.1	

### **Análisis de sobrerrepresentación de conjunto de genes y matriz de asociación**

A partir de los datos de RNA-seq y utilizando el lenguaje de programación R se obtuvo un “Pearson Correlation Coefficient” (PCC) con la función [cor()] entre los lncRNAs DE (FPKM  $\geq$  1 en al menos uno de los tiempos medidos,  $\log_2FC > 2$  o  $< -2$  y un FDR  $< 0.05$  por DESeq2) y cada uno de los genes codificantes en las 10 observaciones (tres del D0, cuatro del D14 y tres del D28) usando el logaritmo base 2 de sus respectivos FPKM. Los genes que no tuvieron cambios en sus niveles de expresión entre los tiempos medidos se eliminaron (D0, D14 y D28). Con el PCC calculado se generó una matriz de correlación entre los lncRNAs DE (632) y los genes codificantes (17,594). Posteriormente, se generó una lista ordenada de PCGs para cada lncRNA DE en base a el PCC. El análisis de sobrerrepresentación de conjunto de genes (GSEA) se llevó a cabo, este arrojó una puntuación normalizada de enriquecimiento (NES) que indica la fuerza y dirección de asociación de los conjuntos de genes (GS) con cada lncRNA. Se obtuvieron 4781 GS significativamente enriquecidos con un FDR  $< 0.05$ . Para realizar la matriz de asociación solamente 13 GS fueron seleccionados por

su importancia en el estado troncal de PSCs y el proceso de diferenciación neuronal dopaminérgica. Se eligieron 25 lncRNAs (10 seleccionados por su alto  $\log_2FC$  y 15 lncRNAs con alto NES en GS seleccionados y altamente correlacionados con genes importantes en la diferenciación dopaminérgica) para generar los mapas de calor con el software y lenguaje de programación R.

### **Construcción de redes de coexpresión de lncRNAs y mRNAs de EFs**

Se identificaron EFs DE en el proceso de diferenciación dopaminérgica usando la base de datos EpiFactors (Medvedeva et al., 2015) ya que contiene una amplia lista de reguladores epigenéticos de humano, sus complejos, blancos y productos abarcando 200 tipos celulares. A partir del PCC se buscaron y filtraron EFs con un  $PCC \geq 0.98$ . Posteriormente, la red de coexpresión se realizó con el “software” Cytoscape v3.8.2 ya que para los fines de este proyecto, es un programa que permite generar redes de interacción de forma intuitiva y rápida (Shannon et al., 2003)

### **Construcción de redes de coexpresión entre lncRNAs y PCGs asociados**

Se usó el software Integrative Genomics Viewer (IGV) v2.9.4 (Thorvaldsdóttir et al., 2013) por su uso fluido e intuitivo para visualizar los datos transcriptómicos previamente obtenidos (Meléndez-Ramírez et al., 2021) con la finalidad de analizar el *locus* de cada uno de los 25 lncRNAs seleccionados e identificar los PCGs que se encuentran en la vecindad de cada lncRNA (< 5 Kb). Posteriormente, se evaluaron y filtraron genes DE y altamente correlacionados con su lncRNA asociado usando el  $PCC > 0.8$ . Finalmente, se hizo una búsqueda de los procesos celulares en los cuales participa cada PCG. Esta información se integró usando el software Cytoscape v3.8.2.

### **Análisis estadístico**

El método de Livak ( $2^{-\Delta\Delta Ct}$ ) se usó para la cuantificación relativa de la expresión por qPCR. La prueba *t* de Student pareada de dos colas se utilizó para comparar la expresión de genes en

los tiempos (D0 y D28). Se usó como gen normalizador *ACTB*, la significancia fue establecida cuando  $p < 0.05$ .

## RESULTADOS

### Identificación de lncRNAs diferencialmente expresados en la diferenciación dopaminérgica implicados en estado pluripotencial de hESCs y diferenciación neural

El perfil transcripcional del proceso de diferenciación dopaminérgica identificó a 370 lncRNAs DE en el intervalo 0-14 D, 109 lncRNAs en el 14-28 D y 480 en el intervalo 0-28 D. Los 30 lncRNAs DE que tuvieron la mayor tasa de cambio tanto a la baja como al alza en el intervalo 0-28 D (ventana de tiempo que comprende el paso del estado troncal de hESCs a la diferenciación neuronal dopaminérgica), se analizaron con la base de datos LNCpedia. En esta base de datos se identificaron lncRNAs reportados previamente asociados a la troncalidad de hESCs y al proceso de diferenciación neuronal (Tabla II y III). Entre estos lncRNAs se encuentran:

1. *ESRG* (*Embryonic Stem cells Related Gene*). Se expresa altamente en células hESCs indiferenciadas. Los niveles de expresión de este transcrito se detectaron por microarreglos donde se diferenciaron las hESCs a un linaje ectodérmico al inducir las con ácido retinoico (RA) (Zhao et al., 2007).
2. *LINC01108* (*lncRNA\_ES1*). Se encuentra sobreexpresado en hESCs e iPSCs en comparación con tejidos somáticos. Los cambios de expresión de este lncRNA se analizaron mediante microarreglos donde al hacer un “Knock-down” en hESCs, estas se diferenciaron espontáneamente hacia células de las tres capas germinales (Ng et al., 2011).
3. *LNCPRESS1*. Está sobreexpresado en hESCs y en iPSCs en comparación con células de las tres capas germinales. Su nivel de expresión se detectó por RNA-seq al depletar este transcrito mediante RNA de interferencia (Jain et al., 2016).
4. *MIR99AHG* (precursor de los microRNAs *miR-10* y *miR-Let-7*). La familia de *miR-10* (asociada a la regulación de los genes *HOX*), se sobreexpresa en células precursoras

neurales derivadas de hESCs. La familia de *miR-Let-7* se expresa en niveles altos en neuronas. La elevación en la expresión de ambos *miRs* se detectó a partir de datos con la técnica de microarreglos al aplicar a células hESCs un protocolo de diferenciación neural con RA y factores neurotróficos (Parsons et al., 2012).

5. *HAGLR (HOXD-AS1)*. Diferencialmente expresado al alza durante la diferenciación neural. La dinámica de expresión de este transcrito se detectó por microarreglos al diferenciar células de neuroblastoma humano (SH-SY5Y) a neuronas al inducir las con RA (Batagov et al., 2013).
6. *SOX1-OT*. Está sobreexpresado en células troncales neurales. Su expresión se cuantificó por RT-qPCR donde se partió de células progenitoras neurales humanas inmortalizadas las cuales se indujeron a diferenciación neural. Además, su nivel de expresión se correlacionó con la expresión del gen *SOX1* (importante en el desarrollo del SNC y el mantenimiento de células troncales neurales) (Ahmad et al., 2017).

### **Validación de lncRNAs seleccionados durante la diferenciación dopaminérgica a partir de hESC por RT-qPCR**

Con el objetivo de corroborar que los lncRNAs DE presentan cambios en su nivel de expresión observados en el perfil transcripcional, se seleccionaron dos lncRNAs DE a la baja (*LINC00678* y *AC019155.2*) y dos DE al alza (*DIO3OS* y *AC018563.1*) para su validación por RT-qPCR (marcados con asterisco en las Tablas II y III, respectivamente). Además, se eligieron los lncRNAs *ESRG* (sobreexpresado en hESCs) y *MIR124-2HG* (sobreexpresado en cerebro) como controles.

El diseño de los cebadores para estos lncRNAs se evaluó por PCR punto final (Figura 13). Posteriormente, se validaron estos lncRNAs por qPCR con tres muestras biológicas independientes para el D0 y lo mismo para el D28 (Figura 14). Al realizar una comparación entre la expresión de estos lncRNAs del intervalo 0-28 D con los datos de secuenciación y la cuantificación relativa obtenida de la qPCR, se obtuvo una  $R^2 = 0.9629$  lo que indica una alta correlación entre los niveles de expresión de los lncRNAs con ambos métodos.

**Tabla II. Los 30 lncRNAs más DE a la baja en el intervalo 0-28 D analizados por LNCpedia.**

Top	Gene ID	lncRNA	log <sub>2</sub> FC	Literatura disponible	Locus (hg38)
1	ENSG00000253507.5	<i>AC104257.1</i>	-9.16	SIN DATOS	chr8:131308545-131317583
2	ENSG00000265992.1	(+) <i>ESRG</i>	-8.91	(Zhao et al., 2007)	chr3:54632122-54639857
3	ENSG00000255666.6	<i>AP000943.2</i>	-8.82	SIN DATOS	chr11:94641907-94651628
4	ENSG00000279481.1	<i>AC104791.2</i>	-8.64	SIN DATOS	SIN DATOS
5	ENSG00000254934.5	(*) <i>LINC00678</i>	-8.54	SIN DATOS	chr11:27599344-27634720
6	ENSG00000275216.2	<i>AL161431.1</i>	-8.46	SIN DATOS	chr13:109269634-109273838
7	ENSG00000236673.5	<i>AL117378.1</i>	-8.38	SIN DATOS	chr6:131901963-131920565
8	ENSG00000219445.2	(*) <i>AC019155.2</i>	-8.38	SIN DATOS	chr7:125229579-125264291
9	ENSG00000260834.1	<i>AC009055.2</i>	-8.21	SIN DATOS	chr16:65190973-65234914
10	ENSG00000279882.1	<i>AC112702.1</i>	-8.09	SIN DATOS	SIN DATOS
11	ENSG00000226673.2	(+) <i>LINC01108</i>	-8.07	(Ng et al., 2011)	chr6:14280139-14285746
12	ENSG00000282024.1	<i>AL009031.1</i>	-8.04	SIN DATOS	chr6:16262682-16264553
13	ENSG00000254339.5	<i>AC064802.1</i>	-8.01	SIN DATOS	chr8:114282136-114287996
14	ENSG00000251292.1	<i>ERVH-1</i>	-8.00	SIN DATOS	chr4:23723262-23733579
15	ENSG00000259048.1	<i>AL392023.2</i>	-7.93	SIN DATOS	chr14:38034287-38194281
16	ENSG00000215866.7	<i>LINC01356</i>	-7.88	SIN DATOS	chr1:112813210-112874625
17	ENSG00000261780.2	<i>LINC02582</i>	-7.86	SIN DATOS	chr18:73324941-73349889
18	ENSG00000283403.1	<i>AC005394.2</i>	-7.80	SIN DATOS	chr19:28435388-28727680
19	ENSG00000282012.1	<i>AP000357.2</i>	-7.79	SIN DATOS	chr22:24691122-24696003
20	ENSG00000253348.1	<i>AC008514.1</i>	-7.55	SIN DATOS	chr5:170767204-170788650
21	ENSG00000256288.1	<i>AC022075.2</i>	-7.46	SIN DATOS	chr12:10358464-10361045
22	ENSG00000283633.1	<i>AP000547.3</i>	-7.44	SIN DATOS	chr22:16601911-16615111
23	ENSG00000230937.11	<i>MIR205HG</i>	-7.38	SIN DATOS	chr1:209428820-209432545
24	ENSG00000280511.1	<i>AL591030.1</i>	-7.38	SIN DATOS	chr6:76561328-76562590
25	ENSG00000278910.4	<i>BANCR</i>	-7.33	SIN DATOS	chr9:69296678-69307100
26	ENSG00000238178.6	<i>AC078993.1</i>	-7.33	SIN DATOS	chrX:16153220-16170869
27	ENSG00000267696.6	<i>ERVK-28</i>	-7.33	SIN DATOS	chr19:27638509-27646483
28	ENSG00000232301.2	(+) <i>LNCPRESS1</i>	-7.26	(Jain et al., 2016)	chr7:101308340-101321149
29	ENSG00000233056.2	<i>ERVH48-1</i>	-7.22	SIN DATOS	chr21:42916803-42925646
30	ENSG00000232973.12	<i>CYP1B1-AS1</i>	-7.12	SIN DATOS	chr2:38082698-38131567

