



UNIVERSIDAD DE CONCEPCION

FACULTAD DE CIENCIAS BIOLÓGICAS

**CARACTERIZACIÓN DE LA EXPRESIÓN ESPACIOTEMPORAL
DE *CITED2* DURANTE EL DESARROLLO EMBRIONARIO EN
*XENOPUS TROPICALIS***

Tesis presentada a la Facultad de Ciencias Biológicas para optar al grado
académico de Magíster en Bioquímica y Bioinformática.

POR: CARLA PAOLA CASTELLI BOERO

Profesora Guía: Dra. Marcela Eliana Torrejón Quezada

Profesora Co-guía: Dra. Estefanía Salomé Tarifeño Saldivia

Concepción, Chile, 2025.

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.

2025, Carla Paola Castelli Boero

AGRADECIMIENTOS

Esta tesis fue financiada por el Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) regular 1180926, y la Vicerrectoría de Investigación y Desarrollo de la Universidad de Concepción (VRID) 2022000483INV.

TABLA DE CONTENIDOS

ÍNDICE DE TABLAS	vii
ÍNDICE DE FIGURAS.....	vii
RESUMEN.....	x
ABSTRACT	xi
INTRODUCCIÓN.....	14
1. Crestas neurales.....	14
2. Inducción y especificación de las crestas neurales.....	16
3. Transición epitelio-mesenquimal en la cresta neural.....	17
4. Migración de las crestas neurales.....	18
5. Neurocristopatías.....	28
6. Señalización de Proteína G heterotrimérica.....	21
7. Ric-8 durante el desarrollo embrionario.....	22
8. Rol de <i>cited2</i> en el desarrollo embrionario.....	25
9. <i>Xenopus</i> como modelo de estudio.....	28
10. Presentación del problema	30
11. Hipótesis.....	32
MATERIALES Y MÉTODOS	35
1. Materiales.....	35
1.1. Oligonucleótidos	35
1.2. Vectores plasmidiales.....	35
1.3. Soluciones	36
1.4. Anticuerpos.....	38

1.5. Datos públicos	38
2. Métodos	38
2.1. Diseño de oligonucleótidos	38
2.2. Fertilización <i>in vitro</i>	39
2.3. Purificación de ARN total desde embriones de <i>X. tropicalis</i>	39
2.4. Reacción de Transcripción Reversa	41
2.5. Reacción de PCR para prueba de oligonucleótidos con polimerasa GoTaq	41
2.6. Reacción PCR con polimerasa de alta fidelidad.....	42
2.7. Purificación de producto PCR.....	43
2.8. Digestión con enzimas de restricción.....	43
2.9. Extracción de ADN desde geles de agarosa.....	44
2.10. Desfosforilación de vector plasmidial	44
2.11. Ligación convencional.....	45
2.12. Ligación en TOPO	45
2.13. Transformación bacteriana	45
2.14. Purificación de ADN plasmidial por minipreparación con buffer TENS.....	46
2.15. Purificación de ADN a pequeña escala	47
2.16. Purificación de ADN plasmidial a gran escala.....	48
2.17. Síntesis de ARNm y sondas de ARN marcadas con digoxigenina	49
2.18. Fijación de embriones con MEMFA	50
2.19. Hibridación <i>in situ</i> en embrión completo.....	51
2.20. Microinyección de ARNm.....	52
2.21. Lisis de embriones.....	52
2.22. Geles de Poliacrilamida SDS-PAGE	53
2.23. Electroforesis en gel de Poliacrilamida SDS/PAGE	54
2.24. Western Blot.....	54
2.25. Extracción de explantes de cresta neural.....	55
2.26. Ensayos de inmunofluorescencia de explantes de cresta neural	55
2.27. Obtención de niveles de expresión de <i>xtcited2</i> y <i>xtric-8a</i> desde data pública de RNA-seq de célula única	56

