



UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales

**Bases genéticas de la evolución del cerebro humano:
estudio comparativo de la regulación transcripcional del factor de
transcripción *NPAS3***

Tesis presentada para optar al título de Doctor de la Universidad de Buenos Aires en
el área Ciencias Biológicas

Alejandro Raúl Cinalli

Director de tesis: Dra. Lucía Florencia Franchini
Consejero de Estudios: Dr. Esteban Hasson

Lugar de trabajo: Instituto de Investigaciones en Ingeniería Genética y Biología Molecular Dr.
Héctor N. Torres.

Fecha de defensa: 25/03/2019

RESUMEN

BASES GENÉTICAS DE LA EVOLUCIÓN DEL CEREBRO HUMANO: ESTUDIO COMPARATIVO DE LA REGULACIÓN TRANSCRIPCIONAL DEL FACTOR DE TRANSCRIPCIÓN *NPAS3*

Se ha propuesto que la divergencia fenotípica entre humanos y chimpancés se debe en gran medida a cambios en la regulación de genes que participan en el desarrollo en lugar de cambios en las secuencias codificantes de proteínas de los genes. Hipotetizamos que la explosión sin precedentes en el repertorio del comportamiento humano se debe a la adquisición de nuevos patrones de expresión de genes preexistentes. Previamente en nuestro laboratorio, se analizó la distribución en el genoma de secuencias genómicas no codificantes (denominados *human accelerated elements*, HAE) que muestran evidencia de evolución acelerada en el linaje humano (aumento en la tasa de sustitución de nucleótidos). Esto permitió identificar la agrupación más grande de HAEs en el genoma humano ubicada dentro del locus del gen que codifica para la proteína *neuronal PAS domain-containing protein 3* (*NPAS3*). *NPAS3* codifica un factor de transcripción bHLH-PAS que se expresa ampliamente en el sistema nervioso en desarrollo del ratón y de humano, y su disfunción se ha asociado con la etiología de la esquizofrenia y el síndrome bipolar en humanos. En el locus de este gen se agrupan 14 HAEs. En este trabajo, realizamos un ensayo comparativo utilizando las secuencias ortólogas de humano y de chimpancé, para determinar si los 14 elementos *NPAS3*-HAEs funcionan como activadores transcripcionales o *enhancers* en líneas de peces cebra y ratones transgénicos. Identificamos dos *NPAS3*-HAEs (HAR202 y HACNS658) en los que la versión humana perdió función como *enhancer* en comparación con la versión de chimpancé, que funciona como un fuerte *enhancer* durante el desarrollo del cerebro. Para determinar si la función *enhancer* se perdió en el linaje humano o lo que ocurrió fue una ganancia de función en el linaje de chimpancé, seleccionamos el elemento HAR202 y realizamos estudios comparativos con secuencias de mono macaco *Rhesus*, ratón y pez cebra. Encontramos que el elemento HAR202 perdió específicamente su función *enhancer* como consecuencia del proceso evolutivo en el linaje humano. Para analizar cuál es el efecto funcional de la pérdida de HAR202 sobre la expresión del gen *NPAS3* generamos

un linaje de ratones mutantes que carece de este elemento en el genoma. Además, identificamos un elemento regulatorio, HACNS96 que ganó función de activador transcripcional en el sistema nervioso central en el linaje humano.

Por otra parte, ha sido ampliamente estudiada la función del gen *NPAS3* en el desarrollo del cerebro y también en el desarrollo de las vías respiratorias. Sin embargo, ha sido postulado que este gen podría participar también en el desarrollo craneofacial ya que personas que presentan una delección en su genoma que incluye al gen *NPAS3* presentan holoprosencefalia. La holoprosencefalia (HPE) es la anomalía del desarrollo más común que afecta el patrón del prosencéfalo humano y las estructuras craneofaciales asociadas. Para investigar si *npas3* está involucrado en el desarrollo de estructuras craneofaciales, realizamos estudios de genética molecular en el pez cebra. El análisis de *npas3* de peces cebra mutantes nulos (*knockout*) y en los que se indujo una marcada reducción de la expresión (*knockdown*) reveló anormalidades craneofaciales críticas que indican que este gen desempeña un papel clave en el desarrollo de estas estructuras. Además, *npas3* demostró ser esencial para el desarrollo del sistema nervioso y del sistema branquial en el pez cebra. Curiosamente, la ablación o reducción de *npas3* afecta la expresión de *sonic hedgehog* (*shha*), un gen maestro del desarrollo que cuando se muta produce holoprosencefalia. Nuestros resultados sugieren el gen *NPAS3* desempeña un papel en el desarrollo craneofacial.

Palabras claves: Evolución acelerada, *NPAS3*, Desarrollo craneofacial, Evolución cerebro.

ABSTRACT

GENETIC BASES OF THE EVOLUTION OF THE HUMAN BRAIN: COMPARATIVE STUDY OF THE TRANSCRIPTIONAL REGULATION OF THE TRANSCRIPTION FACTOR *NPAS3*

It has been proposed that the phenotypic divergence between humans and chimpanzees is largely due to changes in the regulation of genes involved in the development rather than changes in the protein coding sequences of the genes. We hypothesize that the explosion unprecedented in the repertoire of human behavior is due to the acquisition of new patterns of expression of pre-existing genes. Previously in our laboratory, the genome distribution of non-coding genomic sequences (called human accelerated elements, HAE) that show evidence of accelerated evolution in the human lineage (increase in the rate of nucleotide substitution) was analyzed. This allowed to identify the largest group of HAEs in the human genome located within the locus of the gene that codes for the neuronal protein PAS domain-containing protein 3 (*NPAS3*). *NPAS3* encodes a transcription factor bHLH-PAS that is widely expressed in the developing nervous system of the mouse and human, and its dysfunction has been associated with the etiology of schizophrenia and bipolar syndrome in humans. In the locus of this gene, 14 HAEs are grouped. In this work, we conducted a comparative trial using the human and chimpanzee orthologous sequences, to determine if the 14 *NPAS3*-HAEs elements function as transcriptional activators or *enhancers* in lines of zebrafish and transgenic mice. We identified two *NPAS3*-HAEs (HAR202 and HACNS658) in which the human version lost function as an *enhancer* compared to the chimpanzee version, which functions as a strong *enhancer* during brain development. To determine if the *enhancer* function was lost in the human lineage or what happened was a gain of function in the chimpanzee lineage, we selected the HAR202 element and made comparative studies with sequences of Rhesus macaque, mouse and zebrafish monkey. We found that the HAR202 element specifically lost its *enhancer* function because of the evolutionary process in the human lineage. To analyze what is the functional effect of the loss of HAR202 on the expression of the *NPAS3* gene, we generated a lineage of mutant mice lacking this element

in the genome. We also identified a regulatory element, HACNS96 that gained transcriptional activator function in the central nervous system in the human lineage.

On the other hand, the role of the *NPAS3* gene in the development of the brain and in the development of the respiratory tract has been widely studied. However, it has been postulated that this gene could also participate in the craniofacial development since people who present a deletion in their genome that includes the *NPAS3* gene present holoprosencephaly. Holoprosencephaly (HPE) is the most common developmental anomaly that affects the pattern of the human forebrain and the associated craniofacial structures. To investigate if *npas3* is involved in the development of cranio-facial structures, we conducted molecular genetic studies in zebrafish. The analysis of *npas3* of zebrafish mutants' nulls (*knockout*) and in which a marked reduction of expression was induced (*knockdown*) revealed critical skull-facial abnormalities that indicate that this gene plays a key role in the development of these structures. In addition, *npas3* proved to be essential for the development of the nervous system and the branchial system in the zebrafish. Interestingly, the ablation or reduction of *npas3* affects the expression of sonic hedgehog (*shha*), a master gene of development that, when mutated, produces holoprocensephalia. Our results suggest the *NPAS3* gene plays a role in craniofacial development.

Keywords: Accelerated evolution, *NPAS3*, craniofacial development, Brain evolution

AGRADECIMIENTOS

A Lucía Franchini, por abrirme las puertas de su laboratorio.

A Gretel Kamm, Laura Gonzalez, Milena Rosenzvit, cuyos experimentos fueron fundamentales para estos resultados.

A Marcelo Rubinstein y a todo el laboratorio de *Genes, cerebro y conducta*, muy especialmente a Marta Treimun y sus manos, por todos los ratones transgénicos y mutantes producidos.

A los técnicos del acuario y bioterio.

Al CONICET, AGENCIA, INGEBI, y la Facultad de Cs. Exactas y Naturales de la UBA, que pusieron los recursos financieros.

A María Eugenia Gómez-Casati y Aitana Castro, por la colaboración en temas que superaban mi conocimiento.

A los compañeros del grupo, Mati, Fran, Mile, Lean, Lari, Cata, Anita, Álvaro, por ser un grupo que supo ensamblarse a la perfección.

Al laboratorio 222, por hacer del ambiente laboral un lugar sumamente placentero donde se puede trabajar a gusto.

A toda la gente fuera del ámbito científico, familia, amigos, Flor, que sirvieron como base y soporte durante todos estos años.

A Lucía y Flor, quienes por su paciencia y ayuda fueron esenciales en el proceso de elaboración de esta tesis.

ABREVIATURAS

ADN	Ácido desoxirribonucleico
ANC	<i>Accelerated non-coding conserved sequences</i>
ARNm	Ácido ribonucleico mensajero
bHLH-PAS	basic helix-loop-helix, period/aryl hydrocarbon receptor/single minded
CNS	Conserved non-coding sequence
dNTP	<i>deoxynucleotide triphosphates</i>
Dpc	Days post coitum/ Días post coito
Dpf	Days post coitum/ Días post fecundación
Dr	<i>Danio rerio</i> / Pez cebra
EGFP/ EGFP	<i>Enhanced green fluorescent protein: Gen/ Proteína</i>
FT	Factor de transcripción
Gg	<i>Gallus gallus</i> / Pollo
HACNS	<i>Human accelerated conserved non-coding sequence</i>
HAR	<i>Human Accelerated Regions</i>
HAEs	Elementos acelerados en humanos (<i>Human accelerated elements</i>)
HPE	Holoprosencephaly/ Holoprosencefalia
Hpf	Horas post-fecundación
Hn	<i>Homo neanderthalensis</i> / Neandertal
Hs	<i>Homo sapiens</i> / Humano
HTBE	<i>human terminal branch elements</i>
IF	Inmunofluorescencia
ISH/ WMISH	Hibridación <i>in situ</i> (<i>In situ hybridization</i>)/ <i>Whole mount in situ hybridization</i>
Kbp	Kilo pares de bases/ Mil pares de bases
KO	<i>Knockout</i>
<i>lacZ</i>	Gen de <i>Escherichia coli</i> que codifica para la proteína β -galactosidasa
Ma	Millones de años
Mb	Mega pares de bases/ Millón de pares de bases
Mm	<i>Mus musculus</i> / Ratón común
Mmul	<i>Macaca mulatta</i> / Mono macaco Rhesus
MO	Morfolino
NPAS3/ NPAS3	<i>Neuronal PAS domain-containing protein 3: Gen/ Proteína de humano</i>
<i>Npas3</i> / <i>Npas3</i>	<i>Neuronal PAS domain-containing protein 3: Gen/ Proteína de ratón</i>
<i>npas3</i> / <i>Npas3</i>	<i>Neuronal PAS domain-containing protein 3: Gen/ Proteína de pez cebra</i>
PCR	<i>Polymerase Chain Reaction</i>
pb	Pares de bases
Pt	<i>Pan troglodytes</i> / Cimpancé
RefSeq	<i>Reference Sequence</i>
SgRNA	Signal RNA/ ARN guía
SNC	Sistema nervioso central
TAD	Dominio de asociación topológica/ Topologically associating domain
UCSC	<i>University of California, Santa Cruz</i>
wt	<i>wild type</i> / tipo salvaje

INDICE

RESUMEN	II
ABSTRACT.....	IV
AGRADECIMIENTOS.....	VI
ABREVIATURAS	VII
INDICE.....	VIII
Introducción general.....	1
Nuestra herencia primate	2
Historia breve de nuestro linaje.....	3
Evolución del cerebro humano	7
Bases genéticas de la evolución humana	12
Los aportes del campo de la evolución del desarrollo.....	15
El código regulatorio del genoma	17
Los elementos regulatorios	18
Predicción de <i>Enhancers</i>	21
Genes blanco de los <i>enhancers</i>	25
Ensayos funcionales para regiones regulatorias.....	27
Origen y evolución de los <i>enhancers</i>	30
La evolución regulatoria de nuestro genoma	31
Detección de evolución acelerada en el genoma humano	34
Caracterización funcional de los HAEs.....	36
Del genoma al estudio de un gen específico	37
HIPÓTESIS GENERAL	41
OBJETIVO GENERAL	41
Capítulo I.....	43
INTRODUCCIÓN ESPECÍFICA.....	44
HIPÓTESIS DE TRABAJO	55
OBJETIVOS ESPECÍFICOS	55
RESULTADOS Y DISCUSIÓN	56
1RA PARTE: ANÁLISIS FUNCIONAL DE LOS NPAS3-HAES EN PECES CEBRA	56
Comparación de la funcionalidad de las regiones aceleradas humanas entre líneas estables y transitorias de peces cebras transgénicos.....	56

Comparación de la funcionalidad de las regiones aceleradas humanas con la de sus ortólogos de chimpancé	58
Comparación de la funcionalidad de las regiones aceleradas humanas con la de sus ortólogos del mono macaco Rhesus	62
Análisis funcional de distintas secuencias ortólogas de vertebrados del HAR202	65
2DA PARTE: ESTUDIO DE LOS ELEMENTOS ACELERADOS DEL LOCUS DE NPAS3 EN RATONES	68
Comportamiento del elemento acelerado no codificante HAR202Hs durante el desarrollo embrionario en ratones transgénicos	68
Comportamiento del elemento acelerado no codificante HAR202Pt durante el desarrollo embrionario en ratones transgénicos	72
El dominio de asociación topológica TAD que contiene al locus del gen NPAS3 contiene también a los 14 HAEs	75
Estudio de marcas epigenéticas en el locus de NPAS3	78
Generación de ratones mutantes portadores de una delección del elemento HAR202	81
Capítulo II	84
INTRODUCCIÓN ESPECÍFICA	85
<i>Trachealess</i> , el gen ortólogo de NPAS3 en mosca	85
Patrón de expresión de NPAS3 en el sistema nervioso de mamíferos	87
Análisis de ratones mutantes del gen <i>npas3</i>	87
NPAS3, un gen cuya disfunción está asociada al desarrollo de esquizofrenia y otras patologías del sistema nervioso	88
NPAS3 y el desarrollo craneofacial	90
Estudios funcionales en peces cebra: uso de morfolinós	92
HIPÓTESIS DE TRABAJO	94
OBJETIVOS ESPECÍFICOS	94
RESULTADOS	95
Estudios filogenéticos para identificar el ortólogo NPAS3 / <i>trh</i> del pez cebra	95
Expresión de <i>Npas3</i> durante el desarrollo del pez cebra	97
NPAS3 juega una función clave en el desarrollo de pez cebra.	99
El knockout de <i>Npas3</i> produce un fenotipo similar al de los <i>knockdown</i>	105
<i>Npas3</i> es importante para el desarrollo del sistema craneofacial y branquial.	108
NPAS3 afecta la supervivencia en el pez cebra	109
La caída de <i>Npas3</i> afecta el desarrollo del sistema nervioso y la proliferación neural.	110
Expresión genética “río abajo” alterada en knock downs de <i>npas3</i>	113
<i>Npas3</i> participa en el desarrollo de la vejiga natatoria.	119
<i>Npas3</i> actúa en el desarrollo de estructuras craneales en ratones	120
DISCUSIÓN	124
Conclusiones	128

Materiales	133
y métodos.....	133
PRODUCCIÓN DE PECES CEBRA TRANSGÉNICOS	134
Construcción de los transgenes para pez cebra.....	134
Transcripción <i>in vitro</i> del ARN que codifica para la transposasa <i>Tol2</i>	138
Animales y acuario	139
Obtención de embriones para inyección.....	139
Inyección de los embriones con la transposasa.....	140
Generación de líneas de peces cebra transgénicos	141
Análisis del patrón de expresión de la proteína reportera EGFP	141
Genotipificación de los peces	142
PRODUCCIÓN DE RATONES TRANSGÉNICOS.....	143
Construcción de transgenes para ratón	143
Preparación de los transgenes para microinyección pronuclear	143
Animales y bioterio	144
Producción de embriones para microinyección	145
Microinyección pronuclear	146
Transferencia de los embriones inyectados a hembras pseudo preñadas	146
Genotipificación de los ratones transgénicos.....	147
Análisis de la expresión de LacZ en ratones transgénicos	148
HIBRIDACIÓN <i>IN SITU</i> EN CORTES DE CRIÓSTATO	149
HIBRIDACIÓN <i>IN SITU</i> EN EMBRIONES COMPLETOS DE PEZ CEBRA.....	151
INYECCIÓN DE MORFOLINOS.....	153
EDICIÓN DEL GENOMA CRISPR / CAS9 EN PECES CEBRA.....	154
EDICIÓN DEL GENOMA CRISPR / CAS9 EN RATONES	155
Tinción de azul de Alcian de peces	157
Tinción de azul de Alcian de ratones	157
Gel de electroforesis y western blot	157
Inmunofluorescencia	158
ANÁLISIS BIOINFORMÁTICOS.....	159
Adquisición de secuencias	159
Análisis filogenético	159
Bibliografía	161
Apéndice.....	180
Figuras suplementarias	180

Introducción general

Nuestra herencia primate

Homo sapiens es una de las aproximadamente 200 especies vivientes que colectivamente constituyen el orden Primates (Purvis 1995). Este grupo se originó hace aproximadamente 65 millones de años (Tavaré et al. 2002). Los primates se clasifican en dos grandes subórdenes: Strepsirrhini y Haplorrhini. Los strepsirrhinos se conocen comúnmente como prosimios e incluyen a los lemures, lorinos y gálagos. Por otro lado, los haplorrhinos consisten en los tarseros y los antropoideos. Los antropoideos comprenden a los platirrinios (o monos del nuevo mundo, como el mono araña y el mono tití) y los catarrhinos. Dentro de este último grupo se agrupan los cercopitécidos (o monos del viejo mundo como los macacos, los colobos y los babuinos) y los grandes simios u hominoideos (gibones, orangutanes, gorilas, chimpancés y humanos) (Purvis 1995; Begun 2003) (figura 1).

Tanto los primates actuales como los fósiles constituyen un conjunto diverso de especies. Algunas especies de primates están entre las que retuvieron más características ancestrales dentro de la clase de los mamíferos; otras muestran especializaciones morfológicas y de comportamiento completamente novedosas dentro de los mamíferos (Martin 1986; Lewin y Foley 2004; Fleagle 2013). Históricamente, se han estudiado las características de los primates tratando de identificar las autapomorfías de los humanos. Sin embargo, muchas características a veces consideradas específicas de los humanos como la capacidad de caminar erguidos, gran inteligencia y una gran sociabilidad, más que representar una discontinuidad con el resto de los primates, son una extensión de las características de estos como grupo (Martin 1990).

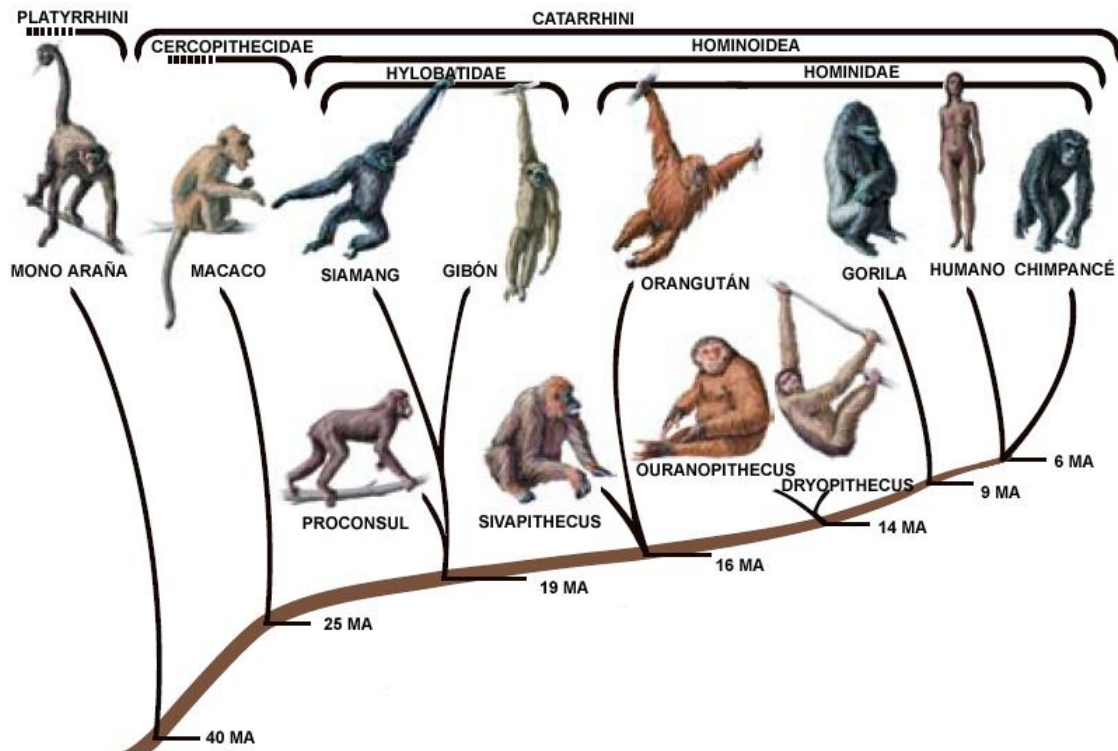


Figura 1. Árbol filogenético que muestra las relaciones de parentesco probables entre distintos grupos de primates actuales y extintos. Abreviatura: MA, millones de años atrás. Modificado de (Begun 2003).

Historia breve de nuestro linaje

Para poder dilucidar los cambios genéticos responsables de los rasgos humanos, es esencial comprender la historia evolutiva de nuestro linaje. La información del registro fósil actualmente disponible sugiere dos patrones importantes en la historia evolutiva del linaje humano. Por un lado, si bien *Homo sapiens* es la única especie viva representante nuestro linaje, desde la divergencia con el chimpancé existió una gran diversidad de formas morfológicas, muchas de las cuales coexistieron durante largos periodos de tiempo. Estas diversas formas fueron interpretadas como múltiples especies (Carroll 2003). Por otro lado, las características morfológicas que diferencian al *H. sapiens* de otros primates no se desarrollaron simultáneamente durante la historia del linaje humano. Por lo contrario, se observa un patrón de evolución en mosaico, donde las diferentes estructuras morfológicas cambiaron a tasas diferentes y en momentos distintos (McHenry 1975; Jurmain et al. 2014).

Las estimaciones actuales sostienen que los primeros homínidos pudieron haber divergido de los primeros representantes del linaje de los chimpancés en algún momento durante el Mioceno tardío o Plioceno temprano, es decir entre 8 y 4 millones de años atrás, en África (Carroll 2003; Lewin y Foley 2004). Si bien *sapiens* es el único representante vivo del linaje Homininae, desde la divergencia con el chimpancé existió una gran diversidad de especies, muchas de las cuales coexistieron durante largos periodos de tiempo. La evolución de los rasgos “modernos” de *H. sapiens* a partir de este antepasado no fue un proceso gradual, lineal y aditivo, sino sumamente complejo (Carroll 2003; Tattersall 2010). Algunos rasgos, como el bipedalismo, surgieron muy tempranamente mientras que otros, como el aumento del tamaño cerebral, no evolucionaron hasta hace relativamente poco tiempo.

Las relaciones de parentesco, orden cronológico, la magnitud y ritmo de la evolución de los caracteres en este linaje se encuentran constantemente sujetos a revisión a medida que ocurren nuevos hallazgos fósiles (Tattersall 2010). La Figura 2 muestra un árbol provisorio donde se resume la diversidad, pero no todas las relaciones de parentesco entre los miembros de la familia Hominidae. Los homínidos más antiguos que se conocen en la actualidad son *Sahelanthropus tchadensis* de Chad, fechado entre 6-7 Ma atrás (Brunet et al. 2005) y *Orrorin tugenensis* de Kenia (Pickford y Senut 2001) (Figura 2). Su tamaño cerebral, ~360 cc, se encuentra dentro del rango que se observa en los chimpancés, y el cráneo porta una cresta supraorbital de gran tamaño, similar a la de los gorilas (Brunet et al. 2005). Sin embargo, la posición y orientación del foramen magnum, sugiere que *Sahelanthropus* presentaba bipedalismo y su columna vertebral sostenía verticalmente como en los humanos modernos en lugar de horizontalmente como en simios y otros cuadrúpedos (Zollikofer et al. 2005). Alrededor de 4 Ma atrás habrían aparecido los primeros miembros del género *Australopithecus*, homínidos bípedos, cuyo cráneo tiene un tamaño que oscila entre 390 y 515 cc, similar a los chimpancés y gorilas (Falk et al. 2000). Esto sugiere que tendrían habilidades cognitivas similares a la de los simios actuales (McHenry 1982). *Australopithecus* fue un género muy exitoso que persistió durante casi tres millones de años (Figura 2).

Hace aproximadamente 2.5 millones de años atrás existió una divergencia importante dentro de un grupo de australopitecinos. Este evento involucró el surgimiento de homínidos con mandíbula y dientes más pequeños. Además, este grupo mostró un incremento significativo en el tamaño del endocráneo, un pequeño aumento del tamaño corporal y

presencia de bipedalismo completo (Leakey, Tobias, y Napier 1964), a diferencia de los australopitecinos que eran bípedos facultativos (Striedter 2005). Era el surgimiento de los primeros representantes del género *Homo*. Este género estuvo compuesto por miembros bastante diversos. Algunos grupos de *Homo* tempranos, como *Homo rudolfensis* y *Homo habilis*, eran relativamente similares a los australopitecinos (Leakey, Tobias, y Napier 1964). En cambio, hace 2 millones de años atrás surgió otra especie de *Homo* en África, de cuerpo y cerebro más grande, al que se denominó *Homo ergaster* (Wood 1996). Equipado con nuevas habilidades, *H. ergaster* fue el primer homínido en migrar fuera de África hacia Asia e Indonesia hace 1.8 millones de años atrás. Fuera de África, *H. ergaster* dio origen a la especie que hoy conocemos como *Homo erectus* (Figura 2; Wood 1996; Striedter 2005; Lewin y Foley 2004).

Esta especie que llegó a habitar gran parte del Viejo Mundo, desde África hasta China e Indonesia desde 1,9 Ma a 0,14Ma. Los fósiles se han encontrado juntamente con restos que hacen suponer que esta especie fue la primera en usar el fuego para cocinar, como así también, en hacer hachas de mano a partir de piedras; algo notable si tenemos en cuenta que su capacidad craneana era de 600 cm³. Era una especie bastante corpulenta, de unos 180 cm y presentaba un marcado dimorfismo sexual, donde las hembras encontradas eran un 25% más pequeñas que los machos.

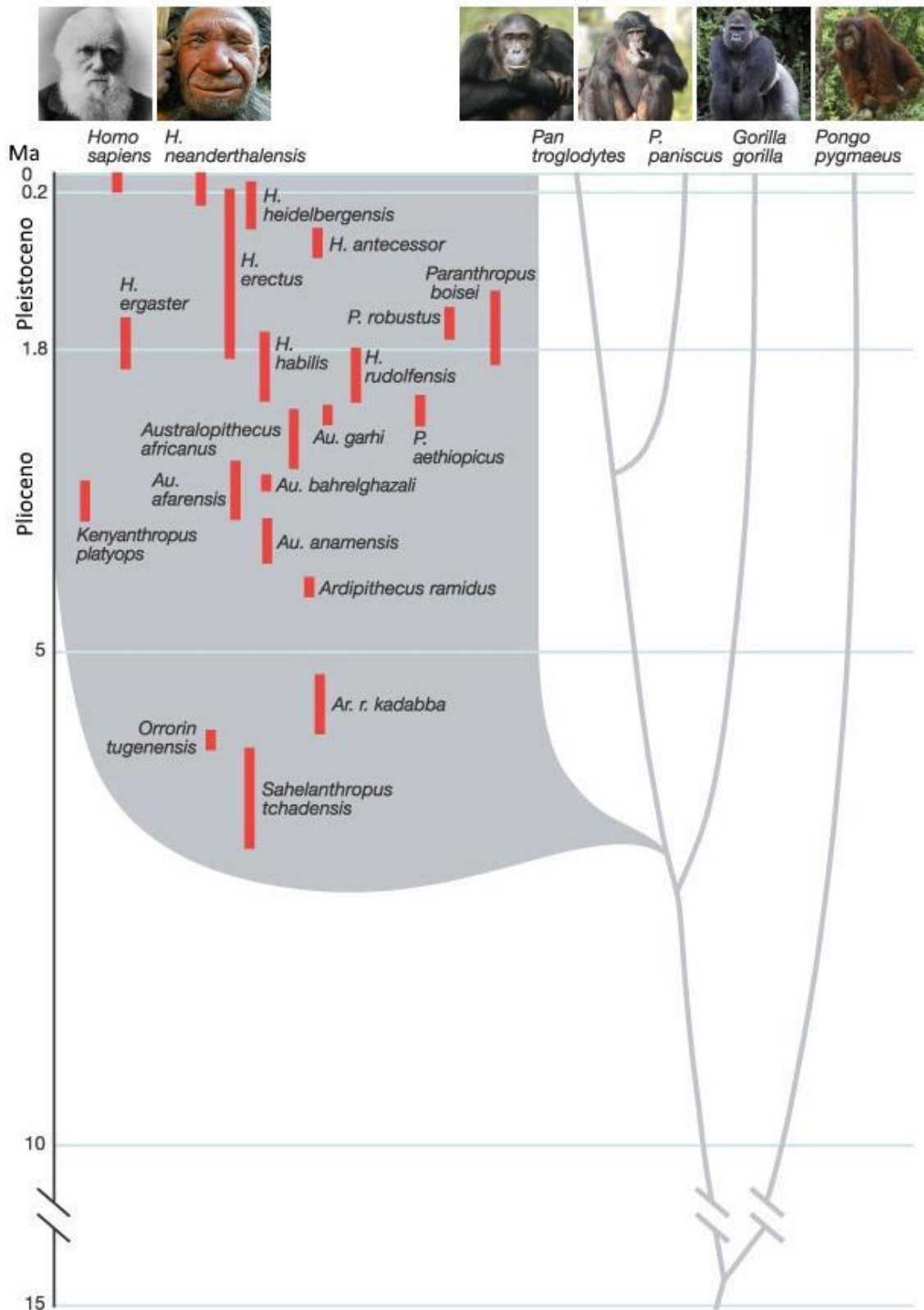


Figura 2. Escala del tiempo y filogenia de los homínidos. Los tiempos aproximados de divergencia derivan de datos moleculares. Las relaciones filogenéticas entre los homínidos son inciertas y se marcan como un sombreado gris. Las barras rojas denotan el lapso de las especies fósiles y/o la incertidumbre de las edades fósiles. Modificada de (Carroll 2003).

Existen varias hipótesis en torno al origen de los humanos modernos. En un extremo, una hipótesis sostiene que la transformación ocurrió gradualmente dentro de todas las poblaciones de *H. erectus*, desencadenando el surgimiento de múltiples poblaciones de humanos modernos en África y Eurasia. Bajo esta mirada, las raíces genéticas de las poblaciones modernas de *H. sapiens* son profundas y se encuentran en el origen de las primeras poblaciones de *H. erectus* (Thorne y Wolpoff 1992). En el otro extremo, una hipótesis alternativa plantea un origen único para *H. sapiens*, en África, seguido de una expansión poblacional hacia el resto del viejo mundo con reemplazo de las poblaciones de humanos arcaicos presentes, es decir *H. ergaster*, *H. erectus*, *H. heidelbergensis* y *H. neanderthalensis*. Bajo este escenario, las raíces genéticas de las poblaciones modernas de *H. sapiens* son más recientes y datan de aproximadamente 200.000-100.000 años atrás (Wilson y Cann 1992; Cann, Stoneking, y Wilson 1987). Cualquiera sea el caso, la evidencia muestra que hace aproximadamente 250.000-200.000 años atrás, *H. erectus* desapareció y apareció un nuevo grupo de homínidos con características consideradas anatómicamente modernas (hueso occipital menos prominente, frente más empinada y una cresta superciliar y sagital reducidas), de gran tamaño cerebral (1.400 cc) y la evidencia arqueológica de un conjunto de herramientas y técnicas de caza más sofisticado. Este grupo son los *H. sapiens* (Figura 1; Striedter 2005).

Evolución del cerebro humano

El sistema nervioso, en particular el cerebro y sus capacidades cognitivas, se encuentran entre los atributos más distintivos y fascinantes de nuestra especie. La manera en la que el sistema nervioso evolucionó en el linaje humano y cómo éste se diferencia del de nuestros parientes primates más cercanos, aún no se conoce detalladamente.

El cerebro humano tiene todas las características básicas de un cerebro de mamífero típico, empezando por una corteza de seis capas (también denominada neocorteza). Además, tiene las características típicas de un cerebro de primate, incluyendo una gran neocorteza, una gran corteza visual y una corteza prefrontal lateral (Figura 3; Preuss 2012). A pesar de estas similitudes generales, el cerebro humano tiene varias características únicas. En primer lugar, los seres humanos tienen el mayor número de neuronas entre los primates, 86 mil millones (Azevedo et al. 2009), en comparación con el número de neuronas de los chimpancés y

gorilas estimado en 28 y 33 mil millones, respectivamente (S Herculano-Houzel y Kaas 2011). Por otro lado, el cerebro humano no es el más grande de la tierra, siendo superado por los cerebros de los cetáceos y elefantes (Roth y Dicke 2005). Sin embargo, aunque el cerebro del elefante tiene cerca de 251 mil millones de neuronas, tan sólo 5.6 mil millones (2.2%) son corticales, mientras que la mayoría se concentra en el cerebelo (97%) (Suzana Herculano-Houzel et al. 2014). En contraste, 20.9% de las neuronas presentes en el cerebro humano son corticales (S. Herculano-Houzel 2012). Por lo tanto, la corteza cerebral humana es, en proporción al tamaño total del cerebro, la más grande, ocupando 84% de toda la masa cerebral, y contiene el mayor número de neuronas, 18 mil millones aproximadamente (Gabi et al. 2016). Hasta recientemente se creía que la evolución humana había supuesto una expansión particular en la corteza prefrontal del cerebro. Sin embargo, esta noción popular ha sido cuestionada recientemente, demostrando que la región prefrontal humana posee aproximadamente el 8% de las neuronas corticales, al igual que los primates no humanos (Gabi et al. 2016).

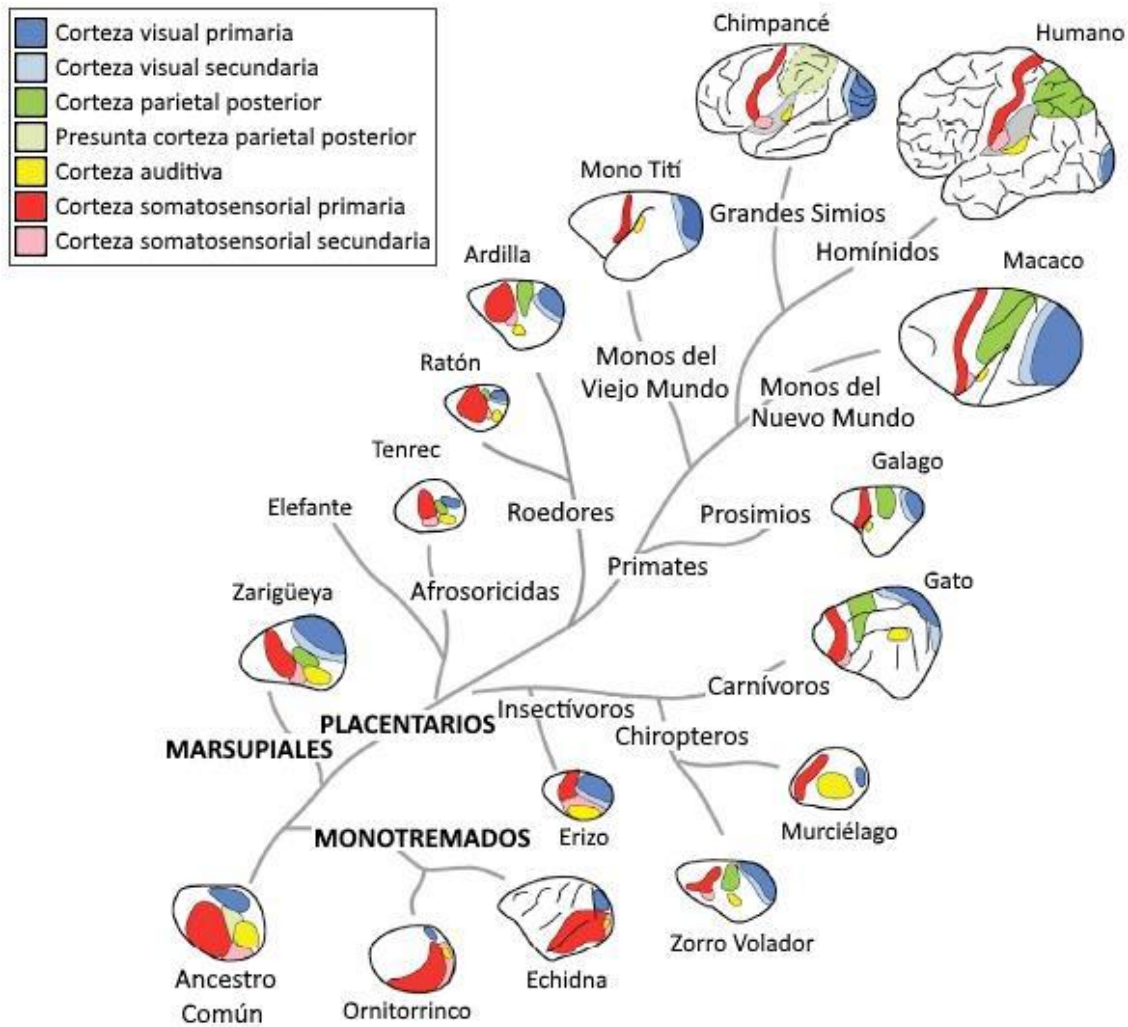


Figura 3. Ilustración de las relaciones filogenéticas entre las subclases y órdenes principales de mamíferos, indicando algunos campos corticales compartidos. Modificado de (Belmonte et al. 2015).

Si comparamos con el conjunto de los mamíferos, los primates (particularmente los diurnos) se caracterizan por presentar un sentido de la visión muy desarrollado, mientras que el sentido del olfato está más reducido (Fleagle 2013). Esto está acompañado por una tendencia evolutiva dentro de este clado a una mayor diurnalidad (Striedter 2005) e incluso en los catarrinos, existió una reinversión de la visión tricrómica (Jacobs et al. 1996; Dulai et al. 1999). Posiblemente como producto de la predominancia del sentido de la visión dentro del grupo, los cerebros de los primates son relativamente más grandes que los encontrados en otros mamíferos de peso similar (Striedter 2005). Esta característica se pone particularmente de manifiesto al comparar las reglas de escalado específicas de clado entre, por ejemplo, roedores y primates de tamaños corporales similares. Por un lado, los cerebros primates tienen una mayor densidad de neuronas en sus cerebros que los cerebros de

roedores de tamaño similar (por ejemplo 40.000 neuronas/mg en la corteza del mono *Aotus trivirgatus* contra 12.000 neuronas/mg en el roedor agutí *Dasyprocta prinnolopha*), mientras que la densidad de células no neuronales en los cerebros de estos dos grupos es comparable (Suzana Herculano-Houzel, Mota, y Lent 2006; Suzana Herculano-Houzel et al. 2007). El mayor número de neuronas por unidad de volumen presumiblemente podría conferir a los primates con mayores capacidades computacionales que roedores de tamaños equivalentes (Herculano-Houzel 2011). De hecho, los primates no humanos son extraordinariamente inteligentes. En la naturaleza, estos individuos dedican mucho tiempo a establecer una red de amistades y observar las alianzas de sus congéneres. Este tipo de interacciones sociales son muy complejas, así como las habilidades cognitivas necesarias para procesar este tipo de información (Lewin y Foley 2004). La evolución progresiva de la inteligencia como forma de vida ha sido considerada como una de las características más destacadas de los primates como grupo (Jolly 1972).

Un aspecto importante de todos los cerebros primates es que tienen cortezas cerebrales inusualmente grandes para el tamaño absoluto del cerebro. Por ejemplo, la corteza cerebral de los humanos es la más grande en tamaño relativo (con un 75,5% - 84,0% de la masa total del cerebro) (Hofman 1988; Rilling y Insel 1999). La corteza cerebral representa el 73,0% de la masa total del cerebro en los chimpancés (Stephan, Frahm, y Baron 1981). Además, con un promedio de 86 mil millones de neuronas y 85 mil millones de células no neuronales (Azevedo et al. 2009), el cerebro humano tiene tantas neuronas como se espera para un cerebro primate genérico de su tamaño, ajustándose a lo predicho por las reglas de escalado del cerebro primate (Gabi et al. 2010; S. Herculano-Houzel 2012). También las densidades neuronales en la corteza y el cerebelo se ajustan a los valores esperables para un cerebro primate (Azevedo et al. 2009; S. Herculano-Houzel 2012). En resumen, el cerebro humano se ajusta a un cerebro primate típico, en lo que respecta a número de células cerebrales totales, distribución de neuronas entre la corteza y el cerebelo y proporción de células neuronales y no neuronales, al menos (S. Herculano-Houzel 2012). Quizás, el tamaño cerebral **absoluto** sea un mejor predictor de las habilidades cognitivas excepcionales de los humanos dentro de los primates (Reader y Laland 2002; Deaner et al. 2007).

Por otro lado, respecto a la evolución del cerebro en el linaje de los homíninos, se observa que el patrón de cambio no fue gradual, sino que hubo algunos momentos de aumento abrupto. Por un lado, los australopitecinos tempranos tenían un volumen endocraneal de

aproximadamente 400 cm³, un volumen similar al hallado en los chimpancés y gorilas actuales (Lewin y Foley 2004; Striedter 2005). Por otro lado, el surgimiento del género *Homo* estuvo acompañado por un incremento abrupto del tamaño de cerebro. Por ejemplo, los rangos de volumen del endocráneo de *Homo habilis* y *H. rudolfensis* se estiman entre los 650-800 cm³ y el volumen del endocráneo de *H. ergaster* y *H. erectus* se estima en 850-1100 cm³ (McHenry 1982). Finalmente, la aparición de *H. sapiens* está asociada con un tercer evento de incremento en el tamaño del cerebro, alcanzando un volumen de entre 1100 cm³ a más de 1400 cm³ (Striedter 2005).

Las estimaciones de la cantidad de neuronas en las diferentes especies de homínidos sugieren que, con 27 a 35 mil millones, el número de neuronas de los australopitecinos tempranos se encontraría dentro del rango de los grandes simios actuales (S Herculano-Houzel y Kaas 2011). En cambio, para los *Homos* tempranos, se predice un número de neuronas de 50 a 60 mil millones de neuronas, y de 62 mil millones de neuronas en promedio para *H. erectus*. Llamativamente, para las especies de humanos arcaicos, *Homo heidelbergensis* y *Homo neanderthalensis* se predice un número de neuronas que caerían dentro del rango de los humanos modernos (Azevedo et al. 2009; S Herculano-Houzel y Kaas 2011; S. Herculano-Houzel 2012). Si el factor determinante para el desarrollo de capacidades cognitivas depende principalmente del número de neuronas, estos resultados sugieren que el potencial cognitivo de estas especies de humanos arcaicos podría haber sido comparables a las halladas en los humanos modernos (S Herculano-Houzel y Kaas 2011).

Para entender la evolución de las habilidades cognitivas únicas de nuestra especie, incluyendo el pensamiento abstracto, la planificación a largo plazo y la extraordinaria capacidad de elaborar un lenguaje complejo, hay que tener en cuenta dos cuestiones desafiantes. La primera es poder relacionar la cognición humana con el número de neuronas, el tamaño del cerebro, la corteza altamente desarrollada y, en particular, las diferencias neuroanatómicas. Por otro lado, la segunda cuestión es conectar los cambios humano-específicos en el ADN con los rasgos característicos de la especie (Franchini y Pollard 2015). Existen diversos estudios que sostienen esta hipótesis, ya que lograron identificar cambios en la morfología y abundancia tanto de neuronas excitatorias e inhibitorias como de células de la glía y diferencias en las conexiones y localización regional de distintos tipos neuronales en el neocórtex de humanos y primates (Sherwood, Subiaul, y

Zawidzki 2008; DeFelipe 2011; Elston et al. 2011; Oberheim, Goldman, y Nedergaard 2012; Raghanti et al. 2016; von Bartheld, Bahney, y Herculano-Houzel 2016).

Bases genéticas de la evolución humana

El estudio de la evolución humana desde la perspectiva de la genética molecular procura hallar los cambios a nivel genómico y genético que potencialmente pudieron haber contribuido con la adquisición de las características particulares de nuestra especie. El advenimiento de la genómica junto con el desarrollo de nuevas tecnologías en los campos de la biología molecular y de la bioinformática han aportado nuevas formas de estudiar la evolución humana. Nuestro conocimiento acerca de los genes que sufrieron selección positiva o bien modificaciones en sus niveles de expresión a lo largo de la evolución de nuestro linaje está creciendo rápidamente. Sin embargo, el mayor desafío reside en poder vincular estos cambios ocurridos a nivel genómico con el surgimiento de características neuroanatómicas novedosas y asociarlas a su vez a capacidades cognitivas concretas (Preuss 2012; Lucía F. Franchini y Pollard 2015). A la hora de querer hilar estos cambios evolutivos, la genómica y genética molecular representan herramientas valiosísimas en este campo que permiten llevar a cabo análisis comparados entre el humano y otros primates (Gibbs et al. 2007; Chimpanzee Sequencing and Analysis Consortium 2005), e incluso de los homínidos extintos (Prabhakar et al. 2006; Prüfer et al. 2014). Además, el desarrollo de las herramientas bioinformáticas necesarias para el análisis de dichas secuencias hizo que se potenciara la capacidad de examinar y encontrar los cambios genéticos específicos de los humanos.

La comparación del genoma humano con el genoma de otras especies ha catalogado estas diferencias genómicas en tres clases generales. Estas se clasifican en grandes alteraciones cromosómicas, inserciones y deleciones pequeñas denominadas *indels* y sustituciones (Figura 4; Vallender, Mekel-Bobrov, y Lahn 2008; Franchini y Pollard 2015).

Las grandes alteraciones cromosómicas involucran modificaciones estructurales, como duplicaciones, deleciones y reordenamientos, en los cromosomas (Figura 4A). Son relativamente raras, pero abarcan varios *loci* debido a su tamaño, que en general se trata de miles de pares de bases (Coe et al. 2014). Las primeras diferencias estructurales entre el genoma de los seres humanos y el genoma del chimpancé fueron halladas usando técnicas

de tinción de bandas de cromatina, donde se descubrió que hubo una fusión de dos cromosomas ancestrales para formar el cromosoma humano 2. También se describieron bandas de heterocromatina constitutiva específicas de los humanos en los cromosomas 1, 9, 16 e Y, e inversiones pericéntricas en los cromosomas 1 y 18 (Yunis y Prakash 1982). Muchas de estas variantes estructurales han alterado la expresión génica de diversos genes del desarrollo. Por ejemplo, la inversión pericéntrica en el cromosoma 1 está asociada con enfermedades neurogenéticas humanas y contiene un aumento en el número de copias de los genes del desarrollo *SRGAP2* (Dennis et al. 2012), *HYDIN* (Doggett et al. 2006) y familias proteicas conteniendo el dominio *DUF1220* (Fortna et al. 2004).

Las inserciones y deleciones pequeñas, *indels*, abarcan aproximadamente 3.5% del genoma humano, la mayoría de las cuales se encuentran en regiones no codificantes (Varki y Altheide 2005; Chimpanzee Sequencing and Analysis Consortium 2005). Los *indels* presentes en estas regiones podrían alterar el fenotipo humano mediante la modificación o la eliminación de regiones regulatorias conservadas (Bae, Jayaraman, y Walsh 2015; Carroll 2005; King y Wilson 1975). Un estudio reciente demostró que 510 regiones altamente conservadas fueron perdidas en el linaje humano, muchas de las cuales eran no codificantes (McLean et al. 2011). Además, los *indels* también pueden afectar a los genes, sus variantes de *splicing*, sus marcos de lectura y sus codones de inicio y terminación (Varki y Altheide 2005). Estos cambios han generado varios pseudogenes (Karro et al. 2007) y transcritos sin sentido humano específicos (Lareau et al. 2007), entre los cuales se encuentra la pérdida de cientos de receptores olfatorios funcionales durante la evolución humana (Gilad, Man, y Glusman 2005).

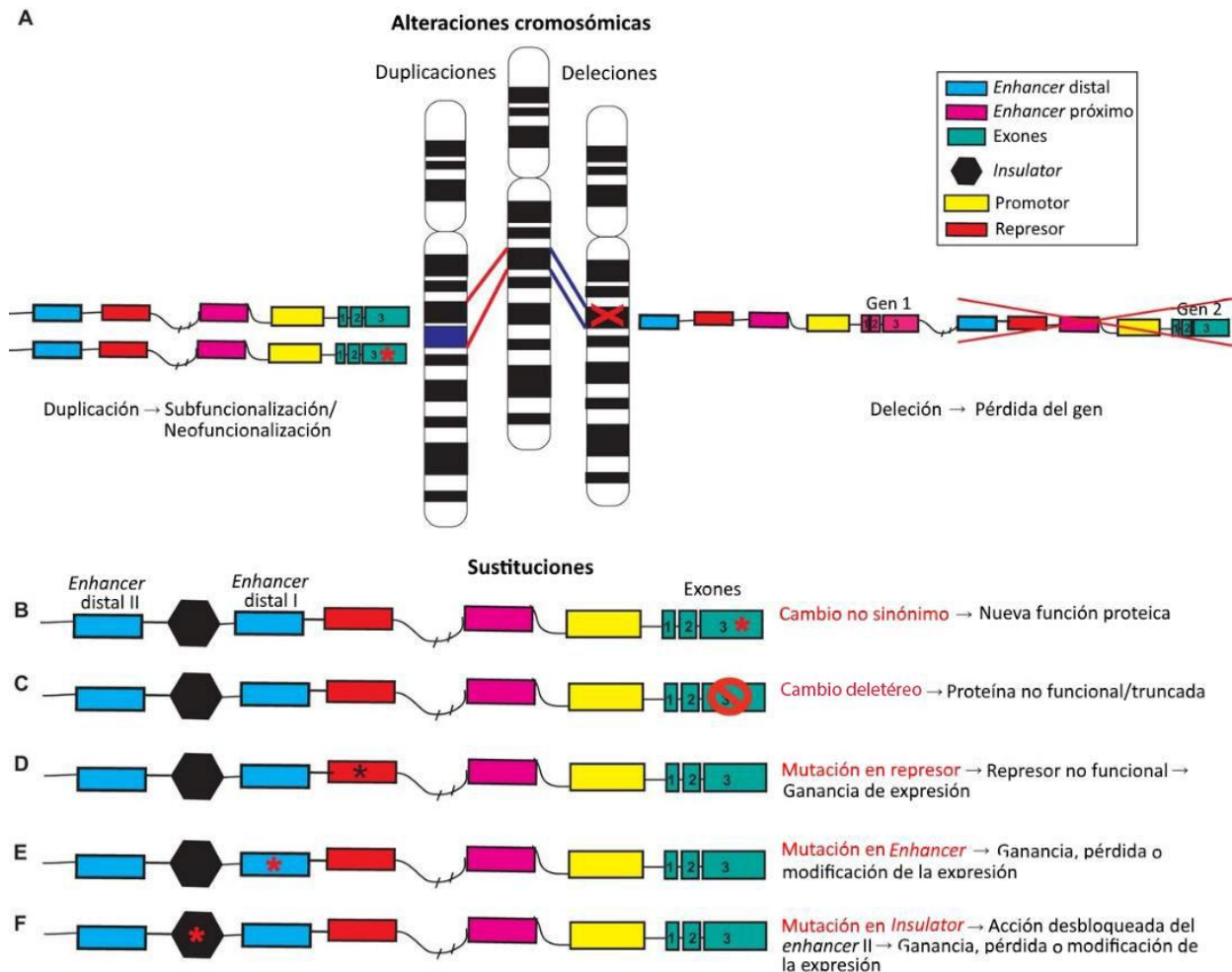


Figura 4. Cambios genéticos y genómicos con posibles consecuencias funcionales. (A) Alteraciones cromosómicas grandes. Las duplicaciones pueden llevar a la aparición de nuevos parálogos que pueden ganar nuevas funciones mediante neofuncionalización o subfuncionalización. También puede haber pérdida de genes debido a mutaciones deletéreas, produciendo pseudogenes. En contraste, las delecciones pueden llevar a la pérdida de genes y elementos regulatorios. (B) Cambios no sinónimos en la región codificante de genes pueden llevar a un cambio en la función proteica. (C) Cambios deletéreos pueden producir proteínas truncadas o pseudogenes. (D) Cambios en los elementos represores pueden llevar a una ganancia de expresión génica. (E) Cambios en los elementos *enhancer* pueden llevar a una ganancia, pérdida o modificación de la expresión génica. (F) Cambios en los elementos *insulators* pueden llevar a una ganancia, pérdida o modificación de la expresión génica como resultado de acciones desbloqueadas de los *enhancers* distales. Modificado de (Franchini y Pollard 2015).

En comparación con el genoma del chimpancé, nuestro pariente homínido más cercano, el genoma del humano difiere en aproximadamente 30 millones de sustituciones, es decir 1.2% del genoma humano, y la mayoría ocurrieron en regiones no codificantes (Figura 4B-F; Varki y Altheide 2005; Chimpanzee Sequencing and Analysis Consortium 2005). Las teorías evolutivas postulan que la mayoría de las sustituciones son neutrales y que es poco probable que hayan producido las características distintivas del ser humano. Sin embargo,

existen algunos genes del desarrollo que han adquirido cambios humano-específicos y se cree que podrían haber tenido algún rol en la evolución humana (O'Bleness et al. 2012; Vallender, Mekel-Bobrov, y Lahn 2008). Por ejemplo, se encontraron cambios en las regiones codificantes con potenciales consecuencias en la evolución del habla humana en el *locus* del gen *Forkhead Box Protein 2 (FOXP2)*. Por un lado, mutaciones en este gen fueron asociadas con defectos severos del habla en familias humanas (Hill y Walsh 2005). Adicionalmente, estudiando el patrón de cambio molecular en este gen en los primates, se identificaron dos sustituciones de aminoácidos que se fijaron en el linaje humano después de la divergencia con el linaje del chimpancé y antes de la divergencia de los Neandertales (Enard et al. 2002). Estudios realizados sobre ratones transgénicos donde se introdujeron las sustituciones específicas de los humanos en el gen *FoxP2* endógeno de ratón mostraron que estas modificaciones produjeron alteraciones en las vocalizaciones de las crías y ocasionaron defectos neurológicos en los circuitos de los ganglios córtico-basales (Enard et al. 2009). Entonces, las sustituciones que tuvieron lugar en la secuencia codificante del gen *FOXP2* en el linaje humano podrían estar vinculadas con cambios en los circuitos de los ganglios córtico-basales, que participan del lenguaje (Enard et al. 2009).

Los aportes del campo de la evolución del desarrollo

El desarrollo es un proceso altamente restringido y estrictamente regulado, orquestado a través de complejas redes regulatorias de genes (Davidson y Erwin 2006). Existe una notable conservación del desarrollo entre organismos evolutivamente distantes como cnidarios, insectos y mamíferos, ya que las vías moleculares que controlan este programa a través de los distintos linajes animales son las mismas. Por ejemplo, los humanos, los gusanos planos y los cnidarios usan los mismos elementos básicos de las cascadas de señalización paracrina, como las vías *Wnt* y *TGFβ* (Finnerty et al. 2004; Carroll 2005). Los factores de transcripción que regulan el desarrollo también están altamente conservados, y tienden a regular los mismos procesos generales en diversas especies, incluyendo la regulación del patrón del eje corporal por los genes *Hox* (Lemons y McGinnis 2006), los órganos sensibles a la luz por *Pax6* (Gehring 2011), formación de la cabeza por homólogos de *Otx* (Yasuoka et al. 2014) y morfología del corazón por *Tinman / Nkx2-5* (Valentine, Jablonski, y Erwin 1999).

La amplia mayoría de los caracteres adquiridos en nuestro linaje son complejos y su origen evolutivo se encuentra relacionado a cambios morfológicos y fisiológicos intrincados que ocurren durante el desarrollo embrionario. Estos cambios morfológicos que ocurren durante el desarrollo son consecuencia de alteraciones en la expresión espaciotemporal de los genes involucrados en este proceso (Carroll 2008). Aquellos cambios de secuencia ocurridos en regiones genómicas que llevaron a la alteración de la maquinaria molecular del desarrollo probablemente facilitaron la evolución de rasgos morfológicos humano-específicos (Carroll 2003; Trinkaus y Shipman 2005). Es por eso por lo que enfoque de los estudios genético-evolutivos debe estar dirigido a entender los cambios a nivel regulatorio como las principales fuentes de variabilidad y diversidad morfológica (Carroll 2008; Rubinstein y de Souza 2013).

Las capacidades cognitivas, el bipedalismo y la producción de sonidos y palabras, son rasgos que tienen su origen en cambios morfológicos que se originan durante el desarrollo embrionario. Es por ello que aquellas preguntas enfocadas al origen evolutivo de dichos cambios deben ser abordadas desde la perspectiva de la biología evolutiva del desarrollo (Hall 2003). Desde este enfoque, es esperable que los mecanismos principales por los cuales se produce la evolución morfológica en los humanos estén relacionados fundamentalmente con alteraciones en la expresión temporal y espacial de genes del desarrollo (debido a cambios en sus regiones regulatorias en *cis*) y no a sustituciones aminoacídicas en los propios genes. Esta idea fue propuesta por King y Wilson en 1975, quienes postularon que las diferencias en secuencias regulatorias, más que en las secuencias proteicas, serían las responsables de las diferencias biológicas entre chimpancé y humanos (King y Wilson 1975), propuesta que se basa en los postulados originales de Jacob y Monod acerca de la importancia del desarrollo en la evolución (Jacob y Monod 1961).

Esta hipótesis se basa en una concepción esencialmente conservadora y modular del proceso de desarrollo que considera que los organismos están contruidos de unidades modulares (Gilbert, Dobyys, y Lahn 2005). La ventaja de las unidades modulares es que permiten que ciertas partes del cuerpo cambien sin interferir con las funciones de las otras partes (Gilbert, Dobyys, y Lahn 2005). Entre ellas, las regiones que forman parte de la regulación de los genes son modulares. Por ejemplo, es muy probable que cambios drásticos en la actividad de un factor de transcripción (que a su vez es responsable de la regulación de muchos genes) tenga efectos pleiotrópicos deletéreos. Por otro lado, modificaciones en la

actividad de un elemento regulatorio en cis discreto que controla la expresión de una única proteína en un tejido determinado es probable que resulte en cambios precisos del patrón de expresión espacial y/o temporal, produciendo cambios fenotípicos específicos sin alterar la mayor parte de las vías del desarrollo (Carroll 2005; Gilbert, Dobyns, y Lahn 2005). Entonces, si un gen particular pierde un elemento regulatorio, sufre una modificación en él o lo gana, el organismo que contiene este alelo particular expresará este gen en diferentes lugares, en diferentes tiempos o a diferentes tasas en relación con aquellos organismos que conservan el alelo original pudiendo derivar en el desarrollo de diferentes morfologías (Gilbert, Dobyns, y Lahn 2005).

El estudio de regiones regulatorias

El código regulatorio del genoma

Mientras que es relativamente sencillo inferir la función de una proteína a partir de su secuencia codificante en el ADN (gracias al relativamente rígido código genético), en general no es posible descifrar las propiedades funcionales de una región no codificante solamente analizando su secuencia. Esto se debe principalmente a que aún no se conoce en detalle el código regulatorio del genoma (Rubinstein y de Souza 2013). Este conocimiento implicaría comprender las relaciones complejas que tienen lugar entre las regiones regulatorias del ADN, factores de transcripción, ARNs regulatorios, vías de señalización y marcas epigenéticas, las que, en su conjunto, determinan la expresión de un gen (Rubinstein y de Souza 2013).

Este relativo desconocimiento del código regulatorio produjo un sesgo dentro de disciplinas tales como la genómica comparativa y la genética evolutiva hacia el análisis y detección de eventos en regiones codificantes (como pueden ser los eventos de duplicación génica y evolución de la secuencia de proteínas). Mientras que, los cambios en regiones no codificantes tendieron a ser ignorados (Carroll 2005). Sin embargo, secuenciado el genoma humano, hoy se sabe que aproximadamente 1,5% de nuestro genoma está compuesto por secuencias que codifican para proteínas y un 45% está compuesto por ADN repetitivo (principalmente secuencias remanentes de retrotransposones). Aproximadamente el 50% restante está compuesto por ADN no codificante antiguamente conocido como ADN basura

(Lander et al. 2001). En la actualidad, existe evidencia creciente que buena parte de estas secuencias no codificantes contienen los elementos regulatorios responsables de generar los diferentes programas de expresión génica (Heintzman y Ren 2009). La clave para comprender las bases moleculares que sustentan la evolución morfológica y fisiológica estará en dilucidar cómo actúan los mecanismos encargados de controlar la expresión génica (Britten y Davidson 1969; Carroll 2005; Wray 2007). Además, este conocimiento será muy importante para conocer cómo estos mecanismos participan en procesos patológicos y defectos del desarrollo, los cuales cada vez más están siendo asociados a mutaciones en regiones regulatorias (Rubinstein y de Souza 2013).

Los elementos regulatorios

Entre los elementos regulatorios mejor caracterizados se encuentran los promotores y las secuencias regulatorias que actúan a distancia como ser: las secuencias activadoras de la expresión o *enhancer*, los inhibidores de la expresión génica o *silencers* y las secuencias aisladoras o *insulators* (Wittkopp y Kalay 2011). La transcripción comienza en las regiones promotoras donde se reclutan las proteínas de la maquinaria basal de transcripción (Lee y Young 2000; Butler y Kadonaga 2002). Se postula que los promotores definen la orientación y el sitio exacto de inicio de la transcripción (Smale y Kadonaga 2003). Por otro lado, los *enhancers* actúan reclutando factores de transcripción determinados, coactivadores y complejos modificadores de la cromatina que en su conjunto, promueven la transcripción desde los promotores blanco (Glass y Rosenfeld 2000; Ong y Corces 2011). Por otro lado, los *silencers* unen proteínas y/o complejos de remodelación de la cromatina que inhiben la expresión de un gen (Glass y Rosenfeld 2000). Finalmente, las secuencias aisladoras proveen un nivel de regulación adicional, previniendo el avance de regiones de heterocromatina sobre regiones de cromatina activa, y/o restringiendo el alcance de un *enhancer* exclusivamente a su/s promotor/es blanco (West, Gaszner y Felsenfeld 2002) (Figura 5).

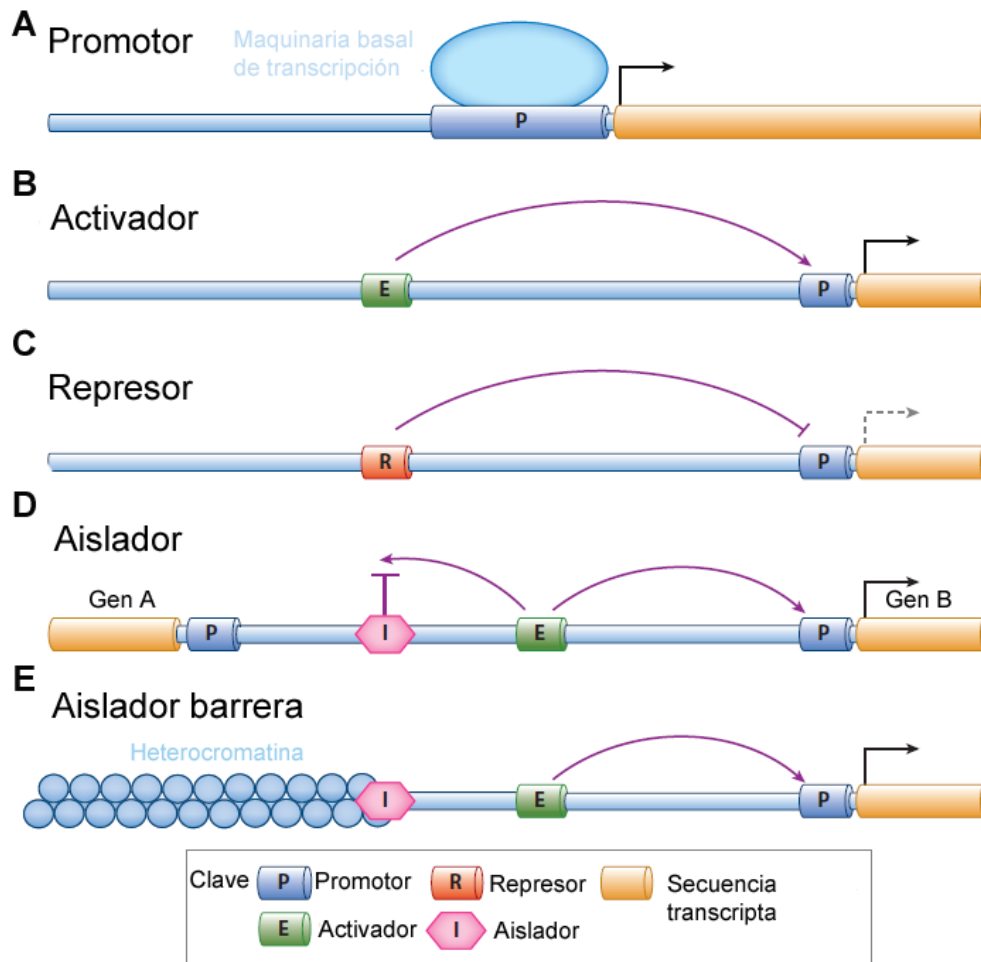


Figura 5. Tipos de elementos regulatorios conocidos. (A) Las secuencias promotoras son los sitios donde se unen los elementos de la maquinaria basal de la transcripción. Los activadores o *enhancers* (B) y los represores o *silencers* (C) median efectos positivos y negativos sobre la transcripción de sus genes blanco por medio de la interacción con los promotores. Los elementos aisladores que median el control del alcance de los *enhancers* (D) bloquean la activación de la transcripción promiscua por parte de un enhancer acotándola exclusivamente a su gen blanco. Existen regiones con función aisladora que funcionan como barrera al avance de las zonas de heterocromatina (E) permitiendo la normal expresión de los genes que controlan. Modificado de Noonan y McCallion 2010.