(+) Controles positivos para pluripotencia

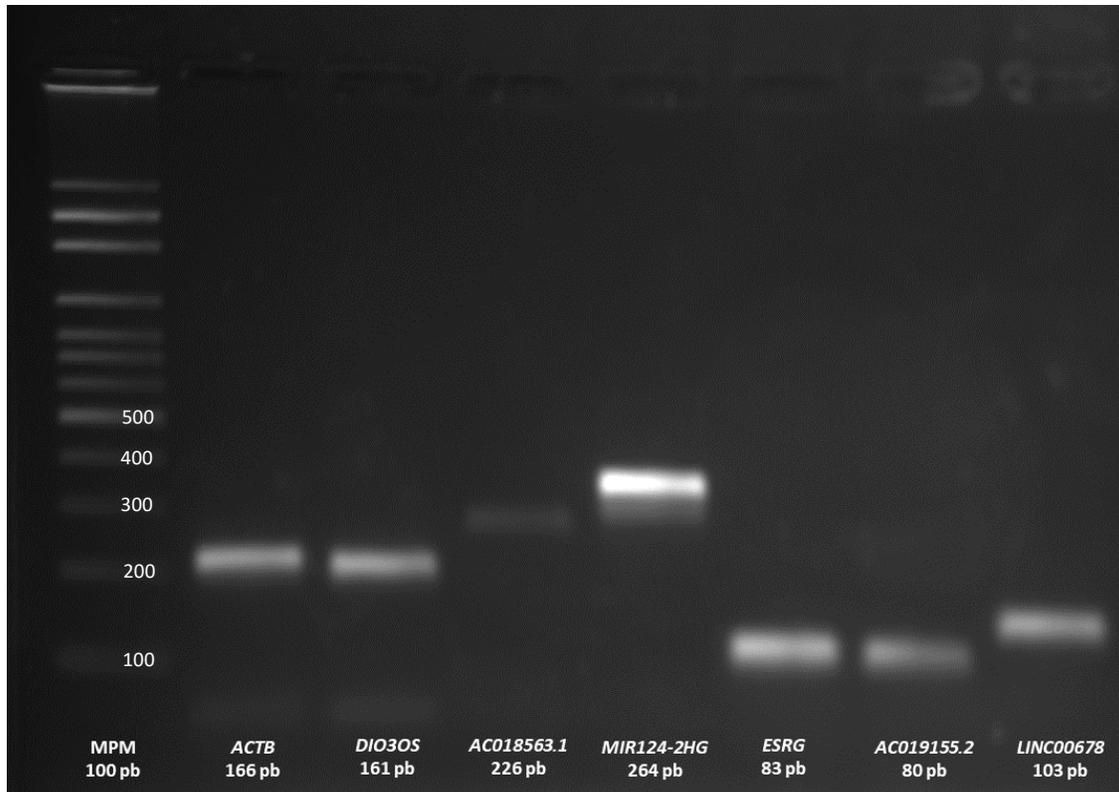
(\*) lncRNAs a evaluar

**Tabla III. Los 30 lncRNAs más DE al alza en el intervalo 0-28 D analizados por LNCpedia.**

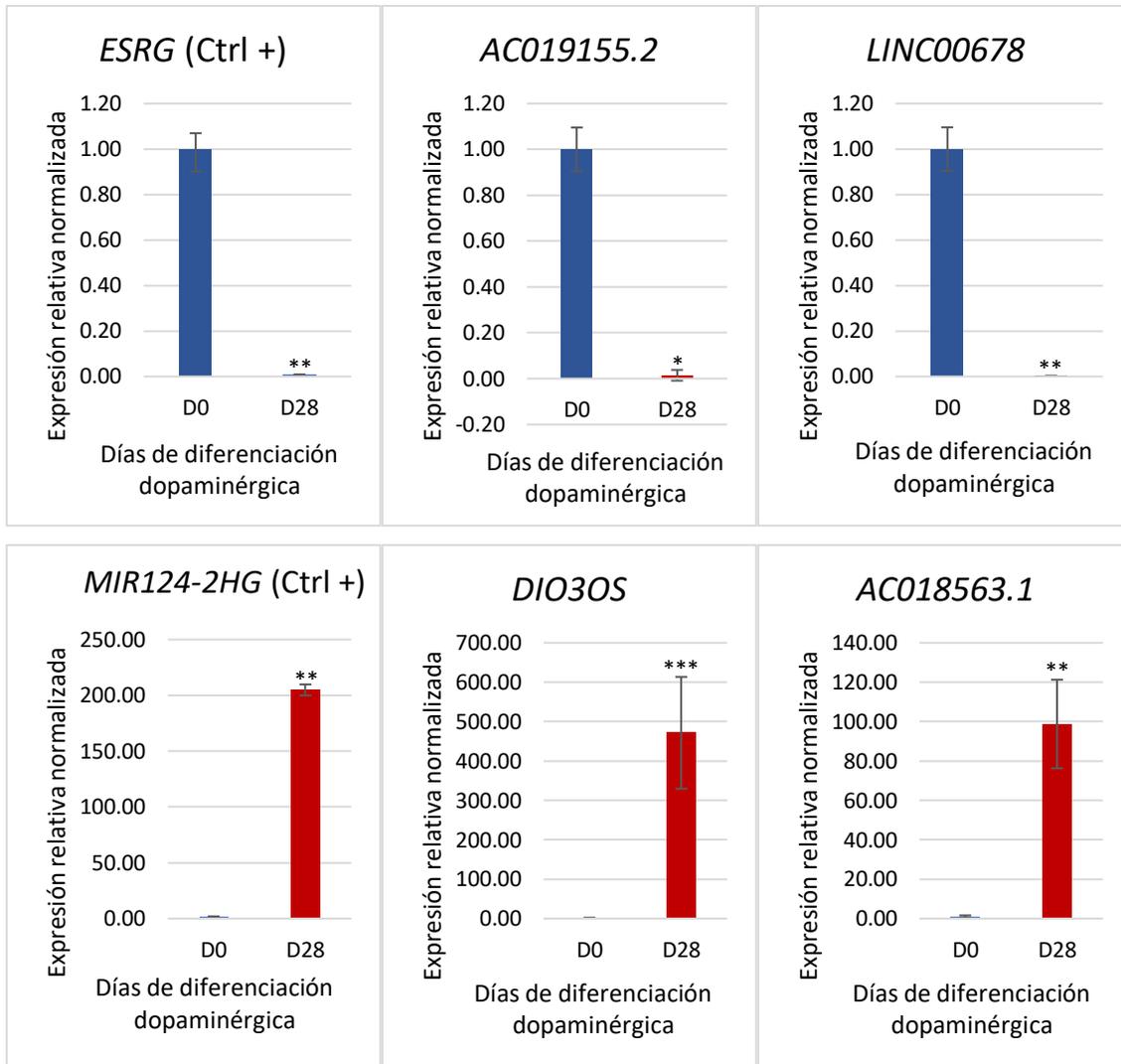
Top	Gene ID	lncRNA	log2FC	Literatura disponible	Locus (hg38)
1	ENSG00000258498.8	(*) <i>DIO3OS</i>	7.70	Correlación alta con <i>DIO3</i>	chr14:101557304-101560431
2	ENSG00000259280.1	(*) <i>AC018563.1</i>	7.12	SIN DATOS	chr15:36881018-36887035
3	ENSG00000254377.5	(+) <i>MIR124-2HG</i>	5.96	Activación de astrocitos	chr8:64373218-64378827
4	ENSG00000215386.12	(+) <i>MIR99AHG</i>	5.85	(Parsons et al., 2012)	chr21:16621782-16643063
5	ENSG00000230148.8	<i>HOXB-AS1</i>	5.79	Glioblastoma	SIN DATOS
6	ENSG00000237149.5	<i>ZNF503-AS2</i>	5.72	Cáncer gástrico	chr10:75403876-75408308
7	ENSG00000227128.4	<i>LBX1-AS1</i>	5.54	SIN DATOS	chr10:101229594-101232605
8	ENSG00000279249.2	<i>AC007614.1</i>	5.51	SIN DATOS	chr16:49287430-49290227
9	ENSG00000224189.7	<i>HAGLR</i>	5.41	(Yarmishyn et al., 2014)	chr2:176176456-176188958
10	ENSG00000262223.7	<i>AC110285.1</i>	5.18	SIN DATOS	chr17:81376022-81385129
11	ENSG00000224243.1	(+) <i>SOX1-OT</i>	4.92	(Ahmad et al., 2017)	chr13:111972310-112108015
12	ENSG00000253552.7	<i>HOXA-AS2</i>	4.91	Leucemia	chr7:27122443-27123818
13	ENSG00000250208.6	<i>FZD10-DT</i>	4.82	SIN DATOS	chr12:130126852-130162365
14	ENSG00000248127.1	<i>AC026774.1</i>	4.81	SIN DATOS	chr5:76081946-76084186
15	ENSG00000279382.1	<i>AC018665.1</i>	4.77	SIN DATOS	SIN DATOS
16	ENSG00000245662.3	<i>LINC02211</i>	4.74	SIN DATOS	chr5:25190682-25306172
17	ENSG00000254226.5	<i>LINC01933</i>	4.64	SIN DATOS	chr5:152195448-152270445
18	ENSG00000245870.3	<i>LINC00682</i>	4.61	SIN DATOS	chr4:41876481-41882611
19	ENSG00000257060.6	<i>AC091078.1</i>	4.57	SIN DATOS	chr15:93312557-93530262
20	ENSG00000247809.7	<i>NR2F2-AS1</i>	4.50	SIN DATOS	chr15:96268772-96327361
21	ENSG00000229520.1	<i>LINC00404</i>	4.46	SIN DATOS	chr13:112106095-112106882
22	ENSG00000261888.1	<i>AC144831.1</i>	4.46	SIN DATOS	chr17:83104255-83106910
23	ENSG00000253661.1	<i>ZFH4-AS1</i>	4.42	SIN DATOS	chr8:76610881-76683278
24	ENSG00000257935.2	<i>LHX5-AS1</i>	4.42	Dif. endodermal	chr12:113472003-113480481
25	ENSG00000225792.1	<i>AC004540.2</i>	4.42	SIN DATOS	chr7:26372144-26376272
26	ENSG00000233429.9	<i>HOTAIRM1</i>	4.39	Abuso de cocaína	chr7:27096112-27100154
27	ENSG00000277954.1	<i>AC092376.2</i>	4.36	SIN DATOS	chr16:79202624-79206739
28	ENSG00000258584.2	<i>FAM181A-AS1</i>	4.35	SIN DATOS	chr14:93925233-93927066
29	ENSG00000231160.9	<i>KLF3-AS1</i>	4.30	SIN DATOS	chr4:38625334-38664883
30	ENSG00000260249.2	<i>AC007608.3</i>	4.27	SIN DATOS	SIN DATOS

