

**CENTRO DE INVESTIGACIÓN Y ESTUDIOS AVANZADOS DEL
INSTITUTO POLITÉCNICO NACIONAL**
UNIDAD ZACATENCO

DEPARTAMENTO DE GENÉTICA Y BIOLOGÍA MOLECULAR

**“Establecimiento de un sistema basado en el método
CRISPR-Cas9 para la identificación *in situ* de complejos
asociados a la cromatina de loci involucrados en la
diferenciación y regeneración muscular.”**

Tesis que presenta:

Lic. en Biotec. Rodolfo Daniel Ávila Avilés

Para obtener el Grado de

Maestro en Ciencias

en la Especialidad de

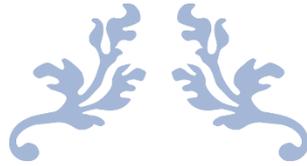
Genética y Biología Molecular

Tutor:

Dr. José Manuel Hernández Hernández

Ciudad de México

Agosto 2020



Resumen

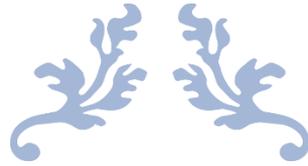


El músculo esquelético adulto se caracteriza por una alta capacidad de regeneración, la cual es dependiente de la presencia y funcionalidad de las células madre musculares, llamadas células satélite (SC). En la homeostasis muscular, las SC están en un estado de reposo; después de un estímulo inducido por daño físico o enfermedad, las SC se activan y vuelven a ingresar al ciclo celular mediado por un complejo proceso de regulación transcripcional, donde Pax7 marca inequívocamente el linaje muscular en SC y modula jerárquicamente la expresión de los factores reguladores miogénicos (MRF), Myf5 y MyoD. A medida que avanza la miogénesis, el factor MyoD induce la expresión de miogenina y MRF4 para culminar con la expresión de proteínas estructurales del músculo esquelético diferenciado.

La expresión diferencial de MRF y su actividad en el proceso de diferenciación y regeneración muscular se controla en múltiples niveles. A nivel epigenético, la combinación de modificaciones en la estructura de la cromatina ejerce un papel dinámico y decisivo en la regeneración muscular. Por ejemplo, la presencia de elementos de regulación transcripcional remota, como potenciadores y promotores, la síntesis de ARN no codificantes, modificaciones postraduccionales de proteínas reguladoras, así como la combinación de factores de transcripción y complejos de remodelación de la cromatina, regulan la expresión génica en un contexto celular dado en la miogénesis. En los últimos años, y gracias a los avances en métodos genómicos como la inmunoprecipitación de cromatina junto con la tecnología de secuenciación de última generación, el mapeo de la organización tridimensional de la cromatina en dominios topológicos, se ha logrado la comprensión de la dinámica de la regulación de expresión genética específica de tejido de forma integral. Sin

embargo, una limitación de los métodos actuales es la necesidad de conocer un objetivo de interés, ya sea un factor de transcripción o un elemento regulador (promotor o potenciador), así como la disponibilidad de reactivos comerciales de calidad suficiente para realizar experimentos bioquímicos y genómica a gran escala.

Con este interés, aquí reportamos el establecimiento de un método para la identificación de complejos de regulación transcripcional de loci musculares de interés en miogénesis, utilizando una variante del sistema CRISPR-Cas9. Esto consiste en la sobreexpresión de una nucleasa inactiva dCas9 biotinilada y RNAs guías (sgRNA) específicos de loci musculares de interés. Después del entrecruzamiento y la fragmentación de la cromatina, la inmunoprecipitación de complejos por estreptavidina permitirá la purificación de proteínas, ARN y ADN presentes en uno o más elementos reguladores. Esto permitirá enfoques proteómicos y genómicos a gran escala para identificar nuevos componentes en la regulación epigenética en el modelo celular C2C12.



Abstract



Adult skeletal muscle is characterized by a high capacity for regeneration, which is dependent on the presence and functionality of muscle stem cells, called satellite cells (SC). In muscular homeostasis, SC are in a state of rest; after a stimulus induced by physical damage or disease, the SCs are activated and re-enter the cell cycle mediated by a complex process of transcriptional regulation, where Pax7 unequivocally marks the muscular lineage in SC and hierarchically modulates the expression of myogenic regulatory factors (MRF), Myf5 and MyoD. As myogenesis progresses, MyoD factor induces myogenin and MRF4 expression to culminate in the expression of differentiated skeletal muscle structural proteins.

The differential expression of MRF and its activity in the process of muscle differentiation and regeneration is controlled at multiple levels. At the epigenetic level, the combination of modifications in the chromatin structure plays a dynamic and decisive role in muscle regeneration. For example, the presence of remote transcriptional regulatory elements, such as enhancers and promoters, synthesis of non-coding RNA, post-translational modifications of regulatory proteins, as well as the combination of transcription factors and chromatin remodeling complexes, regulate gene expression. In recent years, and thanks to advances in genomic methods such as chromatin immunoprecipitation coupled with next generation sequencing technology, as well as chromatin conformation capture based technologies, mapping of the three-dimensional organization of chromatin in topological domains has allowed understanding the dynamics of regulation of tissue-specific gene expression comprehensively. However, a limitation of current methods is the need to know a target of interest, either a transcription factor or a regulatory element (promoter or enhancer), as well as the availability

of commercial reagents of sufficient quality to carry out biochemical experiments and large-scale genomics.

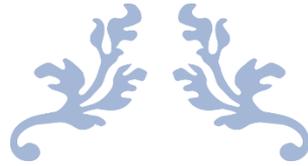
With this interest, we report here the establishment of a method for the identification of transcriptional regulation complexes of muscle loci of interest in myogenesis, using a variant of the CRISPR-Cas9 system. This consists of overexpression of a biotinylated dCas9 inactive nuclease and sgRNA specific to muscle loci of interest. After crosslinking and fragmentation of chromatin, immunoprecipitation of complexes by streptavidin will allow the purification of proteins, RNA and DNA present in one or more regulatory elements. This will allow large-scale proteomic and genomic approaches to identify new components in epigenetic regulation in the C2C12 cell model.

Dedicatoria

A mi Creador

A mi Familia

A mi Compañera de Camino



Agradecimientos

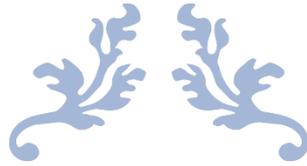


Este proyecto de investigación fue posible gracias al financiamiento otorgado por CONACYT en la beca No. 280518, y a las instalaciones y equipamiento facilitados por el departamento de Genética y Biología Molecular del Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional.

Agradezco al Dr. José Manuel Hernández Hernández por haberme aceptado en su grupo de trabajo, por su apoyo incondicional para llevar a cabo el proyecto, por los consejos y la guía en el camino, y por aceptar y apoyar mis ideas y proyectos. Gracias por su ayuda, sus consejos, sus enseñanzas y su confianza.

También quisiera agradecer a mis asesores el Dr. Oscar Hernández Hernández y el Dr. Bulmaro Cisneros Vega por su ayuda a lo largo de este camino, por sus consejos y observaciones y por alentarme a mejorar.

Agradezco por último y no menos importante a mis padres Rodolfo Ávila Bonilla y Carolina Avilés Ramon por su apoyo incondicional a lo largo de toda mi vida, por su amor, comprensión, ejemplo, esfuerzo y entrega.



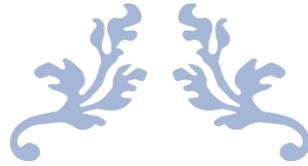
Índice



Resumen	2
Abstract	4
Dedicatoria	6
Agradecimientos	7
Índice	8
Índice de Abreviaturas	11
Índice de Figuras	13
Índice de Tablas	15
Introducción.....	16
1.1 Desarrollo muscular y regeneración.....	16
1.1.1 El músculo esquelético	16
1.1.2 Origen del músculo esquelético.....	17
1.1.3 Células satélites.....	19
1.1.4 Factores reguladores miogénicos (MRF).....	22
1.2 Niveles de organización y regulación de la cromatina en la miogénesis	
23	
1.2.1 Nivel de secuencia.....	24
1.2.2 Nivel de estructura.....	25
1.2.3 Nivel de disposición en 3D	29
1.3 Determinación de elementos regulatorios	31
1.3.1 Métodos basados en la inmunoprecipitación de la cromatina	32

1.3.2 Métodos basados en la detección de la accesibilidad a la cromatina	33
1.3.3 Métodos de Captura de la Conformación de la Cromatina	35
1.3.4 Purificación por afinidad CRISPR de elementos regulatorios (CAPTURE).....	35
Justificación.....	38
Objetivos	39
3.1 Objetivo General.....	39
3.2 Objetivos Particulares	39
Hipótesis.....	40
Metodología.....	41
5.1 Creación de Biblioteca de plásmidos	41
5.1.1 Formación de células calcio competentes.....	41
5.1.2 Transformación por choque térmico de células DH5 α	41
5.1.3 Miniprep para extracción de DNA plasmídico.....	42
5.1.4 Digestión enzimática para comprobación de transformación.	42
5.2 Cultivo de celular.....	42
5.3 Ensayos de diferenciación	43
5.4 Criopreservación de líneas celulares	43
5.5 Curva de Muerte Celular	44
5.6 Ensayo de lipofección	44
5.7 Extracción de RNA total	45
5.8 Retro-transcripción y PCR (RT-PCR).....	45
5.9 Diseño de sgRNAs.....	47
5.10 Clonación de sgRNA.....	47
5.11 Transducción de sgRNA	48
5.12 Análisis comprobatorios del sistema.....	49
5.12.1 Comprobación de especificidad de blanco mediante PCR.....	49

5.12.2	Comprobación de obtención de proteínas asociadas al locus de interés	49
5.12.3	Comprobación de identificación de RNAs asociados al locus de interés	50
Resultados	51
6.1	Establecimiento de una biblioteca de plásmidos	51
6.2	Generación de la línea celular estable C2C12-dCas9/BirA	53
6.3	Diseño de sgRNAs	58
Discusión	60
Perspectivas futuras	65
Conclusiones	67
Anexos	81

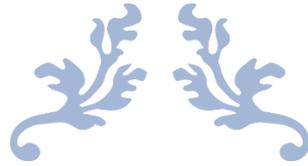


Índice de Abreviaturas



AMP	Ampicilina
ATAC	Assay for Transposase Accesible Chromatine
bHLH	hélice-asa-hélice
BMP4	Proteína morfogenética ósea 4
CAPTURE	Purificación por afinidad CRISPR de elementos regulatorios
Cdkn	Inhibidor de cinasas dependientes de ciclina
cDNA	DNA complementario
CE	Core Enhancer
ChIP	Inmunoprecipitación de la cromatina
CS	Células satélites
DM	Distrofia miotónica
DMD	Distrofia muscular de Duchenne
DMEM	Eagle's minimal essential médium
DNA	Ácido desoxirribonucleico
DRR	Región distal Regulatoria
eRNA	enhancer RNA
FAIRE	Formaldehyde Assisted Isolation of Regulatory elements
FAK	Quinasa de Adhesión Focal
FGF	Factor de crecimiento de fibroblastos
FSHD	Distrofia fascioescapulohumeral
Fzd7	Frizzled-7
H	Histona
HGF	Factor de crecimiento de hepatocitos
IN	insulated neighborhood
K	Lisina

kb	Kilobase
LB	Luria Bertani
lncRNAs	RNAs largos no-codificantes
mb	Megabase
me	Metilación
MIBP	Proteína de unión a la integrina 1
Miog	Miogenina
miRNA	microRNA
MPT	Modificaciones post-traduccionales
MRF	Factores reguladores musculares
mRNA	RNA mensajero
mRNP	Partículas ribonucleoproteicas
Myf5	Factor miogénico 5
MyHC	Cadena pesada de miosina
MyLC	Cadena ligera de miosina
MyoD	Proteína 1 de diferenciación miogénica
NEAT1	Nuclear paraspeckle assembly transcript 1
NO	Óxido nítrico
NOS	Óxido-nítrico sintasa
PCR	Reacción en cadena de la polimerasa
Prmt4	Metil-transferasa de argininas Carn1
PRR	Región proximal regulatoria
RNA	Ácido ribonucleico
RT	Retrotranscripción
SE	Duero Equino
seq	Secuenciación
SFB	Suero Fetal Bovino
sgRNA	RNA guía
Shh	Sonic hedgehog
Spry1	Sprouty 1
TAD	Dominio topológicamente Asociado
TSS	Sitio de inicio de la transcripción
Wnt	Wingless-type

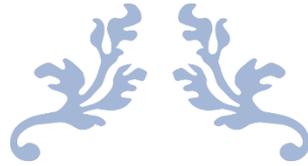


Índice de Figuras



Figura	Página
1 Origen embrionario del músculo esquelético.....	2
2 Miogénesis.....	7
3 Niveles de regulación Transcripcional. A) nivel de secuencia, B) nivel de estructura y C) nivel de disposición en 3D.....	9
4 Métodos para mapear la accesibilidad del genoma.....	19
5 Esquema del método de Purificación por afinidad CRISPR de elementos regulatorios (CAPTURE).....	22
6 Geles de electroforesis al 0.8%, para los plásmidos: a)pEF1a-FB-dCas9-Puro, b)pEF1a-BirA-V5-neo, c)Pslq1651-sgRNA(F+E)-sgGAL4, d)Pslq1651-sgTelomere (F+E), e)psPAX2 y f)N2c/VSVg.....	37
7 Curva de muerte celular para la determinación de la dosis mínima letal a 96 horas para a) puromicina como para b) G418.....	38
8 Prueba de diferenciación de las líneas C2C12 y C2C12-dCas9. A) mioblastos C2C12, b) miotubos con 72 horas de diferenciación de C2C12, c) mioblastos C2C12-dCas9 y d) miotubos C2C12-dCas9 a las 72 horas de diferenciación.....	39

9	Gel de Agarosa de productos de rtPCR para los mRNA de RSP18, Miogenina y la cadena pesada de Miosina; de células C2C12 en estado de mioblastos (MB) y en estado de miotubos (MT) y para células C2C12-dCas9 en las mismas condiciones.....	40
10	Prueba de diferenciación de las líneas C2C12 y C2C12-dCas9/BirA. A) mioblastos C2C12, b) miotubos con 72 horas de diferenciación de C2C12, c) mioblastos C2C12-dCas9/BirA y d) miotubos C2C12-dCas9/BirA a las 72 horas de diferenciación.....	41
11	Gel de Agarosa de productos de rtPCR para los mRNA de MyoD, Miogenina y la cadena pesada de Miosina; de células C2C12 en estado de mioblastos (MB) y en estado de miotubos (MT) y para células C2C12-dCas9/BirA en las mismas condiciones.....	42
12	Gel de Agarosa de productos de rtPCR para los mRNA de RPS18, dCas9 y BirA; de células C2C12 y células C2C12-dCas9/BirA.....	43



Índice de Tablas



Tabla	Página
1 Descripción de digestión enzimática.....	27
2 Descripción de medios de cultivo.....	28
3 SgRNAs diseñados para establecimiento del sistema CAPTURE.....	44
Anexos	
1 Lista de Primers.....	66
2 Lista de sgRNAs diseñados.....	67



Introducción



1.1 Desarrollo muscular y regeneración

1.1.1 El músculo esquelético

En los seres humanos el cuerpo está conformado por más de 600 músculos esqueléticos, aportando alrededor del 28% al 54% del peso corporal total [1]. El músculo esquelético tiene funciones asociadas como: la locomoción, la termogénesis, el mantenimiento de la postura corporal, el metabolismo energético, así como el rol del mantenimiento de las actividades físicas. Por otro lado, también se le han asociado funciones como el control del estado metabólico, la producción y secreción de proteínas con funciones endócrinas [2,3].

Estructuralmente, el músculo esquelético adulto está compuesto por numerosas miofibrillas multinucleadas; unidades funcionales responsables de la contracción muscular, compuestas de la maquinaria contráctil necesaria para funciones fundamentales del organismo. Cada fibra muscular contiene varios cientos a varios miles de miofibrillas. Cada miofibrilla a su vez, está formada por aproximadamente 1500 filamentos de miosina y 3000 filamentos de actina adyacentes entre sí. Así mismo, cada miofibrilla es originada por cientos de células musculares fusionadas tras su diferenciación terminal [4]. Este proceso de diferenciación muscular puede tener dos orígenes temporales; por una parte en el proceso de embriogénesis mediante progenitores embrionarios y por otra parte en la etapa adulta mediado por células madre miogénicas; a partir de ambos orígenes, las células se diferencian primariamente a mioblastos que expresan marcadores específicos de músculo esquelético y se comprometen irreversiblemente al linaje muscular esquelético, los mioblastos se diferencian a miocitos los cuales expresan factores asociados a la diferenciación terminal y se

fusionan formando miotubos, que expresan proteínas estructurales y tejido específico conformando así el músculo esquelético [4].

1.1.2 Origen del músculo esquelético

Entenderemos por miogénesis, al proceso de determinación y formación de las células del tejido muscular [5]. Este proceso se origina del mesodermo paraxial embrionario, y es recordado en el músculo esquelético maduro para su mantenimiento y reparación. El mesodermo se segmenta en estructuras llamadas somitas, donde surge el esclerotoma (que originará hueso y cartílago) y el dermomiótoma (que dará origen a la dermis y a los precursores musculares). La miogénesis embrionaria murina se puede subdividir en varias etapas. La primera es la inducción de las células del dermomiótomo, donde se especifican las células que serán los futuros mioblastos de dos poblaciones: una epiaxial y una hipoaxial. Los mioblastos epiaxiales son inducidos por el estímulo de los factores wingless-type 1 (Wnt1) y Sonic hedgehog (Shh) que producirán la expresión del factor miogénico 5 (Myf5), mientras que los hipoaxiales lo serán por wingless-type 7a (Wnt7a), la proteína morfogenética ósea 4 (BMP4) y Noggin, culminando con la expresión de la proteína 1 de diferenciación miogénica (MyoD) [6] (Figura1).

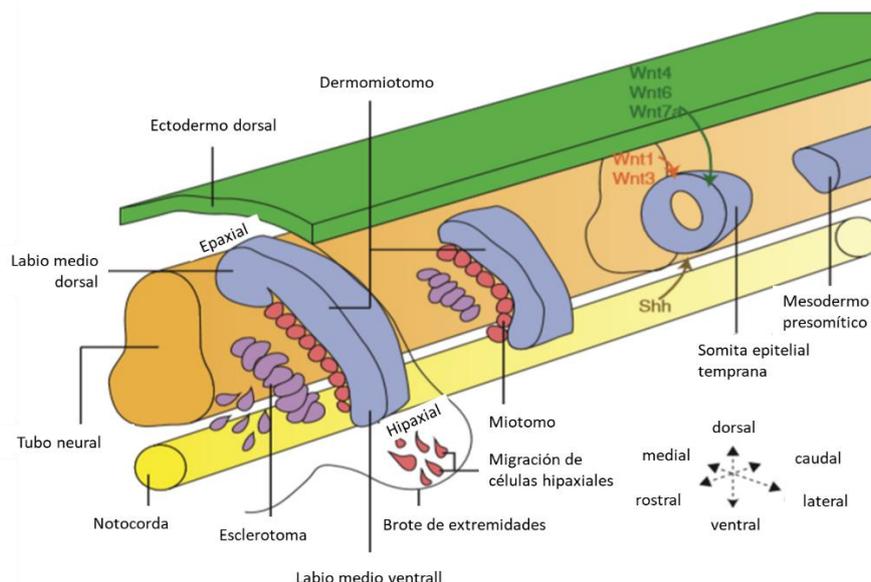


Figura 1. Origen embrionario del músculo esquelético. Se muestra una representación tridimensional a través del embrión durante las etapas tempranas de somitogénesis (final caudal, derecha) y tardía (extremo rostral, izquierda). El mesodermo presomítico se segmenta en bloques de estructuras esféricas, llamadas somitas. Los morfógenos modelan las somitas para

convertirse en el dermomiótomo y el esclerotomo bajo el control de Wnts que emanan del tubo neural (Wnt1 y Wnt3) y el ectodermo de superficie (Wnt4, Wnt6 y Wnt7a) y Sonic Hedgehog (Shh) de la notocorda. Dentro del dermomiótomo, el dominio hipaxial da lugar a los músculos del tronco y las extremidades, incluida una población de células hipaxiales que migran hacia el brote de las extremidades en desarrollo. El dominio epaxial del dermomiótomo da lugar a los músculos dorsales de la espalda. (Ectodermo dorsal, verde; mesodermo, azul; miótomo y células musculares, rojas; tubo neural, naranja; notocorda, amarillo; esclerotomo, morado). [7]

En la segunda etapa los mioblastos proliferan y migran, dicha migración es posibilitada por la disminución de las β -integrinas y la formación de contactos focales mediante la interacción de la proteína de unión a la integrina 1 (MIBP) con la quinasa de Adhesión Focal (FAK) y la Paxilina [8]. Durante la tercera etapa los mioblastos se agregan y diferencian en miocitos maduros, sin embargo, algunos quedan en estado quiescente y son los que darán lugar a las células satélites en el músculo esquelético maduro. En esta etapa las N-Cadherinas y las integrinas $\alpha 7A$, $\alpha 7B$ y $\beta 1D$ se ven incrementadas y la maquinaria del ciclo celular detenida, activando la diferenciación miogénica mediante la activación de la MAPK p38 [6,9].

