



UNIVERSIDAD
NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE MEDICINA
INSTITUTO DE FISIOLÓGIA CELULAR

**ANÁLISIS BIOINFORMÁTICO COMPARATIVO DE
MOTIVOS DE *DNA* EN SECUENCIAS REGULATORIAS
DEL GEN *SCN9A* (Na_v1.7).**

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

LICENCIADA EN NEUROCIENCIAS

P R E S E N T A

ESTEFANÍA SANTANA ROMÁN

DIRECTOR DE TESIS:

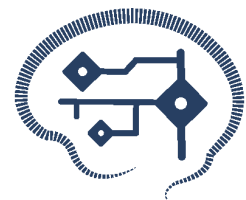
DR. AUGUSTO CÉSAR POOT HERNÁNDEZ

CIUDAD UNIVERSITARIA, CIUDAD DE MÉXICO, 2022.

Facultad de Medicina



Instituto de
Fisiología
Celular



**NEUROCIENCIAS
UNAM**



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

AGRADECIMIENTOS ACADÉMICOS

El presente trabajo se realizó bajo la tutoría del Dr. Augusto César Poot Hernández, en la Unidad de Bioinformática y Manejo de la Información del Instituto de Fisiología Celular de la Universidad Nacional Autónoma de México. Agradezco al Dr. y a los que conforman la UBMI su tiempo, dedicación y apoyo a través de este camino.

A mi comité tutor:

- Dr. Arturo Carlos II Becerra Bracho
- Dr. Fernando Espinosa Camacho
- Dr. Enoch Luis Baltazar
- Dr. Diego Claudio Cortez Quezada

Gracias por sus comentarios, correcciones y consejos, que fueron centrales para acrecentar mi conocimiento y enriquecer este trabajo. Agradezco especialmente la empatía y amabilidad que mostraron a mi petición.

A la Licenciatura en Neurociencias de la Universidad Nacional Autónoma de México. Gracias a todas las personas que la conforman, quienes me han apoyado, orientado y compartido sus conocimientos y experiencias durante estos años. Agradezco el esfuerzo que han realizado para guiarnos en lo que era una nueva experiencia tanto para ustedes como para nosotros.

A la UNAM, por la oportunidad de estudiar una carrera universitaria y de convivir con estudiantes y profesionistas que han sido clave en mi desarrollo académico y personal.

AGRADECIMIENTOS PERSONALES

Deseo dedicar las siguientes líneas a las personas que han estado presentes a lo largo de este recorrido y para quienes resumo mi sentir en una sola palabra: ¡gracias!

A mamá: no he encontrado las palabras exactas que me permitan resumir todo el cariño y agradecimiento que te tengo, los cuales son infinitos. Sin embargo, solo deseo que tengas claro que conozco y reconozco la convicción y valentía con que has transitado estos años, siempre velando por mis hermanos y por mí. Has sido mi mayor soporte, quien me ha impulsado en cada etapa y me ha ayudado a creer en mí misma; gracias por darme el valor de aventurarme a soñar.

Camino tranquila porque tengo la certeza de que siempre vas a mi lado.

Especialmente a papá: me hubiera encantado que me vieras llegar a esta etapa.

Tu curiosidad y gusto por el conocimiento fueron la mejor enseñanza que me diste. En ese entonces aún no sabía que estaba aprendiendo de tu ejemplo, pero ahora me doy cuenta de ello y agradezco que hayamos tenido la oportunidad de coincidir en el mismo espacio. Tu esencia y tu cariño ya han traspasado las barreras del tiempo, puesto que aún llevo de estandarte tus consejos.

A mi hermana y a mi hermano: mis primeros amigos y maestros.

Una de las principales razones por las que he logrado afrontar los momentos difíciles y he disfrutado tanto de la vida es porque ustedes han sido mi compañía. Pienso que llegar a este punto es solo la consecuencia obvia de haber crecido guiada de la alegría, dedicación, inteligencia y bondad de mis hermanos mayores. Gracias por la complicidad, las risas y las enseñanzas.

A mis amigas y amigos, que son una parte esencial y muy especial en mi vida. Pese a que este escrito representa el fin de un camino, lo más importante ha sido la travesía y puedo decir que tuve el gusto de compartirla con ustedes. Esta experiencia no sólo representó un aumento de conocimientos académicos, sino de vida, para lo cual su amistad, los momentos compartidos y consejos jugaron un papel central. Gracias por tantas y tan bonitas anécdotas.

A mi tutor de tesis: agradezco que mi primer acercamiento a un equipo de investigación haya sido bajo la tutoría de una persona con tantos conocimientos y, sobre todo, con tantos valores como lo es usted. Gracias por darme la oportunidad de formar parte de la unidad, por la paciencia, el tiempo dedicado a mi formación y por los consejos, que no solo me ayudaron a crecer académicamente sino también personalmente.

CONTENIDO

ABREVIATURAS	7
ÍNDICE DE TABLAS Y FIGURAS	8
1. RESUMEN	10
2. INTRODUCCIÓN	12
2.1 Nocicepción y dolor	12
2.2 Nociceptores	14
2.3 SCN9A/NAV1.7 y trastornos del dolor	15
2.4 Secuencias de DNA no codificantes	16
2.5 Secuencias motivo	18
2.6 Bases de datos biológicas	18
2.7 Algoritmos y herramientas bioinformáticas en la identificación de secuencias motivo	19
3. JUSTIFICACIÓN	22
4. OBJETIVOS	23
4.1 General	23
4.2 Particulares	23
5. ESTRATEGIA EXPERIMENTAL	24
6. MÉTODOS	25
6.1 Selección y exclusión de secuencias de distintas especies de mamíferos	25
6.2 Obtención de secuencias desde la Base de Datos Ensembl	26
6.3 Alineamiento y comparación de secuencias conservadas	27
6.4 Identificación de secuencias motivo en la región del promotor	28
6.6 Selección y exclusión de enhancers relacionados al gen SCN9A	29
6.7 Análisis e identificación de secuencias motivo dentro de los enhancers seleccionados.	30
6.8 Análisis de la relación entre las secuencias motivo identificadas y los factores de transcripción ya conocidos.	30
7. RESULTADOS Y DISCUSIÓN	31
7.1 Identificación del promotor central del gen SCN9A en distintas especies de mamíferos.	31
7.2 Análisis de motivos en el promotor central	32
7.3 Enhancers relacionados al gen SCN9A	39
7.4 Identificación de enhancers	39
7.5 Análisis de motivos de los enhancers.	41
7.6 Análisis discriminatorio de motivos	48

7.8 Limitaciones	56
8. PERSPECTIVAS	58
9. CONCLUSIONES	63
10. REFERENCIAS	64
11. ANEXO	71

ABREVIATURAS

BD	Base de Datos
DNA	Ácido desoxirribonucleico
DRG	<i>Dorsal Root Ganglia</i>
FT	Factor de Transcripción
GA	Puntuación de Asociación
GH	<i>GeneHancer</i>
GT	Ganglio trigémino
GTE _x	Atlas de Expresión de Tejidos
<i>HOCOMOCO</i>	HOmo sapiens COmprehensive MOdel COllection
ICD	Insensibilidad Congénita al Dolor
nt	nucleótido
pb	pares de bases
RNA	Ácido Ribonucleico
SIT	Sitio de Inicio de la Transcripción
SN	Sistema Nervioso
SNC	Sistema Nervioso Central
SUFT/SUFTs	Sitio/ Sitios de Unión a Factores de Transcripción
SNP	Sistema Nervioso Periférico
TPDE	Trastorno Paroxismal de Dolor Extremo
VGSCs	Canales de Sodio Dependientes de Voltaje

ÍNDICE DE TABLAS Y FIGURAS

Figura 1. Distribución anatómica de la nocicepción y el dolor.	13
Figura 2. Región regulatoria típica de un gen.	17
Figura 3. Técnica general empleada en el descubrimiento de motivos.	20
Figura 4. Estrategia experimental	24
Figura 5. Información filogenética de la clase Mammalia y el orden Eutheria	25
Tabla 1. Cromosoma, longitud y ubicación del gen SCN9A	26
Tabla 2. Identificador de enhancer, valores de Asociación/GenHancer y código de entrada en la BD ENSEMBL	29
Figura 6. Fracción del gen SCN9A y promotor central.	33
Tabla 3. Secuencias motivo del promotor central del gen SCN9A y su valor de <i>e</i> identificados después del análisis con MEME.	34
Figura 7. Identificación de motivos conservados en el promotor central de distintas especies de mamíferos usando el modo “clásico” de <i>MEME Suite</i> .	34
Figura 8. Alineamiento de secuencias del promotor central.	35
Tabla 4. Análisis de los motivos identificados y FTs relacionados.	37
Tabla 5. Principales funciones asociadas a los 12 FTs relacionados al motivo 2 del promotor central.	37
Figura 9. Alineamiento de secuencias y ubicación de cada enhancer, promotor, Sitio de Inicio de la Transcripción, exón 1 y exón 2 en cada especie analizada.	42
Figura 10. Identificación de motivos conservados entre distintas especies de mamíferos usando <i>MEME Suite</i> .	43
Tabla 6. Identificación de motivos conservados en ocho enhancers de distintas especies de mamíferos usando el modo “clásico” de <i>MEME Suite</i> .	44
Figura 11. Descubrimiento de motivos en el enhancer 2 del gen SCN9A usando el modo “clásico” de MEME Suite.	47
Figura 12. Descubrimiento de motivos en el enhancer 6 del gen SCN9A usando el modo “clásico” de MEME Suite.	47
Figura 13. Descubrimiento de motivos en los promotores del gen SCN9A usando el modo “discriminatorio” de <i>MEME Suite</i> .	49
Tabla 7. Secuencias motivo obtenidas mediante <i>MEME</i> y sus distintos	49

alineamientos contra SUFT ya conocidos, utilizando la BD <i>HOCOMOCO</i> .	
Figura 14. Análisis de los niveles de expresión de los FTs relacionados a los SUFTs identificados.	51
Figura 15. Modo “discriminatorio” de <i>MEME</i> Suite en la búsqueda de motivos en el enhancer 6 del gen <i>SCN9A</i> .	52
Tabla 8. Secuencias motivo obtenidas mediante MEME Suite y sus distintos alineamientos contra FT-SUFT ya conocidos, utilizando la BD <i>HOCOMOCO</i> .	52
Figura 16. Análisis evolutivo de la secuencia promotora del gen <i>SCN9A</i> .	60
Figura 17. Análisis evolutivo de la secuencia codificante del gen <i>SCN9A</i>	61
Tabla 9. Distancias patrísticas entre pares secuencias del promotor central del gen <i>SCN9A</i> .	62
Tabla 10. Distancias patrísticas entre pares secuencias de la región codificante del gen <i>SCN9A</i> .	62

1. RESUMEN

El $Na_v1.7$ es un canal de sodio dependiente de voltaje codificado por el gen *SCN9A*, que se expresa principalmente en neuronas sensoriales (nociceptores) del Sistema Nervioso Periférico (SNP). En humanos, la pérdida bialélica de la función lleva a la incapacidad para experimentar dolor, lo que en la clínica se conoce como Insensibilidad Congénita al Dolor asociada a Canalopatías (CIP) (Cox et al, 2010). Sin embargo, dicha mutación es letal en ratones, el principal modelo experimental usado en neurociencias.

El genoma humano está compuesto de regiones codificantes y no codificantes. Éstas últimas representan alrededor de un 98% del total del genoma y comprenden elementos como promotores y *enhancers* (Buroker, 2017). En contraste con lo que antes se pensaba, hoy se sabe que gran parte de la divergencia fenotípica entre especies se debe a cambios en la regulación transcripcional, especialmente a aquella que se presenta en los Sitios de Unión a Factores de Transcripción (SUFT), los cuales se pueden modelar matemáticamente como secuencias motivo.

Aún no es clara la evolución de los elementos reguladores de un gen y su relación con la divergencia fenotípica en la expresión anormal de una proteína (Han et al, 2018). Además, no existen muchos análisis comparativos que analicen estas interacciones entre especies y sus posibles implicaciones en enfermedades que afectan al SN. Por otro lado, existe un número cada vez mayor de secuencias de DNA de distintas especies disponibles en BDs que no aportan información por sí solas sino que deben de ser analizadas con el fin de darles significado en un contexto biológico.

Por lo tanto, en el presente trabajo se realizó un análisis basado en bioinformática y en el estudio de SUFTs de promotores y *enhancers*. Se propuso un estudio comparativo entre distintas especies de mamíferos, con el fin de aportar información y un ejemplo particular que se sume a los datos existentes sobre la evolución de la regulación génica que genera diversidad fenotípica.

El análisis experimental aquí propuesto consistió de la obtención y alineamiento de secuencias correspondientes a 15 especies de mamíferos; la identificación del promotor central y de algunos *enhancers* en cada una de ellas. Posteriormente, se realizó el descubrimiento de motivos en las secuencias no codificantes seleccionadas.

Los resultados obtenidos mostraron las similitudes y diferencias en las secuencias motivo de las distintas especies analizadas. La ganancia o pérdida de SUFTs (representados como secuencias motivo) y los procesos en los que están involucrados podría dar una idea del papel de la evolución de los elementos reguladores de un gen y su relación con la divergencia fenotípica resultado la expresión anormal de una proteína. Por ejemplo, en el fenotipo diferencial entre humano y ratón en el caso particular de pérdida de función del Na_v1.7.

2. INTRODUCCIÓN

2.1 Nocicepción y dolor

Sentir dolor se manifiesta como una experiencia negativa, debido a las señales físicas (y psicológicas) que nos genera. Sin embargo, también es un sistema de alerta que permite a los organismos de distintas especies alejarse de estímulos, eventos u otros individuos que puedan poner en peligro su integridad. Por lo tanto, es necesario para la supervivencia.

De acuerdo a la Asociación Internacional para el Estudio del Dolor, éste es definido como una experiencia sensorial y emocional desagradable, que se asocia con un daño tisular que puede ser real o potencial y que es descrito en términos de dicho daño (Raja *et al*, 2020). Existen dos conceptos fundamentales en la comprensión de la neurobiología de este mecanismo: el dolor en sí mismo y la nocicepción, que aunque están estrechamente relacionados, no son sinónimos.

Nocicepción, es un concepto complejo que se refiere a los mecanismos mediante los cuales un estímulo nocivo es detectado por el Sistema Nervioso Periférico (SNP), codificado y procesado para finalmente emitir una respuesta ya sea a nivel del SNP o del Sistema Nervioso Central (SNC) (Barrot, 2012). Constantemente, los axones de las neuronas aferentes que permiten la detección de estos eventos se encuentran enviando información acerca del entorno a sus cuerpos celulares, que se encuentran localizados en el Ganglio Trigémico (GT) y el Ganglio de la Raíz Dorsal (*Dorsal Root Ganglia*, DRG) (Lopes *et al*, 2017).

Dicho de otro modo, la nocicepción es el procesamiento periférico y central de la información nociva interna o externa que se genera mediante la activación de los nociceptores¹ (National Research Council (US) Committee on Recognition and Alleviation of Pain in Laboratory, 2009).

¹ En la literatura se puede encontrar el uso del término “nociceptor” de dos formas. Por un lado se refiere a las neuronas nociceptivas, es decir, aquellas especializadas en responder a estímulos dañinos. Por el otro lado, algunos autores llaman “nociceptores” a los receptores de las neuronas nociceptivas, que son los encargados de recibir y transducir dicho estímulo. En el presente trabajo se utilizará la primera definición, es decir, el nociceptor como una neurona sensorial especializada en la detección de eventos nocivos.

El dolor, es una reacción subjetiva que requiere de la conciencia y se lleva a cabo a nivel cortical. Comprende la detección, evaluación o percepción de una falla en la protección del cuerpo frente a un estímulo o evento dañinos (reales o potenciales), que están ligados a las experiencias pasadas y recuerdos del individuo (componentes afectivos) (Apkarian, 2019). El dolor, se relaciona con un efecto negativo, por lo que prepara al sujeto para tomar decisiones y buscar soluciones que le permitan evitar la fuente de daño o disminuir dichas sensaciones desagradables.

Resumiendo, la nocicepción permite el procesamiento de la información de un estímulo nocivo (como frío, calor o presión excesivos; lesiones, cortaduras, etc.) desde el SNP hasta el SNC, lo cual se realiza como respuesta a un daño tisular que ha activado los nociceptores internos o externos. La información recibida viajará a través del SNP hasta llegar a la corteza cerebral, donde se dará la percepción consciente y subjetiva del dolor. Todo esto en cuestión de milisegundos (Fig. 1). En conjunto, la nocicepción y el dolor son un sistema de alerta que permite alejarnos de eventos dañinos, lo cual es indispensable en la supervivencia.

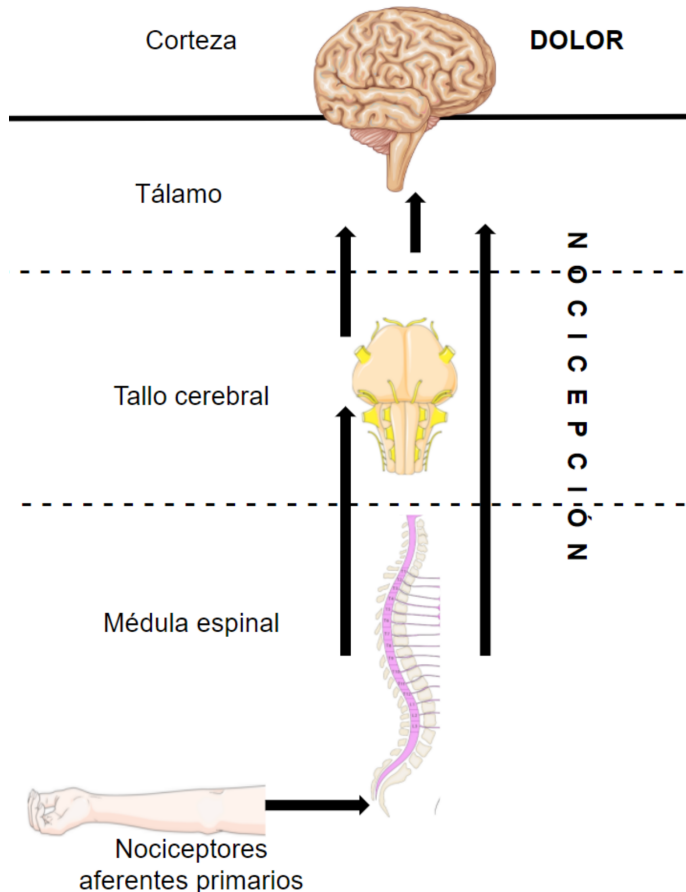


Figura 1. Distribución anatómica de la nocicepción y el dolor.

La nocicepción da inicio en los receptores tisulares mientras que la experiencia del dolor solo se manifiesta hasta que la señal llega a la corteza. Modificado de National Research Council (US) Committee on Recognition and Alleviation of Pain in Laboratory (2009).

2.2 Nociceptores

Los nociceptores son un grupo de células (neuronas) sensoriales especializadas en la transducción de estímulos nocivos. Se clasifican como neuronas pseudounipolares ya que cuentan con una terminación axonal que se divide en una porción periférica (que responde a estímulos potencialmente nocivos) y una rama central, encargada de llevar la información detectada hasta el SNC (a los circuitos de la médula espinal o hasta la corteza) (Dubin & Patapoutian, 2010; Derderian & Tadi, 2021). Los cuerpos celulares de las neuronas nociceptivas que inervan la cabeza, la cara y las estructuras intraorales se encuentran localizadas en el GT, mientras que el DRG recibe las aferencias axonales del resto del cuerpo.

En los axones de los nociceptores se encuentran los Canales de Sodio Dependientes de Voltaje (Voltage-Gated Sodium Channels, VGSCs), los cuales cumplen funciones fundamentales en las neuronas, ya que permiten el inicio y propagación del potencial de acción (Mao *et al*, 2019).

Los VGSCs son complejos proteicos transmembranales formados por una subunidad α (que contiene al poro) y hasta dos subunidades β . A su vez, la subunidad α se compone de cuatro dominios (I-IV), cada uno con seis segmentos transmembranales (S1-S6); el sensor de voltaje se ubica en S4, mientras que S5 y S6 conforman el poro (Yu & Catterall, 2003). Por lo tanto, permite la entrada de iones sodio a la célula como respuesta al cambio de voltaje de la membrana (Gamal El-Din *et al*, 2017); la apertura secuencial de canales de sodio posibilita la creación del Potencial de Acción (Brouwer *et al*, 2014) y la propagación de la información a lo largo de los axones.