2.28. Modelamiento con aprendizaje automático de datos de subpoblaciones de cresta neural.....	58
RESULTADOS	61
1. Objetivo 1: Caracterizar el patrón de expresión espaciotemporal de <i>xtcited2</i> y comparación con el patrón de expresión espaciotemporal de <i>xtric-8a</i>.	61
1.1. Caracterización espaciotemporal de la expresión de transcritos de <i>xtcited2</i> en distintos estadios del desarrollo embrionario.	61
1.2. Análisis bioinformático de niveles de expresión de <i>xtcited2</i> y <i>xtric-8a</i>	69
1.3. Inmunodetección de la proteína de fusión V5+XtCited2 en embrión completo.	76
1.4. Inmunodetección de <i>xtcited2</i> en células de la cresta neural.....	79
2. Objetivo 2: Identificar una subpoblación de cresta neural cardiaca desde datos de RNAseq de célula única de cresta neural.	81
2.1. Búsqueda de clústeres de células de la cresta neural.....	81
2.2. Aplicación del algoritmo <i>Random Forest</i> para identificar los genes más importantes para la clasificación de los diferentes clústeres.	86
2.3. Análisis de expresión diferencial entre clústeres y caracterización de posibles genes marcadores.	89
DISCUSIÓN	96
CONCLUSIONES	108
GLOSARIO.....	110
BIBLIOGRAFÍA	113
ANEXOS.....	139

ÍNDICE DE TABLAS

Tabla 1: Protocolo para preparar geles denaturantes al 12%.de 1.5 mm de espesor.	53
Tabla 2: Hiperparámetros óptimos para el entrenamiento del modelo de <i>Random Forest</i>	84
Tabla 3: Criterios de expresión diferencial utilizados para clasificar genes según el clúster que marcan.....	90

ÍNDICE DE FIGURAS

Figura 1: Esquema general de la inducción de la CN	15
Figura 2: Esquema general de las rutas migratorias que siguen las distintas subpoblaciones de CN.	16
Figura 3: Esquema de los mecanismos involucrados en la migración celular de crestas neurales.....	20
Figura 4: Silenciamiento de <i>ric-8a</i> genera defectos en la migración de las CN y un desarrollo anormal del cartílago craneofacial en <i>Xenopus</i>	24
Figura 5: Inhibición de <i>cited2</i> genera defectos en el desarrollo embrionario de ratón.....	27
Figura 6: Síntesis de sondas para ensayos de hibridación <i>in situ</i>	62
Figura 7: Patrón de expresión de <i>xtcited2</i> en estadios de desarrollo temprano de <i>X. tropicalis</i>	64
Figura 8: Patrón de expresión de <i>xtcited2</i> durante la neurulación de <i>X. tropicalis</i>	66
Figura 9: Patrón de expresión de <i>xtcited2</i> durante la organogénesis de <i>X. tropicalis</i>	68
Figura 10: Expresión de <i>xtcited2</i> en tejidos embrionarios tempranos obtenidos desde ensayos de scRNA-seq.....	72
Figura 11: Expresión de <i>xtcited2</i> en tejidos embrionarios de néurula obtenidos desde ensayos de scRNA-seq.....	74
Figura 12: Expresión de <i>xtcited2</i> y <i>xtric-8A</i> en el desarrollo de la CN craneal.	75
Figura 13: Síntesis de ARNm de V5+XtCited2	77
Figura 14: Inmunodetección de V5+XtCited2	78
Figura 15: Localización subcelular de <i>xtcited2</i>	80
Figura 16: Comparación de agrupamientos generados por <i>clustering</i> Leiden a distintas resoluciones en células de cresta neural de estadio NF18	86

Figura 17: Distribución de genes con Importancia mayor a 0 según el modelo de Random Forest	89
Figura 18: Genes marcadores por clúster.....	92
Figura 19: Análisis de expresión del gen <i>gsk3b</i> en células de cresta neural total de estadio NF18.....	94