(+) Controles positivos para diferenciación neural

(\*) lncRNAs a evaluar



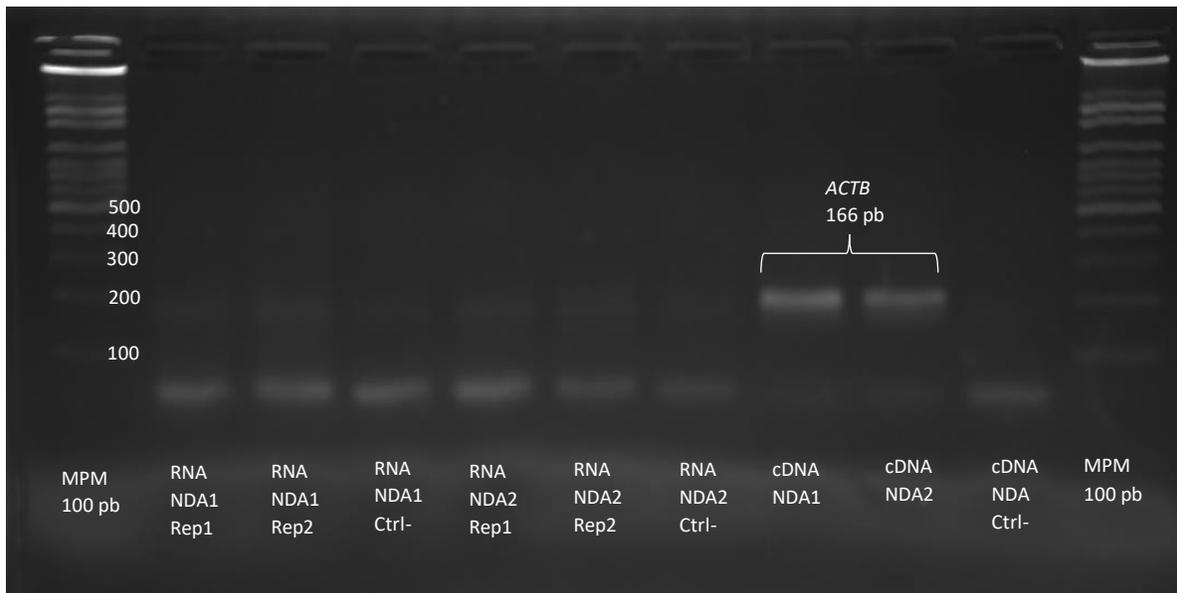
**Figura 13. Productos de PCR punto final de lncRNAs seleccionados para validación.** Para los lncRNAs DE al alza se usó una muestra del D28 y para los lncRNAs DE a la baja se usó una muestra del D0. Se muestra el tamaño del amplicón esperado para cada gen. MPM 100 pb: Marcador de peso molecular de 100 pb, pb: pares de bases.



**Figura 14. Expresión de lncRNAs por RT-qPCR.**

Por el método de Livak ( $2^{-\Delta\Delta Ct}$ ) se obtuvo la expresión normalizada relativa de los lncRNAs usando *ACTB* como gen normalizador en D0 y D28 usando 3 muestras biológicas distintas para cada estadio de diferenciación. Las barras de error indican la desviación estándar. Para identificar diferencias estadísticamente significativas, se usó el estadístico de prueba *t*-Student pareada de dos colas, \*:  $p < 0.05$ , \*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$  comparadas con el control (D0).  $n=3$ . Ctrl +: control positivo.

Para descartar que la detección de los transcritos se atribuye al posible DNA genómico presente en las muestras, se realizó una PCR con 2 muestras de RNA del D28 tratadas con DNasa (NDA1 y NDA2) que fueron usados para la síntesis de cDNA (Figura 15).



**Figura 15. Productos de PCR punto final para *ACTB* con muestras de RNA de D28 (NDA1 y NDA2) tratado con DNasa I.**

Se corrieron dos repeticiones de cada muestra de RNA y sus controles negativos, así como controles positivos usando el cDNA obtenido de la retrotranscripción. Solo en las muestras de cDNA se observa amplicón de *ACTB*. MPM 100 pb: Marcador de peso molecular de 100 pb, Rep1: primera repetición, Rep2: segunda repetición, Ctrl-: Control negativo (sin muestra).

Estos resultados indican que fue posible detectar la presencia de los transcritos seleccionados al confirmarse que los amplicones de cada gen tienen el tamaño esperado y que los cambios en los niveles de expresión de los lncRNAs en el intervalo 0-28 D con RT-qPCR corresponden a los cambios observados en el perfil transcripcional.

### **Inferencia de la función de lncRNAs por análisis bioinformáticos**

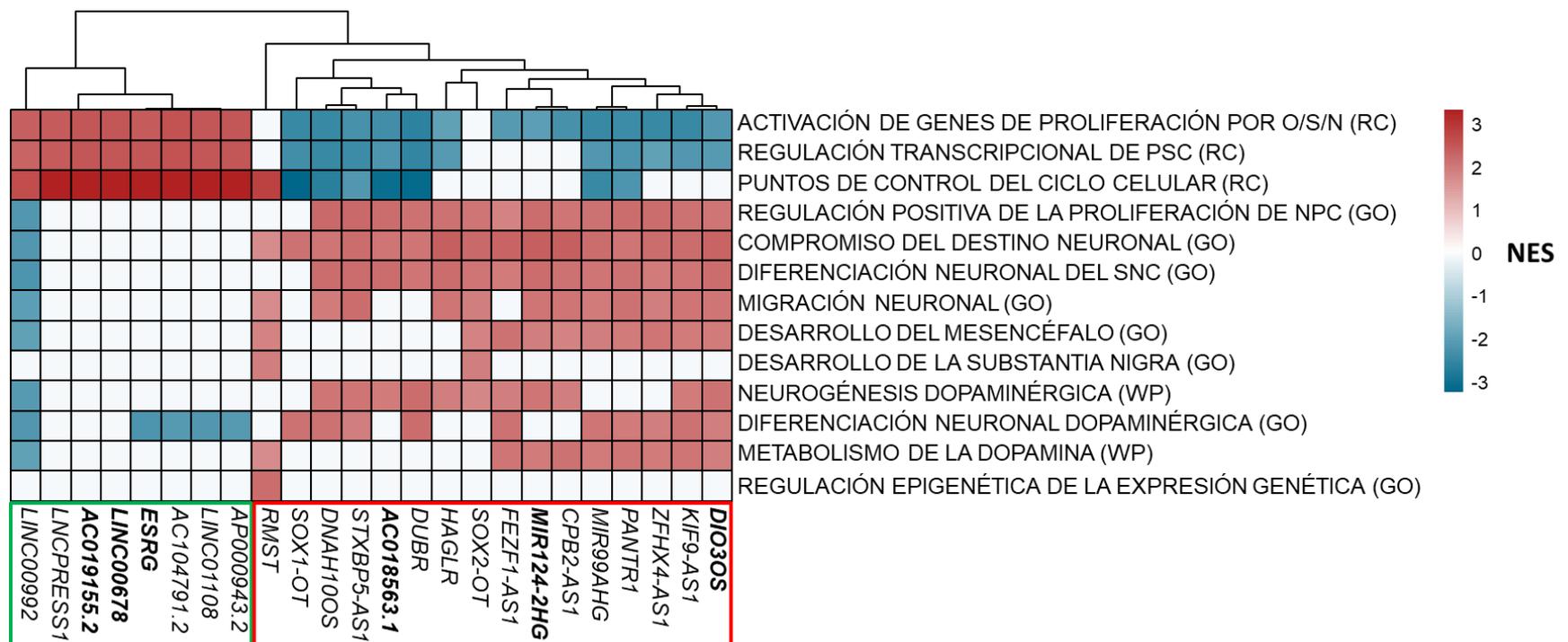
Ya que la función de la gran mayoría de los lncRNAs es desconocida, una forma para predecirla es asociando a estos lncRNAs con una colección de genes involucrados en un proceso biológico, con base en su correlación de sus niveles de expresión. A este método se le conoce como “culpable por asociación”. Para este fin se realizó un análisis de sobrerrepresentación de conjuntos de genes o GSEA el cual, arroja una calificación de

enriquecimiento normalizada (NES) para cada par lncRNA-GS. Así, un NES positivo indica que un GS está sobrerrepresentado positivamente para un lncRNA dado mientras que si tiene un NES negativo no favorece el proceso asociado al GS. En caso de tener un NES  $\approx 0$  indica que el lncRNA no se asocia con este GS.

Se realizó una búsqueda de lncRNAs con alto NES en GS importantes en la neurogénesis y diferenciación dopaminérgica, así como en el desarrollo del mesencéfalo con el objetivo de identificar lncRNAs que posiblemente participen en estos procesos. Se identificó a *HAGLR*, *RMST*, *DUBR*, *KIF9-AS1*, *PANTR1*, *DNAH10OS*, *FEZF1-AS1*, *ZFH4-AS1*, *STXBP5-AS1*, *CPB2-AS1*, *SOX2-OT*, *FAM66B* y *LINC00992*. Además, se eligieron los 13 GS más relevantes en este estudio por su participación en el control del ciclo celular (genes necesarios en el control de la división celular y la transición de las fases del ciclo celular), la regulación transcripcional de PSCs (genes importantes en el mantenimiento de la troncalidad de PSCs), el desarrollo del SNC y la diferenciación neuronal dopaminérgica.