A las secuencias *enhancers* y *silencers* se las definió clásicamente como módulos discretos de ADN capaces de ejercer su actividad independientemente de su orientación respecto del promotor o su posición al sitio de inicio de transcripción (Blackwood y Kadonaga 1998). Los *enhancers* promueven la expresión génica reclutando factores de transcripción específicos, los cuales, a su vez, promueven la descondensación de la cromatina y/o facilitan el ensamblaje de la maquinaria de transcripción del complejo de la ARN polimerasa II (ARNPolII) en los promotores (M. Atchison 2002; Kadonaga 1998; Maston, Evans, y Green 2006). El genoma humano posee entre 1700 y 1900 genes que codifican para factores de

transcripción distintos (Vaquerizas et al. 2009). Estas proteínas usualmente contienen dos dominios distintos, un responsable de reconocer secuencias determinadas sobre el ADN (dominio de unión al ADN) y otro encargado de mediar la actividad regulatoria (generalmente conocido como dominio de transactivación). Uno de los mecanismos principales por los cuales los factores de transcripción actúan es reclutando cofactores encargados de traer proteínas de la maquinaria de remodelación de la cromatina (Kadonaga 1998; Lemon y Tjian 2000). Esta maquinaria es capaz de reposicionar nucleosomas específicos sobre los promotores de forma tal de exponer las secuencias necesarias para el ensamblaje del complejo de la ARNPolIII. Por ejemplo, uno de los complejos remodeladores más estudiados es el SWI/SNF. Este complejo es capaz de modificar de forma no covalente y dependiente de ATP la cromatina, deslizando los nucleosomas sobre el ADN (Kingston y Narlikar 1999). Por otro lado, otro tipo de coactivadores que pueden ser reclutados por los factores de transcripción para promover la expresión génica son las histonas acetil transferasa (HATs). Estos cofactores son capaces de acetilar el dominio N-terminal de las histonas H3 y H4 en residuos específicos. Esta modificación sobre las histonas produce que los nucleosomas se descondensen y colabora con el reclutamiento de factores necesarios para la transcripción (Strahl y Allis 2000). Entre los complejos de HATs más conocidos se encuentran: PCAF, CBP, p300, GCN5, TRRAP, entre otros (Rosenfeld, Lunyak, y Glass 2006). Adicionalmente, los mediadores son un tercer tipo de cofactores reclutados por los factores de transcripción. Estas proteínas facilitan la transcripción por medio de la interacción directa con la ARNPol II (Hatzis y Talianidis 2002) (Figura 6).

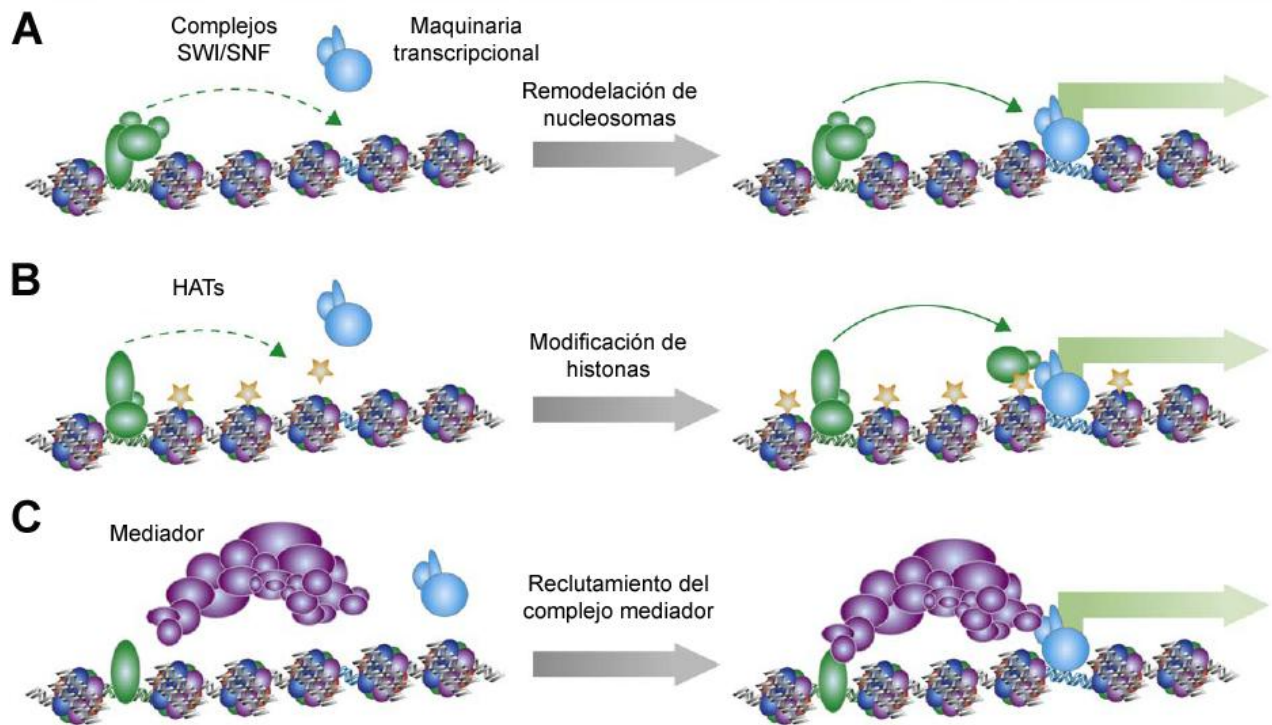


Figura 6. Tres mecanismos clásicos por los cuales pueden actuar los *enhancers*. (A) Los factores de transcripción reclutan maquinaria de remodelación de la cromatina conteniendo los complejos como SWI/SNF (óvalos verdes). Estos complejos son capaces de deslizar los nucleosomas sobre el ADN de los promotores (ADN azul) exponiendo secuencias necesarias para el ensamblaje de la maquinaria de transcripción (óvalos azules). (B) Los factores de transcripción pueden también reclutar complejos de coactivadores con actividad de HAT capaces de acetilar residuos específicos en las histonas (indicado como estrellas amarillas sobre la cromatina). Esta modificación en la cromatina puede facilitar la unión de más factores promotores de la transcripción (óvalos verdes adicionales) o descondensar la cromatina promoviendo la unión de la maquinaria transcripcional (óvalos azules). (C) Los factores de transcripción pueden reclutar complejos mediadores (óvalos violetas) los cuales, a su vez, pueden interactuar directamente con la ARNPolIII para activar la expresión. Modificado de (Heintzman y Ren 2009).

Predicción de *Enhancers*

El modelo actual de regulación génica propone que los *enhancers* contienen en sus secuencias la información necesaria para reclutar los factores de transcripción encargados de controlar temporal y espacialmente la actividad de sus genes blanco. Sin embargo, las secuencias consenso para unión de los factores de transcripción son, en general, altamente degeneradas y se las puede encontrar extensamente distribuidas por todo el genoma (Matys et al. 2003), por lo que la sola predicción de estos sitios es insuficiente para postular la existencia de un *enhancer* (Noonan y McCallion 2010; Wittkopp y Kalay 2011). No obstante, en la actualidad se cuenta con dos grandes aproximaciones metodológicas por las

cuales se puede predecir la existencia de una región regulatoria (Heintzman y Ren 2009). Por un lado, están los métodos que detectan características en la cromatina y en las proteínas unidas al ADN que, a su vez, están asociadas a la presencia de *enhancers*. Paralelamente, están los métodos que se basan en la conservación de las secuencias de ADN a nivel filogenético (Tagle et al. 1988; Bulger y Groudine 2011).

Siguiendo la primera aproximación, los *enhancers* típicamente se localizan en regiones donde hay una disrupción en la estructura de la cromatina. Esta disrupción puede ser puesta en evidencia como una mayor sensibilidad a la digestión con la enzima DNAsa I. Esta hipersensibilidad a la digestión con DNAsa I ocurre debido a que los nucleosomas son desplazados del ADN por la unión de proteínas específicas (como, por ejemplo, los factores de transcripción). De esta forma, el ADN está más expuesto a la digestión (Atchison 1988). Por otro lado, el ADN de regiones *enhancers* presenta modificaciones específicas en las histonas de sus nucleosomas. Dentro de este código se encuentran la monometilación de la histona H3 en su lisina 4 (H3K4Me1) y ausencia de trimetilación (H3K4Me3) (Bulger y Groudine 2011). Otro tipo de evidencia experimental que sugiere la existencia de un *enhancer* es la asociación de regiones del ADN con proteínas específicas (Heintzman y Ren 2009). Una de las asociaciones más estudiadas es el caso de la proteína p300, un coactivador estrechamente asociado a los *enhancers* activos (Heintzman et al. 2007). Por ejemplo, utilizando una estrategia que consistía en inmunoprecipitar ADN utilizando un anticuerpo contra la proteína p300 (marcador de *enhancers* activos), seguido de secuenciación masiva, se encontraron entre 500 y 2500 sitios de unión para p300 en tejido de cerebro de embriones de ratón de 11.5 días *post coitum* (dpc). En casi todos los casos, las regiones halladas resultaron funcionar como *enhancers* en ensayos con ratones transgénicos (Visel et al. 2009). Más recientemente, el Proyecto ENCODE identificó aproximadamente 400.000 elementos con características cromatínicas asociadas con la presencia de *enhancers* en diferentes tipos celulares humanos (ENCODE Project Consortium 2012). La ventaja de este tipo de aproximaciones es que, en teoría, permite simultáneamente identificar un *enhancer* junto con la información sobre el tejido y estadio del desarrollo donde el regulador está activo (Axel Visel et al. 2009). Sin embargo, varios autores consideran que las aproximaciones basadas en el mapeo a gran escala de características de la cromatina proveen indicios sobre la presencia de regiones con potencial capacidad regulatoria pero que deben hacerse necesariamente estudios funcionales

adicionales para comprobar esta función (De Souza, Franchini, y Rubinstein 2013; Graur et al. 2013).

Por otro lado, otro criterio utilizado para detectar posibles *enhancers* (y otras regiones funcionales en el genoma) es la conservación filogenética en regiones no codificantes (Tagle et al. 1988). La hipótesis en la que se basan estos estudios comparativos asume que aquellas regiones funcionales del genoma (tanto que codifican para proteínas como no codificantes) tendrán una tasa de evolución molecular más baja, comparada con la tasa neutra de evolución (que es la tasa a la cual, en teoría, evolucionan las regiones no funcionales). Esto se debe a que solo una pequeña proporción de los cambios que tienen lugar en las regiones con restricciones funcionales pasará a la próxima generación (cambios que resulten selectivamente neutros y/o ventajosos) (M Kimura 1968; Tagle et al. 1988). De acuerdo con la teoría neutralista o del equilibrio mutación-deriva, estos nuevos alelos neutros o ventajosos se fijarán en las poblaciones ya sea por deriva genética o por selección natural, acumulándose, y formarán parte de la divergencia que se analizará en los estudios comparativos. Sin embargo, la gran parte de las sustituciones que ocurren en las regiones funcionales del genoma resultan deletéreas, siendo el destino más probable de estos alelos la eliminación por selección purificadora. Estos cambios no quedan representados en las poblaciones, no aportando a la divergencia. Por otro lado, las mutaciones que ocurren en sitios sin restricciones funcionales tienen por lo general efectos despreciables sobre la aptitud (es decir, son selectivamente neutros) y están libres para fijarse por deriva génica, de allí la mayor tasa de evolución molecular en estas regiones (Kimura 1968; Kimura 1983).

En consecuencia, las regiones genómicas sin restricciones funcionales entre dos especies acumulan sustituciones a tasa neutras y la divergencia aumenta a medida que la distancia evolutiva entre las especies se incrementa. Mientras que, las secuencias ortólogas sujetas a restricciones funcionales en ambas especies también cambian, pero a una tasa menor. Entonces, la conservación filética es un parámetro extremadamente útil para predecir la presencia de un *enhancer* u otro tipo de región funcional en secuencias no codificantes (Tagle et al. 1988). Sin embargo, este método no permite predecir dónde y cuándo ese posible regulador estará activo. También, es importante notar que la conservación entre secuencias ortólogas es un indicador de conservación de la función, pero no lo garantiza (Noonan y McCallion 2010). Por ejemplo, secuencias conservadas entre dos especies pueden tener en cada una función distinta y a la inversa, funciones conservadas entre dos

especies pueden estar codificadas en secuencias con muy poca conservación (Noonan y McCallion 2010). Por ejemplo, se encontró que secuencias regulatorias presentes en el locus RET humano y que no presentan niveles de conservación manifiesta con elementos presentes en el genoma del pez cebra (*Danio rerio*) son capaces de dirigir la expresión a dominios específicos del gen RET endógeno cuando se evalúan en peces cebra transgénicos. Una de las explicaciones para este hallazgo es que existen secuencias ortólogas en pez cebra a las humanas ensayadas pero estas secuencias divergieron más allá del nivel de detección ya sea por pequeños cambios en los sitios de unión para factores de transcripción, rearrreglos de estos sitios u otros cambios evolutivos (Fisher, Grice, Vinton, Bessling, y McCallion 2006). En un ejemplo contrario al anterior, se reportó que los elementos regulatorios ar-D, ar-A y ar-C del gen Sonic hedgehog (Shh) en pez cebra dirigen la expresión a dominios distintos del animal cuando se los compara con sus elementos ortólogos en ratón. Esta diferencia en la actividad de estos *enhancers* es llamativa si se toma en cuenta que las secuencias de estos elementos están altamente conservadas entre estas especies (con aproximadamente un 60% de identidad de secuencia) (Ertzer et al. 2007).

La identificación de regiones funcionales en el genoma por conservación de secuencia se aceleró sustancialmente gracias a que ahora se cuenta con múltiples genomas secuenciados para comparar y métodos sofisticados para evaluar el nivel de conservación de las secuencias (Noonan y McCallion 2010). Los primeros estudios realizados comparando el genoma humano con el de otros vertebrados encontraron miles de secuencias altamente conservadas no codificantes. Estas fueron llamadas *conserved noncoding sequences* (CNSs por sus siglas en inglés) (Siepel et al. 2005; Venkatesh et al. 2006). Estos estudios hallaron que los CNSs no están distribuidos al azar, agrupándose preferencialmente cerca de genes del desarrollo y en desiertos génicos. La gran representación de CNSs asociados a genes del desarrollo puede indicar que, en oposición a genes controlados menos dinámicamente, estos genes requieren un mayor número de regiones regulatorias para poder modular de forma precisa la expresión durante el desarrollo. Además, se encontró que las agrupaciones de CNSs correlacionan con regiones donde la sintenia está conservada entre los vertebrados. Esto podría deberse a que estos elementos regulatorios deben permanecer asociados con sus genes blanco, siendo estas regiones genómicas más resistentes a los rearrreglos (Boffelli, Nobrega, y Rubin 2004; Ahituv et al. 2005; Ovcharenko et al. 2005).

Más recientemente, un estudio mapeó en gran detalle las secuencias de ADN bajo selección purificadora en los mamíferos, utilizando para ello las secuencias completas de 29 genomas, incluido *Homo sapiens*. Este análisis mostró que 4,2% de las secuencias presentes en los genomas de estos mamíferos están conservadas filogenéticamente. Dentro de esta fracción, 68% de estas secuencias son intrónicas e intergénicas y, al menos, un 30% de las secuencias conservadas totales se encuentra en loci con marcas de cromatina típicas de *enhancers* (Lindblad-Toh et al. 2011). Esto indicaría que, aproximadamente dos tercios de todas las secuencias bajo selección purificadora (al menos dentro de los mamíferos) del genoma humano son no codificantes. Entonces, cualquier trabajo que quiera encontrar las bases genéticas de la diversidad morfológica debería considerar a las regiones no codificantes como parte importante de su campo de estudio (Carroll 2005).

Genes blanco de los *enhancers*

Otro de los grandes desafíos cuando se trabaja con *enhancers*, una vez identificados, es determinar cuáles son sus genes blancos. Se puede visualizar lo complejo de este objetivo si se tiene en cuenta que los *enhancers* se pueden encontrar en posición proximal o distal, actuando incluso a millones de pares de bases de distancia del promotor blanco (Lettice et al. 2003). Además, se pueden hallar dentro de intrones u exones de genes vecinos e incluso puede activar genes blancos que se localizan en otros cromosomas (Lomvardas et al. 2006; van Heyningen y Bickmore 2013). No se conoce aún en detalle el mecanismo preciso por el cual estos reguladores pueden actuar sobre tales distancias. Uno de los modelos más populares propone que los elementos regulatorios distales son capaces de reclutar proteínas que interactúan a su vez con la maquinaria de transcripción vía formación de un lazo de cromatina. De esta forma, regulador y promotor quedan en estrecho contacto (Krivega y Dean 2012).

Sin embargo, como primera aproximación para identificar los genes blancos se suele asignar un *enhancer* al gen que tenga físicamente más próximo. No obstante, esta hipótesis se vuelve más sólida si el patrón de expresión del gen candidato y el *enhancer* se superponen (al menos en parte del dominio espacial y temporal de expresión del gen) (Heintzman y Ren 2009). La demostración definitiva requiere de ensayos delecionales o de mutagénesis (ya sea en mutantes naturales o recurriendo a modelos animales) donde se pueda evaluar el

impacto de cambios en las regiones regulatorias sobre el patrón de expresión de su gen blanco.

Uno de los ejemplos más llamativos de *enhancer* distales es el regulador del factor de transcripción Sonic hedgehog (SHH) que controla la expresión de este gen en los miembros posteriores en desarrollo de mamíferos. Este *enhancer* se localiza a más de 1 Mb del sitio de inicio de la transcripción de SHH, en medio de un intrón del gen *Limb region 1* (LMBR1), gen que se encuentra río arriba de SHH (Lettice et al. 2003). La disrupción de este *enhancer* en ratón ocasiona la degeneración de los miembros posteriores por falta de un correcto establecimiento del eje anteroposterior del miembro en desarrollo (Sagai et al. 2005).

El ejemplo anterior pone de manifiesto la importancia de los *enhancer* distales durante el desarrollo y lo complejo de asociar a priori *enhancers* y genes blancos. Incluso, ya se empieza a vislumbrar un nivel de complejidad adicional a la hora de estudiar *enhancers*. Tradicionalmente, se consideró que cada *enhancer* controlaba un único dominio de expresión de su gen blanco. Sin embargo, existe evidencia creciente sobre la gran redundancia de elementos regulatorios que habría en los metazoos, principalmente aquellos asociados a genes del desarrollo (Barolo 2013; Frankel 2012). Por ejemplo, un estudio llevado a cabo en moscas *Drosophila melanogaster* transgénicas, demostró que dos genes blanco del factor de transcripción Dorsal, los genes *brinker* y *sog*, tienen cada uno, al menos, dos *enhancers* capaces de reclutar a Dorsal. Estos dos pares de *enhancers* están activos durante los mismos estadios del desarrollo larval del animal y dirigen la expresión a dominios aparentemente idénticos (Hong, Hendrix, y Levine 2008; Perry et al. 2010). Se postula que esta redundancia de función podría estar asociada a una mayor estabilidad fenotípica (permitiendo resistir a cambios deletéreos en las regiones regulatorias) (Hong, Hendrix, y Levine 2008). Por otro lado, también se sugiere que los *enhancers* con actividad regulatoria parcialmente redundante podrían ser fuente de novedades evolutivas. En este tipo de sistemas, las mutaciones que tuvieran lugar en uno de estos elementos podrían ser mejor tolerados, pudiendo, al mismo tiempo, ser fuente de nuevos patrones de expresión (Hong, Hendrix, y Levine 2008). No obstante, a nivel experimental complejizan aún más los ensayos funcionales ya que, la ausencia de un fenotipo por delección o mutación no necesariamente implica la ausencia de función de ese elemento in vivo.

Ensayos funcionales para regiones regulatorias

Como se detalló anteriormente, en la actualidad se cuenta con métodos refinados para predecir regiones regulatorias. Sin embargo, siempre se deben realizar ensayos *in vivo* para confirmar la presencia de un elemento regulatorio (De Souza, Franchini, y Rubinstein 2013). La mayoría de los trabajos sobre regiones funcionales no codificantes se concentran en el estudio de regiones *enhancer*. Esto es así, principalmente, porque no existen aproximaciones experimentales robustas y aplicables a gran escala para probar regiones, por ejemplo, con función *silencer* y/o *insulator* en vertebrados (Noonan 2009).

Todos los ensayos funcionales para regiones regulatorias requieren de la construcción de transgenes y la posterior introducción de estos en los organismos o cultivos celulares adecuados (Arnone, Dmochowski, y Gache 2004). Idealmente, en los casos que se quiera evaluar cambios regulatorios a nivel filético, los dos elementos candidatos deben ser probados individualmente en las dos especies de las cuales estas regiones provienen. Estos experimentos se pueden realizar construyendo pares recíprocos de organismos modificados genéticamente (Wittkopp y Kalay 2011).

Para estudiar funcionalmente regiones *enhancer*, se requiere, además, de un promotor para mediar su actividad y un gen blanco sobre los cuales actuar. Estos elementos son fundamentales y deben ser tenidos en cuenta a la hora de construir los transgenes. Los transgenes se pueden diseñar de forma tal que contengan al locus genómico completo. En estos casos, el transgén contiene el gen blanco (conteniendo o no un gen reportero o etiqueta en su interior), su propio promotor y el *enhancer* en su ubicación nativa (Wittkopp y Kalay 2011). También se pueden construir transgenes quiméricos en donde el *enhancer* candidato suele ubicarse río arriba de un gen reportero (generalmente en posición proximal). En este segundo caso, se puede utilizar el promotor mínimo del gen blanco (si este se conoce). Por el contrario, si el gen blanco de un *enhancer* no se conoce o no se cuenta con la secuencia del promotor mínimo, se recurren a promotores mínimos que derivan de otros genes e incluso de especies distintas que las que proviene el *enhancer*. Los promotores mínimos son elementos incapaces de activar la expresión de los genes reporteros en ausencia de un *enhancer* funcional y solo brindan las secuencias necesarias para el ensamblaje del complejo de la ARNPolIII (Wittkopp y Kalay 2011). Es importante remarcar que muchos *enhancer* conservan su patrón de expresión aun cuando son

removidos de su contexto genómico normal. Esta autonomía implica que los *enhancers* pueden ensamblar un conjunto completo de actividades bioquímicas que son lo suficientemente robustas como para activar la expresión génica de forma definida (Swanson, Evans, y Barolo 2010).

Una vez construido el transgén, éste debe ser introducido en el modelo experimental adecuado. Existen diferentes métodos para introducir los transgenes dependiendo del modelo con el que se trabaje. Por ejemplo, transfección de células en cultivo (por ejemplo, líneas celulares humanas), transfecciones transitorias en animales transgénicos (como en peces), introducción de material genético extracromosómico (en nematodos), integración al azar en el genoma (como en plantas, peces cebra y ratones) e integración dirigida a loci endógenos determinados (como en moscas). El último nivel de resolución es el reemplazo de loci endógenos y esto puede ser realizado en levaduras o en ratones (Wittkopp y Kalay 2011; Ahituv 2012).

Una de las metodologías más novedosas que se dispone para evaluar la actividad *enhancer* in vivo consiste en un ensayo basado en el uso del transposón Tol2 de Medaka (*Orzias latipes*) para mediar de forma eficiente la transgénesis en pez cebra (*Danio rerio*, K Kawakami y Shima 1999). Esta metodología hace uso de la alta eficiencia de integración de este transposón para movilizar los transgenes al genoma (Kawakami, Shima, y Kawakami 2000). Para evaluar la actividad regulatoria, los elementos candidatos se clonan en un vector universal que contiene las secuencias necesarias para el reconocimiento de la transposasa y un gen reportero (generalmente EGFP) bajo la regulación de un promotor mínimo (como por ejemplo el promotor mínimo del gen cFos de ratón). Este vector se co-inyecta con el ARNm que contienen la secuencia codificante para la transposasa en embriones de pez cebra en el estadio de 1 a 2 células (Kawakami et al. 2004; Fisher, Grice, Vinton, Bessling, Urasaki, et al. 2006). Dentro de las ventajas que cuenta este sistema está la alta tasa de transgénesis alcanzada y la relativamente baja laboriosidad del método (comparado con la producción de ratones transgénicos). Además, permite evaluar la actividad *enhancer* en un modelo de vertebrado y a lo largo del desarrollo embrionario (Fisher, Grice, Vinton, Bessling, Urasaki, et al. 2006). Otra de las ventajas es que los embriones de pez cebra pueden ser tratados con inhibidores químicos para obtener peces transparentes, facilitándose así el análisis de la expresión del gen reportero en embriones intactos (Karlsson, von Hofsten, y

Olsson 2001). Con este método fue posible evaluar la actividad regulatoria de elementos humanos con éxito (Fisher, Grice, Vinton, Bessling, y McCallion 2006).

Cuando se estudian cambios regulatorios a nivel evolutivo que involucran secuencias humanas, nunca es posible hacer los experimentos recíprocos mencionados anteriormente. La única alternativa para evaluar los elementos en un contexto celular nativo consiste en realizar ensayos en líneas celulares humanas (Wittkopp y Kalay 2011). Sin embargo, esta metodología presenta serias limitaciones. Por un lado, debe conocerse el patrón de expresión del *enhancer* para utilizar una línea celular donde ese elemento esté activo. Por el otro, los trabajos sobre cultivos celulares no pueden dar información sobre el patrón temporal de expresión de los elementos analizados en un organismo completo (Ahituv 2012) como es necesario, por ejemplo, cuando se estudian elementos activos durante el desarrollo embrionario. Es por ello que la mejor alternativa en estos casos es realizar estudios donde las diferentes secuencias ortólogas sean evaluadas en una tercera especie modelo (Vallender, Mekel-Bobrov, y Lahn 2008). Por ejemplo, existen numerosos trabajos que analizan la expresión de secuencias regulatorias humanas recurriendo a la utilización en ratones (Pennacchio et al. 2006; Visel et al. 2007) o peces transgénicos (Fisher, Grice, Vinton, Bessling, y McCallion 2006; Visel et al. 2007). En estos casos, se trabaja sobre animales modelos de mamíferos y vertebrados respectivamente, pudiéndose registrar la actividad de un *enhancer* a lo largo del desarrollo embrionario del animal.

Sin embargo, también estos ensayos tienen sus limitaciones. Los estudios realizados con ratones transgénicos son muy costosos y laboriosos. Es por ello por lo que en general, se analiza uno o unos pocos estadios del desarrollo (típicamente a 11.5 dpc). Esto limita la posibilidad de encontrar regiones activadoras de la expresión génica ya que aquellas regiones activas en etapas anteriores o posteriores a la/s ensayada/s serán consideradas inactivas (Noonan 2009). Por otro lado, los estudios en animales transgénicos requieren que las secuencias humanas sean funcionales en contextos celulares heterólogos, como los brindados por ratones o peces. Estos modelos podrían no brindar el entorno necesario para que se recapitule la actividad endógena de un regulador (Noonan 2009). Por ejemplo, diferencias en la afinidad de un factor de transcripción por una secuencia blanco o cambios en el patrón temporal o espacial de un factor de transcripción podrían resultar en un patrón de expresión de un *enhancer* humano en ratón o en pez cebra que no reproduce el patrón endógeno humano (Ritter et al. 2010; Ariza-Cosano et al. 2012).

Origen y evolución de los *enhancers*

Existen diferentes mecanismos que conllevan a la creación de nuevos *enhancers*, entre los cuales se encuentran la inserción de elementos transponibles, mutaciones de novo que generan agrupamientos de sitios de unión para factores de transcripción, y rearrreglos cromosómicos que promueven la adopción de nuevos *enhancers* por parte de los genes (Rubinstein y de Souza 2013). La gran contribución de los elementos transponibles a la composición del genoma (30-50% en mamíferos) proporciona una substancial cantidad de material con el potencial de convertirse en elementos regulatorios funcionales novedosos. Sin embargo, existen pocos casos en los que se ha podido demostrar funcionalmente que la actividad regulatoria fue una innovación mediada por la exaptación de un elemento transponible (De Souza, Franchini, y Rubinstein 2013).

Recientemente, un grupo de investigación definió una lista de regiones regulatorias, definidas por la ocurrencia de la acetilación de la histona H3K27 (Creyghton et al. 2010), activas durante la corticogénesis humana y de ratón para identificar a los *enhancers* que eran probablemente activos en el cerebro anterior ancestral de mamíferos (Emera et al. 2016). De esta manera, dedujeron los orígenes filogenéticos de estos *enhancers*, encontrando que aproximadamente el 20% de ellos surgieron en el linaje ancestral de mamíferos, coincidente con la aparición de la neocorteza cerebral, y todos ellos se encuentran cerca de genes implicados en la migración y señalización celular.

Además, también proporcionaron información acerca de los mecanismos de innovación de la regulación en mamíferos, ya que encontraron que la mayoría de los *enhancers* neocorticales no se originaron mediante exaptación en bloque de transposones. En cambio, los autores presentan un modelo de ciclo de vida, en donde los *enhancers* surgen inicialmente como regiones cortas, débilmente restringidas llamadas 'proto-*enhancers*'. Muchos de estos proto-*enhancers* fueron probablemente perdidos, pero los autores creen que algunos de ellos pudieron servir como puntos de nucleación para dar lugar a *enhancers* nuevos.

Por otro lado, los *enhancers* también pueden evolucionar mediante distintos mecanismos. Una manera drástica para ello es simplemente la delección del *enhancer*, que conlleva a la pérdida de la expresión génica. Otra forma más sutil es a través de mutaciones puntuales y pequeños indels que, o bien crean o eliminan sitios de unión para factores de transcripción

dentro de un *enhancer*, cambiando así su actividad. Por lo tanto, la pérdida de actividad de un *enhancer* puede deberse no sólo a la eliminación de este, sino también a la acumulación de mutaciones puntuales con efectos pequeños que en conjunto causan la pérdida completa del *enhancer* (Rubinstein y de Souza 2013).

En un estudio reciente, un grupo de investigación realizó un seguimiento de la evolución de los *enhancers* y de los promotores activos en hígado a través de 20 especies de mamíferos (Villar et al. 2015). Para ello, utilizaron la ocurrencia de dos histonas claves, altamente utilizadas para identificar promotores y *enhancers*: H3K4me3 (Meissner et al. 2008) y H3K27ac (Creyghton et al. 2010). De esta manera, dedujeron que la mayoría los promotores activos en hígado se encuentran parcial o totalmente conservados a través de las 20 especies. En contraste, los *enhancers* poseen una evolución más rápida que los promotores y observaron que esta es una característica universal del genoma de los mamíferos. Además, descubrieron que los *enhancers* que evolucionaron recientemente surgieron de exaptaciones de ADN ancestral, en vez de expansiones de elementos repetitivos linaje específico.

La evolución regulatoria de nuestro genoma

Tuvieron que pasar muchos años para que se pudiera comenzar a poner a prueba la hipótesis de King y Wilson y emprender el estudio de cambios específicos de los humanos en la regulación de la expresión de los genes (King y Wilson 1975). La hipótesis principal en estos trabajos sostiene que una parte significativa de características exclusivas de nuestra especie debieron haber surgido por medio de modificaciones a nivel regulatorio de la arquitectura génica del desarrollo; es decir, como variaciones específicas de los humanos del programa general, altamente conservado, del desarrollo de los mamíferos (Britten y Davidson 1969; King y Wilson 1975; Carroll 2005; Noonan 2009).

No obstante, identificar regiones regulatorias activas durante el desarrollo embrionario que hayan cambiado su actividad exclusivamente en el linaje humano no es una tarea sencilla. Como los programas genéticos del desarrollo tienden a estar altamente conservados (Gilbert, Dobyns, y Lahn 2005), es esperable que gran parte de las secuencias encargadas de la regulación de estos genes también estén conservadas entre nosotros y el resto de los primates u otros mamíferos. Entonces, para llevar a cabo esta identificación se requieren

métodos específicos con los cuales detectar, a nivel de las secuencias de ADN no codificante, las evidencias dejadas por los procesos evolutivos (Pollard et al. 2006; Prabhakar et al. 2006; C. P. Bird et al. 2007; Lindblad-Toh et al. 2011).

Sin embargo, existen grandes limitaciones a la hora de aplicar al estudio de las regiones no codificantes los métodos más usados para la detección a nivel molecular de procesos de selección positiva. Por ejemplo, en las regiones genómicas no codificantes no es sencillo predecir si una sustitución nucleotídica determinada resulta sinónima o no sinónima (uno de los principios utilizados para la detección de evolución adaptativa en regiones que codifican para proteínas) (Kostka et al. 2012). Esto se debe a que no se conoce en detalle cómo se codifica la información en los elementos regulatorios del genoma. Sumado a lo anterior, es muy probable que los aspectos más importantes de la evolución de los rasgos específicos de nuestro linaje a nivel genético (incluidos varias instancias de expansión de nuestro cerebro) hayan ocurrido en tiempos profundos de la historia evolutiva de nuestra especie (más allá de los 2 millones de años) (Striedter 2005; Vallender, Mekel-Bobrov, y Lahn 2008). Para estos tiempos, la identificación a nivel genético de procesos de evolución adaptativa por detección de polimorfismos actuales no es aplicable (Tattersall 2000). En su conjunto, estas características impiden o limitan enormemente el uso de los algoritmos más frecuentemente usados para la detección de procesos de evolución adaptativa en los genomas (sean estos los basados en la comparación de las tasas de sustitución sinónima y no sinónima en el primer caso o de estudio de frecuencias de polimorfismos actuales, en el segundo).

Dadas las limitaciones anteriormente mencionadas, una forma de predecir los sitios más probables donde pudieron haber ocurrido cambios a nivel de la regulación de los genes es recurrir a la detección de modificaciones significativas en la tasa de evolución molecular en regiones no codificantes, muy conservadas evolutivamente y en un linaje determinado (Pollard et al. 2006; Prabhakar et al. 2006; Kostka et al. 2012). Aplicando esta metodología se pueden poner en evidencia eventos de evolución acelerada; entendiendo por esta al patrón observado en determinadas regiones genómicas para un linaje específico donde ocurrieron más sustituciones nucleotídicas que las que se hubieran esperado a partir de la tasa de evolución de fondo de esa región (Noonan 2009; O'Bleness et al. 2012). Detectar evolución acelerada sobre regiones conservadas del genoma tiene varias ventajas. Por un lado, la conservación es evidencia de restricción funcional y en último lugar, de importancia

biológica. Por el otro lado, el poder de detección de cambios en la tasa de evolución es mayor cuando se compara con una tasa de evolución de fondo baja que cuando se compara con una tasa cercana a la neutra y además, se reducen los costos computacionales ya que sólo se analiza una porción acotada (típicamente del 5-10%) del genoma (Kostka et al. 2012). Sin embargo, esta metodología tiene sus desventajas. Como en estos casos lo que se estudia son cambios en la tasa de evolución molecular es muy difícil identificar los procesos evolutivos responsables de la aceleración. Por ejemplo, en las regiones con evidencia de evolución aceleradas la tasa de sustitución puede ser significativamente más alta que la esperada dada la extrema conservación de esa secuencia en otras especies, pero puede llegar a no exceder la tasa neutral de evolución y entonces podría reflejar un proceso de relajación de la selección purificadora más que un proceso de evolución adaptativa (Kostka et al. 2012). Entonces, se requieren test adicionales para distinguir los procesos evolutivos subyacentes, como, por ejemplo, comparar la tasa de evolución molecular en la región acelerada con la tasa neutra de evolución de esa región, entre otros (Pollard et al. 2006; Lindblad-Toh et al. 2011; Kostka et al. 2012).

Por otro lado, la selección positiva y la relajación de la selección negativa son algunas de las causas que pueden producir un patrón de evolución acelerado, pero no las únicas. Por ejemplo, los cambios en el tamaño o estructuración de las poblaciones pueden llevar a una fijación rápida de los alelos segregantes y producir modificaciones en la tasa de evolución molecular (Kreitman 2000). Adicionalmente, la aceleración de la tasa de sustitución también puede ser el producto de procesos que no involucran necesariamente cambios en el modo de selección o eventos poblacionales (Pollard et al. 2006). Por ejemplo, el proceso evolutivo selectivamente neutro conocido como *GC-biased gene conversion* (gBGC por sus siglas en inglés) es un mecanismo mutagénico causado por la preferencia en las enzimas que reparan el ADN por generar apareamientos entre bases fuertes (GC) frente a los de bases débiles (AT) en sitios donde se hayan generado heteroduplex de ADN durante la recombinación meiótica (Strathern, Shafer, y McGill 1995). Por lo tanto, el proceso de gBGC puede producir un incremento en la probabilidad de fijación de alelos C y G frente a alelos A y T, generando en las zonas afectadas un rápido incremento en tasa de sustitución, comparable en magnitud al producido por un proceso de selección positiva. Sin embargo, hay una diferencia: se observa un gran sesgo en las sustituciones de bases débiles a bases fuertes. Una vez que el proceso de gBGC alcanza un equilibrio decrece la tasa de sustitución (Eyre-

Walker y Hurst 2001; Kostka et al. 2012). Es importante destacar que, si bien el proceso de gBGC se considera selectivamente neutro, existen trabajos que muestran que el incremento del contenido GC de secuencias transcritas puede incrementar su nivel de expresión, lo que podría traer asociado algún valor adaptativo (Kudla et al. 2006). Sin embargo, el rol del proceso de gBGC en la evolución está aún en discusión.

Detección de evolución acelerada en el genoma humano

Hasta el presente cinco grupos han llevado a cabo estudios genómicos a gran escala con el objetivo de detectar regiones genómicas filogenéticamente conservadas que hayan cambiado su tasa de evolución molecular exclusivamente en el linaje humano (Pollard et al. 2006; Prabhakar et al. 2006; C. P. Bird et al. 2007; Bush y Lahn 2008; Lindblad-Toh et al. 2011). Cada uno de estos grupos, de forma independiente, aplicó su propia metodología produciendo como resultado una base de datos de elementos acelerados. El esquema general para la construcción de estas bases de datos se puede resumir en tres pasos:

1- Identificación de regiones genómicas conservadas. Estas regiones pudieron abarcar regiones codificantes y no codificantes (Pollard et al. 2006) o no codificantes exclusivamente (Prabhakar et al. 2006; Bush y Lahn 2008; Lindblad-Toh et al. 2011). La conservación filogenética se estudió eligiendo un grupo de especies de mamíferos y/o vertebrados (pudiendo o no incluir a *Homo sapiens* en este paso), alineando sus genomas y aplicando algún método estadístico para cuantificar el nivel de conservación. Finalmente, de esos alineamientos, cada grupo eligió recuperar aquellas regiones genómicas que superaran algún valor de conservación determinado. De esta forma, se distinguieron regiones con una señal de selección purificadora fuerte y buenos candidatos a ser regiones funcionales donde las sustituciones podrían tener efectos biológicos.

2- Aplicación de un método estadístico para detectar regiones del genoma humano que hayan cambiado su tasa de sustitución nucleotídica. En esta segunda etapa se analizaron exclusivamente los alineamientos que en la etapa anterior resultaron satisfacer el nivel de conservación establecido. Los métodos y tests estadísticos utilizados para la identificación de las regiones con evidencia de evolución acelerada en el genoma humano fueron diversos.

En algunos trabajos se calculó cuan probable era obtener los alineamientos observados para cada región dado un modelo de evolución por selección negativa (o de conservación, donde la tasa de sustitución era la misma, y baja, para todos los linajes) o de evolución acelerada exclusivamente en el linaje humano. Luego, se compararon los ajustes a ambos modelos y se determinó estadísticamente cuanto más probable eran los alineamientos observados dado uno u otro modelo (Pollard et al. 2006; Lindblad-Toh et al. 2011). En otro trabajo, se estimó estadísticamente cuan sorprendente era observar la secuencia humana (posiciones conservadas y no conservadas) dado la tasa de evolución neutra (o de fondo) de esa región y el nivel de conservación filogenético de cada posición en el alineamiento (Prabhakar et al. 2006). Finalmente, un quinto trabajo estimó que cuatro sustituciones específicas de humano en las regiones conservadas de más de 100 pares de bases era estadísticamente significativo como señal de evolución acelerada (C. P. Bird et al. 2007). En resumen, el resultado general de este segundo paso fue la identificación de regiones genómicas con un cambio significativo en la tasa de evolución molecular exclusivamente en el linaje humano.

3- Curado de las regiones identificadas. Una vez identificadas las regiones genómicas con una tasa de evolución molecular significativamente acelerada en el linaje humano, se curaron (manualmente o recurriendo a filtros específicos) para eliminar aquellas regiones que presentaran errores en los alineamientos y/o que estuvieran presentes en regiones anotadas como pseudogenes (entre algunos de los filtros aplicados en estos trabajos). De esta forma, se obtuvieron listas de elementos acelerados altamente curados y ordenados por su valor de verosimilitud (o chi-cuadrado).

Es importante notar que estas metodologías excluyen del análisis a las regiones potencialmente regulatorias que hayan surgido *de novo* en el linaje humano o que no puedan ser detectadas por conservación, ya que una de las condiciones exigidas para todos los elementos es la conservación filogenética. Además, también quedan excluidas regiones no codificantes conservadas donde haya ocurrido relativamente poco cambio (no lo suficiente como para generar un cambio significativo en la tasa de evolución molecular de esa región), sin que esto implique que esos cambios no fueran importantes para la evolución morfológica de nuestra especie. Es por ello por lo que estos trabajos son, en

teoría, capaces de recuperar una fracción del universo de regiones regulatorias con importancia potencial en la evolución regulatoria de nuestro genoma.

Caracterización funcional de los HAEs

Uno de los desafíos cuando se trabaja con HAEs es la caracterización funcional de los mismos. Este objetivo es particularmente complejo dado que la gran mayoría de estos elementos se encuentran en regiones del genoma no caracterizadas funcionalmente y que la conservación por si misma es incapaz de predecir una función (Noonan y McCallion 2010). Sin embargo, varios grupos de investigación han tratado de estudiar estos elementos a nivel funcional.

Dado que los HAEs son secuencias altamente conservadas entre los mamíferos y que mayoritariamente se ubican en regiones que no codifican para proteínas, una de las hipótesis es que actuarían como elementos regulatorios (Capra et al. 2013). Por ejemplo, estudios funcionales realizados sobre el HACNS1 (también denominado HAR2), el HAE con la mayor tasa de evolución molecular del genoma humano (Prabhakar et al. 2006), encontraron que este elemento es capaz de actuar como un *enhancer* activo durante el desarrollo embrionario. Adicionalmente, encontraron que el HACNS1 tiene un dominio de expresión más amplio comparado con los elementos ortólogos de chimpancé y mono Rhesus (ganando expresión en los miembros en desarrollo) (Prabhakar et al. 2008). Sin embargo, aún se desconoce el gen blanco del HACNS1 y/o el impacto funcional del cambio en el patrón de expresión dirigido por este *enhancer*.

Por otro lado, también se realizaron estudios funcionales sobre el HAR1. Este elemento resultó ser un gen de ARN que se expresa durante el desarrollo, principalmente en las neuronas de Cajal-Retzius en la neocorteza. Además, se encontró que el HAR1 (HAR1F) se co-expresa con la proteína reelina, una glicoproteína de matriz extracelular secretada por las neuronas de Cajal-Retzius y que es fundamental para la especificación de las capas de la corteza (Pollard et al. 2006).

En la actualidad existe evidencia adicional que indica que muchas de las regiones no codificantes con una tasa de sustitución alta en el linaje humano exclusivamente tienen la capacidad de actuar como *enhancers* durante el desarrollo embrionario. Por ejemplo, un estudio reciente mostró que 24 HAEs (de entre 29 analizados) resultaron ser capaces de

activar la expresión de un gen reportero en ensayos con ratones transgénicos. Además, 5 de estos *enhancers* mostraron una actividad diferencial entre los elementos humano y chimpancé, sugiriendo que podrían haber participado en la adquisición de nuevos dominios de expresión en alguno (u ambos) de estos linajes (Capra et al. 2013).

Estos resultados sugieren que los HAEs serían una rica colección de secuencias en las cuales buscar las bases genéticas de cambios específicos de los humanos en la regulación de los genes (Capra et al. 2013). Adicionalmente, la alta densidad de HAEs en el locus de NPAS3 sugiere que este gen pudo haber adquirido cambios regulatorios en el linaje de los humanos. Los caracteres morfológicos que pudieron verse afectados por la alta tasa de evolución experimentada por los HAEs presentes en esta región dependerá fuertemente de la función desarrollada por este gen.

Del genoma al estudio de un gen específico

Como adelanta el título, esta tesis se concentra en el estudio del gen *NPAS3*. *NPAS3* es un gen que codifica para un factor de transcripción de la familia *basic helix-loop-helix, period/aryl hydrocarbon receptor/single minded* (bHLH-PAS) (Brunskill et al. 1999). Las proteínas miembros de esta familia son todos reguladores críticos de la expresión génica, capaces de responder frente a diversos estímulos ambientales (incluida luz, oxígeno, voltaje y potencial redox) y de cumplir múltiples roles durante el desarrollo del sistema nervioso, las oscilaciones circadianas y el desarrollo de las tráqueas, entre otras (Massari y Murre 2000; Kewley, Whitelaw, y Chapman-Smith 2004).

Se pueden reconocer dos grandes familias de factores de transcripción bHLH/PAS: los miembros de la clase I son factores que no pueden ni homodimerizar ni heterodimerizar con otros factores de la clase I. Entre este grupo se encuentran NPAS3, las proteínas aryl hydrocarbon receptor (AhR), factores respondedores a hipoxia (HIF; HIF-1, HIF-2, y HIF-3) y proteínas de la familia *single minded* (SIM1 y SIM2). Para formar complejos activos, esta primera clase de proteínas debe dimerizar con proteínas de la clase II. Por otro lado, la clase II de las bHLH/PAS pueden homo y heterodimerizar promiscuamente. Uno de los miembros de esta familia es la proteína *carbon receptor nuclear translocator* (ARNT) (Lindebro, Poellinger, y Whitelaw 1995; Wang et al. 1995; Ema et al. 1996).

Todos los miembros de la familia de proteínas bHLH/PAS se caracterizan por tener un dominio N-terminal rico en aminoácidos básicos por donde la proteína reconoce las secuencias blanco y se unen al ADN (*basic*) adyacente a un dominio de dimerización primario (HLH). Tanto el dominio básico como el HLH son necesarios para la formación de complejos activos sobre el ADN (Bacsi y Hankinson 1996). Además, estos factores de transcripción tienen una o dos secuencias repetitivas degeneradas de aproximadamente 130 aminoácidos cada una (PAS-A y PAS-B) (figura 7, A). Todas las proteínas de la superfamilia bHLH/PAS deben dimerizar para formar complejos funcionales sobre el ADN. En este punto, el dominio PAS juega un rol fundamental. Por un lado, los dominios PAS son capaces de censar cambios ambientales y traducir estos cambios en modificaciones en el patrón de interacción de la proteína. De esta forma, los dominios PAS brindan una interfase de interacción secundaria y determinan con que pareja podrá dimerizar. En el caso de NPAS3, la proteína heterodimeriza con ARNT2 (Wu et al. 2016). Finalmente, hacia el extremo C-terminal, los miembros de esta familia pueden contener diversos motivos de transactivación responsables de la modulación de la expresión de sus genes blanco (Kewley, Whitelaw, y Chapman-Smith 2004; Benjamin S. Pickard et al. 2006) (figura 7, B).

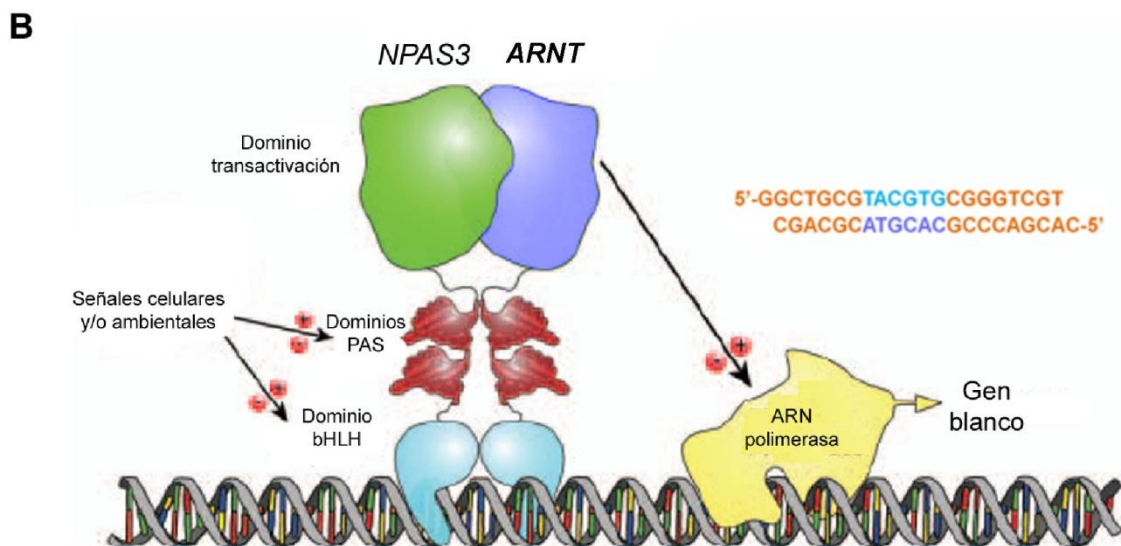
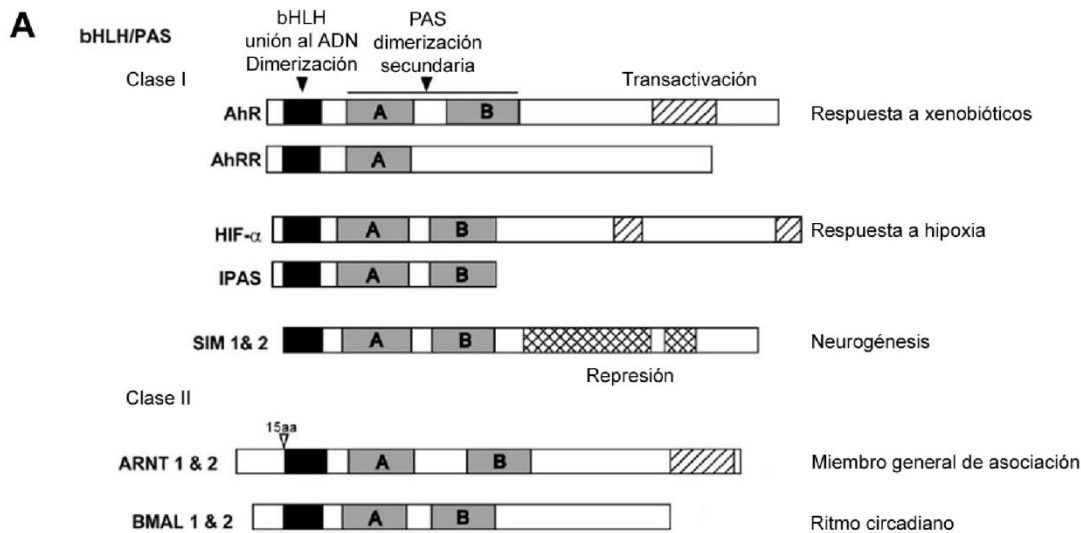


Figura 7. Representación esquemática de la estructura de algunos de los factores de transcripción de la familia bHLH-PAS. (B) Esquema que representa el conocimiento actual de la función de NPAS3. En el dibujo se muestra a NPAS3 unida a su secuencia blanco sobre el ADN formando un heterodímero con ARNT2. Ambas proteínas modulan conjuntamente la expresión de sus genes blancos a través de la interacción con la polimerasa de ARN tipo II. El esquema hace hincapié en la capacidad potencial de NPAS3 de responder frente a ciertos estímulos ambientales aún no identificados. También se ve la secuencia consenso de reconocimiento por parte del heterodímero. Modificado de Kewley, Whitelaw y Chapman-Smith 2004; Pickard et al. 2006.

Actualmente, aparte de NPAS3, se conocen otros tres miembros de la subfamilia NPAS: NPAS1, un gen cuya disfunción se asocia con el desarrollo de esquizofrenia (Erbel-Sieler et al. 2004), NPAS2, que participa de procesos relacionados con la memoria y las oscilaciones circadianas (Reick et al. 2001) y NPAS4, que participa del desarrollo del sistema nervioso (Lin et al. 2008). La comparación de las secuencias aminoacídicas de los miembros

de la familia de proteínas bHLH/PAS muestra que NPAS1 y NPAS3 están íntimamente relacionados, mostrando la mayor identidad de secuencia dentro de los integrantes del subgrupo de proteínas NPAS (Yan et al. 2014). Sin embargo, una diferencia notable entre estos dos genes está en el largo de sus unidades transcripcionales: los once exones de NPAS1 están localizados en un segmento de 20 kpb en el cromosoma 7 del ratón, mientras que los exones de NPAS3 están distribuidos en un locus genómico que abarca 791 kpb del cromosoma 12 de ratón (correspondiente al cromosoma 14 humano) (Erbel-Sieler et al. 2004).

HIPÓTESIS GENERAL

La hipótesis de este proyecto es que la adquisición de nuevos patrones de expresión de genes relacionados con el desarrollo y la función cerebral en el linaje humano habría sido crítica para la evolución neuroanatómica de nuestro cerebro y sus capacidades cognitivas diferenciales. Estos nuevos patrones de expresión estarían codificados en cambios en la secuencia de las regiones regulatorias de genes que se expresan en el sistema nervioso central. Es muy probable que estos nuevos patrones de expresión hayan dotado a nuestros antepasados de ventajas adaptativas que fueron seleccionadas y contribuyen a la evolución de las características diferenciales de nuestra especie.

OBJETIVO GENERAL

El objetivo principal de este proyecto es estudiar las bases genéticas de la evolución del cerebro humano. Para ello se propone realizar un estudio detallado a nivel evolutivo y funcional de regiones genómicas no codificantes, altamente conservadas y que cambiaron su tasa de evolución molecular exclusivamente en el linaje humano que se agrupan en el locus del factor de transcripción *NPAS3*. Evidencias obtenidas en nuestro laboratorio indican que estas regiones podrían contener elementos regulatorios de la expresión génica que cambiaron su actividad exclusivamente en nuestro linaje. Específicamente:

- ❖ Determinar si los *NPAS3*-HAEs tienen la capacidad de actuar como *enhancers* transcripcionales con un análisis comparativo de expresión en ensayos en líneas estables de peces cebra y ratones transgénicos portadores de las regiones ortólogas de humano y chimpancé,
- ❖ Generar y estudiar modelos de animales transgénicos en los cuáles a través de ensayos de ingeniería genética se intentará determinar el impacto funcional de la ausencia o el reemplazo de estas regiones regulatorias tienen en el gen *NPAS3*,
- ❖ Evaluar con mayor detalle la función del gen *NPAS3* en el neurodesarrollo y en el desarrollo craneofacial, aspectos que no han sido completamente estudiados y qué son

relevantes para comprender la importancia de la evolución linaje específica de este gen y su maquinaria regulatoria en humanos.

Capítulo I

INTRODUCCIÓN ESPECÍFICA

Como ya fuera mencionado en la Introducción general, cinco grupos de investigación han llevado a cabo estudios a gran escala con el objetivo de detectar regiones aceleradas en el genoma humano (Bird et al. 2007; Bush and Lahn 2008; Lindblad-Toh et al. 2011; Pollard et al. 2006; Prabhakar et al. 2006). Cada uno de ellos, de forma independiente, aplicó su propia metodología produciendo como resultado una base de datos de elementos acelerados. Cuatro de estas bases de datos fueron utilizadas por la Dra. Kamm en su tesis doctoral realizada en nuestro laboratorio para generar una base de datos no redundante con 1629 elementos que se denominaron *Human Accelerated Elements* (HAEs por sus siglas en inglés, Figura 1, Kamm et al. 2013). Cabe destacar que la base de datos elaborada por Bird y cols. (C. P. Bird et al. 2007) no fue utilizada para la generación de la base de datos HAEs porque la metodología empleada difiere sustancialmente de las metodologías empleadas por el resto de los grupos.

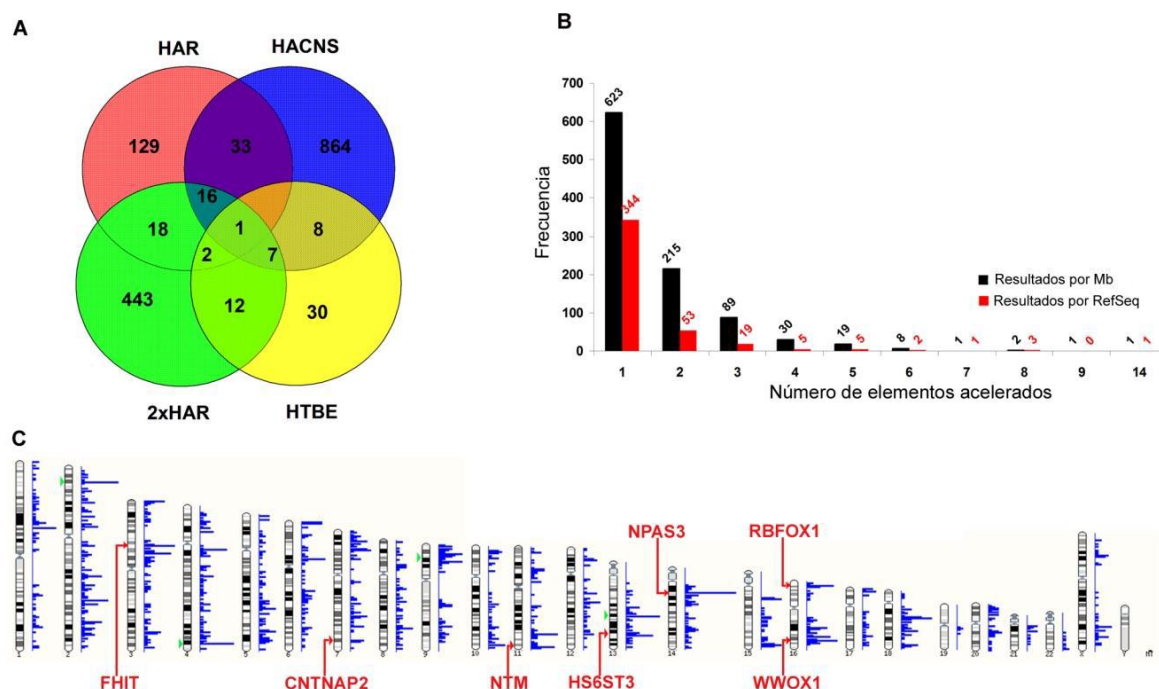


Figura 1. Agrupamiento de HAEs en el genoma humano. (A) Diagrama de Venn mostrando el solapamiento de los HAEs provenientes de las diferentes bases de datos originales, HARs (Pollard, Salama, King, et al. 2006; Pollard, Salama, Lambert, et al. 2006); HACNS (Prabhakar et al. 2006); HTBE (Bush y Lahn 2008); y 2xHARs (Lindblad-Toh et al. 2011). Adicionalmente, hay 65 elementos compartidos entre las bases de datos 2XHAR y HACNS y 3 elementos compartidos por las bases de datos de HARs y HTBE que no se muestran en el diagrama. El número total de HACNS y 2XHARs

expresados en el diagrama es mayor al número de elementos publicados en las bases de datos originales debido a solapamientos múltiples en los *loci* HACNS56/2, HAR20 y HACNS11. (B) Histograma de frecuencia mostrando la distribución de HAEs por gen humano (unidad transcripcional o UT) o por intervalos genómicos de 1 Mb (solo se muestran los resultados para la ventana 2). (C) Esquema donde se muestra con barras azules la densidad de HAEs a lo largo de los cromosomas humanos generado con el *density track* disponible en el sitio Ensembl (www.ensembl.org). Con flechas rojas se muestra la ubicación de las UT con 6 o más HAEs. Con flechas verdes se indica la ubicación de los intervalos genómicos de 1 Mb con 8 o más HAEs en caso de que no hayan sido indicados anteriormente con flechas rojas (Modificado de Kamm 2013).

De este análisis surgió como resultado que el locus del gen *NPAS3* contiene la mayor cantidad de HAEs de todo el genoma humano (Tabla 1). En particular, se encontró que la unidad transcripcional (UT: se la definió desde el comienzo al final de la transcripción incluyendo intrones) de *NPAS3* contiene 15 HAEs no redundantes, todos ellos intrónicos. De ellos, 6 son HARs (HAR21, HAR89, HAR96, HAR173, HAR189 y HAR202), 5 son HACNS (HACNS96, HACNS221, HACNS490, HACNS553 y HACNS658) y 4 son 2xHAR (2xHAR122, 2xHAR142, 2xHAR157 y 2xHAR2). Sin embargo, los elementos HACNS490 y HAR189 se encuentran parcialmente solapados, reduciendo el número de elementos independientes de 15 a 14 (Figura 2).

Tabla 1. Genes y regiones genómicas conteniendo 6 o más HAEs

# HAEs	Posición RefSeq (NCBI36/hg18)	Símbolo Gen	HAEs por kpb
14	chr14:32,478,210-33,340,702	<i>NPAS3</i>	1,62
8	chr3:59,710,076-61,212,173	<i>FHIT</i>	0,53
	chr16:6,009,133-7,703,341	<i>RBFOX1</i>	0,47
	chr16:76,691,052-77,804,065	<i>WWOX</i>	0,72
7	chr11:130,745,581-131,711,926	<i>NTM</i>	0,72
6	chr7:145,444,386-147,749,021	<i>CNTNAP2</i>	0,26
	chr13: 95,541,093-96,289,813	<i>HS6ST3</i>	0,80



Figura 2. Localización del *locus* de *NPAS3* en el cromosoma 14 del genoma humano (NCBI36/hg18) y los elementos acelerados HARs (rojo), HACNS (azul) y 2xHARs (verde) presentes en *locus*.

Respecto a la región codificante, *NPAS3* posee una identidad de secuencia del 95% con el resto de los mamíferos, un 98,7% con los primates, y un 99,8% con los chimpancés con solo dos sustituciones aminoacídicas entre 940 residuos. Utilizando el programa PAML se analizó si *NPAS3* presentaba signo de selección positiva en humanos. Este análisis no detectó señales de selección positiva en nuestro linaje (Kamm, López-Leal, et al. 2013a).

Para investigar si los *NPAS3*-HAEs desempeñan alguna función Kamm et al (2013a) investigaron la capacidad de los 14 HAE de *NPAS3* para actuar como *enhancers* de la transcripción durante el desarrollo del cerebro mediante el uso de un ensayo de expresión de reportero en peces cebra transgénico transitorios. Se clonaron cada uno de los 14 HAEs en un vector de expresión basado en el sistema Tol2 (Kawakami 2005) río arriba de un promotor mínimo de *c-Fos* de ratón (Dorsky, Sheldahl, y Moon 2002) seguido de la secuencia codificante de la proteína fluorescente verde (EGFP; Fisher et al. 2006). Los análisis de expresión realizados a 24 y 48 horas post fecundación (hpf) revelaron que 11 de los 14 *NPAS3*-HAE dirigen patrones de expresión de EGFP distintivos y reproducibles en el sistema nervioso en desarrollo del pez cebra (Figura. 3 y Tabla 2). En conjunto, estos resultados mostraron que la mayoría de los *NPAS3*-HAE pueden actuar como *enhancers* transcripcionales durante el desarrollo del sistema nervioso (Kamm, López-Leal, et al. 2013a). A través de análisis bioinformáticos, se encontró además que todas las sustituciones humanas específicas presentes en cada uno de los 14 *NPAS3*-HAE producen una ganancia o

pérdida de sitios de unión de factores de transcripción (TFBS; Tabla 3). Estos resultados sugieren que los cambios adquiridos en humanos podrían haber contribuido a cambios en el dominio de la expresión espacio-temporal de *NPAS3* en el cerebro humano (Kamm, López-Leal, et al. 2013a).

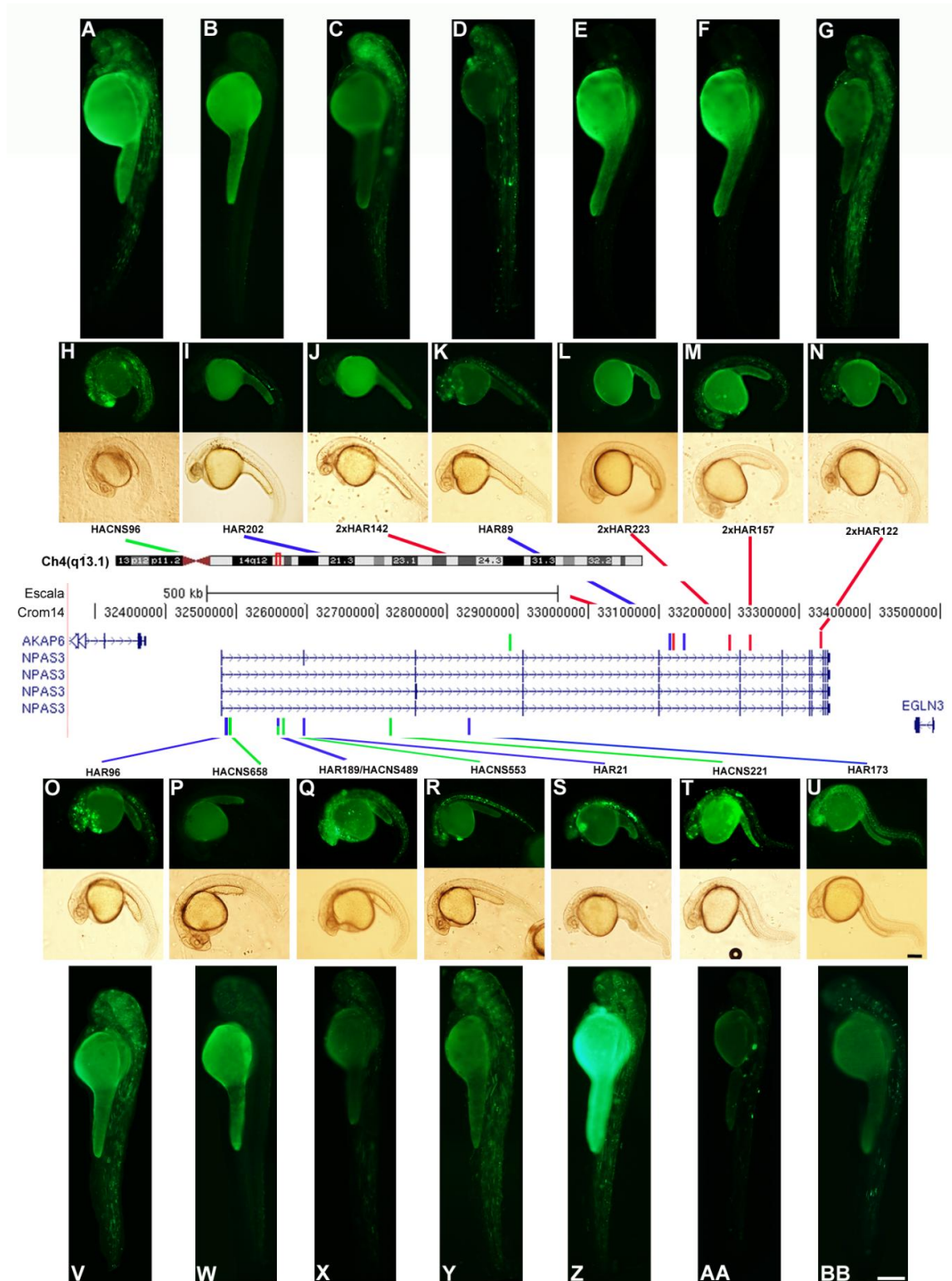


Figura 3. Ensayo de expresión transitoria en peces cebra transgénicos. En el panel central se muestra la localización del *locus* de *NPAS3* en el cromosoma 14 del genoma humano (NCBI36/hg18) en una ventana de 1.1 Mb de longitud de acuerdo con el *UCSC Genome Browser* (www.genome.ucsc.edu). Abajo en barras coloreadas se indica la localización de los 14 HAEs y debajo las variantes de *splicing* de *NPAS3*. Las barras azules indican los HAR (Pollard et al. 2006), las verdes representan los HACNS (Prabhakar et al. 2006) y las rojas muestran los 2xHARs (Lindblad-Toh et al. 2011). Los elementos solapados se muestran con barras de dos colores (azul y verde). Arriba y debajo del panel central se muestran micrografías de peces transgénicos con el patrón representativo de la expresión mosaico de EGFP para cada elemento a 24 hpf (H,I,J,K,L,M,N,O,P,Q,R,S,T y U) y 48 hpf (A,B,C,D,E,F,G,V,W,X,Y,Z y AA). Escala: 250 μ m.

Tabla 2: Análisis de expresión en peces cebra transgénicos a 24 y 48 hpf.