En ratón, la miogénesis embrionaria se da en dos fases u olas de miogénesis: una fase primaria que se presenta hacia los días embrionarios 10.5 (E10.5) a E12.5 y una fase secundaria, hacia los días E14.5 a E17.5.

La primera fase da como resultado la producción de las miofibras primarias, que se derivan de los progenitores Pax3+ en ratón. Estas miofibras primarias forman los primeros miotomas y músculos de las extremidades, proporcionando las plantillas sobre las cuales se construirán los músculos adultos y se caracterizan por la expresión de un conjunto específico de proteínas, como la cadena pesada (MyHC) y ligera (MyLC1) de miosina. Los tejidos embrionarios proporcionan señales inductivas clave para la miogénesis esquelética. Wnt, BMP y Shh representan las principales vías de señalización para la inducción de diferentes destinos somáticos. Se ha descrito que el primer signo de miogénesis es la activación del factor miogénico Myf5 y que el programa miogénico está controlado por una red central de factores de transcripción, que incluye Pax3 y un conjunto de factores reguladores musculares (MRF) que consta de Myf5, MyoD (Myod1), Miogenina y MRF4 (Myf6) [10].

La segunda fase de la miogénesis en el ratón se inicia con la presencia de progenitores miogénicos que expresan los factores transcripcionales Pax3 y Pax7; estos precursores proliferan y se fusionan entre sí o con las fibras primarias y dan lugar a fibras secundarias que expresan marcadores específicos tales como β -enolasa, Nfix y MyLC3. Además, se ha reportado que el factor de crecimiento de hepatocitos (HGF) producido por el mesodermo de la placa lateral es esencial para la correcta migración de los mioblastos y que la señalización de insulina actúa con la señalización de Wnt para promover la miogénesis y la fusión de mioblastos. Además, la señalización del factor de crecimiento de fibroblastos (FGF) promueve la proliferación de mioblastos [10]

1.1.3 Células satélites

El músculo esquelético adulto se caracteriza por tener una alta capacidad de regeneración. Esta capacidad depende de la presencia y funcionalidad de las células troncales musculares, llamadas células satélites (CS). Las CS tienen una alta capacidad de replicación y regeneración, y comprenden del 2 al 7% del total de los núcleos asociados a una miofibrilla en el músculo esquelético adulto [4].

Las CS tienen su origen en el desarrollo embrionario de ratón hacia el día E16.5, procedente de una población celular ubicada en el mesodermo paraxial, adyacente al tubo neural. Mientras la gran mayoría de los precursores musculares entran en diferenciación muscular dando origen a miofibrillas, una población se mantiene indiferenciada y ubicada entre el sarcolema y la matriz extracelular [11].

En homeostasis muscular las CS se encuentran en estado quiescente; con bajos niveles metabólicos y transcripcionales [12]. Tras un estímulo inducido por un daño físico, estímulos de crecimiento o algunas enfermedades, las CS se activan y entran al ciclo celular, para, por una parte, originar células precursoras miogénicas (mioblastos) necesarios para la regeneración muscular y por otro lado para autorrenovarse manteniendo el reservorio de CS [13].

Análisis comparativos de perfiles de expresión génica, han revelado un grupo de cerca de 500 genes que se encuentran sobre-expresados en CS quiescentes, en comparación con células activadas. Entre ellos, se incluyen reguladores negativos de la progresión del ciclo celular como el inhibidor de cinasas

dependientes de ciclina 1B (Cdkn1b) y 1C (Cdkn1c), la proteína supresora de tumores Rb y el factor inhibidor del factor de crecimiento de fibroblastos, sprouty 1 (Spry1). Se sugiere que la expresión de estos factores disminuye a medida que las CS se activan e ingresan al programa de diferenciación [14]. También, se ha descrito que la expresión del factor transcripcional Foxo3 es mayor en las CS quiescentes. En estas, Foxo3 induce la expresión de Notch1, cuya señalización favorece el mantenimiento de la quiescencia celular. De hecho, alteraciones en la señalización de Notch, como los son los factores RBP-J, Hes y Hey, causan activación espontánea de las CS y una disminución de su población total [15]. Adicionalmente, se ha determinado la participación de micro-RNAs en el mantenimiento de la quiescencia de CS. Por ejemplo, Myf5 es blanco del micro-RNA miR-31; juntos son secuestrados en partículas ribonucleoproteicas (mRNP) presentes en CS quiescentes [16-17]. Tras la activación de las CS, los gránulos mRNP son disociados permitiendo la rápida traducción de Myf5 y con ello la progresión del programa de diferenciación muscular. En respuesta a las señales de activación, en las CS activadas se induce la expresión de genes involucrados en la síntesis de RNA y DNA, así como en la progresión del ciclo celular. Adicionalmente, marcadores de superficie que participan en la transducción de señales de activación se expresan en CS activadas, entre ellos, c-Met, FGFR1, FGFR4, syndecan-3 y syndecan-4 [18-21]. Por otro lado, el estrés mecánico producido por el ejercicio o la ruptura de las miofibrillas induce la liberación de TNF-alfa, óxido nítrico (NO) y HGF los cuales actúan como señales de activación; Como consecuencia del daño a la miofibrilla, la producción de NO por la óxido-nítrico sintasa (NOS) dispara la liberación de HGF desde la matriz extracelular, el cual se asocia con su receptor c-Met. Aunque el papel de c-Met en la activación celular es poco conocida, HGF y FGF potencian la señalización de la vía de las MAPK cinasas para modular la proliferación y diferenciación celular. Por ejemplo, se ha observado que la señalización de p38 a/b MAPK es fundamental para la proliferación de las CS [22-23].

Las CS son una población heterogénea con rasgos fenotípicos y funcionales distintos. Indistintamente de su estado metabólico, todas las CS expresan el factor transcripcional Pax7 [13]. Se ha determinado que un 80% de las células Pax7⁺ son a la vez positivas para el MRF Myf5, mientras que el 20% restante

únicamente expresan Pax7. Se ha observado que las células Pax7⁺/Myf5⁻ pueden dar lugar tanto a células Myf5⁺ como Myf5⁻ a través de divisiones asimétricas que suceden a lo largo del eje apical-basal, perpendicular a la fibra muscular. Por otro lado, las células Pax7⁺/Myf5⁺ solo pueden generar células Myf5⁺ a través de divisiones simétricas generadas en el plano longitudinal de la fibra muscular. Estas evidencias sugieren que la población de células Pax7⁺/Myf5⁻ funcionan como verdaderas CS troncales, y las células Pax7⁺/Myf5⁺, más sensibles a ingresar al programa de diferenciación muscular, funcionan como CS miogénicas, destinadas a la reparación de la fibra muscular [24-25].

Los mecanismos moleculares que regulan la expansión simétrica o asimétrica de las CS han sido asociados con la segregación asimétrica de factores citoplasmáticos, así como la distribución diferencial de componentes de la matriz extracelular. Por ejemplo, se ha demostrado que la vía no canónica de señalización por Wnt dirige la expansión de células Myf5⁻ a través de divisiones simétricas. La expresión muscular de Wnt7a incrementa durante el proceso de regeneración y su receptor de membrana Frizzled-7 (Fzd7) es más abundante en las células Myf5⁻. Durante la división de células Myf5⁻, Fzd7 induce la expresión polarizada de Vangl2 en respuesta a Wnt7a para activar la vía de señalización de polaridad planar (PCP) y de esta manera asegurar la expansión de la población de células Pax7⁺/Myf5⁻. Además, Wnt7a/Fzd7 activan la vía de AKT/mTOR en las fibras musculares adultas para inducir hipertrofia muscular, teniendo la función dual de expansión de progenitores musculares y la de incremento de la masa muscular de las fibras diferenciadas [26]. Por otro lado, las CS tienen la capacidad de remodelar su nicho celular para regular su propia expansión. Las células Myf5⁺ activadas durante la regeneración producen grandes cantidades de la glicoproteína de matriz extracelular Fibronectina, la cual incrementa la actividad de Wnt7a durante la regeneración muscular [26-27]. La inhibición de la expresión de Fibronectina en CS aisladas reduce su capacidad de regenerar el nicho de CS después de su trasplante en ratones receptores.

A nivel molecular, distintas redes de regulación transcripcional controlan jerárquicamente el desarrollo y la diferenciación muscular. Pax7 induce la expresión de Myf5 para dar lugar a la diferenciación. La expresión de MRFs

como Myf5 y MyoD aseguran el compromiso miogénico y dirigen el programa irreversible de diferenciación de mioblastos, los cuales proliferan originando progenitores miogénicos. En los progenitores miogénicos, la expresión de Pax7 disminuye y los MRFs Miogenina (Myog) y MRF4 inducen la expresión de genes de diferenciación terminal y proteínas estructurales que permitirán la fusión y la funcionalidad de la miofibrilla madura (Figura 2) [28].

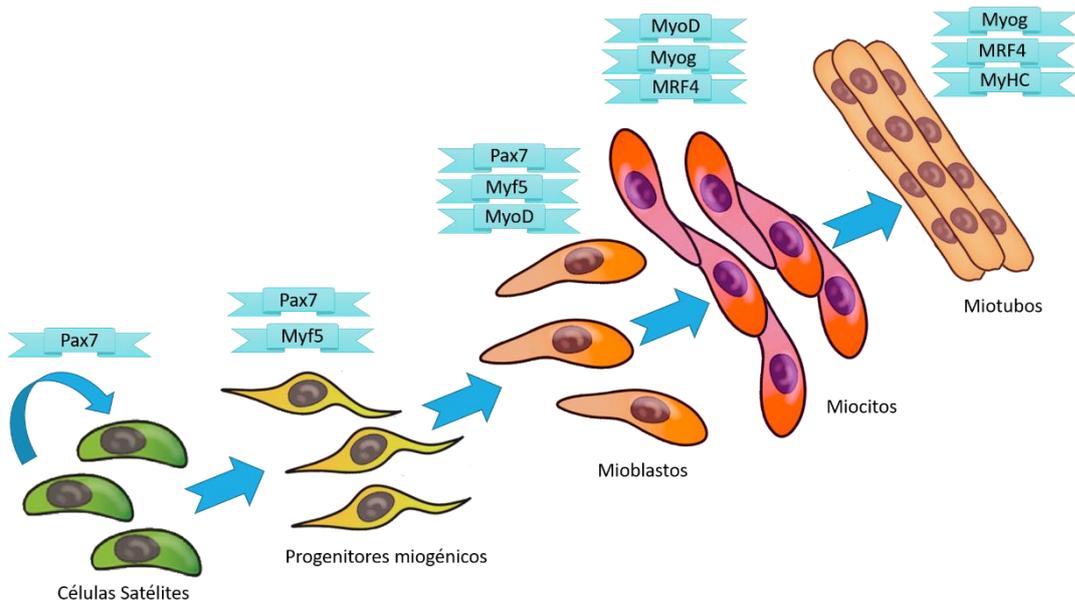


Figura 2. Miogénesis. Se muestra la expresión de genes miogénicos implicados en la activación, diferenciación y fusión de células miogénicas, junto con la transición de células satélite a mioblastos, miocitos, y finalmente fusionándose en miotubos o miofibras

1.1.4 Factores reguladores miogénicos (MRF).

Dado el papel fundamental de los MRFs en la miogénesis mostramos un acercamiento para su comprensión. La familia de los MRFs se constituye por los factores MyoD, Miogenina, Myf-5 y MRF4 (también conocido como Herculin o MYF6), los cuales son proteínas básicas con estructura de hélice-asa-hélice (bHLH) [29]. La estructura proteica de los MRFs está altamente relacionada, estos factores básicos de transcripción de clase II de hélice-asa-hélice (bHLH) y contienen tres dominios conservados: el dominio de transactivación amino terminal con una zona rica en histidina/cisteína; la región central con el motivo bHLH que incluye el dominio básico helicoidal y Hélice I y II; y otro dominio de transactivación en el terminal carboxilo que contiene Hélice III [30]. Los aminoácidos básicos en bHLH son importantes para la unión del DNA, así como para el mantenimiento de la conformación de un dominio topológico asociado

(TAD), que es fundamental para la activación de genes musculares [31] Además, la región básica contiene una secuencia de alanina-treonina llamada código miogénico que permite la activación específica de los genes musculares [31,32]. El dominio bHLH les confiere la capacidad de reconocer las secuencias llamadas cajas E (CANNTG) en las regiones reguladoras de genes blanco; esto se logra tras la heterodimerización con un miembro de la familia de proteínas E expresada de forma ubicua [29].

Por otro lado, diferentes partes de la región promotora de MRF4 pueden ser trans-activadas por todos los otros MRF, pero MRF4 no tiene una función de autoactivación para mejorar su nivel de expresión, sin embargo, se requiere MRF4 para regular negativamente el nivel de Myogenina durante la terminal diferenciación [33].

Durante la fase temprana de la miogénesis, aunque se requieren moléculas inductivas para la regulación positiva de MyoD y Myf5 en la somita y el inicio de la diferenciación del músculo esquelético [34,35], MyoD y Myf5 son esenciales para el establecimiento y mantenimiento de las células precursoras miogénicas. La miogenina es importante para la diferenciación terminal miogénica, y MRF4 está involucrado en la miogénesis temprana y terminal.

1.2 Niveles de organización y regulación de la cromatina en la miogénesis

La expresión diferencial de los MRFs en el proceso de diferenciación y regeneración muscular es controlada a distintos niveles, entre ellos destaca el control de la expresión a nivel transcripcional. Dicho control es logrado mediante la presencia de secuencias reguladoras en *-cis*, como promotores y “enchancers” y los factores asociados a cromatina que controlan el desarrollo y la fisiología mediante la regulación de la transcripción de genes blanco [36]. Esta regulación está relacionada con varios niveles de organización del material genético, como serán detallados en los subsecuentes apartados (Figura. 3).

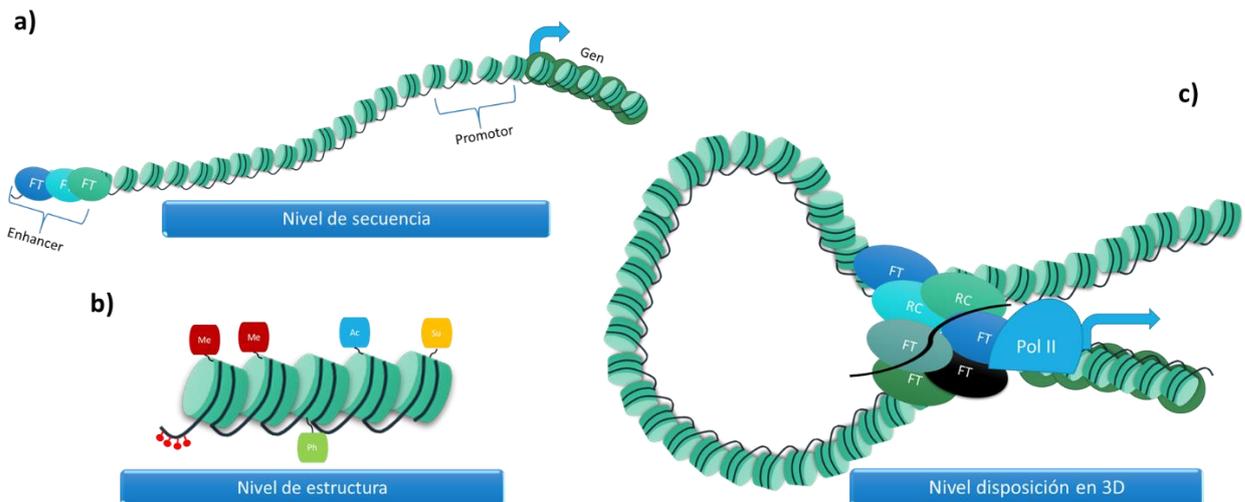


Figura 3. Niveles de regulación Transcripcional. A) nivel de secuencia, B) nivel de estructura y C) nivel de disposición en 3D.

1.2.1 Nivel de secuencia

A nivel de secuencia de nucleótidos, la transcripción de un gen depende de la presencia de promotores, secuencias río arriba del sitio de inicio de la transcripción y que contiene los elementos necesarios para el reclutamiento de la RNA polimerasa II, de factores basales de la transcripción y de factores transcripcionales tejido-específicos. Adicionalmente, los “enhancers” o enhanceres, son secuencias de DNA que pueden estar presentes río arriba o río abajo del sitio de inicio de la transcripción, a distancias que pueden llegar al orden de las megabases con respecto a gen blanco y que promueven la unión de factores de transcripción tejido específicos, remodeladores de la cromatina y RNAs regulatorios [37]. Se ha propuesto que el catálogo de “enhancers” activos en una célula particular, determina la identidad celular durante el desarrollo (Figura 3a).

Generalmente, el contacto directo entre los “enhancers” y los promotores a nivel tridimensional, se genera a través de proteínas estructurales como las cohesinas, lo que permite el establecimiento de asas en donde se enriquece el contenido de maquinaria transcripcional y, por lo tanto, el aumento en las tasas de transcripción [36].

Un papel fundamental en el desarrollo miogénico es el control de la expresión a nivel de promotores y enhancers realizado por MyoD. MyoD interactúa con las proteínas E para formar heterodímeros de unión al DNA. Estos heterodímeros

activan sinérgicamente la transcripción específica de genes miogénicos de una manera definida temporalmente [38]. El complejo MyoD/proteína-E exhibe preferencia por sitios de unión consenso en el DNA, conocidas como cajas E. Las cajas E contienen la secuencia consenso VCASCTGT, y se han encontrado enriquecidas dentro de las regiones promotoras y enhancers asociadas con genes de expresión muscular [39-42]. Además, se ha determinado que tanto *in vivo* [43] como *in vitro* [44,45] MyoD tiene la propiedad de activar la transcripción de genes que contienen múltiples motivos de cajas E. Aunque los mecanismos por los cuales MyoD contribuye a la diferenciación muscular se han centrado principalmente en la regulación de las regiones promotoras de los genes, diferentes experimentos han revelado una dependencia de MyoD para la actividad enhancer. Se han identificado regiones enhancers críticas para el programa miogénico y enriquecidas con elementos E-box, que reclutan MyoD a los genes distales como miogenina, [46,47] MyoD [48], la cadena 1 de miosina [49-51], Ckm [52,53] Myf5 y Mrf4 [54,55], γ -sarcoglicano [56] y Adamts5 [57].

En el contexto muscular, la presencia de un enhancer localizado a -25kb del gen MyoD llamado "Core Enhancer" (CE), regula la transcripción de MyoD. Además de reclutar maquinaria transcripcional, el CE es también transcrito por la RNA pol II en un RNA denominado "enhancer RNA" (eRNA), el cual funciona en cis, ubicándose en el promotor de MyoD y reclutando la desmetilasa de histonas LSD1, permitiendo la apertura de la cromatina y su posterior transcripción [58].