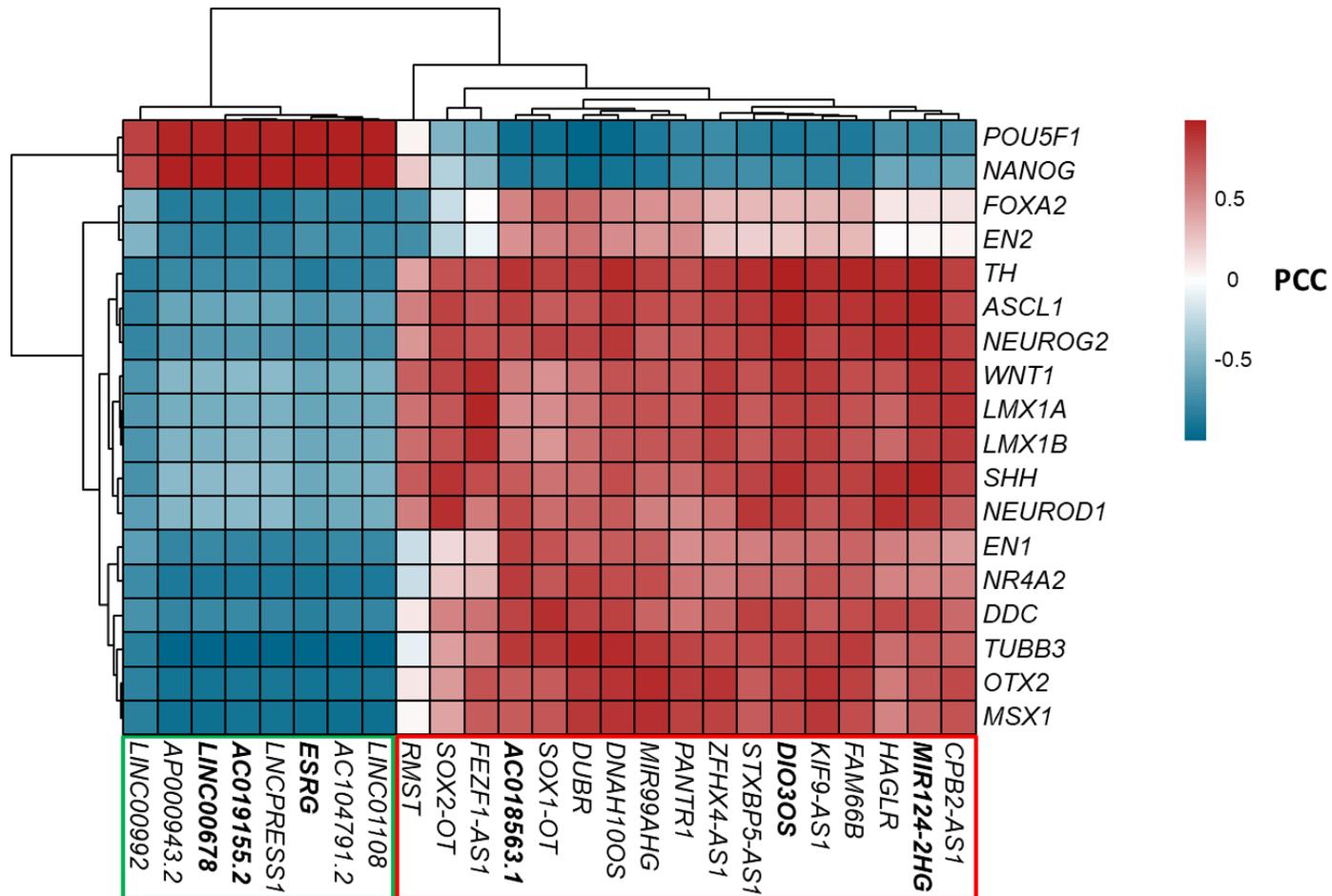
Los lncRNAs DE a la baja en el proceso de diferenciación neuronal dopaminérgica estuvieron enriquecidos en el GS asociado a el control del ciclo celular, la regulación transcripcional de células troncales y la activación de genes implicados en la proliferación por los factores de pluripotencia (*OCT4*, *SOX2* y *NANOG*), mientras que, los lncRNAs DE al alza en este proceso tuvieron un enriquecimiento positivo en GS involucrados en la regulación positiva de la proliferación de NPCs, el compromiso del destino neuronal, diferenciación neuronal y algunos de ellos en la neurogénesis y diferenciación dopaminérgica (Figura 16).

Los patrones de coexpresión de genes involucrados en la pluripotencia y en la neurogénesis dopaminérgica se analizaron para observar con mayor detalle el comportamiento y la asociación con los lncRNAs. Los lncRNAs sobreexpresados en el DO de diferenciación tienen alta correlación positiva con genes de pluripotencia como *POU5F1* (gen que codifica para *OCT4*) Y *NANOG* mientras que los de los lncRNAs DE al alza están correlacionados positivamente con genes importantes en la neurogénesis dopaminérgica. El lncRNA *DIO3OS*, tuvo la correlación positiva más alta con el gen *TH* (PCC = 0.992) el cual, codifica para la enzima que realiza el paso limitante en la síntesis de la dopamina (Figura 17).



**Figura 16. Participación de lncRNAs en procesos biológicos.**

Matriz de asociación entre lncRNAs candidatos (columnas) y GS (filas). Los lncRNAs DE al alza en el intervalo 0-28 D se encuentran dentro del recuadro rojo y en el recuadro verde se señalan los lncRNAs DE a la baja. Los lncRNAs en negritas fueron validados por RT-qPCR. Los GS fueron tomados de “The Molecular Signatures Database” (MSigDB). RC: REACTOME, GO: Gene Ontology, WP: WikiPathways, O/S/N: OCT4/SOX2/NANOG.



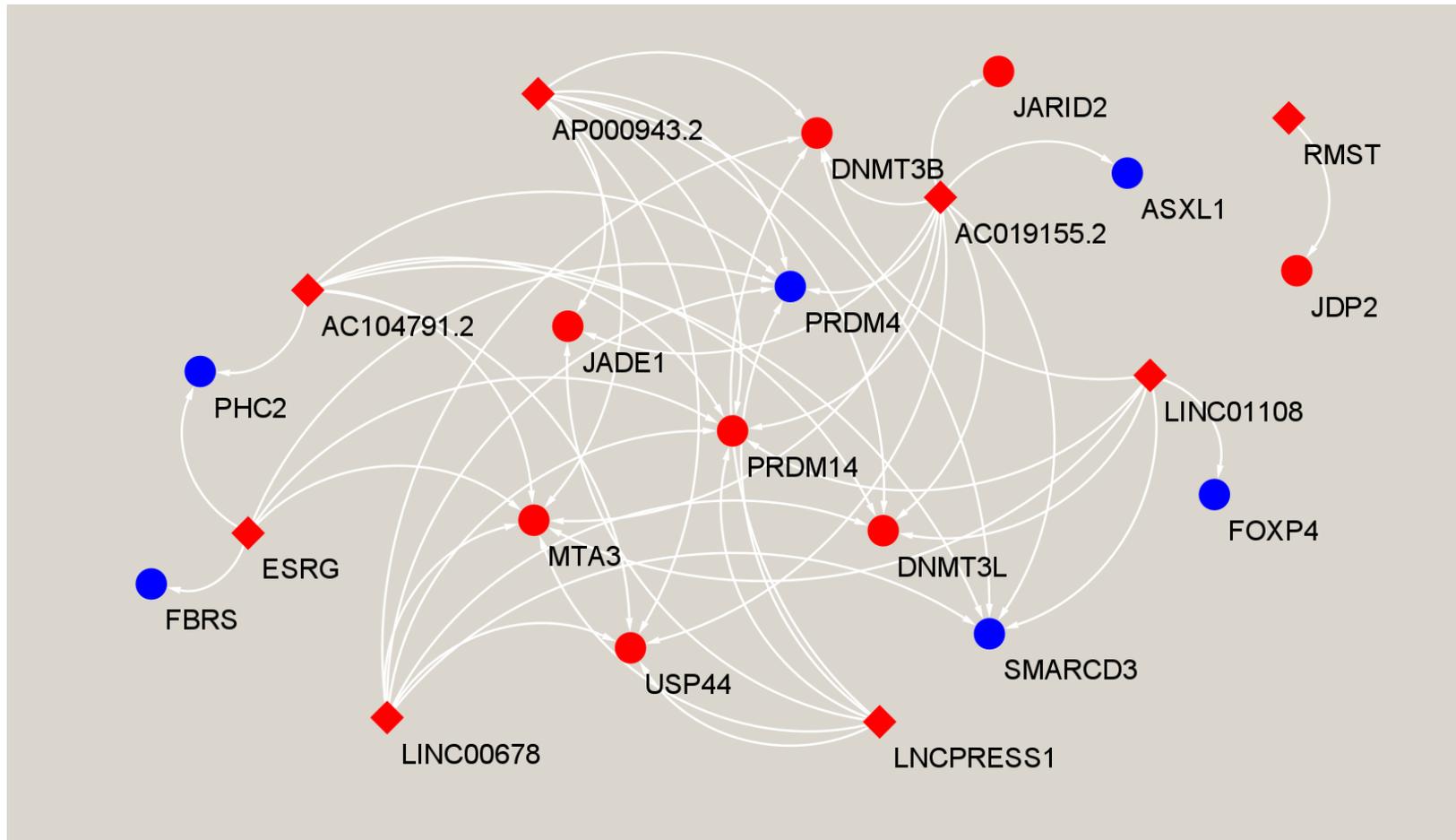
**Figura 17. Correlación de expresión entre lncRNAs candidatos y marcadores de pluripotencia y dopaminérgicos.** Mapa de calor de los lncRNAs seleccionados y marcadores de pluripotencia (*POU5F1* y *NANOG*) y de neurogénesis dopaminérgica en el intervalo 0-28 D. Los lncRNAs DE al alza se encuentran dentro del recuadro rojo y en el recuadro verde se señalan los lncRNAs DE a la baja. Los lncRNAs en negritas fueron validados por RT-qPCR.

## EFs identificados DE en la diferenciación dopaminérgica

Con base en los datos del RNA-seq, se buscaron los genes DE asociados a la regulación epigenética de la expresión genética en los tres intervalos de la diferenciación dopaminérgica (0-14, 14-28 y 0-28 D) utilizando EpiFactors, una base de datos que contiene información de 815 proteínas y 69 complejos que participan en este tipo de regulación (remodeladores de la cromatina, remodeladores de la cromatina dependientes de ATP, variantes y chaperonas de Histonas y proteínas modificadoras de DNA) o cofactores de proteínas que forman complejos con EFs, que son importantes para su actividad (Medvedeva et al., 2015).

Se encontraron 58 EFs en el intervalo 0-14 D, 19 EFs en el intervalo 14-28 D y 65 EFs en el intervalo 0-28 D con un total de 88 EFs distintos entre los tres intervalos.

Con la finalidad de identificar lncRNAs asociados a la regulación epigenética de la expresión genética, se realizó una red de coexpresión de los 25 lncRNAs previamente evaluados por GSEA y los 88 EFs DE en los tres intervalos usando un  $PCC \geq |0.98|$  entre los pares. Solamente los lncRNAs sobreexpresados en el D0 en comparación con el D28 mostraron una alta correlación con EFs. Entre los EFs que están altamente correlacionados positivamente con los lncRNAs DE a la baja en el intervalo 0-28 D se identificaron siete: *PRDM14*, *MTA3*, *DNMT3B*, *USP44*, *DNMT3L*, *JADE* y *JARID2*. Por otra parte, los EFs altamente correlacionados negativamente con estos lncRNAs DE a la baja en el intervalo 0-28 D también fueron siete: *JDP2*, *PRDM4*, *SMARCD3*, *PHC2*, *FBRS*, *ASXL1* y *FOXP4* (Figura 18).



**Figura 18. Los lncRNAs DE a la baja tienen alto nivel de coexpresión con EFs.**  
 Red de coexpresión entre lncRNAs (rombos) y EF (círculos) DE donde se muestran los genes sobreexpresados en rojo y los subexpresados en azul en el D0 (0-28 D). Las conexiones entre genes indican alta correlación ( $PCC \geq |0.98|$ ).

## Regulación *in cis* de lncRNAs

Los lncRNAs pueden regular la expresión de PCGs cercanos a su *locus (in cis)* (revisado en Fatica & Bozzini, 2014). Con el fin de identificar lncRNAs que pueden regular la expresión genética de esta forma se analizó el *locus* de los 25 lncRNAs evaluados anteriormente por GSEA donde se buscaron PCGs altamente correlacionados ( $PCC > 0.8$ ), a la distancia genómica más cercana entre genes ( $\leq 5$  Kb) y solamente se seleccionaron los PCGs DE en el intervalo 0-28 D con el fin de identificar posibles genes regulados por su lncRNA asociado. Para esto se utilizó el programa IGV el cual permite la visualización de los datos previamente obtenidos por el RNA-seq (Figura 19).