Nombre del Transgén	Localización genómica (NCBI36/hg18)	Largo (pb)	N	% (+)	% C	Dominios de expresión consistente							
						T	M	R	O	M E	A F	C	MS
HAR96Hs	chr14: 32486645-32487168	524	87/ 72	82/ 83	39/21	+	+	+	+	-/+	+/-	-	-
HACNS658Hs	chr14: 32487503-32487837	335	116/ 94	25/ 34	-	-	-	-	-	-	-	-	-
HAR189Hs/ HACNS489Hs	chr14: 32558123-32558864	742	71/ 67	80/ 82	42/18	+/-	+/-	+/-	-	+/-	-	-	+/-
HACNS553Hs	chr14: 32566093-32566346	254	94/ 108	63/ 82	50/32	+	+	+	+	+	-	+/-	+/-
HAR21Hs	chr14: 32596637-32596862	298	71/ 64	84/ 86	38/33	+	+	+	+	+	+/-	+/-	+
HACNS221Hs	chr14: 32718813-32719042	230	62/ 56	77/ 68	34/37	+/-	+	+	+/-	+/-	-	-	+/-
HAR173Hs	chr14: 32830590-32830773	184	79/ 69	75/ 74	28/37	+/-	+	+/-	+/-	+/-	-	-	-
HACNS96	chr14: 32888897-32889211	315	95/ 100	78/ 84	26/17	+	+	+	+	+	+/-	+/-	+
HAR202Hs	chr14: 33115156-33115595	440	69/ 75	32/ 17	-	-	-	-	-	-	-	-	-
2xHAR142Hs	chr14: 33118315-33118577	263	94/ 40	81/ 70	27/30	+	+/-	+	+	+	+/-	-	+/-
HAR89Hs	chr14: 33135162-33135937	776	64/ 28	83/ 86	22/36	+	+	+	+/-	+	+/-	-	+
2xHAR223Hs	chr14: 33199815-33200914	1100	86/ 33	42/ 36	-	-	-	-	-	-	-	-	-
2xHAR157Hs	chr14: 33229602-33230232	631	58/ 32	90/ 44	36/37	+/-	+/-	+/-	+/-	+/-	-	-	+/-
2xHAR122Hs	chr14: 33331419-33331887	469	62/ 31	80/ 84	53/45	+	+	+	+	+	-	-	+/-
cFos		-	64/ 114	30/ 33	-	-	-	-	-	-	-	-	-

Nota: Las celdas en verde indican los HAEs que no mostraron actividad *enhancers* y el control (cFos). Las coordenadas genómicas, el largo y el número de sustituciones humanas específicas corresponden a la región conservada completa que contiene al HAE y la que fue usada para generar los transgenes. Referencias: N, número de animales analizados. % (+), porcentaje de embriones con expresión de EGFP. %C, porcentaje de embriones con expresión consistente de EGFP. T, telencéfalo. M, mesencéfalo. R, rombencéfalo. O, ojo. ME, médula espinal. AF, arcos faríngeos. C, corazón, MS, músculo somítico. Cuando los resultados son diferentes entre las dos edades, entre barras se indican los resultados a 24/48 hpf e (+) y (-) indica presencia o ausencia de expresión consistente en la estructura respectivamente.

Gracias a la secuenciación de múltiples genomas humanos fue posible detectar sitios polimórficos de alta frecuencia en diferentes poblaciones humanas. Por ejemplo, el proyecto 1000 genomas secuenció, en su primera fase, 1.092 genomas humanos provenientes de 14 poblaciones de África, Europa, Este de Asia y América (Durbin et al. 2010; 1000 Genomes Project Consortium et al. 2012). Gracias a este análisis se detectaron sitios polimórficos presentes en al menos una población humana en todos los HAEs de *NPAS3* (considerando también las secuencias conservadas que contienen a los HAEs) (tabla 3). En particular, existen al menos cuatro regiones HAEs de *NPAS3* que tienen polimorfismos registrados para los sitios con sustituciones humano-específicas (HAR189, HAR202,

HACNS658 y 2xHAR142) (tabla 3). De los 5 sitios polimórficos en poblaciones humanas ubicados en las posiciones registradas como variantes específicas de humanos en estos 4 elementos, en dos de esos polimorfismos todas las variantes halladas son derivadas respecto a la base homóloga ancestral inferida a partir de los alineamientos con genomas de otros mamíferos (G/T en humanos y A ancestral en polimorfismo rs145548775 del HAR189 y C/G en humanos y T ancestral en el polimorfismo rs114570257 del 2xHAR142). Para los otros tres sitios polimórficos en bases específicas de humanos en los HAEs intrónicos de *NPAS3*, una de las variantes polimórficas corresponde con la base ancestral inferida (C/T en humanos y T ancestral en el polimorfismo rs10134006 del 2xHAR142, A/T en humanos y T ancestral en el polimorfismo rs79570848 del HAR202 y T/C en humanos y C ancestral en el polimorfismo rs10149308 del HACNS658). Las restantes sustituciones específicas de los humanos presentes en los HAEs de *NPAS3* resultan ser no polimórficas.

Adicionalmente, la publicación del genoma de Neandertal (Prüfer et al. 2014) y de Denisovanos (M. Meyer et al. 2012), permitió revisar la datación del origen de las sustituciones específicas de humanos en los HAEs. La gran mayoría de las sustituciones específicas de los humanos resultaron estar compartidas con los genomas de los humanos arcaicos disponibles en la actualidad (Neandertales y Denisovanos). Sin embargo, se encontraron algunas excepciones. Tres pares de bases presentes en tres HAEs (HACNS658, HAR202, 2xHAR157) de *NPAS3* fueron anotadas como derivadas en humanos ya que son diferentes a las bases ortólogas en Neandertales y Denisovanos. Dos de estas bases son polimórficas en las poblaciones humanas actuales (rs10149308 en HACNS658 y rs79570848 en HAR202) y una tercera es fija (base chr14:34160070 en 2xHAR157).

Tabla 3. Análisis de polimorfismos humanos actuales en HAEs de *NPAS3* y secuencias en genomas de humanos arcaicos, y cambios en los sitios de unión para factores de transcripción (TFBS por sus siglas en inglés) entre las secuencias de los HAEs y regiones ortólogas en chimpancé.

HAE	SP	SustHs	D/A	Polimorfismo; número	Humanos Arcaicos	Estado	Conservación con pez cebra	TFBS ganancia (↑) y pérdida (↓)	Sesgo AT→CG
HAR96	7	2	G/A T/A		G T		NO (Rana)	↑YTBP; ↓EVII ↑CREB; ↓RORA ↑ETSF; ↓AP4R ↓PERO	A→G; A→T

HACNS658	6	6	C/T	rc10149308	C	P	NO (Lagarto)	↑LEFF; ↓HOMF	A→G; T→C
			T/C		T/C			↑NR2F; ↓HOXF	T→A; C→T
			G/A		G			↑XBBF; ↓AP1R	T→C; C→T
			A/T		A			↑PERO; ↓STEM	
			T/C		T				
			C/T		C				
HAR189/ HACNS490	10	6	G/C	rs145548775	G	NO (Lagarto)	↑FKHD; ↓PLAG	C→G; T→C	
			C/T		C		↑E2FF; ↓GCMF	T→C; T→G	
			C/T		C		↑HOMF; ↓NR2F	A→G; G→A	
			G/T		G		↑GRHL; ↓TALE		
			G/A		G		↑INRE; ↓TEAF		
			A/G		A		↑CLOX; ↓RREB		
							↑HZIP; ↑BRNF ↑BNCF		
HACNS553	5	4	G/A		G	NO (Pollo)	↑EGRF; ↓GTBX	A→G; T→G	
			G/T		G		↑NSM; ↓CAAT	A→T; A→G	
			T/A		T		↓GATA		
			G/A		G		↓CLOX		
HAR21	2	3	T/G		T	NO (Lagarto)	↑NKXH; ↓BRAC	G→T; A→G	
			G/A		G		↑RU49; ↓DMRT	C→T	
			T/C		T		↓EV11; ↓ABDB ↓TALE; ↓GATA		
HACNS221	3	5	T/C		T	NO (Ornitorrinco)	↑SORY; ↓FKHD	C→T; T→G	
			G/T		G		↑BRN5; ↓EV11	T→G; C→T	
			G/T		G		↑EV11; ↓APIR	T→C	
			T/C		T		↓CREB; ↓OVOL		
			C/T		C		↓PARF; ↓GREF		
HAR173	2	2	C/A		C	NO (Lagarto)	↑PRDF; ↓HBOX	A→C; A→G	
			G/A		G		↑GF11; ↓SORY ↑CLOX; ↓HOXF ↓PLZF		
HACNS96	6	6	G/A		G	SI	↑WHNF; ↓CREB	A→G; A→C	
			C/A		C		↑MOKF; ↓H1FF	T→C; A→G	
			C/T		C		↑HNF1; ↓HASF	T→C; T→G	
			G/A		G		↑GATA; ↓BRAC		
			C/T		C		↑MYBL; ↓ABDB		
			G/A		G		↑HOMF; ↓FKHD		
			C/T		C		↑MEF2; ↓MEF2		
			G/T		G		↑MYT1; ↓CDXF ↑RORA; ↓VTBP ↓MYT1; ↓ARID ↓PAX5; ↓NKXH		
HAR202	8	5	C/T	rs79570848	C	F	SI	↓BCL6	T→C; T→A
			A/T		T			↓STAT	A→G; C→T
			G/A		G				

			T/C G/T	T G				T→G
2xHAR142	4	6			NO (Mamíferos placentarios)		↑CLOX; ↓STEM	A→G; T→C
			G/A	G			↑CREB; ↓ZF02	G→T; A→G
			C/T	rs114570257			↑HIF1; ↓NR2F	A→G; T→C
			T/G				↑HNF1; ↓BCL6	
			G/A				↑HZIP; ↓EVI1	
			G/A				↑MYBL; ↓OCT1	
			C/T	rs183390399			↓HOXF; ↓CART ↓MEF2; ↓SATB ↓PDX1	
HAR89	7	3			SI		↑PAX5; ↓SATB	A→G; A→G
			G/A	G			↑HOMF; ↓OCT1	G→A
			G/A	G			↑MYBL; ↓ABDB	
			A/G	A			↑PERO; ↓HOXF	
							↑NR2F; ↓CART	
							↓NF1F; ↓RUSH	
							↓YY1F	
2xHAR223	13	8	T/C	T	SI		↑CART; ↓CHOP	C→T; T→C
			C/T	C			↑ZF03; ↓EVI1	T→C; A→G
			C/T	C			↑NKXH; ↓SORY	T→C; G→C
			G/A	G			↑SORY; ↓OCT1	T→C; G→T
			C/T	C			↑HOXC; ↓STEM	
			C/G	C				
			C/T	C				
2xHAR157	5	10	T/G	T	SI	F	↑OAZF; ↓PARF	T→C; T→C
			C/T	C			↑VTBP; ↓OCT1	T→C; A→G
			C/T	C			↑HOMF; ↓SORY	C→A; C→G
			G/A	G			↑NKHX; ↓AIRE	A→C; C→A
			A/C	C			↑GF11; ↓RREB	C→T; C→G
			G/C	G			↑HOXC; ↑ZF06	
			C/A	C			↑LEFF; ↑SMAD	
2xHAR122	9	4	A/C	A	SI		↑FKHD	
			T/C	T				
			G/C	G				
			C/T	C			↑MYBL; ↓SORY	T→C; T→C
			C/T	C			↑CEBP; ↓FKHD	A→G; A→G
			G/A	G			↑CARE; ↓IRFF	
			G/A	G			↓SNAP	

Nota: El número de sustituciones humanas específicas corresponden a la región conservada completa que contiene al HAE (Siepel et al. 2005). Adicionalmente, en sesgo AT→CG se muestra el tipo de sustituciones ocurridas en los HAES en el linaje humano. En la última columna y en celdas amarillas se muestran aquellos elementos que presentan un sesgo fuerte (mayor al 60% de las

sustituciones ocurridas) AT→CG (indicio del proceso evolutivo neutral BGC), mientras que en celeste se indica un sesgo CG→AT. Abreviaturas: SP, número de sitios polimórficos en los HAEs en poblaciones humanas actuales de acuerdo al *1000 genomes project phase 1* (track disponible en www.ucsc.edu/) ((Durbin et al. 2010; 1000 Genomes Project Consortium et al. 2012)). Sust Hs, número de sustituciones específicas de los humanos dentro de los HAEs. D/A, base derivada /base ancestral inferida a partir del alineamiento de múltiples genomas de mamíferos en las posiciones que cambiaron exclusivamente en el linaje humano. PHs, alelos en sitios polimórficos en poblaciones humanas actuales en las posiciones que cambiaron exclusivamente en el linaje humano. En humanos arcaicos se indica la base correspondiente inferida en el genoma de humanos arcaicos a partir de la secuenciación de múltiples genomas de Neandertales y Denisovanos (Reich et al. 2010; M. Meyer et al. 2012; Prüfer et al. 2014). En recuadros rojos se indica cuando la base en esta posición fue anotada como derivada en humanos modernos respecto a la base inferida en los humanos arcaicos. En condición se indica si la base derivada en humanos fue detectada en condición polimórfica (P) o fija (F) en los genomas de humanos arcaicos (Prüfer et al. 2014).

Además, en otro trabajo (Kamm, Pisciotto, et al. 2013b), Kamm estudió en particular el 2xHAR142, uno de los 14 NPAS3-HAE, que se encuentra en el quinto intrón del locus del gen NPAS3 (Figura 4). Usando ratones transgénicos, mostró que tan el elemento humano como sus ortólogos de ratón y chimpancé 2xHAR142 se comportan como *enhancers* transcripcionales que conducen la expresión del gen reportero lacZ a un subdominio de expresión NPAS3 similar en el sistema nervioso central del ratón. Curiosamente, el ortólogo 2xHAR142 humano dirige la expresión de lacZ en un patrón de expresión extendido en el sistema nervioso en desarrollo del ratón (Figura 5).

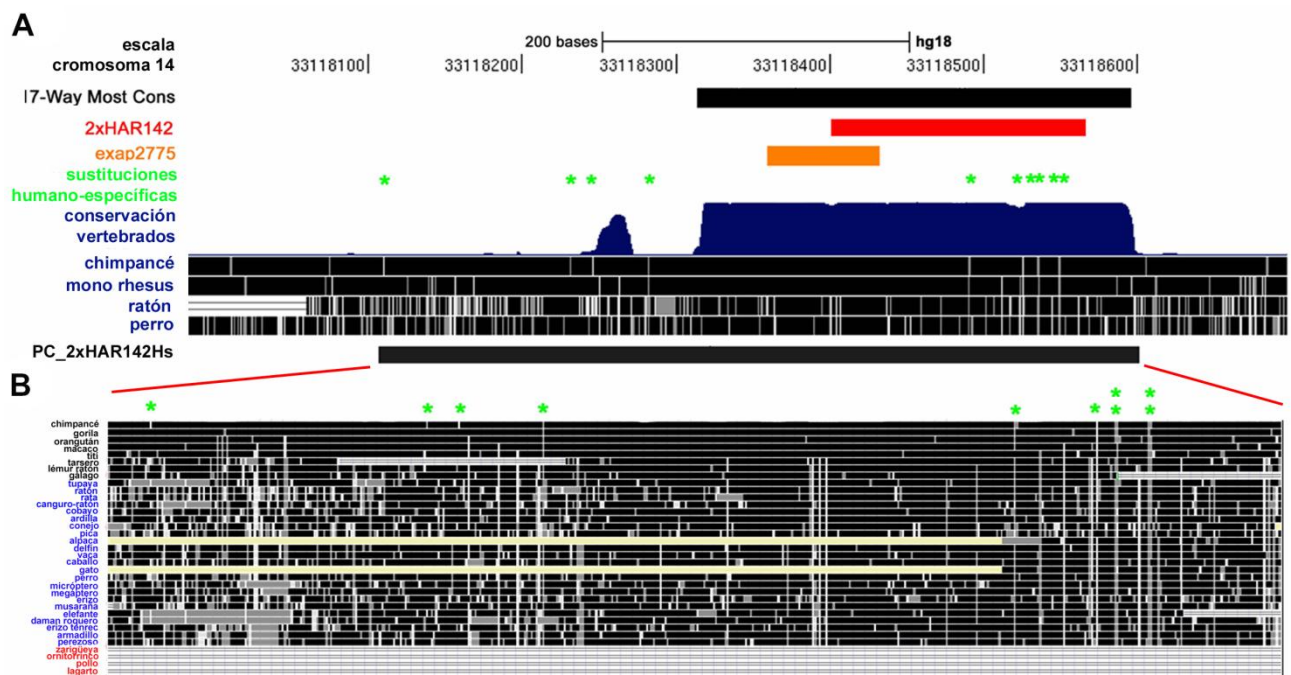


Figura 4. Esquema de la región clonada en el transgén PC_2xHAR142Hs conteniendo la región conservada descrita por el programa PhastCons, el elemento 2xHAR142. (B) Conservación entre vertebrados de la secuencia del elemento PC_2xHAR142Hs de acuerdo con el *USCS Genome Browser*. Con asteriscos verdes se indican las sustituciones humano-chimpancé.

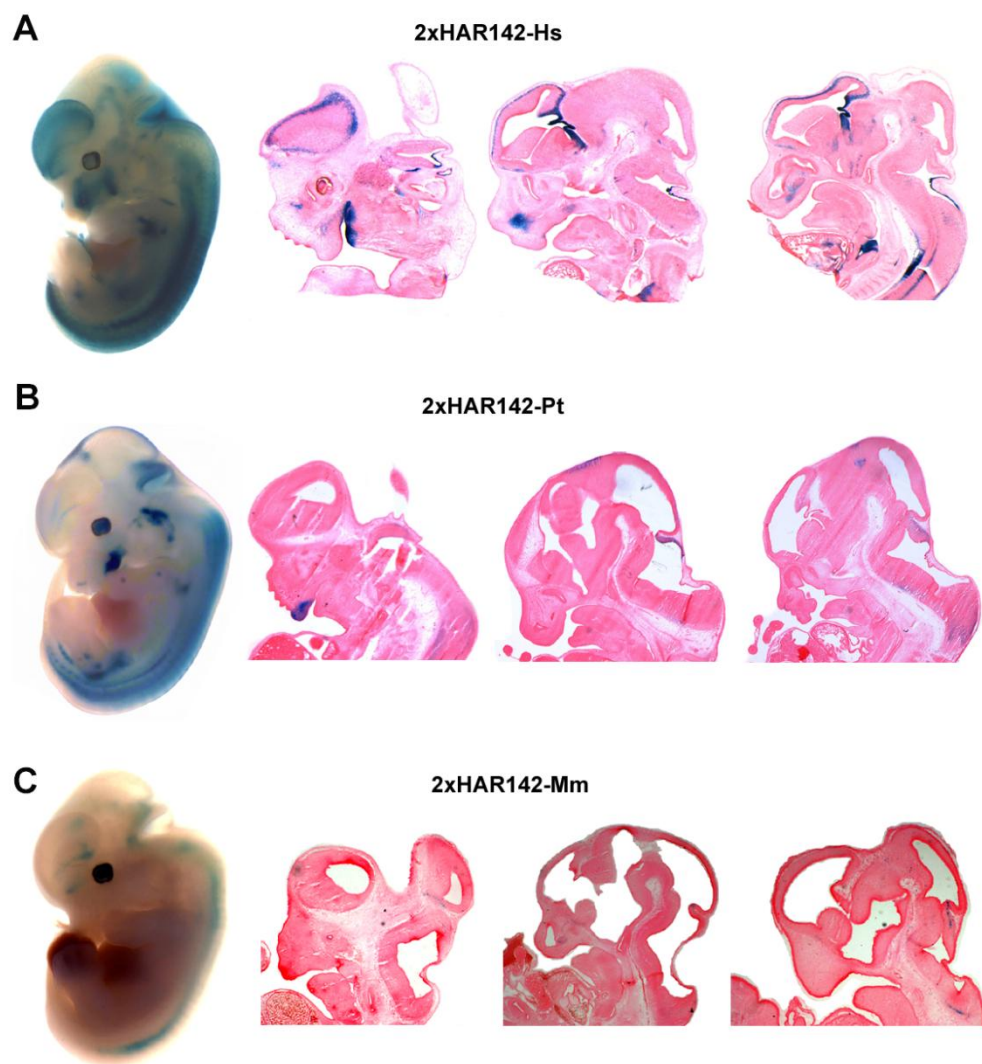


Figura 5. Análisis de expresión comparativo entre los elementos PC_2xHAR142Hs, PC_2xHAR142Pt y PC_2xHAR142Mm en ratones transgénicos. Patrones de expresión de *lacZ* representativos de los elementos (A) PC_2xHAR142Hs, (B) PC_2xHAR142Pt y (C) PC_2xHAR142Mm en embriones completos y en cortes sagitales de ratón a 12.5 dpc.

HIPÓTESIS DE TRABAJO

Las regiones *NPAS3*-HAEs funcionan como *enhancers* transcripcionales y el proceso evolutivo ocurrido en el linaje humano en estas regiones resultó en diferencias funcionales entre los *enhancers* humanos y los *enhancers* ortólogos de nuestro pariente primate más cercano, el chimpancé (*Pan troglodytes*).

OBJETIVOS ESPECÍFICOS

- Investigar si los 14 HAEs presentes en los intrones de *NPAS3* tienen la capacidad de actuar como *enhancers* transcripcionales en ensayos estables en peces cebra transgénicos.
- Estudiar comparativamente la actividad *enhancer* entre los elementos humanos y las regiones ortólogas de chimpancé de los 14 *NPAS3*-HAEs en ensayos de expresión en peces cebra transgénicos.
- Analizar comparativamente la actividad *enhancer* y el patrón de expresión del elemento HAR202, uno de los *NPAS3*-HAEs presentes en el 5to intrón de *NPAS3*, en un modelo mamífero. Para ello, se generarán ratones transgénicos conteniendo la secuencia humana o de chimpancé de este elemento fusionada río arriba de un promotor mínimo dirigiendo la expresión de la proteína reportera *lacZ*.
- Analizar funcionalmente las sustituciones específicas de *H. sapiens* del elemento acelerado HAR202 que surgieron posteriormente a la divergencia entre humanos modernos y Neandertales/ Denisovanos.
- Estudiar si los *NPAS3*-HAEs que fueron identificados como *enhancers* estarían regulando la transcripción del gen *Npas3* endógeno a través del análisis de experimentos de conformación de la cromatina.

RESULTADOS Y DISCUSIÓN

1RA PARTE: ANÁLISIS FUNCIONAL DE LOS *NPAS3*-HAES EN PECES CEBRA

Comparación de la funcionalidad de las regiones aceleradas humanas entre líneas estables y transitorias de peces cebras transgénicos

Cómo ya fuera mencionado en nuestro laboratorio se analizó la capacidad de actuar cómo *enhancers* transcripcionales a los 14 *NPAS3*-HAes en ensayos transitorios en peces cebras transgénicos (Kamm, López-Leal, et al. 2013a). Decidimos analizar de manera más profunda este aspecto, y validar la reproducibilidad de los ensayos, en líneas estables transgénicas de peces cebras. En comparación con la transgénesis transitoria, las líneas estables permiten generar líneas de peces transgénicos con una expresión que es homogénea y no de tipo mosaico. Esto permite estudiar en detalle la reproducibilidad del patrón de expresión del transgén en distintas líneas estables. Por otra parte, el hecho de que la prole puede heredar el transgén en forma mendeliana permite la realización de experimentos en los que se necesitan muchos animales que sean idénticos en su patrón de expresión (Suster et al. 2009).

Para obtener las líneas transgénicas estables se amplificaron mediante PCR las 14 regiones conservadas (detectadas a través del alineamiento de los genomas de múltiples especies de vertebrados usando el programa PhastConst, Siepel et al. 2005), que contienen los HAes presentes en intrones de *NPAS3*, de humano. Estas regiones fueron clonadas en un vector río arriba del promotor mínimo del gen *cFos* de ratón que se halla fusionado a la secuencia codificante de la proteína reportera EGFP. Este vector contiene las secuencias Tol2 (secuencias derivadas de un transposón) que son reconocidas por una transposasa del pez Medaka (*Oryzias latipes*) la cual media la transgénesis en el genoma de pez cebras (Kawakami 2005; Fisher, Grice, Vinton, Bessling, Urasaki, et al. 2006). Estos vectores conteniendo cada una de las regiones *NPAS3*-HAes se inyectaron en embriones de pez cebras de una célula. Los animales obtenidos fueron criados hasta la adultez y cruzados con peces de tipo salvaje o wild type. El patrón de expresión se estudió en embriones obtenidos de estas cruces a 48 hpf. Se analizaron al menos tres líneas estables por cada región estudiada

por cada especie analizada. Se clasificó los dominios de expresión de EGFP utilizando únicamente términos anatómicos generales y comparables entre vertebrados (como ser “médula espinal” o “prosencéfalo”) y se evitó la clasificación dentro de subdominios o tipos celulares (Ariza-Cosano et al. 2012).

Nuestros resultados mostraron que, de los 14 elementos probados, 9 elementos actuaron como *enhancers* de la expresión de EGFP (Figura 6, Tabla 4). De los elementos que dirigieron la expresión de EGFP en líneas estables (HAR21, HAR89, HAR96, HAR173, HACNS96, HACNS221, HACNS490, 2xHAR122, 2xHAR157), las áreas de expresión fueron las mismas que las producidas en las líneas transitorias (Tabla 2). Por otra parte, dos elementos (2xHAR142, HACNS553) mostraron diferencias con respecto a los resultados obtenidos en ensayos transitorios ya que en nuestros ensayos estables estos elementos no mostraron capacidad de dirigir la expresión del gen reportero a ninguna región del pez cebra en desarrollo de manera consistente (Tabla 2; Figura 6).

En particular, el elemento 2xHAR142 produjo un resultado diferente. Si bien la expresión en sistema nervioso en transitorio de EGFP dirigida por este elemento era evidente particularmente a 48 hpf (Figura 3, Tabla 2; Kamm, López-Leal, et al. 2013a), ninguna de las cuatro líneas estables analizadas mostró expresión en ninguna de las edades estudiadas (Figura 6, Tabla 4). Este elemento fue, además, ampliamente estudiado en ratones transgénicos en nuestro laboratorio donde se observó que la expresión dirigida por el elemento humano además de ser fuerte y consistente, mostraba una expansión del territorio de expresión con respecto de sus ortólogos de chimpancé y ratón (Figura 5; Kamm, Pisciotano, et al. 2013b).

Por otro lado, de los 9 elementos capaces de dirigir la expresión de forma consistente en el ensayo de expresión en el pez cebra, solo 4 de ellos (HACNS96, HAR89, 2xHAR157, 2xHAR122) están conservados en todos los vertebrados incluido pez cebra (Tabla 3). Mientras que 2 de los elementos (HAR202, 2xHAR223) que no fueron capaces de activar la expresión de EGFP en este ensayo de líneas estables también están conservados en pez cebra. Por lo que no parecería haber una relación entre la conservación de estos elementos en peces cebras y su capacidad de dirigir la expresión en este modelo.

Queda entonces la duda de cuál sería la razón por la que no se pudieron obtener líneas estables con expresión consistente de elemento que en transitorio logran dirigir la expresión. Si bien se sabe que la expresión transitoria de un gen reportero en las

generaciones F0 inyectadas no siempre es recapitulada por las líneas estables F1 (Clark et al. 1994; Bonifer et al. 1996; Cranston et al. 2001), se supone que el método Tol2 utilizado debería reducir este problema al mínimo (Suster et al. 2009). En las líneas F0 se observa un efecto mosaico de la expresión, esto se debe a que el desarrollo del pez cebra es bastante rápido, pasando el cigoto de un estadio de una célula al siguiente estadio de 2 células en tan solo 45 minutos. Esto puede causar que la inserción de nuestra construcción no ocurra en este estadio de una célula, lo que generaría que no todas las células del futuro pez posean el transgén. Debido a este efecto mosaico, puede ocurrir que las líneas germinales no logren integrar efectivamente el transgén (Berghmans et al. 2008), por esta razón se crían suficientes F0, con el objetivo de tener al menos 3 líneas que sean capaces de pasar el transgén a la descendencia. Los problemas tales como los efectos de posición y el silenciamiento del transgén integrado al genoma de la misma manera que ocurren con los métodos de integración aleatoria de transgén en otros organismos modelo (Ni et al. 2008; Kawakami 2005; Balciunas et al. 2006). Es por esto que es necesario realizar el análisis de al menos tres líneas estables y la expresión en un tejido determinado sólo se reporta cuando es consistente en las tres líneas. Cuando existen discrepancias se deben analizar más líneas estables. En nuestro laboratorio y en otros laboratorios de nuestro instituto se ha observado que de manera frecuente existen discrepancias entre los resultados obtenidos en ensayos transitorios y estables. Este hecho no está profundamente discutido ni estudiado en la literatura, sino que se reduce a comunicaciones personales entre investigadores que utilizan este modelo. Nos proponemos en un futuro cercano realizar una investigación más profunda y sistemática acerca de esta discrepancia entre ensayos y trataremos de dilucidar las bases moleculares de la misma.

Comparación de la funcionalidad de las regiones aceleradas humanas con la de sus ortólogos de chimpancé

Las regiones HAEs, si bien presentan un mayor número de sustituciones en el linaje humano, en general corresponden a regiones altamente conservadas en vertebrados. Por lo que es lícito preguntarse si, por un lado, estas regiones ortólogas conservadas también poseen algún tipo de funcionalidad, y, por otro lado, si la funcionalidad es la misma que la observada en los elementos humanos. Para contestar este interrogante nos planteamos

evaluar la función cómo *enhancers* transcripcionales de las regiones ortólogas de nuestro pariente más cercano, el chimpancé común (*Pan troglodytes*).

Al igual que con las regiones humanas, se generaron también líneas estables de peces cebras transgénicos con las regiones ortólogas de chimpancé de los 14 NPAS3-HAEs usando el sistema de Tol2 y evaluando la expresión de la proteína reportera EGFP.

De las 14 regiones analizadas, 10 (HAR21Pt, HAR89Pt, HAR96Pt, HAR173Pt, HACNS221Pt, HACNS490Pt, 2xHAR122Pt, 2xHAR157Pt, HAR202Pt, HACNS658Pt) dirigieron la expresión de EGFP a distintas regiones anatómicas de los peces cebras. Se observó principalmente expresión en diferentes dominios de SNC en desarrollo (Figura 6; Tabla 4). De estas 10 regiones, los elementos HAR202Pt y HACNS658Pt mostraron una actividad *enhancer* diferencial respecto de sus secuencias ortólogas de humanos a 48 hpf. Los HAR202 de ambas especies ya habían sido analizados en nuestro laboratorio (Kamm et al 2013a), pero la diferencia entre los elementos del HACNS658 de ambas especies es una novedad.

Por otro lado, de las cuatro regiones que no generaron expresión en los peces (HACNS553Pt, HACNS96Pt, 2xHAR142Pt, 2xHAR223Pt), solo HACNS96Pt mostró diferencia con su ortólogo humano que sí dirigía la expresión de EGFP (Figura 6, tabla 4). Mientras que los otros 3 elementos (HACNS553Pt, 2xHAR142Pt, 2xHAR223Pt), al igual que los elementos humanos, no actúan como *enhancer* de la transcripción.

En resumen, nuestros resultados indican que 3 NPAS3-HAEs (HAR202, HACNS658 y HACNS96) son *enhancers* transcripcionales activos en el sistema nervioso en desarrollo del pez cebra y poseen funciones diferenciales entre las secuencias ortólogas de humano y de chimpancé (Figura 6). Si bien estos elementos experimentaron un aumento en su tasa de evolución molecular exclusivamente en el linaje humano comparado con la tasa de evolución más baja presente en el linaje de los chimpancés y otros primates (Pollard et al. 2006), es necesario analizar evidencia adicional proveniente de un tercer linaje cercanamente emparentado, para excluir la posibilidad de que el patrón de expresión observado para el elemento de chimpancé represente el estado derivado.

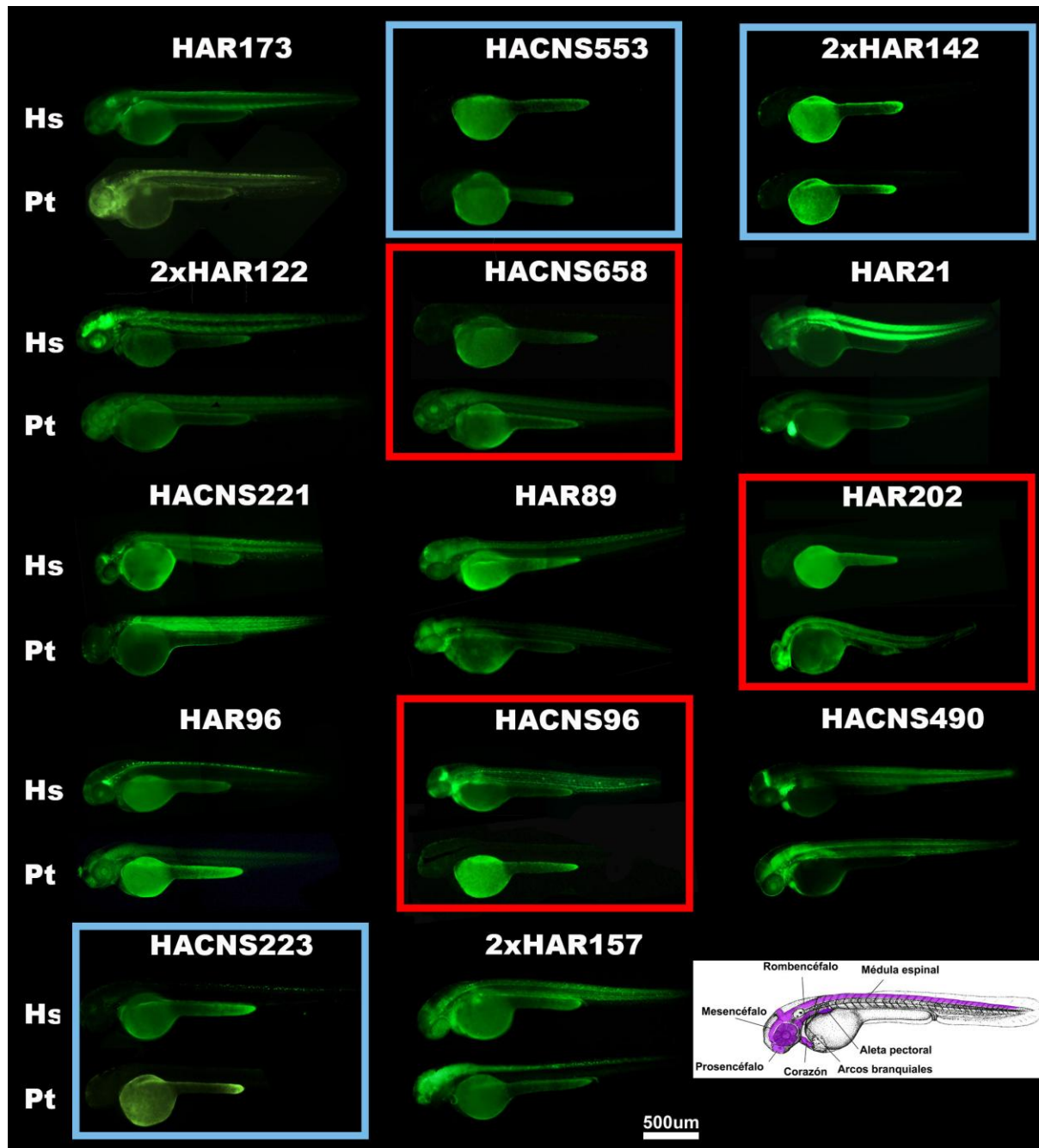


Figura 6. Expresiones comparativas de las líneas transgénicas estables representativas en peces cebras portadoras de secuencias de humano y de chimpancé de los 14 NPAS3-HAEs. El análisis se hizo sobre peces cebras a 48 hpf. Hs: *Homo sapiens*; Pt: *Pan troglodytes*. Recuadro rojo: líneas transgénicas estables que dieron diferencias humano-chimpancé. Recuadro azul: líneas transgénicas estables que no dieron expresión ninguno de los ortólogos. Esquina inferior derecha: esquema de pez cebras a 48hpf con las áreas anatómicas que usamos para analizar los patrones de expresión. Para comparar todas las líneas ver Figuras suplementarias 1, 2, 3, 4.

Tabla 4. Tabla con los resultados de todas las líneas transgénicas conteniendo los NPAS3-HAEs y sus diferentes ortólogos.

Nombre del transgén	# de líneas con expresión de EGFP	Dominios de expresión presentes en 2 o más líneas transgénicas a 48 hpf							
		P	M	R	ME	C	AF	AP	O
HAR96Hs	3/3	-	+	+	+	+	-	+	+
HAR96Pt	3/3	-	+	+	+	+	-	+	+
HACNS658Hs	0/2	-	-	-	-	-	-	-	-
HACNS658Pt	2/2	+	+	+	+	+	+	+	+
HACNS658Mmul	3/3	+	+	+	+	+	+	+	+
HACNS490Hs	2/4	+	+	+	+	+	+	+	+
HACNS490Pt	3/3	+	+	+	+	+	+	+	+
HACNS553Hs	0/4	-	-	-	-	-	-	-	-
HACNS553Pt	0/4	-	-	-	-	-	-	-	-
HAR21Hs	4/4	+	+	+	+	+	+	+	+
HAR21Pt	2/2	+	+	+	+	+	-	+	+
HACNS221Hs	2/2	+	+	+	+	-	+	-	-
HACNS221Pt	3/3	+	+	+	+	+	+	+	+
HAR173Hs	5/5	+	+	+	+	+	+	+	+
HAR173Pt	4/4	+	+	+	+	+	+	+	+
HACNS96Hs	3/3	-	+	+	+	+	+	-	-
HACNS96Pt	0/2	-	-	-	-	-	-	-	-
HACNS96Mmul	4/4	+	+	+	+	+	+	+	+
HAR202Hs	2/4	-	-	-	-	-	-	-	-
HAR202Hn	4/4	+	+	+	+	+	+	+	+
HAR202Pt	4/4	+	+	+	+	-	+	+	+
HAR202Mmul	3/3	+	+	+	+	-	+	+	+
HAR202Mm	3/3	+	+	+	+	-	+	+	+
HAR202Dr	6/6	+	+	+	+	+	+	+	+
2xHAR142Hs	0/4	-	-	-	-	-	-	-	-
2xHAR142Pt	0/4	-	-	-	-	-	-	-	-
HAR89Hs	2/2	+	+	+	+	+	+	+	+
HAR89Pt	3/3	+	+	+	+	+	+	+	+
2xHAR223Hs	1/4	-	-	-	-	-	-	-	-
2xHAR223Pt	0/4	-	-	-	-	-	-	-	-
2xHAR157Hs	3/3	+	+	+	+	+	+	+	+
2xHAR157Pt	3/3	+	+	+	+	+	+	+	+
2xHAR122Hs	3/3	+	+	+	+	+	+	+	+
2xHAR122Pt	3/3	+	+	+	+	+	+	+	+

Nota: En la segunda columna se indica el número de líneas estables que presentan expresión/las líneas que se analizaron. Las celdas en azul indican los resultados con la secuencia de chimpancé y

humanos de aquellos transgenes que no tuvieron expresión reproducible entre 2 o más líneas. Las celdas en rojo aquellas líneas que dieron expresión diferencial entre humano y chimpancé, donde las líneas humanas no presentaban expresión consistente. Las celdas en amarillo aquellas líneas que dieron expresión diferencial entre humano y chimpancé, donde las líneas chimpancé no presentaban expresión consistente. Referencias: P, prosencéfalo. M, mesencéfalo. R, rombencéfalo. ME, médula espinal. C, corazón. AB, arcos faríngeos. AP, aleta pectoral, O, ojo. Con (+) y (-) se indica presencia/ausencia de expresión de EGFP.

Comparación de la funcionalidad de las regiones aceleradas humanas con la de sus ortólogos del mono macaco Rhesus

Para resolver el interrogante surgido en el apartado anterior, procedimos entonces a clonar las regiones ortólogas de macaco Rhesus (*Macaca mulatta*) correspondientes a los elementos HAR202, HACNS658 y HANCS96 y a generar los peces cebras transgénicos para evaluar su actividad. El macaco Rhesus es una especie de primate catarrino de la familia Cercopithecidae cuyo linaje divergió del linaje que llevó a la aparición de las especies de chimpancé y humano hace 25 millones de años (Kumar et al. 1998).

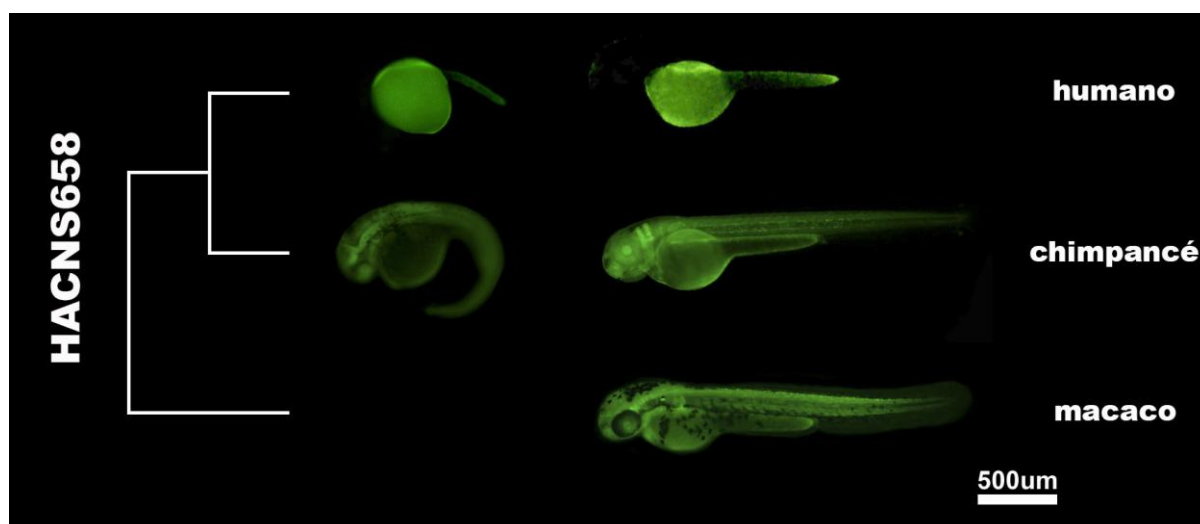


Figura 7. Expresión de EGFP de los distintos elementos ortólogos del HACNS658 a 24 (derecha) y 48 (izquierda) hpf. Se muestra una línea representativa por cada secuencia. Del elemento de macaco queda pendiente el análisis a 24 hpf. Para comparar todas las líneas ver Figura suplementaria 3.

Al evaluar la actividad como activador transcripcional de elemento del HACNS658Mmul observamos que su actividad es muy similar en cuanto al patrón de expresión con respecto a la versión de chimpancé (Figura 7). Dado que observamos la ausencia de expresión dirigida por la secuencia de humano, este resultado sugiere que la elevada tasa de sustituciones

nucleotídicas observada en la región ortóloga en humanos (Figura 8-A) afectaron de alguna manera su capacidad de actuar como activador de la transcripción; actividad que parecería estar conservada, al menos, en los linajes primates analizados.

En cuanto a las líneas de peces transgénicos portadoras del elemento HACNS96Mmul, también observamos actividad *enhancer* (Figura 9). En este caso se diferencia de su ortólogo de chimpancé (HACNS96Pt) el cual no dirige la expresión de EGFP en el pez cebra. Por otra parte, su actividad transcripcional es muy similar a la observada para el elemento humano a 24 hpf. A 48 hpf la actividad de la secuencia de macaco y de humano presentan diferencias en las regiones correspondientes al prosencéfalo y el ojo, donde no se observa actividad en las líneas humanas mientras que en las de macaco sí. Estos resultados sugieren que habría ocurrido un cambio de actividad espacio/temporal del elemento humano bastante específico y puntual, ya que mientras que es capaz de dirigir la expresión en el prosencéfalo a 24 hpf, a las 48 hpf esta expresión desaparece.

Observando las secuencias (Figura 8-B), vemos que entre humano y chimpancé hay 8 sustituciones nucleotídicas de diferencia (6 que son sustituciones específicas humanas y dos específicas de chimpancé). Ha sido determinado que la secuencia humana de este elemento habría sufrido un proceso de conversión génica por sesgo GC (gBGC, ver introducción general, Capra et al 2013); ya que en todas las sustituciones humano-específicas se reemplazaron bases débiles (A, T) a una base fuerte (G, C). Estos datos sugieren que el alto número de sustituciones del elemento HACNS96 en humanos, que ocurrieron dentro de una región altamente conservada en el linaje de los vertebrados fueron inducidas por un proceso evolutivo neutral del tipo gBGC y no habría sido bajo un proceso de selección positiva.

Por otra parte, no sucedió lo mismo con la secuencia ortóloga de chimpancé. Esta sufrió solo dos sustituciones comparada con macaco y humano (Figura 8-B), lo que revela una tasa de sustitución mucho más esperable para este tipo de regiones conservadas. Sin embargo, nuestros resultados con ensayos en peces cebra sugieren que estos cambios fueron suficientes para causar una pérdida de función de la región en chimpancé. Sin embargo, más ensayos para determinar si los cambios rompen sitios de unión de factores de transcripción serán necesarios para poder llegar a conclusiones más firmes.

C

HAR202Hs
HAR202Hn
HAR202Pt
HAR202Mmul
HAR202mm
HAR202Gg
HAR202Dr

10 20 30 40 50 60 70 80 90 100 110 120 130

HAR202Hs
HAR202Hn
HAR202Pt
HAR202Mmul
HAR202mm
HAR202Gg
HAR202Dr

140 150 160 170 180 190 200 210 220 230 240 250 260 270

HAR202Hs
HAR202Hn
HAR202Pt
HAR202Mmul
HAR202mm
HAR202Gg
HAR202Dr

280 290 300 310 320 330 340 350 360 370 380 390 400 410

HAR202Hs
HAR202Hn
HAR202Pt
HAR202Mmul
HAR202mm
HAR202Gg
HAR202Dr

420 430 440

HAR202Hs
HAR202Hn
HAR202Pt
HAR202Mmul
HAR202mm
HAR202Gg
HAR202Dr

Figura 8. (A) Alineamiento del elemento HACNS658 con las secuencias ortólogas de chimpancé y macaco. **(B)** Alineamiento del elemento HACNS96 con las secuencias ortólogas de chimpancé y macaco. **(C)** Alineamiento del elemento HAR202 con las secuencias ortólogas de humanos arcaicos, chimpancé, macaco, ratón, pollo y pez cebra. Las bases conservadas se muestran sombreadas en azul y las sustituciones e inserciones en escala de celeste. Los recuadros rojos muestran las sustituciones humano-específicas. Los recuadros verdes las sustituciones chimpancé específicas. Los recuadros amarillo las sustituciones macaco específica. El recuadro naranja muestra la base polimórfica en humanos modernos y la base presente en humanos arcaicos.

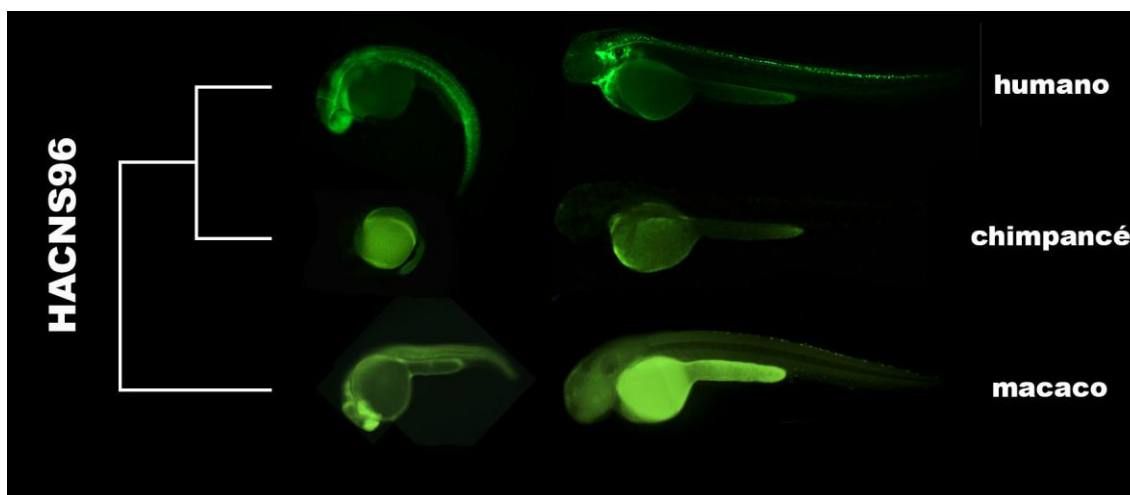


Figura 9. Expresión de EGFP de los distintos elementos ortólogos del HACNS96 a 24 (derecha) y 48 (izquierda) hpf. Se muestra una sola línea representativa por secuencia analizada. Para comparar todas las líneas ver Figura suplementaria 3.

Al evaluar la actividad del HAR202 de macaco (202Mmul) se ve que su actividad es igual a la de su ortólogo de chimpancé, el 202Pt (Figura 10), ya que dirige la expresión de EGFP a las mismas regiones anatómicas, particularmente al encéfalo y medula espinal.

Análisis funcional de distintas secuencias ortólogas de vertebrados del HAR202

El elemento HAR202, tal como se describió en la introducción, es un elemento que se encuentra altamente conservado en el linaje de los vertebrados, incluyendo a peces óseos como el pez cebra (Tabla 4 de la Introducción). Para analizar si distintas secuencias ortólogas de vertebrados modelo funcionan como *enhancer* transcripcionales clonamos los elementos ortólogos del HAR202 de ratón (HAR202mm), pollo (HAR202Gg) y pez cebra (HAR202Dr). Como se ve en la Figura 10,

la función directora de la transcripción del HAR202 se encuentra altamente conservada en los distintos linajes de vertebrados probados. Tanto los elemento ortólogos de ratón, como el de pollo y el de pez cebra tienen funcionalidad de *enhancer*. No solo eso, sino que, además, dirigen a las mismas regiones anatómicas en los dos estadios analizados (24 y 48 hpf). Es decir, que solo el elemento humano (HAR202Hs) es el único que posee una diferencia en su función.

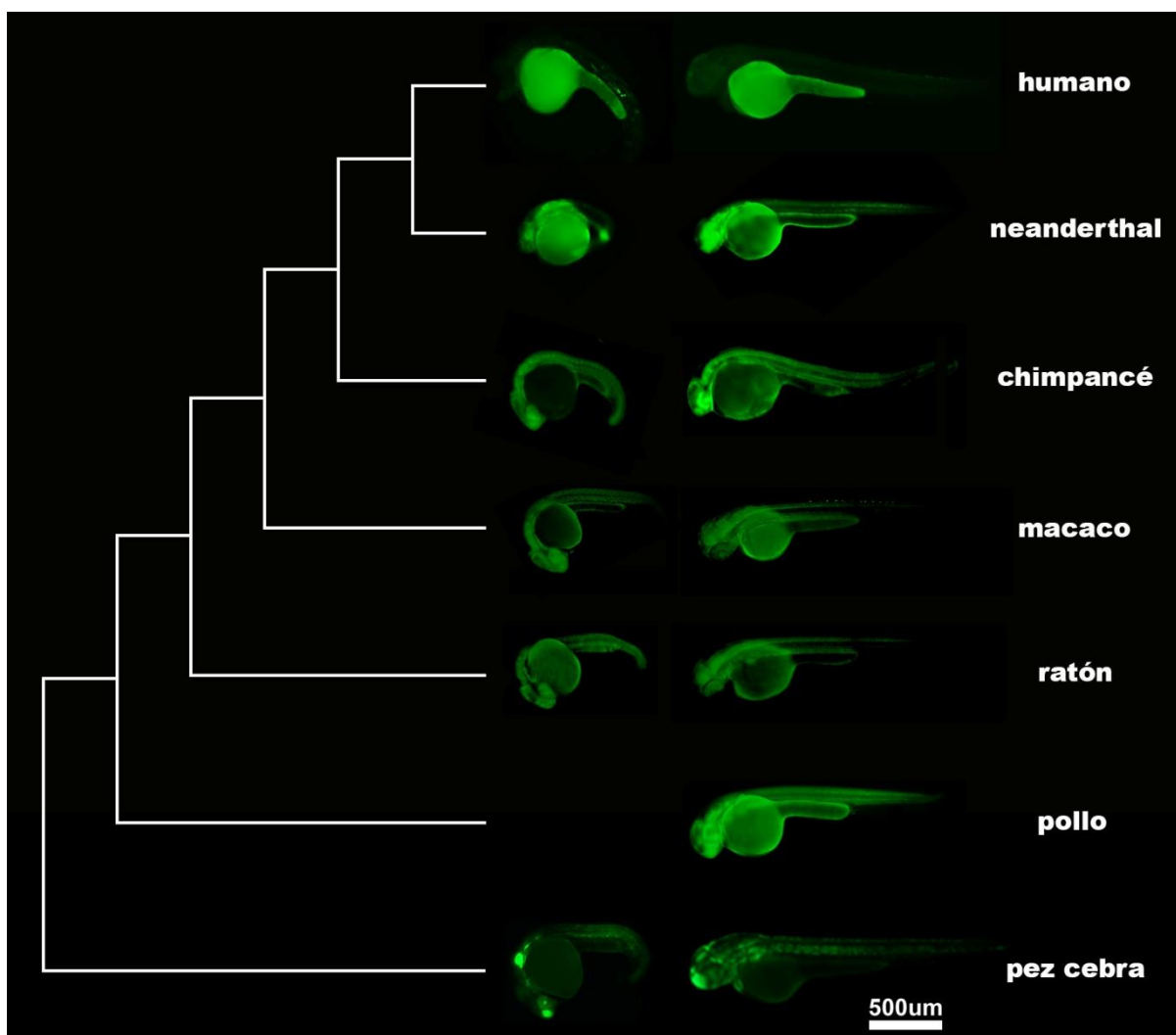


Figura 10. Expresión de EGFP de los distintos elementos ortólogos del HAR202 a 24 (derecha) y 48 (izquierda) hpf. Se muestra una sola línea representativa por secuencia analizada. Para comparar todas las líneas ver Figura suplementaria 4.

La secuencia del HAR202 humano presenta 6 de las 7 sustituciones nucleotídicas que son diferentes a la secuencia de chimpancé que son compartidas con humanos arcaicos. En efecto, la secuencia del HAR202 en humano arcaico (HAR202Hn) presenta una base de diferencia menos con el chimpancé respecto del humano (A/T en

humanos y T ancestral para el polimorfismo rs79570848, Fig. 8-C). Esta base de diferencia entre humano y neandertal es polimórfica en humanos, pero está fijada dentro de los individuos arcaicos secuenciados (Prüfer et al. 2014). Esta base polimórfica se encuentra, principalmente en personas de Asia de este (A=0.795, T=0.205, The genome Aggregation Database Lek et al. 2016); A=0.824, T=0.176, 1000 Genomes Project) y en menor medida en poblaciones americanas (A=0.94 T=0.06, The genome Aggregation Database; A=0.96 T=0.04, 1000 Genomes Project).

A diferencia de los otros elementos probados en el transcurso de este trabajo, el elemento humano arcaico no fue amplificado por PCR por carecer del genoma para usar como molde, por lo que fue sintetizado (GenScript, USA) previamente a ser clonado en la construcción con el promotor mínimo cFos y EGFP.

Otro resultado que se desprende de este experimento es que el elemento humano arcaico (HAR202Hn) sí es capaz de actuar como *enhancer*, lo que resulta más llamativo debido que entre el elemento humano moderno y el del arcaico solo hay una única base de diferencia. Futuros experimentos deberán determinar, cuál es el factor de transcripción que estaría actuando en este lugar, y si la sustitución causa que este factor se deje de pegar o que un factor distinto comience a hacerlo.

2DA PARTE: ESTUDIO DE LOS ELEMENTOS ACELERADOS DEL LOCUS DE *NPAS3* EN RATONES

Comportamiento del elemento acelerado no codificante HAR202Hs durante el desarrollo embrionario en ratones transgénicos

El gen *NPAS3* es un caso notable en la evolución humana. Por un lado, se encontró que este gen contiene el agrupamiento más grande de HAEs en el genoma humano (Kamm, López-Leal, et al. 2013a). Adicionalmente, se mostró que, al menos 9 de los 14 HAEs presentes en intrones de *NPAS3* son capaces de funcionar como *enhancers* en líneas transgénicas estables de pez cebra, particularmente durante el desarrollo del sistema nervioso. Sin embargo, la distancia evolutiva que existe entre el pez cebra y humanos (450 millones de años, Kumar y Hedges 1998) limita la extrapolación a *Homo sapiens* de los resultados obtenidos en el modelo vertebrado. Por esta razón, con el objetivo de explorar el impacto que los cambios filogenéticos experimentados en los *NPAS3*-HAEs pudieron tener sobre la evolución del cerebro humano, el siguiente paso fue estudiar en más detalle la función y origen evolutivo de uno de estos elementos en un modelo mamífero (particularmente, en ratón). De entre los 14 HAEs de *NPAS3*, se seleccionó el elemento HAR202 por lo visto en el apartado anterior. Este elemento comprende 444 pb localizadas en el 5to intrón de *NPAS3* (chr14:34,045,406-34,045,848, NCBI37/hg19; figura 11).

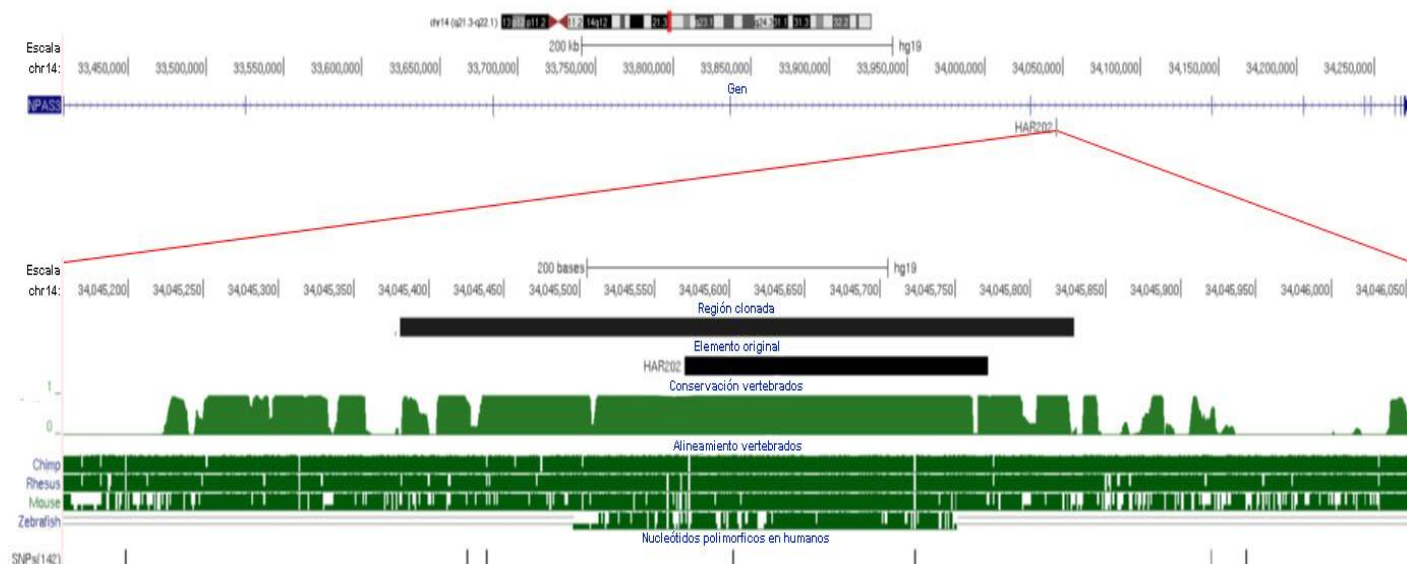


Figura 11. Diagrama y alineamiento de la secuencia del *enhancer* HAR202Hs. Esquema de la región clonada en el transgén HAR202Hs conteniendo la región conservada descrita por el programa PhastCons y el elemento HAR202. Conservación entre vertebrados de la secuencia del elemento HAR202Hs de acuerdo con el *USCS Genome Browser*. Con asteriscos verdes se indican las sustituciones humano-chimpancé.

Con el objetivo de estudiar si el elemento HAR202 tiene la capacidad de actuar como un *enhancer* activo durante el desarrollo embrionario en un modelo mamífero, se amplificaron por PCR 444 pb conteniendo completamente el elemento humano HAR202 (HAR202Hs) y se clonó río arriba del promotor mínimo del gen HSP68 de ratón regulando la expresión de la proteína reportera lacZ (Kothary et al. 1989). El promotor mínimo usando en estos ensayos, en ausencia de elementos regulatorios adicionales, es incapaz de promover la expresión del gen reportero durante el desarrollo embrionario de ratón (Kothary et al. 1989). Además, este sistema ha sido usado exitosamente para estudiar la actividad de *enhancers* clonados río arriba de su secuencia (Poulin et al. 2005; Kamm, López-Leal, et al. 2013b; L. F. Franchini et al. 2011).

Se generaron 4 líneas independientes de ratones transgénicos conteniendo esta construcción que se analizaron a día embrionario E11.5. Además, se estudió la expresión de la proteína reportera a tres edades distintas del desarrollo embrionario del ratón: 10.5, 11.5 y 12.5 dpc en una línea transgénica seleccionada (figura 12-A). A pesar de que se observó cierta variabilidad entre líneas (probablemente debido a diferencias en el número y la posición de la inserción del transgén en el genoma), se encontró que el elemento HAR202Hs tiene la capacidad de actuar como un *enhancer*, activando la expresión de lacZ de forma consistente (en tres de las líneas) en varios

dominios de expresión a E11.5. Particularmente, en 3 de las 4 líneas generadas se observó expresión consistente de la proteína reportera lacZ en diferentes lugares del sistema nervioso en desarrollo, incluyendo el cerebro posterior y la médula espinal, en las tres edades analizadas (Figura 12-A). En todas las edades analizadas se observó expresión de lacZ en el rombencéfalo y en médula espinal. Fuera del sistema nervioso, se observó expresión de lacZ en la zona mesenquimal de los primordios de los miembros, los ojos, entre otros (Figura 12-A). Estos resultados demuestran que el elemento HAR202Hs tiene la capacidad intrínseca de actuar como un *enhancer* activo durante el desarrollo embrionario.

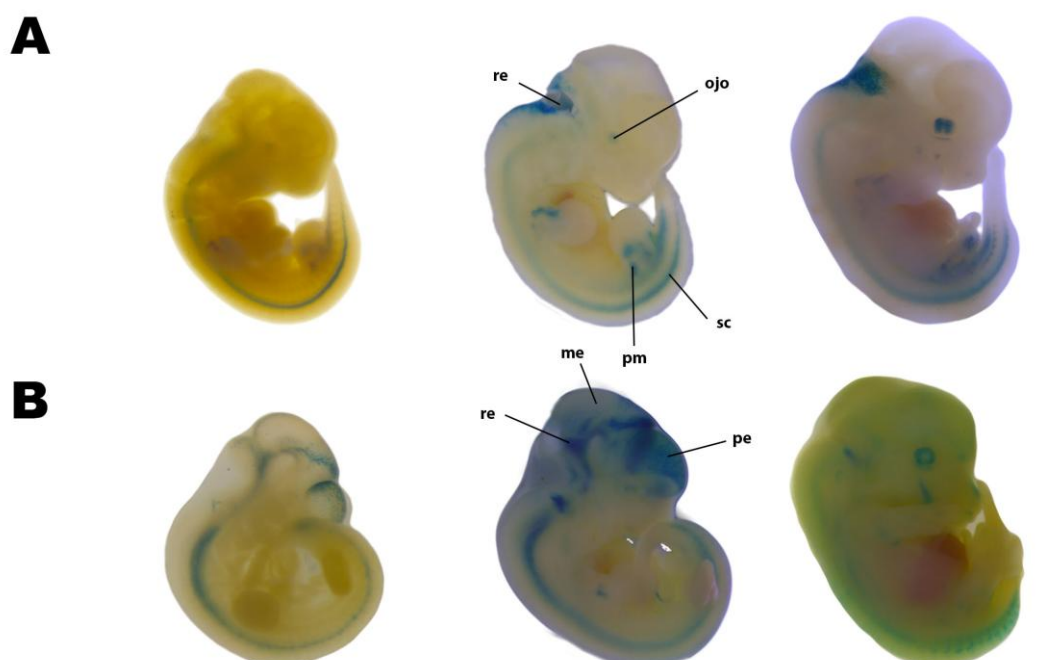


Figura 12. Líneas representativas de ratones transgénicos mostrando el patrón de expresión de lacZ. (A) Línea transgénica representativa de la secuencia HAR202Hs en los estadios embrionarios 10.5 dpc, 11.5 dpc y 12.5 dpc (B) Línea transgénica representativa de la secuencia HAR202Hn en embriones completos a 10.5 dpc, 11.5 dpc y 12.5 dpc. Referencias: (re) rombencéfalo, (pm) primordio mesenquimal del miembro, (me) mesencéfalo, (pe) prosencéfalo, (sc) médula espinal, (dpc) días post coito. La barra representa 300um.

También se clonó el elemento humano arcaico (que coincide con el polimorfismo presente en ciertas poblaciones humanas actuales, HAR202Hn) con la construcción HSP68/lacZ y se generaron en este caso 5 líneas transgénicas independientes y se estudió su expresión en las mismas edades que el HAR202Hs (Figura 12-B).

Sorprendentemente, a pesar de tener una única sustitución respecto al HAR202Hs, el patrón de expresión espacial de este elemento abarca otras regiones adicionales a las ya vistas. Particularmente se vio expresión en el mesencéfalo y en el prosencéfalo a 10.5, 11.5 y 12.5 dpc además de las expresiones ya vistas en el rombencéfalo, médula espinal, ojos, primordio de las extremidades y otros (Figura 12-B). Todas las líneas mostraron una expresión bastante consistente entre ellas. Solo una línea no mostró expresión pese a ser positiva por PCR (Figura 13-B).

Se puede discutir respecto a estas dos secuencias acerca de si la región que nosotros denominamos HAR202Hn, está representando una región ancestral dentro del género *Homo*. Y si la región HAR202Hs, con su única base diferente apareció juntamente con *H. sapiens* hace 200000 años. Lo que genera duda en este caso es el hecho de que entre poblaciones humanas actuales se encuentran personas que poseen esta región supuestamente ancestral. Como vimos, la presencia de este polimorfismo es bastante reducida (4%, The 1000 Genomes) y además ha sido observada principalmente? en poblaciones del Asia y americanas originarias (17% y 4 %, The 1000 Genomes). Podemos generar aquí entonces una hipótesis, que este polimorfismo haya sido reintroducida en *H. sapiens* a partir de hibridaciones con otras especies de *Homo* en algún momento de la historia en que convivieron (Liu et al. 2015). Actualmente, existe suficiente información que documenta la cruce entre neandertales y humanos anatómicamente modernos (HAM, (Sankararaman et al. 2014; Prüfer et al. 2014; Vernot y Akey 2014; Gunz et al. 2019) entre HAM y denisovanos (M. Meyer et al. 2012; Harmon 2012; Browning et al. 2018) y recientemente, de la cruce entre neandertales y denisovanos (Slon et al. 2018). Estos datos sustentan esta hipótesis. Es posible que el uso masivo de técnicas de secuenciación de genomas antiguos que está siendo cada vez más utilizada arrojen más datos que corroboren esta idea.