Así mismo se ha determinado que Pax7 activa la expresión de Myf5 a través de diferentes sitios de unión ubicados en los enhancers de 57.5 kb, 111 kb y 129 kb, río arriba al sitio de inicio transcripcional de Myf5. Mientras que el enhancer de 111 kb impulsa la expresión de Myf5 en células satélite quiescentes, el enhancer de 57.5 kb está más relacionado con las células satélite inactivadas de expresión de Myf5 y los mioblastos proliferantes [59, 60-62]

1.2.2 Nivel de estructura

Un segundo nivel de regulación transcripcional está basado en la organización del DNA en cromatina y nucleosomas. El nucleosoma es la unidad funcional de la cromatina y es una estructura conformada por un octámero de histonas H2A, H2B, H3 y H4. La cromatina es una molécula dinámica susceptible a un gran número de modificaciones y arreglos estructurales. Las principales

modificaciones en la cromatina son: la metilación del DNA, modificaciones post-traduccionales de las histonas como metilación, acetilación, ADP-ribosilación, ubiquitinación y fosforilación. Estas modificaciones son catalizadas por diversas enzimas en regiones específicas del genoma y su combinatoria tiene un efecto funcional directo en diversos procesos celulares (Fig. 3b). Por ejemplo: se han asociado la acetilación, la metilación de la arginina así como la metilación de las lisinas (K) 4, 36 y 79 de la histonas 3 (H3) como marcas de activación transcripcional, mientras las metilaciones en H3K9, H3K27 y H4K20 como modificaciones en las histonas asociadas al silenciamiento de genes. Estas modificaciones median una distribución de la cromatina en dos estados: la cromatina en estado abierto o eucromatina que es transcripcionalmente activa, y la cromatina en estado cerrado o heterocromatina, que es transcripcionalmente no permisiva [9]. Las histonas son el blanco de una gran variedad de modificaciones post-traduccionales (MPT), que se producen principalmente, pero no exclusivamente, en los extremos N-terminales de estas. Su incorporación puede generar patrones de modificaciones (como ejemplo, histonas con lisinas metiladas o acetiladas), los cuales son reconocidos por proteínas específicas (llamadas “lectores”), que contienen dominios especializados de reconocimiento y pueden interactuar diferencialmente con cada una de estas modificaciones o los patrones que ellas formen. El reclutamiento de proteínas “lectoras” de la cromatina, ayuda a la incorporación de actividades enzimáticas que tendrán un efecto en la regulación de la expresión génica [37].

En el proceso de miogénesis la activación transcripcional específica de determinados genes que inducen la diferenciación e identidad muscular es fundamental, uno de los mecanismos más comunes es mediante la hiperacetilación de histonas. Los estudios en mioblastos C2C12 sugieren que la acetilación de histonas es globalmente más alta en el estado indiferenciado, en comparación con las células inducidas a diferenciarse en miotubos [42]. Sin embargo, se ha observado un aumento en la acetilación de histona H4 durante la diferenciación específicamente en genes diana de MyoD, lo que sugiere que la acetilación de histona aumenta en un subconjunto de genes importantes para la diferenciación miogénica [42]. En este sentido y en el contexto muscular, se

ha demostrado que hay dos pasos de activación para la inducción de miogenina: el primero requiere actividad de PCAF para acetilar las histonas centrales de la región promotora de miogenina, así como del factor MyoD; y el segundo requiere la actividad dependiente de p38 del complejo de remodelación SWI/SNF para proporcionar una conformación abierta en la cromatina que asegure la expresión de miogenina [63].

Por otra parte, otras marcas de histonas asociadas a cromatina activa se han descrito en el proceso de diferenciación muscular. En este sentido se ha mostrado que Pax7 recluta al complejo proteínico con actividad metil-transferasa de histonas que incluye a las proteínas Wdr5, Ash2L y MLL2 al enhancer transcripcional localizado a -57kb del sitio de inicio de la transcripción de Myf5. Este complejo traslocado induce la trimetilación en la lisina 4 de la histona 3 lo que activa al enhancer. Para que este reclutamiento ocurra, las argininas contenidas en el dominio N-terminal de Pax7 deben ser metiladas por la metil-transferasa de argininas Carm1 (Prmt4) durante divisiones asimétricas, generando la señal para la expresión de Myf5 en las células que entrarán al programa de diferenciación terminal. El análisis de divisiones asimétricas en cultivos de fibras musculares demostró que la interacción de Pax7-Carm1 se da exclusivamente en CS miogénicas Pax7+/Myf5+, no obstante que la expresión de estos componentes transcripcionales es idéntica en ambas poblaciones, lo que implica la presencia de reguladores epigenéticos adicionales que modulan la actividad transcripcional durante divisiones asimétricas [64].

De manera opuesta cuando se induce a los mioblastos a diferenciarse en miotubos, se disminuye la expresión de genes que impiden el programa de diferenciación. La represión genética está mediada en parte por el complejo represivo polycomb 2, que comprende la histona H3K27 metiltransferasa Ezh2 [65]. Se ha demostrado que el promotor de Pax7 presenta niveles aumentados de trimetilación H3K27 y niveles reducidos de expresión cuando las células son inducidas a diferenciarse [66]. Aunque se requiere Ezh2 para reprimir la expresión de ciertos genes durante la miogénesis, la expresión de Ezh2 disminuye progresivamente ante las señales de diferenciación [67]; lo que sugiere que la expresión de Ezh2 debe desactivarse para permitir la miogénesis adecuada. Se ha demostrado que la expresión de Ezh2 disminuye por la acción

de microRNA específicos (miRNA) que se expresan cuando se induce a las células a diferenciarse como el miRNA 214 Y EL miRNA 26a [68, 69]. En mioblastos indiferenciados, Ezh2 está presente en las regiones reguladoras de genes necesarios para la diferenciación muscular (Myh10, Ckm, Myog), evitando así su expresión [67, 70]. Cuando los mioblastos inician la diferenciación, Ezh2 deja los promotores de genes específicos del músculo.

Por otra parte, un miembro de la familia de las metil-transferasas de argininas, Prmt5 ha sido implicado en la regulación de la función de las CS. En condiciones normales, el reclutamiento de Prmt5 a regiones reguladoras del gen p21, correlaciona con la metilación de la arginina 8 de la histona 3 (H3R8me2), generando un ambiente transcripcionalmente no permisivo que refleja la represión del gen p21 y la inhibición de la proliferación celular, lo que resulta fundamental para el mantenimiento del estado quiescente. El análisis de la actividad de Prmt5 y su participación durante los procesos de divisiones asimétricas se desconoce, sin embargo, es posible que exista especificidad por los blancos de metilación entre Prmt5 y Carm1 durante la activación de CS, lo cual puede estar relacionado con la segregación asimétrica de componentes celulares durante los ciclos de división [71].

Consiguientemente estos factores asociados a la cromatina pueden ser reclutados mediante RNAs largos no-codificantes (lncRNAs), los cuales funcionan como esponjas y andamios atrayendo a elementos regulatorios de la transcripción de un gen [72, 73]. Existen muchos lncRNA asociados a la modificación de histonas mediante el reclutamiento y la interacción con modificadores de histonas. NEAT1 (Nuclear paraspeckle assembly transcript 1) es un lncRNA que acelera la proliferación de mioblastos e inhibe su diferenciación, el primer efecto al inducir la reducción del nivel de expresión del gen inhibidor de la quinasa dependiente de ciclina P21; el segundo efecto NEAT1 suprimir la transcripción de genes miogénicos específicos, como Myog, Myh4 y Tnni2. Diferentes estudios muestran que NEAT1 interactúa con el enhancer de Ezh2, reclutando a Ezh2 para promotores específicos [74].

Malat1 (metastasis-associated lung adenocarcinoma transcript 1) es un lncRNA que se ha asociado a muchos tejidos con altos niveles de expresión, actuando como regulador durante la diferenciación celular [75,76]. En concreto, la

disminución en la expresión de Malat1 induce la aceleración de la diferenciación muscular demostrada por la diferenciación miogénica en células en cultivo. Malat1 induce la diferenciación miogénica mediante el reclutamiento de Suv39h1 a los loci de MyoD; Suv39h1 inserta una trimetilación en la lisina 9 de la histona 3 (H3K9me3), que suprime la transcripción del gen objetivo [75, 77].

Algunos otros lncRNA pueden regular la miogénesis mediante el reclutamiento de complejos de modificación cromosómica para dirigirse a los promotores de genes, como el lncRNA SYISL (like SYNPO2 intron sense-overlapping lncRNA). SYISL está involucrado en la diferenciación de mioblastos; los estudios muestran que la sobreexpresión de SYISL induce la proliferación de mioblastos y reduce la diferenciación, por el contrario, el silenciamiento de SYISL induce la diferenciación muscular y la regeneración muscular después del daño. Mecánicamente, SYISL induce el reclutamiento específico de PRC2 a los promotores de los genes diana p21, MyoG o Myh4; PRC2 inserta H3K27me3 en los genes de los promotores objetivo [78].

A diferencia del mecanismo de represión génica, es posible una posible activación de la transcripción mediada por la modificación de histonas. Myoparr es un lncRNA asociado al promotor de genes de diferenciación específica muscular [79]. Myoparr interactúa físicamente con Ddx17. Ddx17 es un coactivador que induce la transcripción de MyoD. Mecánicamente, Myoparr induce la interacción entre Ddx17 y PCAF; PCAF es una proteína histona acetiltransferasa, que interactúa con genes específicos miogénicos que regulan la transcripción mediante la remodelación de la cromatina inducida por la marca acetil histona [80].

1.2.3 Nivel de disposición en 3D

Un tercer nivel de regulación es el dado por la arquitectura tridimensional de la cromatina, formada por interacciones entre regiones distales y posibilitada por la presencia de elementos regulatorios [72] (Fig. 3c). Distintos estudios han revelado que el genoma eucarionte se organiza en módulos estructurales en torno a 500kb – 1Mb llamados dominios físicos o con mayor frecuencia dominios topológicamente asociados TADs (Topologically Associating Domains) [81-83]. Los contactos entre regiones genómicas que pertenecen a un mismo TAD se producen con mucha mayor frecuencia que los contactos entre regiones que

pertencen a TADs adyacentes. Los TADs transcripcionalmente activos tienden a asociarse preferentemente con otros TADs activos, mientras que los silenciados tienden a asociarse con otros silenciados [84]. La importancia de estos TADs es que no sólo actúan como unidades de organización estructural, sino que parecen también actuar como unidades funcionales. Así los TADs correlacionan con muchas marcas de actividad de la cromatina tales como: modificaciones de histonas [81,83], expresión génica coordinada [82,85], asociación a la lámina nuclear [81] y coordinación de la replicación del DNA [86]. A pesar de todos los estudios llevados a cabo en los últimos años, no se ha podido aclarar la cuestión de si la organización tridimensional del genoma en estos dominios estructurales es una causa o una consecuencia de las funciones genómicas [87-89]. De esta forma, se producen sitios transcripcionalmente activos que derivan en la transcripción de múltiples genes, comúnmente relacionados e involucrados en procesos celulares comunes [72]

Por su parte los “insulated neighborhoods” (INs), son componentes importantes de los TADs, y son regiones del DNA que contienen uno o más genes y cuyos límites están unidos por las proteínas estructurales CTCF y cohesina [90]. Los INs también limitan la regulación genética dentro de sus límites, al albergar interacciones entre elementos reguladores en *cis*, como la comunicación promotor-enhancer [91]. También se ha determinado que las estructuras genómicas más complejas, como los TADs, están generalmente conservadas en los distintos tipos celulares, mientras que las interacciones de cromatina dentro de un IN son específicas y dinámicas del tipo celular [92].

En músculo esquelético se ha sugerido que MyoD puede regular la expresión génica modificando la arquitectura tridimensional del genoma. En este sentido, Dall’Agnese y colaboradores [93] realizaron estudios de trans-diferenciación de fibroblastos primarios humanos a mioblastos y miotubos, mediante la introducción de un transgen de MyoD inducible. Tras esto, se reportó que MyoD se une a casi 50,000 sitios en el genoma en mioblastos y a 80,000 sitios en miotubos diferenciados. Únicamente el 5% de estos sitios de unión a MyoD ubicados en promotores de genes expresados diferencialmente durante la diferenciación. Además de los elementos promotores, se detectó la unión de

MyoD en los sitios de unión a CTCF y las regiones H3K27ac tanto en mioblastos como en miotubos [93].

En este mismo sentido, experimentos de Hi-C mostraron un patrón de correulación de genes dentro de TADs unidos a MyoD, donde el 14% del genoma interactúa en *-cís* dentro de estos elementos durante compromiso y diferenciación miogénica. Esto se logra mediante la unión de MyoD en las interacciones de cromatina que involucran contactos promotor-promotor y promotor-enhancer, lo que indica que MyoD reconecta la arquitectura de cromatina en los promotores, enhancers e INs durante la trans-diferenciación de fibroblastos a músculo esquelético. Este perfil de unión de MyoD y CTCF al DNA determina la expresión génica, mediante una extensa organización de interacciones que involucran elementos genómicos estructurales y reguladores en *cís*, incluidos promotores, enhancers, INs y proteínas asociadas. Por tanto, la reconfiguración de un IN induce interacciones diferenciales mediadas por MyoD que permite reprimir genes no músculo específicos y activar genes específicos del músculo; mostrando que se requiere una expresión constante de MyoD para el mantenimiento del paisaje tridimensional de cromatina, logrando el compromiso y la diferenciación miogénica [93].

En resumen, la presencia de elementos de regulación transcripcional a distancia como enhancers y promotores, la síntesis de RNAs no codificantes, las modificaciones post-traduccionales de proteínas reguladoras, así como la combinatoria de factores de transcripción y complejos remodeladores de la cromatina, responden a las distintas señales extracelulares para ejecutar la expresión de genes en un determinado contexto celular en la miogénesis.

1.3 Determinación de elementos regulatorios

En años recientes y gracias a los avances en los métodos genómicos como la inmunoprecipitación de cromatina acoplada a tecnología de secuenciación de última generación, así como del mapeo de la organización tri-dimensional de la cromatina en dominios topológicos, se ha logrado el entendimiento de la dinámica de regulación de la expresión genética tejido-específica de forma integral. Podríamos distinguir en general diferentes tipos estrategias para la identificación de elementos regulatorios y los complejos proteínicos reguladores asociados: los basados en la inmunoprecipitación de la cromatina, los basados

en la detección de la accesibilidad a la cromatina y los métodos de Captura de la Conformación de la Cromatina.

1.3.1 Métodos basados en la inmunoprecipitación de la cromatina

Las técnicas basadas en la inmunoprecipitación de la cromatina (ChIP), permiten obtener un perfil “*in vivo*” de las regiones de DNA unidas por proteínas específicas. En los experimentos de ChIP, todas las interacciones DNA-proteína son fijadas, las células lisadas y la cromatina fragmentada mediante sonicación en fragmentos de varios cientos de pares de bases. El uso de anticuerpos específicos contra factores de transcripción o contra modificaciones postraduccionales, permite la precipitación de un conjunto de fragmentos específicos de DNA. La identificación y cuantificación a escala genómica de esta librería de fragmentos de DNA precipitados se puede hacer mediante la hibridación, de los mismos, con microarreglos de DNA (ChIP-chip) o mediante el uso de técnicas de secuenciación masiva (ChIP-seq) [94].

La elección de la diana para el anticuerpo usado en este tipo de experimento determinará, por tanto, el tipo de información obtenida. El uso de anticuerpos contra factores de transcripción específicos permitirá la identificación del subconjunto específico de elementos genómicos reguladores controlados por dichos factores de transcripción en el contexto determinado. Por otro lado, si la diana para el anticuerpo es una proteína remodeladora de la cromatina o modificadora de histonas (p300, ATAC, CHD7, Brg1) asociadas al control de los elementos reguladores en general, se obtendría un mapa general de los elementos reguladores [94]. Del mismo modo, se podría obtener este tipo de mapas generales mediante el uso de anticuerpos contra modificaciones postraduccionales de las histonas asociadas con el estado regulatorios específico de elementos cis-reguladores y promotores. Así, por ejemplo, la mono-metilación de la lisina 4 de la histona 3 (H3K4me1) está asociada a elementos reguladores distales y promotores en estado “pre-activado” [95-96], mientras que la acetilación de la lisina 27 de la histona 3 (H3K27ac) se ha asociado a elementos reguladores distales activos [94-95], y la trimetilación de la lisina 27 de la histona 3 (H3K27me3) se ha asociado con elementos reguladores reprimidos [99-100]. Por tanto, el uso de anticuerpos específicos contra este tipo de modificaciones de histonas permite la obtención del conjunto

de secuencias de DNA reguladores que se encuentran en un estado transcripcional particular en las condiciones experimentales concretas.

1.3.2 Métodos basados en la detección de la accesibilidad a la cromatina

La unión de factores de transcripción a elementos reguladores está condicionada a la accesibilidad de la secuencia de DNA. Por tanto, regiones reguladoras interaccionan con nucleosomas híper-dinámicos, que contienen unas variantes específicas de histonas, H3.3 y H2A.Z, que se depositan de manera independiente a la replicación. Esto hace que estas regiones de la cromatina pasen un mayor tiempo desprovistas de nucleosomas y sean más sensibles a nucleasas de DNA o más permeables a transposasas. Estas tres características físicas (desvinculación de los nucleosomas, sensibilidad a DNAasas y accesibilidad a transposasas) también se han usado como principio para la detección a escala genómica de elementos reguladores [94]. A continuación, se describen los métodos más comúnmente usados:

DNAasal-seq: se basa en el uso de la endonucleasa DNAasa I. Esta enzima es capaz de digerir los fragmentos de DNA que no están asociados a histonas independientemente de su secuencia. A estos fragmentos se les conoce como sitios hipersensibles a la DNAasa I. Tras la digestión con la enzima, la purificación de los fragmentos digeridos se puede hacer bien, mediante la biotinización de los extremos generados y la posterior purificación [101] o bien directamente mediante la purificación de la porción de los fragmentos más pequeños [102]. Tras esta purificación, la librería de fragmentos es secuenciada mediante técnicas de secuenciación masiva (Fig. 4a).

FAIRE (Formaldehyde Assisted Isolation of Regulatory elements). Al igual que el ChIP, se basa en la fijación de las interacciones proteínas-DNA mediante el uso de agentes fijadores tipo formaldehído. Tras la fijación, continúa un paso de sonicación. Los fragmentos de DNA que no estuviesen asociados a nucleosomas, pueden ser separados de los que sí lo están, mediante un paso de fenolización. Esta librería enriquecida en fragmentos no asociados a nucleosomas puede secuenciarse mediante técnicas de secuenciación masiva, obteniéndose un perfil de regiones de cromatina abierta [103] (Fig. 4b).

ATAC-seq (Assay for Transposase Accesible Chromatine using sequencing): En este método se emplea una transposasa bacteriana llamada Tn5. Esta transposasa ha sido ensamblada *in vitro* con adaptadores compatibles con las plataformas de secuenciación masiva de DNA, formando un transpososoma. La técnica se basa en el uso de cromatina nativa (no fijada) que se trata con esta Tn5 modificada. La Tn5 producirá una mayor frecuencia de transposición de los cebadores en las regiones desprovistas de nucleosomas (activas) que en las regiones cubiertas por los mismos. De esta forma esta Tn5 es capaz de fragmentar y marcar al mismo tiempo con inserciones un genoma completo. La librería de fragmentos marcados por los cebadores de la Tn5 es amplificada por PCR y secuenciada (Figura 4c). Además, el uso de Tn5 cargada con secuencias modificadas de cebadores permite combinar varias muestras en un mismo proceso de secuenciación. Otra de las principales ventajas de esta técnica es que es capaz de llevarse a cabo con un número muy limitado de células (entre 500 y 50000 células frente a los varios millones que se necesitan en las técnicas anteriores). Por último, esta técnica produce una gran resolución en la señal, lo que permite incluso identificar los sitios de unión de los distintos factores de transcripción [104].