De los 25 lncRNAs solamente seis de ellos tienen un gen asociado con los parámetros establecidos y estos son genes DE al alza en el intervalo 0-28 D. Estos pares son:

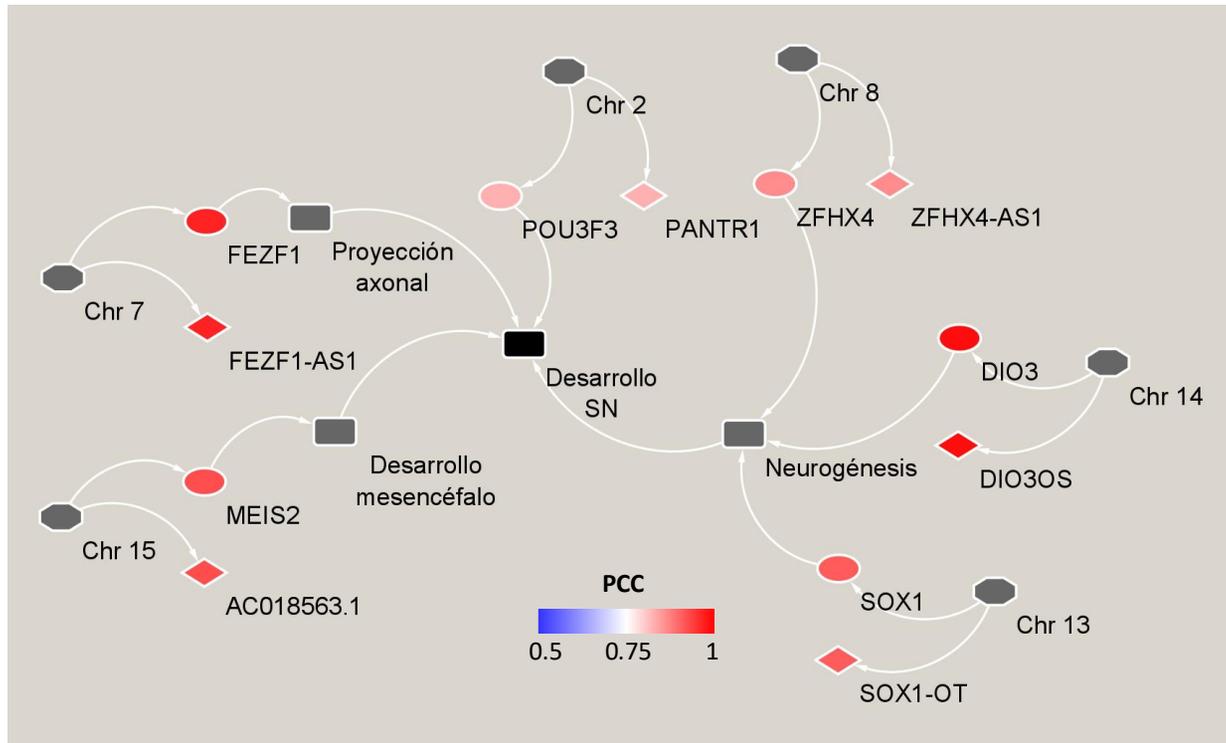
*DIO3-DIO3OS*, *AC018563.1-MEIS2*, *SOX1-SOX1-OT*, *POU3F3-PANTR1*, *FEZF1-FEZF1-AS1* y *ZFHX4-ZFHX4-AS1*. Al analizar el papel funcional de los PCGs asociados, se observó que todos son importantes en el desarrollo del SN (Figura 20).

lncRNA	PCG	Tipo	locus	Dist. entre genes (Kb)	Visualizador de datos genómicos	Nivel de expresión Log <sub>2</sub> FC(FPKM)												
<i>DIO3OS</i>	<i>DIO3</i>	LB	chr14:101,554,707-101,565,663	1		<table border="1"> <caption>Log<sub>2</sub>FC(FPKM) for DIO3OS and DIO3</caption> <thead> <tr> <th>Días de diferenciación</th> <th>DIO3OS (Log<sub>2</sub>FC)</th> <th>DIO3 (Log<sub>2</sub>FC)</th> </tr> </thead> <tbody> <tr> <td>D0</td> <td>-2.5</td> <td>-1.5</td> </tr> <tr> <td>D14</td> <td>1.0</td> <td>3.0</td> </tr> <tr> <td>D28</td> <td>4.5</td> <td>8.0</td> </tr> </tbody> </table>	Días de diferenciación	DIO3OS (Log <sub>2</sub> FC)	DIO3 (Log <sub>2</sub> FC)	D0	-2.5	-1.5	D14	1.0	3.0	D28	4.5	8.0
Días de diferenciación	DIO3OS (Log <sub>2</sub> FC)	DIO3 (Log <sub>2</sub> FC)																
D0	-2.5	-1.5																
D14	1.0	3.0																
D28	4.5	8.0																
<i>AC018563.1</i>	<i>MEIS2</i>	PT	chr15:36,867,227-36,908,987	2.2		<table border="1"> <caption>Log<sub>2</sub>FC(FPKM) for AC018563.1 and MEIS2</caption> <thead> <tr> <th>Días de diferenciación</th> <th>AC018563.1 (Log<sub>2</sub>FC)</th> <th>MEIS2 (Log<sub>2</sub>FC)</th> </tr> </thead> <tbody> <tr> <td>D0</td> <td>-2.5</td> <td>1.5</td> </tr> <tr> <td>D14</td> <td>3.5</td> <td>5.5</td> </tr> <tr> <td>D28</td> <td>6.5</td> <td>7.5</td> </tr> </tbody> </table>	Días de diferenciación	AC018563.1 (Log <sub>2</sub> FC)	MEIS2 (Log <sub>2</sub> FC)	D0	-2.5	1.5	D14	3.5	5.5	D28	6.5	7.5
Días de diferenciación	AC018563.1 (Log <sub>2</sub> FC)	MEIS2 (Log <sub>2</sub> FC)																
D0	-2.5	1.5																
D14	3.5	5.5																
D28	6.5	7.5																
<i>SOX1-OT</i>	<i>SOX1</i>	LI	chr13:111,969,070-112,130,532	0		<table border="1"> <caption>Log<sub>2</sub>FC(FPKM) for SOX1-OT and SOX1</caption> <thead> <tr> <th>Días de diferenciación</th> <th>SOX1-OT (Log<sub>2</sub>FC)</th> <th>SOX1 (Log<sub>2</sub>FC)</th> </tr> </thead> <tbody> <tr> <td>D0</td> <td>-2.5</td> <td>-1.5</td> </tr> <tr> <td>D14</td> <td>2.0</td> <td>3.5</td> </tr> <tr> <td>D28</td> <td>3.5</td> <td>5.0</td> </tr> </tbody> </table>	Días de diferenciación	SOX1-OT (Log <sub>2</sub> FC)	SOX1 (Log <sub>2</sub> FC)	D0	-2.5	-1.5	D14	2.0	3.5	D28	3.5	5.0
Días de diferenciación	SOX1-OT (Log <sub>2</sub> FC)	SOX1 (Log <sub>2</sub> FC)																
D0	-2.5	-1.5																
D14	2.0	3.5																
D28	3.5	5.0																

lncRNA	PCG	Tipo	locus	Dist. entre genes (Kb)	Visualizador de datos genómicos	Nivel de expresión $\text{Log}_2\text{FC}(\text{FPKM})$
PANTR1	POU3F3	PT	chr2:10 4,809,023- 104,858,606	1		
FEZF1-AS1	FEZF1	AS	chr7:12 2,297,895- 122,314,205	0		
ZFHx4-AS1	ZFHx4	AS	chr8:76, 679,219 - 76,869, 285	0		

**Figura 19. Los lncRNAs DE al alza tienen alto nivel de coexpresión con PCGs cercanos.**

Se muestra el *locus* de cada pareja lncRNA-PCG, distancia más corta entre ellos, tipo y *locus*, así como los datos de expresión en cada tiempo. Las barras de las gráficas indican el log<sub>2</sub>(FPKM) para PCGs (rojo) y lncRNAs (azul), mientras que las barras de error representan ± Desviación estándar (SD). LB: lncRNA bidireccional; LI: lincRNA; AS: lncRNA antisentido; PT: processed transcript.



**Figura 20. Participación *in cis* de lncRNAs en el desarrollo del SN.**

Red de coexpresión entre lncRNAs (rombos) y PCGs (elipses) cercanos donde se ilustran los cromosomas (octágonos) donde se encuentra el par lncRNA-PCG, en cuadros grises se indica el proceso biológico en el que está involucrado el PCG asociado, y el cuadro negro señala el proceso biológico general. El color de cada par lncRNA-PCG señala su PCC.

## DISCUSIÓN

### 1. Validación de lncRNAs por RT-qPCR

Los cambios en los niveles de expresión de *LINC00678*, *AC019155.2*, *DIO3OS* y *AC018563.1* en el intervalo 0-28 D reportados en el perfil transcripcional (Meléndez-Ramírez et al., 2021) y los obtenidos en este trabajo con RT-qPCR tienen una alta correlación ( $R^2 > 0.96$ ) por lo que fue posible validar estos lncRNAs.

### 2. Inferencia de la función de lncRNAs en el proceso de diferenciación dopaminérgica por GSEA

Los lncRNAs sobreexpresados en hESCs se encontraron asociados positivamente a la regulación transcripcional de PSC, el control del ciclo celular y la activación de genes de proliferación por OCT4/SOX2/NANOG, los cuales son procesos biológicos importantes en el mantenimiento de PSCs (Thomson, 1998). Entre los lncRNAs que estuvieron enriquecidos positivamente en estos GS y reportados anteriormente se encuentran; *ESRG*; gen altamente expresado en PSCs y regulado por OCT4 (Zhao et al., 2007), *LINC01108*; lncRNA asociado a la pluripotencia de PSCs y regulado por NANOG (Ng et al., 2011) y *LNCPRESS1*; transcrito que favorece la pluripotencia de PSCs (Jain et al., 2016).

Por otro lado, los lncRNAs sobreexpresados en el D28 estuvieron correlacionados positivamente con GS asociados al desarrollo del SNC. Entre los lncRNAs que estuvieron asociados al GS de diferenciación neuronal del SNC se encuentran *MIR124-2HG* (precursor de *miR-124*) y *HAGLR*, previamente reportados como transcritos involucrados en este proceso (Makeyev et al., 2007, Wei et al., 2021) mientras que *RMST* estuvo asociado al desarrollo de la *SNpc*, zona del mesencéfalo donde se localizan los somas de las nDAs que se pierden en la EP y donde se ha reportado que se expresa abundantemente este transcrito (Ng et al., 2011). Además, *RMST* fue el único lncRNA asociado a la regulación epigenética de la expresión genética, acorde con lo reportado en el 2019, donde se observó que este transcrito incrementa la expresión de *DNMT3B* (gen que codifica para la proteína con el mismo nombre y que regula la expresión genética a nivel

epigenético al metilar el DNA *de novo*) al unirse con HuR, estabilizador de mRNAs (Peng et al., 2019).

La alta coincidencia entre los procesos biológicos enriquecidos del GSEA con lo reportado previamente permite inferir que posiblemente los lncRNAs que aún no están caracterizados participen en los procesos celulares en los cuales presentaron una alta correlación.