Existen diez isoformas identificadas de la subunidad α de los VGSCs: $Na_v1.1$ a $Na_v1.9$ y Na_x (Gamal El-Din *et al*, 2017; Wang *et al*, 2017). Estas isoformas cuentan con propiedades funcionales parecidas, pero cuyas diferencias relacionadas con sus patrones de expresión diferencial, son relevantes para su función. En el SNC se expresan en mayor medida $Na_v1.1$, $Na_v1.2$, $Na_v1.3$ y $Na_v1.6$. Por su parte, $Na_v1.4$ es la isoforma característica en músculo esquelético y $Na_v1.5$ en músculo cardiaco. Finalmente, en el SNP predominan $Na_v1.7$, $Na_v1.8$ y $Na_v1.9$, los cuales están relacionados, entre otras funciones, con la

nocicepción (Gamal El-Din *et al*, 2017). Na_x es un canal atípico no dependiente de voltaje, presente en múltiples regiones, entre ellas, el hipocampo, el SNP, el tálamo, el cerebelo, entre otras (Baroni & Moran, 2015).

2.3 *SCN9A/NA_v1.7* y trastornos del dolor

El $Na_v1.7$ está codificado por el gen *SCN9A* y se expresa predominantemente en el SNP. Por ejemplo, en las neuronas simpáticas y en las neuronas nociceptivas presentes a lo largo de toda la trayectoria del DRG (Black *et al*, 2012); también se encuentra en otras estructuras como los axones del epitelio olfativo que proyectan al bulbo olfativo (McDermott *et al*, 2019). En 2004 se identificó que una mutación en el gen *SCN9A* causa eritromelalgia, una condición en la que se presentan episodios de calor y dolor severo que afectan principalmente a las extremidades (Jha *et al*, 2021). Con este antecedente, continuó el estudio de trastornos del dolor y su relación con *SCN9A/NA_v1.7* y fue en 2006 que se identificó que una mutación de ganancia de función es la responsable de una condición hereditaria de dolor extremo llamada Trastorno Paroxismal de Dolor Extremo (TPDE) (Fertleman *et al*, 2006).

Otro de los estudios pioneros en la relación entre $Na_v1.7$ y nocicepción fue llevado a cabo por Nassar *et al* (2004). Contrario a lo que sucede en el caso de la eritromelalgia y el TPDE, los autores notaron que un ratón condicional (sin canales $Na_v1.7$ funcionales en la mayoría de los nociceptores) presentaba ausencia de dolor inflamatorio. Interesantemente, la idea de usar un ratón condicional surgió del descubrimiento de que la delección global (bialélica) de $Na_v1.7$ es letal en ratones. Pronto se reportó la pérdida de función en humanos, sin embargo, no correspondía a un fenotipo letal como en el caso de los ratones sino que causaba un trastorno del dolor del grupo de las canalopatías conocido como Insensibilidad Congénita al Dolor (ICD) (Nassar *et al*, 2004). La principal característica de estos pacientes es la incapacidad para percibir estímulos nocivos, sin que el tacto, la presión o la temperatura no dañinos estén afectados (Emery *et al*, 2016). Es importante destacar que las personas con ICD también presentan anosmia (Gingras *et al*, 2014).

Las semejanzas que se presentan en estos artículos y, aún más importante, las diferencias que ponen de manifiesto (como el hecho de que la pérdida de función de este canal sea

letal en ratón pero no en humano), llevan obligadamente a la formulación de cuestionamientos acerca del uso de modelos animales como sujetos de investigación y la extrapolación a humanos de los resultados obtenidos. Por otro lado, exponen la importancia de conocer las relaciones evolutivas entre distintas especies que conducen a fenotipos diferentes.

2.4 Secuencias de *DNA* no codificantes

¿Cuántos genes tenemos? fue una de las principales cuestiones que perduraron durante algún tiempo en el campo de la investigación científica. Como consecuencia, los estudios sobre genética humana se centraron únicamente en genes que codifican proteínas o solo en los exones. Por ejemplo, dentro de los principales objetivos del Proyecto Genoma Humano (Hood & Rowen, 2013), destacan:

1. Identificar todos los genes en el ADN humano.
2. Determinar las secuencias de los 3 mil millones de pares de bases nitrogenadas que componen el *ADN* humano.

Sin embargo, en el 2003 (13 años después de su inicio) uno de los resultados más sobresalientes del proyecto puso de manifiesto que la mayoría de las secuencias de *ADN* en humanos no codifican directamente proteínas (98%). La mayoría de la información se encuentra fuera de los exones codificantes y ahora se sabe que está involucrada en la regulación de la expresión génica (Palazzo & Gregory, 2014).

Por otro lado, mediante técnicas de secuenciación se ha mostrado que un 25-50% de regiones codificantes están relacionadas con mutaciones asociadas a enfermedades. Sin embargo, solo estudiar regiones que codifican proteínas, todavía deja sin determinar la etiología génica de un gran porcentaje de padecimientos (Scacheri *et al*, 2015). Por esto, es importante el estudio de las regiones no codificantes como posibles sitios de variabilidad génica

El *DNA* no codificante engloba varios tipos de elementos reguladores como promotores y *enhancers* (Fig. 2). Un promotor está compuesto por el centro del promotor (*core promoter*) y elementos de regulación cercanos o promotor proximal (*proximal promoter*).

El centro del promotor es una región que se localiza al inicio de un gen y que sirve como punto de acoplamiento de la maquinaria transcripcional y de ensamblaje del Complejo de Preiniciación de la Transcripción. También es el lugar que define el Sitio de Inicio de la Transcripción (SIT) y la dirección de ésta. En cuanto al promotor proximal, éste incluye a las secuencias circundantes río arriba inmediatas al centro del promotor (Maston *et al*, 2006).

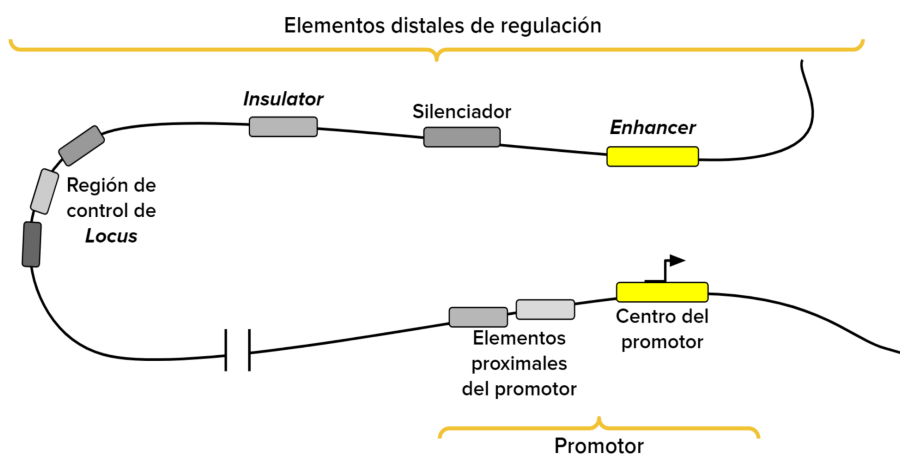


Figura 2. Región reguladora típica de un gen.

Está compuesta por un promotor (central y proximal) y elementos de regulación distales, como los *enhancers*. Modificado de Maston *et al*, 2006.

Los *enhancers* pertenecen al grupo de elementos distales de regulación e indican quién, dónde, cuándo y cuánto se debe de expresar un gen. Es decir, cuáles son los genes que serán expresados en un tipo celular específico, cuándo lo harán y los niveles de tal expresión (Scacheri & Scacheri, 2015). Los *enhancers* pueden regular la transcripción de un gen de dos formas principales (Peng & Zhang, 2018):

- *Enhancers* múltiples → un gen objetivo
- Un *enhancer* → múltiples genes objetivo

Es decir, la actividad transcripcional de un gen es regulada por el efecto de diversos *enhancers* (actuando de forma sinérgica, aditiva, redundante y/o jerárquica). Por otro lado, un solo *enhancer* puede tener efectos reguladores sobre más de un gen (Krijger & de Laat, 2016)

2.5 Secuencias motivo

Los genomas contienen gran cantidad y variedad de patrones estructurados, entre los que se encuentran los Sitios de Unión a Factores de Transcripción (SUFT), que pueden ser representados matemáticamente como Secuencias Motivo o Motivos de *DNA* (Boeva, 2016). Es decir, los SUFTs son patrones de *DNA* cortos y repetitivos que indican sitios de unión específicos para nucleasas o Factores de Transcripción (FT) o que pueden cumplir funciones a nivel de *RNA*. Por otro lado, los Motivos de *DNA* son modelos generados por métodos computacionales que indican secuencias sobrerrepresentadas o conservadas en un gen (D'haeseleer, 2006). Existen muchas Secuencias Motivo predichas sin un Factor de Transcripción asociado, ya sea porque aún no se conoce la relación entre ellos o porque dicho Motivo no es un SUFT. En otras palabras, un SUFT puede describirse como una Secuencia Motivo pero no todas las Secuencias Motivo predichas computacionalmente son SUFTs.

Como se mencionó, los SUFT indican sitios de unión específicos para FTs. Los FTs son proteínas que regulan la expresión génica al unir, de forma directa o indirecta, su Dominio de Unión al DNA con la secuencia objetivo. Es decir, reconocen SUFTs específicos, ya sea en la región del promotor o de manera distal en los *enhancers* (Boeva, 2016).

Algunos autores (Dermitzakis & Clark, 2002; Tuğrul *et al*, 2015) han indicado que con el fin de comprender las diferencias fenotípicas entre especies, poblaciones e individuos, es indispensable el estudio de la evolución de los elementos génicos de regulación. También han planteado que los SUFTs, al ser clave en la regulación de la expresión de un gen, deben estar relacionados a la variación fenotípica (Tuğrul *et al*, 2015) y podrían ser usados en la identificación funcional de elementos reguladores que enlacen las dinámicas evolutivas de distintas especies (Dermitzakis & Clark, 2002).

2.6 Bases de datos biológicas

Las Bases de Datos (BD) de acceso gratuito son de gran importancia para la bioinformática. Éstas contienen información (tanto de humano como de otras especies) sobre proteínas, DNA, RNA, elementos reguladores, variantes, etc. Son una herramienta esencial que permite a los investigadores de ciertas áreas diseñar experimentos e interpretar

sus resultados, así como compartir los de interés con el resto de la población científica (Baxevanis, 2011).

Dentro del gran número y variedad de BD existentes, *Ensembl* es un proyecto bioinformático enfocado en genomas (Hubbard *et al*, 2002; Howe *et al*, 2021), su anotación y la difusión de datos genómicos de una gran cantidad y variedad de vertebrados. Uno de los usos de una BD como *Ensembl* (junto con otras herramientas) es comparar secuencias genómicas e inferir la historia evolutiva de genes y genomas. Otra ventaja que presenta esta BD es su uso alrededor del mundo por compañías e instituciones educativas, así como en dispositivos que van desde supercomputadoras a laptops (Howe *et al*, 2021).

Otra BD ampliamente conocida por bioinformáticos y científicos en general es *GeneCards* (Stelzer *et al*, 2016), que comprende una colección de genes, enfermedades, variantes, proteínas, células y rutas celulares de humano. Como se mencionó previamente, el entendimiento de la regulación del genoma es uno de los principales desafíos científicos, para esto, los *enhancers* son uno de los principales blancos de investigación. Una de las ventajas de *GeneCards* es que ha incorporado *GeneHancer*, una BD de *enhancers* encontrados en el humano, así como los genes objetivo que se han identificado y relacionado a éstos (Fishilevich *et al*, 2017).

2.7 Algoritmos y herramientas bioinformáticas en la identificación de secuencias motivo

Como ya se mencionó en el apartado 2.4, una Secuencia Motivo es un patrón corto y repetitivo de nucleótidos de *DNA* que cumplen con variadas funciones biológicas. Uno de los problemas que se ha planteado la bioinformática es el descubrimiento de Motivos, es decir, la identificación y localización de las secuencias que pueden ser funcionalmente relevantes, como las que corresponden a un SUFT. Ésta no es una tarea sencilla, ya que los SUFTs suelen medir entre 6-12 pb, además, las secuencias intergénicas son muy largas y variables (Hashim *et al*, 2019), lo que puede representar un problema metodológico y de análisis. La técnica general de descubrimiento de SUFTs consta de tres principales etapas: pre procesamiento, descubrimiento y post procesamiento (Fig. 3).

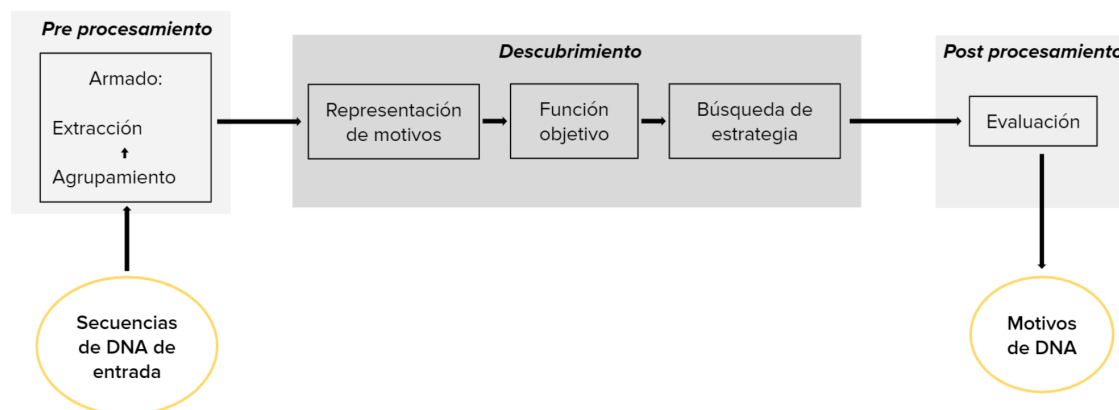


Figura 3. Técnica general empleada en el descubrimiento de motivos.

Consta de tres principales etapas: pre procesamiento, descubrimiento y post-procesamiento. Modificado de Hashim *et al*, 2019.

Existen diversas aproximaciones que han utilizado los tres pasos de la técnica con el fin de dar solución al problema de descubrimiento de motivos. En resumen, hay cuatro enfoques principales (Hashim et al, 2019):

1. Numerativo: Consiste en el conteo y comparación de las frecuencias de oligonucleótidos en todos los posibles modelos de Motivos. Fue la primera en utilizarse y, a su vez, se divide en simple, de agrupación, de corte, entre otros.
2. Probabilístico: La probabilidad de que un conjunto de nucleótidos sea una Secuencia Motivo se obtiene de la multiplicación de cada una de las probabilidades de que un nucleótido esté presente en una posición determinada dentro de la secuencia. Es uno de los más aceptados y se subclasifica en avanzado, determinista, estocástico y otros.
3. Inspirados en la naturaleza: Enfoque que simula el comportamiento de insectos u otros animales con el objetivo de encontrar solución al problema del descubrimiento de Motivos. Está compuesto por algoritmos de inteligencia de enjambre, bioinspirados, basados en fenómenos físicos o químicos, evolutivos, entre otros.
4. Combinatorio: Las herramientas combinatorias son las más recientes y consisten en la unión de múltiples algoritmos, su desempeño depende de la combinación seleccionada.

El método probabilístico más popular es *MEME* (*Multiple Expectation–Maximization for Motif Elicitation*) de *MEME Suite* (Bailey *et al*, 2009; Bailey *et al*, 2015). Primero, *MEME* encuentra una secuencia motivo inicial y después utiliza una serie de pasos de un algoritmo determinista (Esperanza-Maximización) con el objetivo de mejorar el motivo identificado inicialmente. Este proceso ocurre hasta que los valores de una matriz peso posición ya no mejoran o hasta que se ha alcanzado el número de iteraciones establecido (Bailey *et al*, 2015; Hashim *et al*, 2019).

3. JUSTIFICACIÓN

Aún no es clara la evolución de los elementos reguladores de un gen y su relación con la divergencia fenotípica en la expresión anormal de una proteína (Han *et al*, 2018). Además, no existen muchos análisis comparativos que analicen estas interacciones entre especies y las posibles implicaciones en enfermedades que afectan al SN. Por otro lado, existe un número cada vez mayor de secuencias de *DNA* de distintas especies disponibles en BDs que no aportan información por sí solas sino que deben de ser analizadas con el fin de darles significado en un contexto biológico.

Por lo tanto, en el presente trabajo se propone un análisis basado en bioinformática y en el estudio de SUFTs de promotores y *enhancers*. Se propone un estudio comparativo entre distintas especies de mamíferos, con el fin de aportar información y un ejemplo particular que se sume a los datos existentes sobre la evolución de la regulación génica que genera diversidad fenotípica.

4. OBJETIVOS

4.1 General

- Realizar un análisis comparativo de secuencias no codificantes asociadas al gen *SCN9A* en distintos mamíferos e identificar posibles Sitios de Unión a Factores de Transcripción (SUFT) dentro de elementos reguladores (como promotores y *enhancers*) que puedan estar implicados en la regulación de la expresión diferencial de éste.

4.2 Particulares

1. Seleccionar especies de mamíferos de los tres órdenes de la clase *Mammalia* y excluir aquellas que no cuenten con datos de la secuencia del gen *SCN9A* en la BD *Ensembl*.
2. Descargar de la BD *Ensembl* las secuencia del gen *SCN9A* de cada uno de los mamíferos incluidos en el estudio, así como las regiones río arriba al sitio de inicio de la transcripción.
3. Alinear todas las secuencias obtenidas usando *VISTA Tools*, comparar las similitudes y diferencias entre éstas e identificar las regiones de interés así como los sitios conservados.
4. Identificar Secuencias Motivo o SUFTs en la región del promotor central usando *MEME Suite*.
5. Identificar los *enhancers* relacionados al gen *SCN9A* anotados en la base de datos *GeneCards*
6. Identificar y analizar secuencias motivo dentro de los *enhancers* asociados al gen *SCN9A*.
7. Realizar un análisis de la relación entre los Sitios de Unión a Factores de Transcripción identificados y los Factores de Transcripción ya conocidos. Comparar los resultados obtenidos entre especies.

5. ESTRATEGIA EXPERIMENTAL

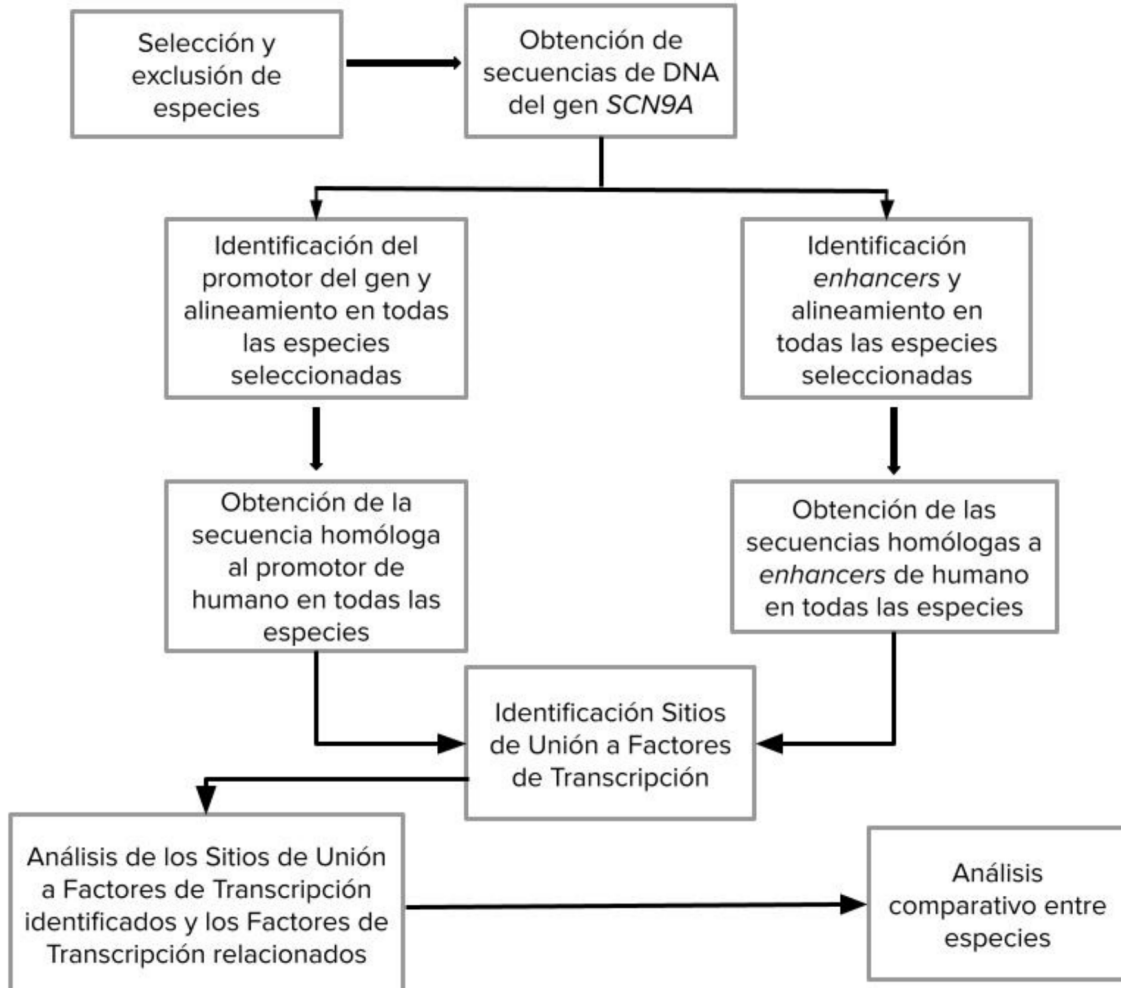


Figura 4. Estrategia experimental empleada en la realización de este proyecto.