RESUMEN

Las crestas neurales (CN) son un tejido transitorio que surge durante el desarrollo embrionario de vertebrados, inducido por señales del mesodermo, neuroectodermo y ectodermo no neural. Tras su inducción, las CN experimentan una transición epitelio-mesenquimal, adquiriendo motilidad y migrando hacia diversas regiones del embrión para diferenciarse en una amplia gama de tipos celulares, donde destacamos estructuras craneofaciales y cardíacas. Disrupciones en el desarrollo de las CN podrían generar una o varias neurocristopatías, un grupo de patologías, síndromes y malformaciones congénitas derivadas de las CN. Previamente en nuestro laboratorio, se ha demostrado que Ric-8A, un factor intercambiador de nucleótidos de guanina, juega un papel crucial en la migración de las células de la (CN), y su pérdida de función está asociada a malformaciones craneofaciales. Por otro lado, hemos identificado a Cited2, un transactivador transcripcional con un rol importante en el desarrollo embrionario, como posible interactor de Ric-8A. Sin embargo, poco se conoce tanto de su función como del patrón de expresión de Cited2 en *Xenopus*. En este trabajo de tesis, se busca caracterizar la expresión espaciotemporal de *cited2* en el desarrollo embrionario en *Xenopus tropicalis*, y compararla con la de *ric-8a* con el fin de entender a futuro su posible interacción funcional para la formación de CN. Mediante hibridación *in situ* de embrión completo, se detectó expresión de *xtcited2* desde estadio NF 9 hasta NF 45. En

etapas tempranas, se observa principalmente en el polo animal y el tejido mesodermal, restringiéndose luego al tubo neural, CN, cerebro y corazón. Análisis de inmunofluorescencia de explantes de CN demostró localización nuclear, lo cual es consistente con su función biológica. La comparación con *xtric-8a*, mostró que ambos genes se expresan en tejidos neurales y de CN, siendo *xtcited2* más ampliamente distribuido y *xtric-8a* restringido al tubo neural, CN y tejidos derivados. Estos resultados se complementaron con datos de RNA-seq de célula única disponibles públicamente, donde se evaluó la expresión de ambos genes en tejidos que se derivan a CN craneal y se observó un patrón idéntico de ambos genes en este contexto. Además, se analizó un set de datos de CN total en estadio NF 18 con el fin de identificar una subpoblación de CN cardiaca. Se aplicaron técnicas bioinformáticas de reducción de dimensionalidad y búsqueda de clústeres, junto a un modelo de clasificación de Random Forest para evaluar la importancia de genes. Se identificaron 3 clústeres bien definidos con potenciales marcadores por clúster. Sin embargo, no se logró identificar una subpoblación cardiaca con certeza debido a limitaciones con el pipeline utilizado. Entre estas, se incluye el uso de un set de datos correspondiente a un estadio temprano del desarrollo (NF18), la alta similitud transcripcional entre células de la CN, y la influencia dominante de un único gen (*gsk3b*) sobre el modelo de clasificación, lo que podría haber apantallado las diferencias entre otros genes.

ABSTRACT

Neural crest cells (NC) are a transient, multipotent cell population that arises during vertebrate embryonic development, induced by signals from the mesoderm, neuroectoderm, and non-neural ectoderm. Following their induction, NCs undergo an epithelial-mesenchymal transition, gaining motility and migrating to various regions of the embryo, where they differentiate into a wide range of cell types, including craniofacial and cardiac structures. Disruptions in NC development can lead to neurocristopathies, a group of congenital disorders, syndromes, and malformations derived from NC cells. Previously, our lab studied Ric-8A, a guanine nucleotide exchange factor with a crucial role on NC migration, with its loss of function associated with craniofacial malformations. Cited2 a transcriptional coactivator with an important role in embryonic development, has been identified as a potential Ric-8A interactor. However, little is known about both its function and its expression pattern in *Xenopus*. This thesis aims to characterize the spatiotemporal expression of *cited2* during the embryonic development of *Xenopus tropicalis* and compare it with that of *xtric-8a* to analyze their potential functional relationship in NC formation. Whole-mount *in situ* hybridization was used to detect *xcited2* expression from stage NF 9 to NF45. At early stages, expression was primarily observed in the animal pole and mesodermal tissue, later becoming restricted to the neural tube, NC, brain, and heart. Immunofluorescence analysis of NC explants demonstrated nuclear

localization consistent with its biological function. Comparison with *xtric-8a*, showed that both genes are expressed in neural and NC tissues, with *xtcited2* exhibiting a broader distribution and *xtric-8a* being more specifically restricted to the neural tube, NC, and its derivatives. These results were complemented with publicly available single-cell RNA-seq data, evaluating the expression of both genes in tissues that give rise to cranial NC, where both showed an identical expression pattern in this context. Additionally, a total NC dataset at stage NF 18 was analyzed to identify a cardiac NC subpopulation. Dimensionality reduction and clustering techniques were applied, along with a Random Forest classification model to assess gene importance. Three well-defined clusters with potential marker genes per cluster were identified. However, a cardiac subpopulation could not be reliably identified due to limitations with the used pipeline, including the use of data corresponding to an early developmental stage (NF18), the high transcriptional similarity between NC cells, and the dominant influence of a single gene (*gsk3b*) on the classification model, which may have masked differences involving other relevant genes.