Hasta el momento, entonces, se pudo reproducir en ensayos con ratones transgénicos parcialmente el resultado observado en los ensayos de expresión de peces cebra. Si bien en el caso del HAR202Hn se reprodujo bastante bien lo observado en peces, en el caso de la secuencia HAR202Hs, no pudimos reproducir los resultados en ensayos de peces cebra. En efecto, mientras que en el pez cebra no observamos expresión consistente, en ratones en tres de las cuatro líneas analizadas pudimos documentar expresión consistente el gen reportero dirigido por la secuencia HAR202Hs. Pudimos

por otra parte, reproducir que ambos elementos dirigen la expresión de manera expresión diferencial. La única sustitución nucleotídica presente en el elemento humano moderno, si bien no modifica su función *enhancer*, es capaz de modificar las áreas en donde es reconocido por los factores de transcripción presentes en el ratón.

Comportamiento del elemento acelerado no codificante HAR202Pt durante el desarrollo embrionario en ratones transgénicos

Con el objetivo de determinar si la evolución acelerada en el linaje humano experimentada por el elemento HAR202 (HAR202Hs) pudo llevar a un cambio funcional en un modelo mamífero, se comparó el patrón de expresión producido por este *enhancer* con los generados por regiones genómicas que contienen a los elementos ortólogos en chimpancé. Para ello, se generaron y analizaron a 11.5 dpc 7 líneas independientes de ratones transgénicos conteniendo la región ortóloga de 444 pb del chimpancé (Figura 13-C). Este elemento, que lleva el elemento ortólogo en chimpancé del HAR202 (HAR202Pt) se clonó río arriba del promotor mínimo HSP68 de ratón, dirigiendo la expresión de la proteína reportera lacZ.

El análisis del patrón de expresión de lacZ del elemento HAR202Pt a 11.5 dpc se encontró que esta construcción produce un patrón de lacZ con varias similitudes con el patrón observado para el elemento HAR202Hs (figura 13-A), pero también se observaron diferencias entre las líneas. Por ejemplo, los dominios de expresión del elemento HAR202Pt incluyen la médula espinal y el rombencéfalo de manera similar a lo encontrado para el elemento humano (figura 13-C). Por fuera del SNC, se encontró que el elemento HAR202Pt dirige la expresión a los primordios de los miembros y a los ojos. En una línea en particular, además, se observó una fuerte expresión en el corazón, y en otra se ve expresión en los miembros, pero en el área apical del mismo adicionalmente a la región mesenquimal que se ve en las otras líneas.

Por otro lado, encontramos 2 líneas que no mostraban ninguna expresión del gen reportero, a pesar de tener el transgén en su genoma (determinado por PCR). Esta ausencia de expresión podría deberse a qué se insertaron en regiones silenciosas del genoma que inhiben su expresión.

Es interesante contrastar estos resultados con los obtenidos en el análisis realizado con peces cebra transgénicos (ver apartado anterior). Por un lado, observamos que el elemento humano es capaz de dirigir la expresión consistente del gen reportero en tres de las cuatro líneas transgénicas de ratón analizadas (Figura 13-A), mientras que en pez cebra las líneas analizadas no mostraron expresión consistente (Figura 10). Por otra parte, el ortólogo de chimpancé, que presentaba una fuerte expresión de EGFP en prosencéfalo, mesencéfalo, rombencéfalo y cuerda espinal en el pez (Figura 10), en el ratón sólo es capaz de dirigir la transcripción de la proteína lacZ en medula espinal (Figura 13-C). Esta diferencia en los resultados obtenidos podría explicarse por una parte por la distancia filogenética entre los modelos de vertebrados utilizados que presentan 450 millones de años de distancia evolutiva. Esto podría resultar en una diferencia en la expresión de los distintos factores de transcripción necesarios para activar la expresión del elemento HAR202 en cada sistema (Ariza-Cosano et al. 2012).

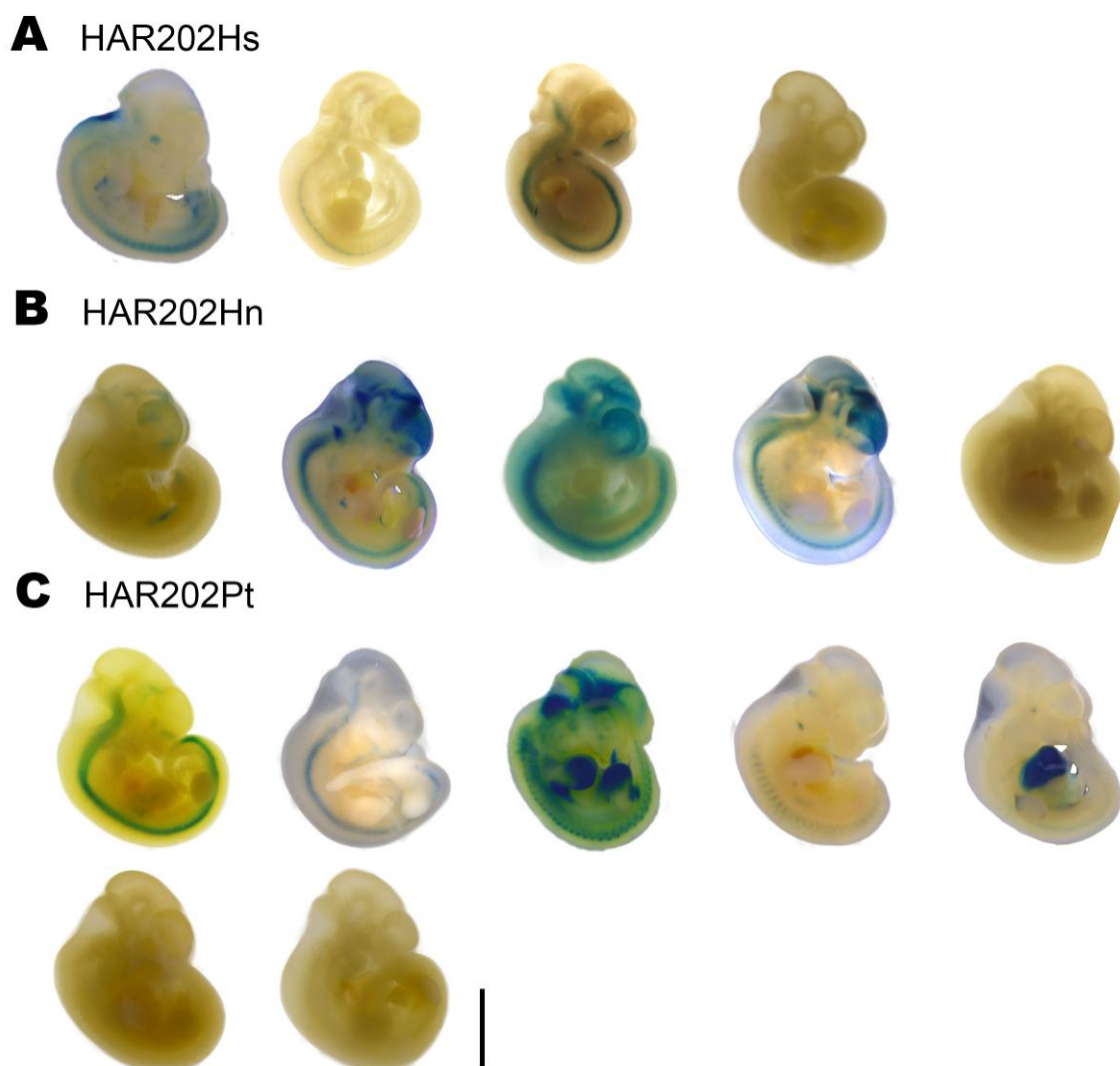


Figura 13. Análisis de expresión comparativo entre los elementos HAR202Hs, HAR202Pt y HAR202Mm en ratones transgénicos. Líneas independientes mostrando los patrones de expresión de lacZ de los elementos (A) HAR202Hs, (B) HAR202Pt y (C) HAR202Mm en embriones enteros de ratón a 11.5 dpc. Referencias: (re) rombencéfalo, (pm) primordio mesenquimal del miembro, (me) mesencéfalo, (pa) primordio apical del miembro, (sc) medula espinal, (gd) ganglios dorsales, (c) corazón, (af) arcos faríngeos. La barra representa 300um.

También es necesario agregar que las construcciones que fueron microinyectadas en los ratones difieren en su diseño respecto a las microinyectadas en peces cebra. No sólo en el gen reportero utilizado (lacZ en ratones, EGFP en peces), sino, y principalmente, en el promotor mínimo (Hsp68 en ratones, cFos en peces). El promotor mínimo del gen cFos de ratón consiste en 94 pb 5' del gen que incluye al TATA box (Dorsky, Sheldahl, y Moon 2002), y fue ampliamente utilizado para generar construcciones transgénicas para probar regiones regulatorias, ya que es incapaz de

dirigir la expresión por sí mismo (Kawakami 2005). El promotor del gen Hsp68 de ratón consiste en la región -664 a +113 del locus, también es ampliamente utilizado y se conoce que es incapaz de dirigir la expresión de lacZ sin una región regulatoria (Pennacchio et al. 2006). Si se sabe que este promotor es inducible por estrés con calor, pero en ningún momento del protocolo los ratones fueron sometidos a este estrés (Kothary et al. 1989). Quizás lo conveniente en este caso sea repetir los experimentos en ratón usando esta vez la misma construcción que usamos en peces cebra para descartar esta variable.

Estudiando la región conservada del *enhancer* HAR202Hs se encontró que tiene un 99,8% de identidad de secuencia con el elemento ortólogo HAR202Hn (una única sustitución específica en la región altamente conservada definida por el programa *PhastCons*, Figura 8-C), contra un 98,4% de identidad de secuencia con la región ortóloga presente en el elemento HAR202Pt (5 sustituciones específicas de los humanos, y 3 sustituciones específicas de chimpancé, figura 8-C) y 99,8% de identidad con la secuencia ortóloga HAR202Hn. Esta información sugiere que las sustituciones acumuladas en esta secuencia después de la separación de los linajes que dieron origen a los chimpancés y a los humanos podrían ser responsables de la adquisición de nuevos territorios de expresión por parte del elemento HAR202Hs, particularmente en la corteza en desarrollo. Analizando los cambios en los sitios predichos de unión para factores de transcripción (tabla 3), se observa que entre las secuencias ortólogas de chimpancé y humano del elemento HAR202 existen múltiples cambios (ya sea ganancias o pérdidas) en los sitios probables de unión para estas proteínas. Se requerirá trabajo adicional para estudiar cual/es de estos sitios (u otros aún no identificados) son las responsables del cambio heterotópico en el patrón de expresión de estos dos *enhancers*.

El dominio de asociación topológica TAD que contiene al locus del gen *NPAS3* contiene también a los 14 HAEs

Nos propusimos analizar si existen evidencias que sugiera que los NPAS3-HAEs estarían en efecto interactuando con el promotor de este gen y regulando su expresión. En nuestra búsqueda de posibles elementos reguladores de *NPAS3* que muestren una

evolución específica del ser humano, surge la duda acerca de si los HAEs localizados en los intrones del gen *NPAS3*, los cuales ya conocemos que son capaces de actuar como *enhancers* de la transcripción, son *enhancers* de este gen; es decir, si los *NPAS3*-HAEs que poseen función de *enhancer* dirigen la expresión de *NPAS3*. Se ha demostrado que el genoma está organizado en estructuras jerárquicas en 3D (Jin et al. 2013; Rao et al. 2014) que probablemente desempeñan un papel clave en la regulación de genes. Los TADs (*topologically associating domain*, por sus siglas en inglés) son estructuras cromosómicas en 3D que, se piensa, permiten acercar los elementos reguladores distales al promotor de cada gen al que regulan (Jin et al. 2013; Mifsud et al. 2015; Sanyal et al. 2012). Las técnicas basadas en la Captura de Conformación de Cromatina (3C) han comenzado a revelar detalles moleculares de la organización nuclear en células eucarióticas (Dekker et al. 2002). Hoy en día está claro que cada cromosoma ocupa un espacio separado en el núcleo en Interfase, conocido como "territorio cromosómico", que se divide en vecindarios o compartimentos distintos (Lieberman-Aiden et al. 2009). Dentro de cada compartimento, los dominios de asociación topológica (TAD) limitan las interacciones de la cromatina (Dixon et al. 2012). Dentro de cada TAD, las interacciones de la cromatina entre los elementos reguladores cis distales se producen de una manera dependiente del tipo de célula y tejido para permitir la modulación de la actividad del promotor por los *enhancers* (Dryden et al. 2014).

Para obtener evidencia acerca de la interacción de los *NPAS3*-HAEs y el promotor de *NPAS3* analizamos la región genómica donde se encuentra ubicado el gen de *NPAS3* utilizando datos de experimentos de Hi-C de distintos órganos y tejidos y de distintos estadios del desarrollo. La técnica de Hi-C permite cuantificar las interacciones entre todos los pares de fragmentos posibles simultáneamente en un ensayo de captura de la cromatina, extendiendo los resultados obtenidos con la técnica de 3C dónde se usa la PCR para cuantificar las interacciones solo entre dos fragmentos específicos. La secuenciación del material producido por ensayos de Hi-C produce mapas de interacciones de todo el genoma. Estos mapas son los que definen los límites de los TADs

Para determinar el TAD que incluye al gen *NPAS3* utilizamos datos generados a partir de Hi-C de tejidos (adulto: hipocampo, hígado, bazo, pulmón, ovario, corteza

prefrontal, Schmitt et al. 2016; fetal: corteza en desarrollo, Won et al. 2016 y células (linfocitos B, progenitores neuronales, células madre embrionarias, Schmitt et al. 2016) humanos.

Encontramos que el TAD que incluye al gen *NPAS3* se extiende por casi dos megabases en el cromosoma 14 del genoma humano (Hg19; chr14:33040001-35000000). Al cruzar esta información con nuestra base de datos de HAEs, encontramos que hay 16 HAEs en el *NPAS3*-TAD, de los cuales 14 corresponden a nuestros HAEs estudiados (Fig. 14). Los otros 2 HAEs, el HAR144 y el HACNS780, se ubican en regiones intergénicas. Como se ha demostrado que la mayoría de las interacciones ocurren dentro del mismo TAD (Dixon et al 2013), es probable que estos HAEs interactúen con uno de los sitios de inicio de la transcripción *NPAS3* (TSS) para regular su expresión (Fig. 14). Sin embargo, también observamos que hay otros genes ubicados en el mismo TAD que *NPAS3* (*AKAP6*, parcialmente y *EGLN3*). Es de notar que estos genes poseen una sintenia sumamente conservada, pudiendo rastrear su origen hasta peces cartilaginosos. En el caso de *AKAP6*, su TSS se ubica por fuera del *NPAS3*-TAD, por lo que podríamos descartar una posible regulación por parte de los *NPAS3*-HAEs como *enhancers* del gen *AKAP6*. Por otra parte, el TSS del gen *EGNL3* se ubica dentro del TAD y no podemos descartar que los *NPAS3*-HAEs regulen al gen *EGNL3*.

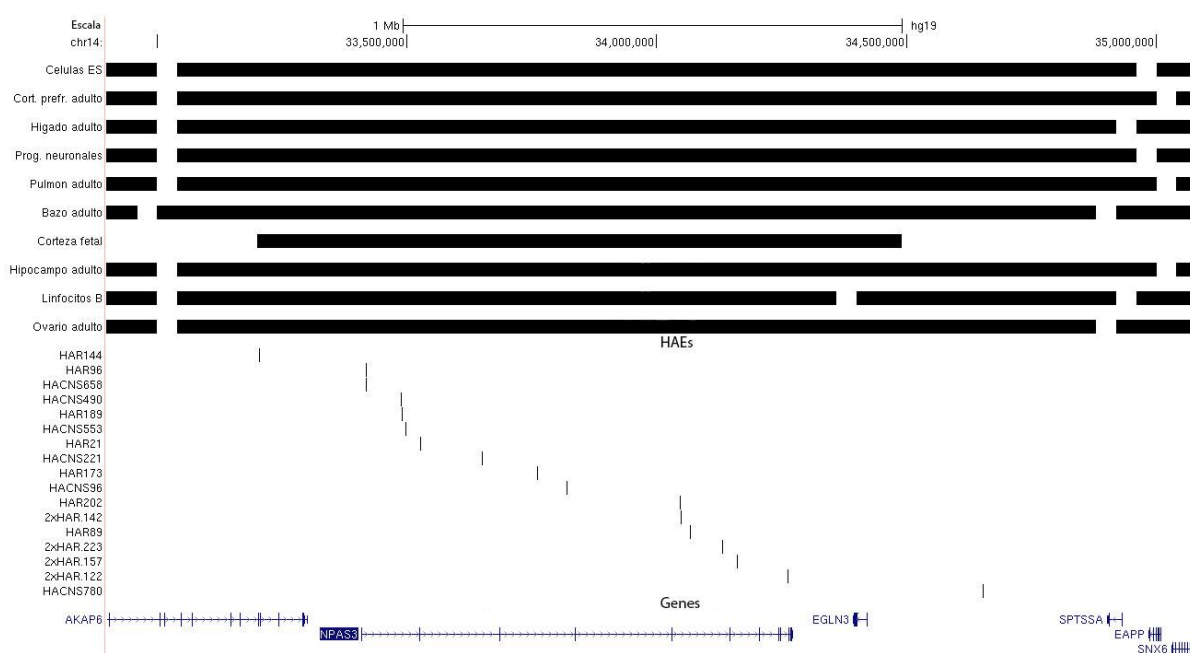


Figura 14. *NPAS3*-TAD contiene 16 HAEs. Diagrama que muestra el locus de *AKAP6*, *NPAS3* y *EGLN3* en el cromosoma 14 (14q13.1) del genoma humano (hg19; modificado después de la

captura desde el navegador del genoma UCSC). Se observa la ubicación de los 14 HAEs intrónicos estudiados y los 2 intergénicos. Toda la región está ubicada dentro del mismo dominio asociado topológicamente, según lo determinado mediante los datos de conformación cromosómica de células madre embrionarias (ES), progenitores neuronales, linfocitos; hipocampo, hígado, pulmón, baso, ovario, corteza prefrontal adultos (Schmitt et al 2016); corteza en desarrollo fetal (Won et al. 2016).

Estudio de marcas epigenéticas en el locus de NPAS3

Mucho de lo visto hasta este punto muestra que varios de los *NPAS3*-HAEs funcionan como *enhancer* en los distintos ensayos de expresión probados. Por lo que, lo siguiente que quisimos ver, fue si estos elementos regulatorios se comportan de igual manera en un sistema de expresión endógena juntamente con todo su entorno. Para ello nos enfocamos en la tarea de buscar distintas marcas epigenéticas relacionadas con funciones regulatorias que podrían, de esta manera, darnos una idea de cómo se comportan estos elementos.

Para eso buscamos en bases de datos de marcas epigenéticas disponibles, información acerca de marcas compatibles con la función en la regulación transcripcional para los 14 elementos *NPAS3*-HAEs. Utilizamos información de bases de datos producidas a partir de chip-seqs de Histona 3 trimetilada en la lisina 4 (H3K4me3, señal de cercanía de regiones promotoras), Histona 3 acetilada en la lisina 27 (H3K27ac, señal de cercanía de *enhancers*), la exonucleasa DNase I (señal de proteínas unidas al ADN), el represor transcripcional CTCF (señal de sitio de unión al ADN de la proteína *insulator* CTCF) y la proteína p300 (proteína que se une a *sitios enhancers*, Tabla 5, Visel, Rubin, y Pennacchio 2009; Nord et al. 2013).

Tabla 5. Señales de marcas epigénéticas encontradas en distintas bases de datos.

	Tipo de marca	Tipo de células o tejido	Referencia
HAR96	DNaseI	célulasH7-hESC	Encode
HACNS658	DNaseI	célulasH7-hESC	Encode
HACNS490	H3K27ac	mesencéfalo embrión ratón	Encode
	DNaseI	prosencéfalo embrión ratón, retina embrión ratón	Encode
HACNS553	H3K4me3	neuroesferas derivadas de eminencia ganglionar fetal humana	Encode

	H3K27ac	músculo fetal humano, cerebro humano adulto, fibroblastos, prosencéfalo adulto ratón	Encode, Nord et al, 2013
	DNaseI	músculo fetal humano, mesencéfalo humano adulto, tallo cerebral humano adulto, fibroblastos, timo fetal humano, médula espinal fetal humana, astrocitos, cerebro fetal humanos, miembros fetales humano, prominencia facial embrionaria humano, fotorreceptores retina humano, pulmón fetal humano	Encode
HAR21			Encode
HACNS221			Encode
HAR173	DNaseI	cerebro fetal humano (56-58,72 y 76 días), mesencéfalo adulto humano, miembros feto humano	Encode
	H3K4me3	cerebro fetal humano (semana 20)	Encode
	H3K27ac	células gliales radiales humanas, prosencéfalo embrión ratón (E11,5-E15,5)	Encode, Nord et al, 2013
HACNS96	DNaseI	cerebro fetal humano, riñón fetal humano, medula espinal fetal humano, retina fetal humano, primordio miembro embrión ratón (E11,5), retina embrión ratón (E14,5), prosencéfalo ratón embrión (E14,5), tubo neural ratón embrión (E11,5), prominencia facial embrión (E11,5)	Encode
	p300	Corteza cerebral fetal humana (semana 20)	Visel et al, 2014
	H3K27ac	miembro embrión ratón (E11,5-E13,5), rombencéfalo ratón (P0), prosencéfalo ratón adulto (P7, P21, P56)	Encode, Nord et al, 2013
HAR202	DNaseI	células ES humanas, intestino ratón adulto, miembro embrión ratón (E14,5), primordio miembro embrión ratón (E11,5)	Encode
	CTCF	células K562 (linfoma)	Encode
	p300	Corteza cerebral fetal humana (semana 20)	Visel et al, 2014
2xHAR142			Encode
	H3K4me3	BE2C (neuroblastoma)	Encode
	H3K27ac	células ES humanas, células iPS	Encode
HAR89	DNaseI	corteza cerebelar humano adulto (78 y 84 años), cerebro fetal humano (85, 96, 101, 104, 105, 109, 112, 117, 12 dpc), medula espinal fetal humana (59, 72, 87, 113dpc)	Encode
2xHAR223	DNaseI	cerebro fetal humano (56-58, 72, 76 dpc)	Encode
2xHAR157			Encode
2xHAR122	H3K4me3	Humano: hipocampo adulto (73 años), aorta torácica adulto (37-54 años), sustancia nigra adulto (81 años), ovario adulto (30 años), lóbulo temporal adulto (75 años), gastrocnemio medial adulto (54 años), esfínter gastroesofágico adulto (53 años), núcleo caudado adulto (81 años), arteria coronaria adulta (53 años), fibroblasto de pulmón, cingulate gyrus adulta (75 años), angular gyrus adulta (75 años)	Encode
	H3K27ac	Humano: hipocampo adulto (73 años), sustancia negra adulto (81 años), núcleo caudado adulto (81 años), lóbulo temporal adulto (75 años), cingulada gyrus adulta (75 años), área frontal media (46 años y 81 años)	Encode

DNaseI	Humano: medula espinal fetal (59, 72, 87,89, 96, 105, 106, 113 días), pulmón fetal (54 días y 58 días), riñón fetal (85 días), médula oblongata adulta (78 años y 84 años), músculo fetal (113 días), riñón fetal (115 días), timo fetal (97 días), cerebro medio adulto (78 años y 84 años), testículo fetal, ovario, cerebro fetal (56-58 días), mesodermo cardíaco fetal, piel del cuerpo fetal (82 días), retina fetal (74 días y 85 días), estómago adulto (96 días), estómago fetal (96 días), extremidad fetal (56 días), bazo fetal (112 días), pons adulto (78 años), células madre neuronales, glándula suprarrenal fetal (85 días), natural killers de células adulto (34 años), gyrus temporal superior adulto (84 años), intestino grueso fetal (103 días), T-helper adulta (33 años), B célula adulta (43 años). Ratón: miembros embrión (E11,5, E14,5), mesencéfalo embrión (E14,5), miembro inferior embrión(E11,5), retina embrión (E14,5), cerebro embrión (E18,5), prominencia facial embrionaria (E11.5), rombencéfalo embrión (E11.5), primordio miembro anterior embrión (11.5 días), mesodermo embrión (11.5 días), cerebro medio embrión (11.5 días), tubo neural embrión (11.5 días)), cerebro posterior embrión(14,5 días), cerebro anterior posnatal (0 días), retina postnatal (7 días), cerebro anterior embrión (14,5 días), cerebro medio posnatal (0 días), timo postnatal (0 días), adipocito, pulmón embrión (14,5 días), células de Müller posnatales (12 días), corazón postnatal (0 días), estómago postnatal (0 días), pulmón postnatal (0 días), retina adulto (8 semanas), prominencia facial embrionaria (14,5 días), fibroblasto (8 semanas)	Encode
CTCF	Humano: Células epiteliales del plexo coroideo, células B adulto, fibroblastos de encía adultos (24 años), fibroblastos prepucios postnatales, monocitos, células epiteliales del esófago, células epiteliales mamarias, fibroblastos de arteria pulmonar, fibroblastos de piel de adultos (22 años), fibroblasto de pulmón fetal (12 semanas), fibroblasto de pulmón, células epiteliales pigmentarias retina, fibroblasto de piel de la pierna fetal (12 semanas), células epiteliales del túbulo proximal, fibroblasto de pulmón adulto (45 años), fibroblasto cardíaco, músculo esquelético adulto (22 años), astrocito de la médula espinal, fibroblasto de la dermis, célula muscular cardíaca, osteoblasto, fibroblasto de glándula mamaria, astrocito, fibroblasto de pulmón (11 años y 45 años), queratinocito, fibroblasto cutáneo, células epiteliales renales, células epiteliales bronquiales, células epiteliales mamarias (50 años), astrocito del cerebelo, bazo adulto, riñón adulto (22 años, 27 años y 35 años), páncreas adulto. Ratón: rombencéfalo posnatal (0 días), hígado embrión (14,5 días), riñón postnatal (0 días), pulmón postnatal (0 días), mesencéfalo posnatal (0 días), intestino postnatal (0 días), hígado postnatal (0 días), pulmón embrión (14,5 días), corazón adulto (8 semanas), estómago postnatal (0 días), corazón postnatal (0 días), intestino delgado (8 semanas), hígado (8 semanas), cerebro anterior postnatal (0 días)), fibroblasto embrionario (13,5 semanas), cerebelo (8 semanas), bulbo olfativo (8 semanas), riñón (8 semanas), médula ósea (8 semanas), testículo (8 semanas), pulmón (8 semanas), placa cortical (8 semanas), timo (8 semanas), macrófago de médula ósea (8 semanas)	Encode

De los distintos elementos acelerados, sobresalen por la cantidad y el tipo de marcas el HANCS96, el HANCS553, el HAR202, el HAR89 y el 2xHAR122. En efecto, el elemento HACNS96 presenta marcas que correlacionan con actividad *enhancer* como H3K27ac en células gliales radiales humanas, prosencéfalo del embrión de ratón (E11,5-E15,5), o p300 en corteza cerebral fetal humana (semana 20) (Tabla 5). Esto parecería indicar que tanto en ratón como en humano esta región funcionaría como *enhancer*. Estos datos sumados a la evidencia de nuestro laboratorio que indica que este elemento habría ganado función *enhancer* en el linaje humano lo convierte en un candidato ideal para un estudio detallado a través de la generación de ratones transgénicos y mutantes.

Del mismo modo las marcas epigenéticas en el elemento HAR202 (H3K27ac: miembro de embrión de ratón (E11,5-E13,5), rombencéfalo de ratón (P0), prosencéfalo de ratón adulto (P7, P21, P56), Tabla 5) apoyan nuestros resultados que indican que este elemento se comporta como un fuerte activador. Sin embargo, experimentos de p300 indican que este elemento uno este factor en la corteza humana de que estadio (Tabla 5). Nuestros experimentos no muestran que este elemento sea activo en este tejido durante el desarrollo embrionario, por lo menos el polimorfismo mayoritario (HAR202Hs). Mas experimentos de este tipo serán necesarios para estudiar esta discrepancia.

Generación de ratones mutantes portadores de una delección del elemento HAR202

Considerando los resultados de los ensayos de expresión obtenidos en peces cebrá y en ratones transgénicos portadores de las secuencias de HAR202, nos planteamos investigar la función del HAR202 endógeno actuando en su propia localización en el genoma en la transcripción del gen *NPAS3*. Nuestros ensayos en peces cebrá mostraron que el elemento HAR202 ortólogo en ratón (HAR202mm, Figura 10) actúa como *enhancer* de la transcripción. Por lo que decidimos generar un ratón mutante para la región conservada que contiene al ortólogo del HAR202 de ratón. Para ello se hizo uso de la técnica de edición génica CRISPR/Cas9. Se diseñaron y generaron guías de ARN dirigidas a los extremos de esta región conservada (Figura 15).

Una solución conteniendo las guías diseñadas juntamente con el ARNm de la enzima Cas9 fue microinyectada en cigotos de ratón y posteriormente transferidos al útero de hembras pseudopreñadas donde se desarrollaron hasta el nacimiento. Los animales fueron genotipificados y observamos que hubo generación de ratones mutantes portadores de la delección (Figura 15).

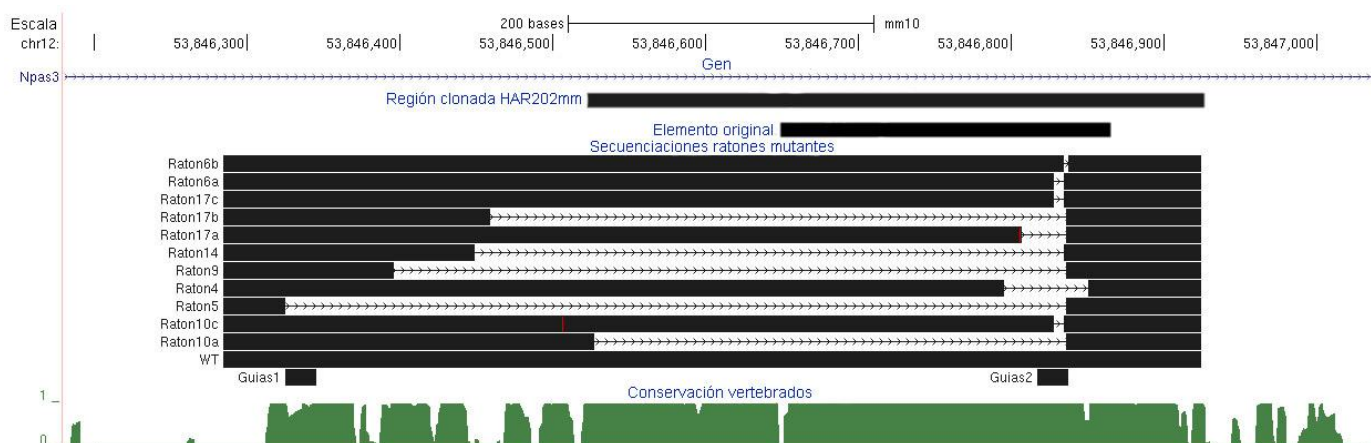


Figura 15. Región del gen de Npas3 de ratón que fue eliminada. Se ven la región clonada HAR202mm, la región acelerada en humanos del HAR202 tal como fuera identificado originalmente (Pollard et al 2006), las guías utilizadas (Guia1 y Guia2) y los resultados de las secuenciaciones de los ratones microinyectados nacidos comparadas con la secuenciación de un ratón de tipo salvaje (WT).

De 17 ratones nacidos, 7 poseían algún tipo de mutación en su genoma en la región esperada. Puede verse que las diferentes deleciones no fueron del mismo tamaño (Figura 15). Eso puede deberse a que no todas las guías son iguales de eficientes a la hora de señalarle a la enzima CAS9 donde debe de cortar. En este caso se ve que a Guía 2, fue mucho más eficaz que la 1.

Otra particularidad de este resultado es que se ven más de un tipo de deleción para ciertos ratones. Los ratones 6 y 10 presentan ambos dos tipos de deleciones diferentes. El ratón 17, sorprendentemente, presenta 3, indicando que posee 3 alelos distintos. La explicación que le damos a esto es que la enzima Cas9 debió de haber cortado más de una vez y en distintos estadios embrionario; es decir en un estadio de una célula y, al menos, en un estadio de dos células causando que el animal sea un mosaico, que presenta células con distintos genotipos (efecto mosaico).

De los 7 ratones mutantes sólo el ratón 5 tuvo la deleción esperada para las guías diseñadas. Por esta razón conservamos a este ratón para generar nuestra colonia.

Este ratón (macho) no tuvo problemas evidentes para alcanzar la madurez sexual e, incluso reproducirse. Para evitar posibles problemas de efecto mosaico (las secuenciaciones se realizan a partir de ADN extraído de muestras de tejido de la oreja del ratón, lo que nada nos dice del genotipo de las células germinales), se cruzó con un ratón hembra de tipo salvaje o *wild type* y se conservaron los ratones F1 que poseían la deleción esperada. A su vez, estos ratones F1 heterocigotas para la deleción fueron

cruzados entre ellos y se observó que también se pudieron reproducir sin problemas y criar a su prole (F2) hasta la edad del destete. Las camadas F2 contienen generalmente en la misma camada ratones de tipo *wild type* (*NPAS3*^{+/+}), heterocigotas (*NPAS3* ^{Δ HAR202/+}) y mutantes (*NPAS3* ^{Δ HAR202/ Δ HAR202}) para la región ortóloga en ratón del HAR202 y son los utilizados para la realización de los experimentos funcionales. Observamos que, en su aspecto exterior los ratones portadores de la mutación son ratones normales que no presentan ninguna particularidad a simple vista, son capaces de reproducirse y de cuidar a sus crías. Por lo que, al menos a grandes rasgos, la falta de este elemento, que sabemos que es regulador de la transcripción por los ensayos de peces cebra, no parecería estar afectando la vida normal de estos ratones. Estos ratones están siendo investigados a partir de ensayos de hibridación in situ, PCR cuantitativas, inmunohistoquímica, etc. para explorar la función del HAR202 en la expresión del gen *NPAS3*. Estos ensayos son tendientes a detectar algún tipo de fluctuación en los niveles y áreas de expresión de *NPAS3* en los ratones portadores de la mutación en relación con ratones de tipo salvaje.

Capítulo II

INTRODUCCIÓN ESPECÍFICA

***Trachealess*, el gen ortólogo de *NPAS3* en mosca**

Los genes parálogos en vertebrados *NPAS3* y *NPAS1* serían aparentemente ortólogos del gen *Trachealess* (*Trh*) de *Drosophila melanogaster*, compartiendo 45% y 42% de identidad de secuencia aminoacídica con *Trh*, respectivamente (Brunskill et al. 1999; Pieper et al. 2005) (figura 1, A y B). A diferencia de los genes *NPAS3* y *NPAS1*, se conoce relativamente bien la función de *Trh*. Este gen, que abarca aproximadamente 30 kpb del cromosoma 3L de *D. melanogaster* (Marygold et al. 2012), es un factor de transcripción de la familia bHLH-PAS esencial para el desarrollo del sistema respiratorio de *D. melanogaster* (Wilk, Weizman, y Shilo 1996; Chung, Chavez y Andrew 2011). Este sistema está compuesto por un conjunto de tubos interconectados (o tráqueas) que surgen desde las placas traqueales, un grupo de células ectodérmicas que se forman a cada lado de diez segmentos embrionarios (abarcando desde el segundo segmento torácico hasta el octavo segmento abdominal). Las células de cada placa se invaginan y dan lugar a la formación del árbol traqueal. La estructura básica se asemeja a la de los pulmones, con tubos formados por células epiteliales que se ramifican por los cuales se transportan oxígeno directamente a los órganos internos (Chung, Chavez y Andrew 2011).

En embriones de mosca, la expresión de *Trh* se puede detectar inicialmente en los núcleos de las células que darán origen a las tráqueas, luego en las células en migración que formarán los tubos del sistema traqueal y finalmente en las tráqueas ya formadas (estadios embrionarios 11-15). Adicionalmente, se observa expresión en las células de los ductos de las glándulas salivares (estadio 13) (Wilk, Weizman, y Shilo 1996; Chung, Chavez, y Andrew 2011). Estudios realizados en embriones de mosca mutantes mostraron que, en ausencia de *Trh*, no se forman las placas de la tráquea y no ocurre la invaginación desde los primordios de las tráqueas conduciendo a una completa ausencia del sistema respiratorio de la mosca (Wilk, Weizman y Shilo 1996). Por otro lado, la expresión ectópica de *Trh* tiene la capacidad de inducir la formación de placas traqueales ectópicas. El inicio de la expresión de *Trh* depende de estímulos

externos, pero una vez iniciada su expresión, entra en un circuito de autorregulación que garantiza el mantenimiento de los niveles de expresión de este factor de transcripción (Wilk, Weizman, y Shilo 1996). Adicionalmente, *Trh* regula la expresión de todos los genes conocidos que intervienen en la formación de tráqueas en la mosca (entre ellos el gen ortólogo en mosca de los *Fibroblast growth factor receptors* de vertebrados, *Breathless*) (Chung, Chavez, y Andrew 2011). Por otro lado, entre los estadios embrionarios 13-16 comienza la expresión de *Trh* en sistema nervioso central. En particular, se detecta expresión del ARN mensajero en el cordón nervioso ventral y en el cerebro de las moscas (Tomancak et al. 2002, 2007). Sin embargo, se desconoce la función de este gen a nivel del sistema nervioso en desarrollo de la mosca.

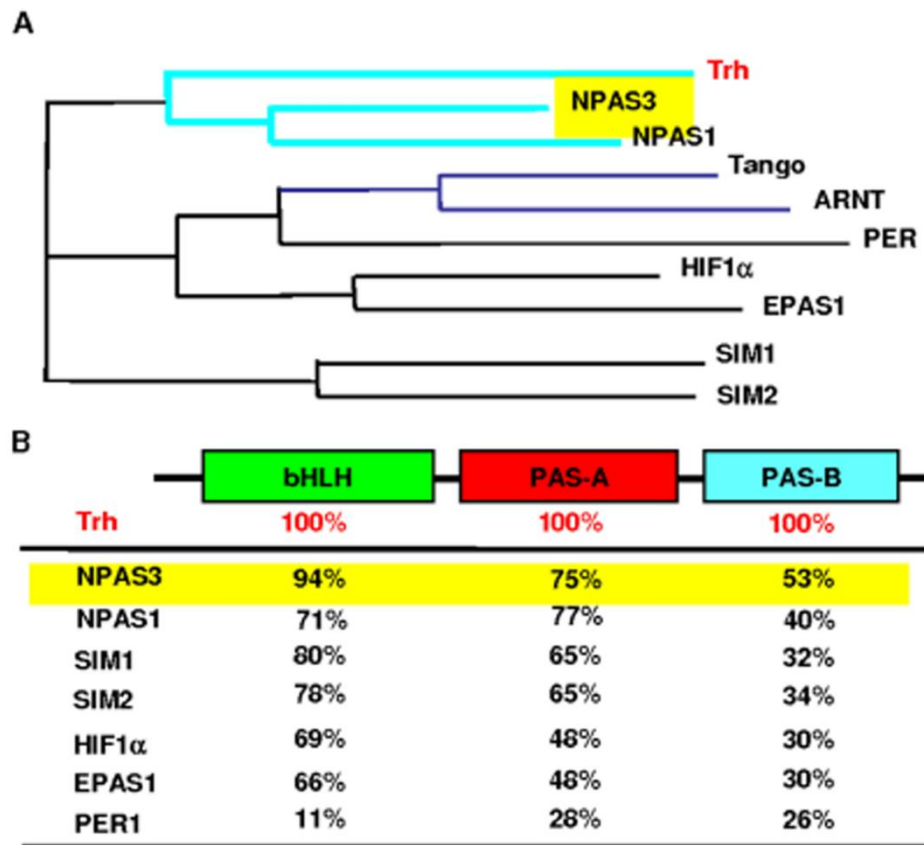


Figura 1. NPAS3 y NPAS1, los homólogos murinos del gen de *Drosophila* *Trh*. (A) Árbol filogenético de los genes de la familia bHLH-PAS: los genes humanos NPAS3 y NPAS1 están más relacionados con el gen *Trh* de *D. melanogaster* mientras que la proteína ARNT está más relacionado con el gen *Tango* de la mosca de la fruta. Adicionalmente, se muestran las relaciones con otras proteínas de la familia tales como PER, HIF1α, EPAS1 y SIM1/2. (B) Porcentajes de identidad de secuencia entre los miembros de la familia bHLH-PAS: NPAS3 comparte mayor similitud de secuencia aminoacídica con *Trh* en los dominios bHLH, PAS-A y PAS-B. Modificado de (Zhou et al. 2009).

Patrón de expresión de *NPAS3* en el sistema nervioso de mamíferos

En mamíferos *NPAS3* se expresa abundantemente en el sistema nervioso en desarrollo tanto en ratones como en humanos (Brunskill et al. 1999; Gould y Kamnasaran 2011), donde se observan altos niveles de expresión en el tubo neural, neuroepitelio y, en etapas posteriores, en la capa neopallial de la corteza. Fuera del sistema nervioso, se observa expresión de *NPAS3* en pulmones, corazón, miembros, piel y riñones (Brunskill et al. 1999; Zhou et al. 2009). Ya en adultos, se observa expresión de *NPAS3* en los núcleos de pequeñas interneuronas inhibitorias uniformemente distribuidas a lo largo de todo el sistema nervioso central. Resultados obtenidos de combinar la detección de *NPAS3* con la de distintos marcadores en el sistema nervioso de ratón muestran que *NPAS3* se expresa casi exclusivamente en neuronas que sintetizan el neurotransmisor inhibitorio ácido gamma-aminobutírico (GABA). Mientras que no se detectó la presencia de *NPAS3* en células gliales. Por otro lado, el número de células que expresan *NPAS3* excede, casi en un orden de magnitud, al de las células que expresan *NPAS1* (Erbel-Sieler et al. 2004).

Análisis de ratones mutantes del gen *npas3*

Estudios realizados sobre ratones mutantes donde se truncó la expresión de *Npas3* (*Npas3*^{-/-}) mostraron diversos defectos en el sistema nervioso y en otros sistemas (Erbel-Sieler et al. 2004; Brunskill et al. 2005; Zhou et al. 2009). Por un lado, los ratones *Npas3*^{-/-} presentan una baja viabilidad. Dentro de las primeras horas después del nacimiento, las crías *Npas3*^{-/-} muestran un intenso jadeo y se vuelven cianóticas. Para las 48 horas de vida, el 95% de estas crías muere, mientras que el 100% de sus hermanos heterocigotas para la versión truncada del gen u homocigotos para el locus no afectado, sobreviven. La causa de la alta letalidad de las crías *Npas3*^{-/-} se atribuye principalmente a defectos severos en el desarrollo de los alvéolos con un posterior desarrollo de bronquiectasia (una condición donde ocurre una dilatación anormal e irreversible del árbol bronquial, Zhou et al. 2009). Por otro lado, en las camadas que fueron cuidadas por madres mutantes *Npas3*^{-/-} también se observó una alta mortalidad

de las crías (independientemente de su genotipo en el locus de *Npas3*) debido a carencias en el cuidado parental (Erbel-Sieler et al. 2004; Brunskill et al. 2005).

Entre otra de las características, los ratones *Npas3*^{-/-} pesan un 20% menos que sus hermanos de genotipo salvaje o heterocigotos y esta diferencia en el peso se mantiene durante toda la vida (Erbel-Sieler et al. 2004; Brunskill et al. 2005). Entre las anormalidades morfológicas halladas en los ratones *Npas3*^{-/-} se incluye alteraciones en el patrón de foliación del cerebelo, hipoplasia y defectos de conectividad del cuerpo calloso, ensanchamiento del tercer y cuarto ventrículo y del ventrículo lateral, reducción de tamaño del hipocampo anterior, entre otros (Brunskill et al. 2005). Con respecto a los defectos comportamentales, se encuentra que los ratones *Npas3*^{-/-} desarrollan comportamientos motores estereotipados (Erbel-Sieler et al. 2004), presentan hiperactividad, defectos en la coordinación motora y en la memoria de reconocimiento y muestran una respuesta alterada frente a situaciones de ansiedad (Brunskill et al. 2005). También se encontraron perturbaciones en las vías de señalización dopaminérgicas, glutamatérgicas y serotoninérgicas (Brunskill et al. 2005). Estos estudios ponen de manifiesto el rol esencial de *Npas3* tanto a nivel estructural como funcional del cerebro.

***NPAS3*, un gen cuya disfunción está asociada al desarrollo de esquizofrenia y otras patologías del sistema nervioso.**

Actualmente, diferentes estudios en humanos asocian ciertas variantes genéticas presentes en el locus que contiene a *NPAS3* con una mayor susceptibilidad a desarrollar esquizofrenia y otras enfermedades neurológicas en humanos (Benjamin S. Pickard et al. 2006). La primera evidencia de dicha asociación se reportó al hallar que una translocación recíproca entre los cromosomas 9 y 14 que ocasionaba la disrupción la secuencia de *NPAS3* entre los exones 3-5 pero no afectaba a ningún gen sobre el cromosoma 9, estaba presente en una familia afectada de esquizofrenia severa (figura 2, A). La madre y la hija afectadas tenían la translocación t(9;14)(q34;q13) pero no estaba presente en el padre no afectado. De expresarse, la predicción del producto proteico de este gen truncado sería una proteína conteniendo el dominio de interacción con el ADN bHLH y de dimerización PAS (Benjamin S. Pickard et al. 2006)

pero carente del dominio C-terminal de transactivación (y posiblemente de alguno de los dominios PAS también) (Kamnasaran et al. 2003; Ben S. Pickard et al. 2005) (figura 2, B). Como la haploinsuficiencia de *NPAS3* no tiene evidentes consecuencias funcionales sobre el fenotipo de ratones mutantes a nivel del sistema nervioso (Erbel-Sieler et al. 2004), una de la hipótesis que surgen es que el cerebro humano, a diferencia del de ratón, podría ser más sensible a la disminución en la carga génica de *NPAS3*. Sin embargo, sí se observaron defectos en la estructura alveolar de ratones heterocigotas para el alelo truncado de *Npas3* (Zhou et al. 2009). Una segunda posibilidad sostiene que la translocación observada en la familia afectada de esquizofrenia podría generar la expresión de una versión de NPAS3 capaz de actuar como dominante negativa. Dado que los miembros de la familia bHLH-PAS deben dimerizar con sí mismos o con otros miembros de la familia para poder ejercer su función, la expresión de una proteína capaz de unirse al ADN y dimerizar (por medio de sus dominios básicos y HLH-PAS, respectivamente) pero carente de sus dominios regulatorios podría ocasionar el secuestro de los miembros proteicos necesarios para la dimerización formando complejos no funcionales sobre el ADN (Erbel-Sieler et al. 2004; Benjamin S. Pickard et al. 2006).

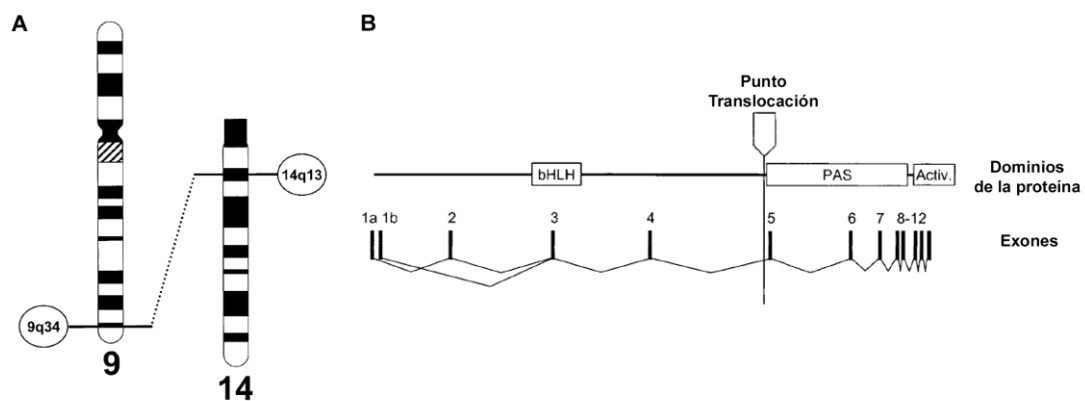


Figura 2. Rearreglo genómico en el gen *NPAS3* presente en el cromosoma humano 14 en una familia diagnosticada con esquizofrenia severa. (A) Ideograma representando la translocación recíproca $t(9;14)(q34.11;q13.1)$ observada en la madre e hija afectadas de esquizofrenia. (B) Esquema indicando el sitio del rearreglo cromosómico afectando al locus de *NPAS3*. Adicionalmente, se muestran los dominios funcionales codificados por los exones de *NPAS3*. Modificado de (Pickard et al. 2005).

Por otro lado, recientemente, se identificó una mutación puntual (un cambio nucleotídico G por A) en la región codificante de *NPAS3* que co-segregaba con esquizofrenia en una familia humana. La mutación identificada está en el exón 8 del gen y da lugar a un cambio aminoacídico valina por isoleucina en la posición 304 (V304I). Esta mutación estaba presente en los tres miembros afectados de esquizofrenia de esta familia y ausente en los miembros restantes. Ensayos funcionales mostraron que, cuando se expresó la proteína *NPAS3*-V304I en neuronas primarias de ratón hubo una disminución significativa del crecimiento de las neuritas respecto de la expresión de la proteína salvaje, brindando evidencia del impacto funcional de esta mutación podría tener (Yu et al. 2014).

Adicionalmente, existen varios trabajos que reportan la asociación de diversos polimorfismos dentro del locus de *NPAS3* con enfermedades psiquiátricas, principalmente esquizofrénica y desorden bipolar (Huang et al. 2010). Por ejemplo, un estudio de caso-control encontró asociación entre múltiples alelos presentes en intrones de *NPAS3* con una mayor probabilidad de desarrollar esquizofrenia y desorden bipolar (Pickard et al. 2009). Más recientemente, un estudio de asociación de escala genómica (GWAS, por *genome-wide association studies*) de variación con desorden bipolar, encontró que 10 polimorfismos de sitio único presentes en intrones de *NPAS3* estaban positivamente asociados con esta condición (Weber et al. 2011).

***NPAS3* y el desarrollo craneofacial**

La holoprosencefalia (HPE) es la malformación estructural más común del cerebro anterior humano y también afecta el desarrollo de estructuras craneofaciales. HPE aparece en 1 cada 250 gestaciones, pero en solo 1 de cada 8,000 nacimientos vivos (Lacbawan et al. 2009). Se han mapeado varios loci para HPE a sitios cromosómicos específicos y, en algunos casos, se han identificado los genes responsables. HPE8 (Holoprosencefalia locus 8) se asigna a 14q13 y se define entre los marcadores D14S49 y AFM205XG5 (Figura 3, Kamnasaran et al. 2005). Las personas que llevan deleciones en HP8 presentan una amplia gama de defectos en el sistema nervioso central que van desde la holoprosencefalia alobar (la forma más grave, en la que el cerebro no se separa) hasta una morfología del SNC casi normal (Kamnasaran et al. 2005). Además,

la morfogénesis facial también se ve afectada y las personas suelen presentar microftalmia, coloboma (malformación ocular), dismorfia facial (Kamnasaran et al. 2005), falta de tabique nasal y labio y paladar hendido con agenesia de premaxila y vómer en casos extremos (Segawa et al. 2007). Tres genes ubicados en esta región, *Sorting Nexin 6* (SNX6), *NPAS3* y *C14ORF11* (*Chromosome 14 Open Reading Frame 11*), se han señalado como posibles candidatos para la HPE. Como ya ha sido mencionado, el factor de transcripción *NPAS3*, se expresa principalmente en el cerebro humano (Gould y Kamnasaran 2011) y ratón (Brunskill et al. 1999). En dos de los tres modelos de ratones *knockout* para *Npas3* reportados en los últimos años (Brunskill et al. 2005; Erbel-Sieler et al. 2004; Zhou et al. 2009) se observaron problemas de desarrollo cerebral, sin embargo, ninguno de estos estudios informó, por el contrario, anomalías craneofaciales en ratones deficientes en *Npas3*, lo que arroja dudas sobre el *NPAS3* como un gen candidato para la HPE ligada a HP8.



Figura 3. Región del cromosoma 14 en donde figuran, con líneas verdes verticales, los marcadores moleculares que definen el área propuesta para el locus de HPE8 (Kamnasaran et al., 2005).

En este sentido, el pez cebra es una herramienta poderosa para estudiar el desarrollo del cráneo y la cara en vertebrados (Schilling y Le Pabic 2009), ya que varios genes clave que participan en el desarrollo de estructuras craneofaciales se han descubierto inicialmente usando este modelo animal (Yelick y Schilling 2002). Por lo tanto, en este estudio, buscamos determinar la importancia funcional del ortólogo de *NPAS3* de pez cebra (*npas3*) en el desarrollo craneofacial.

Estudios funcionales en peces cebra: uso de morfolinos

Una forma rápida de estudiar el rol de un gen particular durante el desarrollo embrionario consiste en reducir los niveles de expresión de su ARNm. Un método particularmente eficiente y sencillo es utilizar morfolinos antisentido para inhibir o reducir los niveles de expresión de un determinado mensajero en peces cebra (Nasevicius y Ekker 2000). Los morfolinos son derivados sintéticos del ADN con dos modificaciones principales: 1) el anillo de desoxirribosa de 5 átomos de carbono es reemplazado por un anillo morfolino con 4 átomos de carbono, uno de nitrógeno y uno de oxígeno; 2) el enlace aniónico fosfodiéster es reemplazado por una unión no iónica fosforodiamidato (Bedell, Westcot, y Ekker 2011) (figura 4). Los morfolinos consisten en oligonucleótidos de aproximadamente 25 bases que actúan vía complementariedad de bases con el ARNm blanco. En pez cebra se utilizan dos tipos principales de morfolinos (MOs): los morfolinos que bloquean la traducción y los morfolinos que bloquean el procesamiento del mensajero (particularmente el *splicing*). El primer tipo de morfolino (MOs traduccional) se une por complementariedad de bases con la región 5' no traducida (5'UTR por sus siglas en inglés) del mensajero blanco impidiendo el ensamblado del complejo ribosomal, y, por lo tanto, previniendo la traducción. Por otro lado, el mecanismo de acción de los morfolinos de procesamiento (MOs de *splicing*) consiste en bloquear la maduración del pre-ARNm inhibiendo la formación de los complejos del spliceosoma (Bill et al. 2009). Los diferentes tipos de MOs tienen usos complementarios a la hora de estudiar la función de los genes en pez cebra. Por un lado, los MOs de *splicing* solo inhiben los mensajeros sintetizados de novo por el embrión. Mientras que, los MOs de traducción son capaces de inhibir tanto los mensajeros cigóticos como los maternos, cubriendo de esta forma, fenotipos tal vez enmascarados por los mensajeros maternos (Bill et al. 2009). En general, los fenotipos producidos por los MOs se analizan durante los tres primeros días del desarrollo embrionario del pez cebra. Sin embargo, muchos estudios reportan que los efectos de los MOs pueden perdurar hasta los 5 días (Kimmel et al. 2003).

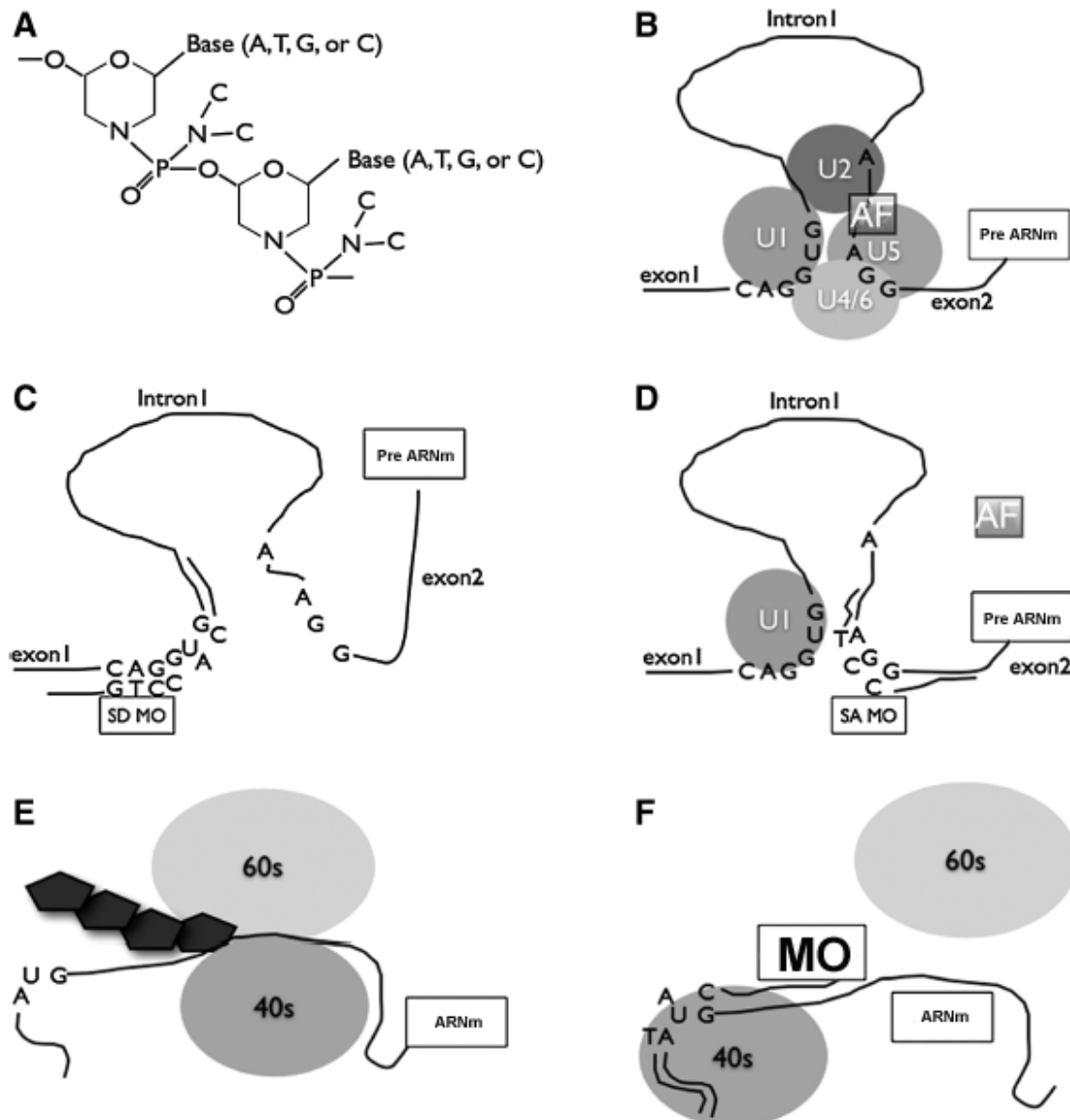


Figura 4. Oligonucleótidos antisentido de morfolidos y sus mecanismos de acción. (A) Los MOs están compuestos por unidades de anillos morfolino unidos por enlaces fosforodiamidato y las mismas bases que tiene el ADN. El mecanismo de acción de los MOs consiste en: (B) impedir el splicing normal o (E) bloquear la traducción. Los MOs diseñados contra el sitio donador (SD) de splicing inhiben la unión del complejo U1, impidiendo la formación del lazo y finalmente promoviendo la incorporación del intrón (C). La incorporación de un intrón frecuentemente produce la finalización prematura de la traducción y la degradación del mensajero por el sistema de nonsense-mediated decay. Los MOs diseñados contra el sitio aceptor (SA) funcionan teóricamente previniendo la unión de la proteína U2AF, necesaria para reclutar el complejo de splicing U2 (D). Los MOs traducionales se unen al sitio ATG o a las secuencias 5'UTR de los mensajeros impidiendo de esta forma el escaneo de la subunidad ribosomal 40S y/o formación del complejo ribosomal completo (F). Figura modificada de Bill et al. 2009.

HIPÓTESIS DE TRABAJO

El gen *NPAS3* juega un papel clave en el desarrollo y funcionamiento del sistema nervioso en mamíferos. Además, ha sido postulado que también participa en el desarrollo craneofacial pero nunca ha sido probado. La hipótesis de trabajo de este capítulo propone que el gen *NPAS3* *participa al mismo tiempo del desarrollo del sistema nervioso y del desarrollo craneofacial.*

OBJETIVOS ESPECÍFICOS

- Estudiar la función de *NPAS3* en pez cebra, un modelo vertebrado que ha sido ampliamente utilizado para estudiar genes que participan en el desarrollo craneofacial y del sistema nervioso. Para ello se realizarán experimentos de genética reversa utilizando una serie de morfolinós dirigidos contra el mensajero del gen y también la generación de peces mutantes carentes de este gen a través de la técnica de CRISPR/Cas9.
- Comparar mutantes del gen *NPAS3* en peces cebra con mutantes en ratones generados también a través de la técnica de CRISPR/Cas9.

RESULTADOS

Estudios filogenéticos para identificar el ortólogo *NPAS3* / *trh* del pez cebra

Identificamos al supuesto gen ortólogo de *NPAS3* en el genoma del pez cebra como el gen *ENSDARG00000109820* ubicado en el cromosoma 17 (Chr17:9248874- 9952341). Con el fin de analizar la relación filogenética entre el supuesto ortólogo de *NPAS3* encontrado en el pez cebra y los genes previamente identificados de mamíferos, vertebrados e invertebrados *NPAS3/trh*, realizamos un análisis filogenético completo. Nuestro análisis filogenético muestra que el ortólogo de *NPAS3* identificado en el pez cebra (*ENSDARG00000079182; npas3*) agrupa con los genes previamente identificados como *NPAS3* de los vertebrados. Además, en el pez cebra el ortólogo del gen *NPAS1* como el gen *ENSDARG00000015876; npas1* (Figura 5 y Figura S5). Más aún, nuestro árbol filogenético (Fig. 5 y Fig. S5) muestra que *NPAS3* y *NPAS1* se derivan de una duplicación de *trh* que probablemente se produjo en la duplicación del genoma completo que ocurrió en el linaje de vertebrados, hace unos 500 millones de años (Dehal y Boore 2005). Como muchos genes, ambos genes ortólogos de *trh*, *NPAS1* y *NPAS3*, se conservaron en el genoma de los vertebrados probablemente para cumplir diferentes funciones (Stanco et al. 2014). Como solo se ha encontrado un ortólogo de *trh* en el céfalocordado *Branchiostoma floridae* (Li, Lu, y Yu 2014), esto respalda la idea de que el evento de duplicación que generó estos genes ocurrió después de la división de las ramas que conducen a los vertebrados.

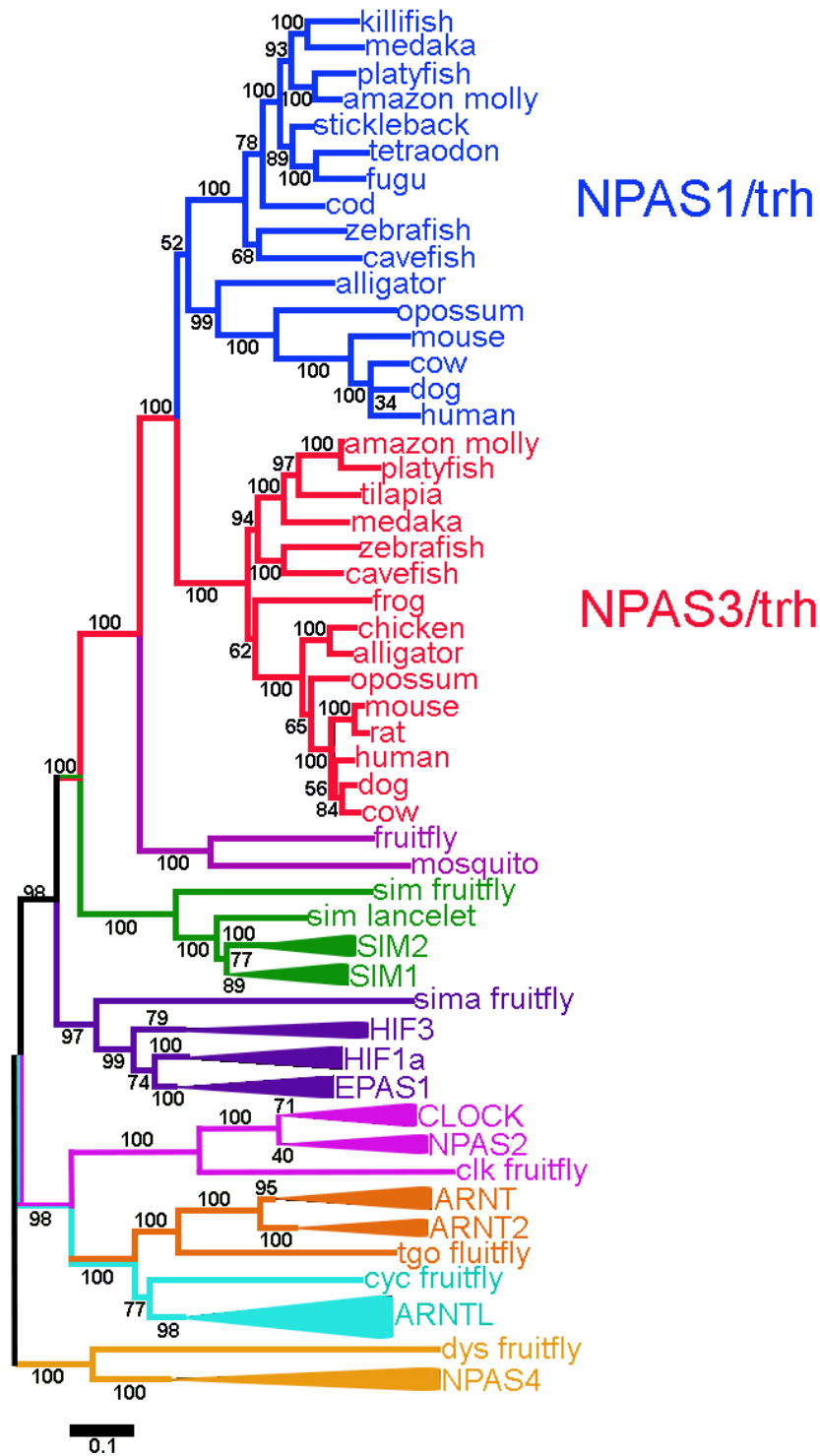


Figura 5. Árbol filogenético de ortólogos y parálogos de NPAS3 de la familia de proteínas bHLH-PAS. Reconstrucción filogenética usando el método *Neighbor-joining* de la familia NPAS3 / *trh* basada en la alineación de secuencia múltiple de longitud completa de 123 especies utilizando el método de sustitución de aminoácidos de Jones-Taylor-Thornton (JTT). Para la alineación de la secuencia, se aplicó un método de eliminación parcial (corte de cobertura del 95%) para dar cuenta de los sitios faltantes. Se llevó a cabo una prueba de *bootstrap* de

filogenia (1000 repeticiones). Las ramas de vertebrados de los parálogos NPAS3 / trh se colapsaron para mayor claridad (ver árbol completo en la Figura suplementaria 5). Barra de escala: 0.1 sustitución / sitio.