6. MÉTODOS

6.1 Selección y exclusión de secuencias de distintas especies de mamíferos

La selección de especies a analizar se realizó con base en la información filogenética obtenida de *Tree of life web project* (<http://tolweb.org/Mammalia/15040/1995.01.01>), de donde se seleccionó la clase *Mammalia*, con el fin de incluir representantes de cada orden, es decir, Monotremata, Marsupialia y Eutheria (Fig. 5a). Debido a la ausencia de secuencias del gen *SCN9A* de marsupiales tanto en *Ensembl* como en *GeneBank*, se excluyó dicho orden. Posteriormente, como el grupo más numeroso fue el *Eutheria* (mamíferos placentarios), se eligieron especies representantes de cada uno de sus subórdenes (Fig. 5b). Algunos de éstos se excluyeron ya que no existen representantes en las BDs analizadas o falta el gen de interés.

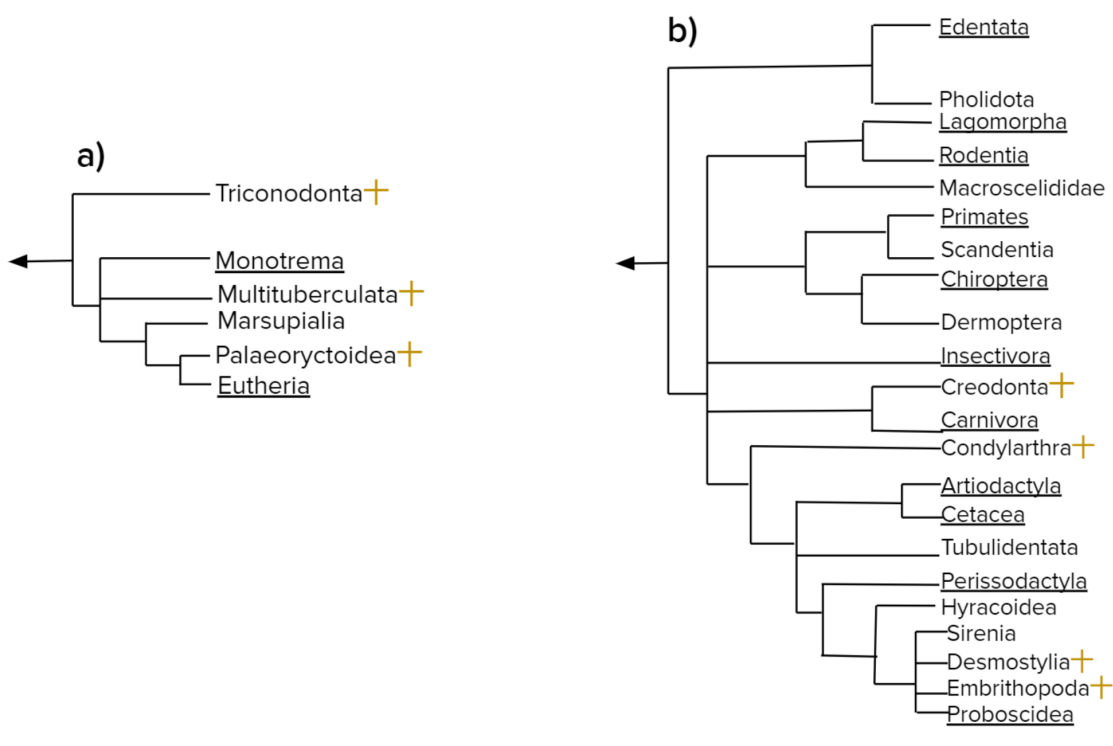


Figura 5. Información filogenética de la clase *Mammalia* y el orden *Eutheria*

a) Órdenes de la clase *Mammalia* incluidos en el presente estudio. Se excluyó al orden *Marsupialia* debido a que en la BD usada no se encontró el gen de interés. **b)** Debido a que *Eutheria* es el grupo con más representantes, la mayoría de las especies utilizadas (subrayado) se seleccionaron de este conjunto. Las cruces amarillas hacen referencia a los grupos extintos. Modificado de *The Tree of Life Web Project* (2005).

Dos especies que se habían incluido inicialmente (ornitorrinco y armadillo) se eliminaron del estudio ya que las regiones relevantes que sirvieron para realizar las distintas

comparaciones no están secuenciadas adecuadamente. Finalmente, las secuencias de las especies presentadas a continuación, son las que se eligieron para continuar con el análisis: *Homo sapiens* (humano), *Felis catus* (gato), *Bos taurus* (vaca), *Canis lupus familiaris* (perro), *Tursiops truncatus* (delfín), *Loxodonta africana* (elefante), *Equus caballus* (caballo), *Macaca mulatta* (macaco), *Pteropus vampyrus* (murciélago), *Heterocephalus glaber* (ratopín), *Mus musculus* (ratón), *Sus scrofa domestica* (cerdo), *Ursus maritimus* (oso polar), *Oryctolagus cuniculus* (conejo), *Rattus norvegicus* (rata).

6.2 Obtención de secuencias desde la Base de Datos *Ensembl*

Las secuencias de *DNA* del gen *SCN9A* para cada especie se obtuvieron de la base de datos *Ensembl* (Howe *et al.*, 2021). Se descargaron un total de 15 secuencias del gen *SCN9A* correspondientes a 15 especies de mamíferos. Cada una de ellas contiene a los intrones y exones del gen y 180,000 pares de bases (pb) río arriba del inicio del primer exón codificante. En la tabla 1 se muestra la localización y longitud de cada secuencia elegida, así como el cromosoma al que pertenecen.

A continuación se muestran los pasos que se siguieron en la obtención de secuencias:

1. Ingresar a <https://www.ensembl.org/index.html> desde cualquier servidor web.
2. En el apartado *Search*, buscar y seleccionar la especie de interés, así como escribir el nombre del gen de interés (*SCN9A*).
3. Exportar la secuencia de interés en formato *FASTA*, incluyendo intrones, exones y el número de nucleótidos río arriba del SIT previamente seleccionados.
4. Descargar en un archivo de texto

Especie/ID	Chr	Coordenadas	Especie	Chr	Coordenadas
Homo sapiens (humano) ENSG00000169432.18	2	166,195,185-166,376,001	Pteropus vampyrus (murciélago) ENSPVAG00000002662.1	8	57,031,736-57,107,365
Felis catus (gato) ENFCAG00000000647.6	1	157,823,201-157,923,727	Heterocephalus glaber (ratopín) ENHGLG00100005815.1	1	3,997,887-4,085,449
Bos taurus (vaca) ENBTAG000000002425.6	2	29,996,676-30,088,178	Mus musculus (ratón) ENSMUSG000000075316.11	2	66,310,424-66,465,306
Canis lupus familiaris (perro)	36	11,290,625-11,	Sus scrofa domestica (cerdo)	1	36,171-133,760

ENSCAFG00000011438.5		383,255	ENSSSCG00000015913.4		
Tursiops truncatus (delfin) ENSTTRG00000003797.1		9,418-92,813	Ursus maritimus (oso polar) ENSUMAG000000013160.1	1	15,218,430-15,318,302
Loxodonta africana (elefante) ENSLAFG00000003123.4	3	61,042,161-61,145,632	Oryctolagus cuniculus (conejo) ENSOCUG00000005623.4	7	105,663,854-105,768,147
Equus caballus (caballo) ENSECAG00000022007.3	18	46,262,385-46,350,792	Rattus norvegicus (rata) ENSRNOG00000006639.6	3	52,583,951-52,664,209
Macaca mulatta (macaco) ENSMMUG00000018510.4	12	53,798,041-53,907,865			

Tabla 1. Cromosoma, longitud y ubicación del gen *SCN9A* (y secuencias no codificantes relacionadas) en la BD *ENSEMBL*. Junto a cada especie seleccionada se muestra el ID de entrada para la BD *ENSEMBL* (negritas).

5.3 Alineamiento y comparación de secuencias conservadas

Con el fin de identificar regiones de interés dentro de las secuencias previamente seleccionadas, se realizó un análisis comparativo de las mismas con ayuda del servidor *VISTA Tools* en su versión *mVISTA* (Mayor *et al*, 2000). El principal análisis se realizó al promotor central del gen *SCN9A*, que no se conoce en otras especies pero se encuentra identificado en el genoma del humano (Diss *et al*, 2007). Por lo tanto, se llevó a cabo una alineación del promotor central de humano con el resto de las secuencias, con el fin de identificar la región homóloga en cada mamífero seleccionado.

El procedimiento fue el siguiente:

1. Ingresar a <https://genome.lbl.gov/vista/index.shtml>
2. Seleccionar la opción *mVISTA* y elegir el número de secuencias a analizar. *mVISTA* realiza alineaciones de secuencias de hasta megabases de dos o más especies (Frazer *et al*, 2004).
3. Ingresar cada secuencia de cada especie en formato *FASTA*, etiquetar cada documento y añadir un archivo de anotación en formato *FASTA* (ubicación y sitios de interés conocidos en humanos) que servirá como referencia al alinear con el resto de las especies.

4. Visualizar los resultados y extraer las regiones de interés, es decir, las regiones homólogas al promotor en humano y aquellas que estén conservadas por arriba del nivel 70% por cada 100 pb.

6.4 Identificación de secuencias motivo en la región del promotor

Como se mencionó en la introducción, en el problema del descubrimiento de motivos de DNA se han propuesto algunas herramientas bioinformáticas basadas en uno o la combinación de dos algoritmos de análisis.

Dentro de la aproximación probabilística *MEME* de *MEME Suite* (Bailey *et al*, 2009; Bailey *et al*, 2015), es una herramienta de tipo determinista que usa el método de Esperanza-Maximización con el objetivo de mejorar una matriz de peso posición a través de un cierto número de iteraciones.

1. Desde un servidor web ingresar a <https://meme-suite.org/meme/>
2. Seleccionar *MOTIF DISCOVERY* → *MEME*
3. Elegir el modelo clásico o discriminatorio según sea el caso.
4. Cargar un archivo tipo *FASTA* que contenga la región de interés seleccionada previamente del programa *VISTA Tools* para cada una de las especies.
5. Configurar el resto de las opciones de interés (distribución, número de motivos y tamaño, ver abajo)

MEME Suite es una herramienta que permite la identificación de motivos entre secuencias. La distribución de los motivos encontrados se realiza de tres maneras: cero o una ocurrencia por secuencia (*Zero or One Occurrence per Sequence -zoops-*), una ocurrencia por secuencia (*One Occurrence Per Sequence -oops-*) o cualquier número de repeticiones (*Any Number of Repetitions -any-*). Se eligió trabajar con *zoops* ya que esta opción es recomendada cuando se sospecha que algunos motivos pueden no estar presentes en todas las secuencias. Además, se pidió identificar 5 motivos en un rango de 6-12 nt en las secuencias de interés. Por otro lado, se utilizaron dos métodos de descubrimiento de motivos: clásico y discriminatorio. El modo clásico descubre motivos enriquecidos en un conjunto de secuencias de entrada, mientras que el discriminatorio encuentra motivos enriquecidos en un *set* respecto a un segundo conjunto. En la sección de “Resultados y Discusión” se amplía el concepto y se explica cuando se utilizó un método o el otro.

6.6 Selección y exclusión de enhancers relacionados al gen *SCN9A*

Los *enhancers* son secuencias que tienen un papel fundamental en la regulación de la expresión tejido específica de un gen, muchas veces desde una distancia lejana al promotor y de forma independiente a la orientación (*Jin et al*, 2011).

Para identificar los posibles *enhancers* que regulan al gen *SCN9A* se siguieron los siguientes pasos:

1. Búsqueda de los *enhancers* asociados al gen *SCN9A* de humano almacenados en la base de datos *GeneCards* (*Stelzer et al*, 2016). En *GeneCards*, se encuentra una colección de 40 *enhancers* relacionados al gen de interés, sin embargo, solo se seleccionaron aquellos con mayor puntuación de Asociación (GA) o de *GeneHancer* (GH). El valor de GH es calculado con base en una combinación de anotaciones, a saber, del número de fuentes que hacen referencia a éste, de la puntuación de la fuente y de los FT y SUFTs relacionados. Por su parte cada GA se representa por la puntuación del *GeneHancer*, de los métodos que le dan soporte y de la lista de todos los genes asociados a éste (*Fishilevich et al*, 2017). De este modo, se excluyeron 32 *enhancers* y se seleccionaron los 8 con mayor calificación en las dos medidas antes mencionadas (Tabla 2).

Número de enhancer e ID	GH score	GA score	Entrada en <i>ENSEMBL</i>	Número de enhancer e ID	GH score	GA score	Entrada en <i>ENSEMBL</i>
1.-GH02J166371	2.2	272.9	chr2:166371201-166371600	5.-GH02J166344	0.9	10.6	chr2:166343402-166345399
2.-GH02J166493	1.6	6.2	chr2:166493668-166494979	6.-GH02J166233	0.9	0.1	chr2:166233202-166235399
3.-GH02J166410	1.2	6.9	chr2:166410401-166414000	7.-GH02J166181	0.9	0.1	chr2:166181401-166182012
4.-GH02J166229	1.1	0.1	chr2:166229602-166231286	8.-GH02J166360	0.8	15.4	chr2:166360541-166362200

Tabla 2. Identificador de *enhancer*, valores de Asociación/*GenHancer* y código de entrada en la BD *ENSEMBL* para cada uno de los ocho *enhancers* incluidos en el análisis.

6.7 Análisis e identificación de secuencias motivo dentro de los enhancers seleccionados.

La identificación y procesamiento de los *enhancers* seleccionados se realizó de forma similar a lo descrito anteriormente para el caso del promotor, siguiendo los pasos listados a continuación:

1. Identificar los ocho *enhancers* en cada especie y descargar las secuencias correspondientes de la BD *Ensembl*.
2. Realizar un alineamiento de secuencias en formato FASTA para cada *enhancer*, incluyendo todas las especies (*VISTA Tools*).
3. Incluir en una anotación, el lugar en que se ubica cada enhancer en humano y, tomando al humano como referencia, determinar las regiones homólogas en el resto de las secuencias (*MEME*).

6.8 Análisis de la relación entre las secuencias motivo identificadas y los factores de transcripción ya conocidos.

Luego de la identificación de los motivos conservados entre especies, se analizó si estas regiones corresponden a SUFTs conocidos. Para esto, se estudió cada motivo encontrado tanto en las secuencias promotoras como en los *enhancers*, usando la herramienta *Tomtom* de *MEME Suite 5.1.1*. *Tomtom* compara uno o más motivos contra una base de datos de motivos ya conocida. La base de datos seleccionada fue *HOCOMOCO Human and Mouse* (v11 core) (HOMO sapiens COMprehensive MODEL COLLECTION) que contiene modelos de sitios de unión a factores de transcripción, 680 factores de transcripción para humano y 453 para ratón (Kulakovskiy *et al*, 2017).

Partiendo de la idea de que los métodos de descubrimiento de motivos clásico y discriminatorio proporcionaron distintos resultados, se realizó una búsqueda de cada uno de los SUFTs encontrados para cada secuencia motivo. Se descartaron aquellos que, tras su inspección en el Atlas de Expresión de Tejidos del servidor *The Genotype-Tissue Expression* (GTEx) (Keen & Moore, 2015) no mostraron una expresión marcada en el SN.

7. RESULTADOS Y DISCUSIÓN

7.1 Identificación del promotor central del gen *SCN9A* en distintas especies de mamíferos.

Con el objetivo de identificar el promotor central del gen *SCN9A* en los distintos mamíferos, se utilizó la información reportada por Diss *et al* (2007), quienes encontraron el centro de la región promotora (*core promoter*) del gen *SCN9A* de humano ~64,000 pb río arriba del SIT. Éste comprende una región de 345 nucleótidos (nt) localizado en el inicio del exón 1 (-266 a +80). De este modo, se usó esta información para buscar esa ubicación dentro de la comparación realizada con *VISTA Tools*. Una vez identificada, se obtuvieron las secuencias correspondientes de ~345 pb para cada especie, que coincidieron con la posición encontrada en el humano.

VISTA Tools es una herramienta bioinformática que ha planteado un método de identificación de elementos funcionales en los genomas mediante el alineamiento y comparación de múltiples secuencias de *DNA* que provienen de diferentes especies (Frazer *et al*, 2004). Las ventajas que plantea *VISTA Tools* frente a otros programas de su tipo es que posee un algoritmo de alineamiento eficiente y proveé una interfaz de visualización de resultados fácil de interpretar, razones por las que se eligió utilizar esta herramienta con el fin de realizar alineaciones entre secuencias.

En la figura 6a se muestra la comparación de las secuencias de los distintos mamíferos con el gen de humano realizada con el programa *VISTA Tools*. Cada fila representa la región del gen *SCN9A* seleccionada para cada especie y el alto de las “montañas” representa el nivel de conservación. El segmento presentado abarca la secuencia río arriba al inicio de este gen, las regiones no codificantes conservadas (rojo) y los tres primeros exones (morado), de los cuales, el primero es no codificante. Dentro de las secuencias no codificantes, el UTR se muestra en color azul. El promotor se localiza entre el final del UTR y el inicio del primer exón no codificante (Figura 6b), abarcando estas dos regiones. Nótese la falta de sitios conservados (flechas rojas) en la región que se alineó con la correspondiente al promotor de humano en las secuencias de armadillo y ornitorrinco, mismas que, después de confirmar la falta de sitios conservados, fueron eliminadas del

resto de los análisis. También resalta la menor similitud en general que se puede observar en las secuencias de rata y ratón.

7.2 Análisis de motivos en el promotor central

El programa *MEME* de *MEME Suite* descubre motivos nuevos sobrerrepresentados estadísticamente en las secuencias de entrada, los cuales se presentan como matrices de probabilidad, es decir, describe qué tan probable es que una determinada letra se presente en una posición específica dentro de un patrón de la secuencia (Bailey *et al.*, 2015). Se usó *MEME* para encontrar motivos en el promotor central del gen *SCN9A* de todas las especies. *MEME*, devuelve tantos motivos como sean requeridos (5 en este caso) y para cada uno de ellos presenta un valor de e (e-value) (Tabla 3), un sitio y un ancho. En el caso de la significancia estadística, *MEME* generalmente encuentra primero los motivos más significativos (bajo valor de e). Por otro lado, en el descubrimiento de motivos, devuelve valores de p (p-value), la secuencia completa del motivo identificado así como el inicio y fin del mismo. En la figura 7 podemos identificar estas características, por ejemplo, se presentan cinco cajas de distinto color para cada uno de los motivos encontrados, así como su posición relativa dentro de las secuencias. Las líneas sobre las que estas cajas se encuentran, indican el largo de la secuencia.

La representación de un motivo realizada por *MEME* es en realidad una secuencia de consenso. Es decir, el programa construye un motivo a través del uso repetido de matrices peso posición y como resultado se obtiene una letra que constituye el nucleótido más frecuente en todas las secuencias de entrada (Liljas, 2013). Por lo tanto, el modelo de motivo final puede diferir de las secuencias reales de cada especie. En la figura 8 se muestra el alineamiento de los promotores centrales, la secuencia motivo correspondiente a cada especie y el motivo consenso entre secuencias, así como su ubicación dentro del promotor.

En el ratón y la rata solo se identificaron cuatro de los cinco motivos requeridos. El motivo faltante, en contraste con el humano, es el número 2 (Figura 7, caja de color azul). Según el valor de e calculado, este es el segundo motivo más significativo del análisis (Tabla 3).

Como se mencionó, existen marcadas diferencias fenotípicas entre el humano y el ratón debido a la pérdida de función del producto del gen *SCN9A*. En el humano, el resultado es unacanalopatía llamada Insensibilidad Congénita al Dolor (ICD), mientras que dicho genotipo es letal en ratones.