INTRODUCCIÓN

1. Crestas neurales.

Las crestas neurales (CN) son un tipo de tejido transitorio que aparece durante el desarrollo embrionario de vertebrados, inducidas por señales moleculares provenientes del mesodermo, neuroectodermo y ectodermo no neural (Figura 1). Una vez que se genera el cierre del tubo neural, en la etapa de neurulación, las CN pasan por un proceso de transición epitelio-mesenquimal (TEM) y delaminación, tras el cual procederán a migrar desde la región dorsal del tubo neural a distintas zonas del embrión, diferenciándose en una amplia gama de tipos celulares, incluyendo melanocitos, estructuras craneofaciales, células neuronales y gliales, del sistema nervioso, entre otras (Gilbert, 2005). Dependiendo de su destino celular, pueden ser categorizadas en 4 subpoblaciones: craneal, cardíaca, troncal y/o sacral-vagal (Mayor y Theveneau, 2013; Vega-Lopez y col., 2018). Las CN craneales, también llamadas cefálicas, migran dorso-lateralmente y de forma colectiva para generar gran parte del cartílago, hueso y músculo craneofacial (Gilbert, 2005). Por su parte, las CN cardíacas son una subregión de las craneales, ubicadas cerca del tercer somita y que migran, en su lugar, hacia el tercer, cuarto y quinto arco faríngeo y al tracto de salida cardíaco embrionario, donde formarán todo el tejido conectivo-muscular de las grandes arterias y tabique cardíaco (Gilbert, 2005; Scholl y Kirby, 2009; Yamagishi, 2021) (Figura 2).