### 3. Red de coexpresión entre lncRNAs y EFs

Debido a la enorme relevancia de la regulación epigenética de la expresión de genes relacionado con el mantenimiento del estado pluripotencial de PSCs, como en el proceso de diferenciación celular (Godini et al., 2018), se realizó una red de coexpresión entre lncRNAs seleccionados y EFs DE. Notablemente, solo los lncRNAs sobreexpresados en el D0 en comparación con el D28 tuvieron una alta correlación con EFs ( $PCC = |0.98|$ ). Los EFs que tuvieron una correlación positiva están involucrados en el mantenimiento de ESCs. JARID2, DNMT3L y DNMT3B interactúan con PRC2, el cual es importante para el balance entre la diferenciación de células troncales y la proliferación al colocar la marca represiva de Histonas H3K27me3 lo que conlleva a la compactación de la cromatina y la regulación de la expresión de genes importantes en el destino celular en el desarrollo (Laugesen et al., 2016). JARID2 promueve el reclutamiento de PRC2 a la secuencia promotora de los genes blanco y además estabiliza la marca represiva H3K27me3 (Fiorenzano et al., 2019). Las DNMTs son enzimas que metilan el DNA, entre ellas se encuentra la DNMT1, se encarga de metilar la hebra nueva del DNA cuando se replica mientras que DNMT3A y DNMT3B son catalíticamente activas en comparación con DNMT3L (cofactor de DNMT3s). Estas se encuentran presentes en altos niveles en etapas tempranas del desarrollo para regular la expresión de genes específicos (Gao et al., 2020). En particular, *DNMT3B* se encuentra sobreexpresado en el estado pluripotencial de PSCs y a pesar de que su expresión es ubicua, en el cerebro se encuentra en niveles muy bajos (Bongso et al., 2004). De hecho, el “knockdown” de *DNMT3B* está involucrado en el establecimiento del patrón mesencefálico temprano al aumentar la expresión de *FOXA2*, *EN1* y *TH* (Martins-Taylor et al., 2012). Es posible que la disminución de la

expresión de *DNMT3B* en el D28 pueda favorecer la expresión de genes importantes en la neurogénesis dopaminérgica.

Por otro lado, los EFs altamente correlacionados negativamente con los lncRNAs participan en la diferenciación celular (*JDP2*, *PRDM4*), desarrollo embrionario (*PHC2*) y diferenciación neural/neuronal (*SMARCD3*) (The UniProt consortium, 2021) <https://www.uniprot.org/>.

Estos resultados, aunque coinciden con estudios anteriores donde se identifican lncRNAs que regulan epigenéticamente las PSCs al asociarse con EFs (Chen et al., 2020), no permiten dilucidar si la expresión diferencial de los EFs es causa o consecuencia de los lncRNAs, ni tampoco es posible asociarlos explícitamente a procesos epigenéticos debido a la naturaleza del análisis. Sin embargo, ya que los niveles de expresión entre lncRNA-EF están altamente correlacionados, es posible que participen en un mismo proceso biológico o vía regulatoria, aunque son necesarios estudios complementarios que permitan corroborarlo.

#### 4. Regulación *in cis* de lncRNAs

Los lncRNAs se expresan de forma dinámica y específica dependiendo de su contexto celular. La evidencia sugiere que lncRNAs pueden regular la expresión de genes que se encuentran cercanos a su *locus* y que, en muchos casos estos PCGs son importantes en el neurodesarrollo (Mercer et al., 2010). Es por ello que se generó una red de coexpresión *in cis* donde solamente se consideraron pares lncRNA-PCG con un alto nivel de correlación ( $PCC > 0.8$ ) ya que según la bibliografía, esta regulación tiende a ser extremadamente positiva en comparación con la regulación *in trans* (Derrien et al., 2012). Estos pares lncRNA-PCG mostraron una correlación muy alta y positiva:

*DIO3-DIO3OS* ( $PCC = 0.98$ ), *FEZF1-FEZF1-AS1* ( $PCC = 0.96$ ), *AC018563.1-MEIS2* ( $PCC = 0.92$ ), *SOX1-SOX1-OT* ( $PCC = 0.91$ ), *ZFHX4-ZFHX4-AS1* ( $PCC = 0.86$ ) y *POU3F3-PANTR1* ( $PCC = 0.82$ ).

*DIO3OS* fue el lncRNA que tuvo los cambios más dinámicos al alza en el proceso de diferenciación dopaminérgica y, sorprendentemente sus niveles de coexpresión con su PCG asociado, *DIO3* y con *TH* fueron muy altos ( $PCC > 0.98$ ), por este motivo se procedió a investigar si existía un vínculo

entre *DIO3* y *TH*. Por su parte, *DIO3* codifica para la yodotironina desyodasa tipo 3 (*DIO3*), enzima encargada de inactivar a la hormona tiroidea T3 regulando así esta vía de señalización importante en muchos procesos que van desde el metabolismo hasta la neurogénesis y el desarrollo neural a través de ventanas de tiempo específicas al unirse a su receptor nuclear y regular la expresión genética (Liu & Brent, 2018). Sin embargo, ya que esta vía de señalización tiene efectos muy variados dependiendo del contexto celular, no fue posible hacer ninguna inferencia mediante la información transcripcional. No obstante, estudios se ha reportado que al exponer precursores neurales multipotenciales a T3, se incrementa significativamente la diferenciación a oligodendrocitos (Ahlgren et al., 1997) o al exponer células C17.2 (línea transgénica neuronal de ratón) con T3, la expresión de genes involucrados en la neurogénesis y el desarrollo de proyecciones neuronales disminuía drásticamente, entre ellos *Sox2* y *Nr4a2*. NR4A2 es un TF importante en la diferenciación y mantenimiento de neuronas dopaminérgicas del mesencéfalo durante el desarrollo al regular la expresión de *Slc6a3* (DAT), *Slc18a2* (VMAT), *Drd2* y *Th* (Chatonnet et al., 2013). Esta evidencia sugiere una posible regulación de *DIO3OS* en la expresión de *TH* al regular los niveles de expresión de *DIO3*. A pesar de que la correlación de expresión entre *DIO3OS* y *NR4A2* es positiva, no es muy alta (PCC = 0.67), lo que sugiere que el aumento de *TH* que se observa puede ser efecto de otros mecanismos de regulación sobre este gen. Además, el par *DIO3OS-DIO3* se encuentran en un *locus* improntado que se ha reportado genera ncRNAs y proteínas asociadas con el neurodesarrollo y a desórdenes neurodegenerativos (Mo et al., 2015) por lo que *DIO3OS* puede estar involucrado en la diferenciación neuronal dopaminérgica.

El lncRNA *FEZF1-AS1* se ha asociado a la proliferación celular en distintos tumores incluido el glioblastoma, considerado el tumor cerebral más agresivo (Luo et al., 2020), mientras que su proteína asociada, FEZF1 está involucrada en la proyección axonal y en el patrón rostrocaudal del diencéfalo (The UniProt consortium, 2021). En el caso de *MEIS2*, proteína altamente correlacionada con el lncRNA *AC018563.1* se ha reportado que participa en el desarrollo del mesencéfalo (The UniProt consortium, 2021).

El par *SOX1-OT-SOX1*, tiene un papel en el proceso de diferenciación neural (Ahmad et al., 2017). Con respecto al par *ZFHX4-AS1-ZFHX4* existe poca bibliografía, solo la proteína está involucrada en la diferenciación neural y muscular (The UniProt consortium, 2021).

Finalmente, *PANTR1* es un lncRNA cuyo ortólogo en ratones, *lincBRN1a*, regula la transición entre la troncalidad de NSCs y la diferenciación (Carelli et al., 2019), mientras su gen asociado, *POUF3*, codifica para el TF con su mismo nombre el cual está involucrado en el desarrollo del SNC (Snijders Blok et al., 2019). Si bien estos resultados son interesantes debido a que se relacionan con el desarrollo del SN central, solamente *DIO3* mostró una asociación al fenotipo dopaminérgico.

En el caso de los lncRNAs candidatos que no mostraron asociación con EFs o genes *in cis*, no se descarta que puedan regular *in trans*, sin embargo, son necesarios estudios específicos de cada uno para su validación funcional.

Entre las limitantes a considerar en este proyecto se encuentran las siguientes:

- Solamente se secuenciaron transcritos con cola de poliA por RNA-seq. Por ende, no fue posible detectar e identificar posibles lncRNAs con actividad reguladora que no tengan esta característica (Zhang et al., 2014)
- Solo fueron evaluados tres tiempos de la diferenciación dopaminérgica (D0, D14 y D28), por lo que existen tiempos intermedios donde no se cuenta con información ni tampoco de tiempos posteriores al D28 en los cuales se expresan genes con funciones específicas del tipo celular (La Manno et al., 2016)
- Aunque los análisis bioinformáticos son herramientas que permiten inferir posibles resultados, estos dependen de la información proporcionada, además, no consideran la totalidad de las variables presentes en un sistema biológico por ello es imprescindible elaborar experimentación *in vivo* o *in vitro* para corroborar esta información.

## CONCLUSIÓN

De acuerdo con el análisis bioinformático de los lncRNAs DE en el proceso de diferenciación neuronal dopaminérgica a partir de hESCs, los lncRNAs DE a la baja que se validaron (*LINC00678* y *AC019155.2*) en el intervalo 0-28 D están asociados al control del ciclo celular, la regulación transcripcional de PSCs y la proliferación celular. Además, su expresión estuvo altamente correlacionada positivamente con EFs que regulan el mantenimiento de la pluripotencia, mientras que con EFs que modulan la diferenciación celular, estuvieron altamente correlacionadas pero negativamente.

Por otro lado, *DIO3OS* y *AC018563.1*, los lncRNAs DE al alza en el intervalo 0-28 D que se validaron en este trabajo posiblemente favorecen procesos de neurogénesis y/o diferenciación neuronal dopaminérgica.

Con base en lo anterior, *LINC00678*, *AC019155.2*, *DIO3OS* y *AC018563.1* son candidatos para su caracterización funcional.

El estudio de estas biomoléculas permitirá un mejor entendimiento de los mecanismos implicados en la toma de decisiones entre el mantenimiento de la troncalidad de PSCs o la transición hacia el compromiso del destino celular.

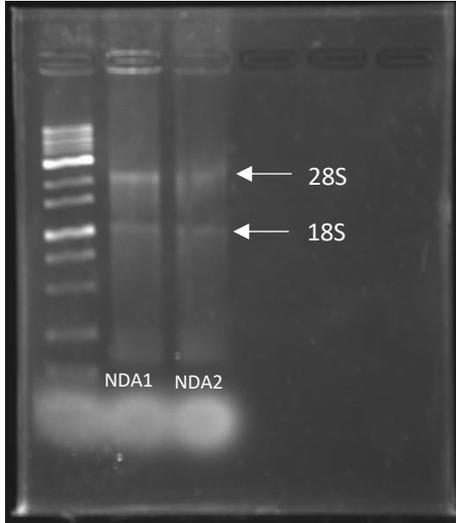
## PERSPECTIVAS

Identificar TFs que pueden regular la expresión de los lncRNAs de interés.

Caracterizar funcionalmente los lncRNAs validados por qPCR (*LINC00678*, *AC019155.2*, *DIO3OS* y *AC018563.1*) mediante análisis de expresión en compartimentos celulares, estudios de ganancia/pérdida de función y ensayos que permitan demostrar su interacción con biomoléculas como ensayos de “RNA pull-down” o técnicas de inmunoprecipitación.