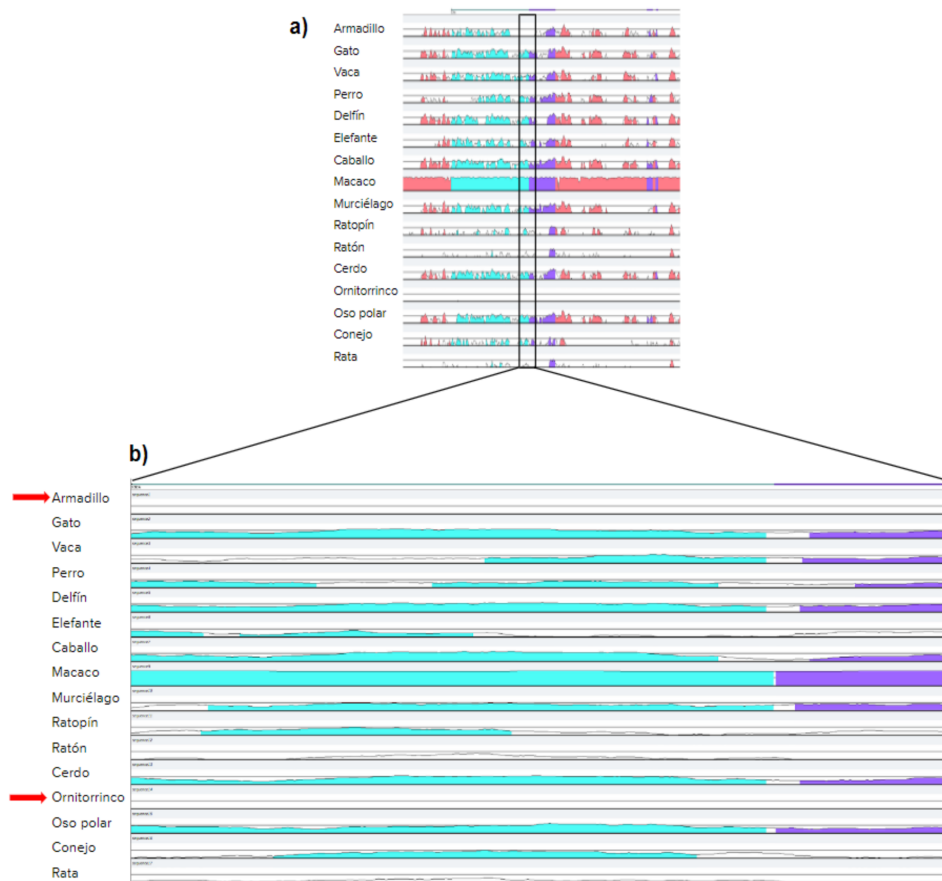


Figura 6. Fracción del gen *SCN9A* y promotor central.

a) Fracción del gen *SCN9A* en donde se observa la región UTR (azul) y algunos exones (morado), para cada una de las especies. **b)** Región estimada de la zona reportada como secuencia promotora del gen *SCN9A* por Diss *et al* (2007) la cual corresponde a una región de 345 pb, ubicada al inicio del exón 1 (-266 a +80). Nótese la falta de regiones conservadas en las secuencias de armadillo y ornitorrinco, mismas que fueron excluidas de análisis posteriores.






Secuencia motivo	Valor de e
1.- 	$9.6e^{-021}$
2.- 	$6.3e^{-017}$
3.- 	$2.3e^{-015}$
4.- 	$5.6e^{-014}$
5.- 	$1.7e^{-012}$

Tabla 3. Secuencias motivo del promotor central del gen *SCN9A* y su valor de e identificados después del análisis con *MEME*.

Los motivos más significativos corresponden a un bajo valor de e . El tamaño de las letras coincide con el nivel de alineamiento con las secuencias de entrada, es decir, con la probabilidad de que ese nucleótido se presente en una posición específica dentro de la secuencia.

Na_v1.7 cumple la misma función en ambas especies. Sin embargo, la conservación funcional de la expresión génica, podría no significar la conservación de elementos reguladores en secuencias no codificantes, pudiendo ser importantes en el entendimiento de fenotipos diferenciales entre especies debido a mutaciones de pérdida de función.

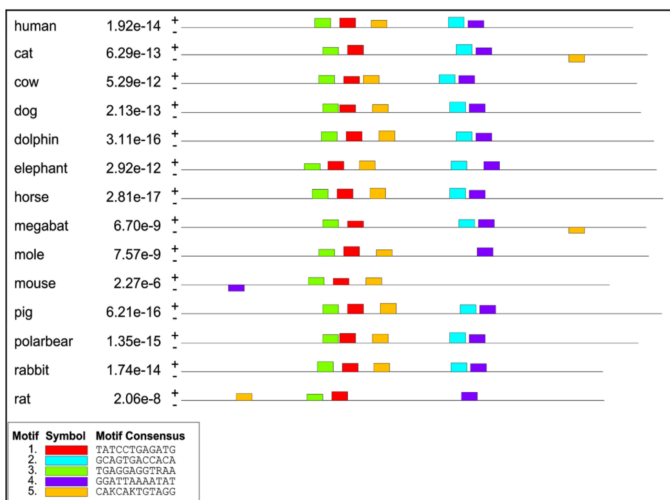


Figura 7. Identificación de motivos conservados en el promotor central de distintas especies de mamíferos usando el modo “clásico” de *MEME Suite*.

Las especificaciones dadas al programa fueron la búsqueda de 5 motivos con una longitud de entre 6 y 12 nt que tuvieran como máximo una repetición por secuencia (zoops). De izquierda a derecha: ID de la secuencia motivo, valor de p , ubicación de los motivos. Mole = ratopín (*naked mole rat*). Por el gran parecido con el humano, en este análisis se obvió la secuencia de macaco.

También se puede notar que, en comparación con el humano y la mayoría de las especies analizadas, los motivos numerados como 4 y 5 se encuentran en una posición distinta en ratón y rata respectivamente, además de que se observan en la cadena complementaria (motivo 4 de ratón) (Figura 7, cajas de color morado y amarillo). En el resto de las especies, con excepción de la rata, que es otro modelo murino característico, el número y ubicación de las secuencias motivo es muy parecido. En cuanto a la orientación de la secuencia motivo, Lis & Walther (2006) se cuestionaron si un mismo motivo podría tener efectos diferenciales en la regulación de la expresión génica dependiendo de si se encontraba en la hebra 5'-3' o en la 3'-5'. Estudiaron este fenómeno al mapear la orientación de motivos cercanos al SIT en los genes nucleares de *Arabidopsis thaliana*, mediante métodos estadísticos y bioinformáticos, y encontraron que independientemente de su orientación, la presencia de motivos se asoció con efectos pronunciados en la expresión de los genes en estudio. Es decir, es el evento de unión FT-SUFT el que provoca cambios en la regulación de un gen independientemente de su orientación y que la ubicación del mismo, en cualquiera de las dos cadenas no tuvo efectos funcionales significativos.

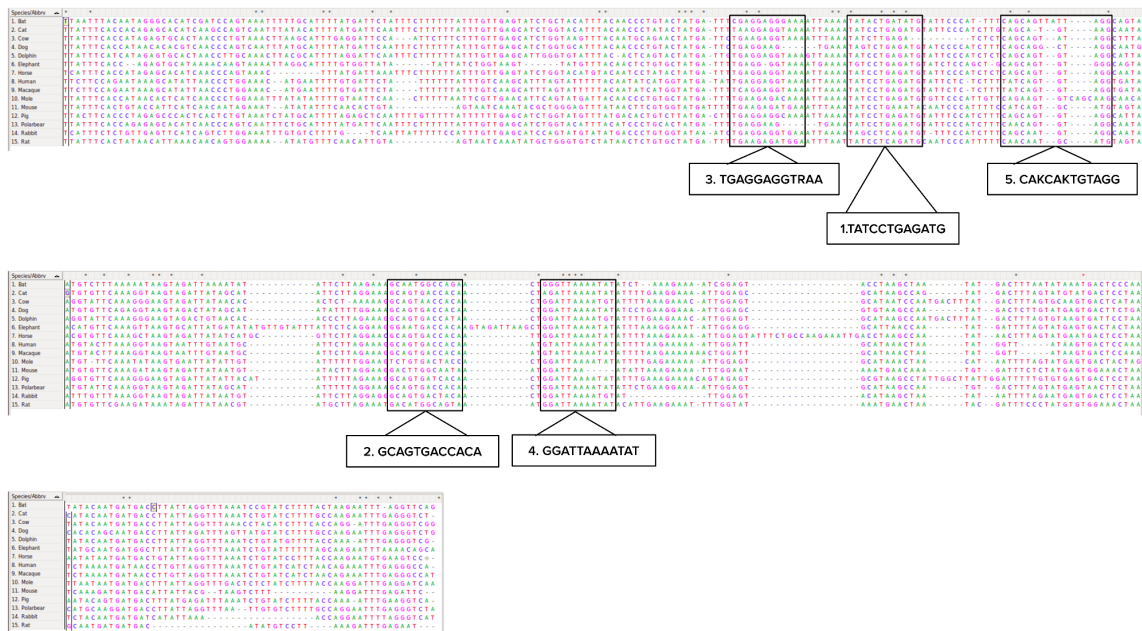


Figura 8. Alineamiento de secuencias del promotor central. Las secuencias correspondientes al promotor central del gen *SCN9A* se alinearon una vez más usando la herramienta de alineamientos múltiples *MUSCLE* incluida en *MEGA X*. En los recuadros negros se muestra la secuencia real de los motivos consenso obtenidos de *MEME*.

Luego de la identificación de los motivos conservados entre especies se analizó si estos corresponden a Sitios de Unión a Factores de Transcripción (SUFT) ya reportados en la literatura. Para esto, se estudió cada motivo por separado usando la herramienta *Tomtom* de *MEME Suite 5.1.1* (Bailey *et al*, 2015). *Tomtom* compara uno o más motivos contra una base de datos de motivos ya conocida. La base de datos seleccionada fue *HOCOMOCO Human (v11 core)* (HOmo sapiens COmprehensive MOdel COllection) que contiene modelos de sitios de unión a factores de transcripción: 680 factores de transcripción para humano y 453 para ratón (Kulakovskiy *et al*, 2017). El resultado de este análisis se muestra en la tabla 4. El quinto motivo no arrojó alineamientos en esta comparación, es decir, o se trata de un SUFT aún no identificado en la literatura o es un falso positivo.

En la tabla 5 se resume la función de cada uno de los FTs relacionados al motivo 2, que está ausente en las secuencias de rata y ratón del análisis realizado a los promotores. Las funciones de regulación descritas son variadas y sin algún patrón aparente.

Recientemente, Ha *et al* (2022) estudiaron la divergencia evolutiva de las relaciones reguladoras entre los FT y los genes diana en módulos funcionales, concluyendo que existe un recableado de las redes reguladoras de genes, lo cual contribuye a las diferencias que se encuentran cuando, por ejemplo, se utilizan ratones como modelos en la reproducción de enfermedades humanas y no se obtienen los mismos fenotipos. Es decir, pese a que las secuencias entre humano y ratón se puedan encontrar conservadas, la divergencia funcional entre genes ortólogos ha surgido frecuentemente durante la evolución. La ausencia del motivo 2 en esta comparación, podría sugerir una pérdida o ganancia de un SUFT en ratón y humano respectivamente y ser un ejemplo concreto del proceso descrito por estos autores.

Aún no es claro como la evolución de secuencias reguladoras impacta en la divergencia fenotípica. En este caso, en el fenotipo diferencial resultado de la expresión anormal de *Na_v1.7* entre humano y ratón. Sin embargo, en línea con lo propuesto por Ha *et al* (2022) y tomando en cuenta los resultados aquí presentados, una posible explicación es que la ganancia o pérdida de un SUFT o secuencia motivo es el evento que precede al recableado de la redes reguladoras, las cuales se encontraron específicas a cada especie.

Adicionalmente, los autores proponen que elementos como promotores y *enhancers* también están involucrados e interactúan entre sí en el proceso de reconexión de secuencias reguladoras. El análisis aquí propuesto permite observar en qué especies se han conservado los motivos localizados, dando una idea de cuáles secuencias motivo y en qué especies ha habido un proceso de divergencia, pese a que estén altamente emparentadas. Entender estos mecanismos de expresión fenotípica diferencial es relevante en la traslación de resultados usando modelos animales como ratones. Adicionalmente, el análisis bioinformático que se llevó a cabo, podría servir de precedente con el fin de guiar las investigaciones que usan experimentos de laboratorio y modelos animales con el fin de reducir la búsqueda, especificar las condiciones y ahorrar tiempo, dinero y recursos.

Secuencia motivo o SUFT	Número de alineamientos significativos con la base de datos	Factores de Transcripción alineados
1 (TATCCTGAGATG)	11	STAT6, PRDM6, TP53, TP73, THRB, SIX2, P63, SIX1, SRF, Kaiso (ZBTB33), TCF7L1
2 (GCAGTGACCACA)	12	NR1H3, ATF1, CREM, RXRB, CBF, CREB1, VDR, FOXA2, RUNX1, NR1H4, FOXM1, ZNF264
3 (TGAGGAGGTRAA)	31	SPIB, IRF4, BCL11A, SPI1, IRF8, RARA, E2F7, ZNF708, ZBTB17, ZNF586, TBX21, ETV5, DUX4... Solo se muestran los primeros 13
4 (GGATTAATAATAT)	13	ZNF586, DUX4, POU1F1, MEF2A, ALX1, NFATC4, FOXQ1, PGR, TWIST1, HOXC9, OTX2, ZNF329
5 (CAKCAKTGTAGG)	0	

Tabla 4. Análisis de los motivos identificados y FTs relacionados.

Identificador de cuatro de las cinco secuencias motivo encontradas por la herramienta *MEME* y que arrojaron alineamientos de SUFT de la base de datos *HOCOMOCO Human (v11 core)*.

Factor de Transcripción (gen/proteína)	Papel biológico
NR1H3	Forma un heterodímero con el receptor X de retinoides y regula la expresión de genes diana que contienen elementos de respuesta a los retinoides (<i>GeneCards</i> , 2022).
ATF1	Influye en los procesos fisiológicos celulares mediante la regulación de la

	expresión de genes diana río abajo, que están relacionados con el crecimiento, la supervivencia y otras actividades celulares (<i>GeneCards</i> , 2022).
CREM	El gen codifica un factor de transcripción (ZIPb) que se une al elemento de respuesta AMPc, presente en muchos promotores virales y celulares. Es un componente importante de la transducción de señales mediada por AMPc durante el ciclo espermatozoario. El uso de sitios de inicio de traducción y promotores alternativos permite que este gen ejerza especificidad espacial y temporal para la capacidad de respuesta de AMPc. Se han encontrado múltiples variantes de transcripción empalmadas alternativamente que codifican varias isoformas diferentes para este gen, algunas de las cuales funcionan como activadores y otras como represores de la transcripción (<i>GeneCards</i> , 2022).
RXRβ	Forma homodímeros con los receptores de ácido retinoico, hormona tiroidea y vitamina D, aumentando tanto la unión al ADN como la función transcripcional en sus respectivos elementos de respuesta (<i>GeneCards</i> , 2022).
CBFB	La proteína codificada por este gen es la subunidad beta de un factor de transcripción heterodimérico de unión al núcleo, regula genes específicos de hematopoyesis y osteogénesis. La subunidad beta es una subunidad reguladora que no se une al ADN sino que mejora alostéricamente la unión del ADN por la subunidad alfa, ya que el complejo se une al sitio central de varios <i>enhancers</i> y promotores (<i>GeneCards</i> , 2022).
CREB1	Miembro de la familia de proteínas de unión al ADN con cremallera de leucina (<i>leucine zipper</i>). Induce la transcripción de genes en respuesta a la estimulación hormonal de la vía AMPc (<i>GeneCards</i> , 2022).
VDR	Miembro de los receptores de hormonas nucleares de FTs inducidos por ligandos. Las diana río abajo del receptor de vitamina D3 están principalmente involucradas en el metabolismo mineral, aunque éste regula una variedad de otras vías metabólicas, como las involucradas en la respuesta inmune y el cáncer (<i>GeneCards</i> , 2022).
FOXA2	Miembro de la clase <i>forkhead</i> de proteínas de unión al ADN. Estos factores nucleares de hepatocitos son activadores transcripcionales de genes específicos del hígado, como la albúmina, también interactúan con la cromatina. Miembros similares de la familia en ratones tienen funciones en la regulación del metabolismo y en la diferenciación del páncreas y el hígado (<i>GeneCards</i> , 2022).
RUNX1	Forma un complejo con el cofactor CBFB. Este complejo proporciona estabilidad a la proteína RUNX1 que participa en la generación de células madre hematopoyéticas y en su diferenciación en líneas mieloides y linfoides. Se ha demostrado que la pérdida de la función de RUNX1 altera la diferenciación entre las líneas mieloides y linfoides, lo que a menudo resulta

	en el desarrollo de leucemia (<i>GeneCards</i> , 2022).
NR1H4	Este gen codifica un FT activado por ligando que funciona como un receptor de ácidos biliares, cuando se une a los ácidos biliares se une al ADN y regula la expresión de genes implicados en la síntesis y el transporte de ácidos biliares (<i>GeneCards</i> , 2022).
FOXM1	Activador transcripcional implicado en la proliferación celular, regulando la expresión de varios genes del ciclo celular. También juega un papel en la reparación de roturas de ADN participando en la respuesta del punto de control de daños en el ADN (<i>GeneCards</i> , 2022).
ZNF264	Este gen codifica una proteína con dedos de zinc. Las proteínas con dedos de zinc a menudo se localizan en el núcleo, se unen a los ácidos nucleicos y regulan la transcripción. Involucrado en la regulación transcripcional (<i>GeneCards</i> , 2022).

Tabla 5. Principales funciones asociadas a los 12 FTs relacionados al motivo 2 del promotor central.

Se resumen las funciones de FT de cada proteína/gen asociadas al motivo 2, después del análisis con *MEME Suite*. La información presentada se extrajo de la BD *GeneCards*.

7.3 *Enhancers* relacionados al gen SCN9A

La mayor parte del genoma de los mamíferos está conformado por regiones no codificantes como los *enhancers*. Los *enhancers* son secuencias que tienen un papel fundamental en la regulación de la expresión tejido específica de un gen, muchas veces desde una distancia lejana al promotor y de forma independiente a la orientación (Jin *et al*, 2011).

Como parte del análisis, se realizó la identificación de los posibles *enhancers* asociados a *SCN9A* de humano. Para lograrlo, se inició con una búsqueda de los *enhancers* de este gen en la BD *GeneCards* (Stelzer *et al*, 2016). *GeneCards*, cuenta con una recopilación de 40 *enhancers* relacionados al gen de interés; de éstos, se seleccionaron aquellos con mayor puntuación de Asociación (GA) o de *GeneHancer* (GH), 8 en total. Después de obtener la ubicación de cada uno de ellos en el genoma de humano, se descargaron las secuencias correspondientes a cada *enhancer* seleccionado (desde la BD *Ensembl*).

7.4 Identificación de *enhancers*

Como se mencionó en la sección de “Métodos”, cada una de las secuencias incluye ~180 000 pb río arriba del SIT, se eligió abarcar esta longitud con el objetivo de aumentar la probabilidad de encontrar las regiones no codificantes de interés. La identificación y procesamiento de los *enhancers* seleccionados se realizó de forma similar a lo realizado

para el caso del promotor. Es decir, se identificaron los *enhancers* de relevancia, se buscaron las zonas pertenecientes a cada uno de los ocho *enchancers* seleccionados en el genoma de humano, se obtuvieron las secuencias correspondientes y se llevó a cabo un alineamiento utilizando *VISTA Tools*.

Posteriormente, se determinaron las regiones homólogas en el resto de las especies (alineando con referencia a la secuencia de humano) (Figura 9). Los resultados obtenidos de este alineamiento exponen algunas de las similitudes y diferencias entre las secuencias no codificantes de las especies seleccionadas.

En la figura 9 se muestran los *enchancers* de acuerdo al orden de aparición en el genoma. Por otro lado, en orden ascendente en términos de longitud en pb, el *enhancer* 1 (GH02J166371) es el más pequeño, con 400 pb. Éste se localizó en todas las secuencias y el alineamiento correspondió aproximadamente al tamaño de la secuencia en humano. Sin embargo, la longitud del elemento conservado (70% por cada 100 pb, azul) fue menor en las secuencias de ratopín, rata y ratón (*enhancer* 1, figura 9), de forma similar a lo encontrado en el promotor en las secciones anteriores.

Siguiendo el orden ascendente en términos de longitud en pb, respecto al *enhancer* 6 (GH02J166233), que en humano y macaco cuenta con 599 pb, se alineó por arriba del 70% por cada 100 pb solo con 5 especies más (gato, perro, elefante, caballo y oso polar). Sin embargo, estas secuencias alineadas fueron de menor longitud a la esperada (~200 pb), el resto de las especies no superaron el umbral de 70% por cada 100 pb que usa *VISTA Tools* como medida de conservación (*enhancer* 6, figura 9).

En tercer lugar, el *enhancer* 8 (GH02J166360), de 799 pb en humano, se localizó en casi todas las especies. Estuvo ausente en la secuencia correspondiente a la rata (*gap*) y solo alineó con 5 pb en el elefante; la longitud del elemento conservado (azul) en el ratón fue menor en comparación con el resto de las especies (*enhancer* 8, figura 9).

El enhancer número 3 (GH02J166410), de una longitud de 1311 pb en el humano, estuvo ausente en el alineamiento de la rata y conejo; aunado a esto, la longitud del elemento conservado fue menor en ratopín y ratón (*enhancer 3*, figura 9).