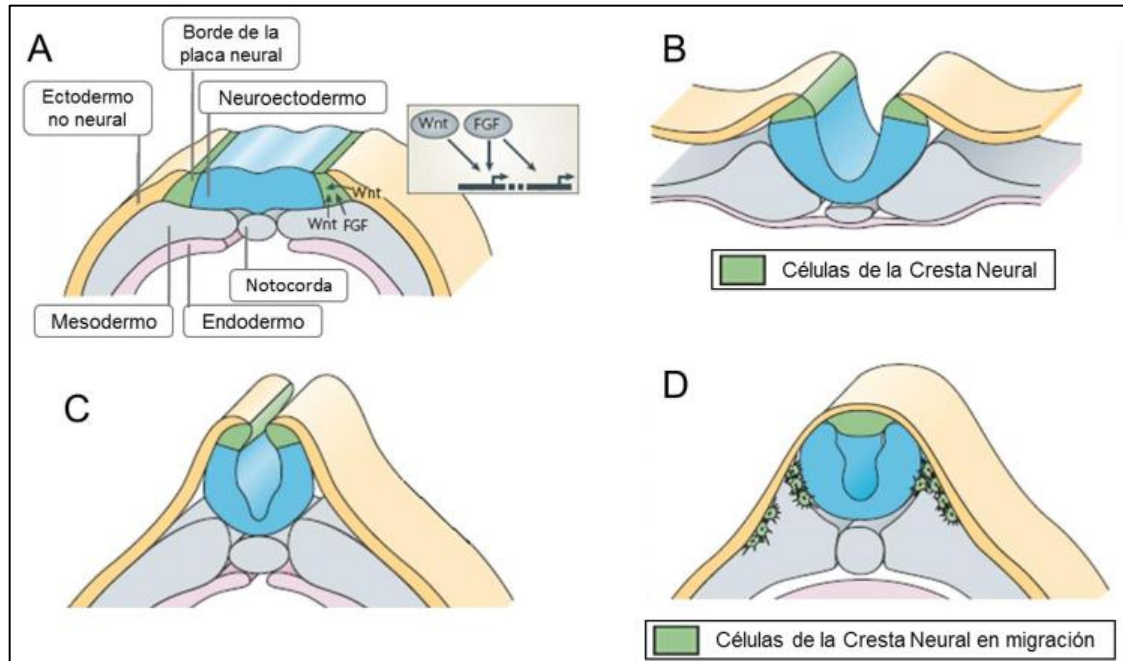


Figura 1: Esquema general de la inducción de la CN. (A) Inducción de la CN comienza en el borde de la placa neural producto de señales moleculares provenientes del ectodermo, neuroectodermo y mesodermo. Durante la neurulación del embrión, los pliegues laterales de la placa neural se elongan y fusionan (B, C), permitiendo así el cierre del tubo neural (D). Durante este proceso, las CN pasan por la TEM y, posteriormente, inician su migración colectiva. (Modificado de Sauka-Spengler y Bronner-Fraser, 2008)

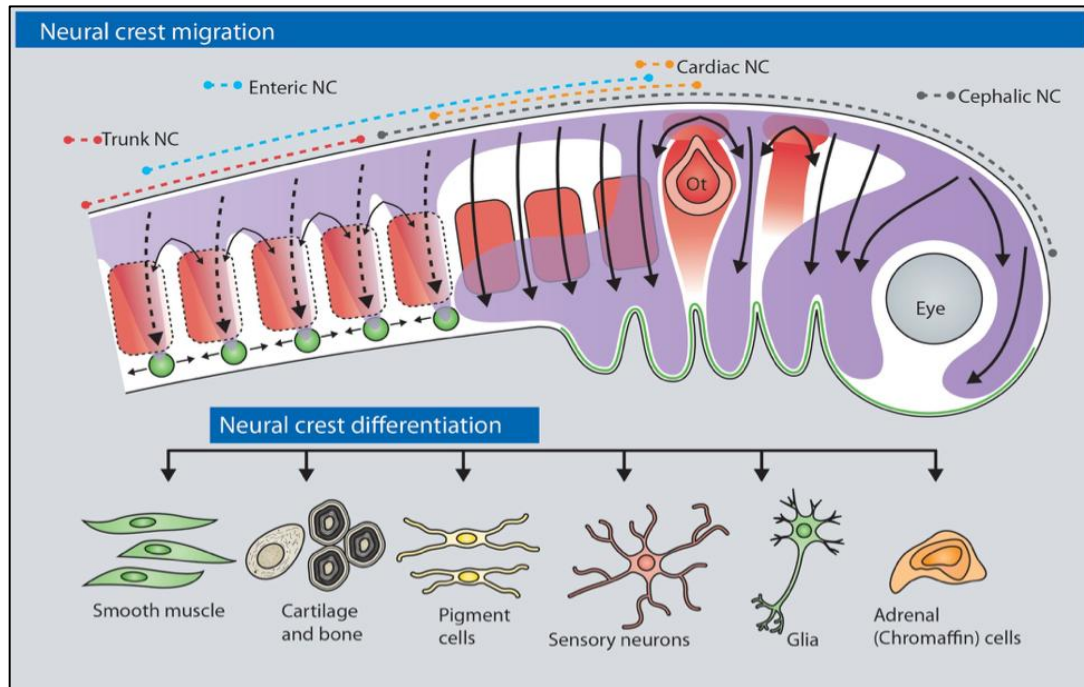


Figura 2: Esquema general de las rutas migratorias que siguen las distintas subpoblaciones de CN. Las distintas subpoblaciones de CN se categorizan según sus rutas migratorias y los tejidos específicos a los que darán origen. Entre estos, podemos encontrar músculo liso, cartílago y hueso, células pigmentadas, neuronas sensoriales, glías, células adrenales, entre otras. Gris: CN cefálica o craneal; amarillo: CN cardíaca; rojo: CN troncal; celeste: CN sacral-vagal o entérica (Modificado de Mayor y Theveneau, 2013).