Expresión de *Npas3* durante el desarrollo del pez cebra

Para describir en detalle el patrón de expresión de *npas3* durante el desarrollo del pez cebra, realizamos hibridaciones *in situ* de cuerpo entero (*whole mount in situ hybridization*, WMISH) e hibridaciones *in situ* en cortes de criostato (ISH) desde 3 horas después de la fecundación (hpf) hasta 5 días después de la fecundación (dpf) usando una ribosonda para *npas3* específica descrita anteriormente (Kamm, López-Leal, et al. 2013a). Nuestros resultados muestran que *npas3* se expresa desde muy temprano en las larvas de pez cebra. De 3 hpf (1-k células) a 24 hpf *npas3* se expresa ubicuamente en las larvas en desarrollo (Fig. 6). Un día después (48 hpf), el patrón de expresión se restringe a estructuras y tejidos específicos (Fig. 6 G-I). A 48 hpf, *npas3* está altamente expresado en todo el sistema nervioso central, incluido el cerebro y la médula espinal (Fig. 6 G, H). También observamos la expresión en el ojo, el corazón y particularmente en las estructuras craneofaciales, incluidos los arcos branquiales y el techo de la boca (Fig. 6 G, I). A los 3 dpf, *npas3* continuó expresándose abundantemente en el techo bucal y los arcos faríngeos, y también en el corazón y las aletas pectorales (Fig. 6 J, L). En el SNC, la expresión del ARNm de *npas3* estaba presente en el telencéfalo, el diencefalo (incluyendo el tálamo y el hipotálamo), el tectum óptico y el rombencéfalo (a saber, en el bulbo raquídeo, Fig. 7 A-F) y también en la médula espinal (Fig. 7 G-H). También observamos alta expresión en el ojo (Fig. 7 A, B). Fuera del SNC, observamos alta expresión en el corazón (Fig. 7 C, D). La expresión de *Npas3* demostró ser muy estable, ya que se expresa abundantemente dos días después (5 dpf) en el cerebro y también en los arcos branquiales, ojos y corazón (Fig. 6 M-O). En el SNC en desarrollo, las áreas de alta expresión incluyen el telencéfalo, el tálamo, el hipotálamo, la tectum óptica, la médula y la médula espinal (Fig. 7 I-O). Además, detectamos la expresión sostenida de *npas3* en los ojos, arcos branquiales (Fig. 6 M, O y Fig. 7 I-K), aletas pectorales (Fig. 7 N) y corazón (Fig. 7 L). En esta etapa, también notamos *npas3* ARNm en la vejiga natatoria (Fig. 7 O).

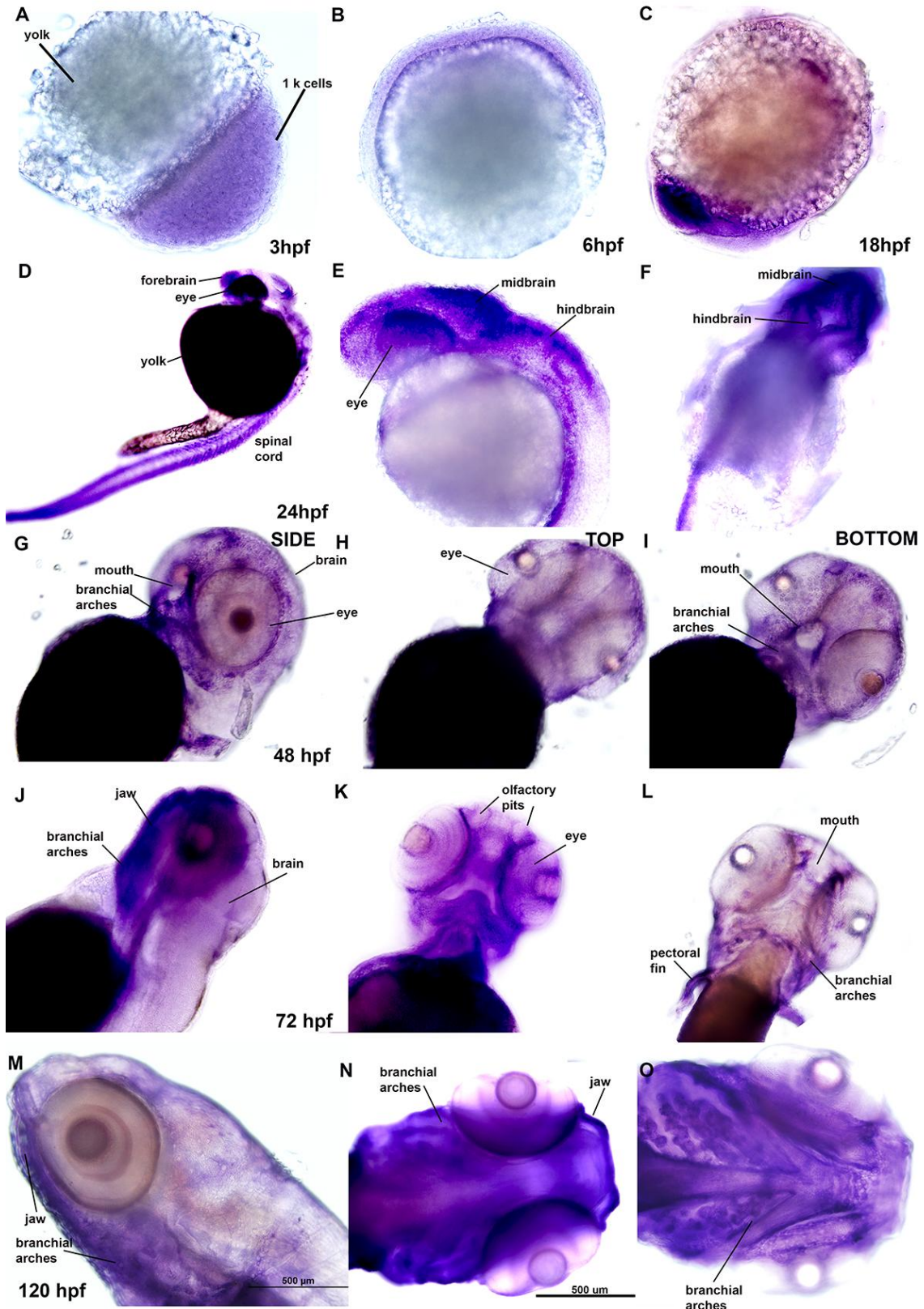


Figura 6. Patrón de expresión del gen *npas3* del pez cebra en diferentes etapas de desarrollo. (A-D) Microfotografías representativas de ensayos de hibridación in situ de montaje completo que muestran el patrón de expresión del gen *npas3* en vista lateral (A-D), dorsal (G, J, M) y ventral (I, L, O) desde 3 (A), 6 (B) 18 (C), 24 (D, E, F), 48 (G, H, I), 72 (J, K, L) y 120 hpf (M, N, O).

La expresión de *npas3* en los arcos branquiales en desarrollo, el corazón y la vejiga natatoria se muestran con puntas de flecha negras, flechas negras y puntas de flecha rojas, respectivamente.

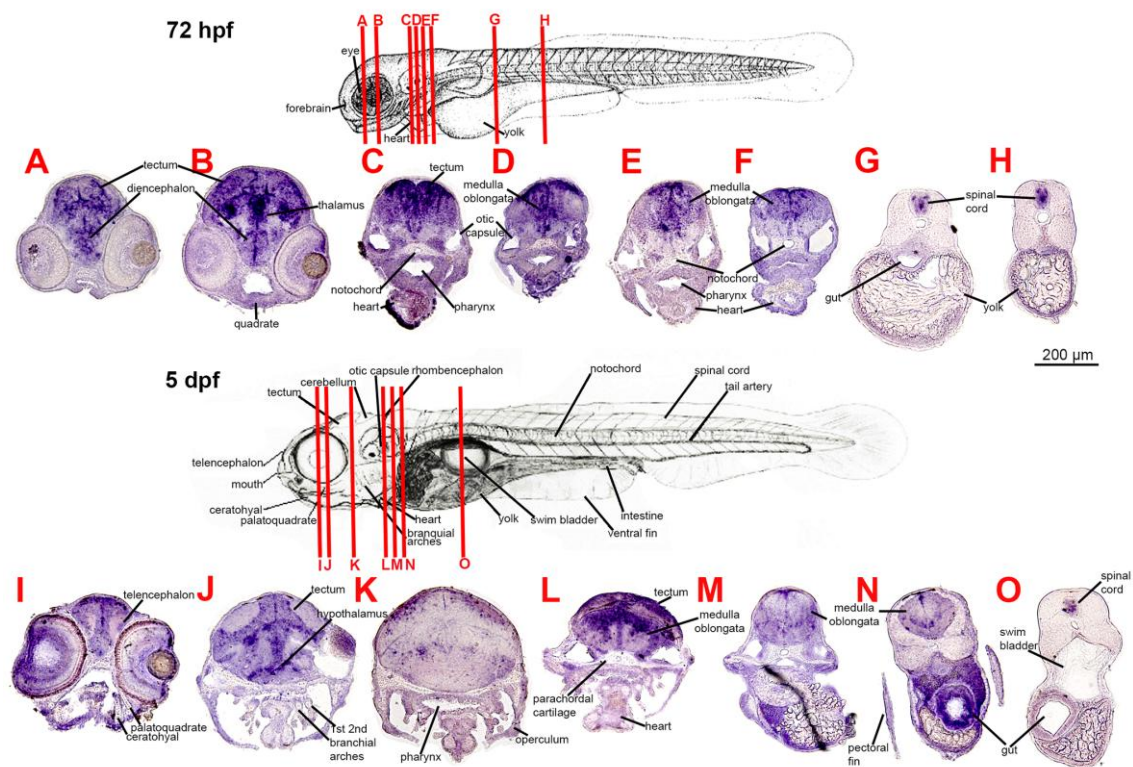


Figura 7. Expresión de *npas3* en el sistema nervioso central y estructuras craneofaciales en el pez cebra. Arriba: Dibujo de pez cebra de 72 hpf que indica (barras rojas) la ubicación aproximada de las secciones coronales que se muestran a continuación. (A-H) Microfotografías representativas de ensayos de ISH en cortes de criostato coronal que muestran el patrón de expresión del gen *npas3* a 3 dpf. Se observa alta expresión en el cerebro (A- E) y la médula espinal (F, G, H). Además, *npas3* se expresa en el ojo, el intestino, los arcos branquiales, la boca, etc. En el medio: Dibujo de 120 hpf de larvas de pez cebra que indican (barras rojas) la ubicación aproximada de las secciones coronales que se muestran a continuación. (I-O) Microfotografías representativas de ensayos de ISH en cortes de criostato coronal que muestran el patrón de expresión del gen *npas3* a 3 dpf. La barra de escala representa 200 μ m.

***NPAS3* juega una función clave en el desarrollo de pez cebra.**

Para estudiar el papel de *npas3* en el desarrollo de peces, buscamos modificar el nivel de expresión de este gen *in vivo* (*knockdown*). Con este objetivo, diseñamos y probamos dos oligonucleótidos antisentido específicos contra ARNm de *npas3*. El primer morfolino (MO1) se diseñó para unir el codón de inicio del ARNm *npas3* (morfolino bloqueante de la traducción) y el segundo oligonucleótido (MO2) se dirigió al sitio de corte y empalme (*splicing*) entre el exón 3 y el intrón 4 de este ARNm

(morfolino bloqueador del *splicing*) (Fig. 8 A). Comenzamos inyectando diferentes concentraciones de soluciones de morfolino en embriones de 1-2 células. Encontramos que la menor concentración probada (0.75 μ M) no afectó la supervivencia de embriones inyectados con MO1 o MO2 en comparación con animales no inyectados (denominados de tipo salvaje; WT) o embriones inyectados con un volumen igual de solución de tampón (control) a 24 dpf (fracción media de supervivencia de los grupos embrionarios a 24 hpf $\pm$ SEM: MO1 = $0,88 \pm 0,04$, n = 290; MO2 = $0,79 \pm 0,11$, n = 300; control = $0,94 \pm 0,02$, n = 280; WT = $0,92 \pm 0,01$, n = 290; $P > 0.05$ para todos los contrastes; prueba de Kruskal-Wallis seguida de las pruebas posteriores de Dunn).

La eficacia de los MOs de *npas3* fue evaluada por varios enfoques. Diseñamos dos pares de primers para evaluar posibles anomalías en el *splicing* del exón 4, sitio de acción de MO2 (figura 8 A). Los experimentos de RT-PCR usando estos dos conjuntos diferentes de primers revelaron la aparición de una banda de 100 pb más pequeña que la del corte y *splicing* correcto, probablemente correspondiente a la exclusión del exón 4 (Figura 8 B, flechas). Esta banda anómala fue específica para embriones tratados con el MO que interfiere con el *splicing*. Además, los ensayos de transferencia Western que usan un anticuerpo contra *Npas3* mostraron la presencia de una banda inmunorreactiva al peso molecular esperado (~ 100 kDa) en las dos condiciones de control ensayadas (Fig. 8 C, carriles 1 y 2). Sorprendentemente, no hemos podido observar niveles de proteína detectables en muestras de larvas inyectadas con MO1 (Fig. 8 C, carril 3). MO2 tuvo un efecto similar pero la reducción en la proteína *Npas3* no fue tan completa (Fig. 8 C, carril 4). Estos resultados indican una caída efectiva del gen objetivo por ambos MO, siendo el MO1 el más eficiente.

Posteriormente, se realizó un análisis fenotípico de larvas de 3 y 5 dpf tratadas con morfolinos *npas3* 0,75 μ M (larvas morfantes). Observamos que una proporción considerable de larvas morfantes mostraba agrandamiento de su pericardio (50% y 33% de animales inyectados MO1 o MO2 respectivamente) y defectos craneofaciales severos (91% y 66% de animales tratados con MO1 o MO2) (Fig. 9 B, C y Fig. 10 A, B) a 3 dpf. A 5 dpf, encontramos que el 50% de las larvas tratadas continuaron mostrando una expansión de su pericardio (Fig. 9 C y Fig. 10A). Además, una gran mayoría de las larvas inyectadas desarrollaron hidrocefalia (75% y 60% de los embriones inyectados con MO1 y MO2 correspondientemente) (Fig. 9 B y C). Además, alrededor del 75% de

los morfantes mostraban defectos en la mandíbula y otros elementos craneofaciales (Fig. 9 C y Fig. 10 B) y una gran mayoría de los peces inyectados con MOs no inflaron su vejiga natatoria (94% y 88% de los embriones inyectados con MO1 y MO2 correspondientemente) (Fig. 9 C, flechas rojas y Fig. 10 D). Por el contrario, estos defectos rara vez o nunca se vieron en animales no inyectados o en peces inyectados solo con solución tampón a 3 o 5 dpf (Fig. 9 B, C y Fig. 10 A-D).

Para probar la especificidad, los embriones se trataron con una dosis subóptima de cada MO (es decir, un nivel bajo que no provoca un fenotipo en sí mismo), solo o en combinación, y su fenotipo analizado. Solo cuando ambos MOs fueron coinyectados a una concentración no efectiva, los embriones exhibieron características previamente vistas en los morfantes *npas3* (Fig. 11 y Fig. 5 B), incluyendo cola curvada, defectos craneofaciales, edema pericárdico, ojos pequeños e hidrocefalia.

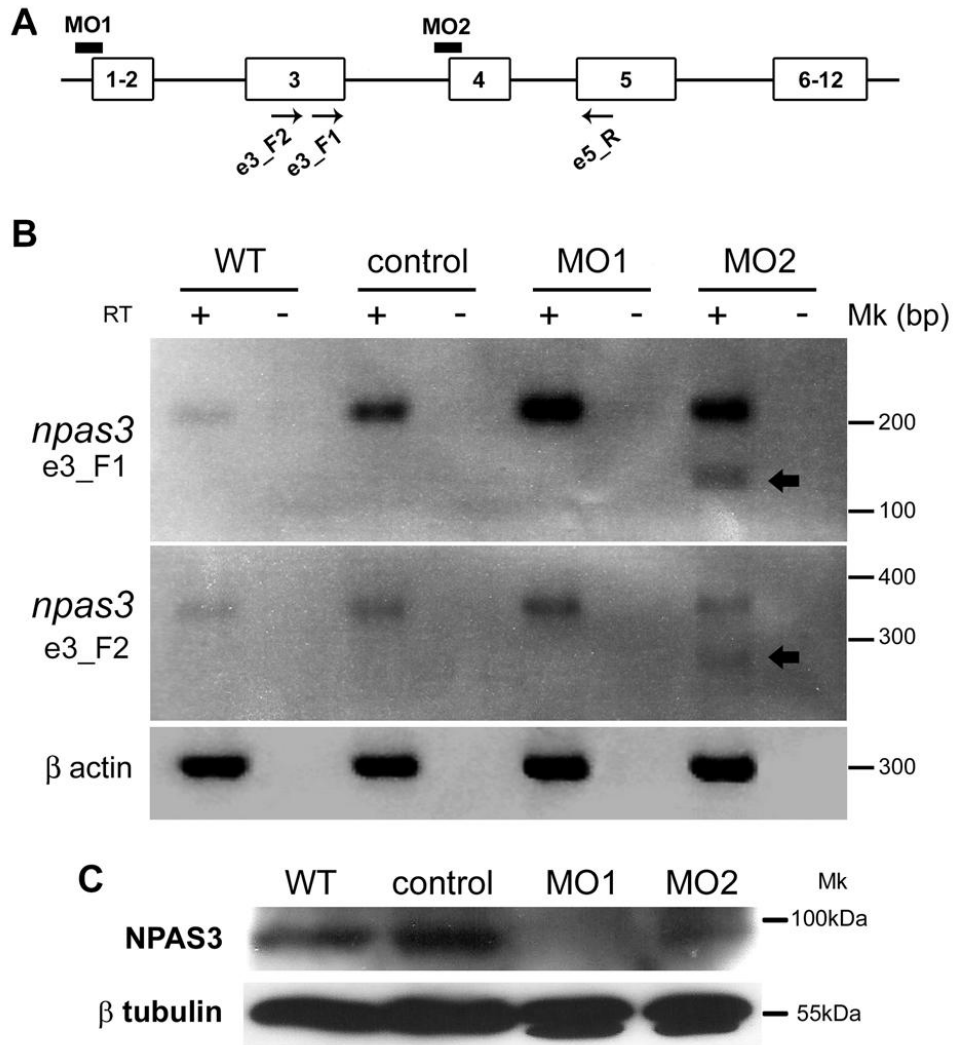


Figura 8. Efecto del tratamiento con MO en el splicing y expresión del ARNm de *npas3*. (A) Representación esquemática del gen *npas3*. Las flechas representan sitios de recocido para *primers* diseñados para detectar anomalías en el splicing del exón 4. (B) Análisis de RT-PCR de *npas3* y β -actina en larvas de 3-dpf inyectadas con MO o solución de control. Las flechas denotan un producto de una banda de menor peso molecular de lo esperado. Cada muestra se analizó en presencia (+) o ausencia (-) de retrotranscriptasa. Los marcadores moleculares se muestran a la derecha. (C) Ensayo de transferencia Western de la expresión de la proteína *npas3* en embriones tratados con 3-dpf. Beta-tubulina se empleó como control de carga. Los marcadores de peso molecular se representan a la derecha.

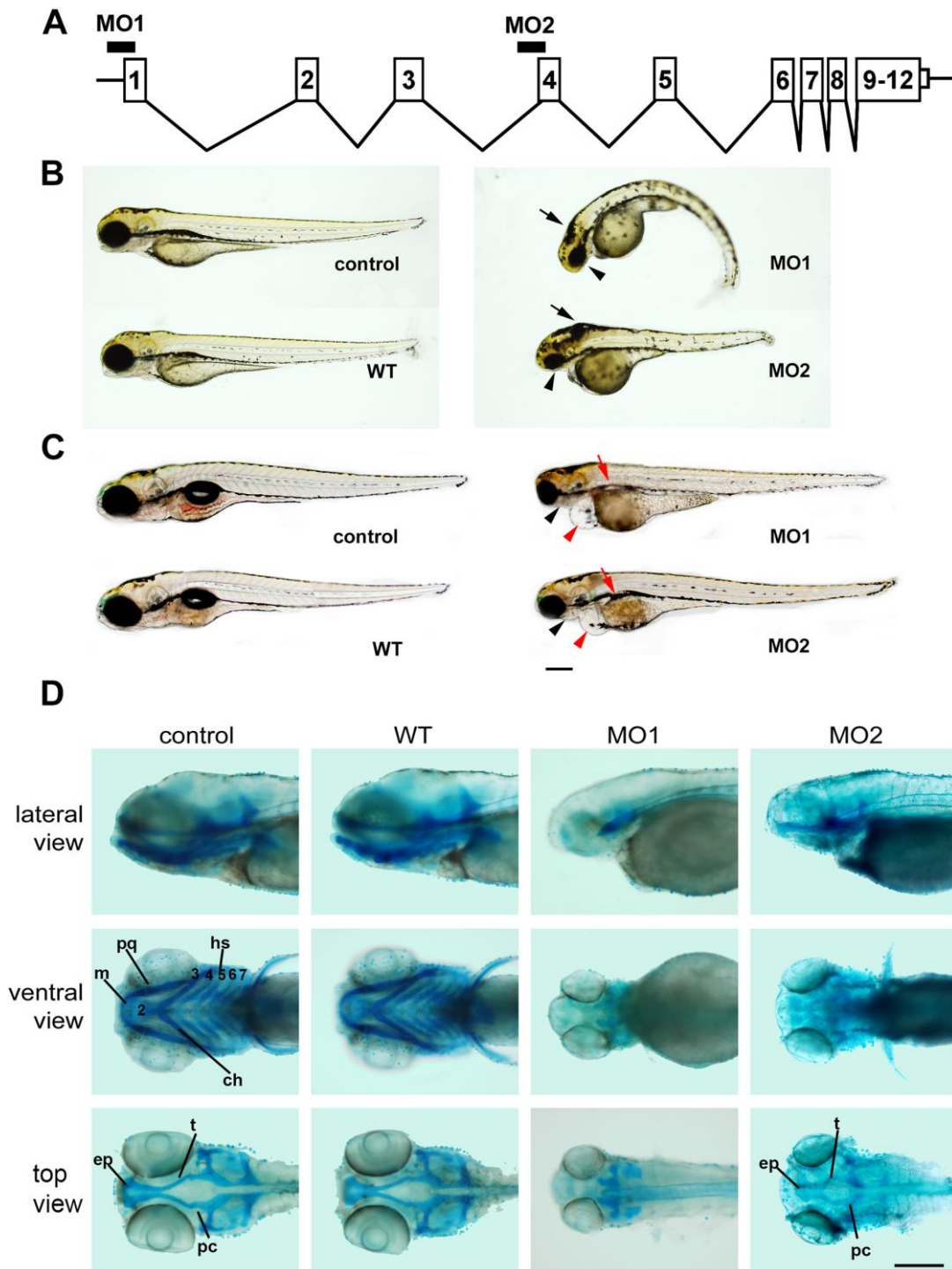


Figura 9. Defectos morfológicos de larvas tratadas con morfolidos *npas3*. (A) Dibujo esquemático del gen *npas3* del pez cebra de acuerdo con la base de datos de *Ensembl* (www.ensembl.org) que muestra el sitio de unión predicho para MO1 y MO2 en el ARNm de *npas3*. (B) Microfotografías representativas de larvas morfantes *npas3* tratadas con $0.75 \mu\text{M}$ de MO1 o MO2, control (control) o peces no inyectados (WT) a 3 (B) o 5 (C) dpf. Las flechas negras indican ventrículos cerebrales agrandados e hidrocefalia. Las puntas de flecha negra muestran defectos craneofaciales. La ampliación del pericardio está indicada con puntas de flecha rojas y la vejiga natatoria inflada por debajo se representa con flechas rojas. (D) Microfotografías representativas que muestran tinción de cartílago azul de alcian en morfantes *npas3* y animales de control a 5 dpf en lateral (panel superior) y vistas ventrales (panel central)

donde es posible apreciar los arcos faríngeos o después de disecar larvas para exponer el neurocráneo (panel inferior). ep: placa etmoidal; t: trabécula; pc: paracordales; m: cartílago de Meckel; pq: palatocadrado; hs: hiosimpléctico; ch: ceratohial; los números indican los arcos branquiales o los cartílagos ceratobranquiales. Barra de escala 250µm.

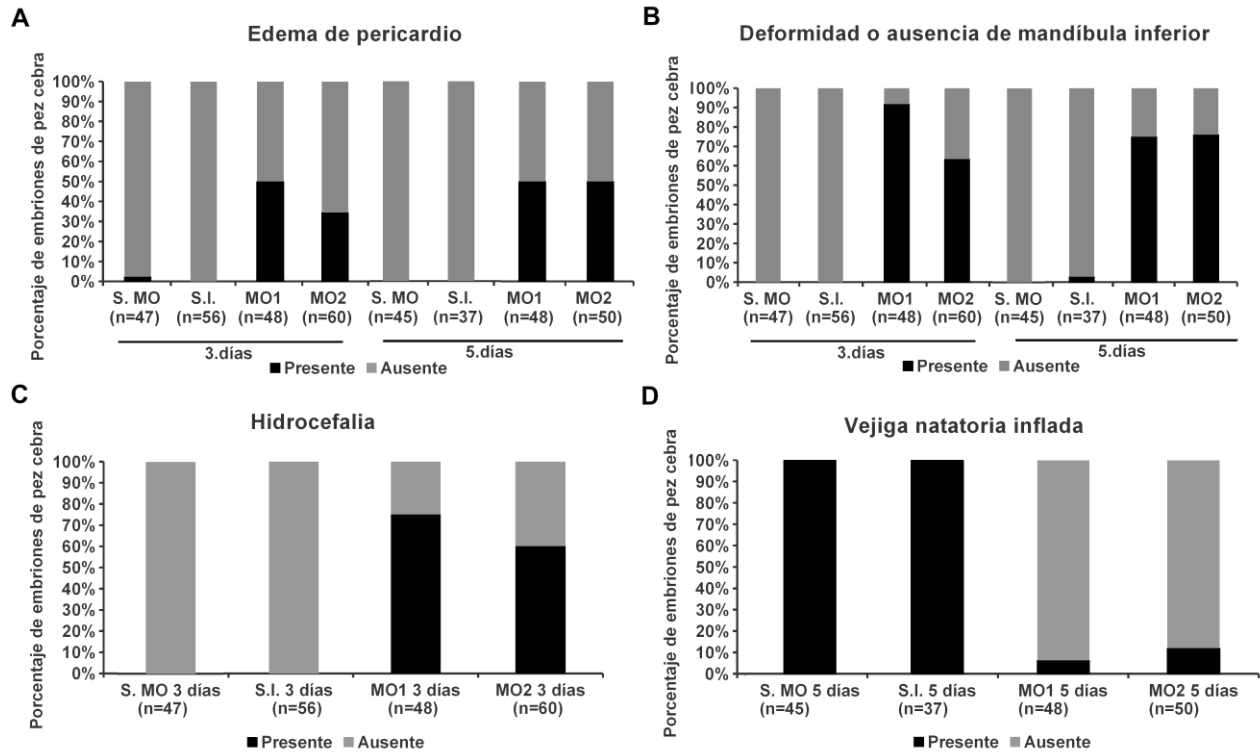


Figura 10. Análisis cuantitativo de defectos morfológicos de larvas tratadas con morfolidos npas3. (A) Porcentaje de larvas que muestran agrandamiento de su pericardio y defectos craneofaciales (B) a 3 y 5 dpf. (C) Porcentaje de larvas que muestran hidrocefalia a 3 dpf y vejiga natatoria inflada por debajo de 5 dpf (D).



Figura 11. Especificidad de los morfolinós *npas3*. Fenotipo de larvas de pez cebra 3-dpf inyectadas con dosis subóptimas (sub) de MO1, MO2 o una combinación de ambas. Las condiciones de control también se muestran. Las flechas negras denotan ventrículos cerebrales agrandados e hidrocefalia. Las puntas de flecha negra muestran defectos craneofaciales mientras que la ampliación del pericardio está indicada con puntas de flecha rojas. Barra de escala 250µm.

El *knockout* de *Npas3* produce un fenotipo similar al de los *knockdown*.

Recientemente, la edición del genoma se ha convertido en fácilmente disponible gracias a la implementación en biología molecular de la tecnología derivada del sistema inmune procarionte, conocido como CRISPR/Cas9. Este método permite una modificación genética específica y precisa y se ha aplicado con éxito a varios modelos animales, incluidos ratones, ratas, ranas, gusanos y peces cebra (Hwang et al. 2013). Con el fin de estudiar más a fondo la función de *npas3*, diseñamos varias guías únicas de ARN (sgRNAs) dirigidos a los mismos exones que los morfolinós, es decir, el exón 1 y el exón 4 (Fig. 12 A y Fig. 13 A). La microinyección de estos ARN guía condujo a la aparición de larvas de 3 dpf con ojos pequeños, cola curva y edema cardíaco (Fig. 12 C y Fig. 13), características que seguían siendo evidentes 2 días después (Fig. 12 D y Fig. 13). Todas estas anomalías se parecían mucho a las observadas en los embriones de morfantes de *npas3* (Fig. 9 B, C). Un análisis detallado del ADN genómico de embriones mutantes reveló la presencia de deleciones en *npas3*, varias de las cuales se predice que alterarán el marco de lectura abierto del ARNm ya que incluyen al codón de inicio y pueden conducir a una proteína truncada (figura 12 B).

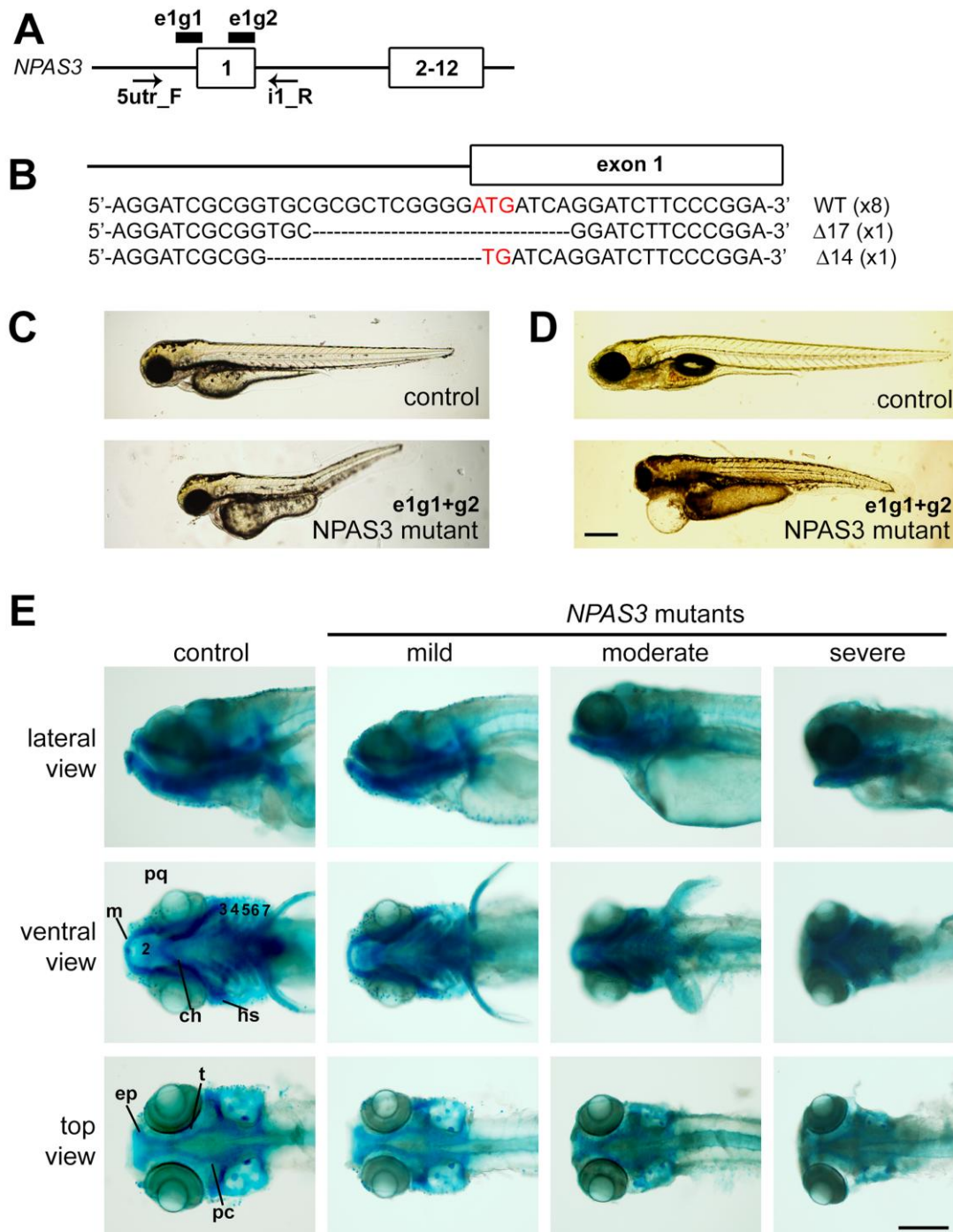


Figura 12. El *knockout* de *Npas3* induce un fenotipo similar al de los MOs. (A) Diagrama esquemático de las guías de ARN utilizadas para generar mutaciones en el gen *npas3*. Ninguno de ellos está dibujado a escala. (B) Los resultados del análisis de genotipo realizado con primers de PCR indicados en A. de larvas mutantes representativas que muestran deleciones que afectan a la región ATG de *npas3* detectadas después de la comparación con la secuencia WT. (C) Embriones mutantes en 3 y (D) 5-dpf que muestran un fenotipo anormal. Los animales de control también se representan en la parte superior. (E) Tinción de azul de alcian de la larva mutante. Los embriones (5 dpf) muestran diferentes niveles de gravedad en el fenotipo que se pueden apreciar en las vistas lateral (panel superior) y ventral (panel central) o después de la disección para exponer el neurocráneo (panel inferior). ep: placa etmoidal; t: trabecula; pc: paracordales; m: cartílago de Meckel; pq: palatoc cuadrado; hs: hiosimpléctico; ch: ceratohial;

los números indican los arcos branquiales o cartílagos ceratobranquiales. Barra de escala 250µm.

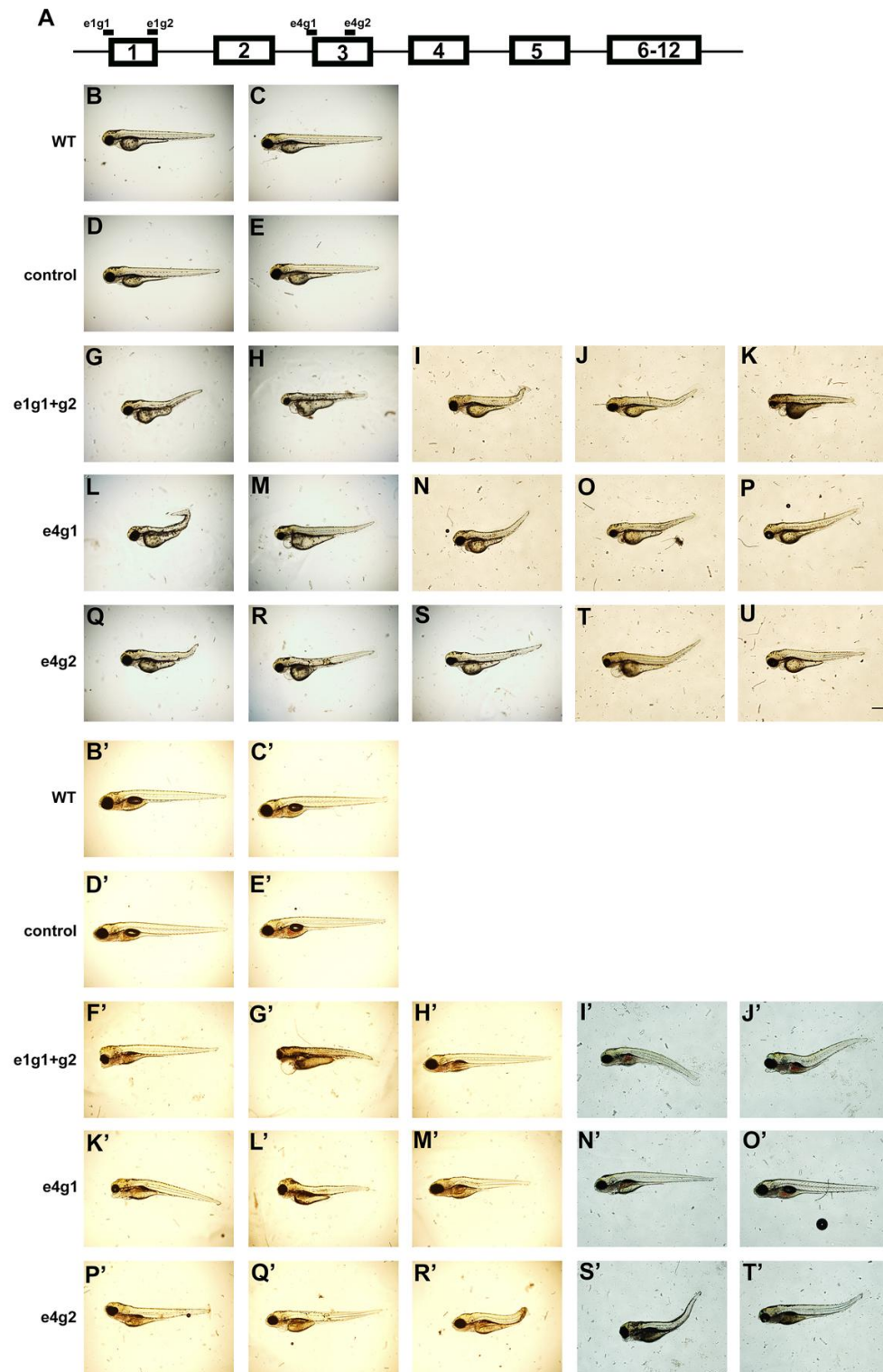


Figura 13. Embriones mutantes 3-dpf y 5-dpf que muestran un fenotipo anormal. (A) Esquema del gen *npas3* del pez cebra que indica la ubicación aproximada de las guías en los exones 1 y 4. Larvas 3-dpf (GN) wildtype (B, C), tratado con vehículo (D, E), tratado con e1g1 + 2 (G-H), tratado con e4g1 (LP) y tratado con e4g2 (QU). Larvas de 5-dpf (G'-N ') wildtype (B,' C'),

tratadas con vehículo (D,' E '), tratadas con e1g1 + 2 (G'-K'), tratadas e4g1 (L ' -P ') y tratado con e4g2 (Q'-U'). Barra de escala 250µm.

***Npas3* es importante para el desarrollo del sistema craneofacial y branquial.**

Observamos que la expresión de *npas3* era particularmente alta en el desarrollo de arcos branquiales y estructuras craneofaciales de larvas de pez cebra (Fig. 6 J-O y Fig. 7 B, I-K). Se ha demostrado que las personas humanas portadoras de deleciones en la región genómica de NPAS3 (conocida como HPE8) presentan problemas de desarrollo del sistema nervioso así como defectos craneofaciales (Kamnasaran et al. 2005; Segawa et al. 2007). Teniendo en cuenta esta información, decidimos estudiar con más detalle el desarrollo del tejido cartilaginoso presente en las estructuras craneofaciales, incluyendo la mandíbula y los arcos faríngeos realizando tinción con azul de alcian (N. C. Bird y Mabee 2003; Solomon et al. 2003) en peces *knockdown* and *knockout* de *npas3*. Las larvas morfantes de 5 dpf mostraron atrofas y malformaciones severas en sus elementos les cartilaginosos, incluida la ausencia casi total de todos los arcos faríngeos (Figura 9 D). Por otro lado, los componentes cartilaginosos del neurocráneo se vieron menos afectados por los MOs de *npas3*, se formaron estas estructuras, pero con alteraciones en su morfología (Fig. 9 D). Las malformaciones fueron más dramáticas en peces tratados con MO1, donde no se formaron las placas trabecular y etmoidal. En el pez cebra tratado con MO2, la placa etmoidal y la trabécula están presentes, aunque son sensiblemente más cortas (figura 9 D). La tinción con azul de Alcian de *npas3*-KO mostró la aparición de mutantes con diferentes grados de compromiso con respecto al desarrollo de arcos craneofaciales y branquiales (Fig. 12 D). Algunos embriones mostraban estructuras ligeramente más pequeñas con conservación de la forma (fenotipo leve) mientras que otros describían anormalidades serias (fenotipo severo). En individuos severamente afectados, no se observan arcos mandibulares o posteriores (Fig. 12 D). En el neurocráneo de fenotipos mutantes leves y moderados observamos que la placa etmoidal y la trabécula están presentes, pero son muy cortas (Fig. 12 D) de forma similar a las observadas en peces tratados con MO2 (Fig. 9 D).

NPAS3 afecta la supervivencia en el pez cebra.

La determinación de la tasa de supervivencia mostró que una gran proporción (~ 70%) de larvas inyectadas con el morfolino bloqueador de la traducción (MO1) murió durante el quinto día después de la fecundación (dpf), mientras que interferir con el splicing (MO2) condujo a una mortalidad más tardía ya que alrededor del 70% de las larvas murieron entre 8 y 9 dpf (Fig. 14 A). Entre el día 6 y el día 10 la mayoría de los peces morfantes mueren ya que no pudieron comer (Fig. 14 A). Con respecto a los embriones *npas3* mutantes, las curvas de mortalidad fueron menos pronunciadas, con aproximadamente el 50% de las larvas que mueren a 5 dpf (Fig. 14 B). Se ha demostrado que los ratones *Npas3*-null muestran una mortalidad perinatal muy alta ya que el 95% de ellos mueren en las primeras 48 horas después del nacimiento (Zhou et al. 2009). Sin embargo, en otro modelo de ratón *knockout* de *Npas3* se ha indicado que la viabilidad se ve gravemente afectada, pero no se informan los porcentajes de mortalidad (Stanco et al. 2014). En otros modelos de ratones no hay informes de una mayor mortalidad, aunque los problemas de nidación y cuidado de las crías redujeron severamente la viabilidad de los cachorros (Brunskill et al. 2005; Erbel-Sieler et al. 2004).

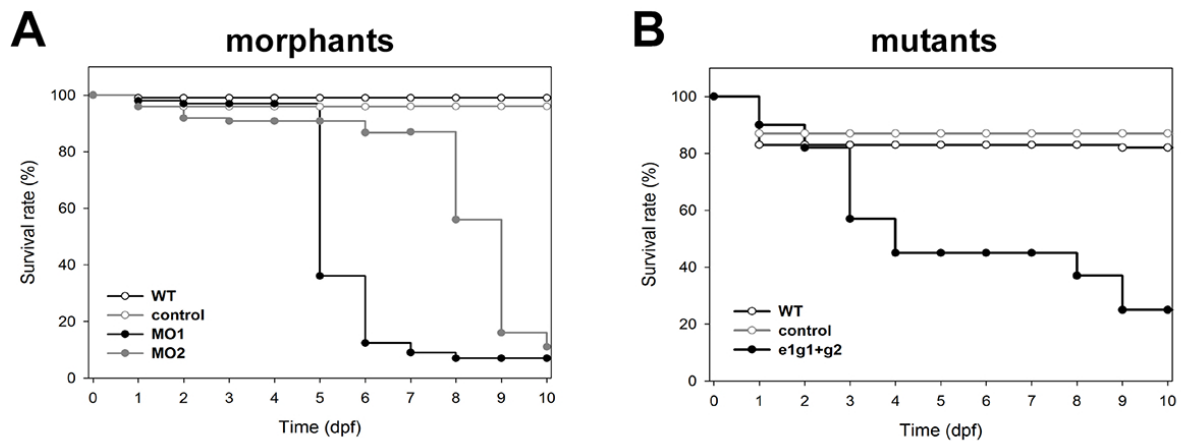


Figura 14. Tasa de supervivencia de embriones alterados. (A) Curva de supervivencia de animales inyectados con MO o (B) sgRNA dirigidos hacia *npas3*. Las condiciones de control también se muestran.

La caída de *Npas3* afecta el desarrollo del sistema nervioso y la proliferación neural.

Con el fin de caracterizar aún más el papel de *npas3* en el desarrollo del sistema nervioso, analizamos la expresión del marcador de desarrollo de neuronas HuC/D en peces tratados con morfolinós *npas3*. Descubrimos que, en comparación con los peces control, los peces inyectados con MO de traducción mostraron una disminución dramática en la expresión de HuC/D a 3 y 5 dpf en la región dorsal del mesencéfalo, incluyendo el tectum óptico, el límite del cerebro medio y posterior (que incluye el cerebelo), el rombencéfalo y la retina (Fig. 15 y Fig. 16).

Dado que la proteína *Npas3* de ratón se ha implicado en la regulación de la proliferación de progenitores neuronales en el hipocampo y también en el telencéfalo (Pieper et al. 2005; Stanco et al. 2014), evaluamos la posibilidad de que los embriones morfolinos tuvieran alteraciones en este proceso. Se analizaron las neuronas proliferantes marcando conjuntamente las células en fase M (fosfohistona H3) y un marcador neuronal (HuC/D). La cuantificación del número de células reveló una reducción considerable en las neuronas que se dividen en aquellos animales a los que se les inyectó el MO que interfiere con la traducción (MO1, Fig. 15).

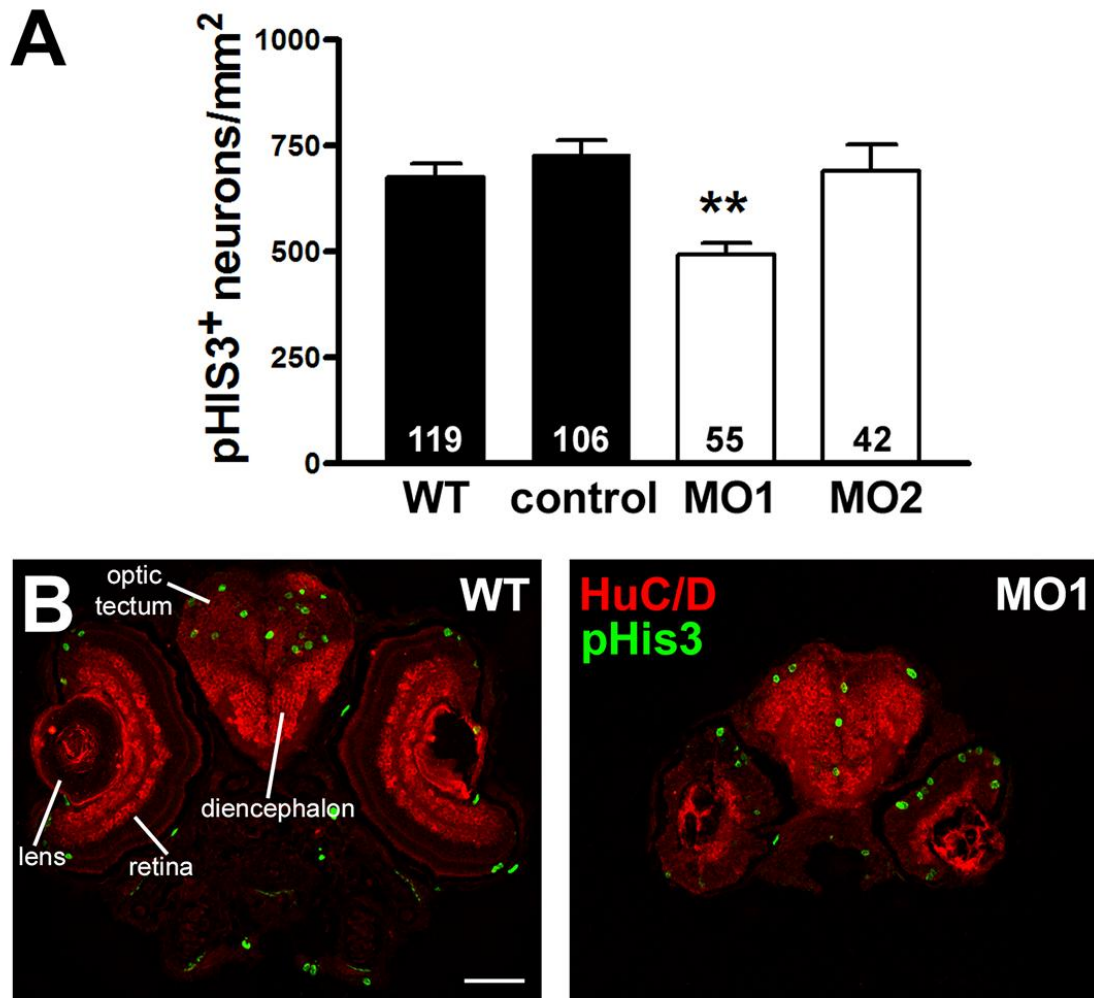


Figura 15. Efecto de la disminución en la expresión de Npas3 sobre la proliferación neural. (A) Cuantificación de neuronas positivas a fosfohistona H3 (pHis3⁺) en cortes de criostato de embriones inyectados de 3 dpf **: $p < 0,01$ en comparación con el grupo WT (ANOVA de una vía seguido de prueba de comparación múltiple Newman-Keuls). (B) Imágenes confocales representativas de animales de control o inyectados con MO que muestran marcaje de fosfohistona H3 (células verdes, proliferantes) y HuC / D (rojo, neuronas). Los números en la parte inferior de las barras especifican la cantidad de datos analizados. Barra de escala 50 μ m.

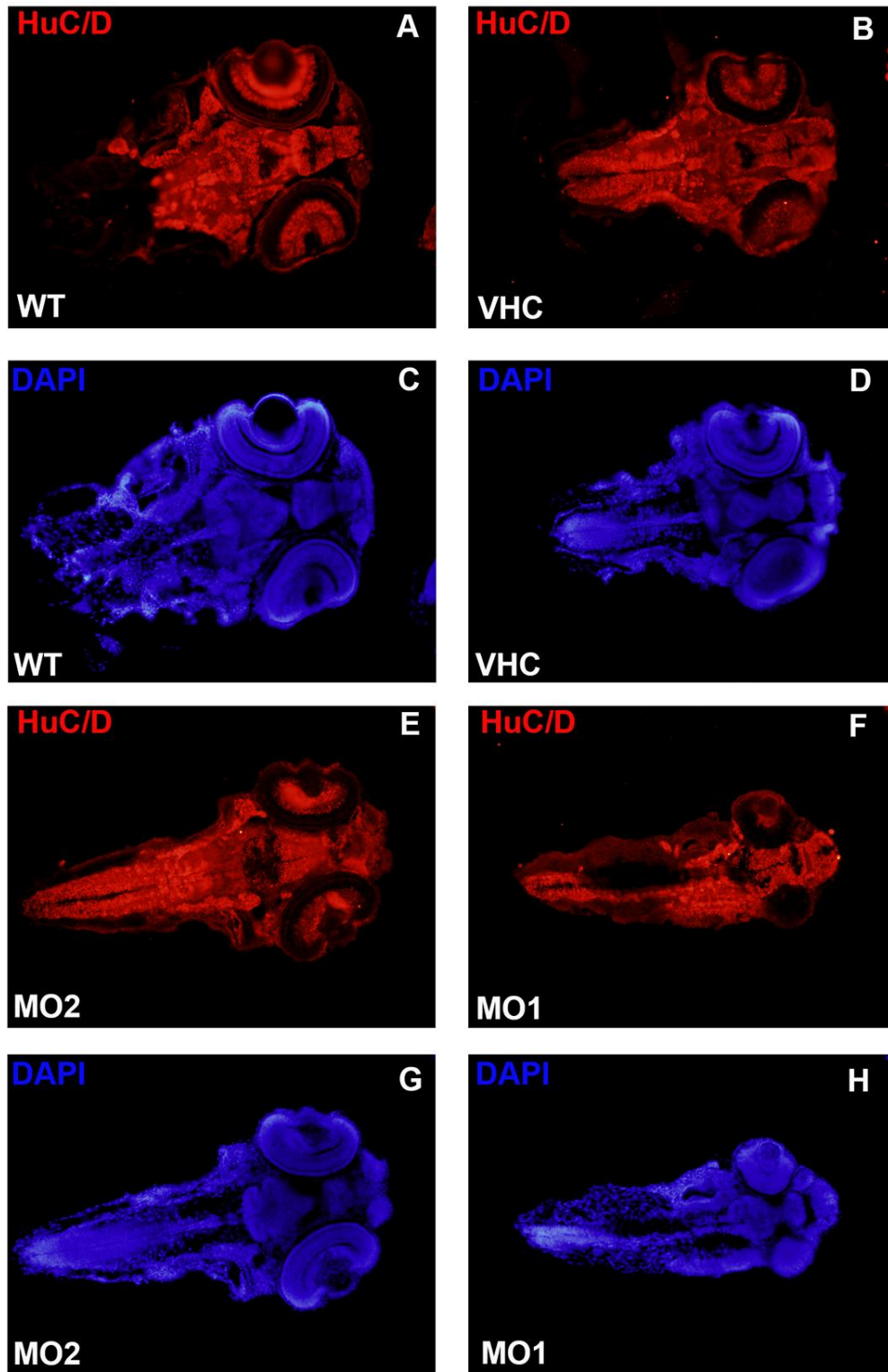


Figura 16. Efectos del *knockdown* de *Npas3* en el desarrollo del cerebro. Imágenes de microscopía de fluorescencia representativa de secciones coronales de criostato de larvas de pez cebra 5 dpf wt(A, C), tratadas con vehículo (B, D), tratadas con el morfolino de splicing (MO2; E, G) tratadas el morfolino de traducción (MO1; F, H), mostrando marcación de HuC/D (A, B, E, F; rojo, neuronas) o DAPI (C, D, G, H, azul, núcleo).

Expresión genética “río abajo” alterada en *knockdowns* de *npas3*

Para explorar los mecanismos subyacentes a los defectos fenotípicos observados en el pez cebra *npas3-knockdown* realizamos estudios de hibridación *in situ* de varios genes previamente informados como objetivos de *NPAS3* / *trh*. De hecho, se ha informado que los efectos traqueales están mediados por la regulación de las vías de FGF en *D. melanogaster* (Ohshiro, Emori, y Saigo 2002; Ohshiro y Saigo 1997).

Para identificar la vía que está regulada por *npas3* en el pez cebra, analizamos la expresión de los receptores de FGF *fgfr1*, *fgfr3* y *fgfr4*. En peces morfantes de *npas3* observamos el aumento y la deslocalización de la expresión de *fgfr1* en el sistema nervioso en desarrollo y los arcos branquiales en comparación con los peces tratados con vehículo y de tipo salvaje a 3 y 5 dpf (Fig. 17). Por otro lado, la expresión de *fgfr3* y *fgfr4* no se vio afectada por la reducción de *npas3* a 3 dpf (Fig. 18).

Además, analizamos la expresión de los ligandos *fgf3*, *fgf8* y *fgf10a*. Encontramos que *fgf3*, *fgf8* y *fgf10a* se expresan fuertemente en estructuras en las que *npas3* está altamente expresado, como el neurocráneo y el viscerocráneo, particularmente en los arcos branquiales a 3 y 5 dpf (figura 19). Sin embargo, la expresión *fgf3*, *fgf8* y *fgf10a* permanece sin modificar en el pez cebra tratado (MO1 y MO2) en comparación con el control (WT y control) a 3 y 5 dpf (figura 19).

Extensos estudios biológicos y genéticos moleculares han demostrado que las proteínas Sprouty de *D. melanogaster* y mamíferos inhiben la señalización del receptor tirosina quinasa río abajo del FGF (Hanafusa et al. 2002; Sasaki et al. 2001; Sivak, Petersen, y Amaya 2005; Tsang y Dawid 2004) y se ha demostrado un aumento de los niveles de ARNm de Sprouty4 en el cerebro de ratones *Npas3* KO (Pieper et al. 2005). Luego, analizamos la expresión de los genes que codifican *sprouty 1* y *4* (*spry1*; *spry4*) en el pez cebra tratado y controlado con MO de *npas3*. Observamos que *spry1* y *4* se expresan ampliamente en todas las estructuras donde *npas3* se expresa en etapas de desarrollo analizadas, como arcos faríngeos, cerebro, boca, etc. Sin embargo, no observamos diferencias en la expresión de los genes mencionados anteriormente en morfante y tipo salvaje peces en las etapas analizadas (Fig. 20).

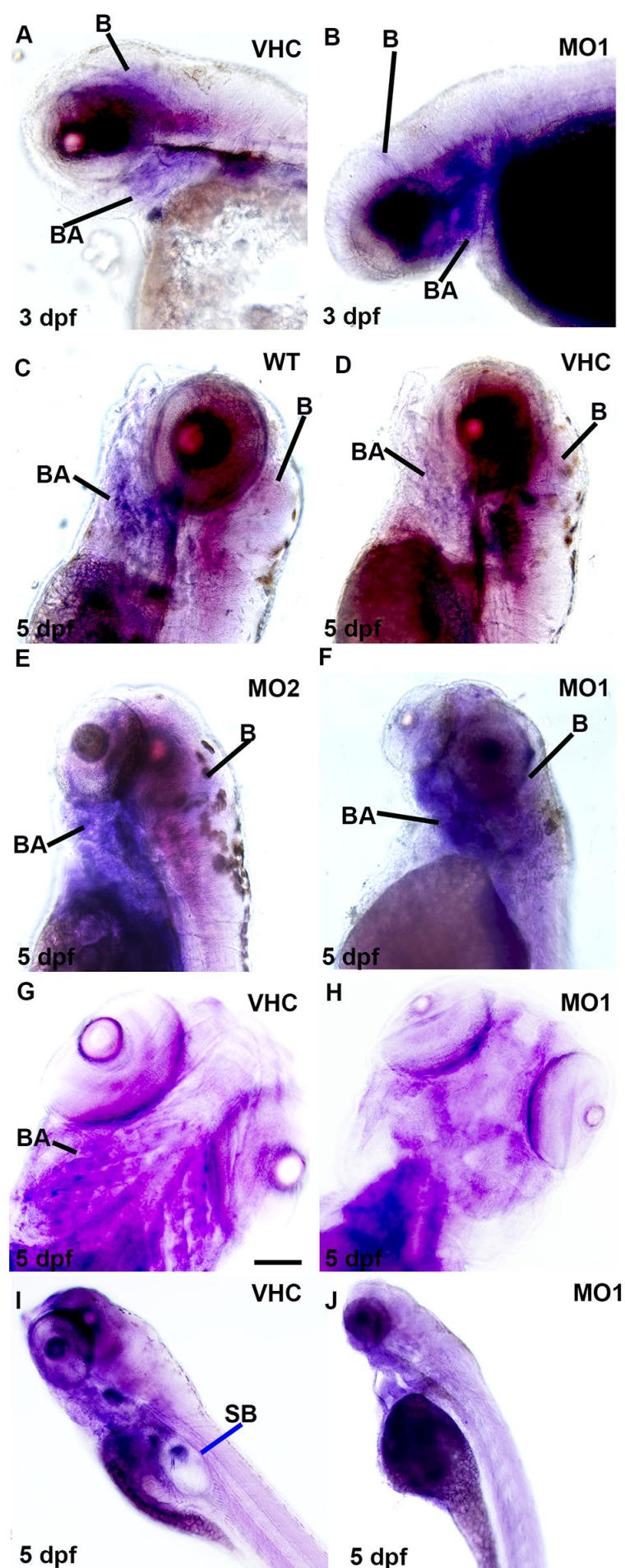


Figura 17. Npas3 regula el patrón de expresión de fgfr1 durante el desarrollo del pez cebra. Expresión de fgfr1 a 3 dpf (A, B) y 5 dpf (CJ) en el embrión tratado con el morfolino de splicing (E; MO2) y el tratado con morfolino de traducción (B, F, H, J; MO1), control (A, D, G, I) y larvas de pez cebra no tratadas (C).

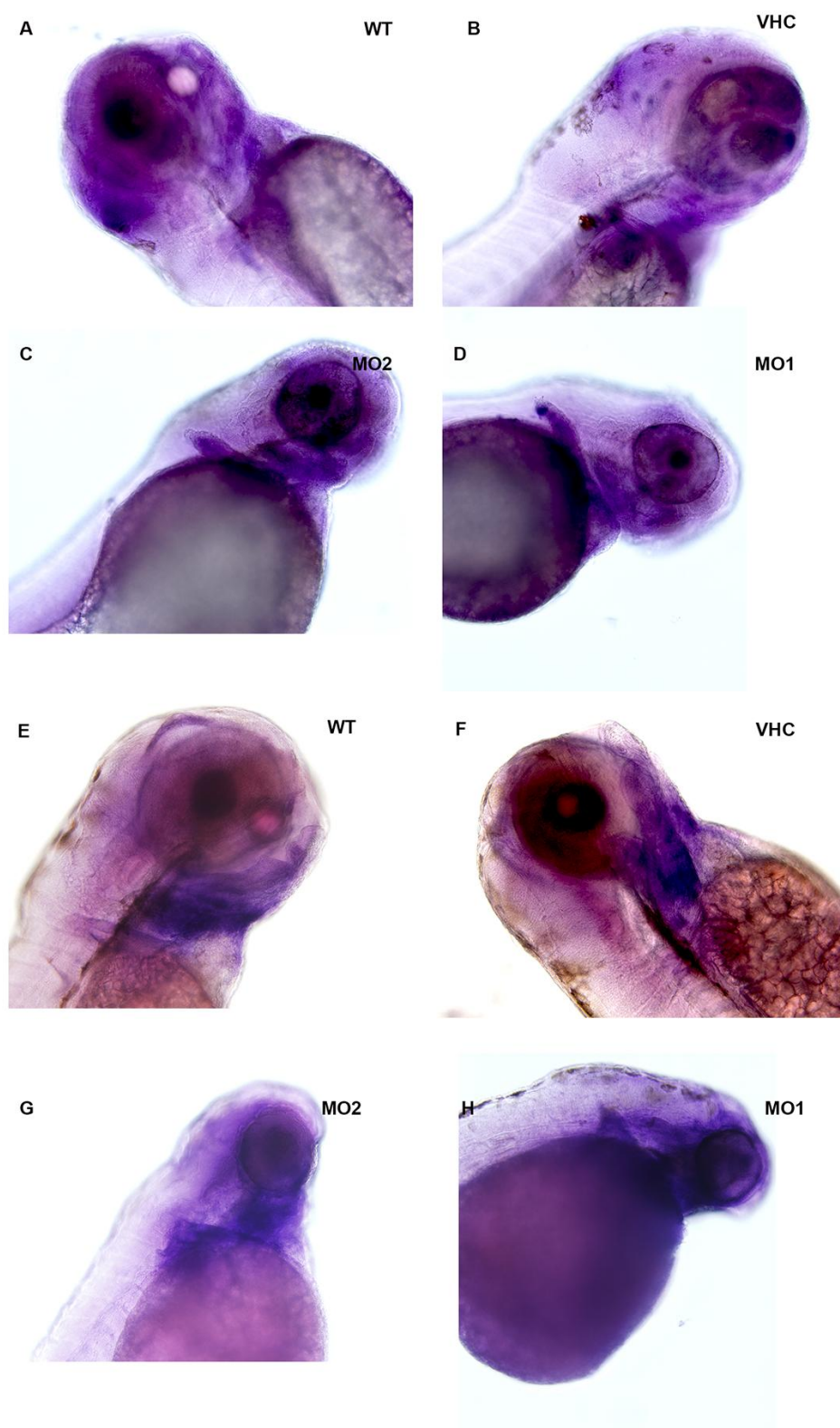


Figura 18. La caída de Npas3 no altera el patrón de expresión de *fgfr3* y *fgfr4*. Expresión de *fgfr4* (A, B, C, D) y *fgfr3* (E, F, G, H) a 72 hpf en el vehículo no tratado (A, E), control (B, F) del embrión tratado con el morfolino de splicing (C, G; MO2) y la traducción morfolino tratado (D, H; MO1) larvas de pez cebra.

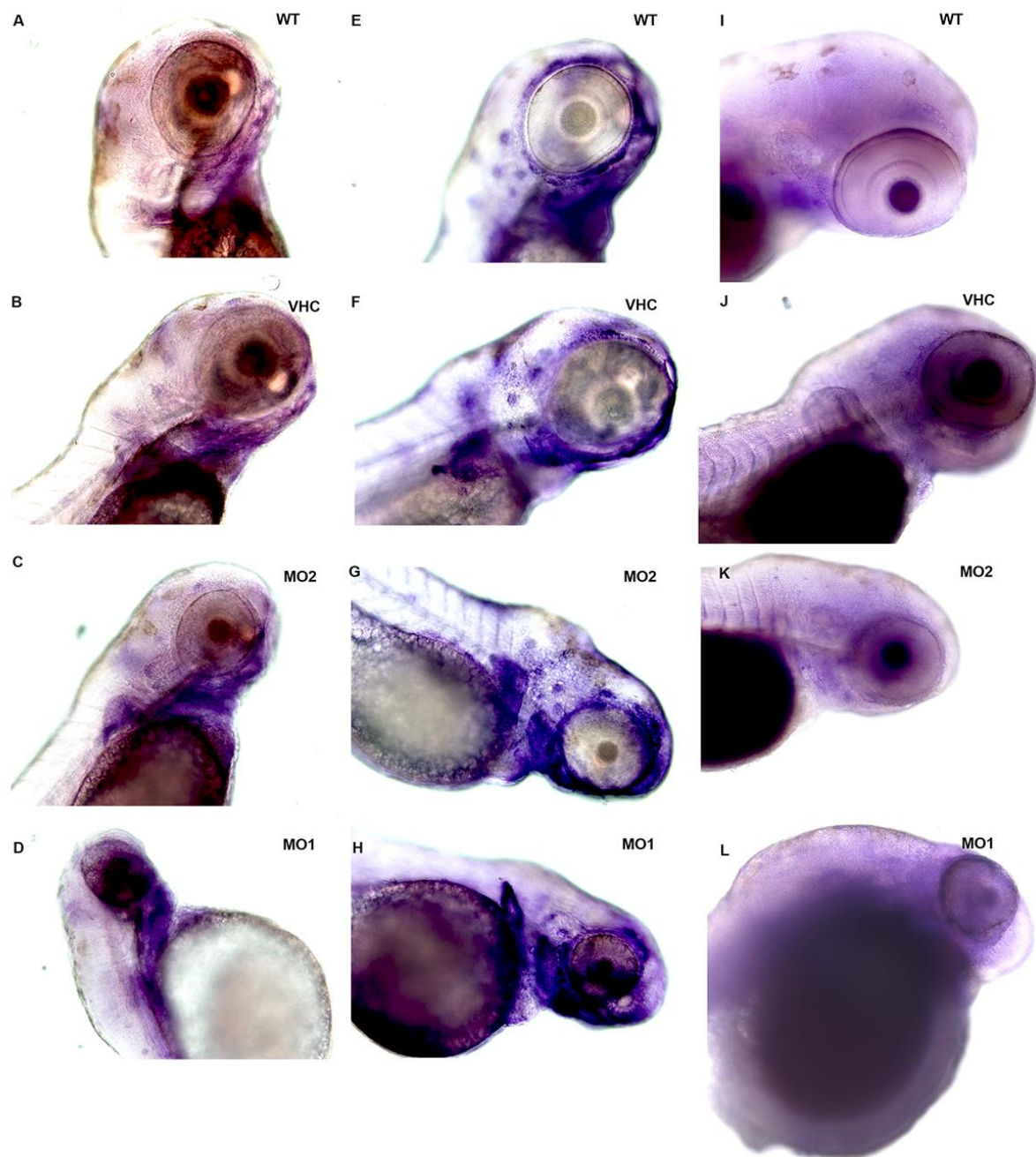


Figura 19. La disminución en la expresión de *npas3* no altera el patrón de expresión de *fgf3*, *fgf8* y *fgf10a*. Expresión de *fgf10a* (A, B, C, D) y *fgf3* (E, F, G, H) y *fgf8* (I, J, K, L) a 72 hpf en vehículo de control no tratado (A, E, I) (B, F, J) embrión tratado con el morfolino de splicing (C, G, K; MO2) y el tratado con morfolino de traducción (D, H, L; MO1).