El elemento de conservación para el *enhancer 7* (GH02J166181) (1399 pb) no fue localizado en las secuencias correspondientes a rata, ratón y conejo; además, éste resultó mucho más corto en el ratopín (*enhancer 7*, figura 9).

Por otro lado, en el *enhancer 2* (GH02J166493), que en humano cuenta con 1659 pb, se encontraron regiones de alineamiento homólogas en todas las especies. Sin embargo, la longitud de los elementos conservados fue menor para el caso de los tres roedores, en contraste con el resto de las especies (*enhancer 2*, figura 9).

El *enhancer 4* (GH02J166229) se identificó en casi todas las especies (exceptuando al murciélago), encontrando elementos conservados más cortos en ratopín, rata, ratón y conejo (*enhancer 4*, figura 9), esto en contraste con la secuencia de humano, que tiene una longitud de 1997 pb.

Por último, el *enhancer 5* (GH02J166344), tuvo una longitud de 3599 pb en humano y se localizó en el alineamiento de todas las especies seleccionadas. De manera similar a lo sucedido con el *enhancer 4*, las regiones alineadas más pequeñas fueron las de rata y ratón, seguidas de las de ratopín y conejo (*enhancer 5*, figura 9).

En general, se observó que las secuencias correspondientes a *enhancers* en roedores y lagomorfos se parecen menos a las del humano y el resto de los mamíferos. Aunque es cierto que este planteamiento tiene sus excepciones y, tomando como base el antecedente que se tiene en la literatura acerca de las diferencias fenotípicas entre el humano y el ratón debido a la pérdida de función del gen *SCN9A*, resulta importante el análisis y comparación de estas regiones entre estos animales.

7.5 Análisis de motivos de los *enhancers*.

Una vez identificadas las regiones correspondientes a los ocho *enhancers* en humano y alineadas al resto de las especies, se realizó una búsqueda de motivos en todas las regiones

de interés, incluyendo todas las secuencias. En el *enhancer* número 8 se excluyó la secuencia de elefante ya que alineó con solo 5 pb, tamaño que es menor al permitido por *MEME*.

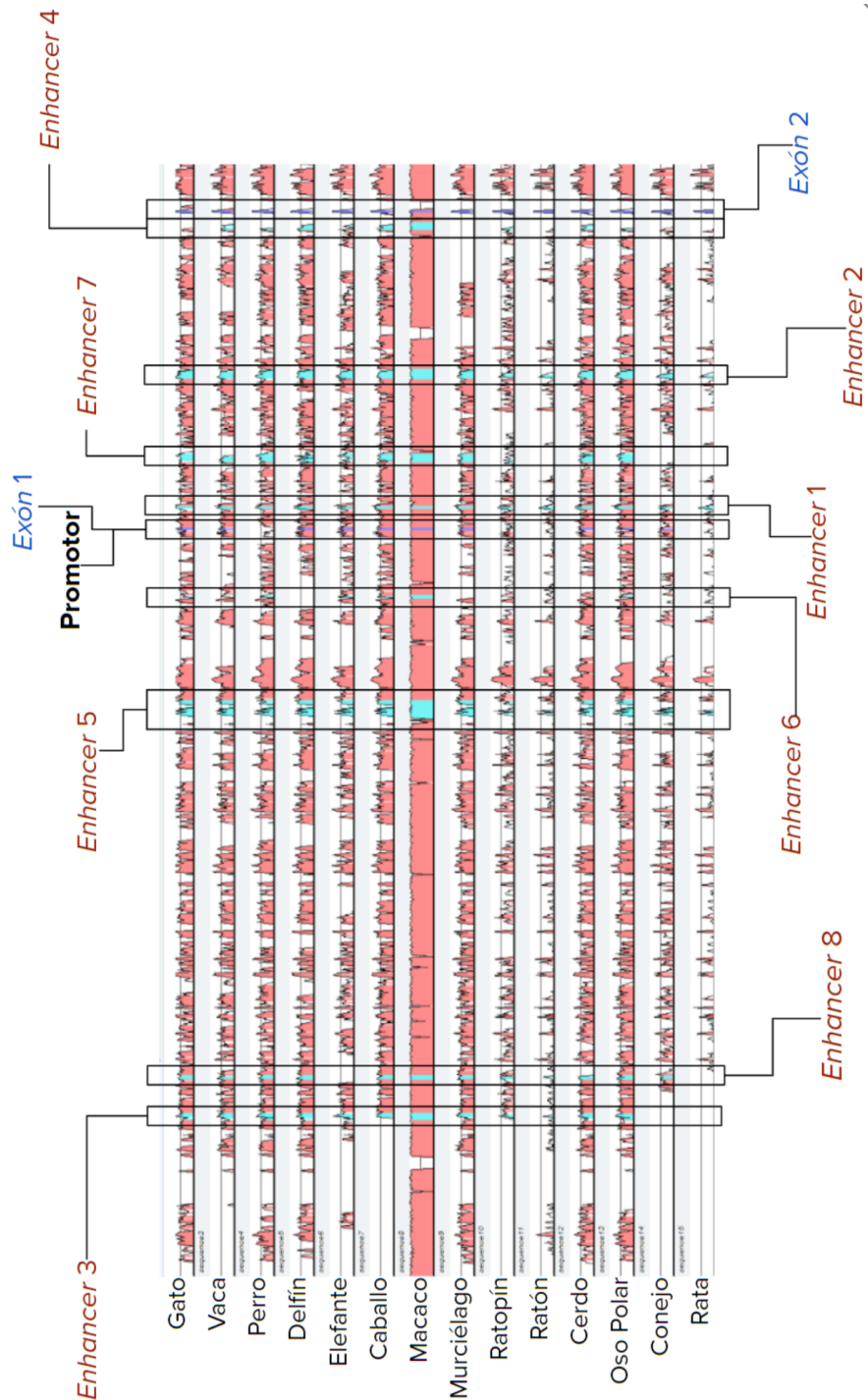


Figura 9. Alineamiento de secuencias y ubicación de cada *enhancer*, promotor, Sitio de Inicio de la Transcripción, exón 1 y exón 2 en cada especie analizada. El alineamiento se realizó con *Vista Tools*, usando la secuencia correspondiente a Humano como referencia.

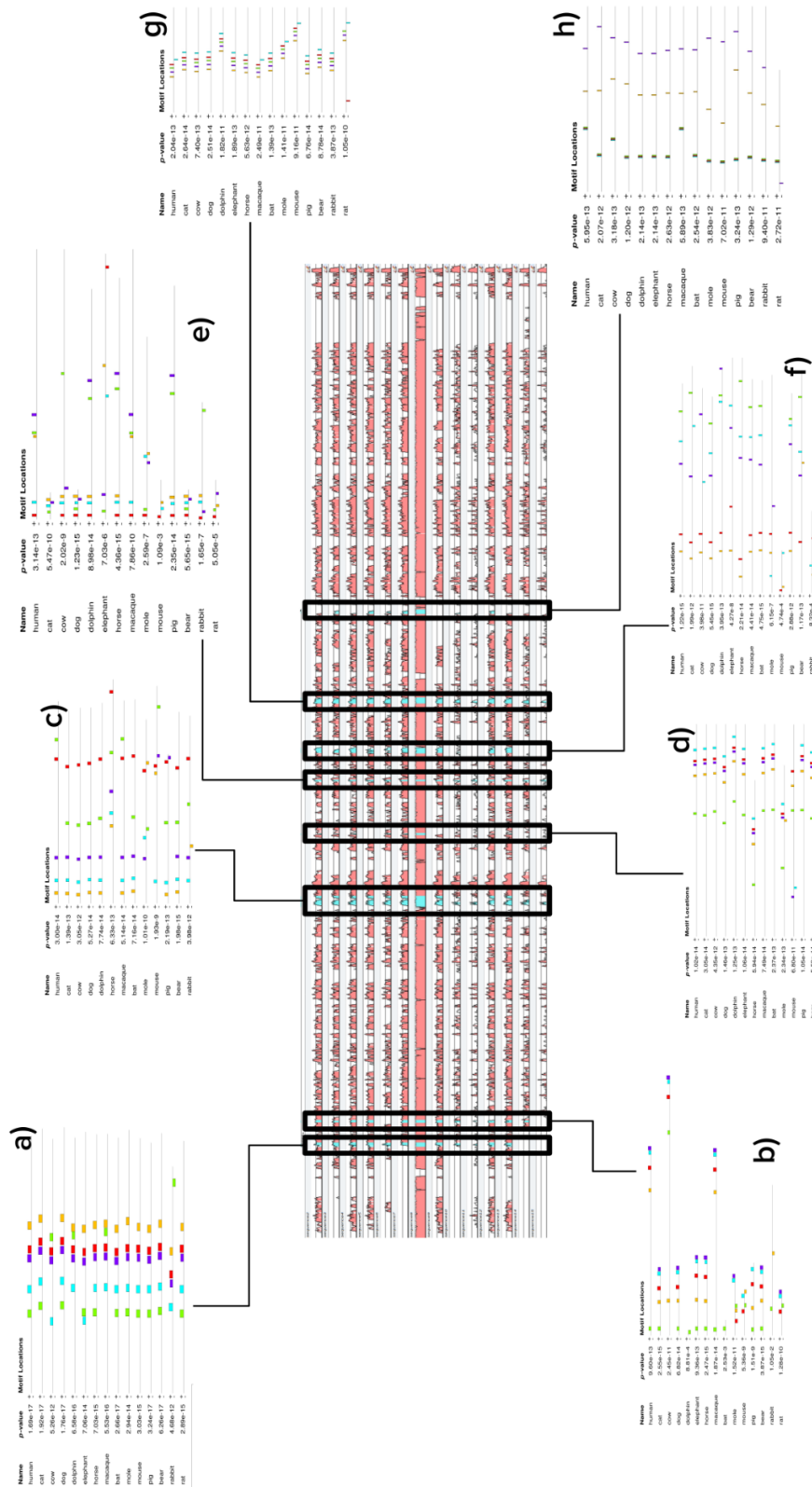


Figura 10. Identificación de motivos conservados entre distintas especies de mamíferos usando *MEME Suite*. La búsqueda se realizó para cada uno de los ocho *enhancers* seleccionados. Las especificaciones dadas

al programa fueron la búsqueda de 5 motivos con una longitud de entre 6 y 12 nt que tuvieran como máximo una repetición por secuencia. a) motivos identificados en el *enhancer* 3; b) motivos identificados en el *enhancer* 8; c) motivos identificados en el *enhancer* 5; d) motivos identificados en el *enhancer* 6; e) motivos identificados en el *enhancer* 1; f) motivos identificados en el *enhancer* 7; g) motivos identificados en el *enhancer* 2; y h) motivos identificados en el *enhancer* 4. En el apartado “Anexo” se encuentra una vista ampliada de cada uno de los motivos identificados para cada *enhancer*.

N° de <i>enhancer</i> ID	Longitud en humano (pb)	Motivos identificados (Consenso)
1 (GH02J166371)	400	1. AAATGCCAGACC 2. CTCTTGGCAGCT 3. GGAAGGAAATGG 4. TGGGTGGTRAGG 5. GAATGCCTTGAA
2 (GH02J166493)	1659	1. YCTGCTCAGTCC 2. CTCTCACTCTGG 3. GCCTAGGCACTG 4. CAGGAAATRTGC 5. TTAAGTTCTGGA
3 (GH02J166410)	1311	1. TAGCAAGTAAGT 2. CTGCCAGCCTTT 3. CAAAAACAGCTT 4. CTGGTGGAGACC 5. AAGAGAAGCCCA
4 (GH02J166229)	1997	1. CCTTCTCCTCCC 2. GAAATTTCCATC 3. CTGCTTCACATC 4. GCCAAGATAACA 5. ATGAGGTCTGAA
5 (GH02J166344)	3599	1. AATGGATGGATC 2. CAGGCAACATTT 3. TATGAATTCAGC 4. CTTCAGGGCTTC 5. CCTCTGAGTATG
6 (GH02J166233)	599	1. TCTCCAGACAAG 2. ATGAGGAGTTTT 3. ATTGAAGCAHDG 4. CCTGTCCTCAAC

		5. TTGAGTTGGGG
7 (GH02J166181)	1399	1. ACACTGAAAGTC 2. TGAGGAATCTCC 3. CCCAKCCYYTG 4. GTKDCCTTTTGG 5. CACTCACTKCCT
8 (GH02J166360)	799	1. CAACARTGAAGG 2. ATCTTTCTGTCT 3. TCTCAGAGGCC 4. TCCTAGATTTTC 5. CTGCTTGCTTTA

Tabla 6. Identificación de motivos conservados en ocho *enhancers* de distintas especies de mamíferos usando el modo “clásico” de MEME Suite. Las especificaciones dadas al programa fueron la búsqueda de 5 motivos con una longitud de entre 6 y 12 nt que tuvieran como máximo una repetición por secuencia (zoops). Este análisis se repitió ocho veces, una por cada *enhancer* seleccionado. En paréntesis se muestra el ID de entrada de cada *enhancer* para la BD *GeneCards*.

Tal como ocurrió en la búsqueda de SUFTs en la secuencia del promotor, en los resultados que se presentarán a continuación (figura 10), se eligió trabajar con *zoops* ya que esta opción es recomendada cuando se sospecha que algunos motivos pueden no estar presentes en todas las secuencias. Además, se pidió identificar 5 motivos en un rango de 6-12 nt para cada *enhancer*. El resto de las opciones se dejaron en sus valores por omisión. En este análisis, cada uno de los cinco motivos requeridos para cada *enhancer* resultaron significativos (<0.001). En la tabla 6 se resumen las secuencias consenso de los motivos para cada *enhancer* obtenidas de este análisis.

No se encontraron diferencias de interés entre las secuencias correspondientes al *enhancer* 1 (400 pb en humano). Los motivos etiquetados, según su valor de *e*, como 2 y 3 fueron los únicos que se localizaron en un lugar distinto en algunas de las especies. El motivo 2 se encontró en la segunda posición en la mayoría de los animales, excepto en la vaca, el delfín, el elefante y el conejo (al inicio). Por otro lado, el motivo 3 se ubicó al inicio de la secuencia en la mayoría de las especies y en una posición distinta a la esperada en la vaca, el delfín y el macaco. Finalmente, el motivo 3 del conejo se localizó ocupando una posición distinta a la de humano y en la cadena complementaria (Figura 10a).

En el *enhancer 2* (1659 pb en humano) llama la atención que de los 5 motivos requeridos solo se encontraron 4 en el caso de la vaca, el elefante y el ratopín (motivo 3 ausente). Aunado a esto, las secuencias de rata, ratón y conejo son notablemente más cortas que el resto y solo cuentan con 3; los motivos faltantes en estas dos especies son los etiquetados como 3 y 4 (Figuras 10f y 11).

El *enhancer 3* (1311 pb en humano) solo mostró diferencias en los motivos 2 y 5 de las secuencias de ratón y ratopín, respectivamente. En el ratón, el motivo 2 se encuentra en una posición distinta y en la cadena complementaria, mientras que en el ratopín, el motivo 5 también se encuentra en la cadena complementaria respecto al humano (Figura 10d).

En el *enhancer 4* (1997 pb en humano) es importante destacar que el motivo 2 se encuentra en la secuencia complementaria sólo en el caso del humano, el macaco, el ratopín, el ratón y la rata (Figura 10g).

Los motivos localizados en el *enhancer 5* (3599 pb en humano) son regulares a lo largo de las secuencias, la única diferencia reside en el motivo 4 de la rata, que además de ubicarse en la cadena complementaria, está en una posición distinta a la del resto de las especies y la longitud de la secuencia es la menor en comparación con las demás (Figura 10h).

El tamaño de las secuencias del *enhancer 6* (599 pb en humano) varía ampliamente entre las especies. En el delfín y el murciélago sólo se encuentra el motivo 3; en el ratón falta el motivo 4 y en la rata el motivo 5; el conejo solo alineó con dos motivos (3 y 5). En el ratopín, el ratón, el conejo y la rata se localizaron algunos motivos en la cadena complementaria, caso distinto lo mostrado en humano y el resto de las especies (Figura 10b y 12).

Las secuencias correspondientes al *enhancer 7* (1399 pb en humano) mostraron una alta variación entre las especies, ya sea en la posición de los motivos como en la longitud de las secuencias. La rata y el gato alinearon 4 motivos, mientras que el ratón solo 3 (Figura 10e).

Por último, respecto al humano, los motivos 4 y 5 del *enhancer* 8 (799 en humano) fueron localizados en las secuencias de ratopín, ratón, cerdo y conejo en una posición distinta a la tomada como referencia en el caso de humano (Figura 10c).

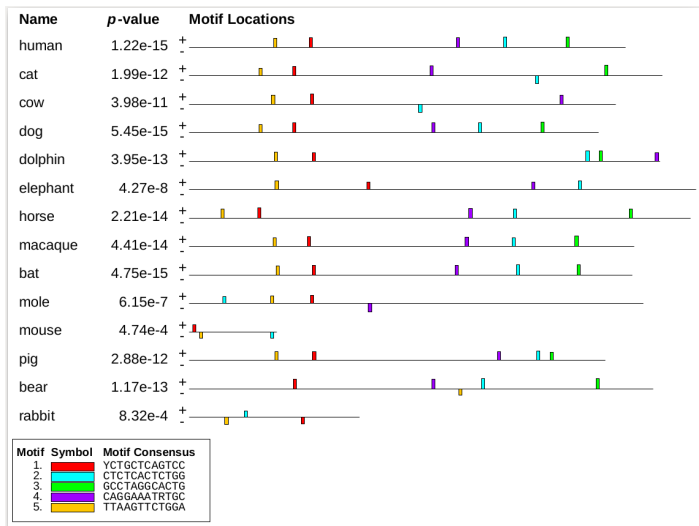


Figura 11. Descubrimiento de motivos en el *enhancer* 2 del gen *SCN9A* usando el modo “clásico” de *MEME Suite*.
Mole = ratopín (*naked mole rat*).

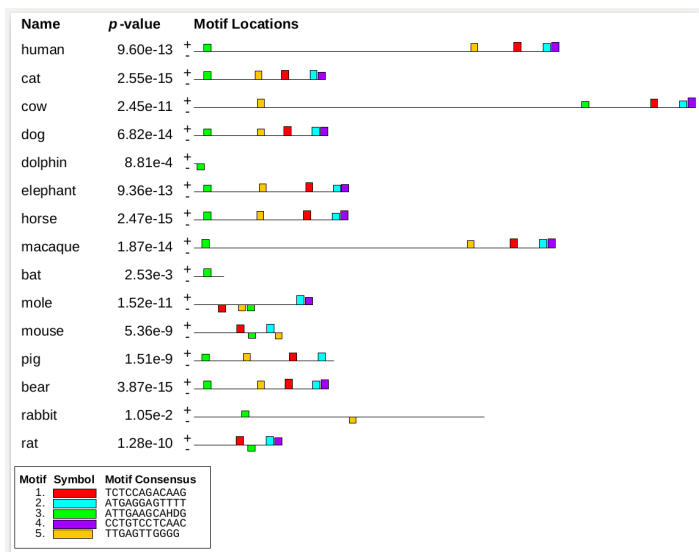


Figura 12. Descubrimiento de motivos en el *enhancer* 6 del gen *SCN9A* usando el modo “clásico” de *MEME Suite*.
Mole = ratopín (*naked mole rat*).

En resumen, además de las diferencias en cuanto al tamaño (pb) de cada uno de los *enhancers*, se observaron distintas secuencias motivo identificadas para cada uno de estos. La posición, número y orientación de cada motivo varió con frecuencia dependiendo de la especie. Sin embargo, estos cambios fueron más marcados en los roedores. Como se mencionó, los promotores y *enhancers* están involucrados e interactúan entre sí en el proceso de reconexión de secuencias reguladoras. Este análisis nos ayuda a identificar

las similitudes y diferencias en la conservación de motivos en cada *enhancer* entre especies.

7.6 Análisis discriminatorio de motivos

En el análisis anterior se utilizó el modo “clásico” de descubrimiento de motivos en *MEME*. Sin embargo, el modo “discriminatorio” permite utilizar la información biológica de una secuencia control o negativa con el fin de encontrar motivos en las secuencias positivas, lo que direcciona el proceso de descubrimiento de motivos a las secuencias en estudio y es recomendable cuando se buscan SUFTs en regiones de naturaleza no aleatoria (Bailey et al, 2010).

El ratón, que es el principal modelo experimental que se usa en neurociencias, pertenece al orden *Rodentia*, junto con otros animales incluidos en este estudio, como la rata y el ratopín. El grupo hermano del orden *Rodentia* es *Lagomorpha*, cuyo exponente más popular es el conejo (*Tree of life web project*, 2005). Estos dos grupos conforman el clado de los *Glires*. Por lo anterior, y debido a que en el análisis de motivos de enhancers se observaron diferencias entre las secuencias de humano y, en algunos casos, entre las del ratón, la rata, el ratopín y el conejo, se decidió realizar un análisis discriminatorio con todas las secuencias seleccionadas, eligiendo como control al clado de los *Glires*, por su relevancia como modelos experimentales y debido a la expresión diferencial del gen *SCN9A* entre humanos y ratones. Este nuevo descubrimiento de motivos se realizó tanto en la secuencia del promotor, como en el enhancer 6, que fue uno de los que mostró marcadas diferencias entre los *Glires* y el resto de las especies.