2. Inducción y especificación de las crestas neurales.

La inducción de las CN inicia durante la etapa de gastrulación, en la cual se formarán las tres capas germinales (ectodermo, mesodermo y endodermo) y se

generarán los primeros movimientos celulares masivos (Gilbert, 2005). Las vías de señalización BMP, Wnt, FGF y Notch han sido descritas como cruciales para este proceso (Stuhlmiller y García-Castro, 2012). Un ejemplo de ello es el modelo de tres pasos, el cual consiste en un primer paso donde la represión de la vía Wnt permite la especificación de la placa neural; en el segundo paso, niveles intermedios de BMP y activación de Wnt son necesarios para la especificación de las células de la CN; y, en el tercer paso, ambas vías son requeridas para la mantención de la identidad de la CN (Steventon y Mayor, 2012).

En paralelo a estas vías de señalización, también se ha descrito la participación de múltiples factores de transcripción (FTs). Entre ellos podemos encontrar a los genes de especificación de placa neural, como *tfap2*, *msx1*, *zic1*, *pax3/7*, *dlx5/6*, *gata2/3* y *foxi1/2* (Khudyakov y Bronner-Fraser, 2009; Meulemans y Bronner-Fraser, 2004; Simões-Costa y Bronner, 2015); genes de especificación de CN, como *foxd3*, *snai1/2*, *twist* y *sox8/9* (Prasad y col., 2012; Simões-Costa y Bronner, 2015); y genes asociados a TEM, como *snail1/2*, *twist*, *foxd3*, *sip1/zeb2* y *sox9/10* (Prasad y col., 2012; Simões-Costa y Bronner, 2015).

3. Transición epitelio-mesenquimal en la cresta neural.

Posterior a su inducción, la CN pasan por un proceso de TEM caracterizado por un cambio en la expresión de E-cadherina a N-cadherina, ambas cruciales en la polaridad y adhesión celular, causando que las células de la CN pierdan sus características epiteliales y ganen motilidad, permitiendo así el inicio de la

migración colectiva (Huang y col., 2016; Kerosuo y Bronner-Fraser, 2012; Nieto, 2013; Scarpa y col., 2015; Theveneau y Mayor, 2012a). De forma interesante, varios de los FTs asociados a la especificación de la CN también participan en su TEM y posterior migración, como *snai1/2* y *twist*, *foxd3*, *sox9/10*, entre otros (Martik y Bronner, 2017; Piacentino y col., 2020; Simões-Costa y Bronner, 2015). Por otra parte, la vía Wnt también cumple un rol importante en la TEM, pues se ha descrito que esta vía de señalización es necesaria para modular el intercambio en la expresión de las cadherinas (Chalpe y col., 2010; Kerosuo y Bronner-Fraser, 2012; Piacentino y col., 2020; Taneyhill y col., 2007).

4. Migración de las crestas neurales.

El control de la dirección y movimiento colectivo ocurre mediante interacciones célula-célula, las cuales, a su vez, son reguladas por múltiples vías de señalización y mecanismos de acción. Cuando dos células chocan, se genera una señal que inhibe la formación de protrusiones en el sitio de contacto, lo cual, a su vez, estimula la formación de protrusiones en la dirección contraria (Figura 3). A este fenómeno lo llamamos inhibición de locomoción por contacto (CIL), descrito como una repolarización inducida por contacto célula-célula que cambia la dirección de migración y está regulada por la familia de las proteínas Rho GTPasas (Ridley y col., 2003; Roycroft y Mayor, 2016; Theveneau y Mayor, 2012b), junto a una participación importante de N-cadherina y proteínas de la vía WNT/PCP (Carmona-Fontaine, Matthews, y Mayor, 2008; Scarpa y col., 2015;

Theveneau y col., 2010). Gracias a este proceso, las protrusiones sólo pueden extenderse hacia el lado libre de células; es decir, en la dirección apropiada para la migración (Carmona-Fontaine, Matthews, y Mayor, 2008; Carmona-Fontaine, Matthews, Kuriyama, y col., 2008; Mayor y Theveneau, 2013). Por otro lado, para evitar que las células se dispersen sin sentido, existe un mecanismo opuesto a CIL: la co-atracción (CoA), encargada de mantener un movimiento migratorio cohesivo (Figura 3). C3a, un péptido del sistema del complemento, junto a su receptor C3aR han sido descritos como los reguladores principales de la CoA en CN mediante la activación de Rac-1, proteína G pequeña que se inhibe durante CIL (Carmona-Fontaine y col., 2011; Carmona-Fontaine, Matthews, y Mayor, 2008). Así, CIL y CoA trabajan en conjunto para mantener una migración cohesiva. Por otro lado, la migración está fuertemente influenciada por quimioatrayentes, fenómeno llamado Quimiotaxia. Las CN polarizadas responden a gradientes de Sdf1 (*stromal cell-derived factor 1*) que secretan las placodas, células guía de la migración, gracias a que expresan su receptor correspondiente, Cxcr4 (Figura 3). *In vivo*, la combinación de estos tres mecanismos genera el llamado “*chase and run*”, donde las placodas secretan Sdf1 para atraer a las CN, induciendo la formación de protrusiones en su dirección, pero se alejan por CIL al entrar en contacto con éstas (Carmona-Fontaine y col., 2011; Carmona-Fontaine, Matthews, Kuriyama, y col., 2008;