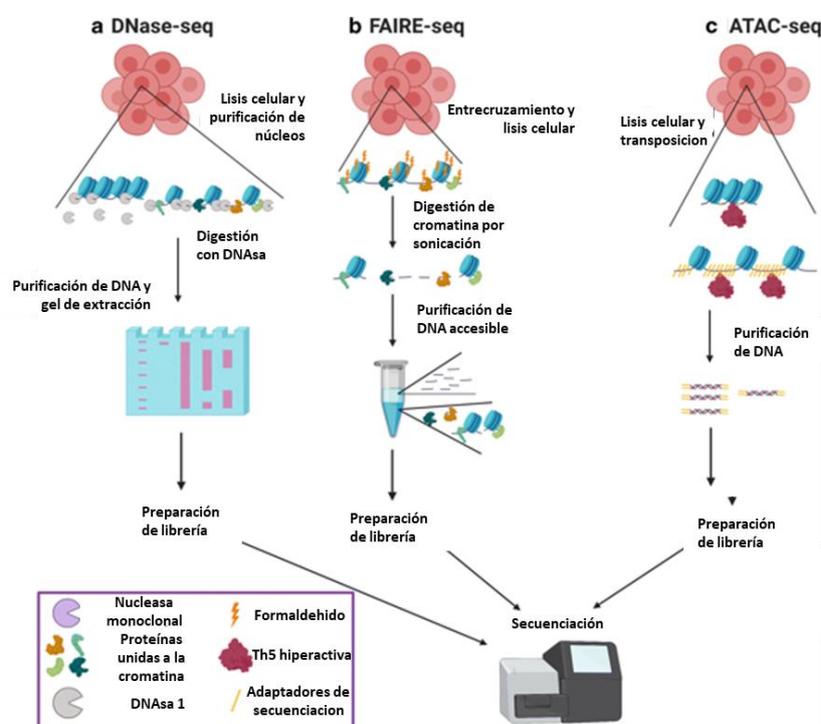


Figura 4. Métodos para mapear la accesibilidad del genoma. A) DNase-seq se basa en la digestión preferencial de regiones de cromatina que no están protegidas por proteínas unidas,

dejando atrás regiones accesibles que se conocen como sitios hipersensibles a DNasa I (DHS). B FAIRE-seq depende de la reticulación de las proteínas que interactúan con la cromatina en el DNA utilizando formaldehído. C. ATAC-seq se basa en la transposasa Tn5 hiperactiva para insertar adaptadores de secuenciación en regiones accesibles del genoma. Después de la transposición, el DNA genómico puede aislarse y amplificarse por PCR, luego someterse a una secuenciación profunda [105]

1.3.3 Métodos de Captura de la Conformación de la Cromatina

Dada la enorme cantidad de elementos enhancers y su capacidad de actuar a largas distancias, es fundamental poder asociar dichos elementos con sus promotores diana y conocer cuál es el mecanismo por el que son capaces de influir en la actividad de promotores que frecuentemente se encuentran a grandes distancias. Estos métodos han permitido estudiar el grado o frecuencia de interacción entre distintas regiones de la cromatina [106]. Los elementos reguladores de un promotor en particular tenderán a estar localizados en regiones de la cromatina que posean una frecuencia de interacción significativa. En 2002 Dekker y cols [107] desarrollaron la tecnología de Captura de la Conformación de la Cromatina (3C). Esta técnica bioquímica permite analizar la frecuencia de contactos entre dos regiones del genoma en una población celular. Desde entonces, se han desarrollado métodos genómicos derivados del 3C que, en comparación con métodos basados en la microscopía, permiten un análisis sistemático de la organización del DNA a una resolución mucho mayor, como lo son los métodos 3C, 4C, 5C, HiC [106].

Los mapas o matrices de contacto (representaciones gráficas de la frecuencia de interacción entre distintos fragmentos de DNA a lo largo de un locus o genoma completo) nos ayudan a entender las reglas que determinan la organización y funcionamiento del genoma en el interior de la célula. Es importante tener en cuenta que las conclusiones que se derivan de estos métodos son probabilistas y no deterministas. Esto es, no se puede inferir la organización o funcionamiento del genoma de una célula individual, sino un comportamiento medio en una población celular.

1.3.4 Purificación por afinidad CRISPR de elementos regulatorios (CAPTURE)

Todas las técnicas mencionadas anteriormente han permitido el entendimiento de la dinámica de regulación de la expresión genética tejido-específica de forma

integral. Sin embargo, una limitante de los métodos actuales es la necesidad de conocer un blanco de interés, ya sea un factor de transcripción o un elemento regulador (promotor o enhancer) así como de la disponibilidad de reactivos comerciales de calidad suficiente para realizar experimentos bioquímicos y genómicos a gran escala [72]. Resulta entonces evidente, la necesidad de estudiar la composición de los complejos reguladores de la expresión de genes de interés biomédico en el contexto muscular, con el objetivo de identificar proteínas, sus modificaciones post-traduccionales, RNAs regulatorios, así como los contactos -intra e intercromosomales que, en su conjunto, proporcionen información sobre nuevos actores en la regulación epigenética que permitan su exploración en diversos contextos de enfermedad.

En este sentido, se ha propuesto un nuevo sistema para el estudio a gran escala de componentes de complejos de cromatina, basado en la tecnología CRISPR-Cas9, denominado purificación por afinidad CRISPR de elementos regulatorios (CAPTURE) [108-109]. El sistema contempla la utilización de un RNA guía (sgRNA) complementario al locus de interés y una dCas9, catalíticamente inactiva la cual contiene un sitio aceptor de biotina. A través de la introducción del sgRNA a las células blanco, éste se posiciona en el locus de interés para reclutar a la proteína dCas9, formando el complejo sgRNA-dCas9-biotinilada. Tras la fragmentación de la cromatina por sonicación, la precipitación de complejos mediante la interacción biotina-estreptavidina permite el aislamiento y el análisis de los componentes en ese locus particular; desde análisis proteómicos para la identificación de péptidos y sus modificaciones post-traduccionales, hasta la identificación de RNAs regulatorios y contactos entre enhancers y promotores [108-109] (Figura 5).

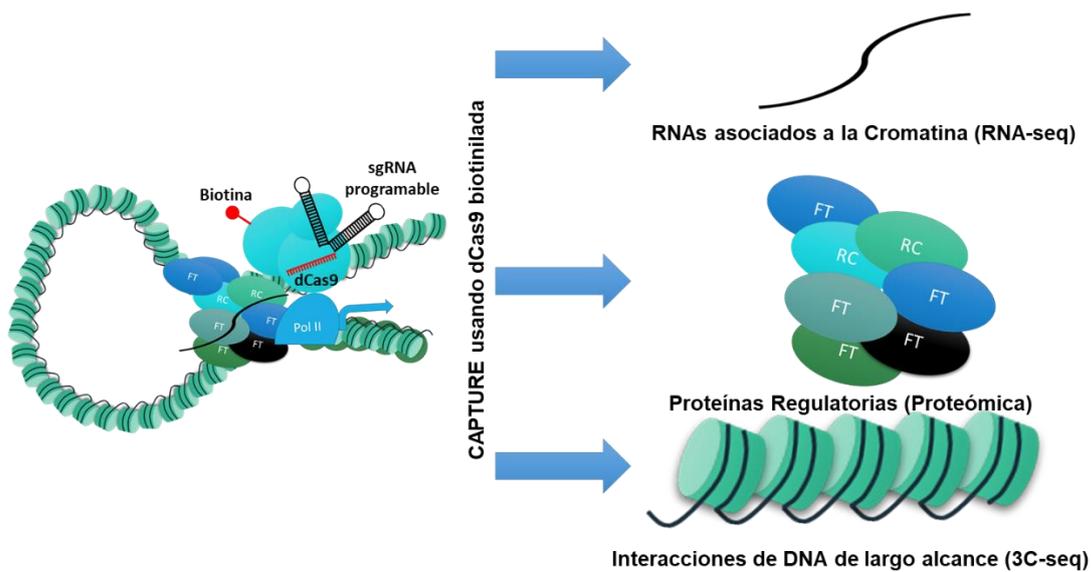


Figura 5. Esquema del método de Purificación por afinidad CRISPR de elementos regulatorios (CAPTURE).

Entre las ventajas del empleo de este sistema, podemos mencionar [108-109]:

- Alta sensibilidad y afinidad entre la biotina y la estreptavidina ($k_d=10^{-14}$ mol/L) superior al uso de anticuerpos.
- Al evitarse el uso de anticuerpos se permite reducir significativamente las interacciones no específicas.
- Debido a la alta estabilidad de la interacción biotina-estreptavidina se permite una rápida purificación y la eliminación de contaminantes.
- Este sistema es adaptable a múltiples blancos mediante la manipulación del sgRNA o a análisis múltiples a un mismo sistema mediante conjugaciones entre dCas9/sgRNA.



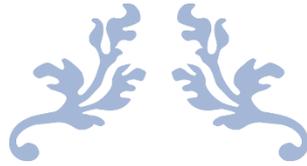
Justificación



Las secuencias regulatorias en *cis* y los factores reguladores en *trans*, controlan el desarrollo y la regeneración muscular mediante la regulación de la expresión; el conocimiento de su existencia, su disposición espacial y su estructuración mediante interacciones de largo alcance en la cromatina, y son importantes dado que puede proveer información valiosa sobre los mecanismos de regulación y su actividad en un locus y una condición de interés médico.

Dada la importancia del conocimiento de dichos factores se hace patente la necesidad de aplicación de nuevas tecnologías, como la propuesta por el sistema CAPTURE, en modelos musculares que permitan el entendimiento de la regulación epigenética en procesos como la embriogénesis, la diferenciación y la regeneración muscular en estados homeostáticos así como su exploración en enfermedades específicas como la distrofia muscular de Duchenne (DMD), la distrofia fascioescapulohumeral (FSHD) y la distrofia miotónica (DM). También, es plausible la exploración de componentes regulatorios en otros contextos como padecimientos metabólicos, envejecimiento o cáncer.

Aunado a lo anterior, la utilización de modelos celulares de músculo, como las células C2C12, se hace necesaria dada su utilidad en el estudio de procesos de proliferación y diferenciación *in vitro*, lo cual permitirá la identificación de complejos regulatorios a lo largo del proceso de diferenciación muscular.



Objetivos



3.1 Objetivo General

Implementar un método basado en el sistema CRISPR/Cas9 con el fin de identificar nuevos complejos asociados a la cromatina presentes en genes involucrados en la regeneración y diferenciación muscular.

3.2 Objetivos Particulares

- Diseñar sgRNAs dirigidos hacia elementos promotores de los genes MyoD, Pax7, y como controles negativos el gen GAL4 y regiones teloméricas.
- Clonar dichos sgRNA en un vector de sobre-expresión y confirmar la integridad de su secuencia mediante secuenciación tipo Sanger.
- Generar líneas celulares C2C12 que expresen de manera estable la proteína dCas9 y el donador de biotina BirA (C2C12-dCas9/BirA).
- Transducir células C2C12-dCas9/BirA con los sgRNAs diseñados para los blancos específicos.
- Comprobar el sistema implementado mediante el análisis de complejos asociados en los genes MyoD y Pax7, mediante western blot, RT-PCR y PCR y determinar su especificidad mediante la comparación con los controles negativos GAL4 y elementos teloméricos.



Hipótesis



El sistema CAPTURE permitirá la identificación de los elementos en *cis* regulatorios y factores en *trans* asociados a loci específicos implicados en la regeneración y diferenciación muscular, en el modelo *in vitro* de diferenciación muscular C2C12.



Metodología



5.1 Creación de Biblioteca de plásmidos

5.1.1 Formación de células calcio competentes

Para cumplir dicho objetivo se realizó el crecimiento de células DH5 α en un mililitro de medio LB por 12 hrs a 37°C y 200rpm; acto seguido se tomó el cultivo y se realizó una dilución 1 a 100 en medio LB incubándose por 4 hrs a 37°C y 200rpm; transcurrido el tiempo se incubaron en hielo durante 20 min. Posteriormente el cultivo se centrifugó a 4000rpm por 10 min; tras decantar el sobrenadante, la pastilla se resuspendió en 20ml de una solución de CaCl₂ 0.1M, se incubó en hielo por 30 min y se centrifugó bajo las condiciones antes descritas. Por último, se decantó el sobrenadante y las bacterias fueron resuspendidas en 5mL de una solución CaCl₂ 0.1M con 15% de glicerol. Las bacterias calcio competentes fueron criopreservadas a -80°C o utilizadas directamente para realizar la transformación.

5.1.2 Transformación por choque térmico de células DH5 α

Para dicho fin se ocuparon los siguientes plásmidos:

- pEF1 α -BirA-V5-neo (#100548-Addgene)
- pEF1 α -FB-Cas9-puro (#100547-Addgene)
- pSLQ1651-sgTelomere(F+E) (#51024)
- pSLQ1651-sgRNA(F+E)-sgGal4 (#100549)
- VSV-G (#8454)
- psPAX2 (#12260)

La transformación consistió en lo siguiente. Se tomaron 100 μ L de Bacterias DH5 α calcio competentes a las cuales se le añadió 200ng del plásmido correspondiente, dejándose incubar en hielo por 30 min. Acto seguido se puso

la mezcla a 42°C por 30 segundos y posteriormente en hielo por 2 min. Tras el choque térmico se añadió 1mL de medio LB dejándose incubar por 90min a 37°C con agitación constante a 200rpm; transcurrido este tiempo se sembraron por extensión en 2 placas de medio LB con ampicilina adicionada como selección (LB+AMP). Las placas fueron incubadas a 37°C por 15 hrs, y revisadas las colonias resistentes a ampicilina como testigo de la transformación.

5.1.3 Miniprep para extracción de DNA plasmídico.

Para dicho efecto se escogieron 4 colonias aisladas para cada plásmido antes mencionado, y se realizó un subcultivo de estas en 4mL de medio LB+AMP a 37°C y 200rpm por 15 hrs. El procedimiento de extracción del DNA plasmídico se realizó utilizando el kit AxyPrep™ Plasmid Miniprep Kit siguiendo las instrucciones provistas por el proveedor. Tras la extracción, la concentración del plásmido fue determinada mediante medición en Nanodrop.

5.1.4 Digestión enzimática para comprobación de transformación.

Con el fin de verificar la identidad de cada plásmido, se realizó una digestión enzimática utilizando las siguientes enzimas:

Plásmido	Enzima 1	Enzima 2	Buffer
pEF1 α -BirA-V5-neo (#100548)	BamHI	XbaI	CutSmart
pEF1 α -FB-Cas9-puro (#100547)	BamHI	XbaI	CutSmart
pSLQ1651-sgTelomere(F+E) (#51024)	BstXI	XhoI	3.1
pSLQ1651-sgRNA(F+E)-sgGal4 (#100549)	BstXI	XhoI	3.1
VSV-G (#8454)	EcoRI		FD
psPAX2 (#12260)	EcoRI		FD

Tabla 1. Descripción de digestión enzimática.

250ng de cada plásmido fueron digeridos con 10 unidades cada enzima a un volumen de 50 μ L de agua MiliQ, dicha reacción se incubó a 37°C durante 2 horas. Posterior al tiempo de reacción se corrió una electroforesis en gel de agarosa 0.8% utilizando bromuro de etidio para verificar el corte esperado por el tamaño de banda.

5.2 Cultivo de celular

El cultivo de las líneas celulares utilizadas durante la realización del trabajo experimental, se realizó bajo las siguientes condiciones:

Línea celular	Medio de cultivo para proliferación	Medio de cultivo en diferenciación
C2C12	DMEM + 10% SFB + 1% PSA	DMEM + 2% SE + 1% PSA
HEK293T	DMEM + 10% SFB + 1% PSA	-----
C2C12-dCas9	DMEM + 10% SFB + 1% PSA + 0.8µg/ml Puromicina	DMEM + 2% SE + 1% PSA + 0.8µg/ml Puromicina
C2C12-dCas9/BirA	DMEM + 10% SFB + 1% PSA + 0.8µg/ml Puromicina + 600µg/ml G418	DMEM + 2% SE + 1% PSA + 0.8µg/ml Puromicina + 600µg/ml G418
C2C12-dCas9/BirA/sgTelo		
C2C12-dCas9/BirA/sgGal4		
C2C12-dCas9/BirA/sgDRR		
C2C12-dCas9/BirA/sgPRR		
C2C12-dCas9/BirA/sgTSS		

Tabla 2. Descripción de medios de cultivo. DMEM: Dulbecco`s Modified Eagle`s Medio con 4.5 g/L de Glucosa y L-Glutamina (Santa Cruz Biotechnology, Inc); SFB: Suero Fetal Bobino (Biowest); PSA: Penicilina-Streptomycin-Amfotericina B (10,000 IU/ml) (MP Biomedicals, LIC).

Los cultivos celulares se mantuvieron en incubación a 37°C en una atmosfera de 5% de CO₂. Los pases celulares fueron realizados a menos de un 80% de confluencia. Para dicho objetivo, fue retirado el medio de cultivo respectivo de las células mediante aspiración, acto seguido las células fueron lavadas utilizando PBS estérily tras la eliminación del PBS fue añadida tripsina (1x Tripsin-EDTA, MP Biomedicals, LIC) hasta cubrir la superficie de la placa e incubadas por 5 min a 37°C; acto seguido las células fueron recuperadas y centrifugadas a 2000 rpm durante 5 min, posteriormente se eliminó el sobrenadante y las células fueron resuspendidas en el medio de cultivo específico antes señalado, y sembradas en las placas correspondientes.

5.3 Ensayos de diferenciación

Para los ensayos de diferenciación 60,000 células fueron sembradas en placas de 6 pozos con el medio de cultivo de proliferación correspondiente, se mantuvieron en incubación a 37°C en una atmosfera de 5% de CO₂ durante 24 horas. Acto seguido se retiró el medio de proliferación y se añadió el medio de cultivo de diferenciación (Detallado en la tabla 1).

5.4 Criopreservación de líneas celulares

Para la realización de dicho proceso, las células fueron cultivadas hasta alcanzar un 80% de confluencia, el medio de cultivo fue retirado por aspiración; el cultivo celular fue lavado utilizando PBS, posteriormente fue retirado el PBS y fue adicionado Tripsina, tras 5 min de incubación las células fueron recuperadas y centrifugadas a 2000 rpm durante 5 min a 37°C, acto seguido, el sobrenadante fue descartado y las células resuspendidas en SFB con un 10% de Dimetil Sulfóxido (JT Baker). Posteriormente las células fueron transferidas a un criovial y preservadas en refrigeración a -80°C.

5.5 Curva de Muerte Celular

Para determinar la concentración mínima necesaria para la selección por antibióticos, 40,000 células C2C12 fueron sembradas en una placa de 6 pozos, y se mantuvieron en incubación toda la noche a condiciones estándar. Posteriormente se retiró el medio de cada pozo, se lavó con PBS y se agregó medio de selección a las siguientes concentraciones: para la selección con puromicina se ocuparon 0.6, 0.8, 1, 1.2 y 1.4 µg/ml; por su parte para la selección con G418 se ocuparon concentraciones de 400, 600, 800, 1000 y 1200 µg/ml, considerando un control en cada caso. Posteriormente las células fueron incubadas durante 4 días y transcurrido este tiempo se prosiguió a la tripsinización y la determinación del número total de células vivas.

5.6 Ensayo de lipofección

La transfección del plásmido pEF1α-FB-Cas9-puro se realizó mediante lipofección en células C2C12 en estado de proliferación. Para dicho objetivo, fueron sembradas 2,000 células/pozo en una placa de 24 pozos y mantenidas con el medio de proliferación correspondiente en incubación a las condiciones antes descritas durante 24 horas; posteriormente fue retirado el medio de cultivo de las células y se agregó DMEM sin suero. A continuación se agregó por goteo la siguiente mezcla: 2µg del plásmido pEF1α-FB-Cas9-puro + 6µl de Lipofectamina 2000 (Invitrogen) + 150µL de DMEM (Mezcla previamente realizada y mantenida en incubación por 30 min a temperatura ambiente). Acto seguido las células fueron incubadas a 37°C en una atmosfera de 5% de CO₂ por 5.5 horas. Transcurrido el tiempo de incubación el medio fue cambiado por el medio de proliferación correspondiente y las células fueron incubadas por 48 horas. Posterior a su incubación el medio de proliferación fue retirado y se añadió

un medio de selección (DMEM + 10% SFB + 1% PSA + 1µg/ml Puromicina) y las células incubadas por un periodo de 96 horas, tras el cual se determinó si la transfección fue eficiente.

Para la generación de la línea celular estable C2C12-dCas9/BirA se siguió el mismo protocolo descrito anteriormente, pero utilizando el plásmido pEF1α-BirA-V5-neo, las células C2C12-dCas9 y DMEM + 10% SFB + 1% PSA + 1µg/ml Puromicina + 800µg/ml G418 como medio de selección.