En el caso del promotor se identificaron 5 SUFTs, 4 de ellos con 12 pb y uno con 8 pb (Figura 13a). Es importante destacar que, solo en los roedores no se encontraron todas las secuencias motivo requeridas, obteniendo 4 SUFTs en el ratopín y la rata y 3 en el ratón.

Recordando que, el motivo etiquetado con el número 1 es aquel que tiene el valor de e más bajo y, por lo tanto, la mayor significancia, es necesario puntualizar que éste, a diferencia del resto de los roedores, no se encontró en el ratón; tampoco estuvo presente el motivo 3 (Figura 13b). Por su parte, el SUFT faltante en la secuencia del ratopín fue el 2, siendo

también la única especie en la que no se alineó este motivo. Por último, en el caso de la rata el motivo ausente fue el 4. El único motivo que alineó con todas las especies fue el 5.

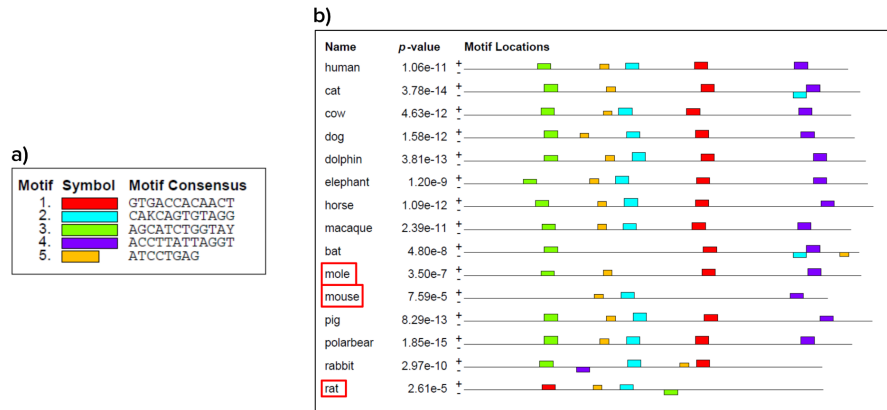


Figura 13. Descubrimiento de motivos en los promotores del gen *SCN9A* usando el modo “discriminatorio” de *MEME Suite*.

Nótese que el largo de las secuencias, el número de SUFTs y la orientación, pueden cambiar entre especies. Mole = ratopín (*naked mole rat*).

Luego de la identificación de los motivos conservados entre especies, se analizó si estas regiones corresponden a SUFTs conocidos. Para esto, se analizó cada motivo encontrado en las secuencias promotoras usando la herramienta *Tomtom* de *MEME Suite 5.1.1*. La base de datos seleccionada fue *HOCOMOCO Human and Mouse* (v11 core).

Se realizó una nueva búsqueda de cada uno de los FT encontrados para cada secuencia motivo, tomando como base la idea de que los modos clásico y discriminatorio proporcionaron distintos resultados. En el modo discriminatorio se buscan motivos enriquecidos en una secuencia respecto a la otra, es decir, específicos a un conjunto de datos. Posteriormente, se descartaron aquellos que, tras su inspección en GTE_x (Keen & Moore, 2015) no mostraron una expresión marcada en el SN (rojo, Tabla 7) y resaltando aquellos que únicamente se expresan en el SN (amarillo, Tabla 7).

Secuencia motivo	Alineamientos significativos con la base de datos	FT relacionados alineados con el SUFT
1 (GTGACCACAAC T)	15	PEBB_HUMAN.H11MO.0.C, RUNX1_HUMAN.H11MO.0.A, RXRB_HUMAN.H11MO.0.C, RUNX3_HUMAN.H11MO.0.A, PEBB_MOUSE.H11MO.0.C, THA_MOUSE.H11MO.1.C, RXRB_MOUSE.H11MO.0.C, MEIS1_MOUSE.H11MO.1.A,

		RUNX3_MOUSE.H11MO.0.A, VDR_HUMAN.H11MO.1.A, ZN250_HUMAN.H11MO.0.C, THB_HUMAN.H11MO.1.D	THA_HUMAN.H11MO.1.D, VDR_MOUSE.H11MO.1.A, ATF2_HUMAN.H11MO.2.C,
2 (CAKCAGTGTAG G)	1	ZSC31_HUMAN.H11MO.0.C	
3 (AGCATCTGGTA Y)	32	TWST1_MOUSE.H11MO.1.B, PTF1A_HUMAN.H11MO.1.B, ASCL2_HUMAN.H11MO.0.D, TWST1_MOUSE.H11MO.0.B, NDF2_MOUSE.H11MO.0.A, BHE23_HUMAN.H11MO.0.D, ATOH1_HUMAN.H11MO.0.B, BHA15_HUMAN.H11MO.0.B, FIGLA_HUMAN.H11MO.0.D, NGN2_MOUSE.H11MO.0.C, NDF1_MOUSE.H11MO.0.A, ZBT18_MOUSE.H11MO.0.D, NDF2_HUMAN.H11MO.0.B, MYF6_MOUSE.H11MO.0.C, MYF5_MOUSE.H11MO.0.D, OLIG2_HUMAN.H11MO.1.B, OLIG1_HUMAN.H11MO.0.D	TWST1_HUMAN.H11MO.1.A, PTF1A_MOUSE.H11MO.1.A, ASCL2_MOUSE.H11MO.0.C, TWST1_HUMAN.H11MO.0.A, OLIG3_HUMAN.H11MO.0.D, ATOH1_MOUSE.H11MO.0.B, BHA15_MOUSE.H11MO.0.A, NGN2_HUMAN.H11MO.0.D, NDF1_HUMAN.H11MO.0.A, ZBT18_HUMAN.H11MO.0.C, MSGN1_MOUSE.H11MO.0.C, MYF6_HUMAN.H11MO.0.C, BHE22_HUMAN.H11MO.0.D, ITF2_HUMAN.H11MO.0.C, OLIG2_MOUSE.H11MO.1.A,
4 (ACCTTATTAGGT)	11	ZN134_HUMAN.H11MO.1.C, ZN708_HUMAN.H11MO.0.C, NOTO_HUMAN.H11MO.0.D, DUX4_HUMAN.H11MO.0.A, HXD13_HUMAN.H11MO.0.D, SRF_MOUSE.H11MO.0.A	ZN134_HUMAN.H11MO.0.C, SRF_HUMAN.H11MO.0.A, LBX2_HUMAN.H11MO.0.D, DMBX1_HUMAN.H11MO.0.D, HXD13_MOUSE.H11MO.0.D,
5 (ATCCTGAG)	9	STAT6_HUMAN.H11MO.0.B, MAFG_HUMAN.H11MO.1.A, MAFF_HUMAN.H11MO.0.B, STAT6_MOUSE.H11MO.0.A, MAFF_MOUSE.H11MO.0.A	SIX4_MOUSE.H11MO.0.C, MAFG_MOUSE.H11MO.1.A, MAFF_HUMAN.H11MO.1.B, SIX2_HUMAN.H11MO.0.A,

Tabla 7. Secuencias motivo obtenidas mediante MEME y sus distintos alineamientos contra SUFT ya conocidos, utilizando la BD HOCOMOCO.

En rojo se muestran los FT que, después de una inspección en *GTE_x* no mostraron expresión en el SN; en amarillo se encuentran aquellos que sólo tienen expresión en el SN o que es más prominente en relación a otros tejidos; sin resaltar se muestran los alineamientos SUFT-FT que se expresan tanto en el SN como en otros tejidos.

Como se mencionó, usando el método discriminatorio de descubrimiento de motivos, en el ratón no se encontraron los motivos 1 y 3 del promotor central. Una vez realizada la búsqueda de expresión en tejidos, es importante notar que en el motivo 3 se encuentran los FT conocidos que, tanto en humano como en ratón, solo tienen expresión en el SN y no se encuentran en otros tejidos (Figura 14 a y b) o alineamientos SUFT-FT expresados principalmente en el SN pero con una leve expresión en otros tejidos corporales (Figura 14 c y d). Como se discutió con anterioridad, este resultado también podría sugerir que durante la evolución, ocurrió un proceso de recableado, en el cual se conservó la función de la proteína entre especies, pero existe una expresión fenotípica diferencial del canal Na_v1.7, que responde a la historia evolutiva de las secuencias reguladoras específica a cada

especie. En el caso del *enhancer 6*, en la rata y el ratón solo se localizaron 3 motivos en el modo discriminatorio (Figura 15).

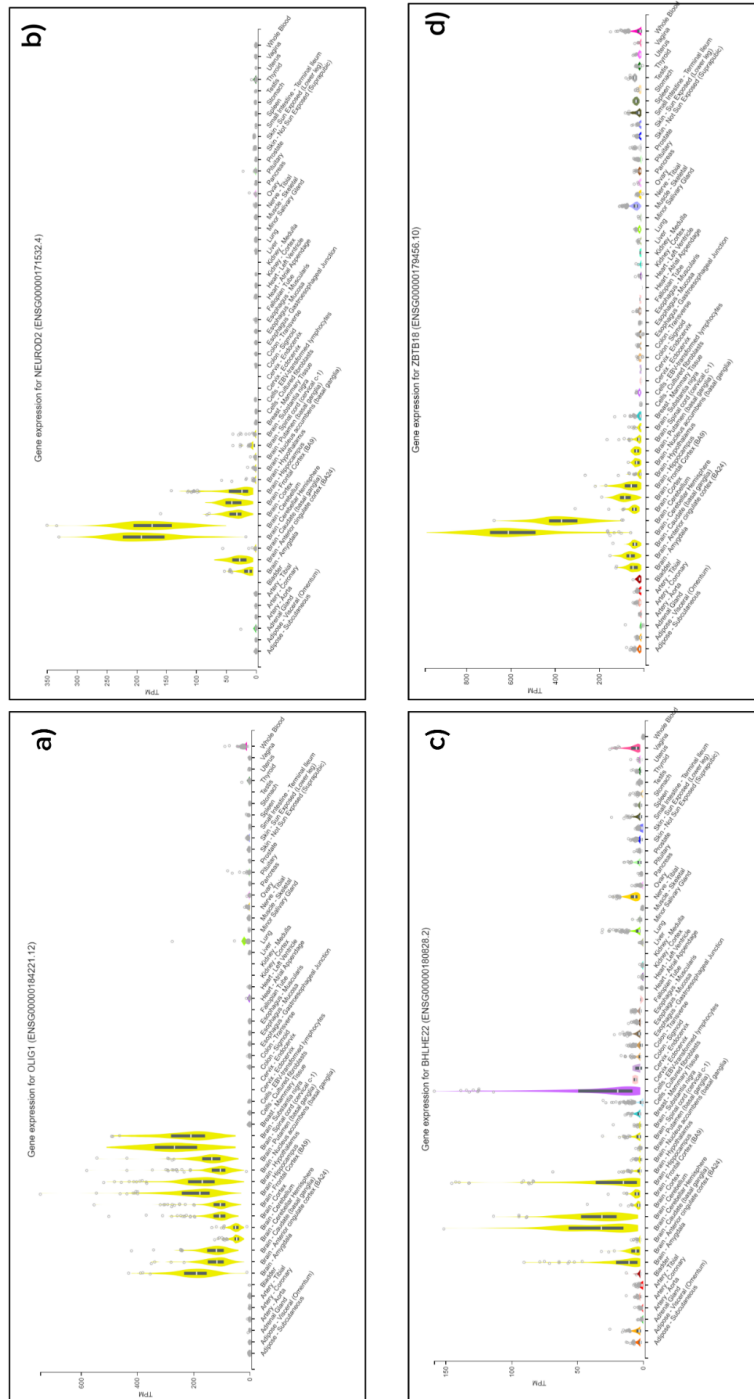


Figura 14. Análisis de los niveles de expresión de los FTs relacionados a los SUFTs identificados

NEUROD1, NEUROD2, ZBTB18 y BHLHE22 son cuatro distintos FTs que coincidieron con SUFTs ya conocidos al analizarse contra la base de datos *HOCOMOCO*. El motivo número 3 del *enhancer 6*, llamó especialmente la atención ya que éste se alineó con 10 FT de ratón o humano que están presentes solo en el SN (a

y b) o mayormente en el SN pero cuenta con expresión reducida en otros tejidos (c y d). Las líneas representan la expresión de cada FT en cada tejidos, el amarillo corresponde al SN. Imágenes tomadas de *GTEx*.

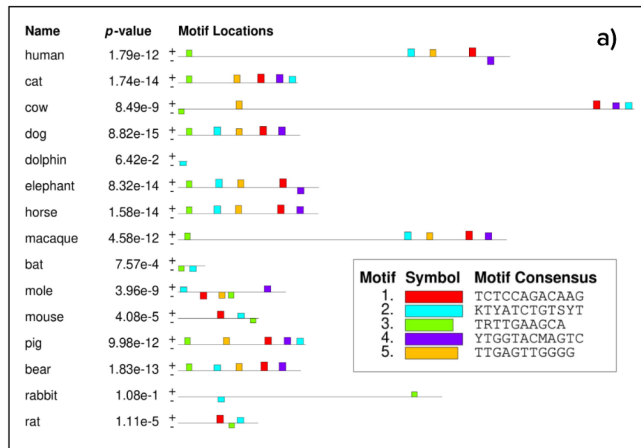


Figura 15. Modo “discriminatorio” de MEME Suite en la búsqueda de motivos en el enhancer 6 del gen SCN9A.

Posteriormente, se realizó la búsqueda con *HOCOMOCO*, esta vez dando como entrada 3 de los 5 motivos identificados en las secuencias del *enhancer 6*. Se eligió este número de secuencias motivo debido a que los motivos 1, 4 y 5 estuvieron ausentes en una o más especies de *Glires*, el resto de ellos estuvo presente en todas las secuencias. Una vez identificados los SUFT conocidos que corresponden con los motivos conservados encontrados en el alineamiento de *MEME*, se utilizó *GTEx* con el fin de descartar aquellos SUFT que no mostraron una expresión marcada en el SN (rojo, Tabla 8). En este caso, los motivos 4 y 5 solo se alinearon con 2 FT-SUFT que tienen expresión, entre otros tejidos, en el SN (Tabla 8). Que el número de secuencias motivo encontradas haya sido de solo 2 podría indicar una especificidad que sería importante estudiar en futuros análisis.

Secuencia motivo	Alineamientos significativos con la base de datos	SUFT relacionados
1 (TCTCCAGACAAG)	14	P53_HUMAN.H11MO.0.A, P73_MOUSE.H11MO.0.B, P63_HUMAN.H11MO.0.A, HSF1_HUMAN.H11MO.1.A, ZNF8_HUMAN.H11MO.0.C, KAISO_MOUSE.H11MO.1.B, HAND1_MOUSE.H11MO.1.C, STAT1_HUMAN.H11MO.0.A P73_HUMAN.H11MO.0.A, STAT6_HUMAN.H11MO.0.B, P63_MOUSE.H11MO.0.B, P53_MOUSE.H11MO.0.A, KAISO_HUMAN.H11MO.1.A, HAND1_HUMAN.H11MO.1.D,
4 (YTGGTACMAGTC)	2	ZN329_HUMAN.H11MO.0.C, ID4_HUMAN.H11MO.0.D

5 (TTGAGTTG GGG)	4	RREB1_HUMAN.H11MO.0.D, EGR3_HUMAN.H11MO.0.D,	RREB1_MOUSE.H11MO.0.D, EGR3_MOUSE.H11MO.0.D
------------------------	---	---	--

Tabla 8. Secuencias motivo obtenidas mediante *MEME Suite* y sus distintos alineamientos contra FT-SUFT ya conocidos, utilizando la BD *HOCOMOCO*.

En rojo se muestran los FT-SUFT que, después de una inspección en *GTEx* no mostraron expresión en el SN; sin resaltar se muestran los FT-SUFT que se expresan tanto en el SN como en otros tejidos.

Como se ha mencionado, el modo clásico descubre motivos enriquecidos en un conjunto de secuencias de entrada, mientras que el discriminatorio encuentra motivos enriquecidos en un set respecto a un segundo conjunto. Por lo tanto, los motivos resultantes en cada análisis fueron distintos, sin embargo, en algunos casos, el mismo FT se asoció a motivos distintos obtenidos por estos dos medios (STAT6, SIX2, P63, SRF, RXRB, VDR, RUNX1, DUX4).

Por lo anterior, los SUFTs son un buen punto de partida en la comprensión de la divergencia de los elementos de regulación (como promotores y *enhancers*), además, se han propuesto como un mecanismo plausible en la evolución de la regulación génica, generando diversidad fenotípica (Tuğrul, 2015). Sin embargo, la divergencia de secuencias reguladoras sin un cambio obvio en la regulación del gen al que pertenecen es un reto en el entendimiento de la regulación génica en sí y en cómo esos mismos cambios pueden conducir a divergencias fenotípicas (Venkataram & Fay, 2010).

La estrategia de análisis aquí propuesta, es una forma sencilla y asequible de identificación de posibles SUFTs conservados y no conservados entre especies. Algunos autores han planteado que la evolución de secuencias reguladoras es un proceso dinámico que involucra la ganancia y pérdida de SUFTs (recambio o *turnover*) (Moses *et al*, 2006; Otto *et al*, 2009; Venkataram & Fay, 2010). En el caso comparativo que se ha estado discutiendo, la ausencia del motivo 2 en rata y ratón (respecto al humano), también podría ser una prueba de la hipótesis de recambio de SUFTs como mecanismo de la evolución de secuencias reguladoras. Es importante resaltar que la ganancia o pérdida de SUFTs puede alterar o no el fenotipo resultante de la expresión de un gen (Moses *et al*, 2006).

Recordando que un *enhancer* consiste en grupos de SUFTs, que se unen a FTs específicos en un tipo celular (Panigrahi & O'Malley, 2021) y son parte de la red de elementos

reguladores que controlan espaciotemporalmente el inicio de la transcripción y la expresión de un gen. El proceso de transcripción dará inicio cuando ocurra la interacción entre los *enhancers* y los promotores, de ahí la importancia de estudiar estos elementos reguladores.

Las mutaciones genéticas a nivel de *enhancers* podrían interferir en la función de la regulación transcripcional. Por ejemplo, causando una pérdida de SUFTs que pudieran ser necesarios en la unión con FTs específicos y, por lo tanto, modificar las interacciones con el promotor (Carullo & Day, 2018). Por lo tanto, resulta necesario el análisis de los mecanismos moleculares que coordinan la unión entre un *enhancer* y un promotor.

Un modelo ampliamente aceptado en la relación *enhancer*-promotor es el de interacciones mediadas por *loops*, en donde un FT reconoce un SUFT específico en una secuencia *enhancer* y conecta directamente con otro FT que ha establecido contacto con el promotor. Por lo tanto, un papel central del reconocimiento SUFT-FT es el establecimiento de contactos directos entre el promotor y un *enhancer* (Monfils & Barakat, 2021), pese a que éstos se encuentren distantes entre sí.

Anteriormente, se discutió cómo la evolución de las secuencias reguladoras debido a la ganancia o pérdida de SUFTs en el promotor central, podría estar implicada en la diversidad fenotípica, incluso sin un cambio obvio en la proteína del gen al que regulan. Por lo tanto, es posible suponer que no solo los cambios en secuencias reguladoras sino las interacción misma entre dos (o más) elementos reguladores podría ser otro de los mecanismos evolutivos de regulación génica que pudieran conducir a cambios fenotípicos resultados de la pérdida de función del producto codificado, como es el caso de ICD en humanos y el fenotipo letal en ratones. En el caso aquí expuesto, resultaría un planteamiento novedoso para comenzar a entender el fenotipo diferencial entre dos especies cercanas evolutivamente (humano y ratón) debido a la pérdida de función de Na_v1.7. Sin embargo, en primer lugar, aún es una incógnita el proceso exacto sobre cómo interaccionan los *enhancers* y promotores para permitir la expresión de un gen a distintos niveles. Por otro lado, la metodología aquí propuesta no permite estudiar esta relación, pero admite notar algunas diferencias y similitudes en los SUFTs tanto en el promotor

como en los *enhancers* asociados al gen *SCN9A*, así como identificar los FTs involucrados, lo que podrían ser un punto de partida en futuros planteamientos experimentales.

En resumen, la conectividad tridimensional entre los *enhancers* y los promotores está sustentada por el número de reconocimientos de secuencias motivos, ya sea iguales o complementarias, que permiten el correcto acoplamiento entre estas dos regiones no codificantes. Por lo tanto, cambios evolutivos en las secuencias motivo o del reconocimiento FT-SUFT de los promotores y/o de los *enhancers* podrían explicar los distintos niveles de robustez frente a mutaciones de *SCN9A* en las distintas especies y además, la diferencias de expresión fenotípica dentro de un mismo grupo como lo son los mamíferos placentarios (*Eutheria*). Es decir, aunque en todas estos organismos $Na_v1.7$ cumple la misma función de nociceptor, las variaciones en el fenotipo resultado de la pérdida de función de este nociceptor podrían estar relacionadas a la evolución de distintos elementos reguladores y las interacciones que mantienen entre sí.