Roycroft y Mayor, 2016; Shellard y Mayor, 2016; Szabó y Mayor, 2015; Theveneau y col., 2010, 2013; Theveneau y Mayor, 2012b).

En nuestro laboratorio, se ha descrito el rol de la señalización de proteína G heterotrimérica durante la migración de la CN en *Xenopus*, particularmente de la proteína accesorio Ric-8A, un factor intercambiador de nucleótido de guanina (GEF), cuya función está asociada a la migración celular a través de la regulación de la dinámica de adhesiones focal y la respuesta quimiotáctica. (Figura 3) (Fuentealba y col., 2013; Leal y col., 2018; Toro-Tapia y col., 2017, 2018).

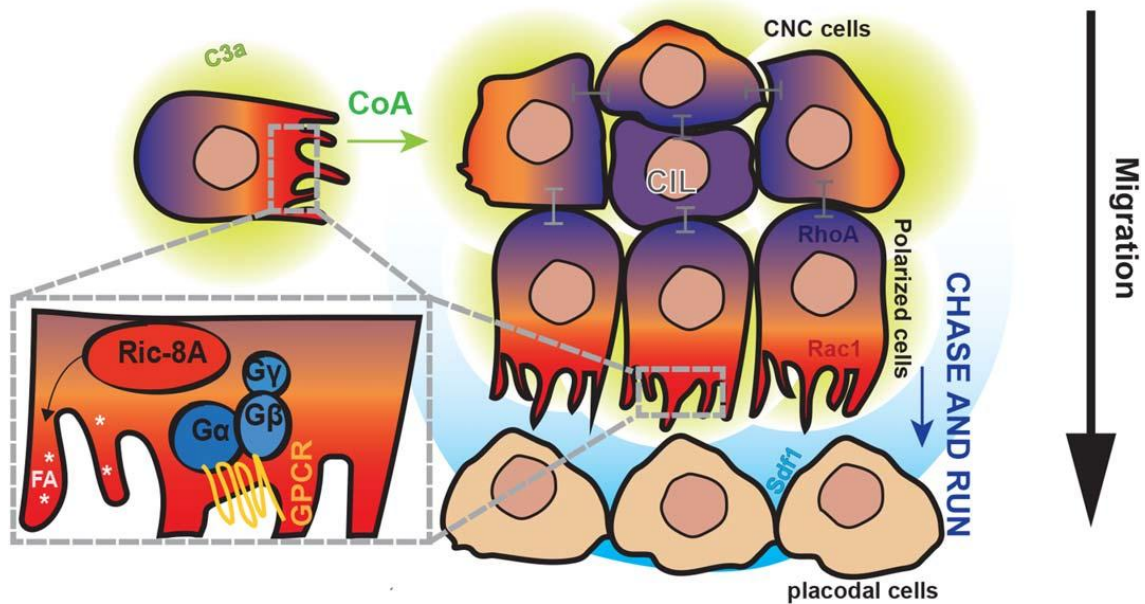


Figura 3: Esquema de los mecanismos involucrados en la migración celular de crestas neurales. La co-atracción (flecha verde, inducida por C3a) e inhibición por contacto (barras grises, regulada por RhoA) trabajan en conjunto

para generar un movimiento en dirección a las placodas, las cuales secretan el quimioatrayente Sdf1 y producen el fenómeno de “*chase and run*” (flecha azul). Además, observamos las diferencias de polarización dentro del grupo de células: en rojo a morado, la activación de Rac-1 permite la formación de protrusiones (rojo), las cuales se encuentran inhibidas en las zonas de contacto célula-célula, donde RhoA se encuentra activo (morado). Además, observamos presencia de Ric-8A y la proteína G heterotrimérica en el borde director libre de células de las CN, regulando su polarización adecuada durante la migración (Toro-Tapia y col., 2017).