5.7 Extracción de RNA total

Para la extracción del RNA total, las células correspondientes fueron cultivadas hasta alcanzar un 80% de confluencia en placas de 6 pozos, posteriormente fue removido el medio de cultivo en el que se encontraban, y agregado 1mL de Trizol (ambion) e incubado por 5 min, posteriormente fue homogenizado y centrifugado a 12,000 g por 10 min, acto seguido se recuperó el sobrenadante y se añadieron 200µl de Cloroformo (MERK) se homogenizó, se incubó por 3 min y se centrifugó a 12,000g por 15 min a 4°C; acto seguido se transfirió la fase acuosa a un nuevo tubo y se añadieron 500µl de isopropanol (VWR), se homogenizó, se incubó por 10 min y se centrifugó a 12,000g por 10 min a 4°C. Posteriormente se retiró el sobrenadante y la pastilla fue resuspendida en 500 µl de una solución de etanol al 75% (Sigma) y centrifugado a 12,000g por 5 min a 4°C. Por último, se descartó el sobrenadante y la pastilla se resuspendió en 100µl de Agua libre de nucleasas. Se determinó su concentración e integridad mediante nanodrop y mediante una electroforesis en gel de agarosa al 2%.

5.8 Retro-transcripción y PCR (RT-PCR)

Para la realización de la RT-PCR, primeramente, se procedió a realizar la retro-transcripción de la muestra correspondiente, para tal efecto se realizó una mezcla A que contenía:

- 1.5 µl de Oligo dT (Invitrogen)
- 1 µg de RNA total
- 1 µl de dNTP mix (10nM) (Invitrogen)
- Agua libre de nucleasas hasta ajustar a 12 µl

Esta mezcla fue calentada a 65°C por 5 min, posteriormente se colocó en hielo y se añadió una segunda mezcla que contenía:

- 4 µl de Buffer 5X (Invitrogen)
- 2 µl de DTT (Invitrogen)
- 1 µl de Agua libre de nucleasas

Y se incubó a 37°C por 2 min, posteriormente se añadió 1 µL de la retro-transcriptasa M-MLV (200 U/µl) y se incubó a 37°C por 50 min y después a 70°C por 10 min.

Tras la retro-transcripción se procedió a la realización de la amplificación por PCR. Para esto se preparó la siguiente mezcla:

- 2.5 µl del Buffer para PCR 10X (Invitrogen)
- 0.75 µl de MgCl₂ (50mM) (Invitrogen)
- 0.1 µl de Taq DNA polimerasa (5 U/µl) (Invitrogen)
- 0.5 µl dNTPs (10 mM) (Invitrogen)
- 0.75 µl de Primer Forward (10 µM) (OligoT4)¹
- 0.75 µl del Primer Reverse (10 µM) (Oligo T4)¹
- 2 µl de cDNA obtenido de la retrotranscripción
- 17.5 µl de Agua libre de nucleasas

Dicha mezcla se sometió a una incubación a 94°C por 2 min, posteriormente a 30 ciclos de:

- 94°C por 45 segundos
- 62°C por 35 segundos
- 72°C por 90 segundos

Posteriormente se sometió a 72°C por 7 minutos para realizar la extensión final. Los productos de PCR fueron verificados mediante electroforesis en gel de agarosa a 1.5%.

¹ La lista completa de los primers ocupados se encuentra en la tabla 1 de la sección de anexos.

5.9 Diseño de sgRNAs

Para la obtención de posibles sgRNA se ocupó la herramienta bioinformática CRISPR Finder del Sanger Institute Genome Editing (<https://wge.stemcell.sanger.ac.uk/>) [110] considerando las regiones PRE y Promotora para el locus del gen Pax7 y DRR, PRR y cercanas al sitio de inicio de transcripción para el locus del gen MyoD, del genoma de ratón NCBI37/mm9; para tal caso se obtuvieron determinados sgRNAs que flanquearan dichas regiones antes mencionadas, cuidando que la predicción de efectos fuera de blanco fueran los mínimos para todos los casos.

5.10 Clonación de sgRNA

Para la clonación de los respectivos sgRNA, se realizó una reacción de PCR considerando la siguiente mezcla:

- 5 µl del Buffer para PCR 10X (Invitrogen)
- 0.2 µl de Taq DNA polimerasa (5 U/µl) (Invitrogen)
- 1 µl dNTPs (10 mM) (Invitrogen)
- 2 µl de Primer sgRNA (10 µM)(OligoT4)²
- 2 µl del Primer Reverse universal (10 µM) (Oligo T4)²
- 1 µl pSLQ1651-sgRNA(F+E)-sgGal4 (1ng)
- Agua libre de nucleasas hasta aforar 50 µL

Se utilizó el siguiente programa de amplificación:

- 1 ciclo de 2 min a 98°C
- 30 ciclos:
 - 10 segundos a 95°C
 - 15 segundos a 60°C
 - 30 segundos a 72°C
- 1 ciclo de 2 min a 72°C
- Mantenimiento a 4°C

Tras la amplificación el producto de PCR se corrió en una electroforesis en gel de agarosa a 1.5 %. El producto de PCR fue purificado utilizando el kit Qiagen

² La lista completa de los primers ocupados para la clonación de los sgRNAs específicos se encuentra en la tabla 1 de la sección de anexos.

PCR purification kit siguiendo las indicaciones de fábrica, y resuspendido en 15 µl de buffer de elución. Posterior a la purificación tanto el plásmido pSLQ1651-sgRNA(F+E)-sgGal4 como el producto de PCR fueron restringidos enzimáticamente utilizando las enzimas BstXI (NEB) y XhoI (NEB) en el buffer 3.1 (NEB) incubando a 37°C por toda la noche. El producto de PCR digerido fue purificado utilizando el kit MinElute PCR purification y siguiendo las indicaciones del fabricante. Por su parte el plásmido pSLQ1651-sgRNA(F+E)-sgGal4 digerido fue separado mediante una electroforesis en gel de agarosa al 0.8% y el esqueleto carbonatado purificado. Posteriormente se realizó una ligación del plásmido pSLQ1651-sgRNA(F+E)-sgGal4 con el amplicón de PCR utilizando la ligasa T4 (NEB) y su respectivo Buffer de ligación. Tras la ligación fueron transformadas bacterias DH5α calcio competentes como ya ha sido descrito, fue extraído el DNA plasmídico y comprobada la clonación mediante secuenciación tipo Sanger.

5.11 Transducción de sgRNA

Para tal efecto fueron sembradas 50,000 células HEK293T en una placa de 6 pozos y se incubaron con el medio de proliferación correspondiente a 37°C en una atmósfera con 5% de CO₂ por 24 horas. Posteriormente se retiró el medio de cultivo y se agregó una mezcla con los siguientes componentes:

- 0.5 µg del plásmido VSV-G
- 1 µg del plásmido psPAX2
- 1.5 µg del plásmido pSLQ1651-sgRNA
- 6 µl de Lipofectamina 2000 (Invitrogen)
- Aforando a 250 µl con medio DMEM

Se dejó incubando por 5.5 horas a 37°C en una atmósfera con 5% de CO₂, posteriormente se aspiró el medio y se agregó medio de proliferación, antes descrito y se incubó por 48 horas a 37°C en una atmósfera con 5% de CO₂. Transcurrido este tiempo se recuperó el medio de cultivo, se filtró ocupando un filtro de 0.45µm y se agregó al medio Polybrene (10µg/mL), incubándose por 10 minutos. Posteriormente la mezcla fue añadida como medio sobre células C2C12-dCas9/BirA previamente crecidas a un 50% de confluencia. Las células

fueron incubadas a 37°C en una atmosfera con 5% de CO₂ por 24 horas, se aspiró el medio con viriones y se agregó medio de proliferación.

5.12 Análisis comprobatorios del sistema.

Para comprobar la eficiencia y especificidad del CAPTURE, las células que expresan establemente los blancos deseados, así como el control positivo para Telómeros y el control negativo dirigido a GAL4, son entrecruzadas mediante formaldehído, seguido por la digestión de la cromatina mediante sonicación, y la purificaron las regiones de interés mediante perlas magnéticas con estreptavidina. La comprobación de la composición de los complejos regulatorios asociados a los loci de interés es detallada a continuación.

5.12.1 Comprobación de especificidad de blanco mediante PCR

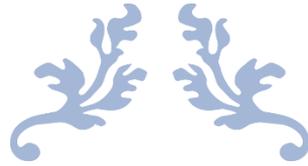
Primeramente, se determina la especificidad del blanco al cual está dirigido el RNA guía. Para tal objetivo se cosechan células C2C12-dCas9/BirA/sgRNA, de las cuales se recuperan los núcleos y se extrae la cromatina mediante lisis nuclear. Posterior a esto se fragmenta la cromatina y se recupera mediante perlas magnéticas acopladas a estreptavidina. Posteriormente la cromatina capturada es sometida a un tratamiento con RNAsa y Proteinasa K para recuperar solamente el DNA. El DNA purificado se ocupa en una reacción de PCR hacia las regiones DRR, PRR y Promotoras del locus de MyoD ocupando los primers descritos en la tabla 1 de la sección de anexos y la metodología antes mencionada para PCR.

5.12.2 Comprobación de obtención de proteínas asociadas al locus de interés

En segundo término, se determina la especificidad del blanco comprobando la presencia de proteínas asociadas a cada locus. Para tal objetivo se cosechan células C2C12-dCas9/BirA/sgRNA, de las cuales se recuperan los núcleos y se extrae la cromatina mediante lisis nuclear. Posterior a esto se fragmenta la cromatina y se recupera mediante perlas magnéticas funcionalizadas con estreptavidina. Posteriormente la cromatina capturada es sometida a un procedimiento de purificación de proteínas. Las proteínas son utilizadas para realizar un ensayo de westernblot dirigido hacia MyoD para comprobar las proteínas de los loci DRR, PRR y TSS y de TERF2 para el locus telomérico.

5.12.3 Comprobación de identificación de RNAs asociados al locus de interés

En último término, se determina la especificidad del blanco al cual está dirigido el RNA guía mediante la determinación de la presencia de RNAs asociados. Para tal objetivo se cosechan células C2C12-dCas9/BirA/sgRNA, de las cuales se recuperan los núcleos y se extrae la cromatina mediante lisis nuclear. Posterior a esto se fragmenta la cromatina y se recupera mediante perlas magnéticas funcionalizadas con estreptavidina. Posteriormente la cromatina capturada es sometida a un tratamiento con DNAsa y Proteinasa K 20 para recuperar solamente el RNA. El RNA purificado se ocupa para realizar una reacción de rtPCR hacia los enhancer RNA: Ce^{RNA} y DRR^{RNA} y Promotoras ocupando los primers descritos en la tabla uno de la sección de anexos y la metodología antes mencionada para rtPCR



Resultados



6.1 Establecimiento de una biblioteca de plásmidos

Se logró establecer una biblioteca de plásmidos contenida en bacterias DH5 α bajo la metodología anteriormente descrita. Los plásmidos contemplados en la biblioteca de plásmidos son: pEF1a-FB-dCas9-Puro el cual contiene la región codificante para la enzima dCas9, pEF1a-BirA-V5-neo que tiene como inserto a la biotín-ligasa BirA, Pslq1651-sgRNA(F+E)-sgGAL4 y Pslq1651-sgTelomere (F+E) que codifican para los sgRNAs dirigidos hacia Gal4 (control negativo) y hacia regiones teloméricas (control positivo) y los plásmidos psPAX2 y N2c/VSVg que son plásmidos empaquetadores para la formación de viriones. Con el objetivo de poder confirmar la identidad de cada plásmido, tras la extracción del DNA plasmídico de las bacterias transformadas se realizó una caracterización por restricción; obteniendo el patrón de digestión esperado.

Para el caso del plásmido pEF1a-FB-dCas9-Puro se digirió con las enzimas XbaI y BamHI que realizan dos cortes en el plásmido, produciendo la formación de dos fragmentos; uno de 7501pb y otro de 4178pb que corresponde al inserto codificante para la proteína dCas9 (Figura 6a). De igual forma el plásmido pEF1a-BirA-V5-neo fue digerido con las enzimas XbaI y BamHI, las cuales permiten la formación de dos fragmentos; uno de 6112pb y otro de 976pb, este último corresponde a la región codificante para la proteína BirA (Figura 6b).

Por otra parte, los plásmidos Pslq1651-sgRNA(F+E)-sgGAL4 y Pslq1651-sgTelomere (F+E) fueron digeridos con las enzimas BstXI y XhoI, que libera un fragmento de 118pb dejando el esqueleto carbonatado de 8190pb (Figuras 6c y 6d).

Por otra parte, los plásmidos psPAX2 y N2c/VSVg, necesarios para la generación de partículas lentivirales fueron digeridos con la enzima EcoRI que

realiza dos cortes en cada plásmido; para el caso del plásmido psPAX2 se formaron dos fragmentos, uno de 6323pb y otro de 4370pb (Figura 6e), por su parte la digestión del plásmido N2c/VSVg da lugar a un fragmento de 5689pb y otro de 1574pb (Figura 6f).

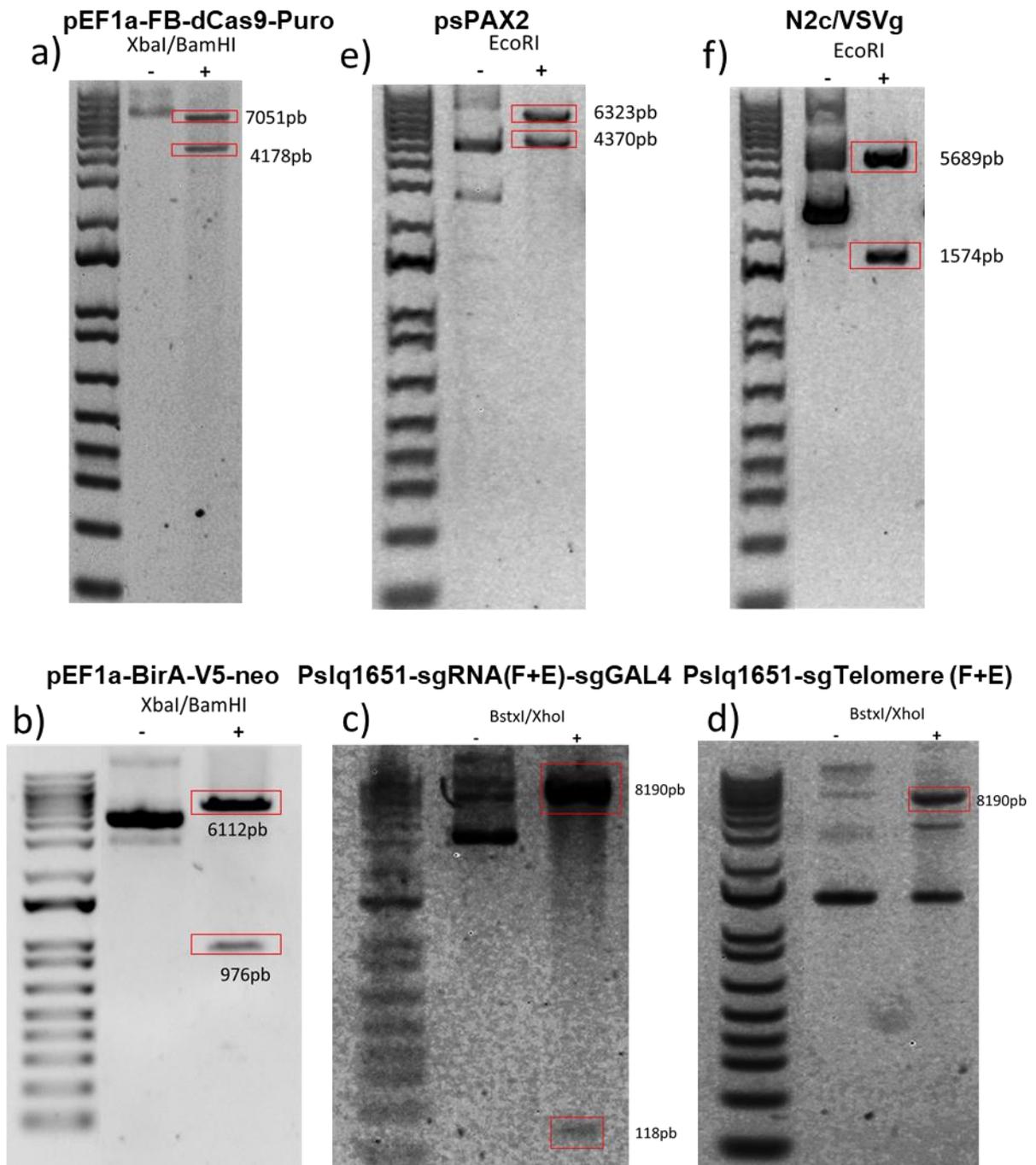


Figura 6: Geles de electroforesis en gel de agarosa al 0.8%. **a)** pEF1a-FB-dCas9-Puro, **b)** pEF1a-BirA-V5-neo, **c)** Pslq1651-sgRNA(F+E)-sgGAL4, **d)** Pslq1651-sgTelomere (F+E), **e)** psPAX2 y **f)** N2c/VSVg. Cada gel está organizado de la siguiente manera: carril 1 corresponde a

la marca de peso molecular, carril 2 es el plásmido en estado nativo y carril 3 es el plásmido digerido con enzimas de restricción.

Con esto fue concluido el primer objetivo particular del presente trabajo: Generar una biblioteca de plásmidos de: pEF1 α -BirA-V5-neo, pEF1 α -FB-Cas9-puro, pSLQ1651-sgTelomere(F+E), pSLQ1651-sgRNA(F+E)-sgGal4, VSV-G y psPAX2.

6.2 Generación de la línea celular estable C2C12-dCas9/BirA

Para la implementación del Sistema CAPTURE en la línea celular C2C12, se realizó la transfección por lipofección de los plásmidos pEF1 α -FB-dCas9-Puro y pEF1 α -BirA-V5-neo. El plásmido pEF1 α -FB-dCas9-Puro que codifica para la proteína dCas9 contiene un gen de resistencia a puromicina, mientras que el plásmido pEF1 α -BirA-V5-neo que codifica para BirA contiene un gen de resistencia a G418; ambos genes de resistencia sirvieron como selección para obtener las células que expresaran establemente lo codificado en cada plásmido. En este sentido, se determinó previo a la lipofección la dosis mínima letal a 96 horas para las células C2C12 mediante un ensayo de muerte celular, que tuvo el comportamiento mostrado en la figura 4 determinando las dosis utilizadas para selección serían 1 μ g/ml de Puromicina y 800 μ g/ml de G418 (Fig. 7).

Determinación de dosis mínima letal en células C2C12

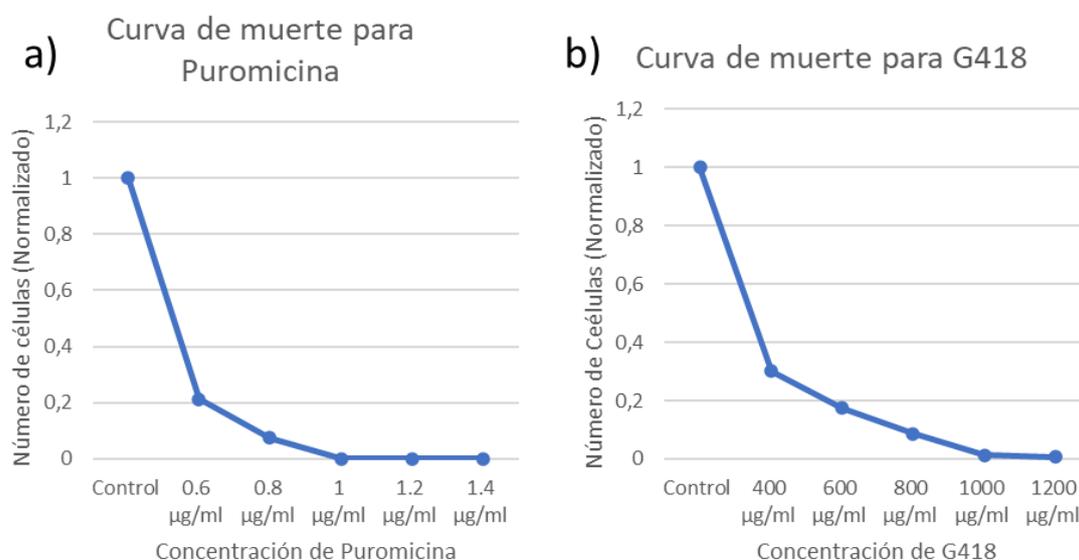


Figura 7: Curva de muerte celular para la determinación de la dosis mínima letal a 96 horas para **a)** puromicina como para **b)** G418

Posteriormente se realizó la lipofección de células C2C12 con el plásmido pEF1a-FB-dCas9-Puro como se describe en la metodología, generando la línea celular C2C12-dCas9. Posteriormente se realizó un ensayo de diferenciación de la línea C2C12-dCas9 tomando como control la línea C2C12 sin transfectar, para determinar que a pesar de la transfección las células mantienen su identidad y su capacidad de diferenciación a miotubos.