7.8 Limitaciones

Con lo anteriormente expuesto, se buscó plantear una metodología que permita ejemplificar la posible evolución de elementos reguladores que pueden conducir a diferencias fenotípicas entre especies relacionadas, sin un cambio en la función normal de la proteína codificada pero con diferencias notables cuando se habla del fenotipo anormal. Sin embargo, el ejemplo seleccionado y la estrategia bioinformática utilizada plantean ciertas limitaciones que se discuten a continuación.

En 2007 Ahmad *et al*, estudiaron la presencia de una mutación de pérdida de función en Na_v1.7 en una familia. Además de ICD (que se acompaña de anosmia), los integrantes no presentaban alguna otra alteración y parecían sanos (Ahmad *et al*, 2007). Este hecho llamó la atención debido a que se sabía que, la misma deficiencia global de Na_v1.7 es letal en ratones (Nassar *et al*, 2004). Una de las posibles explicaciones es que esto se debe a los cuidados parentales y médicos presentes en humanos, lo que podría ser una de las razones por las que un paciente con ICD no presente muerte neonatal y llegue a la vida adulta.

Por otro lado, estos autores compararon los patrones de expresión de Na_v1.7 en distintos tejidos de ratas, ratones, macacos y humanos, usando hibridación *in situ*. Como era de esperarse, se detectaron fuertes señales de hibridación (en todas las especies) en los cuerpos celulares de distintos ganglios como el de la raíz dorsal, lugar en donde se encuentran los cuerpos celulares de los nociceptores. Sin embargo, también obtuvieron prominentes señales de hibridación en los núcleos paraventricular y supraóptico del hipotálamo de roedores, en contraste con los primates, en donde estas señales no estuvieron presentes. Adicionalmente, detectaron expresión de mRNA en la hipófisis y la glándula adrenal de ratones y ratas, no así en humanos y macacos. Estos resultados muestran diferencias en los patrones y niveles de expresión de Na_v1.7 entre roedores y primates, mostrando un aumento de la expresión en el SNC de roedores, lo que podría ser otra de las explicaciones de la letalidad de la delección global de Na_v1.7 en ratones pero no en humanos.

Por otro lado, existen más mecanismos de regulación de secuencias no codificantes que no se discutieron aquí pero que podrían ser relevantes. Osterwalder *et al* explican que,

múltiples enhancers con actividades similares cercanos a la misma región génica (redundancia), confieren robustez fenotípica contra mutaciones de pérdida de función. Por lo tanto, la redundancia de enhancers, y sus SUFTs, en los genomas de mamíferos podría ser una forma efectiva de amortiguador regulatorio.

Además, es importante considerar que no solo existen interacciones directas SUFT-FT, sino también indirectas, mediadas por otros FTs. Por último, es necesario realizar estudios posteriores en los que se tomen en cuenta escalas temporales y el papel de la epigenética al momento de estudiar la evolución de secuencias no codificantes.

8. PERSPECTIVAS

A lo largo de la realización de este proyecto se planteó un seguimiento del mismo con un enfoque práctico de los resultados obtenidos. Como se ha comentado, el ratón y la rata son los principales modelos experimentales utilizados en los proyectos de investigación en neurociencias. Por lo tanto, se propone ampliar la investigación del uso de protocolos basados en bioinformática con el fin de aplicarlos en la elección de la especie específica a cada fenotipo que se desee reproducir para su estudio. En este caso, la distancia evolutiva medida en número de sustituciones por sitio (Figuras 16 y 17, barra de escala) es mucho mayor en rata y ratón respecto a cualquier otra especie, ya sea en la secuencia del promotor del gen *SCN9A* o en la secuencia codificante (Tablas 9 y 10). Análisis de este tipo podrían orientar la selección de los sujetos de experimentación más adecuados según el objetivo de investigación, tomando en cuenta un enfoque basado en divergencias evolutivas. En este caso, las filogenias aquí presentadas son una primera aproximación, en proyectos futuros deberán considerarse aspectos como las diferencias en distancias evolutivas en las regiones codificantes y no codificantes, el uso de filogenias en regiones no codificantes y el número de secuencias (especies) incluidas.

Por otro lado, en este trabajo se presentó una comparación entre distintas especies de mamíferos placentarios. Sin embargo, el planteamiento que se propone acerca de la evolución de las secuencias no codificantes y su relación con la expresión diferencial es robusto y se puede estudiar entre especies aún más relacionadas entre sí, como los roedores.

Tanto en el ratón (*Mus musculus*), en la rata topo desnuda o ratopín (*Heterocephalus glaber*), como en el ratón saltamontes (*Onychomys torridus*) el Na_v1.7 cumple la función de nociceptor que se ha estudiado en el resto de los vertebrados. Sin embargo, estas tres especies están sometidas a distintas presiones ambientales. Por ejemplo, el ratopín vive bajo tierra, en donde se acumulan grandes cantidades de CO₂; se sabe que altas concentraciones de CO₂ pueden resultar dolorosas al convertirse en ácido en los tejidos (Smith *et al*, 2011). Por otro lado, el ratón saltamontes vive en el desierto, en donde convive con el escorpión o alacrán de corteza, que es el más venenoso de América del Norte.

Lo interesante de estos ejemplos es que el ratopín presenta un fenotipo conocido como Insensibilidad ácida (ausencia de nocicepción ácida), lo que le permite vivir en lugares con altas concentraciones de CO₂. Por su parte, el ratón saltamontes no solo convive, sino que se alimenta del alacrán de corteza, cuyas neurotoxinas causan comportamientos de dolor intenso en ratones de laboratorio (Lewin, 2013). Las distintas formas de nocicepción en estas tres especies se han relacionado, entre otros, a Na_v1.7 y a *SCN9A*. Por lo tanto, con lo propuesto en este trabajo, sería útil investigar las regiones no codificantes asociadas al gen *SCN9A* en estas especies, que pertenecen al mismo orden y están altamente emparentados. El objetivo es entender un poco más acerca de la evolución de las secuencias reguladoras no codificantes y lo que podrían significar las distintas interacciones entre SUFT-FT, así como la ganancia y/o pérdida de SUFTs.

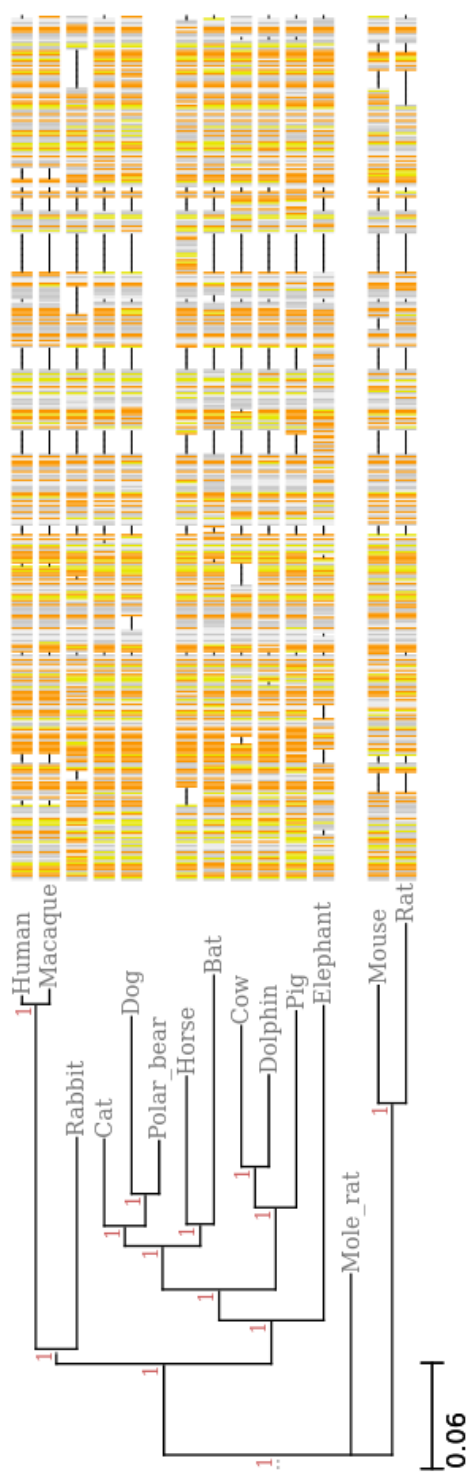


Figura 16. Análisis evolutivo de la secuencia promotora del gen *SCN9A*. Se realizó un análisis filogenético de las secuencias correspondientes al promotor central del gen *SCN9A*. Se muestra un árbol sin raíz obtenido mediante el método de Máxima Verosimilitud, la barra de escala representa el número de sustituciones por sitio (ver tabla). El número que se encuentra en algunos nodos indica el valor de apoyo de cada rama (0-1).

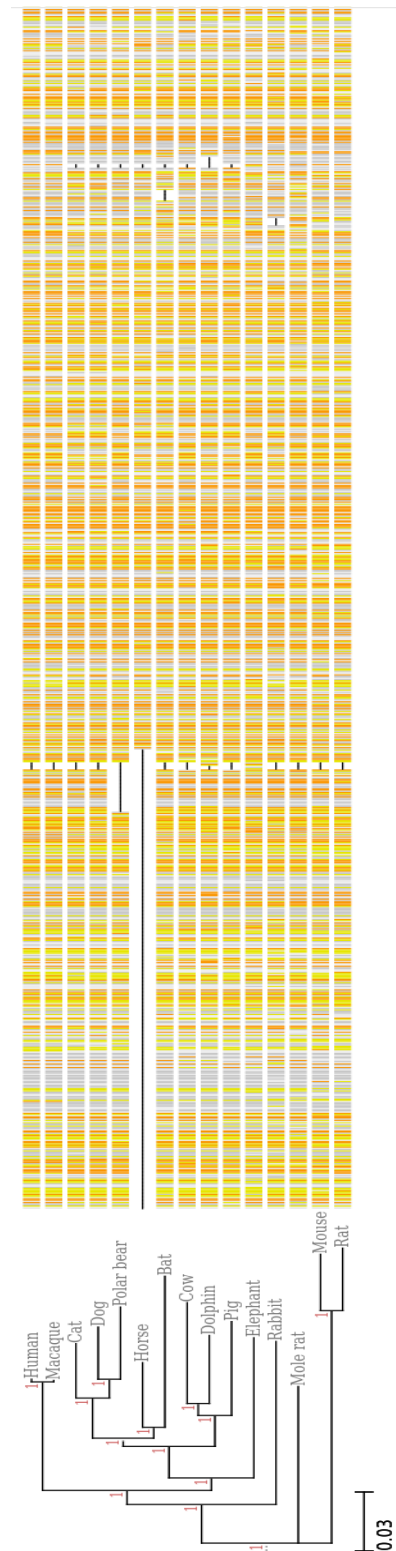


Figura 17. Análisis evolutivo de la secuencia codificante del gen *SCN9A*. Se realizó un análisis filogenético de las secuencias correspondientes a la región codificante del gen *SCN9A*. Se muestra un árbol sin raíz obtenido mediante el método de Máxima Verosimilitud, la barra de escala representa el número de sustituciones por sitio (ver tabla). El número que se encuentra en algunos nodos indica el valor de apoyo de cada rama (0-1).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1.Humano															
2.Gato	0.3241														
3.Vaca	0.3879	0.2281													
4.Perro	0.3918	0.1652	0.2958												
5.Delfin	0.3602	0.2004	0.1307	0.2681											
6.Elefante	0.4028	0.2754	0.3392	0.3432	0.3115										
7.Caballo	0.3594	0.1531	0.2634	0.2208	0.2357	0.3107									
8.Macaco	0.0125	0.3279	0.3917	0.3956	0.3640	0.4065	0.3632								
9.Murciélago	0.4163	0.2100	0.3203	0.2777	0.2926	0.3677	0.2230	0.4201							
10.Ratopin	0.3594	0.2734	0.3372	0.3411	0.3095	0.3521	0.3087	0.3580	0.3656						
11.Ratón	0.5155	0.4347	0.4985	0.5024	0.4708	0.5133	0.4700	0.5193	0.5269	0.3646					
12.Cerdo	0.3974	0.2376	0.2115	0.3054	0.1838	0.3487	0.2729	0.4012	0.3298	0.3467	0.5080				
13.Oso polar	0.3220	0.0954	0.2260	0.1294	0.1983	0.2733	0.1510	0.3258	0.2079	0.2713	0.4326	0.2355			
14.Conejo	0.3163	0.2457	0.3095	0.3134	0.2818	0.3243	0.2810	0.3200	0.3379	0.2758	0.4371	0.3190	0.2436		
15.Rata	0.5502	0.4694	0.5332	0.5371	0.5055	0.5481	0.5047	0.5540	0.5616	0.3994	0.1670	0.5427	0.4673	0.4718	

Tabla 9. Distancias patrísticas entre pares secuencias del promotor central del gen *SCN9A*.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1.Humano															
2.Gato	0.1099														
3.Vaca	0.1292	0.1053													
4.Perro	0.1167	0.0547	0.1121												
5.Delfin	0.1147	0.0909	0.0744	0.0977											
6.Elefante	0.1153	0.1181	0.1374	0.1249	0.1229										
7.Caballo	0.1015	0.0739	0.0970	0.0807	0.0825	0.1098									
8.Macaco	0.0018	0.1094	0.1286	0.1161	0.1142	0.1147	0.1010								
9.Murciélago	0.1396	0.1120	0.1350	0.1187	0.1206	0.1478	0.0950	0.1391							
10.Ratopin	0.1397	0.1526	0.1718	0.1594	0.1574	0.1579	0.1442	0.1391	0.1823						
11.Ratón	0.1974	0.2103	0.2296	0.2171	0.2151	0.2157	0.2020	0.1969	0.2400	0.1957					
12.Cerdo	0.1214	0.0975	0.0906	0.1043	0.0761	0.1296	0.0892	0.1208	0.1272	0.1640	0.2218				
13.Oso polar	0.1255	0.0635	0.1209	0.0547	0.1065	0.1337	0.0895	0.1250	0.1276	0.1682	0.2259	0.1131			
14.Conejo	0.1260	0.1389	0.1582	0.1457	0.1437	0.1443	0.1305	0.1255	0.1686	0.1573	0.2151	0.1504	0.1545		
15.Rata	0.2001	0.2130	0.2323	0.2198	0.2178	0.2184	0.2047	0.1996	0.2427	0.1984	0.0524	0.2245	0.2286	0.2178	

Tabla 10. Distancias patrísticas entre pares secuencias de la región codificante del gen *SCN9A*.

9. CONCLUSIONES

Los resultados obtenidos después del descubrimiento y análisis de motivos con la metodología planteada en este proyecto muestran diferencias en las secuencias motivo entre especies, específicamente entre humano y ratón/rata.

1. Se observaron motivos ausentes en las secuencias de rata y ratón tanto en el centro del promotor como en algunos *enhancers*.
2. Pese a que en estas especies $Na_v1.7$ cumple la misma función, la conservación funcional de la expresión génica, podría no significar la conservación de elementos reguladores en secuencias no codificantes, pudiendo ser importantes en el entendimiento de fenotipos diferenciales entre especies debido a mutaciones de pérdida de función.
3. El caso y la metodología aquí planteados sirven como ejemplo de las hipótesis sobre ganancia y pérdida de SUFTs, que se han propuesto como un mecanismo en la evolución de la regulación génica, generando diversidad fenotípica.
4. El estudio comparativo de las relaciones evolutivas entre especies es de utilidad en el entendimiento del fenotipo diferencial entre ellas.
5. Análisis de este tipo podrían orientar el uso de animales como sujetos de experimentación, tomando en cuenta un enfoque basado en divergencias evolutivas. Además los análisis basados en bioinformática pueden suponer un ahorro de tiempo y dinero en la planeación de protocolos experimentales.

10. REFERENCIAS

- Ahmad, S., Dahllund, L., Eriksson, A. B., Hellgren, D., Karlsson, U., Lund, P.-E., Meijer, I. A., Meury, L., Mills, T., Moody, A., Morinville, A., Morten, J., O'Donnell, D., Raynoschek, C., Salter, H., Rouleau, G. A., & Krupp, J. J. (2007). A stop codon mutation in SCN9A causes lack of pain sensation. *Human Molecular Genetics*, 16(17), 2114-2121. <https://doi.org/10.1093/hmg/ddm160>
- Animals, N. R. C. (US) C. on R. and A. of P. in L. (2009). Mechanisms of Pain. En *Recognition and Alleviation of Pain in Laboratory Animals*. National Academies Press (US). <https://www.ncbi.nlm.nih.gov/books/NBK32659/>
- Apkarian, A. V. (2019). Definitions of nociception, pain, and chronic pain with implications regarding science and society. *Neuroscience letters*, 702, 1-2. <https://doi.org/10.1016/j.neulet.2018.11.039>
- Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W., & Noble, W. S. (2009). MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Research*, 37(Web Server issue), W202-W208. <https://doi.org/10.1093/nar/gkp335>
- Bailey, T. L., Johnson, J., Grant, C. E., & Noble, W. S. (2015). The MEME Suite. *Nucleic Acids Research*, 43(Web Server issue), W39-W49. <https://doi.org/10.1093/nar/gkv416>
- Barrot, M. (2012). Tests and models of nociception and pain in rodents. *Neuroscience*, 211, 39-50. <https://doi.org/10.1016/j.neuroscience.2011.12.041>
- Baxevanis, A. D. (2011). The Importance of Biological Databases in Biological Discovery. *Current Protocols in Bioinformatics*, 34(1). <https://doi.org/10.1002/0471250953.bi0101s34>
- Black, J. A., Frézel, N., Dib-Hajj, S. D., & Waxman, S. G. (2012). Expression of Nav1.7 in DRG neurons extends from peripheral terminals in the skin to central preterminal branches and terminals in the dorsal horn. *Molecular Pain*, 8, 82. <https://doi.org/10.1186/1744-8069-8-82>
- Boeva, V. (2016). Analysis of Genomic Sequence Motifs for Deciphering Transcription Factor Binding and Transcriptional Regulation in Eukaryotic Cells. *Frontiers in Genetics*, 7. <https://www.frontiersin.org/article/10.3389/fgene.2016.00024>

- Derderian, C., & Tadi, P. (2022). Physiology, Withdrawal Response. En *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK544292/>
- Dermitzakis, E. T., & Clark, A. G. (2002). Evolution of Transcription Factor Binding Sites in Mammalian Gene Regulatory Regions: Conservation and Turnover. *Molecular Biology and Evolution*, 19(7), 1114-1121. <https://doi.org/10.1093/oxfordjournals.molbev.a004169>
- D'haeseleer, P. (2006). What are DNA sequence motifs? *Nature Biotechnology*, 24(4), 423-425. <https://doi.org/10.1038/nbt0406-423>
- Diss, J. K. J., Calissano, M., Gascoyne, D., Djamgoz, M. B. A., & Latchman, D. S. (2008). Identification and characterization of the promoter region of the Nav1.7 voltage-gated sodium channel gene (SCN9A). *Molecular and Cellular Neuroscience*, 37(3), 537-547. <https://doi.org/10.1016/j.mcn.2007.12.002>
- Dubin, A. E., & Patapoutian, A. (2010). Nociceptors: The sensors of the pain pathway. *The Journal of Clinical Investigation*, 120(11), 3760-3772. <https://doi.org/10.1172/JCI42843>
- Emery, E. C., Luiz, A. P., & Wood, J. N. (2016). Nav1.7 and other voltage-gated sodium channels as drug targets for pain relief. *Expert Opinion on Therapeutic Targets*, 20(8), 975-983. <https://doi.org/10.1517/14728222.2016.1162295>
- Fertleman, C. R., Baker, M. D., Parker, K. A., Moffatt, S., Elmslie, F. V., Abrahamsen, B., Ostman, J., Klugbauer, N., Wood, J. N., Gardiner, R. M., & Rees, M. (2006). SCN9A Mutations in Paroxysmal Extreme Pain Disorder: Allelic Variants Underlie Distinct Channel Defects and Phenotypes. *Neuron*, 52(5), 767-774. <https://doi.org/10.1016/j.neuron.2006.10.006>
- Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Kohn, A., Twik, M., Safran, M., Lancet, D., & Cohen, D. (2017a). GeneHancer: Genome-wide integration of enhancers and target genes in GeneCards. *Database: The Journal of Biological Databases and Curation*, 2017, bax028. <https://doi.org/10.1093/database/bax028>
- Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Kohn, A., Twik, M., Safran, M., Lancet, D., & Cohen, D. (2017b). GeneHancer: Genome-wide integration of enhancers and target genes in GeneCards. *Database:*