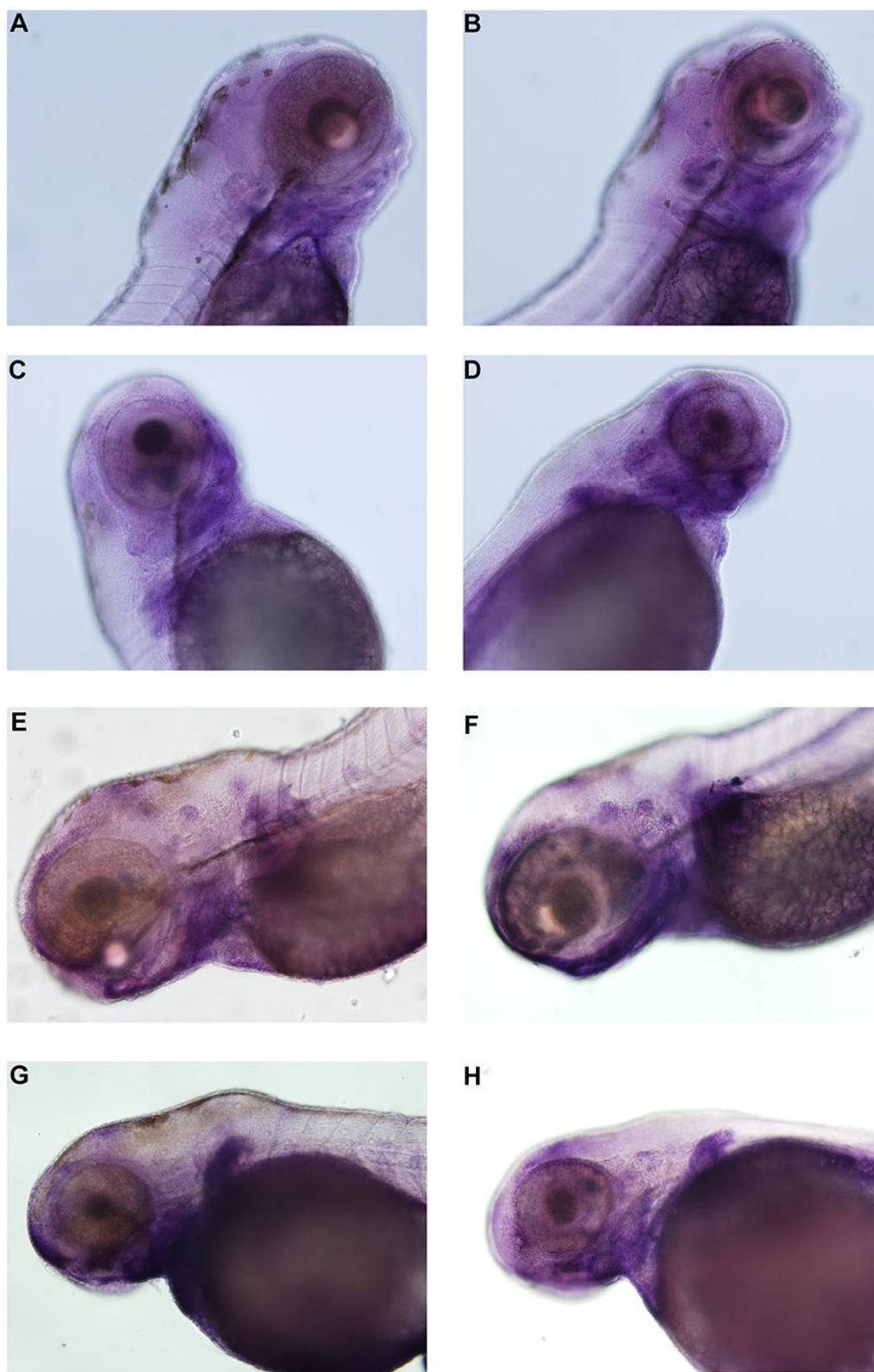


Figura 20. La caída de *Npas3* no altera el patrón de expresión de *spry1* y *spry4*. Expresión de *spry1* (A, B, C, D) y *spry4* (E, F, G, H) a 72 hpf en el no tratado (A, E), vehículo de control (B, F) embrión tratado con el morfolino de splicing (C, G) y el tratado con morfolino de traducción (D, H).

Por otro lado, analizamos la expresión en pez cebra del ortólogo de Sonic hedgehog (*shha*). A las 24 hpf, *shha* se expresa en el hipotálamo (HYP), la zona limitans intratálamica (ZLI), la placa del piso (FP) y la placa basal (BP). Las puntas de flecha negras (Fig. 21, C y D) indican una reducción de la expresión de *shha* en la presunta ZLI de los peces morfantes, esta reducción es más evidente en peces tratados con el morfolino bloqueador de la traducción (MO1, Fig. 21, D). También observamos una expansión de la expresión de *shha* en la región entre la presunta ZLI y el hipotálamo (HYP) (flechas rojas Fig. 21, C y D). También observamos un aumento en la expresión de *shha* en el neuroectodermo dorsal (NE), endodermo faríngeo ventral (PE) y ectodermo oral (OE, Fig. 21, K y L) en peces tratados con MO, en comparación con controles y peces WT (Fig. 21, I y J). También se observa en una vista ventral, que debido a la ausencia de arcos branquiales hay una localización errónea de la boca en desarrollo (Fig. 21, M-P).

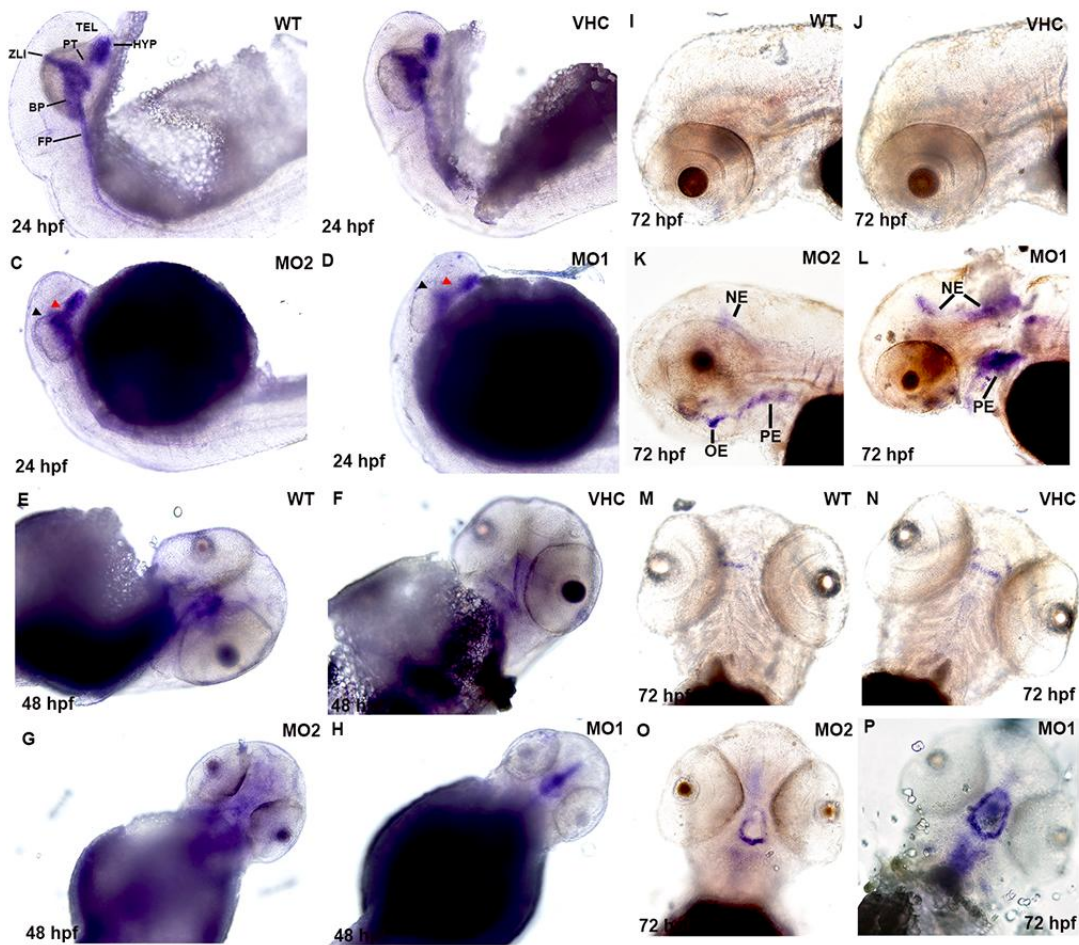


Figura 21. Npas3 regula el patrón de expresión de *shha* durante el desarrollo del pez cebra. Expresión de *shha* a 24 (A, B, C, D), 48 (E, F, G, H) y 72 hpf (I, J, K, L, M, N, O, P) en el embrión

tratado con el morfolino de splicing (C, G, K, O; MO2) y del tratado con morfolino de traducción (D, H, L, P), control (B, F, J, N) y larvas de pez cebra no tratadas (A, E, I, M). A 24 hpf, *shha* se expresa en el hipotálamo (HYP), zona limitans intratalámica (ZLI), placa del suelo (FP), placa basal (BP). Las puntas de flecha negras en C y D indican una reducción de la expresión de *SHH* en la presunta ZLI de peces morfantes; las puntas de flecha rojas en C y D indican una reducción de la expresión de *shha* entre la presunta ZLI y el hipotálamo (HYP). Las flechas azules indican *shha* una expansión temporal de la expresión en el neuroectodermo (NE), el endodermo faríngeo (PE) y el ectodermo oral (OE) en el splicing de morfolino tratado (K, O; MO2) y la traducción morfolino tratada (L, P; MO1) en larvas de pez cebra de 72 hpf.

***Npas3* participa en el desarrollo de la vejiga natatoria.**

Observamos que la gran mayoría de los peces MO inyectados no inflaron su vejiga natatoria (94% y 88% de los embriones inyectados con MO1 y MO2 correspondientemente) (Figura 8, C y Figura 9 D). También observamos la ausencia del desarrollo de la vejiga natatoria en los mutantes *npas3* (Fig. 12 D y Fig. 13). El análisis de ISH muestra que *npas3* se expresa en la vejiga natatoria a 5 dpf (figura 7, O). La vejiga natatoria de teleósteo es un saco lleno de gas ubicado en la parte dorso-anterior de la cavidad corporal que funciona como un órgano hidrostático (Finney et al. 2006; Pelster 2004). Se ha postulado que la vejiga natatoria es un homólogo del pulmón de tetrápodos (Darwin 1859; Farmer 1997a; Longo, Riccio, y McCune 2013). Si bien el desarrollo pulmonar se ha estudiado bien, el mecanismo molecular del desarrollo de la vejiga natatoria no se caracteriza en lo esencial. Aunque los datos morfológicos que respaldan la homología entre el pulmón y la vejiga natatoria son abundantes, los datos moleculares, que actualmente faltan, proporcionarían una fuerte evidencia de apoyo para la relación evolutiva entre la vejiga y los pulmones. Se ha informado que *NPAS3* / *trh* es importante para el correcto desarrollo de los sistemas involucrados en el intercambio gaseoso en otros animales, incluido el sistema traqueal en las moscas (Chung, Chavez, y Andrew 2011; Wilk, Weizman, y Shilo 1996) y en los pulmones de mamíferos (Zhou et al. 2009). Nuestros datos sugieren que *npas3* también podría ser importante para el desarrollo de la vejiga natatoria en peces.

***Npas3* actúa en el desarrollo de estructuras craneales en ratones**

Considerando los resultados obtenidos en peces cebra, surgió la pregunta acerca de por qué no se reportaron modificaciones en el desarrollo craneofacial en los diferentes trabajos donde se generaron ratones mutantes para el gen *Npas3* (Brunskill et al. 2005; Erbel-Sieler et al. 2004; Zhou et al. 2009). Por lo que decidimos generar nuestro propio ratón mutante. Para ello se hizo uso de la técnica de edición génica CRISPR/Cas9. Se diseñaron y generaron guías de ARN dirigidas a los extremos del exón 3, que codifica para el dominio bHLH de unión al ADN. De esta manera esperábamos generar un ratón que exprese una proteína *NPAS3* trunca sin dominio de unión a ADN (Figura 22).

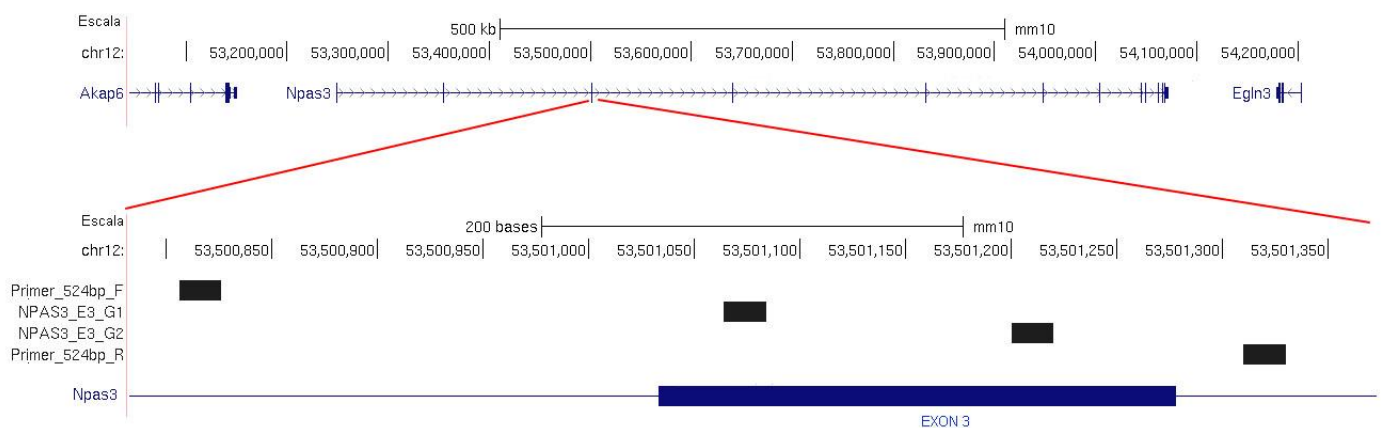


Figura 22. Región del gen de *Npas3* de ratón que fue eliminada. Se ven las guías utilizadas dentro del exón 3 (NPAS3_E3_G1 y G2) y los primers utilizados para genotipificar los ratones mutantes (Primer_524bp_F y R).

Una solución conteniendo las guías diseñadas juntamente con el ARNm de la enzima CAS9 fue microinyectadas en cigotos de ratón y posteriormente transferidos al útero de hembras pseudopreñadas donde se los dejó desarrollar hasta los 14.5 dpc. Se eligió esta edad primero porque los ratones no son viables luego del nacimiento, mueren cianóticos a las pocas horas (Zhou et al. 2009). Se intentó dejar nacer una camada, pero obtuvimos el mismo resultado. Por eso para observar un resultado de manera más o menos rápida decidimos utilizar esta estrategia. Los animales fueron

genotipificados y observamos que hubo generación de ratones mutantes portadores de la delección (Figura 23).

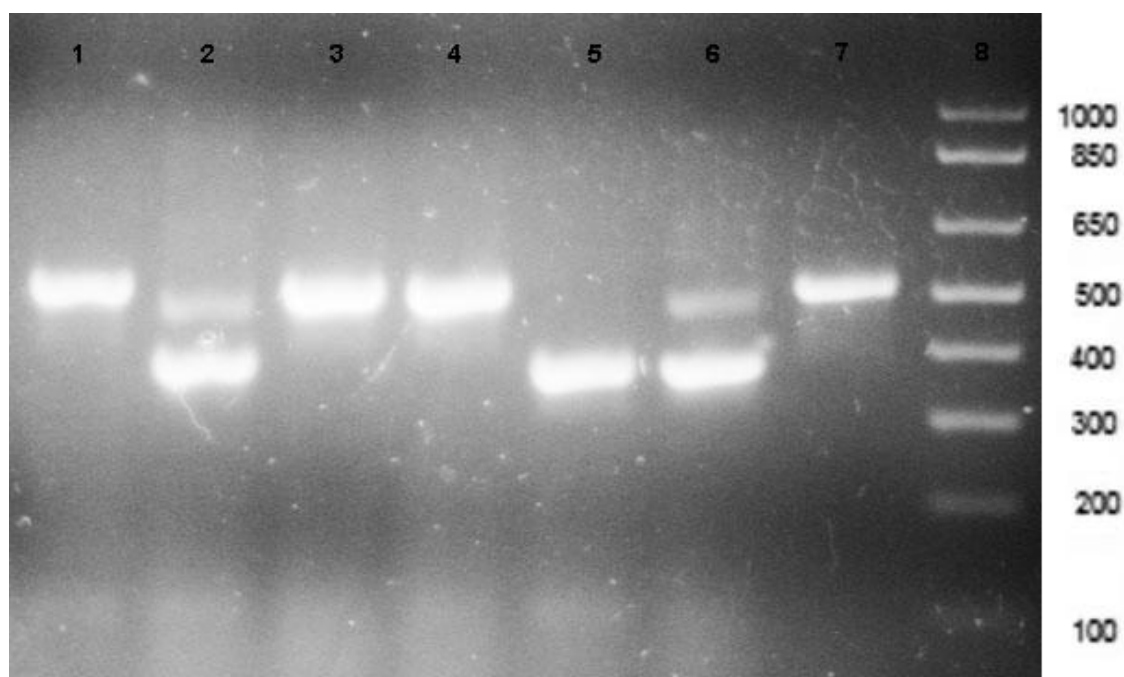


Figura 23. Productos de PCR para genotipificación de los ratones 14.5 dpc microinyectados con las guías y la enzima CAS9. Carril 1, 3 y 4 corresponden a ratones donde no hubo mutación, ya que la banda obtenida corresponde en tamaño al wt (~520pb, carril 7). Carril 2 y 6 corresponden a ratones heterocigotas ya que se ven dos bandas, una del tamaño wt y la otra de ~380pb, que era lo esperado por las guías utilizadas. Carril 5, ratón homocigota para la delección esperada.

Los embriones portadores de la mutación en homocigosis no presentaban un fenotipo anormal en su aspecto morfológico exterior, pero observamos que presentaban un tamaño ligeramente mayor al de sus hermanos heterocigotas y sin mutación.

Para poder analizar que ocurría con el desarrollo craneofacial, sometimos a los embriones a un tratamiento de tinción con azul de Alcian. Este colorante se utiliza para marcar cartílago cómo ya fuera descrito en los experimentos con peces.

Observando los resultados de la tinción, no observamos en los embriones portadores de la mutación (*Npas3* $-/-$, *Npas3* $+/-$) defectos en el hueso dentario, como así tampoco en el maxilar en comparación con los embriones de tipo salvaje (*Npas3* $+/+$; Fig. 24, D, I, N). Si podemos apreciar una anomalía en los huesos del cráneo posterior en los embriones portadores de la mutación (Fig. 24, E, J, O). Lamentablemente solo obtuvimos un único ratón homocigota para la delección esperada por lo que hasta no

obtener una muestra de mayor número de animales no podremos sacar conclusiones. Sin embargo, este experimento nos permitió poner a punto la técnica de generación de ratones mutantes y también la técnica de tinción de azul de Alcian para la detección de defectos craneofaciales.

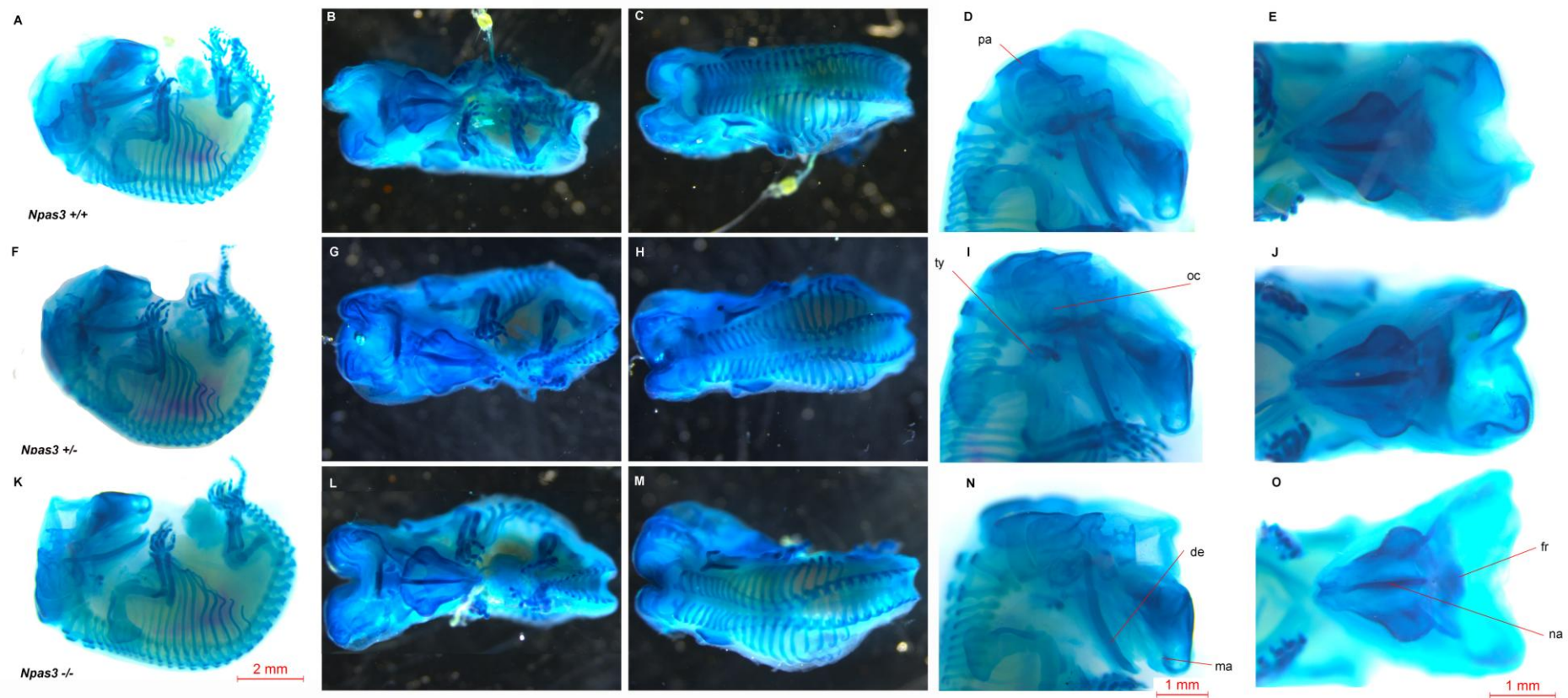


Figura 24. Tinción de Azul de Alcian de los ratones 14.5 dpc microinyectados con las guías y la enzima CAS9. (A-E) Ratón sin delección. (F-J) Ratón heterocigota para la delección. (K-O) Ratón homocigota para la delección. (pa) parietal, (ty) cartilago tiroideo, (oc) capsula otica, (de) dentario, (ma) maxilar, (fr) frontal expiciente, (na) huesos nasales.

DISCUSIÓN

En este trabajo, caracterizamos funcionalmente el ortólogo *NPAS3 / trh* de pez cebra que habíamos identificado recientemente (Kamm, Pisciotano, et al. 2013a). Primero, analizamos la relación filogenética entre el gen identificado como el ortólogo de *npas3* de pez cebra (*ENSDARG00000079182*) y otros miembros de la familia de factores de transcripción HLH-PAS que incluye *NPAS3 / trh*. Este estudio sumado a una búsqueda exhaustiva en el genoma del pez cebra nos permitió definir qué efectivamente el gen *ENSDARG00000079182* es el ortólogo del gen *NPAS3* y qué este gen no se haya duplicado en el genoma del pez cebra. Encontramos además que los genes *NPAS3* y *NPAS1* de vertebrados se derivan de una duplicación de *trh* que probablemente ocurrió en la duplicación del genoma completo que ocurrió en el linaje de vertebrados, hace unos 500 millones de años (Dehal y Boore 2005).

Mostramos que *npas3* se expresa desde muy temprano durante el desarrollo del pez cebra primero de forma ubicua y luego de una manera más restringida en el sistema nervioso, el corazón, las estructuras craneofaciales y las branquias en desarrollo. Estudios anteriores han demostrado que *Npas3* se expresa en el sistema nervioso durante el desarrollo del ratón y también en el adulto (Brunskill et al. 1999; Erbel-Sieler et al. 2004). Además, se ha demostrado que *Npas3* se expresa en el corazón en desarrollo, el oído interno, las glándulas salivales, la dermis y el riñón en el ratón (Brunskill et al. 1999). En los seres humanos, *NPAS3* se expresa en el cerebro en desarrollo (Gould y Kamnasaran 2011). En pollos, el gen *NPAS3* se expresa en el sistema nervioso y el corazón en desarrollo (Shin et al. 2010). Por otro lado, *trh* se expresa en el sistema traqueal en la mosca de la fruta en todo el desarrollo y su función es clave para el desarrollo de este sistema (Wilk, Weizman, y Shilo 1996). El presente estudio proporciona evidencia de que *npas3* es clave para el desarrollo del sistema respiratorio en el pez cebra, ya que la ausencia o reducción de la expresión de *npas3* afecta el desarrollo de las branquias. Además, nuestros resultados sugieren que *npas3* sería importante para el desarrollo de la vejiga natatoria, y la evidencia obtenida indicaría que los programas de desarrollo que subyacen al desarrollo de los pulmones y

la vejiga natatoria se habrían conservado a lo largo de la evolución de los vertebrados. Se ha informado que *Npas3* es crítico para el desarrollo pulmonar, ya que se ha demostrado que en un modelo de ratón que carece de *Npas3* donde se eliminó el dominio PAS-A, hubo un 95% de mortalidad perinatal debido a insuficiencia respiratoria (Zhou et al. 2009).

Nuestros resultados proporcionan evidencia de respaldo al hallazgo de que, en todos los modelos de vertebrados estudiados hasta ahora, el gen *NPAS3* es importante para el desarrollo del sistema nervioso, además de tener una función clave en el desarrollo del sistema respiratorio, incluidas las branquias en peces y los pulmones en mamíferos. Dado que se ha demostrado que *trh* es importante en el desarrollo traqueal en la mosca de la fruta, pero no en la formación del sistema nervioso, esta evidencia sugiere que, en algún momento durante la evolución de los vertebrados, *NPAS3* adquirió una nueva función en el desarrollo del sistema nervioso.

En los seres humanos, las deleciones en el cromosoma 14 que afectan a *NPAS3* se relacionaron con la holoprosencefalia y la región genómica que incluye *NPAS3* se denominó HP8 porque las personas portadoras de deleciones en esta región presentan defectos cerebrales y craneofaciales que caracterizan la HPE (Kamnasaran et al. 2005). En este trabajo, al utilizar los métodos de generación de *knockdown* y *knockout*, mostramos que la ausencia o reducción de la expresión de *npas3* afecta gravemente el desarrollo normal y conduce a una viabilidad muy reducida en peces cebras. Además, demostramos que *npas3* es clave para el desarrollo normal del sistema nervioso y también para la formación de estructuras craneofaciales. Aunque a través de la generación de modelos de ratón mutante *Npas3* se ha demostrado que *Npas3* es importante para el desarrollo del sistema nervioso y los pulmones, no se informaron defectos craneofaciales (Brunskill et al. 2005; Erbel-Sieler et al. 2004; Stanco et al. 2014; Zhou et al. 2009). Se informaron problemas de desarrollo cerebral en dos de los modelos de ratón, que incluyen un cíngulo y una corteza retrosplenial anormales y reducción de tamaño del hipocampo anterior (Brunskill et al. 2005; Erbel-Sieler et al. 2004). También se ha informado de que los ratones que carecen de *Npas3* mostraron una neurogénesis reducida en el hipocampo (Pieper et al. 2005) y también en la eminencia ganglionar (Stanco et al. 2014). Además, los ratones *Npas3*^{-/-} tenían una densidad reducida de interneuronas corticales y volumen reducido de ganglios

corticales y basales (Stanco et al. 2014). En este trabajo demostramos que *npas3* se expresa fuertemente en estructuras craneofaciales, incluyendo los arcos faríngeos del viscerocráneo y el neurocráneo en el pez cebra en desarrollo. Además, mostramos aquí que las larvas tratadas con morfolinós (*knockdown*) y con deleciones en el gen *npas3* (*knockout*) de 5-dpf exhibieron atrofas y malformaciones graves en sus elementos viscerocraneales cartilaginosos, incluida la ausencia casi total de todos los arcos faríngeos. Por otro lado, los componentes cartilaginosos del neurocráneo se vieron menos afectados por las *npas3*-MO ya que estas estructuras se formaron, pero con alteraciones en su morfología. Por lo tanto, en este trabajo, mostramos que *npas3* juega un papel clave en el desarrollo de estructuras craneofaciales, que se derivan parcialmente de la cresta neural en el pez cebra.

Con el fin de identificar los mecanismos moleculares subyacentes a los efectos fenotípicos observados en los peces que carecen de *npas3*, analizamos la expresión de los genes que se ha informado que están regulados por *NPAS3* / *trh* en otros organismos. En particular, analizamos la expresión de los genes que participan en la vía de *FGF*. Encontramos un aumento y una mala localización de la expresión de *fgfr1* en el sistema nervioso en desarrollo y en los arcos branquiales de los peces morfantes *npas3* en comparación con los peces tratados con vehículo y los de tipo salvaje a 3 y 5 dpf. No detectamos cambios robustos en otros genes en las etapas analizadas, aunque nuestros resultados no excluyeron las posibilidades de que otros genes de la vía del *FGF* estén mediando los efectos de *npas3* en el desarrollo del pez cebra. Por otro lado, analizamos la expresión de *shha* en el pez cebra tratado con MO. Observamos diferencias en la localización del neuroectodermo dorsal y la reducción de la expresión de *shha* en el neuroectodermo ventral a 24 hpf. Más adelante, encontramos que a 48 hpf hay un aumento en la expresión de *shha* en el ectodermo oral en peces tratados con MO. También observamos que hay una expresión aumentada de *shha* en estructuras viscerocraneales y neurocraneales y también en la boca en desarrollo de animales inyectados con MO1 o MO2 de 72 hpf, en comparación con los controles y los peces de tipo salvaje. Estos resultados muestran por primera vez que el *npas3* controla el patrón de expresión de *shha* en el cerebro en desarrollo y en las estructuras craneofaciales derivadas de la cresta neural. Se ha demostrado previamente que los ratones nulos *Npas3* muestran una reducción en la expresión de

Shh en los pulmones (Zhou et al. 2009). También se ha demostrado que *SHH* desempeña una función clave en el desarrollo de estructuras craneofaciales que incluyen la cara anterior y el cráneo (Cordero et al. 2004). La pérdida de *Shh* o *Smoothened* (*Smo*: otro gen de la misma familia que *Shh* que participa en la vía de señalización *hedgehog*) produce una ausencia casi completa de formación de cartílago (Brand et al. 1996; Chen, Burgess, y Hopkins 2001). Además, se ha demostrado que *Shh* desempeña un papel clave en el desarrollo del cerebro anterior, incluido el establecimiento de la línea media ventral, la separación de los ojos y el patrón de telencéfalo (Chiang et al. 1996; Shimamura y Rubenstein 1997). *SHH*, localizado en la región cromosómica 7q36 (OMIM600725), se incluye entre los cuatro principales genes causantes de HPE (Belloni et al. 1996; Erich Roessler et al. 1996; E. Roessler et al. 1997).

Nuestros resultados preliminares obtenidos con ratones mutantes para el gen *Npas3* parecen mostrar defectos en el desarrollo del cráneo, aunque no observamos a la edad del desarrollo analizadas defectos craneofaciales. Si bien en ninguno de los tres modelos de ratones mutantes de *Npas3* generados hasta el momento (Brunskill et al. 2005; Erbel-Sieler et al. 2004; Zhou et al. 2009) se reportaron defectos craneofaciales desconocemos si este aspecto no fue analizado o no fue encontrado. Nuestros experimentos fueron diseñados específicamente para estudiar defectos craneofaciales a través de la tinción de cartílago. Sin embargo, el bajo número de animales mutantes obtenidos previene de sacar conclusiones sobre este aspecto. La generación y análisis de más mutantes se encuentra en marcha en nuestro laboratorio.

En resumen, nuestro estudio revela por primera vez que *npas3* desempeña un papel crucial en el desarrollo de las estructuras craneofaciales, lo que sugiere que *NPAS3* podría ser el gen responsable de las malformaciones cerebrales y craneofaciales que muestran a las personas portadoras de una delección en la región genómica HP8. Sin embargo, los experimentos a completarse en un modelo de mamífero serán claves para ampliar conclusiones sobre este aspecto. Además, encontramos que en el pez cebra dónde la expresión de *npas3* se halla sensiblemente disminuida, la expresión del gen *shha* se ve alterada. Estos resultados sugieren que los defectos en la formación del cerebro y de las estructuras craneofaciales observados en los peces que carecen de *npas3* podrían ser una consecuencia de las alteraciones en la expresión de *shha*.

Conclusiones

En la primera parte de esta tesis nos planteamos estudiar de manera comparada y en experimentos de transgénesis estable en peces cebra la función como *enhancers* transcripcionales los elementos acelerados NPAS3-HAEs. Observamos que en comparación con experimentos anteriores de transgénesis transitoria realizados previamente en el laboratorio con los ortólogos humanos de los NPAS3-HAEs, nuestros experimentos en líneas transgénicas estables muestran una recapitulación parcial de los resultados. En efecto, mientras que en experimentos transitorios 11 de los 14 NPAS3/HAEs se comportaban como *enhancers* transcripcionales durante el desarrollo, en experimentos estables sólo 9 lo hacían de la misma manera. Estos resultados, sumados a otros ya obtenidos en el laboratorio en otras líneas de trabajo indican fuertemente que los experimentos en líneas transitorias deben ser tomados con precaución y que más experimentos serán necesarios para determinar las bases moleculares de esta discrepancia.

Identificamos, además, que las secuencias ortólogas de chimpancé de los elementos HAR202 y HACNS658 se comportan como fuertes activadores de la transcripción en claro contraste a la falta de expresión que se observa al analizar las secuencias humanas de estos elementos. Además, encontramos que el elemento HACNS96 de chimpancé no es funcional en ensayos de transcripción a diferencia del elemento humano que se comporta como un *enhancer* fuerte durante el desarrollo del sistema nervioso. Para analizar si los cambios funcionales observados ocurrieron en el linaje humano o en el de chimpancé, realizamos estudios con otra especie de primate antropoide que tuvo un antepasado en común con humanos y chimpancé hace aproximadamente 25 millones de años, el macaco Rhesus. Nuestros resultados indican que en el caso de los HAR202 y HACNS658 la funcionalidad de los elementos estaba conservada en macacos y chimpancé, pero no en humanos. Por otra parte, las secuencias de humano y de macaco del elemento HACNS96, eran funcionales, mientras que la secuencia de chimpancé no se comporta como *enhancer* transcripcional.

Para profundizar nuestros estudios evolutivos y funcionales seleccionamos el elemento HAR202 y probamos la funcionalidad como activadores transcripcionales las regiones ortólogas de ratón, pollo y pez cebra, encontrando que todas se comportan como *enhancers* fuertes en el sistema nervioso durante el desarrollo del pez cebra. Dado que este elemento tiene una base polimórfica en los humanos actuales que es compartida con los humanos ancestrales neandertales y denisovanos (Matthias Meyer et al. 2012; Prüfer et al. 2014), generamos secuencias portadoras de ese cambio nucleotídico ancestral y encontramos que esta secuencia a la que denominamos HAR202Hn se comporta como fuerte activador transcripcional en claro contraste con la secuencia que es mayoritaria en poblaciones actuales de humanos.

Quisimos estudiar también la funcionalidad del elemento HAR202 en un sistema mamífero más parecido al humano. Realizamos entonces experimentos de transgénesis en ratón utilizando en este caso el sistema de promotor mínimo/reportero con el vector hsp68/*lacZ*. Nuestros resultados fueron bastante contrastantes respecto a lo obtenido en pez cebra. En efecto, la región de chimpancé (HAR202Pt) solo dirige la expresión, en el ratón, a la médula espinal, mientras que en peces cebra observamos actividad en todo el sistema nervioso en desarrollo. Además, el elemento humano (HAR202Hs), no solo era capaz de dirigir la expresión a diferencia de lo visto en peces cebra, sino que además lo hacía de manera similar a lo observado al elemento de chimpancé HAR202Pt; es decir, su principal área de expresión es la médula espinal. Por otra parte, el elemento humano arcaico HAR202Hn, fue capaz de recapitular la expresión vista en peces cebra ya que en ratones dirige la expresión al encéfalo (prosencéfalo, mesencéfalo, rombencéfalo), médula espinal y a los primordios de los miembros. Estos resultados contrastantes entre lo obtenido en peces cebra y ratones requerirán que realicemos nuevos ensayos utilizando en ratones el mismo vector que utilizamos para los experimentos en peces cebra, para descartar que las diferencias se deban a los distintos promotores utilizados que pueden no ser funcionales con determinados activadores según ha sido reportado en la bibliografía (Ariza-Cosano et al. 2012).

Por último, nos interesaba conocer el funcionamiento del elemento HAR202 *in vivo* usando la maquinaria propia de expresión del ratón. Los ensayos en peces cebra de este elemento y las evidencias de marcas epigenéticas (Axel Visel et al. 2009; ENCODE Project Consortium 2012; Nord et al. 2013) apoyan la hipótesis de que este elemento regula la expresión de *Npas3*. Para ello generamos un ratón mutante para este elemento (*Npas3*^{ΔHAR202/+} y *Npas3*^{ΔHAR202/ΔHAR202}). Estos ratones mutantes, se desarrollan y se reproducen de manera aparentemente normal, pero en lo inmediato serán estudiados en profundidad para analizar si este enhancer regula la expresión de gen *Npas3*.

En la segunda parte de esta tesis (Capítulo 2) nos propusimos ampliar el conocimiento qué se tiene acerca de las funciones del gen *NPAS3* en el desarrollo. Mediante el uso de morfolidos, inyectados en peces cebra, evaluamos lo que ocurre durante el desarrollo cuando la expresión de *npas3* está disminuida. Observamos que cuándo se interrumpe la traducción de *npas3* se genera un fenotipo más severo que cuándo el *splicing* normal del gen es alterado. Nuestros resultados indican que la disminución de la expresión del gen *npas3* produce serios problemas en el desarrollo que son incompatibles con la viabilidad de los animales. Principalmente observamos la falta de mandíbula y otros defectos craneofaciales. Además, observamos problemas en el desarrollo del encéfalo y de los ojos. Uno de los efectos más interesantes fue la falta

de desarrollo en la vejiga natatoria. Experimentos con ratones *knockout* (Zhou et al. 2009) mostraron que la falta de *Npas3* causa un problema de desarrollo completo de los pulmones. Ha sido postulado que este órgano en tetrápodos es homólogo a la vejiga natatoria de peces (Darwin 1859; Farmer 1997b). Estos resultados indicarían que el gen *npas3* participa en una vía de señalización común en el desarrollo de estos dos órganos. Sin embargo, problemas en el desarrollo de la vejiga natatoria son comunes en experimentos utilizando morfolidos que afectan el desarrollo y más experimentos serán necesarios para determinar si es un problema específico de los peces tratados con morfolidos dirigidos al gen *npas3*.

En este trabajo probamos exitosamente el método de edición génica de CRISPR/cas9 en peces cebras. Nuestras guías diseñadas contra exones del gen *npas3* lograron producir deleciones en las regiones blanco que afectaron su expresión y generaron en los peces un fenotipo similar al obtenido en los experimentos de *knockdown* con morfolidos.

Probamos, además, mediante experimentos de hibridación *in situ* en peces completos como están afectados otros genes del desarrollo cuya expresión ha sido mostrada que es regulada por *Npas3* en ratones (Zhou et al. 2009). Observamos, que las expresiones de los genes *fgfr1* y *shha* están deslocalizada y severamente reducidas. Sugiriendo que estos genes estén siendo regulados directa o indirectamente por *npas3*. Ha sido mostrado que *npas3* aumenta/disminuye la expresión de *fgfr1* y *shh* en ratones en el pulmón (Zhou et al. 2009). *Shh* participa en el desarrollo craneofacial, es uno de los principales genes causantes de HPE (Belloni et al. 1996; E. Roessler et al. 1997)

Por último, para corroborar si los defectos craneofaciales vistos en peces cebras se recapitulan en mamíferos, generamos un ratón mutante portador de una deleción del exón 3 de *Npas3*, que codifica para la región bHLH de unión. Esto genera una proteína trunca. Este ratón fue evaluado a 14.5 dpc, y en principio, no parece tener efectos morfológicos tan profundos como en peces, pero se ve una diferencia en el cráneo en desarrollo en comparación con ratones normales para la expresión de *Npas3*. Deleciones cromosómicas en humanos que incluyen al gen *NPAS3* generan diversos fenotipos craneofaciales que van desde efectos leves a graves en su desarrollo (Kamnasaran et al. 2005). Estos experimentos se hayan en un estado preliminar y un mayor número de animales generados y analizados serán necesarios para concluir acerca del efecto de *npas3* en el desarrollo craneofacial en mamíferos.

En resumen, nuestro trabajo nos permitió identificar dos elementos *NPAS3*-HAEs que perdieron función como enhancers transcripcionales en comparación con las secuencias de nuestro pariente primate más cercano vivo, el chimpancé. Uno de estos elementos el HAR202, presenta también diferencias funcionales entre las secuencias que tienen de la mayoría de los humanos modernos y una proporción de la población que es portadora de la versión de humanos ancestrales. Más experimentos y análisis

serán necesarios para determinar si estos elementos tienen un efecto sobre la transcripción del gen *NPAS3*. Además, será interesante determinar si la versión ancestral presente en un porcentaje de la población fue reintroducida en *H. sapiens* mediante cruza con los denominados humanos ancestrales, neandertales y denisovanos.

Además, encontramos que el gen *npas3*, además de participar en el desarrollo del cerebro y de los pulmones en mamíferos también participa en el desarrollo de estructuras craneofaciales en peces. Es necesario determinar si también participa del desarrollo craneofacial en ratones lo que permitirá identificar a *npas3* cómo el gen responsable de los defectos craneofaciales que se han reportado en humanos portadores de deleciones que incluyen a este gen y a otros genes. Nuestros resultados amplían el conocimiento acerca de las funciones de este factor de transcripción en el desarrollo de los vertebrados.

Materiales y métodos

PRODUCCIÓN DE PECES CEBRA TRANSGÉNICOS

Construcción de los transgenes para pez cebra

Las regiones conservadas *PhastCons* que contienen los HAEs de *NPAS3* (y las respectivas regiones ortólogas) se clonaron a partir de ADN genómico humano, de chimpancé, macaco (*Macaca mulatta*), ratón, pez cebra, pollo en el vector de entrada pENTR/D (Invitrogen, Thermo Fisher). Las secuencias de los primers usados fueron las que se detallan a continuación. Al extremo 5' de los *primers forward* se les agregó la secuencia CACC que permite el clonado direccionado de los insertos en el vector pENTR/D

TgHAR21Hs_primer_F	5'- <u>CACCT</u> CATCTCAAGGGCTTGCTTT-3'
TgHAR21Hs_primer_R	5'-TGCTTGTGTACCACATGCTTT-3'
TgHAR21Pt_primer_F	5'- <u>CACCA</u> ATCATCTCAAGGGCTTGCTTT-3'
TgHAR21Pt_primer_R	5'-TGTGCTTGTGTACCACATGCTTT-3'
TgHAR89Hs_primer_F	5'- <u>CACCA</u> CTGTGAGACCCACCAAAGG-3'
TgHAR89Hs_primer_R	5'-GGTTTGGTTTGATCGAATGG-3'
TgHAR89Pt_primer_F	5'- <u>CACC</u> CTACTGTGAGAACCACCAAAGG-3'
TgHAR89Pt_primer_R	5'-GGTTTGGTTTGATCAAATGG-3'
TgHAR96Hs_608_primer_F	5'- <u>CACCT</u> TGTGTATGTCTATCTAAAATGGAAAA-3'
TgHAR96Hs_608_primer_R	5'-AAGAAGAAACAGCATTCAAGAAA-3'
TgHAR96Pt_608_primer_F	5'- <u>CACCT</u> TGTGTATGTCTATCTAAAATGGAAAA-3'
TgHAR96Pt_608_primer_R	5'-AAGAAGAAACAGCATTCAAGAAA-3'
TgHAR173Hs_primer_F	5'- <u>CACCT</u> GAGGATGAAAGAAATGAATTT-3'
TgHAR173Hs_primer_R	5'-TTCCGTAATACCTTCTGAGTGTT-3'
TgHAR173Pt_primer_F	5'- <u>CACCT</u> GAGGATGAAAGAAATGAATTT-3'
TgHAR173Pt_primer_R	5'-GCTGACAGTACCGTTATTCGTA3'
TgHAR202Hs_primer_F:	5'- <u>CACCC</u> ACCAGATTCCCTTTGCATT-3'
TgHAR202Hs_primer_R:	5'-CAAACAGCAGGAGACGCTAA-3'
TgHAR202Pt_primer_F:	5'- <u>CACCA</u> CTACTGACTGTAATTTACTTT-3'
TgHAR202Pt_primer_R:	5'-CAAACAGCAGGAGACGCTAA-3'
TgHAR202Mmul_primer_F:	5'- <u>CACCG</u> TCATAAACACACAGTTTA-3'
TgHAR202Mmul_primer_R:	5'-CAGGCAGGGATATTTCAATT-3'
TgHAR202mm_primer_F:	5'- <u>CACCG</u> CACACAGTTTATGGTACCG-3'
TgHAR202mm_primer_R:	5'-CCTCCCCGCCACAGGCAGGG-3'
TgHAR202Gg_primer_F:	5'- <u>CACCA</u> TGCCCTGCTTGTGAGAAA-3'
TgHAR202Gg_primer_R:	5'-AATGGAAGTTTTGTGGGCAC-3'
TgHAR202Dr_primer_F:	5'- <u>CACCA</u> GTGTTGTGGTCTGGCTGCT-3'
TgHAR202Dr_primer_R:	5'-GCGTCTGTTATGTGCTGCTC-3'

TgHACNS96HPt__primer_F	5'- <u>CACCCGTTTCCCACTTTGGTACA</u> -3'
TgHACNS96Pt__primer_R:	5'-ATTTCCCTCTGCCTGGTCTT-3'
TgHACNS96Mmul_415_primer_F	5'- <u>CACCCGTTTCCCACTTTGGTACA</u> -3'
TgHACNS96Mmul_415_primer_R	5'-ATTTCCCTCTGCCTGGTCTT-3'
TgHACNS221Hs_primer_F	5'- <u>CACCGCACATGAATCAGTCTTTGTCA</u> -3'
TgHACNS221Hs_primer_R	5'-GCAGCTATTTCTGATGCTGAG-3'
TgHACNS221Pt_primer_F	5'- <u>CACCGCACATGAATCAGTCTTTGTCA</u> -3'
TgHACNS221Pt_primer_R	5'-GCAGCTATTTCTGATGCTGAG-3'
TgHACNS658Hs_primer_F	5'- <u>CACCAATCAGCAGCCATCAGTTTA</u> -3'
TgHACNS658Hs_primer_R	5'-TCCCCGCTCCCTTTAGATG-3'
HACNS658Pt_F	5'- <u>CACCAATCAGCAGATCAGTTTA</u> -3'
HACNS658Pt_R	5'-GAAGAGTCACAGTAAGAAAT-3'
HACNS658Mmul_F	5'- <u>CACCAATCAGCAGATCAGTTTA</u> -3'
HACNS658Mmul_R	5'-GAAGAGTCACAGTAAGAAAT-3'
2xHAR122Hs_F	5'- <u>CACCTGGATGTGGCACTTTCAGAG</u> -3'
2xHAR122Hs_R	5'-TATTACTTGGCTGCGGATCG-3'
2xHAR122Pt_F	5'- <u>CACCTGGATGTGGCACTTTCAGAG</u> -3'
2xHAR122Pt_R	5'-TATTACTTGGCTGCGGATCG-3'
TgHACNS490Hs_primer_F	5'- <u>CACCGGAAGTGTGTTTGATTTCGTACC</u> -3'
TgHACNS490Hs_primer_R	5'-GGCAAAATCAAAAAGTTAATTTCA-3'
TgHACNS490Pt_primer_F	5'- <u>CACCCCTCTTGCCCAGATTCATAACATTC</u> -3'
TgHACNS490Pt_primer_R	5'-CCCCAGGTTTGTAAATCGGATGG-3'
TgHACNS553Hs_primer_F	5'- <u>CACCTAGCCTCTAGCCTGGACCAC</u> -3'
TgHACNS553Hs_primer_R	5'-GAAAAGGCTGGAAAATCAGG-3'
TgHACNS553Pt_primer_F	5'- <u>CACCAGGCTTAATCCTTTGTAGCTAAATT</u> -3'
TgHACNS553Pt_primer_R	5'-GAAAAGGCTGGAAAATCAGG-3'
Tg2xHAR142Hs_primer_F	5'- <u>CACCTGCAGAAATCAATCTCTACTGG</u> -3'
Tg2xHAR142Hs_primer_R	5'-CCCCTAATTAAATGCACAACCTGA-3'
Tg2xHAR142Pt_primer_F	5'- <u>CACCTGCAGAAATCAATCTCTACTGG</u> -3'
Tg2xHAR142Pt_primer_R	5'-CCCCTAATTAAATGCACAACCTGA-3'
Tg2xHAR157Hs_primer_F	5'- <u>CACCGGCAGAAACCCTAGTTTCTCAA</u> -3'
Tg2xHAR157Hs_primer_R	5'-CCTCCACATCCCACAATTTTC-3'
Tg2xHAR157Pt_primer_F	5'- <u>CACCGGCAGAAACCCTAGTTTCTCAA</u> -3'
Tg2xHAR157Pt_primer_R	5'-CCTCCACATCCCACAATTTTC-3'
Tg2xHAR223Hs_primer_F	5'- <u>CACCAGCCCAGGACTTCCAAGAGT</u> -3'
Tg2xHAR223Hs_primer_R	5'-AACCCCTTCAGTTTTGCCTTA-3'
Tg2xHAR223Pt_primer_F	5'- <u>CACCAGCCCAGGACTTCCAAGAGT</u> -3'
Tg2xHAR223Pt_primer_R	5'-AACCCCTTCAGTTTTGCCTTA-3'

Los fragmentos fueron amplificados en un termociclador (Biorad) usando la polimerasa con actividad correctora de errores Pfx (Invitrogen) junto con la polimerasa altamente

procesiva Taq (producida en el laboratorio) en una relación 1:1 y 0,2 mM de MgSO_4 . El programa usado fue el siguiente:

- 1) Desnaturalización inicial: 95 °C por 5 minutos.
- 2) Desnaturalización: 95 °C por 30 segundos.
- 3) *Annealing*: 60 °C por 30 segundos.
- 4) Extensión: 68 °C por 30 segundos a 1 minuto 30 segundos según el tamaño del fragmento (considerando una actividad polimerasa de 1000 bases por minuto).
- 5) 29 ciclos de los pasos 2 a 4.
- 6) Extensión final: 10 minutos a 68 °C.

Los fragmentos deseados (de largos comprendidos entre ~ 200 a 1100 pares de bases) fueron aislados en gel de agarosa. Se recuperó el ADN purificando las bandas con el kit para extracción de gel QIAquick Gel Extraction Kit (Qiagen). Los insertos resultantes se clonaron direccionadamente por medio del sistema TOPO (Invitrogen) en el vector pENTR/D, que a su vez funciona como vector de entrada para el sistema Gateway (Invitrogen). El sistema TOPO utiliza las enzimas topoisomerasas de *Vaccinia* para recombinar el ADN doble cadena del producto de PCR a los extremos libres del vector pENTR/D. La direccionalidad del clonado está garantizada por la complementariedad de bases del extremo 3' GTGG del vector con la secuencia CACC presente en el inserto (figura 1).

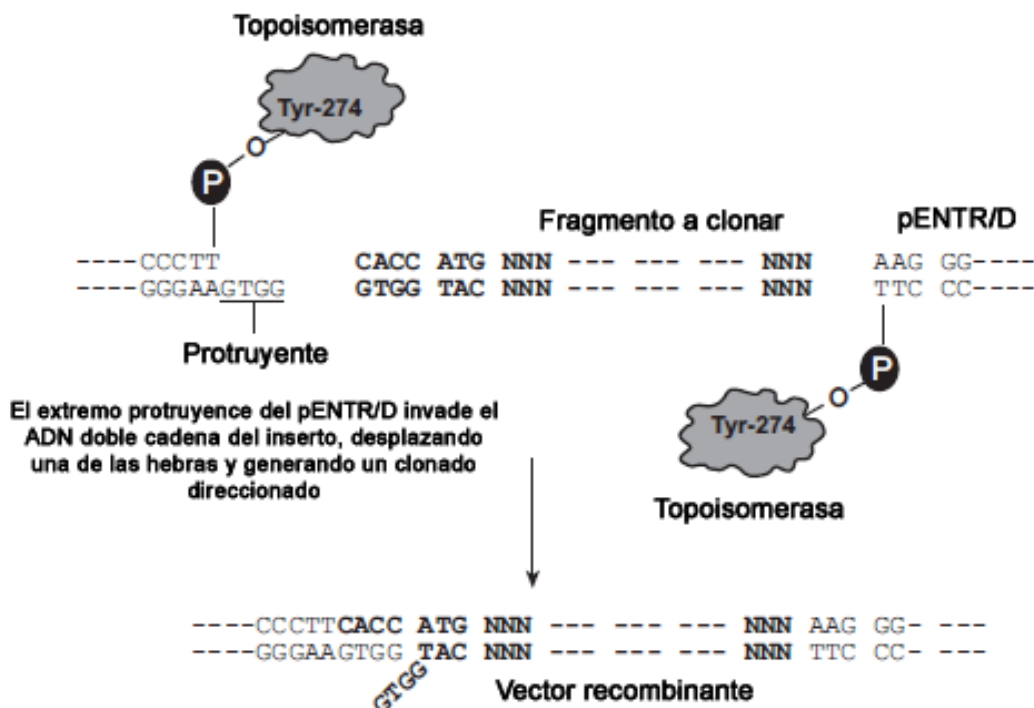


Figura 1: Esquema clonado en el vector pENTR/D por medio del sistema TOPO (Invitrogen, Thermo Fisher). Modificado del manual del uso del vector pENTR/D.

Una vez secuenciados e identificados los clones que contenían los insertos deseados, se realizó el subclonado hacia el vector destino pXIG_cFosGW por medio del sistema Gateway (Invitrogen). Brevemente, el sistema Gateway utiliza las enzimas recombinantes integrasa y excisionasa proveniente del bacteriófago Lambda y la proteína factor de integración de hospedador de *E. coli*. Esta mezcla de enzimas, conocidas como clonasa, catalizan *in vitro* la recombinación entre el inserto presente en el vector de entrada (pENTR/D en nuestro caso) y el vector destino. La recombinación ocurre de forma direccionada entre los sitios attL presentes en el vector de entrada (flanqueando el inserto) y las secuencias attR del vector destino.

En este sistema, las bacterias que incorporan vectores-destino no recombinantes no crecen debido a que, en estas, el vector destino no tiene interrumpida la expresión del gen letal *ccdB*. Así mismo, tampoco crecen las bacterias que incorporan el vector entrada debido a que la resistencia de este vector es kanamicina y la selección se realiza en LB sólido con ampicilina (figura 2). Todos los clonados se realizaron utilizando las cepas de *E. coli* TOP10 (Invitrogen). Esta cepa es susceptible a la proteína letal que interfiere con la girasa de *E. coli* codificada por el locus *ccdB*. Por el contrario, la amplificación del vector pXIG_cFosGW se realizó en la cepa DB3.1 que es resistente al producto del gen *ccdB*.

El vector destino pXIG_cFosGW contiene el promotor mínimo del gen *cFos* de ratón (sin actividad en ausencia de región activadora de la transcripción) río arriba de la secuencia codificante de la proteína reportera fluorescente *Enhanced Green Fluorescent Protein* (EGFP por sus siglas en inglés) (Fisher et al. 2006b). Además, flanqueando el sitio de clonado Gateway, el promotor mínimo y el gen reportero, este vector contiene las secuencias de reconocimiento del transposón *Tol2* del pez medaka (*Orzias latipes*). Estas secuencias son necesarias para el movimiento del transgén y la generación de peces transgénicos (Kawakami et al. 2004).

Se obtuvieron los transgenes para inyección por medio de mini preparaciones de ADN plasmídico de los clones HAE_pXIG_cFosGW deseados. Para ello se utilizó el sistema de

purificación de ADN por columnas *Wizard Plus SV Minipreps DNA Purification System* (Promega).

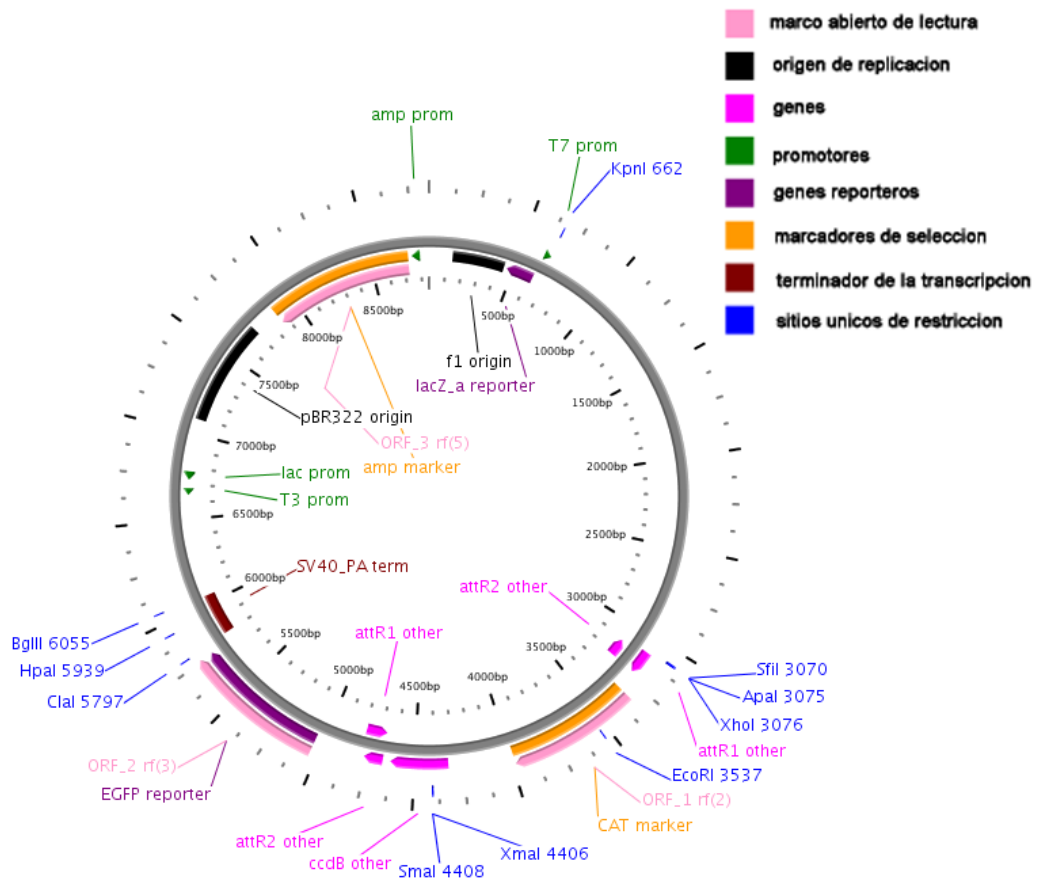


Figura 2: Mapa del vector pXIG_cFosGW (Fisher et al. 2006b).

Transcripción *in vitro* del ARN que codifica para la transposasa *Tol2*

Se preparó una midi preparación de ADN plasmídico del vector pCS-Tp con el *HiSpeed Plasmid Purification kit* (Qiagen). Este vector contiene la secuencia codificante para la transposasa de medaka *Tol2*. Se linealizaron 20 µg de plásmido con la enzima de restricción *NotI* (Invitrogen) en un volumen final de 200 µl. Una vez verificada la linealización completa del vector, se degradó la enzima de restricción con proteinasa K (100 µg ml⁻¹.) Se purificó la preparación de ADN por dos extracciones orgánicas; la primera con 1 volumen de fenol:cloroformo:alcohol isoamílico (25:24:1), la segunda con 1 volumen cloroformo. La fase acuosa resultante se precipitó con acetato de sodio (0.3M) e isopropanol (1 volumen.) El precipitado se lavó con etanol 70% y finalmente se resuspendió en agua libre de nucleasas.

La transcripción *in vitro* se realizó con la ARN polimerasa de fago *sp6* (Roche). Una vez finalizada la reacción de síntesis, el ADN molde se degradó con *DNaseI* y se precipitó el ARN con acetato de litio (0,4 M) y 2 volúmenes de etanol frío. Se resuspendió el ARN en agua libre de nucleasas, se cuantificó y se llevó a una concentración de $175 \text{ ng } \mu\text{l}^{-1}$.

Animales y acuario

Los peces usados en este estudio pertenecen a cepa AB (*Zebrafish* International Resource Center, Universidad de Oregon, Oregon, USA) y a la cepa exocriada obtenida del laboratorio de la Dra. Nora Calcaterra (IBR, Rosario). Los animales se criaron en condiciones estándar en nuestro acuario automatizado modelo ZF0601 (Aquatic Habitats) ubicado en el INGEBI-CONICET. Este sistema mantiene el agua en condiciones óptimas para la cría de los peces mediante diferentes tratamientos: aireación, desgaseado, filtración mecánica (*filter pad*), filtración biológica de amoníaco y nitritos (*Siporax* y *Kaldnes*), control de temperatura, segunda filtración mecánica (50 μm *clartridge*), filtración química por medio de carbón activado y esterilización del agua por UV. Sintéticamente, los animales se mantuvieron en peceras de policarbonato de 3 litros de capacidad, a una densidad máxima de 6 peces adultos por tanque con un flujo de agua de 9 litros por hora (6 recambios del agua de los tanques por hora). Entre otras condiciones, los animales se criaron en un ciclo regular de 14 horas de luz y 10 horas de obscuridad, a $28 \pm 1^\circ\text{C}$, $7 \pm 1,5 \text{ mg/L}$ de densidad de oxígeno, $600 \pm 100 \mu\text{S}$ de conductividad del agua y $7 \pm 1,5$ de pH. Los peces adultos recibieron una dieta de alimento balanceado comercial (Zeigler) dos veces al día suplementada con artemias (INVE).

Obtención de embriones para inyección

El día previo a las inyecciones, se armaron los apareos. Para ello, se usaron unas peceras más pequeñas que las usadas para cría donde hay dos compartimientos separados por una placa divisora de plástico transparente. La base de estas peceras tiene una superficie con ranuras que permite separar los huevos fecundados de los

peces progenitores. Se colocó un macho y una hembra en cada una de las peceras de apareo, dejándolos separados toda la noche por medio de la placa divisoria.

El día de la inyección, poco después del comienzo del ciclo de luz, se removió la división que separa machos y hembras en las peceras de apareo; esto pone en contacto a los animales. En general dentro de los primeros 30 minutos de interacción, comenzaron las primeras puestas de huevos. Las mismas estuvieron compuestas por aproximadamente 150 a 200 embriones cada una.

Inyección de los embriones con la transposasa

Los embriones obtenidos fueron colectados en una placa de Petri de 10 mm y se mantuvieron a 22 °C en medio de embriones (5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl₂, 0.33 mM MgSO₄, 0.00001% peso/volumen de azul de metileno) hasta el momento de inyección.

Se preparó una mezcla fresca de inyección conteniendo:

ADN plasmídico del transgén HAE_ <i>cFos</i> EGFP	125 ng
ARN transposón Tol2 (175 ng μl^{-1})	1 μl
KCl 1 M	0,5 μl
Rojo fenol (2% en agua)	0,5 μl
Agua destilada c.s.p.	5 μl

Los peces transgénicos se generaron microinyectando 1 nl de la mezcla de inyección en el vitelo de embriones en estadio de 1-2 células. Se utilizaron pipetas preparadas a partir de capilares de borosilicato y estiradas con un estirador vertical de micropipetas (*Vertical Pipette Puller*, Model 720, KOPF, David Kopf Instruments, Tujunga, California, USA). Para esto, con la ayuda de una pipeta Pasteur de vidrio, se colocaron los embriones ordenadamente sobre el canto de un portaobjetos apoyado sobre una placa de Petri. Se absorbió el líquido excedente para dejar adheridos los embriones al vidrio del portaobjetos. Las inyecciones se generaron con un microinyector Narishige-IM300 Microinjector (Narishige) conectado a un tanque de N₂ y utilizando un micromanipulador para manejar la pipeta de inyección bajo una lupa Leica wild M3C (Leica).

Luego de la inyección, se identifican los embriones inyectados por la marca roja dejada por el rojo fenol en el vitelo y se retiraron los embriones dañados y/o no inyectados. Se colocaron los embriones inyectados por placa de Petri de 10 mm y se los dejó crecer a 28 °C en medio de embriones hasta la edad deseada en una incubadora. Se dejaron embriones sin inyectar para controlar la viabilidad de los embriones.

Generación de líneas de peces cebras transgénicos

Para la generación de líneas de peces cebras transgénicos, se inyectaron entre 200 a 600 embriones con el transgén de interés como se explicó anteriormente. Los embriones se mantuvieron en placas de Petri con medio de embriones en una incubadora a 28 °C. Al día 5, los embriones fueron trasladados a peceras de cría y mantenidos sin flujo de agua. Se los alimentó dos a tres veces por día con una dieta de larvas especialmente formulada para estadios larvales y juveniles (Larval AP-100, Ziegler, USA). Al décimo día se conectó la pecera de cría al sistema de recirculado de agua bajo un régimen de flujo de agua lento y únicamente durante la noche. Tres días después se suplementó la alimentación de las crías de pez cebra con artemias frescas (INVE). Pasados los quince días, se conectó el tanque de cría al sistema de recirculación de agua a tiempo completo. A los 21- 30 días, se pasaron los animales juveniles a tanques para peces adultos. A los dos meses, los animales entraron en régimen de alimentación de adultos.

Una vez alcanzada la madurez sexual (3 a 4 meses de edad), los animales fueron separados por sexo. Para analizar la expresión de proteína reportera EGFP, se cruzaron los animales adultos inyectados con animales de la línea AB (*wild type*).

Análisis del patrón de expresión de la proteína reportera EGFP

Se utilizó la metodología propuesta por el proyecto CONDOR (Woolfe et al. 2007) para estudiar *in vivo* la actividad *enhancer* de los HAEs. En aquellos casos donde se analizaron embriones a 48 hpf, a las 24 hpf se agregó al medio de embriones 0.1mM de 1-fenil-2- tiourea (Sigma) para prevenir la pigmentación de los animales. Brevemente, se analizó la expresión de EGFP a 24-30 y 48-54 hpf en embriones G0

decorionizados. Para ello, previamente se anestesiaron los embriones con 10 gotas de una solución 0,4% de triclaína (Sigma) por placa de 10 mm. Se montaron los animales en metil celulosa 3% p/v en agua (Sigma) y se registró la expresión de la proteína reportera EGFP usando un microscopio OLYMPUS BX51 acoplado a una cámara digital OLYMPUS DP72 y los filtros adecuados.

Se generaron histogramas de expresión por edad registrando el porcentaje de embriones positivos para la expresión de EGFP en cada una de las siguientes regiones anatómicas (como fueran propuestas en CONDOR): telencéfalo, mesencéfalo, rombencéfalo, médula espinal, ojo, notocorda, corazón, arcos faríngeos/branquiales-faringe, músculo somítico y aletas. La expresión de EGFP por fuera de estos dominios se registró como “otra”. Embriones con expresión de EGFP en 6 o más dominios simultáneamente fueron anotados como con “expresión ubicua”.

Se generaron al menos dos líneas de expresión consistente por transgén.

Genotipificación de los peces

En el caso de los animales que no produjeron expresión de la proteína reportera EGFP, se verificó que la transgénesis haya sido correcta extrayendo el ADN genómico de embriones de 5 días. Para ello, se incubaron las muestras 1-2 horas en agitación a 55°C en 50 µL de *buffer* de lisis (Tris 1 M pH 8.8, EDTA 0.5 M, Tween 20 10% y proteinasa K (20 mg/mL)). Luego de la digestión, se centrifugaron las muestras a velocidad máxima por 20 minutos para eliminar restos de proteínas y SDS. Se cuantificó el ADN con NanoDrop y se diluyó a una concentración de 50 ng/µL. La presencia del transgén de interés fue detectada por medio de una PCR con *primers* específicos para la secuencia de interés.