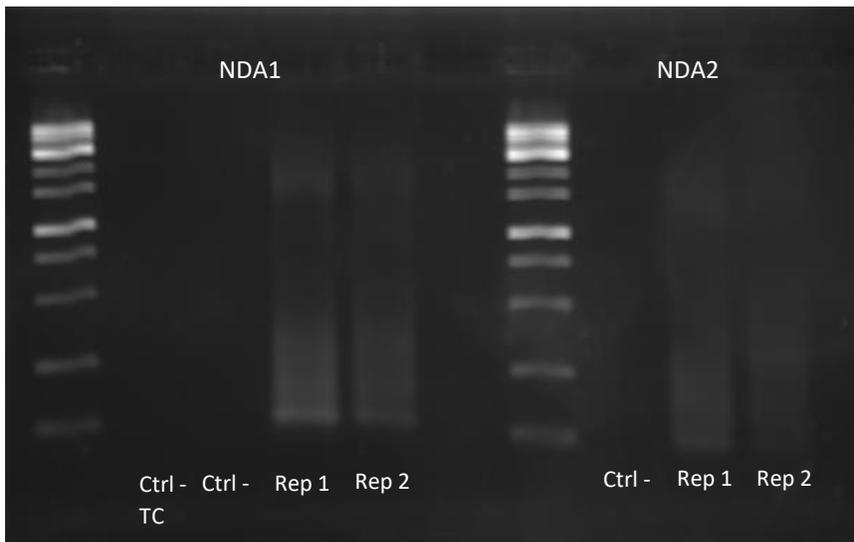
## ANEXOS

### Gel de electroforesis de muestras de RNA del D28 (NDA1 y NDA2) tratadas con DNasa I



Se señalan las bandas correspondientes a las subunidades ribosómicas 28S y 18S, indicativo de muestras de RNA en óptimo estado.

### Gel de electroforesis de muestras de cDNA del D28



Se realizó una RT de dos muestras biológicas del D28 (NDA1 y NDA2). Ctrl- TC: Control negativo sin termociclador, Ctrl -: Control negativo, Rep 1: Repetición 1 con carga de cDNA, Rep 2: Repetición 2 con carga de cDNA.

## BIBLIOGRAFÍA

- Ahlgren, S. C., Wallace, H., Bishop, J., Neophytou, C., & Raff, M. C. (1997). Effects of thyroid hormone on embryonic oligodendrocyte precursor cell development *in vivo* and *in vitro*. *Molecular and Cellular Neurosciences*, 9(5–6), 420–432. <https://doi.org/10.1006/mcne.1997.0631>
- Ahmad, A., Strohbuecker, S., Tufarelli, C., & Sottile, V. (2017). Expression of a *SOX1* overlapping transcript in neural differentiation and cancer models. *Cellular and Molecular Life Sciences*, 74(22), 4245–4258. <https://doi.org/10.1007/s00018-017-2580-3>
- Allis, C. D., & Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nature Reviews Genetics*, 17(8), 487–500. <https://doi.org/10.1038/nrg.2016.59>
- Aron, Liviu (2010). 'Genetic analysis of dopaminergic neuron survival: GDNF/Ret signaling and the Parkinson's disease-associated gene DJ-1'. Dissertation, LMU München: Faculty of Biology. <https://edoc.ub.uni-muenchen.de/11787/>
- Banta, J. A., & Richards, C. L. (2018). Quantitative epigenetics and evolution. *Heredity*, 121(3), 210–224. <https://doi.org/10.1038/s41437-018-0114-x>
- Batagov, A. O., Yarmishyn, A. A., Jenjaroenpun, P., Tan, J. Z., Nishida, Y., & Kurochkin, I. V. (2013). Role of genomic architecture in the expression dynamics of long noncoding RNAs during differentiation of human neuroblastoma cells. *BMC Systems Biology*, 7(Suppl 3), 1–18. <https://doi.org/DOI: 10.1186/1752-0509-7-S3-S11>
- Beaulieu, J. M. (2011). The physiology, signaling, and pharmacology of dopamine receptors. *Pharmacological Reviews*, 63(1), 184–187. <https://doi.org/10.1124/pr.110.002642.182>
- Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. L., & Lander, E. S. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem

- Cells. *Cell*, 125(2), 315–326. <https://doi.org/10.1016/j.cell.2006.02.041>
- Best, J. A., Nijhout, H. F., & Reed, M. C. (2009). Homeostatic mechanisms in dopamine synthesis and release: a mathematical model. *Theoretical Biology and Medical Modelling*, 6(21), 1–20. <https://doi.org/10.1186/1742-4682-6-21>
- Bongso, A. (2004). The transcriptome profile of human embryonic stem cells as defined by SAGE. *Stem Cells*, 22(1), 51–64. <http://onlinelibrary.wiley.com/doi/10.1634/stemcells.22-1-51/full>
- Carelli, S., Giallongo, T., Rey, F., Latorre, E., Bordoni, M., Mazzucchelli, S., Gorio, M. C., Pansarasa, O., Provenzani, A., Cereda, C., & Di Giulio, A. M. (2019). HuR interacts with *lincBRN1a* and *lincBRN1b* during neuronal stem cells differentiation. *RNA Biology*, 16(10), 1471–1485. <https://doi.org/10.1080/15476286.2019.1637698>
- Chambers, S. M., Fasano, C. A., Papapetrou, E. P., Tomishima, M., Sadelain, M., & Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nature Biotechnology*, 27(3), 275–280. <https://doi.org/10.1038/nbt.1529>
- Chatonnet, F., Guyot, R., Benoît, G., & Flamant, F. (2013). Genome-wide analysis of thyroid hormone receptors shared and specific functions in neural cells. *Proceedings of the National Academy of Sciences of the United States of America*, 110(8), E766–75 . <https://doi.org/10.1073/pnas.1210626110>
- Chen, J., Wang, Y., Wang, C., Hu, J. F., & Li, W. (2020). lncRNA functions as a new emerging epigenetic factor in determining the fate of stem cells. *Frontiers in Genetics*, 11(March), 1–12. <https://doi.org/10.3389/fgene.2020.00277>
- Consortium, T. U. (2021). UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Research*, 49(D1), D480–D489. <https://doi.org/10.1093/nar/gkaa1100>
- Corso-Díaz, X., Jaeger, C., Chaitankar, V., & Swaroop, A. (2018). Epigenetic control of gene regulation during development and disease: A view from the retina. *Progress in Retinal and*

*Eye Research*, 65(July), 1–27. <https://doi.org/10.1016/j.preteyeres.2018.03.002>

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D. G., Lagarde, J., Veeravalli, L., Ruan, X., Ruan, Y., Lassmann, T., Carninci, P., Brown, J. B., Lipovich, L., Gonzalez, J. M., ... Guigó, R. (2012). The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Research*, 22(9), 1775–1789. <https://doi.org/10.1101/gr.132159.111>

Duran, R. C. D., Yan, H., Zheng, Y., Huang, X., Grill, R., Kim, D. H., Cao, Q., & Wu, J. Q. (2017). The systematic analysis of coding and long non-coding RNAs in the sub-chronic and chronic stages of spinal cord injury. *Scientific Reports*, 7(July), 1–15. <https://doi.org/10.1038/srep41008>

Fan, Y., & Ng, S. (2020). Replacing what's lost: a new era of stem cell therapy for Parkinson's disease. *Translational Neurodegeneration*, 9(2), 1–10. <https://doi.org/https://doi.org/10.1186/s40035-019-0180-x>

Fatica, A. & I. B. (2014). Long non-coding RNAs: new players in cell differentiation and development. *Genetics*, 15(1), 7-21. <https://doi.org/10.1038/nrg3606>

Fiorenzano, A., Pascale, E., Patriarca, E. J., Minchiotti, G., & Fico, A. (2019). lncRNAs and PRC2: coupled partners in embryonic stem cells. *Epigenomes*, 3(3), 1–20. <https://doi.org/10.3390/epigenomes3030014>

Gale, E., & Li, M. (2008). Midbrain dopaminergic neuron fate specification: Of mice and embryonic stem cells. *Molecular Brain*, 1(8), 1–10. <https://doi.org/10.1186/1756-6606-1-8>

Gao, L., Emperle, M., Guo, Y., Grimm, S. A., Ren, W., Adam, S., Uryu, H., Zhang, Z., Chen, D., Yin, J., Dukatz, M., Anteneh, H., Jurkowska, R. Z., Lu, J., Wang, Y., Bashtrykov, P., Wade, P. A., Wang, G. G., Jeltsch, A., & Song, J. (2020). Comprehensive structure-function characterization of DNMT3B and DNMT3A reveals distinctive *de novo* DNA methylation mechanisms. *Nature Communications*, 11(3355), 1–14. <https://doi.org/10.1038/s41467-020-17109-4>

- Godini, R., Lafta, H. Y., & Fallahi, H. (2018). Epigenetic modifications in the embryonic and induced pluripotent stem cells. *Gene Expression Patterns*, 29(January), 1–9.  
<https://doi.org/10.1016/j.gep.2018.04.001>
- Hu, W., Alvarez-Dominguez, J. R., & Lodish, H. F. (2012). Regulation of mammalian cell differentiation by long non-coding RNAs. *EMBO Reports*, 13(11), 971–983.  
<https://doi.org/10.1038/embor.2012.145>
- Jain, A. K., Mccarthy, R., Li, W., Yang, L., Barton, M. C., Jain, A. K., Xi, Y., Mccarthy, R., Allton, K., Akdemir, K. C., & Patel, L. R. (2016). *LncPRESS1* is a p53-regulated lncRNA that safeguards pluripotency by disrupting SIRT6-mediated de-acetylation of histone H3K56. *Molecular Cell*, 64(5), 967–981. <https://doi.org/10.1016/j.molcel.2016.10.039>
- Jarroux, J., Morillon, A., & Pinskaya, M. (2017). History, discovery, and classification of lncRNAs. *Advances in Experimental Medicine and Biology*, 1008, 1–46. [https://doi.org/10.1007/978-981-10-5203-3\\_1](https://doi.org/10.1007/978-981-10-5203-3_1).
- Jessell, T. M., & Sanes, J. R. (2000). Development: The decade of the developing brain. *Current Opinion in Neurobiology*, 10(5), 599–611. [https://doi.org/10.1016/S0959-4388\(00\)00136-7](https://doi.org/10.1016/S0959-4388(00)00136-7)
- Kriks, S., Shim, J. W., Piao, J., Ganat, Y. M., Wakeman, D. R., Xie, Z., Carrillo-Reid, L., Auyeung, G., Antonacci, C., Buch, A., Yang, L., Beal, M. F., Surmeier, D. J., Kordower, J. H., Tabar, V., & Studer, L. (2011). Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. *Nature*, 480(7378), 547–551.  
<https://doi.org/10.1038/nature10648>
- La Manno, G., Gyllborg, D., Codeluppi, S., Nishimura, K., Salto, C., Zeisel, A., Borm, L. E., Stott, S., R. W., Toledo, E. M., Villaescusa, J. C., Lönnerberg, P., Ryge, J., Barker, R. A., Arenas, E., & Linnarsson, S. (2016). Molecular diversity of midbrain development in mouse, human, and stem cells. *Cell*, 167(2), 566–580. <https://doi.org/10.1016/j.cell.2016.09.027>
- Latg, G., Poulet, C., Bours, V., Josse, C., & Jerusalem, G. (2018). Natural antisense transcripts: Molecular mechanisms and implications in breast cancers. *International Journal of*

*Molecular Sciences*, 19(1), 123. <https://doi.org/10.3390/ijms19010123>

Laugesen, A., Westergaard Hojfeldt, J., & Helin, K. (2016). Role of the Polycomb Repressive Complex 2 (PRC2) in transcriptional regulation and cancer. *Cold Spring Harbor Perspectives in Medicine*, 2, 1–20. <https://doi.org/doi:10.1101/cshperspect.a026575>