- The Journal of Biological Databases and Curation*, 2017, bax028.
<https://doi.org/10.1093/database/bax028>
- Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M., & Dubchak, I. (2004). VISTA: Computational tools for comparative genomics. *Nucleic Acids Research*, 32(Web Server issue), W273-W279. <https://doi.org/10.1093/nar/gkh458>
- Gingras, J., Smith, S., Matson, D. J., Johnson, D., Nye, K., Couture, L., Feric, E., Yin, R., Moyer, B. D., Peterson, M. L., Rottman, J. B., Beiler, R. J., Malmberg, A. B., & McDonough, S. I. (2014). Global Nav1.7 Knockout Mice Recapitulate the Phenotype of Human Congenital Indifference to Pain. *PLOS ONE*, 9(9), e105895. <https://doi.org/10.1371/journal.pone.0105895>
- Gompel, N., Prud'homme, B., Wittkopp, P. J., Kassner, V. A., & Carroll, S. B. (2005). Chance caught on the wing: Cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature*, 433(7025), 481-487. <https://doi.org/10.1038/nature03235>
- Ha, D., Kim, D., Kim, I., Oh, Y., Kong, J., Han, S. K., & Kim, S. (2022). Evolutionary rewiring of regulatory networks contributes to phenotypic differences between human and mouse orthologous genes. *Nucleic Acids Research*, 50(4), 1849-1863. <https://doi.org/10.1093/nar/gkac050>
- Han, S. K., Kim, D., Lee, H., Kim, I., & Kim, S. (s. f.). *Divergence of Noncoding Regulatory Elements Explains Gene-Phenotype Differences between Human and Mouse Orthologous Genes*. 15.
- Hashim, F. A., Mabrouk, M. S., & Al-Atabany, W. (2019). *Review of Different Sequence Motif Finding Algorithms*. 11(2), 19.
- Hood, L., & Rowen, L. (2013). The Human Genome Project: Big science transforms biology and medicine. *Genome Medicine*, 5(9), 79. <https://doi.org/10.1186/gm483>
- Howe, K. L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Azov, A. G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., Davidson, C., Dodiya, K., El Houdaigui, B., Fatima, R., ... Flicek, P. (2020). Ensembl 2021. *Nucleic Acids Research*, 49(D1), D884-D891. <https://doi.org/10.1093/nar/gkaa942>
- Hubbard, T., Barker, D., Birney, E., Cameron, G., Chen, Y., Clark, L., Cox, T., Cuff, J., Curwen, V., Down, T., Durbin, R., Eyraas, E., Gilbert, J., Hammond, M.,

- Huminięcki, L., Kasprzyk, A., Lehtaslaiho, H., Lijnzaad, P., Melsopp, C., ... Clamp, M. (2002). The Ensembl genome database project. *Nucleic Acids Research*, 30(1), 38-41.
- Jha, S. K., Karna, B., & Goodman, M. B. (2022). Erythromelalgia. En *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK557787/>
- Jin, F., Li, Y., Ren, B., & Natarajan, R. (2011). Enhancers. *Transcription*, 2(5), 226-230. <https://doi.org/10.4161/trns.2.5.17712>
- Krijger, P. H. L., & de Laat, W. (2016). Regulation of disease-associated gene expression in the 3D genome. *Nature Reviews Molecular Cell Biology*, 17(12), 771-782. <https://doi.org/10.1038/nrm.2016.138>
- Kulakovskiy, I. V., Vorontsov, I. E., Yevshin, I. S., Sharipov, R. N., Fedorova, A. D., Rumynskiy, E. I., Medvedeva, Y. A., Magana-Mora, A., Bajic, V. B., Papatsenko, D. A., Kolpakov, F. A., & Makeev, V. J. (2018). HOCOMOCO: Towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Research*, 46(Database issue), D252-D259. <https://doi.org/10.1093/nar/gkx1106>
- Lewin, G. R. (2013). Natural Selection and Pain Meet at a Sodium Channel. *Science*, 342(6157), 428-429. <https://doi.org/10.1126/science.1244375>
- Liljas, L. (2013). Consensus Sequences. En S. Maloy & K. Hughes (Eds.), *Brenner's Encyclopedia of Genetics (Second Edition)* (pp. 163-164). Academic Press. <https://doi.org/10.1016/B978-0-12-374984-0.00325-9>
- Lis, M., & Walther, D. (2016). The orientation of transcription factor binding site motifs in gene promoter regions: Does it matter? *BMC Genomics*, 17, 185. <https://doi.org/10.1186/s12864-016-2549-x>
- Lopes, D. M., Denk, F., & McMahon, S. B. (2017). The Molecular Fingerprint of Dorsal Root and Trigeminal Ganglion Neurons. *Frontiers in Molecular Neuroscience*, 10, 304. <https://doi.org/10.3389/fnmol.2017.00304>
- McDermott, L. A., Weir, G. A., Themistocleous, A. C., Segerdahl, A. R., Blesneac, I., Baskozos, G., Clark, A. J., Millar, V., Peck, L. J., Ebner, D., Tracey, I., Serra, J., & Bennett, D. L. (2019). Defining the Functional Role of NaV1.7 in Human Nociception. *Neuron*, 101(5), 905-919.e8. <https://doi.org/10.1016/j.neuron.2019.01.047>

- Mao, W., Zhang, J., Körner, H., Jiang, Y., & Ying, S. (2019). The Emerging Role of Voltage-Gated Sodium Channels in Tumor Biology. *Frontiers in Oncology*, 9. <https://www.frontiersin.org/article/10.3389/fonc.2019.00124>
- Maston, G. A., Evans, S. K., & Green, M. R. (2006). Transcriptional Regulatory Elements in the Human Genome. *Annual Review of Genomics and Human Genetics*, 7(1), 29-59. <https://doi.org/10.1146/annurev.genom.7.080505.115623>
- Models behind the mystery of establishing enhancer-promoter interactions | Elsevier Enhanced Reader.* (s. f.). <https://doi.org/10.1016/j.ejcb.2021.151170>
- Moses, A. M., Pollard, D. A., Nix, D. A., Iyer, V. N., Li, X.-Y., Biggin, M. D., & Eisen, M. B. (2006). Large-Scale Turnover of Functional Transcription Factor Binding Sites in *Drosophila*. *PLoS Computational Biology*, 2(10), e130. <https://doi.org/10.1371/journal.pcbi.0020130>
- Nassar, M. A., Stirling, L. C., Forlani, G., Baker, M. D., Matthews, E. A., Dickenson, A. H., & Wood, J. N. (2004). Nociceptor-specific gene deletion reveals a major role for Nav1.7 (PN1) in acute and inflammatory pain. *Proceedings of the National Academy of Sciences of the United States of America*, 101(34), 12706-12711. <https://doi.org/10.1073/pnas.0404915101>
- Osterwalder, M., Barozzi, I., Tissières, V., Fukuda-Yuzawa, Y., Mannion, B. J., Afzal, S. Y., Lee, E. A., Zhu, Y., Plajzer-Frick, I., Pickle, C. S., Kato, M., Garvin, T. H., Pham, Q. T., Harrington, A. N., Akiyama, J. A., Afzal, V., Lopez-Rios, J., Dickel, D. E., Visel, A., & Pennacchio, L. A. (2018). Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature*, 554(7691), 239-243. <https://doi.org/10.1038/nature25461>
- Otto, W., Stadler, P. F., López-Giraldéz, F., Townsend, J. P., Lynch, V. J., & Wagner, G. P. (2009). Measuring Transcription Factor–Binding Site Turnover: A Maximum Likelihood Approach Using Phylogenies. *Genome Biology and Evolution*, 1, 85-98. <https://doi.org/10.1093/gbe/evp010>
- Palazzo, A. F., & Gregory, T. R. (2014). The Case for Junk DNA. *PLoS Genetics*, 10(5), e1004351. <https://doi.org/10.1371/journal.pgen.1004351>
- Panigrahi, A., & O'Malley, B. W. (2021). Mechanisms of enhancer action: The known and the unknown. *Genome Biology*, 22(1), 108. <https://doi.org/10.1186/s13059-021-02322-1>

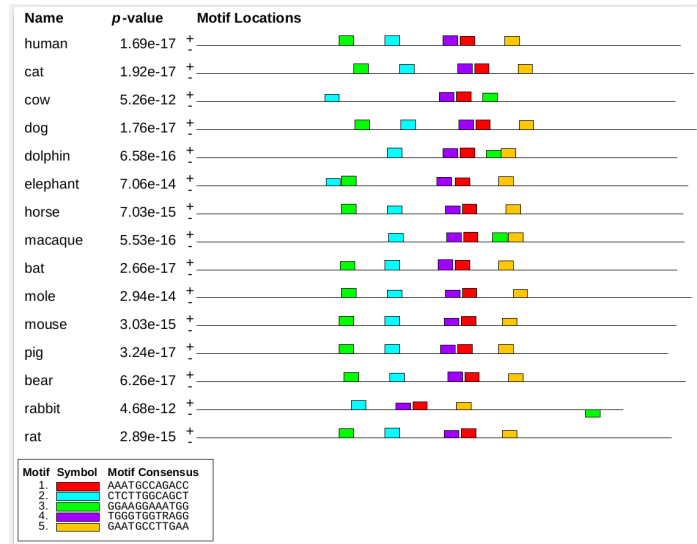
- Peng, Y., & Zhang, Y. (2018). Enhancer and super-enhancer: Positive regulators in gene transcription. *Animal Models and Experimental Medicine*, 1(3), 169-179. <https://doi.org/10.1002/ame2.12032>
- Prud'homme, B., Gompel, N., Rokas, A., Kassner, V. A., Williams, T. M., Yeh, S.-D., True, J. R., & Carroll, S. B. (2006). Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature*, 440(7087), 1050-1053. <https://doi.org/10.1038/nature04597>
- Raja, S. N., Carr, D. B., Cohen, M., Finnerup, N. B., Flor, H., Gibson, S., Keefe, F., Mogil, J. S., Ringkamp, M., Sluka, K. A., Song, X.-J., Stevens, B., Sullivan, M., Tutelman, P., Ushida, T., & Vader, K. (2020). The Revised IASP definition of pain: Concepts, challenges, and compromises. *Pain*, 161(9), 1976-1982. <https://doi.org/10.1097/j.pain.0000000000001939>
- Scacheri, C. A., & Scacheri, P. C. (2015a). Mutations in the non-coding genome. *Current opinion in pediatrics*, 27(6), 659-664. <https://doi.org/10.1097/MOP.0000000000000283>
- Scacheri, C. A., & Scacheri, P. C. (2015b). Mutations in the non-coding genome. *Current opinion in pediatrics*, 27(6), 659-664. <https://doi.org/10.1097/MOP.0000000000000283>
- SMART. (s. f.). Servier Medical Art. Recuperado 18 de enero de 2022, de <https://smart.servier.com/>
- Smith, E. St. J., Omerbašić, D., Lechner, S. G., Anirudhan, G., Lapatsina, L., & Lewin, G. R. (2011). The Molecular Basis of Acid Insensitivity in the African Naked Mole-Rat. *Science*, 334(6062), 1557-1560. <https://doi.org/10.1126/science.1213760>
- Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., Stein, T. I., Nudel, R., Lieder, I., Mazor, Y., Kaplan, S., Dahary, D., Warshawsky, D., Guan-Golan, Y., Kohn, A., Rappaport, N., Safran, M., & Lancet, D. (2016). The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Current Protocols in Bioinformatics*, 54(1). <https://doi.org/10.1002/cpbi.5>
- Tuğrul, M., Paixão, T., Barton, N. H., & Tkačik, G. (2015). Dynamics of Transcription Factor Binding Site Evolution. *PLoS Genetics*, 11(11), e1005639. <https://doi.org/10.1371/journal.pgen.1005639>

- Venkataram, S., & Fay, J. C. (2010). Is Transcription Factor Binding Site Turnover a Sufficient Explanation for Cis-Regulatory Sequence Divergence? *Genome Biology and Evolution*, 2, 851-858. <https://doi.org/10.1093/gbe/evq066>
- Yu, F. H., & Catterall, W. A. (2003). Overview of the voltage-gated sodium channel family. *Genome Biology*, 4(3), 207. <https://doi.org/10.1186/gb-2003-4-3-207>

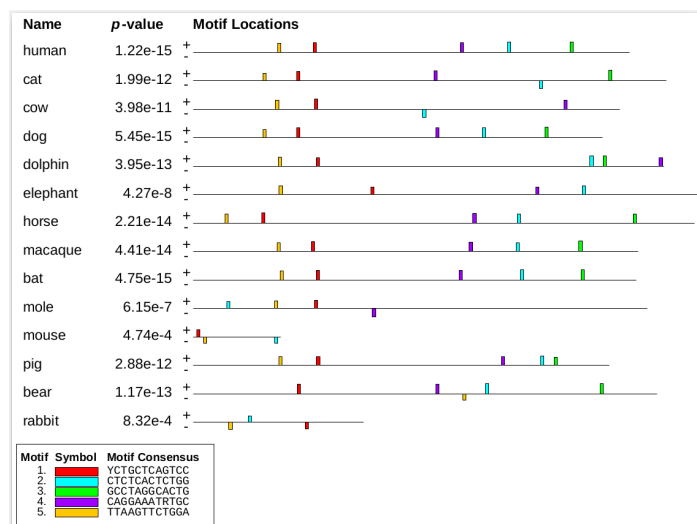
11. ANEXO

Identificación de motivos conservados entre distintas especies de mamíferos usando *MEME Suite*.

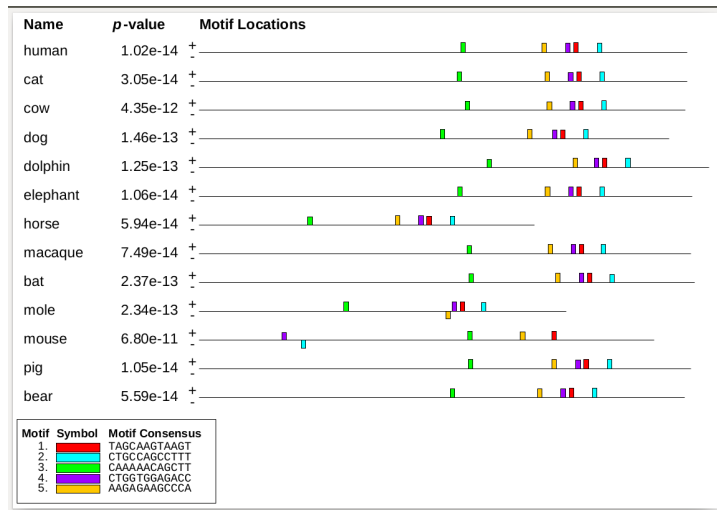
La búsqueda se realizó para cada uno de los ocho *enhancers* seleccionados. Las especificaciones dadas al programa fueron la búsqueda de 5 motivos con una longitud de entre 6 y 12 nt que tuvieran como máximo una repetición por secuencia.



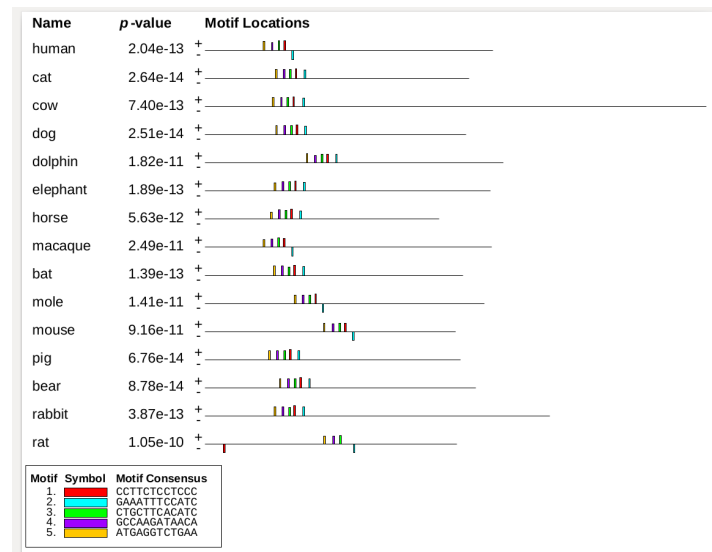
Enhancer 1



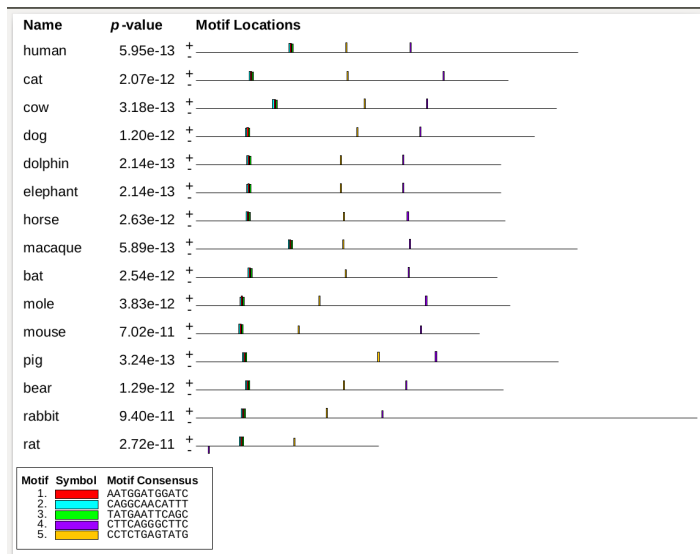
Enhancer 2



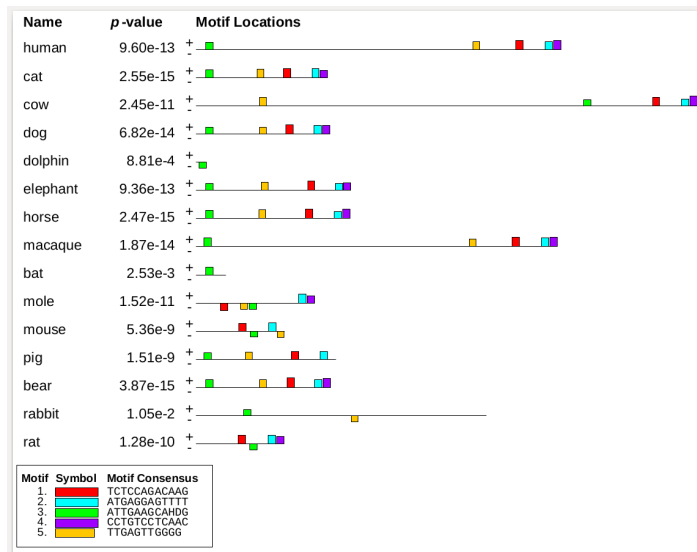
Enhancer 3



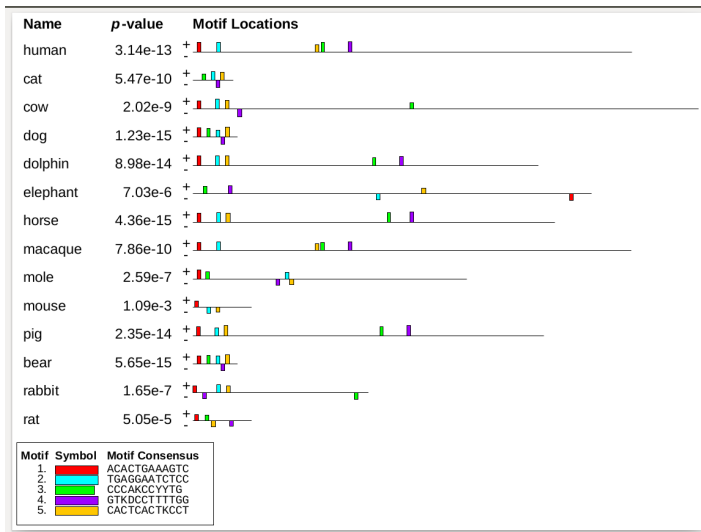
Enhancer 4



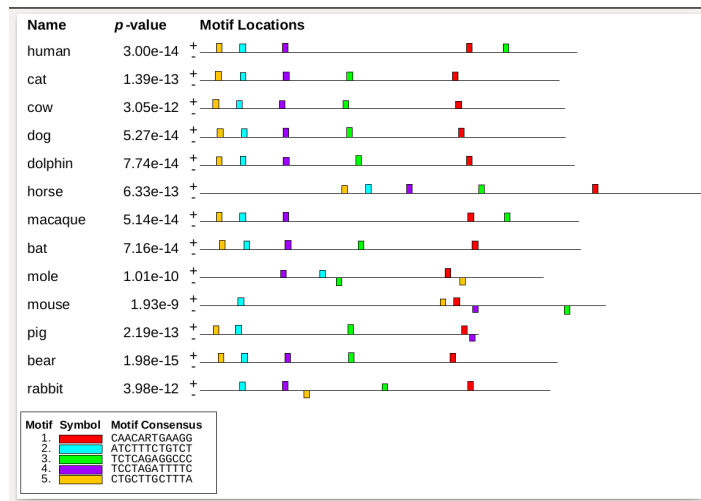
Enhancer 5



Enhancer 6



Enhancer 7



Enhancer 8