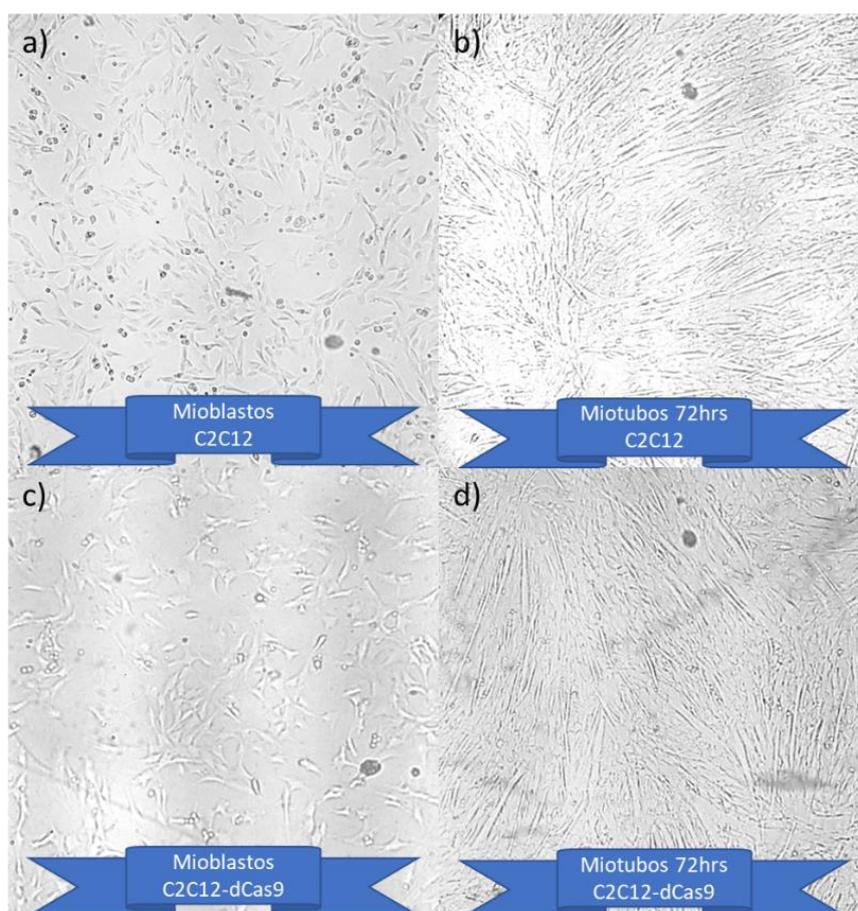


Figura 8. Prueba de diferenciación de las líneas C2C12 y C2C12-dCas9. A) mioblastos C2C12, b) miotubos con 72 horas de diferenciación de C2C12, c) mioblastos C2C12-dCas9 y d) miotubos C2C12-dCas9 a las 72 horas de diferenciación.

Tras una inspección morfológica (Figura 8) acorde a un proceso de diferenciación normal se procedió a la extracción del RNA total como se describe en la metodología y a la realización de la determinación de la expresión de genes de diferenciación, mediante RT-PCR. En la figura 9 se muestra un gel de agarosa al 1.5%, de los productos de PCR para los RNAm de los genes de Miogenina y de la cadena pesada de Miosina (MyHC); los cuales solo amplificaron en las

condiciones de diferenciación, tanto de C2C12 como de C2C12-dCas9; teniendo como control la amplificación del mRNA correspondiente a la proteína ribosomal s18 (RPS18), de expresión constitutiva.

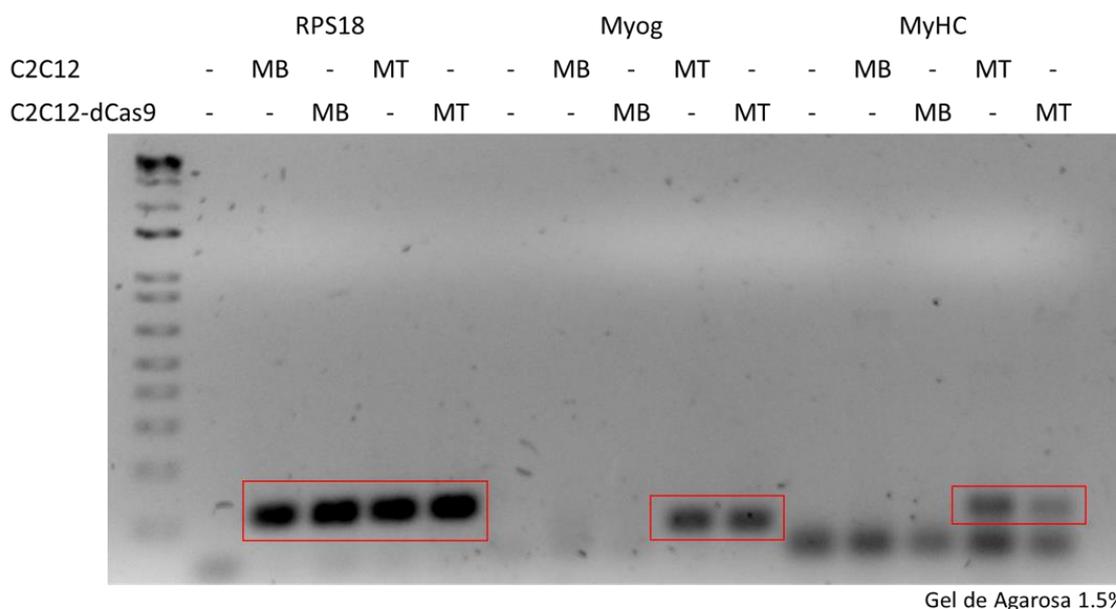


Figura 9. Gel de Agarosa de productos de rtPCR para los mRNA de RSP18, Miogenina y la cadena pesada de Miosina; de células C2C12 en estado de mioblastos (MB) y en estado de miotubos (MT) y para células C2C12-dCas9 en las mismas condiciones.

Posteriormente se tomó la línea celular estable C2C12-dCas9 para ser lipofectada con el plásmido pEF1a-BirA-V5-neo, generando la línea celular C2C12-dCas9/BirA, que expresa establemente tanto la proteína dCas9 como la biotín-ligasa BirA. Tras la obtención de dicha línea, se realizó un ensayo de diferenciación para confirmar que la línea celular mantiene su identidad y su habilidad de diferenciarse a miotubos, como se muestra en la figura 10.

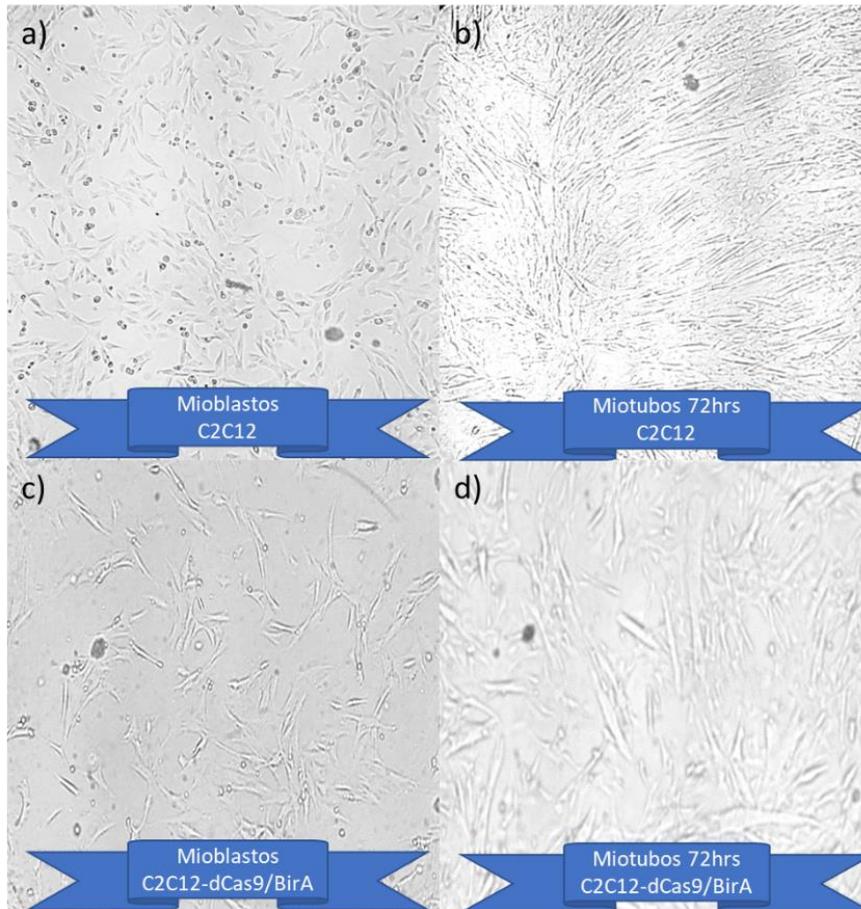


Figura 10. Prueba de diferenciación de las líneas C2C12 y C2C12-dCas9/BirA. A) mioblastos C2C12, b) miotubos con 72 horas de diferenciación de C2C12, c) mioblastos C2C12-dCas9/BirA y d) miotubos C2C12-dCas9/BirA a las 72 horas de diferenciación.

Así mismo, se determinó la expresión de los genes específicos de diferenciación muscular miogenina y MyHC mediante RT-PCR, los cuales amplificaron únicamente en los estadios de miotubos, tanto para las células C2C12 como para las células C2C12-dCas9/BirA. Adicionalmente se determinó la expresión del gen MyoD, característico de la identidad muscular y el cual se expresó en todas las condiciones analizadas (Fig. 11).

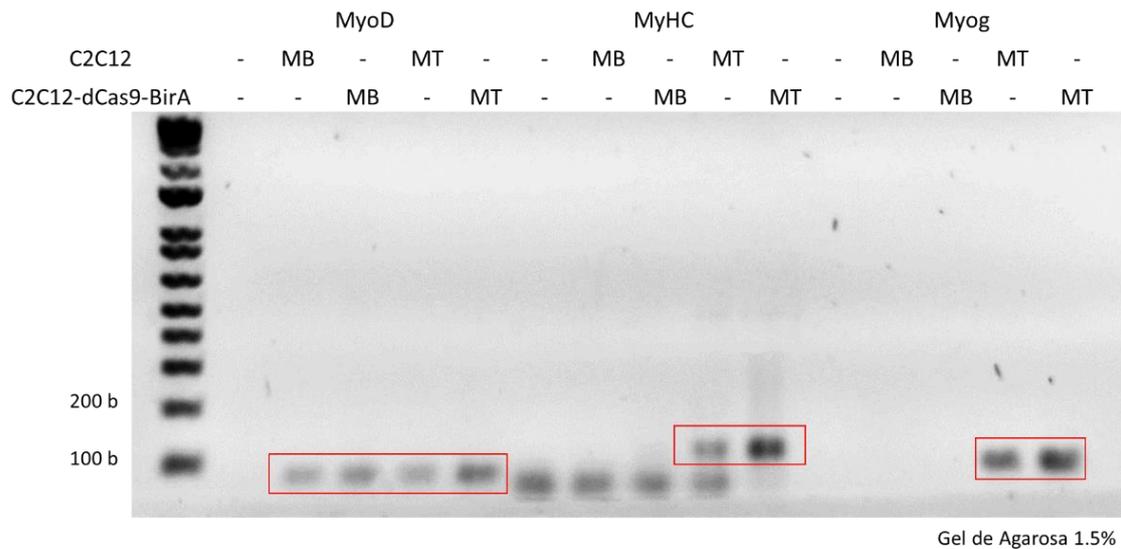


Figura 11. Gel de Agarosa de productos de rtPCR para los mRNA de MyoD, Miogenina y la cadena pesada de Miosina; de células C2C12 en estado de mioblastos (MB) y en estado de miotubos (MT) y para células C2C12-dCas9/BirA en las mismas condiciones.

Para comprobar la efectiva transfección y expresión estable de los elementos que conforman el sistema CAPTURE (dCas9 y BirA), se realizó un ensayo de RT-PCR para detectar la expresión de los mRNAs de dCas9, así como de la biotín-ligasa BirA. Se detectó su amplificación específica en las células C2C12-dCas9/BirA, mostrando que los genes insertados en las células mediante los plásmidos están siendo transcritos; así mismo se tomó como control positivo el mRNA de la proteína ribosomal s18 en el que vemos una amplificación tanto en las células C2C12 como en las C2C12-dCas9/BirA (Figura 12)

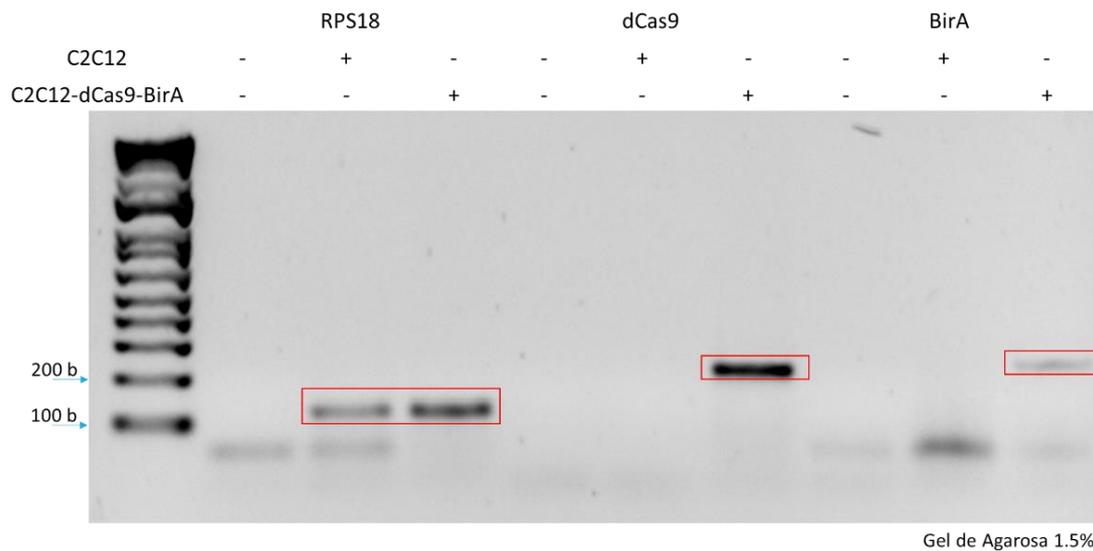


Figura 12. Gel de Agarosa de productos de rtPCR para los mRNA de RPS18, dCas9 y BirA; de células C2C12 y células C2C12-dCas9/BirA.

Con estos ensayos abordamos y concluimos el objetivo particular número 2, que enuncia: Generar líneas celulares C2C12 que expresen de manera estable la proteína dCas9 y el donador de biotina BirA (C2C12-dCas9/BirA).

6.3 Diseño de sgRNAs

El sistema CAPTURE contempla el posicionamiento de una proteína dCas9 mediante sgRNAs programables al locus de interés. En este sentido y bajo los objetivos propuestos para el presente trabajo se obtuvieron sgRNAs utilizando CRISPR Finder del Sanger Institute Genome Editing (WGE), para las regiones PRE y Promotora para el locus del gen Pax7; y DRR, PRR y cercanas al sitio de inicio de transcripción (TSS) para el locus del gen MyoD que en extenso son mostradas en la tabla suplementaria 2.

La herramienta CRISPR Finder ha sido utilizada en reiteradas ocasiones para el diseño de sgRNAs con blancos muy variados [111-113]. Esta herramienta guía la búsqueda de sgRNAs utilizando Genoverse, un navegador genómico rápido, abierto y extensible; de esta manera se visualizan los sitios CRISPR individuales y alineados con el genoma de ratón (en el caso específico al cual se aplicó) [110]. Así mismo se calculan los potenciales fuera del objetivo de todo el genoma de todos los sitios CRISPR dentro del genoma de ratón y los exomas humanos,

mediante un escaneo cada 23 bases del genoma de referencia, buscando un CC como las dos primeras bases (que indica un sitio PAM en la cadena inversa) o un GG como las dos bases finales (sitio PAM en la cadena líder). El potencial fuera del objetivo se encuentra al comparar directamente la secuencia CRISPR con todas las demás coincidencias posibles en el genoma, con hasta 4 pb de desajuste [110]. De esta manera se busca asegurar la especificidad del blanco a ser flanqueado mediante CAPTURE.

De los candidatos propuestos se delimitaron a 3 sgRNAs con las características que se muestran en la tabla 3. Estos candidatos fueron sintetizados en un oligo sentido necesario para la clonación en el vector PSLQ como se define en la sección metodológica, que contemplan un sitio de corte para la enzima de restricción BstXI (CCANNNNNNTGG) y la región de reconocimiento (subrayada).

ID	Secuencia	Distancia a TSS de MyoD	Blanco
MyoD- PRR- sgRNA	ggagaaCCACCTTGTGG <u>TAGCCAAGT</u> GCTACCGCGTA GTTTAAGAGCTATGCTGGAAACAGCA	242 bases rio arriba	región proximal regulatoria (PRR)
MyoD- DRR- sgRNA	ggagaaCCACCTTGTGG <u>GAGACATCCCAAAAAT</u> GACTT GTTTAAGAGCTATGCTGGAAACAGCA	3903 bases rio arriba	región distal regulatoria (DRR)
MyoD- TSS- sgRNA	ggagaaCCACCTTGTGG <u>TTAAGACGACTCTCACGGCT</u> GTTTAAGAGCTATGCTGGAAACAGCA	127 bases rio abajo	TSS de MyoD

Tabla 3. SgRNAs diseñados para establecimiento del sistema CAPTURE.



Discusión



En la actualidad se ha logrado un entendimiento de la dinámica de la regulación de la expresión génica mediante la identificación de elementos regulatorios, lo cual ha sido posible gracias a la implementación de métodos basados en la inmunoprecipitación de la cromatina, en el análisis de accesibilidad a la cromatina y en la captura de la Conformación de la Cromatina; como ha sido revisado en la sección introductoria.

De igual forma se ha mencionado el gran potencial de la técnica CAPTURE en correspondencia con las técnicas antes mencionadas, y se ha determinado que su mayor ventaja, ante estas, es la posibilidad de la determinación locus específica de los elementos regulatorios y la no necesidad de conocer ningún factor directamente implicado en el locus de interés; reflejando que solo el conocimiento de la secuencia es suficiente para el aislamiento, purificación y determinación de los elementos regulatorios del locus de interés.

En este sentido, Michiel Vermeulen y Jérôme Déjardin [114], nos mencionan que en las pasadas cuatro décadas poco menos de treinta artículos han sido publicados refiriendo a técnicas que permitan la purificación y caracterización de un locus genómico específico. De estas, resalta el desarrollo de estrategias que utilizan la secuencia de DNA única de un locus como una característica discriminatoria para la purificación específica del locus. Estas técnicas podrían ser agrupadas en: hibridación directa de sondas de captura a DNA cromatinizado, purificación basada en la unión de proteínas específicas de secuencia, enfoques de etiquetado de proximidad y purificación basada en la nucleasa Cas9 inactiva.

La hibridación directa de sondas de captura a DNA cromatinizado se basa en la especificidad de la secuencia de DNA para el aislamiento de cromatina [115]. La

proteómica de los segmentos de cromatina aislados (PICh) fue el primer método informado que combina la purificación de locus y la espectroscopia de masas. PICH no requiere conocimiento previo sobre las proteínas que se unen al locus de interés o ingeniería genética de las células diana y se basa en la hibridación de sondas sintéticas de ácido nucleico (ácidos nucleicos "bloqueados" y RNA modificado) que son complementarios al locus de interés. Estas sondas se unen al DNA con una especificidad y estabilidad muy altas y albergan el análogo de biotina, que puede captularse en perlas conjugadas con estreptavidina y eluirse con biotina. Esto permite la liberación específica de híbridos sonda-cromatina en solución, mientras que una cantidad significativa de cromatina no diana permanece unida a la matriz de estreptavidina [115-116]. Sin embargo, a diferencia de CAPTURE sus limitaciones son: la dificultad de obtener sondas PICH, las sondas de captura PICH solo funcionan en el extremo 3' del fragmento diana, se requiere la desnaturalización del dsDNA objetivo lo que no es provisto por las sondas, y solo proporciona un factor de enriquecimiento de hasta 10,000 veces, por lo que no es apropiado para estudiar la composición de loci pequeños de copia única de genomas de mamíferos.