PRODUCCIÓN DE RATONES TRANSGÉNICOS

Construcción de transgenes para ratón

Se clonaron a partir de ADN genómico humano, de ratón o de chimpancé las regiones ortólogas que contienen el HAE de *NPAS3* HAR202. Las secuencias de los primers usados fueron las mismas que para peces.

Se realizaron subclonados desde los vectores PENTR/D al vector final HSP68/LacZ. Este vector deriva del plásmido pBluescript y contiene el sitio múltiple de clonado río arriba del promotor mínimo del gen HSP68 de ratón regulando la expresión de la proteína reportera LacZ (Figura 3). Se confirmó por restricción y secuenciación la orientación de los clones elegidos. En el caso de las secuencias de humano y chimpancé, estas fueron subclonadas direccionadamente al vector pHSP68/LacZ GW (Poulin et al. 2005) por medio del sistema Gateway como se indicó anteriormente.

Preparación de los transgenes para microinyección pronuclear

Los transgenes para microinyecciones se obtuvieron digiriendo 20 µg de ADN plasmídico con la enzima Sal I. Esta enzima tiene dos sitios de corte en el plásmido pBluescript_HSP68 que aíslan el transgén del vector pBluescript. Se aisló la secuencia del transgén en gel de agarosa y se purificó el ADN por columnas Wizard Plus SV Minipreps DNA Purification System (Promega). Se realizó una segunda purificación usando columnas de intercambio iónico Elutip-D (Schleicher & Schuell). Luego, se recuperó el ADN precipitando con 10% v/v de NaAcO 3M, pH 5,2 y 2,5 volúmenes de etanol 100%. Se realizaron dos lavados con etanol 70% y se resuspendió en buffer de microinyección (Tris-HCl 5mM, pH 7,4; EDTA 0,1mM). La concentración final de ADN de transgén en la solución de microinyección fue de 500 moléculas/pl.

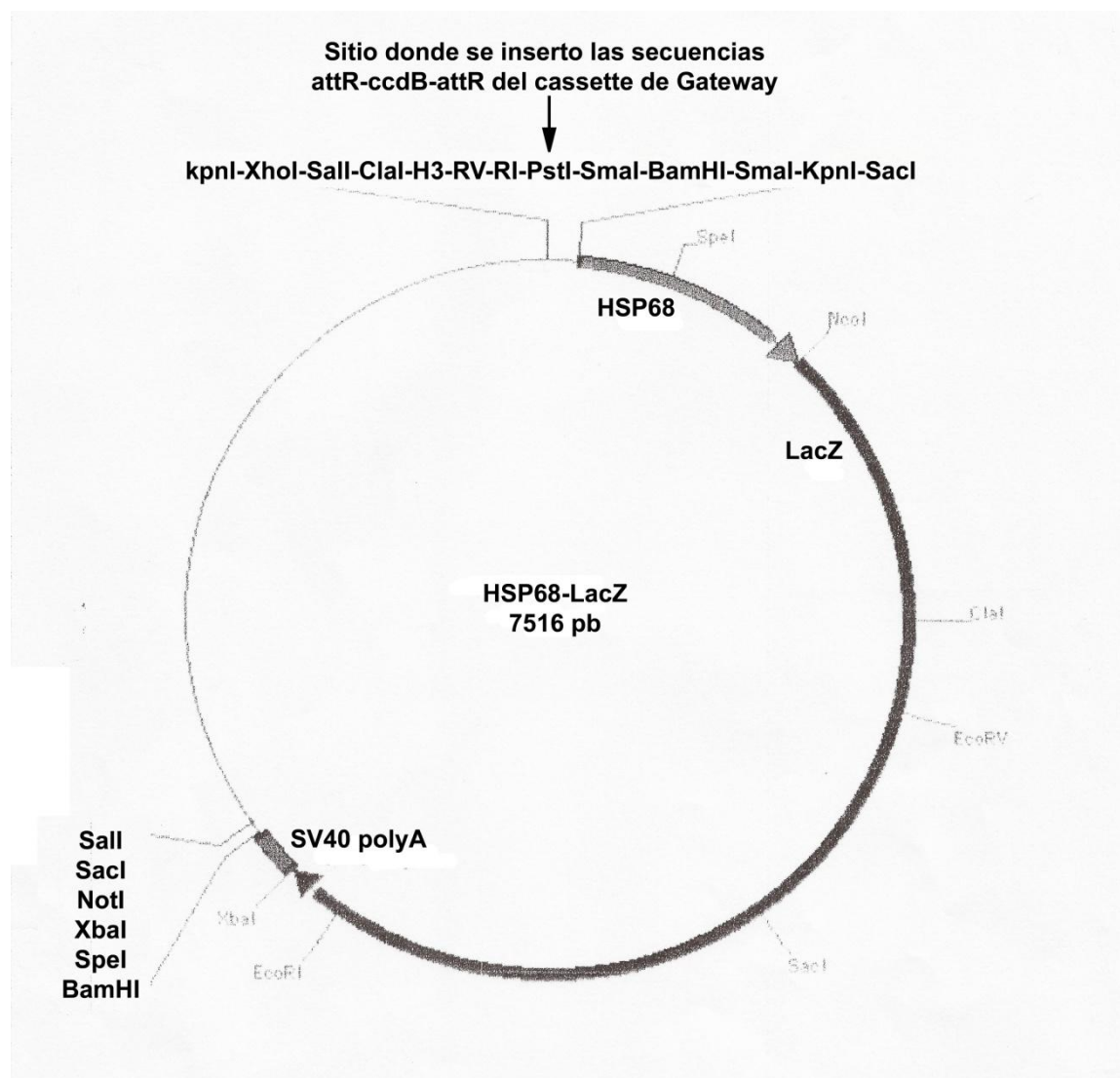


Figura 3: Mapa del vector pHSP68/ LacZGW (Poulin et al. 2005).

Animales y bioterio

Los ratones se alojan en el bioterio que se encuentra en el subsuelo del complejo obligado y que es mantenido por técnicos especializados de los institutos INGEPI e IBYME dependientes del CONICET. Los animales se mantienen en jaulas independientes que cuentan con un sistema de ventilación por inyección de aire filtrado al interior de cada jaula. El sistema de ventilación también extrae aire de las jaulas y lo elimina fuera de la sala de bioterio. Los animales están bajo un régimen de luz/oscuridad de 12 horas, cuentan con una dieta balanceada comercial y la temperatura de la sala donde se encuentran está regulada (20-22 °C).

Producción de embriones para microinyección

Para obtener embriones aptos para microinyección se utilizan animales de la primera generación de la cepa híbrida FVB. Esta cepa se obtiene por cruce de hembras FVB con machos FVB. Hembras FVB prepúberes de 4 y 5 semanas de edad son súper ovuladas mediante un tratamiento hormonal. Este tratamiento consiste en una primera inyección de 5UI i.p. de PMS (*pregnant mare serum*, Sigma) a las 15 hs, cuatro días antes de la recolección de embriones. Dos días después, a las 13 hs, se realiza una segunda inyección de 5 UI i.p. de hGC (*human corionic gonadotrophin*, Sigma.)

Inmediatamente luego de la última inyección, las hembras súper ovuladas se ponen en jaulas de apareo con machos FVB. A la mañana siguiente, las hembras copuladas por el macho (se verifica mediante la presencia de un tapón vaginal de semen) son sacrificadas por dislocación cervical. Se aíslan los oviductos y se colocan en una placa de Petri de 35 mm con medio de cultivo M2 (Sigma, NaCl 88 mM, KCl 4,7 mM, $\text{CaCl}_2 \cdot \text{H}_2\text{O}$ 1,7 mM, KH_2PO_4 1,2 mM, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 1,2 mM, glucosa 5,5 mM, piruvato de sodio 0,33 mM, lactato de sodio 19,8 mM, EDTA 0,1 mM, BSA 0,3%, penicilina 120 UI/ml, sulfato de estreptomicina 50 ug/ml, rojo fenol 0,001%, HEPES 25 mM y pH 7,4). Mediante un par de agujas de 30 G y bajo una lupa, se desgarraron los oviductos para liberar los ovocitos fecundados. Los mismos fueron aspirados con una pipeta Pasteur previamente estirada sobre llama. Los ovocitos junto con las células de granulosa que los rodean fueron colocados en una gota de 90 ul de medio M2 con 1% de hialuronidasa (Sigma). Una vez disgregadas las células de granulosa, se lavaron los embriones mediante 5 - 6 pasajes en medio M2. Finalmente, los embriones listos para la microinyección se mantuvieron en grupos de 15 a 20 en medio M16 (Sigma, NaCl 88 mM, KCl 4,7 mM, KH_2PO_4 1,2 mM, NaHCO_3 23 mM, glucosa 5,5 mM, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 1,2 mM, ácido láctico 4,8 mM, piruvato de sodio 0,2 mM, lactato de sodio 19,8 mM, EDTA 0,1 mM, BSA 0,3%, penicilina 120 UI/ml, sulfato de estreptomicina 50 $\mu\text{g}/\text{ml}$, rojo fenol 0,001%, HEPES 25 mM y pH 7,4) en un incubador a 37 °C y 5% de CO_2 .

Microinyección pronuclear

Los ratones transgénicos fueron generados por microinyección pronuclear de cigotas FVB utilizando micropipetas sujetadoras y de microinyección previamente preparadas a partir de capilares de borosilicato y estiradas con un estirador vertical de micropipetas (Vertical Pipette Puller, Model 720, KOPF, David Kopf Instruments) y un micromanipulador (Leica). Los ovocitos fecundados fueron colocados en una gota de medio M2 ubicada en el medio de la depresión de un portaobjetos cóncavo. Se seleccionaron para inyección aquellos ovocitos que tuvieran ambos pronúcleos visibles utilizando un microscopio invertido con óptica de contraste de interferencia diferencial (FS Labovert, Leica). Sosteniendo el ovocito con la pipeta sujetadora, se inyectó 1 pl de solución de inyección estéril (10 mM Tris, pH 7.5, 0.1 mM EDTA, 100 mM NaCl) conteniendo el ADN del transgén de interés en una concentración de 500 moléculas de ADN por pl. Luego de la microinyección, los embriones se mantuvieron en estufa a 37 °C 10 y 5% de CO₂ hasta el momento de la transferencia.

Transferencia de los embriones inyectados a hembras pseudo preñadas

La noche previa a la microinyección de las cigotas se armaron 12 apareos entre dos hembras jóvenes (de entre 3 y 6 meses de edad) y un macho vasectomizado. A la mañana siguiente, se separaron las hembras pseudo preñadas; que fueron aquellas copuladas por los machos estériles (se verifica nuevamente mediante la presencia de un tapón vaginal de semen). Dado que las hembras utilizadas para pseudo preñar estaban ciclando de forma natural (con un ciclo estral que dura 4/5 días), se calcula que una hembra de cada 4/5 estaba en estro y, por lo tanto, en condiciones de ser copulada.

Finalizada la microinyección, las hembras pseudo preñadas fueron anestesiadas con avertina (300 mg/kg i.p.). Por medio de una incisión lateral, se expuso el oviducto. Luego de desgarrar la membrana que recubre el ovario y el oviducto, se transfirieron 20 a 30 embriones microinyectados. Para ello se introdujo en el ámpula de la hembra una pipeta Pasteur de vidrio adecuadamente estirada a llama cargada con los embriones en medio M2. Finalizada la maniobra se cerró la incisión con sutura.

Las microinyecciones y transferencias a hembras pseudo preñadas fueron realizadas por la técnica Marta Treimun en el laboratorio del Dr. Marcelo Rubinstein (INGEBI-CONICET).

Genotipificación de los ratones transgénicos

Los ratones nacidos de las hembras transferidas fueron apartados de las madres sustitutas a los 20-21 días y separados por sexo. El ADN genómico se extrajo de una muestra de tejido de la oreja. Este material se obtuvo mediante un sacabocado y fue el tejido de descarte que quedó luego del proceso de marcación numérica de los animales para su identificación en el bioterio.

Las muestras de oreja fueron incubadas 12-14 hs en agitación a 55 °C en 127 µl de buffer de digestión (Tris-HCl 50 mM; EDTA 100 mM; SDS 0,5%; 0,5 mg/ml de proteinasa K, pH 8). Luego de la digestión se agregó 40 µl de buffer de precipitación de proteínas (acetato de potasio 3M; pH 5,5). Se centrifugó a máxima velocidad por 10 minutos para eliminar restos de pelos, proteínas y SDS. Se precipitó el ADN con 1 volumen de 2-propanol dejando 2 horas a -20 °C. Luego de centrifugar 10 minutos a máxima velocidad, se lavó el precipitado con 200 µl de etanol 70%. Una vez evaporado el alcohol, se resuspendió el ADN en 20 µl de agua destilada. La cuantificación del ADN se realizó con NanoDrop. La presencia del transgén de interés fue detectada por medio de una PCR con los mismos oligonucleótidos que usamos para levantar los elementos y con otro PCR para lacZ con los siguientes fragmentos:

LacZ-F 5'-TTTCCATGTTGCCACTCGC;

LacZ-R 5'-AACGGCTTGCCGTTAGCA

En todos los casos, se utilizó una polimerasa de ADN *taq* de Invitrogen (Thermo, Brasil).

El programa de amplificación se detalla a continuación:

- 1) Desnaturalización inicial: 95 °C por 5 minutos.
- 2) Desnaturalización: 95 °C por 30 segundos.
- 3) *Annealing*: 60 °C por 30 segundos.
- 4) Extensión: 72 °C por 1 minuto 30 segundos.
- 5) 35 ciclos de los pasos 2 a 4.
- 6) Extensión final: 10 minutos a 72°C.

Productos de amplificación se corrieron en un gel de agarosa 1,5% con bromuro de etidio (0,1 ug/ul de agarosa) y se visualizaron usando luz UV.

Todos los animales fueron tratados respetando los principios de la Guía para el cuidado y uso de animales de laboratorio publicada por el Instituto Nacional de Salud de los Estados Unidos (NIH).

Análisis de la expresión de LacZ en ratones transgénicos

Se analizaron embriones de ratón a diferentes edades del desarrollo embrionario (10.5 a 14.5 dpc). Para ello, se armaron apareos entre un macho adulto positivo para el transgén y dos hembras FVB. A la mañana siguiente, las hembras copuladas por el macho (se verifica mediante la presencia de un tapón vaginal de semen) se separaron del macho y se reservaron en una jaula individual hasta el día deseado de sacrificio (dependiendo de la edad de los embriones a analizar). El día de la recolección de los embriones, las hembras adultas fueron sacrificadas por dislocación cervical. Por medio de una incisión ventral, se aislaron los úteros y se colectaron los embriones. Los embriones fueron lavados en PBS y fijados en una solución de paraformaldehído 4% en PBS a TA y agitación suave. Los tiempos de fijación dependieron de la edad de los embriones: 30 minutos para los embriones 10.5-11.5 dpc, 1 hora para los embriones 12,5 dpc y 1 hora 30 minutos para los embriones 14.5 dpc. Luego de la fijación, se realizaron tres lavados de 30 minutos cada uno en *rinse buffer* (fosfato de sodio 100 mM, MgCl₂ 2 mM, deoxicolato de sodio 0.01%, NP-40 0.02%, pH 7.3). Finalmente, los embriones se dejaron 12-14 horas a 37°C en solución de coloración (*rinse buffer* + 1 mg/ml de 5-bromo-4-cloro-3-indolil-β-D-ácido glucurónico (*X-Gal*), 2,12 mg/ml de ferrocianida de potasio y 1.64 mg/ml de ferricianida de potasio). Al día siguiente, se lavaron los embriones en PBS por 10 minutos. Los embriones fueron clarificados en glicerol 20% en PBS por 12-14 horas a 37 °C y fotografiados en una lupa Leica DFC320. De ser necesario, se realizó un segundo paso de clarificación en glicerol 50% en agua por 12-14 horas a 37 °C.

HIBRIDACIÓN *IN SITU* EN CORTES DE CRIÓSTATO

Las secuencias de las sondas para detección del mensajero de *npas3/Npas3* fueron previamente amplificadas y clonadas en nuestro laboratorio mediante PCR utilizado como molde una solución 1:50 de ADN copia de cerebro de ratón o de ADN copia de embriones de pez cebra de 24 y 48 hpf. Las secuencias de los primers utilizados fueron las siguientes:

Para clonar las sondas contra el ARN mensajero de *NPAS3* de pez cebra:

Npas3_Dr_E2_241_F: 5'-CACCCAGGCTCTGAGGAAGGAGAA-3'

Npas3_Dr_E2_241_R: 5'-CTTTGACGGAGGTGTTGG-3'

Npas3_Dr_EGF_451_F: 5'-CACCTCCGGTTCCTGCTCTAC-3'

Npas3_Dr_EGF_451_R: 5'-TGCGATCCAGAGTCTGTGTG-3'

La amplificación se realizó utilizando la polimerasa de alta fidelidad Pfx (Invitrogen). El programa de amplificación se detalla a continuación:

- 1) Desnaturalización inicial: 95 °C por 5 minutos.
- 2) Desnaturalización: 95 °C por 30 segundos.
- 3) *Annealing*: 64 °C por 30 segundos.
- 4) Extensión: 68 °C por 30 segundos.
- 5) 29 ciclos de los pasos 2 a 4.
- 6) Extensión final: 10 minutos a 72 °C.

Los productos de PCR fueron analizados en gel de agarosa 1,5% y las bandas de interés se aislaron directamente desde el gel. El ADN se purificó utilizando un kit para purificación de ADN desde gel de agarosa (QIAquick Gel Extraction Kit, Qiagen). Luego, los fragmentos de PCR conteniendo la secuencia de la sonda de ratón fueron clonados en el vector pGEM_Teasy (Promega) o en el vector pENTRD (Invitrogen) los fragmentos necesarios para obtener las sondas de pez cebra. Se confirmó la identidad de las secuencias clonadas por secuenciación de los clones aislados con los primers M13F.

Se realizaron midipreparaciones de ADN plasmídico de los clones seleccionados conteniendo las secuencias de interés utilizando el kit HiSpeed Plasmid Purification (Qiagen). Se digirieron 30 µg de ADN plasmídico con las enzimas SpeI (sonda de ratón) o NotI (sonda de pez cebra) por 16 horas a 37 °C. Se confirmó la digestión total de los plásmidos corriendo 1 µl de la solución de digestión en gel de agarosa. Las enzimas

fueron eliminadas utilizando las columnas de purificación de enzimas Micropure-EZ (Millipore) y el ADN se precipitó utilizando 2 volúmenes de etanol y 0,2 volúmenes de acetato de sodio 3 M. Luego de centrifugar, el precipitado se lavó con etanol 70% en agua y el ADN se resuspendió en 30 µl de agua libre de nucleasas. La concentración final de ADN se cuantificó en NanoDrop (Thermo Scientific).

Para la síntesis de las sondas, se utilizó como molde 1 µg de ADN plasmídico linealizado por reacción. Las síntesis y marcación del ARN de las sondas de secuencia antisentido se realizó utilizando el kit DIG RNA Labeling Mix (Roche) y la enzima ARN polimerasa T7 (Roche). El ARN resultante se purificó utilizando el *RNeasy mini kit* (Qiagen). Para obtener la sonda sentido contra el mensajero *Npas3* de ratón, el plásmido se linealizó con la enzima BamHI y la síntesis del ARN se realizó utilizando la enzima Sp6 (Promega). Los productos de síntesis resultantes se analizaron en un gel de agarosa 1% y se cuantificaron en NanoDrop (Thermo scientific). Los largos resultantes de las sondas fueron de aproximadamente de 460 pb para la sonda contra el mensajero de *Npas3* de ratón y de 240 pb (sonda E2) y 450 pb (sonda EG) para las sondas contra el mensajero de *NPAS3* de pez cebra.

Los embriones de ratón o pez cebra de las edades deseadas se fijaron en una solución de paraformaldehído 4% en PBS (buffer fosfato de sodio 20 mM, NaCl 150 mM pH 7.4) por 14-16 horas a 4 °C. Al día siguiente, se lavó el tejido con PBS durante todo el día. Posteriormente, las muestras se equilibraron en una solución de sacarosa al 10% en PBS por 14-16 horas a 4 °C. Los embriones se embebieron en una solución de sacarosa 10%, gelatina 10% en PBS y se congelaron en isopentano enfriado a -55 °C con hielo seco. El material se guardó a -70 °C hasta el momento del corte.

El tejido se cortó en un criostato (Leica) a 18-25 µm los embriones de ratón y a 14-16 µm los embriones de pez cebra, dependiendo de la edad. Los cortes obtenidos se adhirieron a portaobjetos de vidrio Superfrost Plus (Fisher Brand).

Los vidrios se dejaron secar a temperatura ambiente por 30 minutos. Posteriormente el tejido se fijó en una solución de paraformaldehído 4% en PBS por 15 minutos y las muestras se lavaron en PBS 3 veces por 15 minutos. Se realizó un paso de acetilación en una solución con 3,2 ml de trietanolamina, 420 µl ácido clorhídrico 36%, 600 µl de anhídrido acético y 234 ml de agua libre de nucleasas por 10 minutos en agitación suave. Posteriormente, el tejido se permeabilizó tratándolo con una solución de tritón-

X100 1% en PBS por 30 minutos. Paso seguido los cortes se lavaron en PBS y se realizó un paso de pre-hibridación con 400 µl de solución de hibridación (formamida 50%, dextrán sulfato 10%, ARNt de levaduras 1 mg/ml, solución de Denhart's 1 mg/ml en PBS preparado en agua libre de nucleasas) por corte a temperatura ambiente por 4 horas en cámara humidificada. Luego, la solución de prehibridación se reemplazó por 200 µl de solución de hibridación por vidrio conteniendo 1 µl de solución de sonda de ARN marcado con digoxigenina. Los vidrios se cubrieron con un cubreobjeto y se colocaron en una cámara humidificada por 14-16 horas a 72 °C.

Al día siguiente, se realizó un primer lavado en buffer citrato de sodio 0,2X (NaCl 0,3 M y 30 mM de citrato de sodio, pH 7) a 72 °C por 1 hora y un segundo lavado en el mismo buffer por 5 minutos a temperatura ambiente. Los cortes se equilibraron en solución B1 (0,1 M Tris, 0,15 M NaCl, Tritón-X100 0,1%, pH 7,5) 3 veces 5 minutos. El bloqueo se realizó en 500 µl por vidrio de solución B1 con 10% de suero fetal bovino por 4 horas a temperatura ambiente en cámara humidificada. Luego, los vidrios se cubrieron con 500 µl de solución de B1 con 10% de suero fetal bovino y anticuerpo anti-digoxigenina acoplado a fosfatasa alcalina 1/3500 (Roche) por 14-16 horas a temperatura ambiente. Al día siguiente, se realizaron tres lavados en B1 por 10 minutos y los cortes se equilibraron en solución B2 (0,1 M Tris, 0,1 M NaCl, 50 mM MgCl₂, Tween 20 0,1%, pH 9,5). La detección del producto coloreado se realizó reemplazando la solución B2 por 300 µl por vidrio de solución B2 con los sustratos de la fosfatasa alcalina nitroblue tetrazodium (NBT, Roche) y 5-bromo-4cloro-3-indolil fosfato (BCIP, Roche). La reacción se finalizó sumergiendo los vidrios en PBS. Los cortes se montaron en Mowiol (polivinil alcohol, Sigma-Aldrich) y se fotografiaron en un microscopio OLYMPUS BX51 equipado con una cámara OLYMPUS DP72.

HIBRIDACIÓN *IN SITU* EN EMBRIONES COMPLETOS DE PEZ CEBRA

Se sacrificaron aproximadamente 50 embriones de pez cebra de 24, 48 hpf o 5 días en agua helada. El tejido se fijó 14-16 horas en una solución de paraformaldehído 4% en PBS a 4 °C. Luego, los embriones se lavaron en BPT (PBS 1X con 0,1% de Tween-20) y se deshidrataron equilibrando los tejidos en soluciones de concentración creciente de

metanol (25%, 50% y 75%). Finalmente, los embriones se pasaron a metanol 100% y se almacenaron hasta el momento de ser usados a -20 °C.

Los embriones se coloraron en placas multi-well y se rehidrataron pasándolos por soluciones de concentración creciente de PBT (25%, 50% y 75%). Se realizaron dos lavados de 5 minutos con PBT a temperatura ambiente. Los tejidos se permeabilizaron en 500 µl de PBT con 10 µg/ml de proteinasa K por 5 minutos para embriones de 24-48 hpf, 10 minutos para embriones de 72 hpf y 30 minutos para embriones de 5 días o más. Se realizó una post fijación en paraformaldehído 4% en PBS por 10 minutos para embriones de 24 hps, 20 minutos para embriones de 48 y 72 hpf y 30 minutos para embriones de 5 días o más. Posteriormente, se realizaron 5 lavados de 5 minutos cada uno en PBT. La prehibridación se realizó por 1 hora a 65 °C en 500 µl de solución de hibridación (50% formamida, buffer citrato 5X, heparina 50 µg/ml, ARNt de tórula 0,5 mg/ml, Tween-20 0,1%). La hibridación se realizó por 14-16 horas a 65 °C en 300 µl de solución de hibridación con 1 µl/ml de sonda de ARN marcada con digoxigenina.

Al día siguiente, se realizaron una serie de lavados en solución de hibridación con concentraciones crecientes de buffer citrato a 65 °C por 5 minutos (0,6X, 1,2X, 2X). Luego, se hicieron 3 lavados de 20 minutos a 65 °C en buffer citrato con Tween-20 0,1%, uno con buffer citrato 0,2X y dos con buffer citrato 0,1X. Finalmente, se hicieron dos lavados de 5 minutos cada uno a temperatura ambiente con buffer citrato 0,12X y 0,06 en PBT. El bloqueo se realizó por 1 hora a temperatura ambiente en 500 µl de solución de bloqueo (2% de suero normal de pollo, 2 mg/ml de BSA en PBT). Se reemplazó la solución anterior por 300 µl de solución de bloqueo fresca con 1/5000 de anticuerpo anti-digoxigenina acoplado a fosfatasa alcalina (Roche) y se incubó 14-16 horas a 4 °C en agitación suave.

Luego, se hicieron 5 lavados de 15 minutos en PBT. El tejido se equilibró en solución B2 y se rebeló siguiendo el protocolo mencionado anteriormente para las *in situ* en corte de crióstato. La reacción se frenó sumergiendo los embriones en agua estéril y se montaron en una solución de glicerol 70% en agua. Los peces cebra fueron fotografiados en un microscopio OLYMPUS BX51 equipado con una cámara OLYMPUS DP72.

INYECCIÓN DE MORFOLINOS

Se utilizaron dos morfolidos, uno que bloquea la traducción (MO1) y uno que bloquea el *splicing* (MO2) de los ARNm de *NPAS3* de pez cebra (Gene Tools). Las secuencias de los morfolidos utilizados en este estudio fueron:

1) Morfolino de traducción contra el ARNm de *NPAS3* de pez cebra:

MO1: GGGAAGATCCTGATCATCCCCGAGC

2) Morfolinos de *Splicing* contra ARNm de *NPAS3* de pez cebra:

MO2: CCCGATTGCTGTGTAGAAAAACACA

3) Morfolinos control

CO: GGGAAGATCCTGATCATCCCCGAGC

Los morfolidos se reconstituyeron en agua destilada libre de ARNsas y se dejaron a una concentración de stock de 1 mM a -20 °C. El día de la inyección se preparó una mezcla fresca conteniendo:

MO 1 mM	4,0 µl
KCl 1 M	0,5 µl
Rojo fenol (2% en agua)	0,5 µl

Para la puesta a punto de los morfolidos, se inyectaron 1, 2, 2,5 o 5 nl de la mezcla anterior por embrión de en el estadio de 1-2 células, para alcanzar las concentraciones finales de morfolino en los embriones de 0,75, 1,5, 1,875 y 3,75 µM respectivamente. Se inyectaron 100 embriones por concentración y tipo de morfolino. Como control, se inyectaron idénticos volúmenes de solución de inyección, pero reemplazando el morfolino por el buffer en el cual se reconstituyó el morfolino (en este caso, agua libre de ARNsas). Se cuantificó el número total de embriones muertos o con signos de necrosis a 24 hpf para todas las condiciones y se estudió la penetrancia del fenotipo del morfolino para aquellas concentraciones que no elevaban la letalidad/necrosis por sobre los valores observados en los embriones no inyectados.

La concentración final utilizada en todos los ensayos fue de 0.75 µM para los dos morfolidos. Esta concentración no produce una reducción significativa en la supervivencia de los embriones a 24 hpf respecto de los embriones no inyectados tratados en iguales condiciones que los inyectados y tiene niveles altos de penetrancia en el fenotipo (por ejemplos, superiores al 90% en el caso de defectos en el inflado de

la vejiga natatoria). En todos los casos, los dos morfolidos se inyectaron en 100-150 embriones de 1-2 células por réplica durante la misma mañana de inyección y utilizando el mismo lote de reproductores (hembras y machos exocriados). Adicionalmente, se inyectaron 100-150 embriones en iguales condiciones (1 nl) con una solución donde el volumen correspondiente al morfolino fue reemplazado por el buffer de disolución. Adicionalmente, se mantuvo y se analizaron 100-150 embriones que no fueron inyectados como control adicional de la viabilidad. Se analizó el fenotipo de cada morfolino a 3 y 5 dpf.

A la edad deseada, los embriones fueron anestesiados como se detalló previamente y montados en metilcelulosa 2%. Los peces cebras fueron fotografiados en un microscopio OLYMPUS BX51 equipado con una cámara OLYMPUS DP72. Las mediciones de largo del cuerpo, diámetro del ojo y longitud de la cabeza fueron realizadas utilizando el programa Image J. Con los resultados de las tres réplicas se estimó un valor medio y desvío estándar. Se evaluó el ajuste de la serie de datos a la distribución Gaussiana mediante *tests* de Kolmogorov-Smirnov, D'Agostino-Pearson y Shapiro-Wilk. También se evaluó si estas series de datos presentaban homogeneidad de varianzas mediante una prueba de Bartlett. Sin embargo, en todos los casos, alguna de las series de datos no se ajustó a una distribución normal con un valor $p \leq 0,05$ y/o no se pudo comprobar homocedasticidad. Entonces, se realizaron pruebas estadísticas no paramétricas. Para ello, se realizaron contrastes múltiples utilizando la prueba de Kruskal-Wallis seguida de la prueba de contrastes de Dunn's. En todos los casos los valores de $p < 0,05$ fueron considerados significativos. El mismo procedimiento se aplicó para comparar los valores promedio de supervivencia entre tratamientos. Todos los análisis estadísticos se realizaron usando el programa *Graph Pad Prism* 4.00 para Windows (GraphPad Software).

EDICIÓN DEL GENOMA CRISPR / CAS9 EN PECES CEBRA

Los guías (sgRNA) dirigidos a qué exón de *npas3* se diseñaron en el sitio CRISPR del MIT (<http://crispr.mit.edu/>). Las secuencias de guía son las siguientes: e1g1 5'-CAGGATCGCGGTGCGCGCTC-3', e1g2 5'-AGGCTTGACGTCTTACCAAC-3', e4g1 5'-CAATCGGGTCCCAGCGGCGG-3', e4g2 5'-GCGGCGGAGCCCCAATGCTA-3'. La

plantilla de ADN para sgRNAs se obtuvo mediante PCR utilizando primers específicos y el plásmido px330 (Addgene) como molde: Fwd primer = 5'TTAATACGACTCACT ATAGN20gttttagagctagaaatagc3'. La secuencia del promotor mínimo de T7 está río arriba de la secuencia de 20 pb (guía) y una secuencia complementaria al vector de expresión; Rev primer = 5'AAAAGCACCGACTCGGTGCC3'. El kit de transcripción MEGAscript T7 (Ambion, ThermoFisher Scientific) se utilizó para sintetizar las sgRNAs. El ARNm de CAS9 se sintetizó con mMESSAGE mMACHINE® T7 Transcription Kit (ThermoFisher) utilizando un plásmido MLM3613 linealizado con endonucleasa de restricción PmeI (NEB) como molde. Se microinyectaron los huevos (etapa de 1 célula) con 1 nl de una solución que contiene 25 ng / l de sgRNA y 250 ng / µl de mRNA Cas9 (40 mM HEPES pH 7,4, KCl 240 mM, 0,5% de rojo fenol). Las guías del exón 1 se inyectaron conjuntamente, mientras que las guías del exón 4 se analizaron por separado.

Para analizar el genotipo de larvas mutantes, se aisló ADN genómico de embriones individuales según (Foley et al., 2009). En resumen, el pez cebra se digirió individualmente con 20 µl de tampón de extracción por PCR (Tris 10 mM pH 8, EDTA 2 mM, Triton X-100 al 0,2% y Proteinasa K 100 ml / ml) ON a 50 ° C. La inactivación de la proteasa se realizó a 95 ° C durante 5 minutos y luego las muestras se hilaron a 1000 g durante 3 minutos. Se utilizó 1ul de sobrenadante para la amplificación por PCR del exón 1 de NPAS3 con los siguientes cebadores: 5utr_F 5'-GAGCTTGAAAGCTTGGAGCG-3' y i1_R 5'-CAGCTTCACTAAACGGCTGG-3'. Los productos de la PCR se clonaron en el vector pGEMTeasy y se secuenciaron 10 clones de cada mutante analizado

EDICIÓN DEL GENOMA CRISPR / CAS9 EN RATONES

Los guías (sgRNA) dirigidos a qué exón de *npas3* se diseñaron en el sitio CRISPR del MIT (<http://crispr.mit.edu/>). Las secuencias de guía son las siguientes:

Para generar el ratón mutante del HAR202: G1-5' ACTACCTACCGTTCTACAAA, G2-5'TCATTTAAGAAGCTACGGCG

Para generar el ratón mutante del Exón 3 del gen *Npas3*: G1-5'GATGCTGCTCGATCCCGCCG, G2-5'TGAGAGACTTTGCCAACCA

La enzima CAS9 es la misma utilizada que para generar peces mutantes.

La siguiente solución fue preparada para luego ser inyectadas en los cigotos de ratón mediante el mismo proceso que para hacer ratones transgénicos:

ARNm Cas9	100uM
Guía1	50uM
Guía2	50uM
H2O	csp 10ul

La genotipificación de los ratones se realizó utilizando los siguientes primers:

Para el ratón mutante del HAR202: 202KO2_F CCTCCCTGTTCTTTGCACTC

202KO2/3_R GCAGGGGAGATCTCAATGTT

Para el ratón mutante del Exón 3 del gen *Npas3*:

NPASE3genotip_524bp F TCATTCTCTGCACAGCTGGA; NPASE3genotip_524bp R
GAGACACAGACAGACCA

Los ratones inyectados con las guías contra el Exón 3 del gen *Npas3* fueron extraídos a 14.5 dpc.

Los ratones inyectados con las guías contra el *PhastCons* que contiene el HAR202 de ratón. De 17 ratones nacidos, 7 poseían algún tipo de mutación, de diferente tamaño, en su genoma en la región esperada. El ratón 5 tuvo la delección deseada exacta para las guías diseñadas. Este ratón fue utilizado para generar nuestra colonia. Este ratón (macho) no tuvo problemas evidentes para alcanzar la madurez sexual e, incluso reproducirse. Para evitar posibles problemas de efecto mosaico (las secuenciaciones se realizan a partir de ADN extraído de muestras de tejido de la oreja del ratón, lo que nada nos dice del genotipo de las células germinales), se cruzó con un ratón hembra de tipo salvaje o *wild type* y se conservaron los ratones F1 que poseían la delección esperada. A su vez, estos ratones F1 heterocigotas para la delección fueron cruzados entre ellos y se observó que también se pudieron reproducir sin problemas y criar a su prole (F2) hasta la edad del destete. Las camadas F2 contienen generalmente en la misma camada ratones de tipo *wild type* ($NPAS3^{+/+}$), heterocigotas ($NPAS3^{\Delta HAR202/+}$) y mutantes ($NPAS3^{\Delta HAR202/\Delta HAR202}$) para la región ortóloga en ratón del HAR202.

Tinción de azul de Alcian de peces

La tinción de cartílago con azul de Alcian se realizó sobre embriones de pez cebra de 5dpf. Los embriones fueron sacrificados por inmersión en agua helada por 30 minutos. Posteriormente, se fijaron durante toda la noche en una solución de paraformaldeído al 4% en 1 X PBS. Luego, se lavaron tres veces por 10 minutos en una solución de tween20 al 0,1% en 1 X PBS. Se tiñeron por 90 minutos en una solución 0,37% HCl, 70% etanol y 0,1% p/v de azul de Alcian (Sigma-Aldrich, catálogo número 05500-10G). Se lavó tres veces por un minuto con una solución 1% HCl y 70% etanol para eliminar el exceso de colorante. Finalmente, los embriones se clarificaron en una solución 50% glicerol y 0,5% KOH durante toda la noche a 37 °C.

Tinción de azul de Alcian de ratones

Se recolectaron embriones después de la eutanasia de hembras embarazadas y se colocaron en PBS 1 x. A E14.5 no es necesario remover la piel o eviscerar el embrión. Se colocó los embriones en viales de centelleo de vidrio que contengan de etanol 70% para fijar durante la noche a 4 ° C. Se reemplazó el EtOH 70% con EtOH 95% durante 1 h y luego con acetona durante la noche a temperatura ambiente. Se sustituyó la acetona con azul de Alcian durante 30 minutos. Se lavó el embrión con EtOH 70% hasta que el líquido dejó de salir azul. Finalmente se transfirieron los embriones a una solución al 50% de glicerol: 50% (1%) de KOH a temperatura ambiente hasta que el tejido se clarificó. Luego se reemplaza la muestra con glicerol 100% para un almacenamiento a largo plazo y para sacar las fotos.

Gel de electroforesis y western blot

Las transferencias de Western se realizaron de acuerdo con los procedimientos estándar. En resumen, los embriones se anestesiaron en hielo y se homogeneizaron mecánicamente en un buffer de lisis (Tris-HCl 50 mM, pH 7,4, NaCl 150 mM, EDTA 1 mM, Tritón X-100 al 1%) que contenían inhibidores de la proteasa. La concentración de proteína se determinó mediante el ensayo de Bradford. Entre 20 y 50 µg de proteínas

se resolvieron mediante electroforesis en gel de poliacrilamida SDS (SDS-PAGE) y se transfirieron a membranas de nitrocelulosa (Nitrobind, Micron Separations Inc., Westborough, Massachusetts, EE. UU.). La unión inespecífica se bloqueó con leche en polvo no grasa al 2% en TBST (Tris-HCl 50 mM, pH 7,6, NaCl 150 mM, Tween-20 al 0,05%) durante 90 minutos a temperatura ambiente. Las membranas se incubaron durante la noche a 4 ° C con anticuerpos primarios (anti-NPAS3 de conejo, 1: 1000, laboratorio de McKnight; anti-tubulina de conejo, 1: 5000, ab6046, Abcam). Todos los anticuerpos se diluyeron en TBST con leche en polvo sin grasa al 2%. Las bandas inmunorreactivas se detectaron mediante quimioluminiscencia utilizando anticuerpos secundarios conjugados con peroxidasa de rábano picante (IgG de cabra anti-conejo, 1: 3000, ab97051, Abcam) y un kit de detección de ECL SignalFire (Cell Signaling Technology, Danvers, Massachusetts, EE. UU.) Siguiendo las instrucciones del fabricante.

Inmunofluorescencia

Las larvas se anestesiaron en hielo y se fijaron durante 1 h en PFA-PBS al 4% (NaCl 137 mM, KCl 2,7 mM, Na₂HPO₄ 10 mM, KH₂PO₄ 2 mM) a 4°C, seguido de crioprotección en 10% de sacarosa-PBS. La inclusión se realizó en gelatina al 10% -sacarosa-PBS al 10% utilizando críomoldes estándar Tissue-Tek (Sakura Finetek) y los bloques se congelaron rápidamente en isopentano frío. El tejido se almacenó a -80 ° C hasta que se seccionó transversalmente en un criostato a 16µm y se recogió en portaobjetos Superfrost Plus (Fisher Scientific).

Las secciones se bloquearon en KPBS (154 mM NaCl, 16 mM K₂HPO₄, 3,6 mM KH₂PO₄) -0,3% de suero de cabra normal Triton X100-2% durante 90 minutos a temperatura ambiente, seguido de una incubación durante la noche a 4 ° C con anticuerpos primarios diluidos en el mismo tampón. Después de lavar en KPBS (dos veces durante 15 minutos), los portaobjetos se incubaron durante 2 horas a temperatura ambiente con anticuerpos secundarios en KPBS, seguido de un lavado en KPBS (dos veces durante 15 minutos) y se montaron en medio de montaje Vectashield Hardset (Vector Labs). Las imágenes se capturaron con una cámara CCD Olympus DP72

acoplada a un microscopio Olympus BX-51 o un microscopio confocal Nikon Eclipse E880.

Se utilizaron los siguientes anticuerpos y concentraciones: anti-histona de conejo H3 fosfo S10 (1: 500; Abcam, ab5176), anti-HuC / D monoclonal de ratón (1:20; Molecular Probes, A21271), anti-cabra conjugada con Alexa-555 IgG de ratón (1: 1000; Molecular Probes, A21422) e IgG de conejo de cabra conjugada con Alexa-488 (1: 1000; Molecular Probes, A11008).

Todas las imágenes del presente trabajo fueron editadas en Adobe Photoshop CS6

ANÁLISIS BIOINFORMÁTICOS

Adquisición de secuencias

Para estudiar la relación filogenética entre los genes de la familia NPAS3 / trh, descargamos de Ensembl (www.ensembl.org) todas las secuencias de codificación de vertebrados disponibles de varios miembros de la familia de proteínas PAS. Solo las secuencias completas con una alta cobertura de secuenciación se mantuvieron y se curaron manualmente utilizando las secuencias de referencia del GenBank NCBI (www.ncbi.nlm.nih.gov), Búsquedas de BLAST de nucleótidos del archivo de rastreo (blast.ncbi.nlm.nih.gov/Blast.cgi) y las pistas de Genome Browser (genome.ucsc.edu) de UCSC, 'Genes and Gene Predictions', que incluyen predictores de marco de lectura abierta como Genscan, N-Scan y Augustus. Las secuencias codificantes de proteínas de insectos, específicamente la mosca de la fruta (*Drosophila melanogaster*) y el mosquito africano de la malaria (*Anopheles gambiae*) se obtuvieron de Ensembl Metazoa (metazoa.ensembl.org/index.html). Las secuencias de codificación obtenidas se alinearon utilizando MUSCLE (Edgar, 2004) implementado en el software MEGA 5.2 (Tamura et al., 2011).

Análisis filogenético

La filogenia de la familia de proteínas PAS se reconstruyó a partir de un subconjunto

representativo de las secuencias reunidas que comprendía una alineación múltiple de 194 secuencias de longitud completa. La reconstrucción de la filogenia se llevó a cabo utilizando el método de Unión de Vecinos (*neighbour joining*, Saitou y Nei, 1987) con el modelo evolutivo de Jones-Taylor-Thornton (JTT) (Jones et al., 1992) como se implementó en MEGA 5.2. Los sitios de datos faltantes se gestionaron con el parámetro de eliminación parcial con un límite de cobertura del 95%, y se llevó a cabo una prueba Bootstrap de 1000 repeticiones.

Bibliografía

- 1000 Genomes Project Consortium, Goncalo R Abecasis, Adam Auton, Lisa D Brooks, Mark A DePristo, Richard M Durbin, Robert E Handsaker, Hyun Min Kang, Gabor T Marth, y Gil A McVean. 2012. «An integrated map of genetic variation from 1,092 human genomes». *Nature* 491 (7422): 56-65. <https://doi.org/10.1038/nature11632>.
- Ahituv, Nadav. 2012. «Gene Regulatory Sequences and Human Disease». En , 1-17. New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-1683-8_1.
- Ahituv, Nadav, Shyam Prabhakar, Francis Poulin, Edward M. Rubin, y Olivier Couronne. 2005. «Mapping cis-regulatory domains in the human genome using multi-species conservation of synteny». *Human Molecular Genetics* 14 (20): 3057-63. <https://doi.org/10.1093/hmg/ddi338>.
- Ariza-Cosano, Ana, Axel Visel, Len A. Pennacchio, Hunter B. Fraser, José L. Gómez-Skarmeta, Manuel Irimia, y José Bessa. 2012. «Differences in enhancer activity in mouse and zebrafish reporter assays are often associated with changes in gene expression». *BMC Genomics* 13 (1). <https://doi.org/10.1186/1471-2164-13-713>.
- Arnone, Maria I, Ivan J Dmochowski, y Christian Gache. 2004. «Using reporter genes to study cis-regulatory elements.» *Methods in cell biology* 74: 621-52. [https://doi.org/10.1016/S0091-679X\(04\)74025-X](https://doi.org/10.1016/S0091-679X(04)74025-X).
- Atchison, M. 2002. «Enhancers: Mechanisms Of Action And Cell Specificity». *Annual Review of Cell and Developmental Biology* 4 (1): 127-53. <https://doi.org/10.1146/annurev.cellbio.4.1.127>.
- Atchison, Michael L. 1988. «Enhancers: Mechanisms of Action and Cell Specificity». *Annual Review of Cell Biology* 4 (1): 127-53. <https://doi.org/10.1146/annurev.cb.04.110188.001015>.
- Azevedo, Frederico A.C., Ludmila R.B. Carvalho, Lea T. Grinberg, José Marcelo Farfel, Renata E.L. Ferretti, Renata E.P. Leite, Wilson Jacob Filho, Roberto Lent, y Suzana Herculano-Houzel. 2009. «Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain». *Journal of Comparative Neurology* 513 (5): 532-41. <https://doi.org/10.1002/cne.21974>.
- Bacsi, Steven G., y Oliver Hankinson. 1996. «Functional characterization of DNA-binding domains of the subunits of the heterodimeric aryl hydrocarbon receptor complex imputing novel and canonical basic helix-loop-helix protein-DNA interactions». *Journal of Biological Chemistry* 271 (15): 8843-50. <https://doi.org/10.1074/jbc.271.15.8843>.
- Bae, Byoung-Il, Divya Jayaraman, y Christopher A Walsh. 2015. «Genetic changes shaping the human brain.» *Developmental cell* 32 (4): 423-34. <https://doi.org/10.1016/j.devcel.2015.01.035>.
- Balciunas, Darius, Kirk J. Wangenstein, Andrew Wilber, Jason Bell, Aron Geurts, Sridhar Sivasubbu, Xin Wang, et al. 2006. «Harnessing a high cargo-capacity transposon for genetic applications in vertebrates». *PLoS Genetics* 2 (11): 1715-24. <https://doi.org/10.1371/journal.pgen.0020169>.
- Barolo, Scott. 2013. «Shadow enhancers» 34 (2): 135-41. <https://doi.org/10.1002/bies.201100121.Shadow>.
- Bartheld, Christopher S. von, Jami Bahney, y Suzana Herculano-Houzel. 2016. «The search for true numbers of neurons and glial cells in the human brain: A review of 150 years of cell counting». *Journal of Comparative Neurology* 524 (18): 3865-95. <https://doi.org/10.1002/cne.24040>.
- Bedell, Victoria M., Stephanie E. Westcot, y Stephen C. Ekker. 2011. «Lessons from morpholino-based screening in zebrafish». *Briefings in Functional Genomics*. <https://doi.org/10.1093/bfpg/elr021>.
- Begun, David R. 2003. «Planet of the Apes». *Scientific American*, 74-83.
- Belloni, E., M. Muenke, E. Roessler, G. Traverse, J. Siegel-Bartelt, A. Frumkin, H.F. Mitchell, et al. 1996. «Identification of Sonic hedgehog as a candidate gene responsible for holoprosencephaly». *Nature Genetics* 14 (3): 353-56. <https://doi.org/10.1038/ng1196-353>.

- Berghmans, Stephane, John P. Kanki, David Langenau, Karl Hsu, Cicely Jette, Thomas Look, y Rodney Stewart. 2008. «Making waves in cancer research: new models in the zebrafish». *BioTechniques* 39 (2): 227-37. <https://doi.org/10.2144/05392rv02>.
- Bill, B R, A M Petzold, K J Clark, L Schimmenti, y S C Ekker. 2009. «A primer for morpholino use in zebrafish». *Zebrafish* 6 (1): 69-77. <https://doi.org/10.1089/zeb.2008.0555>.
- Bird, Christine P., Barbara E. Stranger, Maureen Liu, Daryl J. Thomas, Catherine E. Ingle, Claude Beazley, Webb Miller, Matthew E. Hurles, y Emmanouil T. Dermitzakis. 2007. «Fast-evolving noncoding sequences in the human genome». *Genome Biology* 8 (6): 1-12. <https://doi.org/10.1186/gb-2007-8-6-r118>.
- Bird, Nathan C., y Paula M. Mabee. 2003. «Developmental Morphology of the Axial Skeleton of the Zebrafish, *Danio rerio* (Ostariophysi: Cyprinidae)». *Developmental Dynamics* 228 (3): 337-57. <https://doi.org/10.1002/dvdy.10387>.
- Boffelli, Dario, Marcelo A. Nobrega, y Edward M. Rubin. 2004. «Comparative genomics at the vertebrate extremes». *Nature Reviews Genetics* 5 (6): 456-65. <https://doi.org/10.1038/nrg1350>.
- Bonifer, C., M. C. Huber, U. Jägle, N. Faust, y A. E. Sippel. 1996. «Prerequisites for tissue specific and position independent expression of a gene locus in transgenic mice». *Journal of Molecular Medicine* 74 (11): 663-71. <https://doi.org/10.1007/s001090050070>.
- Brand, M, C P Heisenberg, R M Warga, F Pelegri, R O Karlstrom, D Beuchle, A Picker, et al. 1996. «Mutations affecting development of the midline and general body shape during zebrafish embryogenesis». *Development (Cambridge, England)* 123 (diciembre): 129-42. <http://www.ncbi.nlm.nih.gov/pubmed/9007235>.
- Britten, R J, y E H Davidson. 1969. «Gene Regulation for Higher Cells: A Theory New facts regarding the organization of the genome». *Science* 25 (3891): 349-57. <https://doi.org/10.1016/j.ympev.2006.05.020>.
- Browning, Sharon R., Brian L. Browning, Ying Zhou, Serena Tucci, y Joshua M. Akey. 2018. «Analysis of Human Sequence Data Reveals Two Pulses of Archaic Denisovan Admixture». *Cell* 173 (1): 53-61.e9. <https://doi.org/10.1016/j.cell.2018.02.031>.
- Brunet, Michel, Franck Guy, David Pilbeam, Daniel E. Lieberman, Andossa Likius, Hassane T. Mackaye, Marcia S. Ponce de León, Christoph P. E. Zollikofer, y Patrick Vignaud. 2005. «New material of the earliest hominid from the Upper Miocene of Chad». *Nature* 434 (7034): 752-55. <https://doi.org/10.1038/nature03392>.
- Brunskill, Eric W., Lisa A. Ehrman, Michael T. Williams, Justin Klanke, Daniel Hammer, Tori L. Schaefer, Renu Sah, Gerald W. Dorn, S. Steven Potter, y Charles V. Vorhees. 2005. «Abnormal neurodevelopment, neurosignaling and behaviour in Npas3-deficient mice». *European Journal of Neuroscience*. <https://doi.org/10.1111/j.1460-9568.2005.04291.x>.
- Brunskill, Eric W., David P. Witte, Andrew B. Shreiner, y S. Steven Potter. 1999. «Characterization of Npas3, a novel basic helix-loop-helix PAS gene expressed in the developing mouse nervous system». *Mechanisms of Development* 88 (2): 237-41. [https://doi.org/10.1016/S0925-4773\(99\)00182-3](https://doi.org/10.1016/S0925-4773(99)00182-3).
- Bulger, Michael, y Mark Groudine. 2011. «Functional and Mechanistic Diversity of Distal Transcription Enhancers». *Cell* 144 (3): 327-39. <https://doi.org/10.1016/j.cell.2011.01.024>.
- Bush, Eliot C., y Bruce T. Lahn. 2008. «A genome-wide screen for noncoding elements important in primate evolution». *BMC Evolutionary Biology* 8 (1): 1-10. <https://doi.org/10.1186/1471-2148-8-17>.
- Butler, Jennifer E.F., y James T. Kadonaga. 2002. «The RNA polymerase II core promoter: A key

- component in the regulation of gene expression». *Genes and Development* 16 (20): 2583-92. <https://doi.org/10.1101/gad.1026202>.
- Cann, Rebecca L., Mark Stoneking, y Allan C. Wilson. 1987. «Mitochondrial DNA and human evolution». *Nature* 325 (6099): 31-36. <https://doi.org/10.1038/325031a0>.
- Capra, John A., Genevieve D. Erwin, Gabriel McKinsey, John L.R. Rubenstein, y Katherine S. Pollard. 2013. «Many human accelerated regions are developmental enhancers». *Philosophical Transactions of the Royal Society B: Biological Sciences* 368 (1632): 20130025. <https://doi.org/10.1098/rstb.2013.0025>.
- Carroll, Sean B. 2003. «Genetics and the making of Homo sapiens». *Nature* 422 (April): 849-57.
- . 2005. «Evolution at Two Levels: On Genes and Form». *PLoS Biology* 3 (7): e245. <https://doi.org/10.1371/journal.pbio.0030245>.
- . 2008. «Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution». *Cell* 134 (1): 25-36. <https://doi.org/10.1016/j.cell.2008.06.030>.
- Chen, W, S Burgess, y N Hopkins. 2001. «Analysis of the zebrafish smoothened mutant reveals conserved and divergent functions of hedgehog activity.» *Development (Cambridge, England)* 128 (12): 2385-96. <http://www.ncbi.nlm.nih.gov/pubmed/11493557>.
- Chiang, Chin, Ying Litington, Eric Lee, Keith E. Young, Jeffrey L Corden, Heiner Westphal, y Philip A. Beachy. 1996. «Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function». *Nature* 383 (6599): 407-13. <https://doi.org/10.1038/383407a0>.
- Chimpanzee Sequencing and Analysis Consortium. 2005. «Initial sequence of the chimpanzee genome and comparison with the human genome». *Nature* 437 (7055): 69-87. <https://doi.org/10.1038/nature04072>.
- Chung, SeYeon, Cy Chavez, y Deborah J. Andrew. 2011. «Trachealess (Trh) regulates all tracheal genes during Drosophila embryogenesis». *Developmental Biology* 360 (1): 160-72. <https://doi.org/10.1016/j.ydbio.2011.09.014>.
- Clark, A. J., P. Bissinger, D. W. Bullock, S. Damak, R. Wallace, C. B.A. Whitelaw, y F. Yull. 1994. «Chromosomal position effects and the modulation of transgene expression». *Reproduction, Fertility and Development* 6 (5): 589-94. <https://doi.org/10.1071/RD9940589>.
- Coe, Bradley P, Kali Witherspoon, Jill A Rosenfeld, Bregje W M van Bon, Anneke T Vulto-van Silfhout, Paolo Bosco, Kathryn L Friend, et al. 2014. «Refining analyses of copy number variation identifies specific genes associated with developmental delay». *Nature Genetics* 46 (10): 1063-71. <https://doi.org/10.1038/ng.3092>.
- Cordero, Dwight, Ralph Marcucio, Diane Hu, William Gaffield, Minal Tapadia, y Jill A. Helms. 2004. «Temporal perturbations in sonic hedgehog signaling elicit the spectrum of holoprosencephaly phenotypes». *Journal of Clinical Investigation* 114 (4): 485-94. <https://doi.org/10.1172/JCI19596>.
- Cranston, Anona, Chen Dong, Jill Howcroft, y A. John Clark. 2001. «Chromosomal sequences flanking an efficiently expressed transgene dramatically enhance its expression». *Gene* 269 (1-2): 217-25. [https://doi.org/10.1016/S0378-1119\(01\)00459-0](https://doi.org/10.1016/S0378-1119(01)00459-0).
- Creyghton, M. P., A. W. Cheng, G. G. Welstead, T. Kooistra, B. W. Carey, E. J. Steine, J. Hanna, et al. 2010. «Histone H3K27ac separates active from poised enhancers and predicts developmental state». *Proceedings of the National Academy of Sciences* 107 (50): 21931-36. <https://doi.org/10.1073/pnas.1016071107>.
- Darwin, Charles. 1859. *On the origin of species by means of natural selection, or preservation of favoured races in the struggle for life*. London: John Murray.

- Davidson, E. H., y Douglas H Erwin. 2006. «Gene Regulatory Networks and the Evolution of Animal Body Plans». *Science* 311 (5762): 796-800. <https://doi.org/10.1126/science.1113832>.
- Deaner, Robert O., Karin Isler, Judith Burkart, y Carel Van Schaik. 2007. «Overall brain size, and not encephalization quotient, best predicts cognitive ability across non-human primates». *Brain, Behavior and Evolution* 70 (2): 115-24. <https://doi.org/10.1159/000102973>.
- DeFelipe, Javier. 2011. «The Evolution of the Brain, the Human Nature of Cortical Circuits, and Intellectual Creativity». *Frontiers in Neuroanatomy* 5 (mayo): 29. <https://doi.org/10.3389/fnana.2011.00029>.
- Dehal, Paramvir, y Jeffrey L Boore. 2005. «Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate». Editado por Peter Holland. *PLoS Biology* 3 (10): e314. <https://doi.org/10.1371/journal.pbio.0030314>.
- Dekker, Job, Karsten Rippe, Martijn Dekker, y Nancy Kleckner. 2002. «Capturing chromosome conformation.» *Science (New York, N.Y.)* 295 (5558): 1306-11. <https://doi.org/10.1126/science.1067799>.
- Dennis, Megan Y, Xander Nettle, Peter H Sudmant, Francesca Antonacci, A Tina, Mikhail Nefedov, Jill A Rosenfeld, et al. 2012. «Human-specific evolution of novel SRGAP2 genes by incomplete segmental duplication». *Cell* 149 (4): 912-22. <https://doi.org/10.1016/j.cell.2012.03.033>. Human-specific.
- Dixon, Jesse R., Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, y Bing Ren. 2012. «Topological domains in mammalian genomes identified by analysis of chromatin interactions». *Nature* 485 (7398): 376-80. <https://doi.org/10.1038/nature11082>.
- Doggett, Norman A., Gary Xie, Linda J. Meincke, Robert D. Sutherland, Mark O. Mundt, Nicolas S. Berbari, Brian E. Davy, et al. 2006. «A 360-kb interchromosomal duplication of the human HYDIN locus». *Genomics* 88 (6): 762-71. <https://doi.org/10.1016/J.YGENO.2006.07.012>.
- Dorsky, Richard I., Laird C. Sheldahl, y Randall T. Moon. 2002. «A Transgenic Lef1/ β -Catenin-Dependent Reporter Is Expressed in Spatially Restricted Domains throughout Zebrafish Development». *Developmental Biology* 241 (2): 229-37. <https://doi.org/10.1006/DBIO.2001.0515>.
- Dryden, Nicola H, Laura R Broome, Frank Dudbridge, Nichola Johnson, Nick Orr, Stefan Schoenfelder, Takashi Nagano, et al. 2014. «Unbiased analysis of potential targets of breast cancer susceptibility loci by Capture Hi-C.» *Genome research* 24 (11): 1854-68. <https://doi.org/10.1101/gr.175034.114>.
- Dulai, K S, M von Dornum, J D Mollon, y D M Hunt. 1999. «The evolution of trichromatic color vision by opsin gene duplication in New World and Old World primates.» *Genome research* 9 (7): 629-38. <http://www.ncbi.nlm.nih.gov/pubmed/10413401>.
- Durbin, Richard M., David L. Altshuler, Richard M. Durbin, Gonçalo R. Abecasis, David R. Bentley, Aravinda Chakravarti, Andrew G. Clark, et al. 2010. «A map of human genome variation from population-scale sequencing». *Nature* 467 (7319): 1061-73. <https://doi.org/10.1038/nature09534>.
- Elston, Guy, Ruth Benavides-Piccione, Alejandra Elston, Paul Manger, y Javier Defelipe. 2011. «Pyramidal cells in prefrontal cortex of primates: marked differences in neuronal structure among species». *Frontiers in Neuroanatomy* 5 (febrero): 2. <https://doi.org/10.3389/fnana.2011.00002>.
- Ema, M, M Morita, S Ikawa, M Tanaka, Y Matsuda, O Gotoh, Y Saijoh, et al. 1996. «Two new members of the murine Sim gene family are transcriptional repressors and show different expression patterns during mouse embryogenesis.» *Molecular and cellular biology* 16 (10): 5865-75. <http://www.ncbi.nlm.nih.gov/pubmed/8927054>.
- Emera, Deena, Jun Yin, Steven K. Reilly, Jake Gockley, y James P. Noonan. 2016. «Origin and evolution of developmental enhancers in the mammalian neocortex». *Proceedings of the National Academy of*

- Sciences* 113 (19): E2617-26. <https://doi.org/10.1073/pnas.1603718113>.
- Enard, Wolfgang, Sabine Gehre, Kurt Hammerschmidt, Sabine M Hölter, Torsten Blass, Mehmet Somel, Martina K Brückner, et al. 2009. «A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice.» *Cell* 137 (5): 961-71. <https://doi.org/10.1016/j.cell.2009.03.041>.
- Enard, Wolfgang, Molly Przeworski, Simon E Fisher, y Cecilia S L Lai. 2002. «Nature01026» 418 (August): 869-72. <https://doi.org/10.1038/nature01025>.
- ENCODE Project Consortium, The ENCODE Project. 2012. «An integrated encyclopedia of DNA elements in the human genome.» *Nature* 489 (7414): 57-74. <https://doi.org/10.1038/nature11247>.
- Erbel-Sieler, C., C. Dudley, Y. Zhou, X. Wu, S. J. Estill, T. Han, R. Diaz-Arrastia, Eric W. Brunskill, S. S. Potter, y Steven L. McKnight. 2004. «Behavioral and regulatory abnormalities in mice deficient in the NPAS1 and NPAS3 transcription factors». *Proceedings of the National Academy of Sciences*. <https://doi.org/10.1073/pnas.0405310101>.
- Ertzer, Raymond, Ferenc Müller, Yavor Hadzhiev, Saradavey Rathnam, Nadine Fischer, Sepand Rastegar, y Uwe Strähle. 2007. «Cooperation of sonic hedgehog enhancers in midline expression». *Developmental Biology*. <https://doi.org/10.1016/j.ydbio.2006.11.004>.
- Eyre-Walker, Adam, y Laurence D. Hurst. 2001. «The evolution of isochores». *Nature Reviews Genetics* 2 (7): 549-55. <https://doi.org/10.1038/35080577>.
- Falk, Dean, John C. Redmond, John Guyer, C. Conroy, Wolfgang Recheis, Gerhard W. Weber, y Horst Seidler. 2000. «Early hominid brain evolution: a new look at old endocasts». *Journal of Human Evolution* 38 (5): 695-717. <https://doi.org/10.1006/JHEV.1999.0378>.
- Farmer, Colleen. 1997a. «Did lungs and the intracardiac shunt evolve to oxygenate the heart in vertebrates?» *Paleobiology* 23 (03): 358-72. <https://doi.org/10.1017/S0094837300019734>.
- . 1997b. «Did lungs and the intracardiac shunt evolve to oxygenate the heart in vertebrates?» *Paleobiology* 23 (03): 358-72. <https://doi.org/10.1017/S0094837300019734>.
- Finnerty, J. R., Kevin Pang, Pat Burton, Dave Paulson, y Mark Q Martindale. 2004. «Origins of Bilateral Symmetry: Hox and Dpp Expression in a Sea Anemone». *Science* 304 (5675): 1335-37. <https://doi.org/10.1126/science.1091946>.
- Finney, Jessica L., George N. Robertson, Chantelle A.S. McGee, Frank M. Smith, y Roger P. Croll. 2006. «Structure and autonomic innervation of the swim bladder in the zebrafish (*Danio rerio*)». *The Journal of Comparative Neurology* 495 (5): 587-606. <https://doi.org/10.1002/cne.20948>.
- Fisher, Shannon, Elizabeth A Grice, Ryan M Vinton, Seneca L Bessling, y Andrew S McCallion. 2006. «Conservation of RET regulatory function from human to zebrafish without sequence similarity.» *Science (New York, N.Y.)* 312 (5771): 276-79. <https://doi.org/10.1126/science.1124070>.
- Fisher, Shannon, Elizabeth A Grice, Ryan M Vinton, Seneca L Bessling, Akihiro Urasaki, Koichi Kawakami, y Andrew S McCallion. 2006. «Evaluating the biological relevance of putative enhancers using Tol2 transposon-mediated transgenesis in zebrafish». *Nature Protocols* 1 (3): 1297-1305. <https://doi.org/10.1038/nprot.2006.230>.
- Fleagle, John G. 2013. *Primate adaptation and evolution*. Elsevier/Academic Press.
- Fortna, Andrew, Young Kim, Erik MacLaren, Kriste Marshall, Gretchen Hahn, Lynne Meltesen, Matthew Brenton, et al. 2004. «Lineage-Specific Gene Duplication and Loss in Human and Great Ape Evolution». Editado por Chris Tyler-Smith. *PLoS Biology* 2 (7): e207. <https://doi.org/10.1371/journal.pbio.0020207>.
- Franchini, L. F., R. Lopez-Leal, S. Nasif, P. Beati, D. M. Gelman, M. J. Low, F. J. S. de Souza, y M. Rubinstein. 2011. «Convergent evolution of two mammalian neuronal enhancers by sequential

- exaptation of unrelated retroposons». *Proceedings of the National Academy of Sciences* 108 (37): 15270-75. <https://doi.org/10.1073/pnas.1104997108>.
- Franchini, Lucía F., y Katherine S. Pollard. 2015. «Genomic approaches to studying human-specific developmental traits». *Development* 142 (18): 3100-3112. <https://doi.org/10.1242/dev.120048>.
- Frankel, Nicolás. 2012. «Multiple layers of complexity in cis-regulatory regions of developmental genes». *Developmental Dynamics* 241 (12): 1857-66. <https://doi.org/10.1002/dvdy.23871>.
- Gabi, Mariana, Christine E. Collins, Peiyan Wong, Laila B. Torres, Jon H. Kaas, y Suzanaerculano-Houzel. 2010. «Cellular Scaling Rules for the Brains of an Extended Number of Primate Species». *Brain, Behavior and Evolution* 76 (1): 32-44. <https://doi.org/10.1159/000319872>.
- Gabi, Mariana, Kleber Neves, Carolinne Masseron, Pedro F M Ribeiro, Lissa Ventura-Antunes, Laila Torres, Bruno Mota, Jon H Kaas, y Suzanaerculano-Houzel. 2016. «No relative expansion of the number of prefrontal neurons in primate and human evolution.» *Proceedings of the National Academy of Sciences of the United States of America* 113 (34): 9617-22. <https://doi.org/10.1073/pnas.1610178113>.
- Gehring, Walter J. 2011. «Chance and Necessity in Eye Evolution». *Genome Biology and Evolution* 3 (enero): 1053-66. <https://doi.org/10.1093/gbe/evr061>.
- Gibbs, R. A., J. Rogers, M. G. Katze, R. Bumgarner, G. M. Weinstock, E. R. Mardis, K. A. Remington, et al. 2007. «Evolutionary and Biomedical Insights from the Rhesus Macaque Genome». *Science* 316 (5822): 222-34. <https://doi.org/10.1126/science.1139247>.
- Gilad, Y., Orna Man, y Gustavo Glusman. 2005. «A comparison of the human and chimpanzee olfactory receptor gene repertoires». *Genome Research* 15 (2): 224-30. <https://doi.org/10.1101/gr.2846405>.
- Gilbert, Sandra L., William B. Dobyns, y Bruce T. Lahn. 2005. «Genetic links between brain development and brain evolution». *Nature Reviews Genetics* 6 (7): 581-90. <https://doi.org/10.1038/nrg1634>.
- Glass, Christopher K., y Michael G. Rosenfeld. 2000. «The coregulator exchange in transcriptional functions of nuclear receptors». *Genes and Development* 14 (2): 121-41. <https://doi.org/10.1101/GAD.14.2.121>.
- Gould, P., y D. Kamnasaran. 2011. «Immunohistochemical Analyses of NPAS3 Expression in the Developing Human Fetal Brain». *Journal of Veterinary Medicine Series C: Anatomia Histologia Embryologia* 40 (3): 196-203. <https://doi.org/10.1111/j.1439-0264.2010.01059.x>.
- Graur, Dan, Yichen Zheng, Nicholas Price, Ricardo B R Azevedo, Rebecca A Zufall, y Eran Elhaik. 2013. «On the immortality of television sets: "function" in the human genome according to the evolution-free gospel of ENCODE.» *Genome biology and evolution* 5 (3): 578-90. <https://doi.org/10.1093/gbe/evt028>.
- Gunz, Philipp, Amanda K Tilot, Katharina Wittfeld, Alexander Teumer, Chin Yang Shapland, Theo G M van Erp, Michael Dannemann, et al. 2019. «Neandertal Introgression Sheds Light on Modern Human Endocranial Globularity.» *Current biology : CB* 29 (1): 120-127.e5. <https://doi.org/10.1016/j.cub.2018.10.065>.
- Hall, Brian K. 2003. «Evo-Devo: evolutionary developmental mechanisms.» *The International journal of developmental biology* 47 (7-8): 491-95. <https://doi.org/10.1387/IJDB.14756324>.
- Hanafusa, Hiroshi, Satoru Torii, Takayuki Yasunaga, y Eisuke Nishida. 2002. «Sprouty1 and Sprouty2 provide a control mechanism for the Ras/MAPK signalling pathway». *Nature Cell Biology* 4 (11): 850-58. <https://doi.org/10.1038/ncb867>.
- Harmon, Katherine. 2012. «New DNA analysis shows ancient humans interbred with Denisovans». *Nature*, agosto. <https://doi.org/10.1038/nature.2012.11331>.