Li, T., Chen, B., Yang, P., Wang, D., Du, B., & Kang, L. (2020). Long non-coding RNA derived from lncRNA-mRNA co-expression networks modulates the locust phase change. *Genomics, Proteomics and Bioinformatics*, 18(6), 664–678. <https://doi.org/10.1016/j.gpb.2020.05.001>

Liu, Y. Y., & Brent, G. A. (2018). Thyroid hormone and the brain: Mechanisms of action in development and role in protection and promotion of recovery after brain injury. *Pharmacology and Therapeutics*, 186, 176–185. <https://doi.org/10.1016/j.pharmthera.2018.01.007>

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods*, 25(4), 402–408. <https://doi.org/10.1006/meth.2001.1262>

Luo, L., Zhang, Y., He, H., Chen, C., Zhang, B., & Cai, M. (2020). lncRNA *FEZF1-AS1* sponges *miR-34a* to upregulate *NOTCH-1* in glioblastoma. *Cancer Management and Research*, 12, 1827–1833. <https://doi.org/10.2147/CMAR.S240531>

Makeyev, E. V., Zhang, J., Carrasco, M. A., & Maniatis, T. (2007). The microRNA *miR-124* promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Molecular Cell*, 27(3), 435–448. <https://doi.org/10.1016/j.molcel.2007.07.015>

Martins-Taylor, K., Schroeder, D. I., Lasalle, J. M., Lalande, M., & Xu, R. H. (2012). Role of DNMT3B in the regulation of early neural and neural crest specifiers. *Epigenetics*, 7(1), 71–82. <https://doi.org/10.4161/epi.7.1.18750>

Medvedeva, Y. A., Lennartsson, A., Ehsani, R., Kulakovskiy, I. V., Vorontsov, I. E., Panahandeh, P., Khimulya, G., Kasukawa, T., FANTOM Consortium, & Drabløs, F. (2015). EpiFactors: a

comprehensive database of human epigenetic factors and complexes. *Database: the Journal of Biological Databases and Curation*, 2015(July), 1-10.

<https://doi.org/10.1093/database/bav067>

Meléndez-Ramírez, C., Cuevas-Diaz Duran, R., Barrios-García, T., Giacomán-Lozano, M., López-Ornelas, A., Herrera-Gamboa, J., Estudillo, E., Soto-Reyes, E., Velasco, I., & Treviño, V. (2021). Dynamic landscape of chromatin accessibility and transcriptomic changes during differentiation of human embryonic stem cells into dopaminergic neurons. *Scientific Reports*, 11(1), 1–18. <https://doi.org/10.1038/s41598-021-96263-1>

Mercer, T. R., Qureshi, I. A., Gokhan, S., Dinger, M. E., Li, G., Mattick, J. S., & Mehler, M. F. (2010). Long noncoding RNAs in neuronal-glia fate specification and oligodendrocyte lineage maturation. *BMC Neuroscience*, 11(14), 1–15. <https://doi.org/10.1186/1471-2202-11-14>

Mo, C. F., Wu, F. C., Tai, K. Y., Chang, W. C., Chang, K. W., Kuo, H. C., Ho, H. N., Chen, H. F., & Lin, S. P. (2015). Loss of non-coding RNA expression from the *DLK1-DIO3* imprinted locus correlates with reduced neural differentiation potential in human embryonic stem cell lines. *Stem Cell Research and Therapy*, 6(1), 1–17. <https://doi.org/10.1186/scrt535>

Mohamed, J., Gaughwin, P. M., Lim, B., Robson, P., & Lipovich, L. (2010). Conserved long noncoding RNAs transcriptionally regulated by OCT4 and NANOG modulate pluripotency in mouse embryonic stem cells. *RNA*, 16(2), 324–337. <https://doi.org/10.1261/rna.1441510>

Ng, S., Johnson, R., & Stanton, L. W. (2011). Human long non-coding RNAs promote by association with chromatin modifiers and transcription factors. *The EMBO Journal*, 31(3), 522–533. <https://doi.org/10.1038/emboj.2011.459>

Park, S., Yoo, J. E., Lee, D. R., Jang, J., Cho, M. S., Kim, D. S., & Kim, D. W. (2019). Establishment of PITX3-mCherry knock-in reporter human embryonic stem cell line (WAe009-A-23). *Stem Cell Research*, 39(July), 101499. <https://doi.org/10.1016/j.scr.2019.101499>

Parsons, X. H., Parsons, J. F., & Moore, D. A. (2012). Genome-scale mapping of microRNA

- signatures in human embryonic stem cell neurogenesis. *Molecular Medicine and Therapeutics*, 1(2), 105. <https://doi.org/10.4172/2324-8769.1000105>
- Peng, W. X., Koirala, P., Zhang, W., Ni, C., Wang, Z., Yang, L., & Mo, Y. Y. (2019). LncRNA *RMST* enhances *DNMT3* expression through Interaction with HuR. *Molecular Therapy*, 28(1), 9–18. <https://doi.org/10.1016/j.ymthe.2019.09.024>
- Perry, R. B. T., & Ulitsky, I. (2016). The functions of long noncoding RNAs in development and stem cells. *Development*, 143(21), 3882–3894. <https://doi.org/10.1242/dev.140962>
- Roberts, T. C., Morris, K. V., & Weinberg, M. S. (2014). Perspectives on the mechanism of transcriptional regulation by long non-coding RNAs. *Epigenetics*, 9(1), 13–20. <https://doi.org/10.4161/epi.26700>
- Schuettengruber, B., Bourbon, H. M., Di Croce, L., & Cavalli, G. (2017). Genome regulation by polycomb and trithorax: 70 years and counting. *Cell*, 171(1), 34–57. <https://doi.org/10.1016/j.cell.2017.08.002>
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, I. T. (2003). Cytoscape: A software environment for integrated models. *Genome Research*, 13(22), 426. <https://doi.org/10.1101/gr.1239303.metabolite>
- Snijders Blok, L., Kleefstra, T., Venselaar, H., Maas, S., Kroes, H. Y., Lachmeijer, A., van Gassen, K., Firth, H. V., Tomkins, S., Bodek, S., DDD Study, Öunap, K., Wojcik, M. H., Cunniff, C., Bergstrom, K., Powis, Z., Tang, S., Shinde, D. N., Au, C., Iglesias, A. D., ... Fisher, S. E. (2019). *De novo* variants disturbing the transactivation capacity of POU3F3 cause a characteristic neurodevelopmental disorder. *American Journal of Human Genetics*, 105(2), 403–412. <https://doi.org/10.1016/j.ajhg.2019.06.007>
- Statello, L., Guo, C. J., Chen, L. L., & Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. *Nature Reviews. Molecular Cell Biology*, 22(2), 96–118. <https://doi.org/10.1038/s41580-020-00315-9>

- Steinhart, Z., & Angers, S. (2018). Wnt signaling in development and tissue homeostasis. *Development*, 145(11), 1–8. <https://doi.org/10.1242/dev.146589>
- Takahashi, K., & Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, 126(4), 663–676. <https://doi.org/10.1016/j.cell.2006.07.024>
- Tarsy, D. (2012). Treatment of Parkinson disease: A 64-year-old man with motor complications of advanced Parkinson disease. *JAMA*, 307(21), 2305–2314. <https://doi.org/10.1001/jama.2012.4829>
- Thomson, J. A. (1998). Embryonic stem cell lines derived from human blastocysts. *Science*, 282(5391), 1145–1147. <https://doi.org/10.1126/science.282.5391.1145>
- Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14(2), 178–192. <https://doi.org/10.1093/bib/bbs017>
- Vallerga, C. L., Zhang, F., Fowdar, J., Mcrae, A. F., Qi, T., Nabais, M. F., Zhang, Q., Kassam, I., Henders, A. K., Wallace, L., Montgomery, G., Chuang, Y., Horvath, S., Ritz, B., Halliday, G., Hickie, I., Kwok, J. B., Pearson, J., Pitcher, T., ... Mellick, G. D. (2020). Analysis of DNA methylation associates the the cystine–glutamate antiporter SLC7A11 with risk of Parkinson’s disease. *Nature Communications*, 11(1238), 1–10. <https://doi.org/10.1038/s41467-020-15065-7>
- Volders, P. J., Helsens, K., Wang, X., Menten, B., Martens, L., Gevaert, K., Vandesompele, J., & Mestdagh, P. (2013). LNCipedia: a database for annotated human lncRNA transcript sequences and structures. *Nucleic Acids Research*, 41(Database), D246–D251. <https://doi.org/10.1093/nar/gks915>
- Volpicelli, F., Perrone-Capano, C., Bellenchi, G. C., Colucci-D’Amato, L., & Di Porzio, U. (2020). Molecular regulation in dopaminergic neuron development. Cues to unveil molecular pathogenesis and pharmacological targets of neurodegeneration. *International Journal of*

*Molecular Sciences*, 21(11), 1–19. <https://doi.org/10.3390/ijms21113995>

Wang, Z., He, C., & Shi, J. S. (2020). Natural products for the treatment of neurodegenerative diseases. *Current Medicinal Chemistry*, 27(34), 5790–5828.

<https://doi.org/10.2174/0929867326666190527120614>

Wei, B., Xiao, G. R., Wu, C. L., & Xu, Y. Q. (2021). *HAGLR* promotes neuron differentiation through the *miR-130a-3p-MeCP2* axis. *Open Medicine*, 16(1), 1121–1131.

<https://doi.org/10.1515/med-2021-0301>

Wijeyekoon, R., & Barker, R. A. (2009). Cell replacement therapy for Parkinson's disease. *Biochimica et Biophysica Acta*, 1792(7), 688–702.

<https://doi.org/10.1016/j.bbadis.2008.10.007>

Williams, L. A., Davis-Dusenbery, B. N., & Eggan, K. C. (2012). SnapShot: directed differentiation of pluripotent stem cells. *Cell*, 149(5), 1174–1174.e1.

<https://doi.org/10.1016/j.cell.2012.05.015>

Yoo, A. S., Staahl, B. T., Chen, L., & Crabtree, G. R. (2009). MicroRNA-mediated switching of chromatin-remodelling complexes in neural development. *Nature*, 460(7255), 642–646.

<https://doi.org/10.1038/nature08139>

Zhang, L., Xue, Z., Yan, J., Wang, J., Liu, Q., & Jiang, H. (2019). LncRNA *Riken-201* and *Riken-203* modulates neural development by regulating the *Sox6* through sequestering miRNAs. *Cell Proliferation*, 52(3), 1–14. <https://doi.org/10.1111/cpr.12573>

Zhang, Y., Yang, L., & Chen, L. L. (2014). Life without A tail: New formats of long noncoding RNAs. *International Journal of Biochemistry and Cell Biology*, 54(September), 338–349.

<https://doi.org/10.1016/j.biocel.2013.10.009>

Zhao, M., Ren, C., Yang, H., Feng, X., Jiang, X., Zhu, B., Zhou, W., Wang, L., Zeng, Y., & Yao, K. (2007). Transcriptional profiling of human embryonic stem cells and embryoid bodies identifies

*HESRG*, a novel stem cell gene. *Biochemical and Biophysical Research Communications*, 362(4), 916–922. <https://doi.org/10.1016/j.bbrc.2007.08.081>