Por su parte la purificación basada en la unión de proteínas específicas de secuencia [116], considera que el locus de interés es lo suficientemente abundante y se une significativamente a proteínas que no se encuentran en otras partes del genoma. Por ejemplo, los dominios efectores similares a los activadores de la transcripción (TALE) se han diseñado para unirse a secuencias de DNA específicas, y se etiquetan para inmunoprecipitarlas enriqueciendo el segmento de cromatina asociado [117]. Otra estrategia implica una ingeniería genética significativa de las células diana en donde el locus de interés se modifica para albergar secuencias de DNA específicas que pueden estar estrechamente unidas por proteínas adaptadoras, como Tet [118] o LexA [119]. Después de la inserción de estas secuencias específicas en el locus de interés y la co-expresión de la proteína adaptadora, el locus se enriquece por inmunoprecipitación. Las deficiencias asociadas a estas técnicas con respecto a la técnica CAPTURE es la necesidad de generar proteínas recombinantes específicas para cada locus de interés, o en su segundo caso, la inserción de secuencias específicas para el aislamiento. Esto, además de dificultar la técnica,

puede inducir la pérdida de componentes del complejo regulador asociados al locus de interés importantes para el estudio.

Para las técnicas de etiquetado por proximidad, uno de los esfuerzos más representativos considera una proteína BirA mutante que tiene actividad promiscua y puede biotinar cualquier lisina en las proteínas circundantes. Esta biotina se fusiona con una dCas9, y luego la dCas9 es guiada por RNA guía específico. La biotinización de los factores cercanos permite su captura con estreptavidina para su posterior identificación, por lo que no hay necesidad de estrategias específicas para mantener interacciones proteicas relevantes en las preparaciones de cromatina *in vitro* [120]. Sin embargo, en comparación con el sistema CAPTURE, esta técnica no permite el análisis locus-específico de interacción de la cromatina o de regulación mediada por RNA, así mismo esta técnica se limita a un espacio considerablemente pequeño dado que solo da información de proteínas que se encuentren en la cercanía inmediata al sitio de unión de la dCas9; además, en la mayoría de las moléculas dCas9 expresadas no están asociadas con el locus objetivo, y algunas incluso están significativamente unidas a objetivos no específicos, lo que implica que la gran mayoría del proteoma biotinilado es inespecífico y genera considerable fondo.

La última clase de técnicas son las que emplean la purificación basada en nucleasa Cas9 inactiva, y es la estrategia seguida por la técnica CAPTURE implementada en esta tesis. En los intentos iniciales se utilizó dCas9 guiada al objetivo y purificada por afinidad basada en un solo anticuerpo, sin embargo, en ese ensayo no se evaluó la especificidad a escala genómica y la relevancia biológica de las proteínas que interactúan [121]. De hecho, un dCas9 guiado por un solo locus puede unirse a decenas a cientos de secuencias fuera del objetivo, y esto podría conducir a la co-purificación de cantidades relativamente grandes de regiones irrelevantes. Este problema se resolvió en el método CAPTURE que emplea el uso de una proteína dCas9 biotinilable para purificar loci de cromatina específicos, exhibiendo ventajas como: una alta sensibilidad por la afinidad entre la biotina y la estreptavidina superior al uso de anticuerpos, lo que permite reducir significativamente las interacciones no específicas; debido a la alta estabilidad de la interacción biotina-estreptavidina se permite una rápida purificación y la eliminación de contaminantes. Dado el uso de un RNA guía programable es

posible la determinación de elementos asociados a un locus específico sin el conocimiento de factores unidos en él, además, esta técnica permite la purificación y la determinación no solo proteómica del locus de interés sino de igual forma la determinación de interacciones de largo alcance entre elementos de regulación génica a distancia y la determinación de RNAs no codificantes implicados en la regulación del locus específico [108-109]. Por su parte, estudios de proteómica pueden determinar la presencia de remodeladores de la cromatina, modificaciones postraduccionales a elementos regulatorios o a factores de la transcripción, así como marcas específicas de histonas que induzcan la elucidación de posibles vías de señalización implicadas en la regulación transcripcional específica del locus de interés.

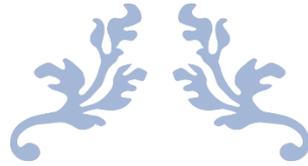
En este sentido, aquí se han reportado avances en la implementación de esta técnica en células C2C12 para la identificación de loci involucrados en la regeneración y diferenciación muscular, en específico hacia las regiones regulatorias cercanas a MyoD y Pax7, con la perspectiva de su futura aplicación en el conocimiento de nuevos componentes de complejos regulatorios asociados a loci de interés biomédico. La implementación de esta técnica requiere la obtención de líneas celulares estables que expresen tanto la biotin-ligasa BirA como dCas9 las cuales han sido generadas como se muestra en los resultados y consta por la presencia específica de los mRNAs respectivos. Por su parte la generación de estas células estables puede inducir la pérdida de la identidad celular o defectos en su proliferación y diferenciación, por lo que mediante los ensayos de diferenciación de las células C2C12-dCas9/BirA fue realizada y comprobada por la expresión de genes de identidad células (MyoD) y los genes de diferenciación Miogenina y MyHC (cadena pesada de Miosina).

Considerando que la variable que determina la implementación eficaz de CAPTURE en células musculares está completamente determinada por la especificidad de los sgRNAs; como se muestra en los resultados, han sido diseñados una serie sgRNAs cuidando que la predicción de efectos fuera de blanco fueran los mínimos para todos los casos, y considerando su unión en los límites de las regiones de interés, de manera que se pueda garantizar la unión específica del complejo.

Además, se han considerado las desventajas que tiene la implementación de la técnica CAPTURE. Dado que dCas9 es una molécula bastante grande, su eficiencia de unión podría verse muy influenciada por la estructura de la cromatina; en donde una cromatina inaccesible puede reducir e inhibir la unión de dCas9, por lo que la determinación del estado en el que se encuentra la cromatina basado en reportes anteriores es fundamental para asegurar la unión [122]. En este sentido los loci propuestos en el presente trabajo, las regiones reguladoras de MyoD y Pax7, son regiones activas y características de células musculares, por lo que se mantienen en conformaciones activas en las células C2C12 empleadas en este ensayo; así se asegura una eficiencia mayor de la unión de dCas9.

Además, la unión de dCas9 podría competir con la unión de complejos biológicamente relevantes [122]. Por lo que atendiendo esta posibilidad los sgRNAs han sido diseñados para no unirse directamente a las regiones reguladoras sino en sus límites y cercanía, tanto río arriba como río debajo de la región de interés, con lo que prevenimos la posibilidad de que la unión de dCas9 intervenga en la unión de factores regulatorios específicos.

Si bien el correcto funcionamiento de CAPTURE en células C2C12 será comprobada hasta la obtención de los complejos regulatorios y la determinación de elementos reportados en la literatura que interaccionan con las regiones regulatorias de nuestro interés, lo presentado en esta tesis representa un avance sustancial en la implementación de sistema CAPTURE.



Perspectivas futuras

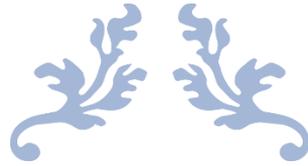


Considerando el grado de avance en el trabajo presentado, como perspectivas se plantea la comprobación del sistema, tal como se muestra en la sección de metodología. En este sentido, la clonación de los sgRNAs diseñados bajo la estrategia metodológica presentada se encuentra en desarrollo, así como la transducción de los sgRNAs a las células C2C12-dCas9/BirA. Con ello, se procederá a la selección y enriquecimiento de células positivas por la expresión de mCherry como gen reportero, mediante citometría de flujo.

Tras la fijación digestión y recuperación de los complejos asociados a la cromatina se procederá a la realización de los análisis comprobatorios, a saber; mediante PCR se analizará la especificidad de la técnica por la región a la cual fue dirigida, utilizando primers que flanqueen las regiones DRR, PRR y la región cercana al sitio de inicio de la transcripción de MyoD; por su parte, a partir de la purificación de RNA total y mediante RT-PCR se comprobará la presencia de RNAs regulatorios asociados con la expresión de MyoD, por ejemplo los enhancers-RNAs ceRNA y DRR-RNA [58,73]; por ultimo mediante western blot dirigido contra MyoD, se comprobará la presencia de esta proteína en las regiones de nuestro interés.

En este sentido se podrá lograr el establecimiento del sistema CAPTURE en células C2C12 dirigido a regiones implicadas en la diferenciación y regeneración muscular, lo que permitirá la determinación y caracterización a nivel proteómico, transcriptómico e interactómico de los loci aquí propuestos. Esto permitirá la comprensión integral de la regulación transcripcional a la que se ven sometidas las células en el proceso de diferenciación muscular.

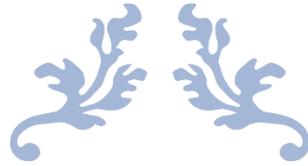
El conocimiento de la presencia y disposición topológica de los factores involucrados en el proceso de regulación transcripcional a través del sistema CAPTURE en modelos musculares, nos permitirá la comprensión de la regulación epigenética en procesos como la embriogénesis, diferenciación y regeneración muscular en estados homeostáticos y su exploración en enfermedades específicas como la distrofia muscular de Duchenne (DMD), la distrofia fascioescapulohumeral (FSHD) y la distrofia miotónica (DM). Además, la exploración de componentes reguladores en otros contextos como condiciones metabólicas, envejecimiento o cáncer es plausible.



Conclusiones



En el presente trabajo se ha logrado la formación de una biblioteca de plásmidos necesarios para el establecimiento del sistema CAPTURE; del mismo modo se han obtenido células C2C12-dCas9-BirA que expresan de manera estable la proteína dCas9 y la donadora de biotinas BirA, estas células C2C12-dCas9-BirA conservan su identidad celular como se comprobó por el análisis de MyoD y los marcadores de diferenciación MyHC y MyoG. Además, se ha logrado el diseño de sgRNAs dirigidos hacia regiones regulatorias de MyoD y Pax7. Por último, se ha determinado la estrategia experimental para la clonación de sgRNAs específicos y su transducción en células C2C12 con la finalidad de establecer su aplicación práctica para la recuperación de complejos asociados a la cromatina y su posterior análisis.



Referencias



1. Zurlo F, Larson K, Bogardus C, Ravussin E. Skeletal muscle metabolism is a major determinant of resting energy expenditure. *J Clin Invest* 1990;86(5):1423e7.
2. Schmidt, M., Schüler, S., Hüttner, S., Eyss, B., Maltzahn, J. (2019). Adult stem cells at work: regenerating skeletal muscle. *Cellular and Molecular Life Sciences*. Volume 76, Issue 13, pp 2559–2570.
3. León A, Melo M, Ramirez V. Role of the myokines production through the exercise. *Journal of Sport and Health Research*, 2012; 4: 157-166.
4. Yucel N, Blau Hm. Skeletal Muscle stem cell. En: *Principles of regenerative Medicine*. 3a edition. 2019
5. Gilbert, SF. (2000) *Developmental Biology*. 6th edición. Sinauer Associates. Massachusetts.
6. Cossu, G.; Borello, U.; (1999) Wnt signaling and the activation of myogenesis in mammals. *EMBO* 18: 6867-6872
7. Ritso, M.; Lin, A.; Hernández-Hernández, J.; Rudnicki, M. (2019) Skeletal Muscle Satellite Cells. En: *Deferring Development*. Taylor and Francis Group.
8. Li, J.; Mayne, R.; Wu, C. (1999) A novel muscle-specific (1-integrin binding protein (MIBP)) that modulates Myogenic Differentiation. *J. Cell Biol.* 147 (7): 1391-1398
9. Musaro, A.; Rosenthal, N. (1999) Maturation of the Myogenic Program is induced by postmitotic expression of Insuline-like Growth factor I. *Mol. Cell Bio.* 19 (4): 3115-3124
10. Chal, J. & Pourquié, O. (2017) Making muscle: skeletal myogenesis. *The Company of Biologist*. doi:10.1242/dev.151035

11. Chal J, Pourquié O. Making muscle: skeletal myogenesis. *The Company of Biologist*. 2017. doi:10.1242/dev.151035
12. Velthoven CTJ, Rando TA. Stem cell quiescence: dynamism, restraint, and cellular idling. *Cell Stem Cell Review* 2019; 24: 213-225
13. Seale, P. et al. Pax7 is required for the specification of myogenic satellite cells. *Cell* 2000; 102: 777–786.
14. Dumont NA, Wang YX, Rudnicki MA. Intrinsic and extrinsic mechanisms regulating satellite cell function. *The company of Biologist* 2015; 142: 15721581
15. Wen, Y. et al. Constitutive Notch activation upregulates Pax7 and promotes the self-renewal of skeletal muscle satellite cells. *Mol. Cell. Biol.* 2012; 32: 2300–2311
16. Cheung TH, Quach NL, Charville GW, Liu L, Park L, Edalati A, et al. Maintenance of muscle stem-cell quiescence by miRNA-489. *Nature*. 2012; 482: 524528
17. Crist CG, Montarras D, Buckingham M. Muscle satellite cells are primed for myogenesis but maintain quiescence with sequestration of Myf5 mRNA targeted by microRNA-31 in mRNP granules. *Cell Stem Cell* 2012; 11: 118-126.
18. Wozniak AC, Pilipowicz O, Yablonka-Reuveni Z, Greenway S, Craven S, Scoott E, Anderson JE. C-Met expression and mechanical activation of satellite cells on cultured muscle fibers. *Journal of histochemistry and cytochemistry* 2003; 11: 1437-1444
19. Templeton TJ, Hauschka SD. FGF-mediated aspects of skeletal muscle growth and differentiation are controlled by a high affinity receptor, FGFR1. *Developmental Biology* 1992; 154: 169-181.
20. Sheehan SM, Allen RE. Skeletal muscle satellite cell proliferation in response to members of the fibroblast growth factor family and hepatocyte growth factor. *Journal of cellular physiology* 1999; 181: 499-506.
21. Cornelison DDW, Filla MS, Stanley HM, Rapraeger AC, Olwin BB. Syndecan-3 and Syndecan-4 specifically mark skeletal muscle satellite cells and are implicated in satellite cell maintenance and muscle regeneration. *Developmental Biology* 2001; 239: 79-94.

22. Ono, Y. et al. Muscle stem cell fate is controlled by the cell polarity protein scrib. *Cell Rep.* 2015; 10: 1135–1148
23. Troy, A. et al. Coordination of satellite cell activation and self-renewal by Par-complex-dependent asymmetric activation of p38alpha/beta MAPK. *Cell Stem Cell* 2012; 11: 541–553.
24. Sambasivan, R. et al. Pax7-expressing satellite cells are indispensable for adult skeletal muscle regeneration. *Development* 2011; 138: 3647–3656.
25. Kuang, S. et al. Asymmetric self-renewal and commitment of satellite stem cells in muscle. *Cell* 2007; 129: 999–1010.
26. Brack, A.S. et al. A temporal switch from notch to Wnt signaling in muscle stem cells is necessary for normal adult myogenesis. *Cell Stem Cell* 2008; 2: 50–59.
27. Carlson, M.E. et al. Imbalance between pSmad3 and Notch induces CDK inhibitors in old muscle stem cells. *Nature* 2008; 454: 528–532.
28. Olguin, H. C. et al. Reciprocal inhibition between Pax7 and muscle regulatory factors modulates myogenic cell fate determination. *J. Cell Biol.* 2007; 177: 769–779.
29. Hernández-Hernández, J M., García-González, E G., Brunab C., Rudnicki M A. (2017) The myogenic regulatory factors, determinants of muscle development, cell identity and regeneration. *Seminars in Cell & Developmental Biology.* 72. 10-18.
30. Zammit, P. (2017). Function of the myogenic regulatory factors Myf5, MyoD, Myogenin and MRF4 in skeletal muscle, sc and regenerative myogenesis. *Seminars in Cell & Developmental Biology.* Volume 72, pp 19-32.
31. Hirai, H., Tani, T., & Kikyo, N. (2010). Structure and functions of powerful transactivators: VP16, MyoD and FoxA. *The International Journal of Developmental Biology*, 54(11-12), 1589–1596. doi:10.1387/ijdb.103194
32. Wang, C., Wang, M., Arrington, J., Shan, T., Yue, F., Nie, Y., ... Kuang, S. (2016). *Ascl2* inhibits myogenesis by antagonizing the transcriptional activity of myogenic regulatory factors. *Development*, 144(2), 235–247. doi:10.1242/dev.138099
33. Asfour, H.A., Allouh, M.Z., & Said, R.S. (2018). Myogenic regulatory factors: The orchestrators of myogenesis after 30 years of

- discovery. *Experimental Biology and Medicine*, 243(2), 118-128.
Doi:10.1177/1535370217749494.
34. Gerhart, J., Elder, J., Neely, C., Schure, J., Kvist, T., Knudsen, K., & George-Weinstein, M. (2006). MyoD-positive epiblast cells regulate skeletal muscle differentiation in the embryo. *The Journal of Cell Biology*, 175(2), 283–292. doi: 10.1083/jcb.200605037
 35. Berkes, C. A., Bergstrom, D. A., Penn, B. H., Seaver, K. J., Knoepfler, P. S., & Tapscott, S. J. (2004). Pbx Marks Genes for Activation by MyoD Indicating a Role for a Homeodomain Protein in Establishing Myogenic Potential. *Molecular Cell*, 14 (4), 465–477. doi: 10.1016/s1097-2765(04)00260-6
 36. Wittkopp, P.J.; Kalay, G. (2012) Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nature*, 13: 59-69
 37. Butanda Ochoa, A.; Guevara Flores, A.; Guevara Fonseca, J.; Matuz Mares, D.; Lara Lemus, R.; Torres Durán, P.V. (2014) Modificaciones en la cromatina y en la herencia epigenética. *Mensaje Bioquímico*, Vol. XLI: 253-288.
 38. Bergstrom DA, Penn BH, Strand A, Perry RL, Rudnicki MA, Tapscott SJ. Promoter-specific regulation of MyoD binding and signal transduction cooperate to pattern gene expression. *Mol Cell* 2002; 9:587-600; PMID:11931766; [http://dx.doi.org/10.1016/S1097-2765\(02\)00481-1](http://dx.doi.org/10.1016/S1097-2765(02)00481-1)
 39. Murre C, McCaw PS, Vaessin H, Caudy M, Jan LY, Jan YN, et al. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* 1989; 58:537-44; PMID:2503252; [http://dx.doi.org/10.1016/0092-8674\(89\)90434-0](http://dx.doi.org/10.1016/0092-8674(89)90434-0)
 40. Blackwell TK, Weintraub H. Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* 1990; 250:1104-10; PMID:2174572; <http://dx.doi.org/10.1126/science.2174572>
 41. Blais A, Tsikitis M, Acosta-Alvear D, Sharan R, Kluger Y, Dynlacht BD. An initial blueprint for myogenic differentiation. *Genes Dev* 2005; 19:55369; PMID:15706034; <http://dx.doi.org/10.1101/gad.1281105>