- Hatzis, Pantelis, y Iannis Talianidis. 2002. «Dynamics of enhancer-promoter communication during differentiation-induced gene activation.» *Molecular cell* 10 (6): 1467-77. [https://doi.org/10.1016/S1097-2765\(02\)00786-4](https://doi.org/10.1016/S1097-2765(02)00786-4).
- Heintzman, Nathaniel D, y Bing Ren. 2009. «Finding distal regulatory elements in the human genome.» *Current opinion in genetics & development* 19 (6): 541-49. <https://doi.org/10.1016/j.gde.2009.09.006>.
- Heintzman, Nathaniel D, Rhona K Stuart, Gary Hon, Yutao Fu, Christina W Ching, R David Hawkins, Leah O Barrera, et al. 2007. «Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome.» *Nature Genetics* 39 (3): 311-18. <https://doi.org/10.1038/ng1966>.
- Herculano-Houzel, S. 2012. «The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost.» *Proceedings of the National Academy of Sciences* 109 (Supplement_1): 10661-68. <https://doi.org/10.1073/pnas.1201895109>.
- Herculano-Houzel, S. 2011. «Not All Brains Are Made the Same: New Views on Brain Scaling in Evolution.» *Brain, Behavior and Evolution* 78 (1): 22-36. <https://doi.org/10.1159/000327318>.
- Herculano-Houzel, S, y J Kaas. 2011. «Gorilla and Orangutan Brains Conform to the Primate Cellular Scaling Rules: Implications for Human Evolution.» *Brain, Behavior and Evolution* 77 (1): 33-44. <https://doi.org/10.1159/000322729>.
- Herculano-Houzel, Suzana, Kamilla Avelino-de-Souza, Kleber Neves, Jairo Porfírio, Débora Messeder, Larissa Mattos Feijó, José Maldonado, y Paul R Manger. 2014. «The elephant brain in numbers.» *Frontiers in neuroanatomy* 8: 46. <https://doi.org/10.3389/fnana.2014.00046>.
- Herculano-Houzel, Suzana, Christine E Collins, Peiyan Wong, y Jon H Kaas. 2007. «Cellular scaling rules for primate brains.» *Proceedings of the National Academy of Sciences of the United States of America* 104 (9): 3562-67. <https://doi.org/10.1073/pnas.0611396104>.
- Herculano-Houzel, Suzana, Bruno Mota, y Roberto Lent. 2006. «Cellular scaling rules for rodent brains.» *Proceedings of the National Academy of Sciences of the United States of America* 103 (32): 12138-43. <https://doi.org/10.1073/pnas.0604911103>.
- Heyningen, V. van, y W. Bickmore. 2013. «Regulation from a distance: long-range control of gene expression in development and disease.» *Philosophical Transactions of the Royal Society B: Biological Sciences* 368 (1620): 20120372-20120372. <https://doi.org/10.1098/rstb.2012.0372>.
- Hill, Robert Sean, y Christopher A. Walsh. 2005. «Molecular insights into human brain evolution.» *Nature* 437 (7055): 64-67. <https://doi.org/10.1038/nature04103>.
- Hofman, Michel A. 1988. «Size and Shape of the Cerebral Cortex in Mammals. II. The cortical volume.» *Brain, Behavior and Evolution* 32 (1): 17-26. <https://doi.org/10.1159/000116529>.
- Hong, J.-W., D. A. Hendrix, y M. S. Levine. 2008. «Shadow Enhancers as a Source of Evolutionary Novelty.» *Science* 321 (5894): 1314-1314. <https://doi.org/10.1126/science.1160631>.
- Huang, Jie, Roy H. Perlis, Phil H. Lee, A. John Rush, Maurizio Fava, Gary S. Sachs, Jeffrey Lieberman, et al. 2010. «Cross-Disorder Genomewide Analysis of Schizophrenia, Bipolar Disorder, and Depression.» *American Journal of Psychiatry* 167 (10): 1254-63. <https://doi.org/10.1176/appi.ajp.2010.09091335>.
- Hwang, Woong Y, Yanfang Fu, Deepak Reyon, Morgan L Maeder, Shengdar Q Tsai, Jeffry D Sander, Randall T Peterson, J-R Joanna Yeh, y J Keith Joung. 2013. «Efficient genome editing in zebrafish using a CRISPR-Cas system.» *Nature Biotechnology* 31 (3): 227-29. <https://doi.org/10.1038/nbt.2501>.
- Jacob, François, y Jacques Monod. 1961. «Genetic regulatory mechanisms in the synthesis of proteins».

- Journal of Molecular Biology* 3 (3): 318-56. [https://doi.org/10.1016/S0022-2836\(61\)80072-7](https://doi.org/10.1016/S0022-2836(61)80072-7).
- Jacobs, Gerald H., Maureen Neitz, Jess F. Deegan, y Jay Neitz. 1996. «Trichromatic colour vision in New World monkeys». *Nature* 382 (6587): 156-58. <https://doi.org/10.1038/382156a0>.
- Jin, Fulai, Yan Li, Jesse R. Dixon, Siddarth Selvaraj, Zhen Ye, Ah Young Lee, Chia-An Yen, Anthony D. Schmitt, Celso A. Espinoza, y Bing Ren. 2013. «A high-resolution map of the three-dimensional chromatin interactome in human cells». *Nature* 503 (7475): 290-94. <https://doi.org/10.1038/nature12644>.
- Jolly, Alison. 1972. *The evolution of primate behavior*. Macmillan. https://books.google.com.ar/books/about/The_Evolution_of_Primate_Behavior.html?id=mUUgAQAAIAAJ&redir_esc=y.
- Jurmain, Robert, Lynn Kilgore, Wenda Trevathan, y Russell L. Ciochon. 2014. *Introduction to Physical Anthropology*. Yolanda Cossio.
- Kadonaga, J T. 1998. «Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines.» *Cell* 92 (3): 307-13. <http://www.ncbi.nlm.nih.gov/pubmed/9476891>.
- Kamm, Gretel B., Rodrigo López-Leal, Juan R. Lorenzo, y Lucía F. Franchini. 2013a. «A fast-evolving human NPAS3 enhancer gained reporter expression in the developing forebrain of transgenic mice.» *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 368 (1632): 20130019. <https://doi.org/10.1098/rstb.2013.0019>.
- Kamm, Gretel B., Francisco Pisciotto, R. Kliger, y Lucía F. Franchini. 2013a. «The Developmental Brain Gene NPAS3 Contains the Largest Number of Accelerated Regulatory Sequences in the Human Genome». *Molecular Biology and Evolution* 30 (5): 1088-1102. <https://doi.org/10.1093/molbev/mst023>.
- Kamm, Gretel B., Francisco Pisciotto, Rafi Kliger, y Lucía F. Franchini. 2013b. «The Developmental Brain Gene NPAS3 Contains the Largest Number of Accelerated Regulatory Sequences in the Human Genome». *Molecular Biology and Evolution* 30 (5): 1088-1102. <https://doi.org/10.1093/molbev/mst023>.
- Kamm, Gretel B., Rodrigo López-Leal, Juan R Lorenzo, y Lucía F Franchini. 2013b. «A fast-evolving human NPAS3 enhancer gained reporter expression in the developing forebrain of transgenic mice.» *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 368 (1632): 20130019. <https://doi.org/10.1098/rstb.2013.0019>.
- Kamnasaran, Deepak, Chih Ping Chen, Koenraad Devriendt, Lakshmi Mehta, y Diane W. Cox. 2005. «Defining a holoprosencephaly locus on human chromosome 14q13 and characterization of potential candidate genes». *Genomics* 85 (5): 608-21. <https://doi.org/10.1016/j.ygeno.2005.01.010>.
- Kamnasaran, Deepak, Muir Wj, Ferguson-smith Ma, y Diane W. Cox. 2003. «Disruption of the neuronal PAS3 gene in a family affected with schizophrenia». *J Med Genet* 40: 325-32.
- Karlsson, Johnny, Jonas von Hofsten, y Per-Erik Olsson. 2001. «Generating Transparent Zebrafish: A Refined Method to Improve Detection of Gene Expression During Embryonic Development». *Marine Biotechnology* 3 (6): 0522-27. <https://doi.org/10.1007/s1012601-0053-4>.
- Karro, John E., Yangpan Yan, Deyou Zheng, Zhaolei Zhang, Nicholas Carriero, Philip Cayting, Paul Harrision, y Mark Gerstein. 2007. «Pseudogene.org: a comprehensive database and comparison platform for pseudogene annotation». *Nucleic Acids Research* 35 (suppl_1): D55-60. <https://doi.org/10.1093/nar/gkl851>.
- Kawakami, Koichi. 2005. «Transposon tools and methods in zebrafish». *Developmental Dynamics* 234

- (2): 244-54. <https://doi.org/10.1002/dvdy.20516>.
- Kawakami, Koichi, A. Shima, y Noriko Kawakami. 2000. «Identification of a functional transposase of the Tol2 element, an Ac-like element from the Japanese medaka fish, and its transposition in the zebrafish germ lineage». *Proceedings of the National Academy of Sciences* 97 (21): 11403-8. <https://doi.org/10.1073/pnas.97.21.11403>.
- Kawakami, Koichi, y A Shima. 1999. «Identification of the Tol2 transposase of the medaka fish *Oryzias latipes* that catalyzes excision of a nonautonomous Tol2 element in zebrafish *Danio rerio*.» *Gene* 240 (1): 239-44. <http://www.ncbi.nlm.nih.gov/pubmed/10564832>.
- Kawakami, Koichi, Hisashi Takeda, Noriko Kawakami, Makoto Kobayashi, Naoto Matsuda, y Masayoshi Mishina. 2004. «A Transposon-Mediated Gene Trap Approach Identifies Developmentally Regulated Genes in Zebrafish». *Developmental Cell* 7 (1): 133-44. <https://doi.org/10.1016/j.devcel.2004.06.005>.
- Kewley, Robyn J, Murray L Whitelaw, y Anne Chapman-Smith. 2004. «The mammalian basic helix-loop-helix/PAS family of transcriptional regulators.» *The international journal of biochemistry & cell biology* 36 (2): 189-204. <http://www.ncbi.nlm.nih.gov/pubmed/14643885>.
- Kimmel, Charles B, Bonnie Ullmann, Macie Walker, Craig T Miller, y Justin G Crump. 2003. «Endothelin 1-mediated regulation of pharyngeal bone development in zebrafish.» *Development (Cambridge, England)* 130 (7): 1339-51. <http://www.ncbi.nlm.nih.gov/pubmed/12588850>.
- Kimura, M. 1968. «Evolutionary rate at the molecular level.» *Nature* 217 (5129): 624-26. <http://www.ncbi.nlm.nih.gov/pubmed/5637732>.
- Kimura, Motoo. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511623486>.
- King, M C, y A C Wilson. 1975. «Evolution at two levels in humans and chimpanzees.» *Science (New York, N.Y.)* 188 (4184): 107-16. <http://www.ncbi.nlm.nih.gov/pubmed/1090005>.
- Kingston, R E, y G J Narlikar. 1999. «ATP-dependent remodeling and acetylation as regulators of chromatin fluidity.» *Genes & development* 13 (18): 2339-52. <http://www.ncbi.nlm.nih.gov/pubmed/10500090>.
- Kostka, Dennis, Melissa J. Hubisz, Adam Siepel, y Katherine S. Pollard. 2012. «The role of GC-Biased gene conversion in shaping the fastest evolving regions of the human genome». *Molecular Biology and Evolution* 29 (3): 1047-57. <https://doi.org/10.1093/molbev/msr279>.
- Kothary, R, S Clapoff, S Darling, M D Perry, L A Moran, y J Rossant. 1989. «Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice.» *Development (Cambridge, England)* 105 (4): 707-14. <http://www.ncbi.nlm.nih.gov/pubmed/2557196>.
- Kreitman, Martin. 2000. «Methods to Detect Selection in Populations With Applications to the Human». *Annual Review of Genomics and Human Genetics* 1 (1): 539-59. <https://doi.org/10.1146/annurev.genom.1.1.539>.
- Krivega, Ivan, y Ann Dean. 2012. «Enhancer and promoter interactions-long distance calls». *Current Opinion in Genetics & Development* 22 (2): 79-85. <https://doi.org/10.1016/j.gde.2011.11.001>.
- Kudla, Grzegorz, Leszek Lipinski, Fanny Caffin, Aleksandra Helwak, y Maciej Zylicz. 2006. «High Guanine and Cytosine Content Increases mRNA Levels in Mammalian Cells». Editado por Laurence D. Hurst. *PLoS Biology* 4 (6): e180. <https://doi.org/10.1371/journal.pbio.0040180>.
- Kumar, Sudhir, y S Blair Hedges. 1998. «A molecular timescale for vertebrate evolution». *Nature* 392 (abril): 917. <https://doi.org/10.1038/31927>.
- Lacbawan, F, B D Solomon, E Roessler, K El-Jaick, S Domene, J I Velez, N Zhou, et al. 2009. «Clinical

- spectrum of SIX3-associated mutations in holoprosencephaly: correlation between genotype, phenotype and function». *Journal of Medical Genetics* 46 (6): 389-98.
<https://doi.org/10.1136/jmg.2008.063818>.
- Lander, E S, L M Linton, B Birren, C Nusbaum, M C Zody, J Baldwin, K Devon, et al. 2001. «Initial sequencing and analysis of the human genome». *Nature* 409 (6822): 860-921.
<https://doi.org/10.1038/35057062>.
- Lareau, Liana F., Maki Inada, Richard E. Green, Jordan C. Wengrod, y Steven E. Brenner. 2007. «Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements». *Nature* 446 (7138): 926-29. <https://doi.org/10.1038/nature05676>.
- Leakey, L S, P V Tobias, y J R Napier. 1964. «A New Species of the Genus Homo From Olduvai Gorge.» *Nature* 202 (abril): 7-9. <http://www.ncbi.nlm.nih.gov/pubmed/14166722>.
- Lee, Tong Ihn, y Richard A. Young. 2000. «Transcription of Eukaryotic Protein-Coding Genes». *Annual Review of Genetics* 34 (1): 77-137. <https://doi.org/10.1146/annurev.genet.34.1.77>.
- Lek, Monkol, Konrad J. Karczewski, Eric V. Minikel, Kaitlin E. Samocha, Eric Banks, Timothy Fennell, Anne H. O'Donnell-Luria, et al. 2016. «Analysis of protein-coding genetic variation in 60,706 humans». *Nature* 536 (7616): 285-91. <https://doi.org/10.1038/nature19057>.
- Lemon, B, y R Tjian. 2000. «Orchestrated response: a symphony of transcription factors for gene control.» *Genes & development* 14 (20): 2551-69.
<http://www.ncbi.nlm.nih.gov/pubmed/11040209>.
- Lemons, D., y William McGinnis. 2006. «Genomic Evolution of Hox Gene Clusters». *Science* 313 (5795): 1918-22. <https://doi.org/10.1126/science.1132040>.
- Lettice, Laura A, Simon J H Heaney, Lorna A Purdie, Li Li, Philippe de Beer, Ben A Oostra, Debbie Goode, Greg Elgar, Robert E Hill, y Esther de Graaff. 2003. «A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly.» *Human molecular genetics* 12 (14): 1725-35. <http://www.ncbi.nlm.nih.gov/pubmed/12837695>.
- Lewin, Roger., y Robert. Foley. 2004. *Principles of human evolution*. Blackwell Pub. Co.
- Li, Kun-Lung, Tsai-Ming Lu, y Jr-Kai Yu. 2014. «Genome-wide survey and expression analysis of the bHLH-PAS genes in the amphioxus *Branchiostoma floridae* reveal both conserved and diverged expression patterns between cephalochordates and vertebrates». *EvoDevo* 5 (1): 20.
<https://doi.org/10.1186/2041-9139-5-20>.
- Lieberman-Aiden, Erez, Nynke L van Berkum, Louise Williams, Maxim Imakaev, Tobias Ragoczy, Agnes Telling, Ido Amit, et al. 2009. «Comprehensive mapping of long-range interactions reveals folding principles of the human genome.» *Science (New York, N.Y.)* 326 (5950): 289-93.
<https://doi.org/10.1126/science.1181369>.
- Lin, Yingxi, Brenda L. Bloodgood, Jessica L. Hauser, Ariya D. Lapan, Alex C. Koon, Tae-Kyung Kim, Linda S. Hu, Athar N. Malik, y Michael E. Greenberg. 2008. «Activity-dependent regulation of inhibitory synapse development by Npas4». *Nature* 455 (7217): 1198-1204.
<https://doi.org/10.1038/nature07319>.
- Lindblad-Toh, Kerstin, David B. Jaffe, Tim Massingham, Hiram Clawson, Or Zuk, Pouya Kheradpour, Jiayu Wen, et al. 2011. «A high-resolution map of human evolutionary constraint using 29 mammals». *Nature* 478 (7370): 476-82. <https://doi.org/10.1038/nature10530>.
- Lindebro, M C, L Poellinger, y M L Whitelaw. 1995. «Protein-protein interaction via PAS domains: role of the PAS domain in positive and negative regulation of the bHLH/PAS dioxin receptor-Arnt transcription factor complex.» *The EMBO journal* 14 (14): 3528-39.
<http://www.ncbi.nlm.nih.gov/pubmed/7628454>.

- Liu, Wu, María Martín-Torres, Yan-jun Cai, Song Xing, Hao-wen Tong, Shu-wen Pei, Mark Jan Sier, et al. 2015. «The earliest unequivocally modern humans in southern China». *Nature* 526 (7575): 696-99. <https://doi.org/10.1038/nature15696>.
- Lomvardas, Stavros, Gilad Barnea, David J. Pisapia, Monica Mendelsohn, Jennifer Kirkland, y Richard Axel. 2006. «Interchromosomal Interactions and Olfactory Receptor Choice». *Cell* 126 (2): 403-13. <https://doi.org/10.1016/j.cell.2006.06.035>.
- Longo, Sarah, Mark Riccio, y Amy R McCune. 2013. «Homology of lungs and gas bladders: Insights from arterial vasculature». *Journal of Morphology* 274 (6): 687-703. <https://doi.org/10.1002/jmor.20128>.
- Martin, R. D. 1986. *Primates: a definition*. Cambridge University Press.
- . 1990. *Primate Origins and Evolution. A phylogenetic reconstruction*. *International Journal of Primatology*. Princeton University Press. <https://doi.org/10.1007/BF02547560>.
- Marygold, Steven J., Paul C. Leyland, Ruth L. Seal, Joshua L. Goodman, Jim Thurmond, Victor B. Strelets, y Robert J. Wilson. 2012. «FlyBase: improvements to the bibliography». *Nucleic Acids Research* 41 (D1): D751-57. <https://doi.org/10.1093/nar/gks1024>.
- Massari, M E, y C Murre. 2000. «Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms.» *Molecular and cellular biology* 20 (2): 429-40. <http://www.ncbi.nlm.nih.gov/pubmed/10611221>.
- Maston, Glenn A., Sara K. Evans, y Michael R. Green. 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>.
- Matys, V, E Fricke, R Geffers, E Gössling, M Haubrock, R Hehl, K Hornischer, et al. 2003. «TRANSFAC: transcriptional regulation, from patterns to profiles.» *Nucleic acids research* 31 (1): 374-78. <http://www.ncbi.nlm.nih.gov/pubmed/12520026>.
- McHenry, H M. 1975. «Fossils and the mosaic nature of human evolution.» *Science (New York, N.Y.)* 190 (4213): 425-31. <http://www.ncbi.nlm.nih.gov/pubmed/809842>.
- . 1982. «The Pattern of Human Evolution: Studies on Bipedalism, Mastication, and Encephalization». *Annual Review of Anthropology* 11 (1): 151-73. <https://doi.org/10.1146/annurev.an.11.100182.001055>.
- McLean, Cory Y., Philip L. Reno, Alex A. Pollen, Abraham I. Bassan, Terence D. Capellini, Catherine Guenther, Vahan B. Indjeian, et al. 2011. «Human-specific loss of regulatory DNA and the evolution of human-specific traits». *Nature* 471 (7337): 216-19. <https://doi.org/10.1038/nature09774>.
- Meissner, Alexander, Tarjei S. Mikkelsen, Hongcang Gu, Marius Wernig, Jacob Hanna, Andrey Sivachenko, Xiaolan Zhang, et al. 2008. «Genome-scale DNA methylation maps of pluripotent and differentiated cells». *Nature* 454 (7205): 766-70. <https://doi.org/10.1038/nature07107>.
- Meyer, M., M. Kircher, M.-T. Gansauge, H. Li, F. Racimo, S. Mallick, J. G. Schraiber, et al. 2012. «A High-Coverage Genome Sequence from an Archaic Denisovan Individual». *Science* 338 (6104): 222-26. <https://doi.org/10.1126/science.1224344>.
- Meyer, Matthias, Martin Kircher, Marie-Theres Gansauge, Heng Li, Fernando Racimo, Sapan Mallick, Joshua G Schraiber, et al. 2012. «A high-coverage genome sequence from an archaic Denisovan individual.» *Science (New York, N.Y.)* 338 (6104): 222-26. <https://doi.org/10.1126/science.1224344>.
- Mifsud, Borbala, Filipe Tavares-Cadete, Alice N Young, Robert Sugar, Stefan Schoenfelder, Lauren Ferreira, Steven W Wingett, et al. 2015. «Mapping long-range promoter contacts in human cells

- with high-resolution capture Hi-C». *Nature Genetics* 47 (6): 598-606.
<https://doi.org/10.1038/ng.3286>.
- Nasevicius, Aidan, y Stephen C. Ekker. 2000. «Effective targeted gene 'knockdown' in zebrafish». *Nature Genetics* 26 (2): 216-20. <https://doi.org/10.1038/79951>.
- Ni, J, Karl J. Clark, SC Fahrenkrug, y Stephen C. Ekker. 2008. «Transposon tools hopping in vertebrates». *Brief Funct Genomic Proteomic* 7 (6): 444-453. <https://doi.org/10.1093/bfpg/eln049>.
- Noonan, James P., y Andrew S. McCallion. 2010. «Genomics of Long-Range Regulatory Elements». *Annual Review of Genomics and Human Genetics* 11 (1): 1-23. <https://doi.org/10.1146/annurev-genom-082509-141651>.
- Noonan, James P. 2009. «Regulatory DNAs and the evolution of human development.» *Current opinion in genetics & development* 19 (6): 557-64. <https://doi.org/10.1016/j.gde.2009.10.002>.
- Nord, Alex S., Matthew J. Blow, Catia Attanasio, Jennifer A. Akiyama, Amy Holt, Roya Hosseini, Sengthavy Phouanavong, et al. 2013. «Rapid and pervasive changes in genome-wide enhancer usage during mammalian development». *Cell* 155 (7): 1521-31.
<https://doi.org/10.1016/j.cell.2013.11.033>.
- O'Bleness, Majesta, Veronica B. Searles, Ajit Varki, Pascal Gagneux, y James M. Sikela. 2012. «Evolution of genetic and genomic features unique to the human lineage». *Nature Reviews Genetics* 13 (12): 853-66. <https://doi.org/10.1038/nrg3336>.
- Oberheim, Nancy Ann, Steven A. Goldman, y Maiken Nedergaard. 2012. «Heterogeneity of Astrocytic Form and Function». En *Methods in molecular biology (Clifton, N.J.)*, 814:23-45.
https://doi.org/10.1007/978-1-61779-452-0_3.
- Ohshiro, T, Y Emori, y K Saigo. 2002. «Ligand-dependent activation of breathless FGF receptor gene in Drosophila developing trachea.» *Mechanisms of development* 114 (1-2): 3-11.
<http://www.ncbi.nlm.nih.gov/pubmed/12175485>.
- Ohshiro, T, y K Saigo. 1997. «Transcriptional regulation of breathless FGF receptor gene by binding of TRACHEALESS/dARNT heterodimers to three central midline elements in Drosophila developing trachea.» *Development (Cambridge, England)* 124 (20): 3975-86.
<http://www.ncbi.nlm.nih.gov/pubmed/9374395>.
- Ong, Chin-Tong, y Victor G. Corces. 2011. «Enhancer function: new insights into the regulation of tissue-specific gene expression». *Nature Reviews Genetics* 12 (4): 283-93.
<https://doi.org/10.1038/nrg2957>.
- Ovcharenko, I., Gabriela G Loots, Marcelo A Nobrega, Ross C Hardison, Webb Miller, y Lisa Stubbs. 2005. «Evolution and functional classification of vertebrate gene deserts». *Genome Research* 15 (1): 137-45. <https://doi.org/10.1101/gr.3015505>.
- Pelster, Bernd. 2004. «pH regulation and swimbladder function in fish». *Respiratory Physiology & Neurobiology* 144 (2-3): 179-90. <https://doi.org/10.1016/j.resp.2004.03.019>.
- Pennacchio, Len A., Nadav Ahituv, Alan M. Moses, Shyam Prabhakar, Marcelo A. Nobrega, Malak Shoukry, Simon Minovitsky, et al. 2006. «In vivo enhancer analysis of human conserved non-coding sequences». *Nature* 444 (7118): 499-502. <https://doi.org/10.1038/nature05295>.
- Perry, Michael W, Alistair N Boettiger, Jacques P Bothma, y Michael Levine. 2010. «Shadow enhancers foster robustness of Drosophila gastrulation.» *Current biology : CB* 20 (17): 1562-67.
<https://doi.org/10.1016/j.cub.2010.07.043>.
- Pickard, B S, A Christoforou, P A Thomson, A Fawkes, K L Evans, S W Morris, D J Porteous, D H Blackwood, y W J Muir. 2009. «Interacting haplotypes at the NPAS3 locus alter risk of schizophrenia and bipolar disorder». *Molecular Psychiatry* 14 (9): 874-84.

- <https://doi.org/10.1038/mp.2008.24>.
- Pickard, Ben S., M.P. Malloy, D.J. Porteous, D.H.R. Blackwood, y W.J. Muir. 2005. «Disruption of a brain transcription factor, NPAS3, is associated with schizophrenia and learning disability». *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 136B (1): 26-32. <https://doi.org/10.1002/ajmg.b.30204>.
- Pickard, Benjamin S., Andrew A. Pieper, David J. Porteous, Douglas H. Blackwood, y Walter J. Muir. 2006. «The NPAS3 gene-emerging evidence for a role in psychiatric illness». *Annals of Medicine* 38 (6): 439-48. <https://doi.org/10.1080/07853890600946500>.
- Pickford, Martin, y Brigitte Senut. 2001. «The geological and faunal context of Late Miocene hominid remains from Lukeino, Kenya». *Comptes Rendus de l'Académie des Sciences - Series IIA - Earth and Planetary Science* 332 (2): 145-52. [https://doi.org/10.1016/S1251-8050\(01\)01528-2](https://doi.org/10.1016/S1251-8050(01)01528-2).
- Pieper, Andrew A., Xinle Wu, Tina W Han, Sandi Jo Estill, Quyen Dang, Leeju C Wu, Sarah Reece-Fincannon, et al. 2005. «The neuronal PAS domain protein 3 transcription factor controls FGF-mediated adult hippocampal neurogenesis in mice».
- Pollard, Katherine S., Sofie R. Salama, Bryan King, Andrew D. Kern, Tim Dreszer, Sol J. Katzman, Adam Siepel, et al. 2006. «Forces shaping the fastest evolving regions in the human genome». *PLoS Genetics* 2 (10): 1599-1611. <https://doi.org/10.1371/journal.pgen.0020168>.
- Poulin, Francis, Marcelo A. Nobrega, Ingrid Plajzer-Frick, Amy Holt, Veena Afzal, Edward M. Rubin, y Len A. Pennacchio. 2005. «In vivo characterization of a vertebrate ultraconserved enhancer». *Genomics* 85 (6): 774-81. <https://doi.org/10.1016/j.ygeno.2005.03.003>.
- Prabhakar, S., J. P. Noonan, S. Paabo, y E. M. Rubin. 2006. «Accelerated Evolution of Conserved Noncoding Sequences in Humans». *Science* 314 (5800): 786-786. doi:10.1126/science.1130738. <https://doi.org/10.1126/science.1130738>.
- Prabhakar, S., I. Plajzer-Frick, Len A. Pennacchio, James P. Noonan, M. Shoukry, D. R. FitzPatrick, Axel Visel, et al. 2008. «Human-Specific Gain of Function in a Developmental Enhancer». *Science* 321 (5894): 1346-50. <https://doi.org/10.1126/science.1159974>.
- Preuss, Todd M. 2012. «Human brain evolution: from gene discovery to phenotype discovery.» *Proceedings of the National Academy of Sciences of the United States of America* 109 Suppl (Suppl 1): 10709-16. <https://doi.org/10.1073/pnas.1201894109>.
- Prüfer, Kay, Fernando Racimo, Nick Patterson, Flora Jay, Sriram Sankararaman, Susanna Sawyer, Anja Heinze, et al. 2014. «The complete genome sequence of a Neanderthal from the Altai Mountains». *Nature* 505 (7481): 43-49. <https://doi.org/10.1038/nature12886>.
- Purvis, A. 1995. «A composite estimate of primate phylogeny.» *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 348 (1326): 405-21. <https://doi.org/10.1098/rstb.1995.0078>.
- Raghanti, Mary Ann, Melissa K. Edler, Alexa R. Stephenson, Lakaléa J. Wilson, William D. Hopkins, John J. Ely, Joseph M. Erwin, Bob Jacobs, Patrick R. Hof, y Chet C. Sherwood. 2016. «Human-specific increase of dopaminergic innervation in a striatal region associated with speech and language: A comparative analysis of the primate basal ganglia». *Journal of Comparative Neurology* 524 (10): 2117-29. <https://doi.org/10.1002/cne.23937>.
- Rao, Suhas S P, Miriam H Huntley, Neva C Durand, Elena K Stamenova, Ivan D Bochkov, James T Robinson, Adrian L Sanborn, et al. 2014. «A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping.» *Cell* 159 (7): 1665-80. <https://doi.org/10.1016/j.cell.2014.11.021>.

- Reader, Simon M, y Kevin N Laland. 2002. «Social intelligence, innovation, and enhanced brain size in primates.» *Proceedings of the National Academy of Sciences of the United States of America* 99 (7): 4436-41. <https://doi.org/10.1073/pnas.062041299>.
- Reich, David, Richard E. Green, Martin Kircher, Johannes Krause, Nick Patterson, Eric Y. Durand, Bence Viola, et al. 2010. «Genetic history of an archaic hominin group from Denisova Cave in Siberia». *Nature* 468 (7327): 1053-60. <https://doi.org/10.1038/nature09710>.
- Reick, M., J A Garcia, C Dudley, y S L McKnight. 2001. «NPAS2: An Analog of Clock Operative in the Mammalian Forebrain». *Science* 293 (5529): 506-9. <https://doi.org/10.1126/science.1060699>.
- Rilling, James K., y Thomas R. Insel. 1999. «The primate neocortex in comparative perspective using magnetic resonance imaging». *Journal of Human Evolution* 37 (2): 191-223. <https://doi.org/10.1006/jhev.1999.0313>.
- Ritter, Deborah I., Qiang Li, Dennis Kostka, Katherine S. Pollard, Su Guo, y Jeffrey H. Chuang. 2010. «The Importance of Being Cis: Evolution of Orthologous Fish and Mammalian Enhancer Activity». *Molecular Biology and Evolution* 27 (10): 2322-32. <https://doi.org/10.1093/molbev/msq128>.
- Roessler, E., E. Belloni, K. Gaudenz, F. Vargas, S. W. Scherer, L.-C. Tsui, y M. Muenke. 1997. «Mutations in the C-Terminal Domain of Sonic Hedgehog Cause Holoprosencephaly». *Human Molecular Genetics* 6 (11): 1847-53. <https://doi.org/10.1093/hmg/6.11.1847>.
- Roessler, Erich, Elena Belloni, Karin Gaudenz, Philippe Jay, Philippe Berta, Stephen W. Scherer, Lap-Chee Tsui, y Maximilian Muenke. 1996. «Mutations in the human Sonic Hedgehog gene cause holoprosencephaly». *Nature Genetics* 14 (3): 357-60. <https://doi.org/10.1038/ng1196-357>.
- Rosenfeld, M. G., Victoria V Lunyak, y Christopher K Glass. 2006. «Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response». *Genes & Development* 20 (11): 1405-28. <https://doi.org/10.1101/gad.1424806>.
- Roth, Gerhard, y Ursula Dicke. 2005. «Evolution of the brain and intelligence». *Trends in Cognitive Sciences* 9 (5): 250-57. <https://doi.org/10.1016/J.TICS.2005.03.005>.
- Rubinstein, Marcelo, y Flávio S.J. de Souza. 2013. «Evolution of transcriptional enhancers and animal diversity.» *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 368 (1632): 20130017. <https://doi.org/10.1098/rstb.2013.0017>.
- S, Prabhakar, Noonan Jp, Pa "bo S, y Rubin Em. 2006. «Accelerated evolution of conserved noncoding sequences in humans». *Science* 314 (November): 786.
- Sagai, T., Masaki Hosoya, Youichi Mizushima, Masaru Tamura, y Toshihiko Shiroishi. 2005. «Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb». *Development* 132 (4): 797-803. <https://doi.org/10.1242/dev.01613>.
- Sankararaman, Sriram, Swapan Mallick, Michael Dannemann, Kay Prüfer, Janet Kelso, Svante Pääbo, Nick Patterson, y David Reich. 2014. «The genomic landscape of Neanderthal ancestry in present-day humans». *Nature* 507 (7492): 354-57. <https://doi.org/10.1038/nature12961>.
- Sanyal, Amartya, Bryan R. Lajoie, Gaurav Jain, y Job Dekker. 2012. «The long-range interaction landscape of gene promoters». *Nature* 489 (7414): 109-13. <https://doi.org/10.1038/nature11279>.
- Sasaki, Atsuo, Takaharu Taketomi, Toru Wakioka, Reiko Kato, y Akihiko Yoshimura. 2001. «Identification of a Dominant Negative Mutant of Sprouty That Potentiates Fibroblast Growth Factor- but Not Epidermal Growth Factor-induced ERK Activation». *Journal of Biological Chemistry* 276 (39): 36804-8. <https://doi.org/10.1074/jbc.C100386200>.
- Schilling, Thomas F, y Pierre Le Pabic. 2009. «Fishing for the signals that pattern the face.» *Journal of*

- biology* 8 (11): 101. <https://doi.org/10.1186/jbiol205>.
- Schmitt, Anthony D., Ming Hu, Inkyung Jung, Zheng Xu, Yunjiang Qiu, Catherine L. Tan, Yun Li, et al. 2016. «A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome». *Cell Reports* 17 (8): 2042-59. <https://doi.org/10.1016/j.celrep.2016.10.061>.
- Segawa, Yoshie, Naoya Itokazu, Akiko Hirose, Shinichiro Nakagawa, y Sachio Takashima. 2007. «A case of partial 14q- with facial features of holoprosencephaly and hydranencephaly.» *Pediatric neurology* 37 (1): 51-54. <https://doi.org/10.1016/j.pediatrneurol.2007.02.010>.
- Sherwood, Chet C., Francys Subiaul, y Tadeusz W. Zawidzki. 2008. «A natural history of the human mind: Tracing evolutionary changes in brain and cognition». En *Journal of Anatomy*. <https://doi.org/10.1111/j.1469-7580.2008.00868.x>.
- Shimamura, K, y J L Rubenstein. 1997. «Inductive interactions direct early regionalization of the mouse forebrain.» *Development (Cambridge, England)* 124 (14): 2709-18. <http://www.ncbi.nlm.nih.gov/pubmed/9226442>.
- Shin, Jiheon, Hye Yun Jeong, Kyung Eun Lee, y Jaesang Kim. 2010. «Isolation and Characterization of Chicken NPAS3.» *Experimental neurobiology* 19 (2): 71-74. <https://doi.org/10.5607/en.2010.19.2.71>.
- Siepel, A., Gill Bejerano, Jakob S Pedersen, Angie S Hinrichs, Minmei Hou, Kate Rosenbloom, Hiram Clawson, et al. 2005. «Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes». *Genome Research* 15 (8): 1034-50. <https://doi.org/10.1101/gr.3715005>.
- Sivak, Jeremy M., Lars F. Petersen, y Enrique Amaya. 2005. «FGF Signal Interpretation Is Directed by Sprouty and Spred Proteins during Mesoderm Formation». *Developmental Cell* 8 (5): 689-701. <https://doi.org/10.1016/j.devcel.2005.02.011>.
- Slon, Viviane, Fabrizio Mafessoni, Benjamin Vernot, Cesare de Filippo, Steffi Grote, Bence Viola, Mateja Hajdinjak, et al. 2018. «The genome of the offspring of a Neanderthal mother and a Denisovan father». *Nature* 561 (7721): 113-16. <https://doi.org/10.1038/s41586-018-0455-x>.
- Smale, Stephen T., y James T. Kadonaga. 2003. «The RNA Polymerase II Core Promoter». *Annual Review of Biochemistry* 72 (1): 449-79. <https://doi.org/10.1146/annurev.biochem.72.121801.161520>.
- Solomon, Keely S, Tetsuhiro Kudoh, Igor B Dawid, y Andreas Fritz. 2003. «Zebrafish foxi1 mediates otic placode formation and jaw development.» *Development (Cambridge, England)* 130 (5): 929-40. <http://www.ncbi.nlm.nih.gov/pubmed/12538519>.
- Souza, Flávio S.J. De, Lucía F. Franchini, y Marcelo Rubinstein. 2013. «Exaptation of transposable elements into novel Cis-regulatory elements: Is the evidence always strong?» *Molecular Biology and Evolution* 30 (6): 1239-51. <https://doi.org/10.1093/molbev/mst045>.
- Stanco, Amelia, Ramón Pla, Daniel Vogt, Yiran Chen, Shyamali Mandal, Jamie Walker, Robert F. Hunt, et al. 2014. «NPAS1 Represses the Generation of Specific Subtypes of Cortical Interneurons». *Neuron* 84 (5): 940-53. <https://doi.org/10.1016/j.neuron.2014.10.040>.
- Stephan, Heinz, Heiko Frahm, y Georg Baron. 1981. «New and Revised Data on Volumes of Brain Structures in Insectivores and Primates». *Folia Primatologica* 35 (1): 1-29. <https://doi.org/10.1159/000155963>.
- Strahl, Brian D., y C. David Allis. 2000. «The language of covalent histone modifications». *Nature* 403 (6765): 41-45. <https://doi.org/10.1038/47412>.
- Strathern, J N, B K Shafer, y C B McGill. 1995. «DNA synthesis errors associated with double-strand-break repair.» *Genetics* 140 (3): 965-72. <http://www.ncbi.nlm.nih.gov/pubmed/7672595>.
- Striedter, Georg F. 2005. *Principles of Brain Evolution*. Sunderland, Massachusetts U.S.A.: Sinauer

- Associates, Incorporated.
- Suster, Maximiliano L., Hiroshi Kikuta, Akihiro Urasaki, Kazuhide Asakawa, y Koichi Kawakami. 2009. «Transgenesis in Zebrafish with the Tol2 Transposon System». En *Methods in molecular biology* (Clifton, N.J.), 561:41-63. https://doi.org/10.1007/978-1-60327-019-9_3.
- Swanson, Christina I., Nicole C. Evans, y Scott Barolo. 2010. «Structural Rules and Complex Regulatory Circuitry Constrain Expression of a Notch- and EGFR-Regulated Eye Enhancer». *Developmental Cell* 18 (3): 359-70. <https://doi.org/10.1016/j.devcel.2009.12.026>.
- Tagle, D A, B F Koop, M Goodman, J L Slightom, D L Hess, y R T Jones. 1988. «Embryonic epsilon and gamma globin genes of a prosimian primate (*Galago crassicaudatus*). Nucleotide and amino acid sequences, developmental regulation and phylogenetic footprints.» *Journal of molecular biology* 203 (2): 439-55. <http://www.ncbi.nlm.nih.gov/pubmed/3199442>.
- Tattersall, Ian. 2000. «Once We Were Not Alone». *Scientific American* 282 (1): 56-62. <https://doi.org/10.1038/scientificamerican0100-56>.
- . 2010. «Human evolution and cognition». *Theory in Biosciences* 129 (2-3): 193-201. <https://doi.org/10.1007/s12064-010-0093-9>.
- Tavaré, Simon, Charles R. Marshall, Oliver Will, Christophe Soligo, y Robert D. Martin. 2002. «Using the fossil record to estimate the age of the last common ancestor of extant primates». *Nature* 416 (6882): 726-29. <https://doi.org/10.1038/416726a>.
- Thorne, A G, y M H Wolpoff. 1992. «The multiregional evolution of humans.» *Scientific American* 266 (4): 76-79, 82-83. <http://www.ncbi.nlm.nih.gov/pubmed/1566033>.
- Tomancak, Pavel, Amy Beaton, Richard Weismann, Elaine Kwan, ShengQiang Shu, Suzanna E Lewis, Stephen Richards, et al. 2002. «Systematic determination of patterns of gene expression during *Drosophila* embryogenesis». *Genome Biology* 3 (12): research0088.1. <https://doi.org/10.1186/gb-2002-3-12-research0088>.
- Tomancak, Pavel, Benjamin P Berman, Amy Beaton, Richard Weismann, Elaine Kwan, Volker Hartenstein, Susan E Celniker, y Gerald M Rubin. 2007. «Global analysis of patterns of gene expression during *Drosophila* embryogenesis». *Genome Biology* 8 (7): R145. <https://doi.org/10.1186/gb-2007-8-7-r145>.
- Trinkaus, Erik, y Pat Shipman. 2005. «Neandertals: Images of ourselves». *Evolutionary Anthropology: Issues, News, and Reviews* 1 (6): 194-201. <https://doi.org/10.1002/evan.1360010605>.
- Tsang, M., y I. B. Dawid. 2004. «Promotion and Attenuation of FGF Signaling Through the Ras-MAPK Pathway». *Science Signaling* 2004 (228): pe17-pe17. <https://doi.org/10.1126/stke.2282004pe17>.
- Valentine, J W, D Jablonski, y D H Erwin. 1999. «Fossils, molecules and embryos: new perspectives on the Cambrian explosion.» *Development (Cambridge, England)* 126 (5): 851-59. <http://www.ncbi.nlm.nih.gov/pubmed/9927587>.
- Vallender, Eric J., Nitzan Mekel-Bobrov, y Bruce T. Lahn. 2008. «Genetic basis of human brain evolution». *Trends in Neurosciences* 31 (12): 637-44. <https://doi.org/10.1016/j.tins.2008.08.010>.
- Vaquerizas, Juan M., Sarah K. Kummerfeld, Sarah A. Teichmann, y Nicholas M. Luscombe. 2009. «A census of human transcription factors: function, expression and evolution». *Nature Reviews Genetics* 10 (4): 252-63. <https://doi.org/10.1038/nrg2538>.
- Varki, A., y Tasha K Altheide. 2005. «Comparing the human and chimpanzee genomes: Searching for needles in a haystack». *Genome Research* 15 (12): 1746-58. <https://doi.org/10.1101/gr.3737405>.
- Venkatesh, Byrappa, Ewen F. Kirkness, Yong-Hwee Loh, Aaron L. Halpern, Alison P. Lee, Justin Johnson, Nidhi Dandona, et al. 2006. «Ancient Noncoding Elements Conserved in the Human Genome».

- Science* 314 (5807): 1892-1892. <https://doi.org/10.1126/science.1130708>.
- Vernot, B., y J. M. Akey. 2014. «Resurrecting Surviving Neandertal Lineages from Modern Human Genomes». *Science* 343 (6174): 1017-21. <https://doi.org/10.1126/science.1245938>.
- Villar, Diego, Camille Berthelot, Sarah Aldridge, Tim F. Rayner, Margus Lukk, Miguel Pignatelli, Thomas J. Park, et al. 2015. «Enhancer Evolution across 20 Mammalian Species». *Cell* 160 (3): 554-66. <https://doi.org/10.1016/j.cell.2015.01.006>.
- Visel, A., S. Minovitsky, I. Dubchak, y L. A. Pennacchio. 2007. «VISTA Enhancer Browser--a database of tissue-specific human enhancers». *Nucleic Acids Research* 35 (Database): D88-92. <https://doi.org/10.1093/nar/gkl822>.
- Visel, Axel, Matthew J. Blow, Zirong Li, Tao Zhang, Jennifer a. Akiyama, Amy Holt, Ingrid Plajzer-Frick, et al. 2009. «ChIP-seq accurately predicts tissue-specific activity of enhancers». *Nature* 457 (7231): 854-58. <https://doi.org/10.1038/nature07730>.
- Visel, Axel, Edward M. Rubin, y Len a. Pennacchio. 2009. «Genomic views of distant-acting enhancers». *Nature* 461 (7261): 199-205. <https://doi.org/10.1038/nature08451>.
- Wang, G L, B H Jiang, E A Rue, y G L Semenza. 1995. «Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O₂ tension.» *Proceedings of the National Academy of Sciences of the United States of America* 92 (12): 5510-14. <http://www.ncbi.nlm.nih.gov/pubmed/7539918>.
- Weber, Heike, Sarah Kittel-Schneider, Alexandra Gessner, Katharina Domschke, Maria Neuner, Christian P Jacob, Henriette N Buttenschon, et al. 2011. «Cross-Disorder Analysis of Bipolar Risk Genes: Further Evidence of DGKH as a Risk Gene for Bipolar Disorder, but also Unipolar Depression and Adult ADHD». *Neuropsychopharmacology* 36 (10): 2076-85. <https://doi.org/10.1038/npp.2011.98>.
- Wilk, R, I Weizman, y B Z Shilo. 1996. «trachealess encodes a bHLH-PAS protein that is an inducer of tracheal cell fates in Drosophila.» *Genes & development* 10 (1): 93-102. <http://www.ncbi.nlm.nih.gov/pubmed/8557198>.
- Wilson, A C, y R L Cann. 1992. «The recent African genesis of humans.» *Scientific American* 266 (4): 68-73. <http://www.ncbi.nlm.nih.gov/pubmed/1566032>.
- Wittkopp, P J, y G Kalay. 2011. «Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence.» *Nat Rev Genet* 13 SRC-: 59-69.
- Won, Hyejung, Luis de la Torre-Ubieta, Jason L. Stein, Neelroop N. Parikshak, Jerry Huang, Carli K. Opland, Michael J. Gandal, et al. 2016. «Chromosome conformation elucidates regulatory relationships in developing human brain». *Nature* 538 (7626): 523-27. <https://doi.org/10.1038/nature19847>.
- Wood, Bernard. 1996. «Human evolution». *BioEssays* 18 (12): 945-54. <https://doi.org/10.1002/bies.950181204>.
- Wray, Gregory A. 2007. «The evolutionary significance of cis-regulatory mutations». *Nature Reviews Genetics* 8 (3): 206-16. <https://doi.org/10.1038/nrg2063>.
- Wu, Dalei, Xiaoyu Su, Nalini Potluri, Youngchang Kim, y Fraydoon Rastinejad. 2016. «NPAS1-ARNT and NPAS3-ARNT crystal structures implicate the bHLH-PAS family as multi-ligand binding transcription factors». *eLife*. <https://doi.org/10.7554/eLife.18790>.
- Yan, Jun, Zhaowu Ma, Xiaopeng Xu, y An-Yuan Guo. 2014. «Evolution, functional divergence and conserved exon-intron structure of bHLH/PAS gene family». *Molecular Genetics and Genomics* 289 (1): 25-36. <https://doi.org/10.1007/s00438-013-0786-0>.
- Yasuoka, Yuuri, Yutaka Suzuki, Shuji Takahashi, Haruka Someya, Norihiro Sudou, Yoshikazu Haramoto,

- Ken W. Cho, Makoto Asashima, Sumio Sugano, y Masanori Taira. 2014. «Occupancy of tissue-specific cis-regulatory modules by Otx2 and TLE/Groucho for embryonic head specification». *Nature Communications* 5 (1): 4322. <https://doi.org/10.1038/ncomms5322>.
- Yelick, Pamela C, y Thomas F Schilling. 2002. «Molecular dissection of craniofacial development using zebrafish.» *Critical reviews in oral biology and medicine : an official publication of the American Association of Oral Biologists* 13 (4): 308-22. <http://www.ncbi.nlm.nih.gov/pubmed/12191958>.
- Yu, L., N. Arbez, L. G. Nucifora, G. L. Sell, L. E. Delisi, C. A. Ross, R. L. Margolis, y F. C. Nucifora. 2014. «A mutation in NPAS3 segregates with mental illness in a small family». *Molecular Psychiatry*. <https://doi.org/10.1038/mp.2012.192>.
- Yunis, J J, y O Prakash. 1982. «The origin of man: a chromosomal pictorial legacy.» *Science (New York, N.Y.)* 215 (4539): 1525-30. <http://www.ncbi.nlm.nih.gov/pubmed/7063861>.
- Zhou, Shutang, Simone Degan, Erin N Potts, W Michael Foster, y Mary E Sunday. 2009. «NPAS3 is a trachealess homolog critical for lung development and homeostasis.» *Proceedings of the National Academy of Sciences of the United States of America* 106 (28): 11691-96. <https://doi.org/10.1073/pnas.0902426106>.
- Zollikofer, Christoph P. E., Marcia S. Ponce de León, Daniel E. Lieberman, Franck Guy, David Pilbeam, Andossa Likius, Hassane T. Mackaye, Patrick Vignaud, y Michel Brunet. 2005. «Virtual cranial reconstruction of *Sahelanthropus tchadensis*». *Nature* 434 (7034): 755-59. <https://doi.org/10.1038/nature03397>.

Apéndice

Figuras suplementarias

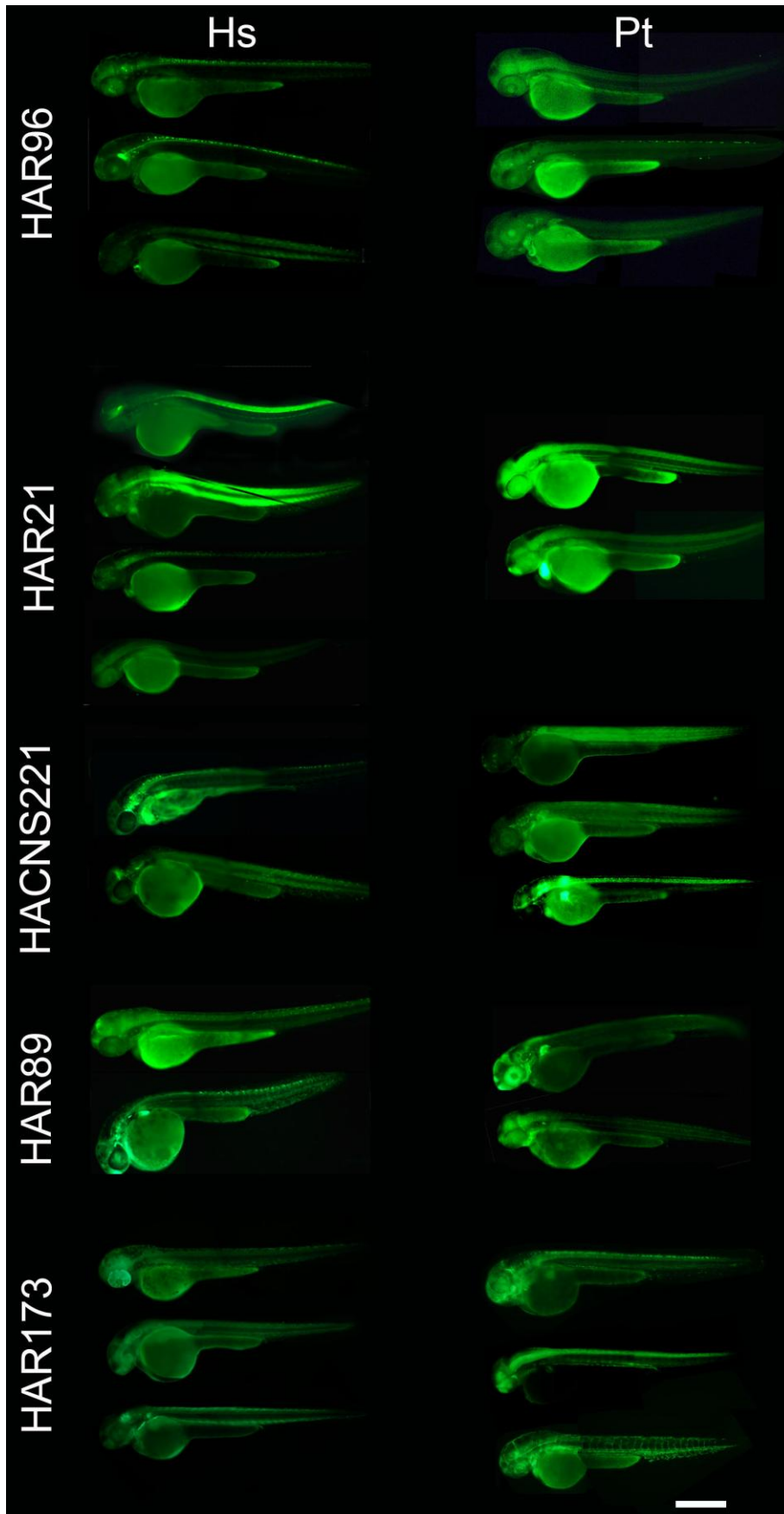


Figura suplementaria 1. Expresiones comparativas de todas las líneas transgénicas estables generadas en peces cebra portadoras de secuencias de humano y de chimpancé de los NPAS3-

HAEs HAR96, HAR21, HACNS221, HAR89, HAR173. El análisis se hizo sobre peces cebra a 48 hpf. Hs: *Homo sapiens*; Pt: *Pan troglodytes*. La barra representa 500um.

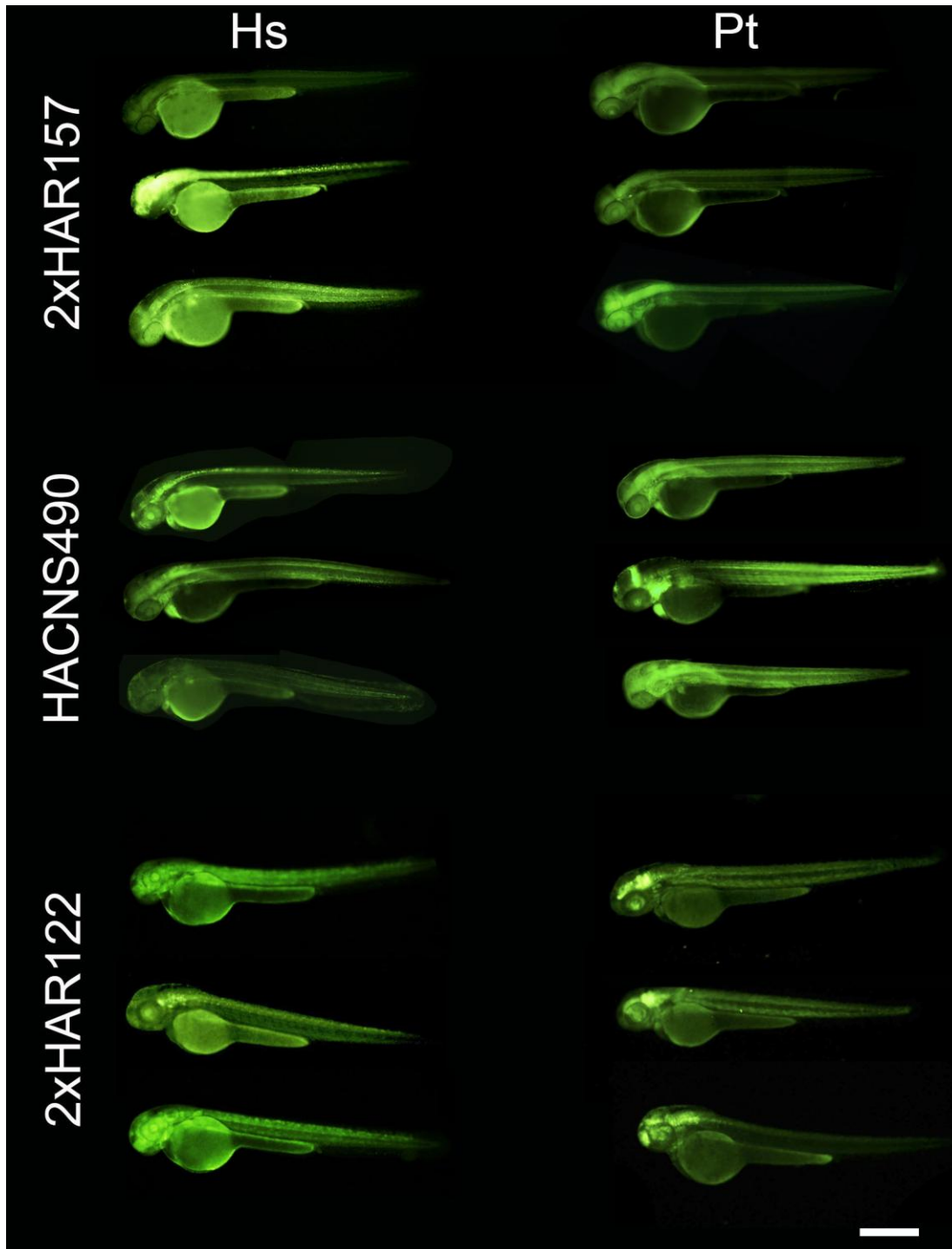


Figura suplementaria 2. Expresiones comparativas de todas las líneas transgénicas estables generadas en peces cebra portadoras de secuencias de humano y de chimpancé de los NPAS3-HAes 2xHAR157, HACNS490, 2x HAR122. El análisis se hizo sobre peces cebra a 48 hpf. Hs: *Homo sapiens*; Pt: *Pan troglodytes*. La barra representa 500um.

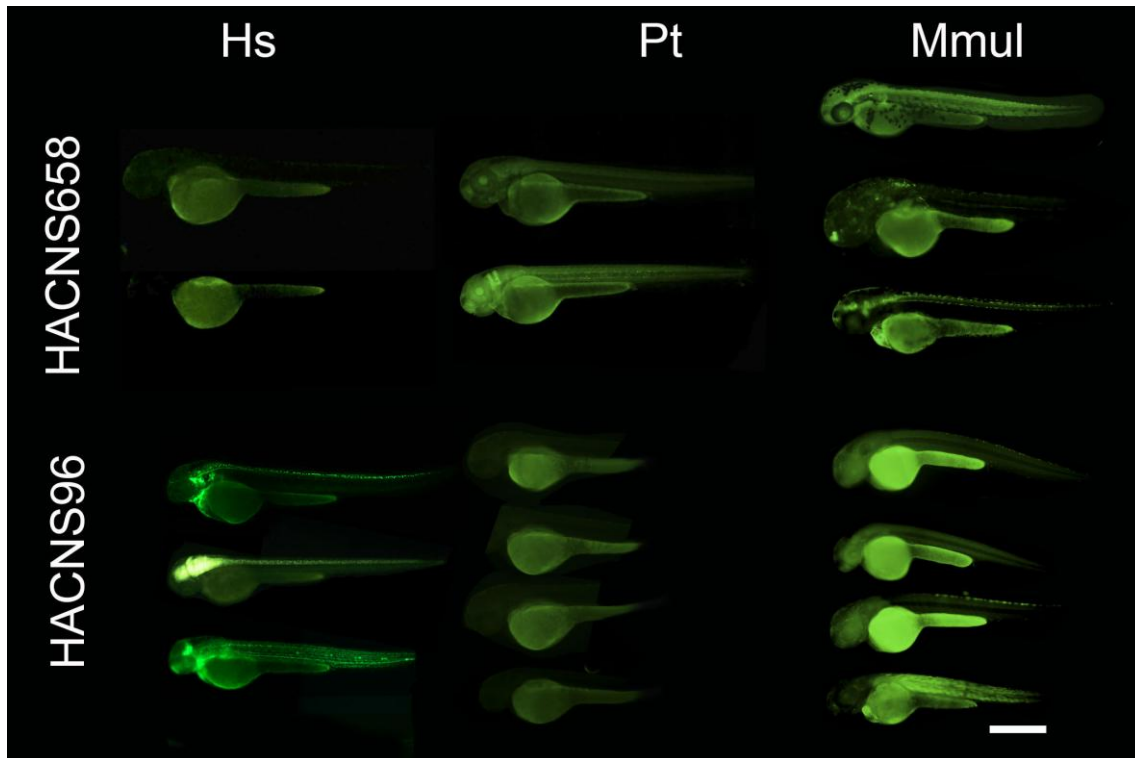


Figura suplementaria 3. Expresiones comparativas de todas las líneas transgénicas estables generadas en peces cebra portadoras de secuencias de humano, de chimpancé y de macaco Rhesus de los NPAS3-HAEs HACNS658, HACNS96. El análisis se hizo sobre peces cebra a 48 hpf. Hs: *Homo sapiens*; Pt: *Pan troglodytes*; Mmul: *Macaca mulatta*. La barra representa 500um.

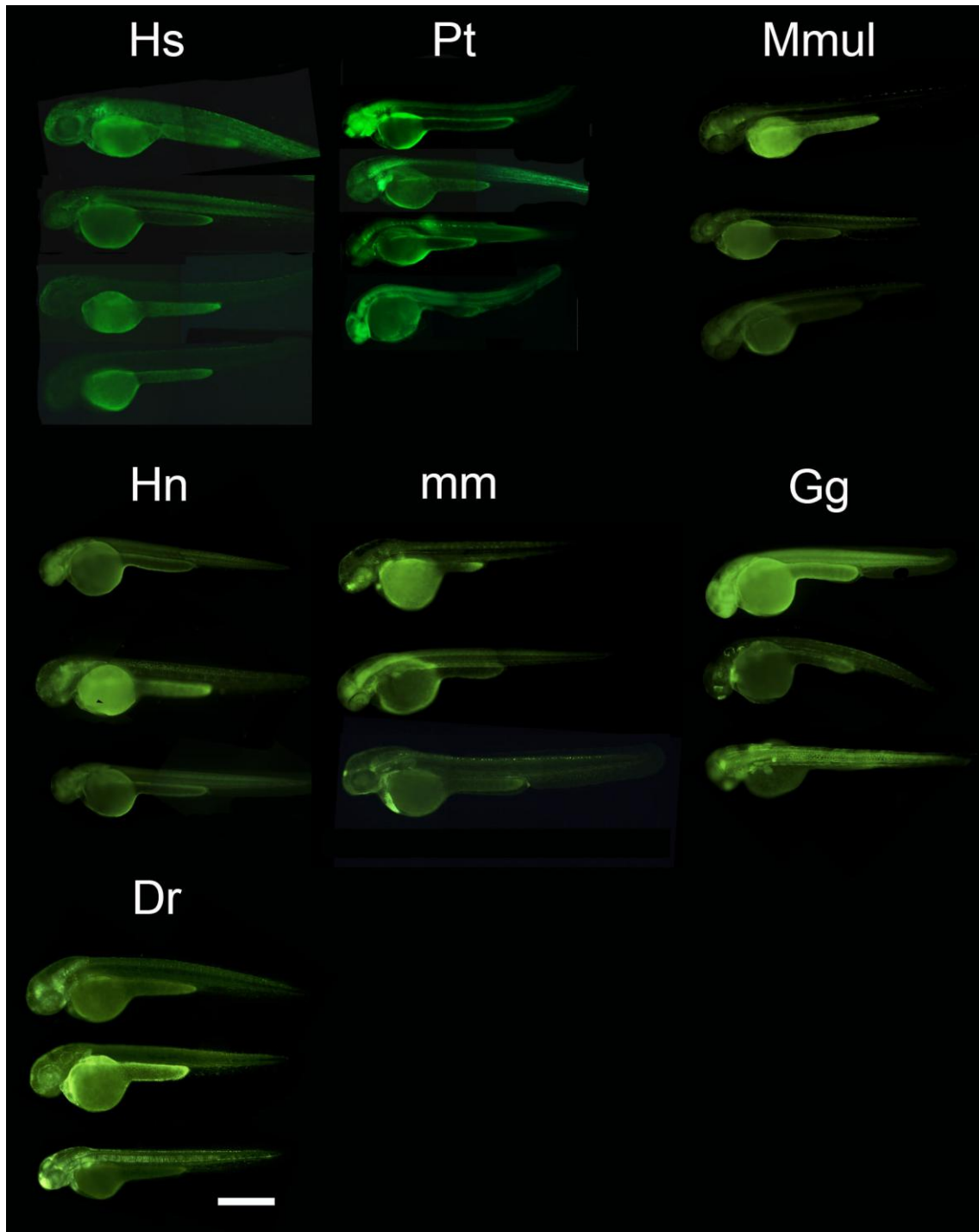


Figura suplementaria 4. Expresiones comparativas de todas las líneas transgénicas estables generadas en peces cebras portadoras de secuencias de humano, de chimpancé y de macaco Rhesus del NPAS3-HAEs HAR202. El análisis se hizo sobre peces cebras a 48 hpf. Hs: *Homo sapiens*; Pt: *Pan troglodytes*; Mmul: *Macaca mulatta*; mm: *Mus musculus*; Gg: *Gallus gallus*; Dr: *Danio rerio*. La barra representa 500um.

Figura suplementaria 5. Árbol filogenético de ortólogos y parálogos de NPAS3 de la familia de proteínas bHLH-PAS. Reconstrucción filogenética usando el método Neighbor-joining de la familia NPAS3 / trh basada en la alineación de secuencia múltiple de longitud completa de 123 especies utilizando el método de sustitución de aminoácidos de Jones-Taylor-Thornton (JTT). Para la alineación de la secuencia, se aplicó un método de eliminación parcial (corte de cobertura del 95%) para dar cuenta de los sitios faltantes. Se llevó a cabo una prueba de bootstrap de filogenia (1000 repeticiones). Barra de escala: 0.1 sustitución / sitio.