42. Cao Y, Yao Z, Sarkar D, Lawrence M, Sanchez GJ, Parker MH, et al. Genome-wide MyoD binding in skeletal muscle cells: a potential for broad cellular reprogramming. *Dev Cell* 2010; 18:662-74; PMID:20412780; <http://dx.doi.org/10.1016/j.devcel.2010.02.014>
43. Davis RL, Weintraub H. Acquisition of myogenic specificity by replacement of three amino acid residues from MyoD into E12. *Science* 1992; 256:102730; PMID:1317057; <http://dx.doi.org/10.1126/science.1317057>
44. Bengal E, Flores O, Rangarajan PN, Chen A, Weintraub H, Verma IM. Positive control mutations in the MyoD basic region fail to show cooperative DNA binding and transcriptional activation in vitro. *Proc Natl Acad Sci U S A* 1994; 91:62215; PMID:8016142; <http://dx.doi.org/10.1073/pnas.91.13.6221>
45. Dilworth FJ, Seaver KJ, Fishburn AL, Htet SL, Tapscott SJ. In vitro transcription system delineates the distinct roles of the coactivators pCAF and p300 during MyoD/E47-dependent transactivation. *Proc Natl Acad Sci U S A* 2004; 101:115938; PMID:15289617; <http://dx.doi.org/10.1073/pnas.0404192101>
46. Yee SP, Rigby PW. The regulation of myogenin gene expression during the embryonic development of the mouse. *Genes Dev* 1993; 7(7A):127789; PMID:8391506; <http://dx.doi.org/10.1101/gad.7.7a.1277>
47. Cheng TC, Wallace MC, Merlie JP, Olson EN. Separable regulatory elements governing myogenin transcription in mouse embryogenesis. *Science* 1993; 261:215-8; PMID:8392225; <http://dx.doi.org/10.1126/science.8392225>
48. Goldhamer DJ, Brunk BP, Faerman A, King A, Shani M, Emerson CP Jr. Embryonic activation of the myoD gene is regulated by a highly conserved distal control element. *Development* 1995; 121:637-49; PMID:7720572
49. Donoghue M, Ernst H, Wentworth B, Nadal-Ginard B, Rosenthal N. A muscle-specific enhancer is located at the 3' end of the myosin light-chain 1/3 gene locus. *Genes Dev* 1988; 2(12B):1779-90; PMID:3240859; <http://dx.doi.org/10.1101/gad.2.12b.1779>
50. Rosenthal N, Berglund EB, Wentworth BM, Donoghue M, Winter B, Bober E, et al. A highly conserved enhancer downstream of the human MLC1/3 locus is a target for multiple myogenic determination factors. *Nucleic Acids*

- Res 1990; 18:623946; PMID:2243772; <http://dx.doi.org/10.1093/nar/18.21.6239>
51. Wentworth BM, Donoghue M, Engert JC, Berglund EB, Rosenthal N. Paired MyoD-binding sites regulate myosin light chain gene expression. *Proc Natl Acad Sci U S A* 1991; 88:1242-6; PMID:1847512; <http://dx.doi.org/10.1073/pnas.88.4.1242>
52. Johnson JE, Wold BJ, Hauschka SD. Muscle creatine kinase sequence elements regulating skeletal and cardiac muscle expression in transgenic mice. *Mol Cell Biol* 1989; 9:3393-9; PMID:2796990
53. Horlick RA, Benfield PA. The upstream musclespecific enhancer of the rat muscle creatine kinase gene is composed of multiple elements. *Mol Cell Biol* 1989; 9:2396-413; PMID:2761536
54. Carvajal JJ, Keith A, Rigby PW. Global transcriptional regulation of the locus encoding the skeletal muscle determination genes Mrf4 and Myf5. *Genes Dev* 2008; 22:265-76; PMID:18198342; <http://dx.doi.org/10.1101/gad.442408>
55. Chang TH, Primig M, Hadchouel J, Tajbakhsh S, Rocancourt D, Fernandez A, et al. An enhancer directs differential expression of the linked Mrf4 and Myf5 myogenic regulatory genes in the mouse. *Dev Biol* 2004; 269:595-608; PMID:15110722; <http://dx.doi.org/10.1016/j.ydbio.2004.02.013>
56. Wakabayashi-Takai E, Noguchi S, Ozawa E. Identification of myogenesis-dependent transcriptional enhancers in promoter region of mouse gamma-sarcoglycan gene. *Eur J Biochem* 2001; 268:94857; PMID:11179961; <http://dx.doi.org/10.1046/j.1432-1327.2001.01954.x>
57. Barthel KK, Liu X. A transcriptional enhancer from the coding region of ADAMTS5. *PLoS One* 2008; 3:e2184; PMID:18478108; <http://dx.doi.org/10.1371/journal.pone.0002184>
58. Scionti, I.; Hayashi, S.; Mourandian, S.; Girard, E.; et al. (2017) LSD1 Controls timely MyoD expression via MyoD Core Enhancer transcription. *Cell Reports*, 18: 1996-2006
59. Soleimani VD, Punch VG, Kawabe Y et al. Transcriptional dominance of Pax7 in adult myogenesis due to high-affinity recognition of homeodomain motifs. *Dev Cell* 2012;22:1208–1220.

60. Buchberger A, Freitag D, RNAold HH. A homeo-paired domain-binding motif directs Myf5 expression in progenitor cells of limb muscle. *Development* 2007;134:1171–1180.
61. Zammit PS, Carvajal JJ, Golding JP et al. Myf5 expression in satellite cells and spindles in adult muscle is controlled by separate genetic elements. *Dev Biol* 2004;273:454–465.
62. Ribas R, Moncaut N, Siligan C et al. Members of the TEAD family of transcription factors regulate the expression of Myf5 in ventral somitic compartments. *Dev Biol* 2011;355:372–380.
63. Li ZY, Yang J, Gao X, et al. Sequential recruitment of PCAF and BRG1 contributes to myogenin activation in 12-O-tetradecanoylphorbol-13-acetate-induced early differentiation of rhabdomyosarcoma-derived cells. *J Biol Chem*. 2007;282(26):18872-18878.
64. Yoh-ichi Kawabe, et al. (2013) Carm1 Regulates Pax7 Transcriptional Activity through MLL1/2 Recruitment during Asymmetric Satellite Stem Cell Divisions. *Cell Stem Cell* 11:3. 333-345
65. Marchesi I, Giordano A, Bagella L. Roles of enhancer of zeste homolog 2: From skeletal muscle differentiation to rhabdomyosarcoma carcinogenesis. *Cell Cycle* 2014;13:516–527.
66. Palacios D, Mozzetta C, Consalvi S et al. TNF/p38a/polycomb signaling to Pax7 locus in satellite cells links inflammation to the epigenetic control of muscle regeneration. *Cell Stem Cell* 2010;7:455–469.
67. Caretti G, Di Padova M, Micales B et al. The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. *Genes Dev* 2004;18: 2627–2638.
68. Juan AH, Kumar RM, Marx JG et al. Mir214-dependent regulation of the polycomb protein Ezh2 in skeletal muscle and embryonic stem cells. *Mol Cell* 2009;36:61–74.
69. Wong CF, Tellam RL. MicroRNA-26a targets the histone methyltransferase Enhancer of Zeste homolog 2 during myogenesis. *J Biol Chem* 2008;283:9836–9843.
70. Seenundun S, Rampalli S, Liu QC et al. UTX mediates demethylation of H3K27me3 at muscle specific genes during myogenesis. *EMBO J* 2010; 29:1401–1411.

71. Zhang, T., Günther, S., Looso, M. et al. Prmt5 is a regulator of muscle stem cell expansion in adult mice. *Nat Commun* 6, 7140 (2015). <https://doi.org/10.1038/ncomms8140>
72. Pombo, A.; Dillon, N. (2015) Three-dimensional genome architecture: players and mechanisms. *Nature Reviews Molecular Cell Biology*, 16: 245-257.
73. Tsai P.F.; Dell'Orso, S.; Rodriguez, J.; Vivanco K.O.; Ko, K.D.; et al. (2018) A muscle-specific enhancer RNA mediates Cohesin recruitment and regulates Transcription in trans. *Molecular Cell*, 71: 129-141
74. Wang, S., Zuo, H., Jin, J. et al. Long noncoding RNA Neat1 modulates myogenesis by recruiting Ezh2. *Cell Death Dis* 10, 505 (2019). <https://doi.org/10.1038/s41419-019-1742-7>
75. West JA, Davis CP, Sunwoo H, et al. The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol Cell*. 2014;55(5):791–802. doi:10.1016/j.molcel.2014.07.012
76. Gutschner, T., Hämmerle, M. & Diederichs, S. MALAT1 — a paradigm for long noncoding RNA function in cancer. *J Mol Med* 91, 791–801 (2013). <https://doi.org/10.1007/s00109-0131028-y>
77. Chen, X., He, L., Zhao, Y. et al. Malat1 regulates myogenic differentiation and muscle regeneration through modulating MyoD transcriptional activity. *Cell Discov* 3, 17002 (2017).
78. Jin JJ, Lv W, Xia P, et al. Long noncoding RNA SYISL regulates myogenesis by interacting with polycomb repressive complex 2. *Proc Natl Acad Sci U S A*. 2018;115(42):E9802–E9811.
79. Hitachi K, Nakatani M, Takasaki A, et al. Myogenin promoter-associated lncRNA Myoparr is essential for myogenic differentiation. *EMBO Rep*. 2019;20(3):e47468.
80. Hitachi K, Tsuchida K. Data describing the effects of depletion of Myoparr, myogenin, Ddx17, and hnRNPK in differentiating C2C12 cells. *Data Brief*. 2019;25:104172.
81. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 2012 Apr 11;485(7398):376-80

82. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*. 2012 Apr 11;485(7398):381-5
83. Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, et al. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell*. 2012 Feb 3;148(3):458-72.
84. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragooczy T, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009 Oct 9;326(5950):28993
85. Le Dily F, Baù D, Pohl A, Vicent GP, Serra F, et al. Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev*. 2014 Oct 1;28(19):2151-62
86. Pope BD, Ryba T, Dileep V, Yue F, Wu W, et al. Topologically associating domains are stable units of replication-timing regulation. *Nature*. 2014 Nov 20;515(7527):402-5
87. Cavalli G, Misteli T. Functional implications of genome topology. *Nat Struct Mol Biol*. 2013 Mar;20(3):2909.
88. de Laat W, Duboule D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature*. 2013 Oct 24;502(7472):499-506.
89. Daniel S. Neems, Arturo G. Garza-Gongora, Erica D. Smith, Steven T. Kosak, (2016) Topologically associated domains enriched for lineage-specific genes reveal expression-dependent nuclear topologies during myogenesis. *Proceedings of the National Academy of Sciences*, 113 (12) E1691-E1700; DOI: 10.1073/pnas.1521826113
90. Heinz, S., Romanoski, C. E., Benner, C., and Glass, C. K. (2015). The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biology* 16, 144–54. doi:10.1038/nrm3949
91. Sun, F., Chronis, C., Kronenberg, M., Chen, X.-F., Su, T., Lay, F. D., et al. (2019). Promoter-Enhancer Communication Occurs Primarily within Insulated Neighborhoods. *Mol Cell* 73, 250-263.e5. doi:10.1016/j.molcel.2018.10.039.
92. Siersbæk, R., Madsen, J. G. S., Javierre, B. M., Nielsen, R., Bagge, E. K., Cairns, J., et al. (2017). Dynamic Rewiring of Promoter-Anchored

- Chromatin Loops during Adipocyte Differentiation. *Mol Cell* 66, 420-435.e5. doi:10.1016/j.molcel.2017.04.010.
93. A Dall'Agnesse, L Caputo, C Nicoletti, J di Iulio, et al. (2019) Transcription factor-directed re-wiring of chromatin architecture for somatic cell nuclear reprogramming toward trans-differentiation. *Molecular cell*. 76:3. 453-472
94. Gómez-Marín, C. Arquitectura de la cromatina de los complejos Six en el desarrollo y la evolución. Tesis Doctoral. Universidad Pablo de Olavide. (2016)
95. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell*. 2007 Jul 13;130(1):77-88.
96. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet*. 2007 Mar;39(3):311-8
97. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, et al. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*. 2011 Feb 10;470(7333):279-83.
98. Bogdanovic O, Fernandez-Miñán A, Tena JJ, de la Calle-Mustienes E, Hidalgo C, et al. Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis. *Genome Res*. 2012 Oct;22(10):2043-53.
99. Akkers RC, van Heeringen SJ, Jacobi UG, Janssen-Megens EM, François KJ, et al. A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in *Xenopus* embryos. *Dev Cell*. 2009 Sep;17(3):425-34
100. Lim EJ, Lu TX, Blanchard C, Rothenberg ME. Epigenetic regulation of the IL-13-induced human eotaxin-3 gene by CREB-binding protein-mediated histone 3 acetylation. *J Biol Chem*. 2011 Apr 15;286(15):13193-204.
101. Crawford GE, Holt IE, Mullikin JC, Tai D, Blakesley R, et al. Identifying gene regulatory elements by genome-wide recovery of DNase hypersensitive sites. *Proc Natl Acad Sci U S A*. 2004 Jan 27;101(4):992-7.

102. Sabo PJ, Humbert R, Hawrylycz M, Wallace JC, Dorschner MO, et al. Genome-wide identification of DNaseI hypersensitive sites using active chromatin sequence libraries. *Proc Natl Acad Sci U S A*. 2004 Mar 30;101(13):4537-42
103. Giresi PG, Kim J, McDaniell RM, Iyer VR, Lieb JD. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res*. 2007 Jun;17(6):877-85
104. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*. 2013 Dec;10(12):1213-8
105. David C. Klein & Sarah J. Hainer (2020) Genomic methods in profiling DNA accessibility and factor localization. *Chromosome Research* volume 28, pages 69–85
106. Gavrillov A, Eivazova E, Priozykova I, Lipinski M, Razin S, et al. Chromosome conformation capture (from 3C to 5C) and its ChIP-based modification. *Methods Mol Biol*. 2009;567:171-88
107. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. *Science*. 2002 Feb 15;295(5558):1306-11.
108. Liu, X.; Zang, Y.; Chen, Y.; Li, M.; Zhou F.; Li, K.; Cao, H.; Ni, M.; Liu, Y.; Gu Z., Dickerson, K.E.; Xie, S.; Hon, G.H.; Xuan, Z.; Zhang, M.Q.; Shao, Z.; Xu, J. (2017) In situ capture of chromatin interactions by biotinylated dCas9. *Cell*, 170: 1028-1043.
109. Liu, X.; Zhang, Y.; Chen, Y.; Li, M.; Shao, Z.; Zhang M.Q.; Xu, J. (2018) CAPTURE: In situ Analysis of chromatin composition of endogenous genomic loci by biotinylated dCas9. *Current protocols in molecular biology*. e64: 1-24.
110. Alex Hodgkins, Anna Farne, Sajith Perera, Tiago Grego, David J. Parry-Smith, William C. Skarnes, Vivek Iyer, WGE: a CRISPR database for genome engineering, *Bioinformatics*, Volume 31, Issue 18, 15 September 2015, Pages 3078–3080
111. McGrath, E., Shin, H., Zhang, L., Phue, J. N., Wu, W. W., Shen, R. F., Jang, Y. Y., Revollo, J., & Ye, Z. (2019). Targeting specificity of

- APOBEC-based cytosine base editor in human iPSCs determined by whole genome sequencing. *Nature communications*, 10(1), 5353. <https://doi.org/10.1038/s41467-019-13342-8>
112. Pravata, V. M., Gundogdu, M., Bartual, S. G., Ferenbach, A. T., Stavridis, M., Ōunap, K., Pajusalu, S., Źordania, R., Wojcik, M. H., & van Aalten, D. (2020). A missense mutation in the catalytic domain of O-GlcNAc transferase links perturbations in protein O-GlcNAcylation to X-linked intellectual disability. *FEBS letters*, 594(4), 717–727. <https://doi.org/10.1002/1873-3468.13640>
 113. Burrage, L. C., Reynolds, J. J., Baratang, N. V., et al. (2019). Biallelic Variants in TONSL Cause SPONASTRIME Dysplasia and a Spectrum of Skeletal Dysplasia Phenotypes. *American journal of human genetics*, 104(3), 422–438. <https://doi.org/10.1016/j.ajhg.2019.01.007>
 114. Vermeulen, M., Déjardin, J. Locus-specific chromatin isolation. *Nat Rev Mol Cell Biol* 21, 249–250 (2020). <https://doi.org/10.1038/s41580-020-0217-0>
 115. Déjardin, J. & Kingston, R. E. Purification of proteins associated with specific genomic Loci. *Cell* 136, 175–186 (2009).
 116. Ide, S. & Dejardin, J. End-targeting proteomics of isolated chromatin segments of a mammalian ribosomal RNA gene promoter. *Nat. Commun.* 6, 6674 (2015).
 117. Byrum, S. D., Taverna, S. D. & Tackett, A. J. Purification of a specific native genomic locus for proteomic analysis. *Nucleic Acids Res.* 41, e195 (2013)
 118. Griesenbeck, J., Boeger, H., Strattan, J. S. & Kornberg, R. D. Affinity purification of specific chromatin segments from chromosomal loci in yeast. *Mol. Cell. Biol.* 23, 9275–9282 (2003).
 119. Fujita, T. & Fujii, H. Direct identification of insulator components by insertional chromatin immunoprecipitation. *PLoS One* 6, e26109 (2011).
 120. Schmidtman, E., Anton, T., Rombaut, P., Herzog, F. & Leonhardt, H. Determination of local chromatin composition by CasID. *Nucleus* 7, 476–484 (2016).
 121. Fujita, T. & Fujii, H. Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding

molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. *Biochem. Biophys. Res. Commun.* 439, 132–136 (2013).

122. Gauchier, M., van Mierlo, G., Vermeulen, M. et al. Purification and enrichment of specific chromatin loci. *Nat Methods* 17, 380–389 (2020). <https://doi.org/10.1038/s41592-020-0765-4>



Anexos



Tabla 1: Lista de Primers.

Primers		Secuencia
Myog	Foward	CAGTGAATGCAACTCCCACA
	Revers	GAGCAAATGATCTCCTGGGTT
RPS18	Foward	AACGGTCTAGACAACAAGCTG
	Revers	AGTGGTCTTGGTGTGCTGAC
MyoD	Foward	TGATGGCATGATGGATTACAGCG
	Revers	GTAGTAGGCGGTGTCGTAGC
MYH	Foward	TCAATGAGATGGAGATCCAGCTGAAC
	Revers	GTCCAGGTGCAGCTGTGTGTCCTTC
dCas9	Foward	CGTGGACGCCATCGTCCCTC
	Revers	CAGTCCGCCCTCTCAGCCT
BirA	Foward	ACTGCGTGA CTGGGGCGTTG
	Revers	CATTTCCGACCCCGGCGACC
MyoD-PRR-sgRNA	ggagaaCCACCTTGTGGT <u>AGCCAAGTGCTACCGCGTAGTTTAAGAGCTA</u> TGCTGGAAACAGCA	
MyoD-DRR-sgRNA	ggagaaCCACCTTGTGG <u>AGACATCCCAAAAATGACTT</u> GTTTAAGAGCTAT GCTGGAAACAGCA	
MyoD-TSS-sgRNA	ggagaaCCACCTTGTGG <u>TTAAGACGACTCTCACGGCT</u> GTTTAAGAGCTAT GCTGGAAACAGCA	
Universal Reverso	ctagtaCTCGAGAAAAAAGCACCGACTCGGTGCCAC	

Tabla 2. Lista de sgRNAs diseñados

Nombre	Secuencia*	Orientación	Inicio	Término	Distancia a TSS	ID
MYOD						
PRR-sg1	TCTCTTGCCCACTTT GTCCT	+	463759 88	463760 10	-486	515075 035
PRR-sg2	CGCGGTAGCACTTG GCTATC	-	463762 27	463762 49	-247	515075 065
PRR-sg3	TAGCCAAGTGCTACC GCGTA	+	463762 32	463762 54	-242	515075 066
DRR-sg1	TAAAGGCTGCTGTTT CCCCA	+	463713 24	463713 46	-5150	515074 323
DRR-sg2	GGCTTGGGGCAGGT GCTAGT	+	463712 74	463712 96	-5200	515074 317
DRR-sg3	GGCAGAGGTGCCTG AGGCTT	+	463712 59	463712 81	-5215	515074 313
DRR-sg4	AGACATCCCAAAAAT GACTT	+	463725 71	463725 93	-3903	515074 494
DRR-sg5	CCCCTGGTCCTGCCT GTGAC	-	463726 21	463726 43	-3853	515074 502
Pax7						
Prom-1	GGTCGAGTTGGAGA TTAAAA	-	139833 727	139833 749	-720	476841 221
Prom-2	TGGAGACTTCTTTTC GTTAT	-	139833 685	139833 707	-678	476841 221
Prom-3	GGGGCCAGTTCGGG AGGGAG	+	139833 659	139833 681	-652	476841 212
Prom-4	TTTTTCCTCAGGGGG TCGCG	+	139833 782	139833 804	-775	476841 230
PRE-1	AAAAACGGGGCTTTG ATTTG	+	139837 397	139837 419	-4390	476841 903
PRE-2	GCTGGGCCAGAGCC ACCACC	+	139837 443	139837 465	-4436	476841 909
PRE-3	ACTGCAGCATCTTTC TTTAC	+	139837 349	139837 371	-4342	476841 894
PRE-4	CCAGAGCGTGGGTG TGGGAA	+	139837 487	139837 509	-4480	476841 924