5. Señalización de Proteína G heterotrimérica.

Con base a lo descrito anteriormente, podemos observar que el desarrollo de la CN está asociado a múltiples procesos dinámicos, donde las células responden y reaccionan a distintas señales y estímulos. Una de las formas en las que las CN logran esto es mediante el sistema de transducción de señales de la proteína G heterotrimérica y sus receptores asociados, los GPCR (Cotton y Claing, 2009).

La proteína G heterotrimérica está constituida por 3 subunidades: α , β y γ . En su vía canónica, la proteína G es activada por los GPCR, como el antes mencionado Cxcr4. Cuando un ligando se une al receptor, la proteína G se activa al intercambiar el GDP de la subunidad $G\alpha$ por un GTP y ésta se separa de la subunidad $G\beta\gamma$; proceso que se revierte una vez que el GTP es hidrolizado de

vuelta a GDP (Gilman, 1987). Si analizamos las características funcionales y estructurales de la subunidad $G\alpha$, podemos subdividir esta familia en $G\alpha_s$, $G\alpha_i/o$, $G\alpha_q$, $G\alpha_{12/13}$ y $G\alpha_v$, las cuales se diferencian según su secuencia y función específica en la señalización durante el desarrollo embrionario y en organismos adultos en distintas especies (Malbon, 2005; Syrovatkina y col., 2016).

6. Ric-8 durante el desarrollo embrionario.

Aparte de la vía canónica antes descrita, la vía de la proteína G heterotrimérica también posee mecanismos alternativos de activación mediante proteínas accesorias, como es el caso de Ric-8, que actúa como GEF (factor intercambiador de nucleótido de guanina) para las $G\alpha$ (Chan y col., 2011; Tall, 2013; Tall y col., 2003). “*Resistance to Inhibitors of Cholinesterase 8*”, abreviado como Ric-8, fue originalmente descrito en mutantes de *C. elegans* que presentaban resistencia a los inhibidores de colinesterasa, de donde proviene su nombre (Miller y col., 2000). A nivel molecular, Ric-8 interactúa con la subunidad $G\alpha$ facilitando el intercambio de GDP por GTP, manteniendo así la vía activa por mayor tiempo (Siderovski y Willard, 2005; Tall, 2013).

En mamíferos, se han descrito dos ortólogos de Ric-8: Ric-8A, que interactúa con $G\alpha_i$, $G\alpha_o$, $G\alpha_q$ y $G\alpha_{13}$, y Ric-8B, que interactúa con $G\alpha_s$ y $G\alpha_q$ (Chan y col., 2011; Tall, 2013; Tall y col., 2003). En nuestro laboratorio, se descubrió la presencia de un homólogo de Ric-8A en *Xenopus*, el cual se expresa durante el desarrollo embrionario en *Xenopus*, principalmente en tejido neural, incluyendo

la CN y sus derivados (Anexo 1 y Anexo 2) (Maldonado-Agurto y col., 2011). Este patrón de expresión es similar al observado para las distintas G α (Anexo 3), lo cual es esperable, al ser sus principales interactores (Fuentealba y col., 2016).

Mediante estudios tanto in vitro como in vivo, se ha determinado que la pérdida función de Ric-8A durante el desarrollo embrionario afecta negativamente la repolarización de las CN y la respuesta al gradiente de Sdf1, lo cual resulta en embriones con fenotipos craneofaciales anómalos y defectos en la diferenciación de células de cartílago (Fuentealba y col., 2013; Leal y col., 2018; Toro-Tapia y col., 2017) (Figura 4). Además, también se determinó que la actividad de Ric-8A es necesaria para la correcta formación del mesodermo dorsal, ectodermo neural y no neural, y la inducción de CN (Maldonado-Agurto y col., 2011; Rodríguez, 2016, Tesis de Magíster; Ruiz, 2024, Tesis de Magíster). Siguiendo esta evidencia, se sugiere que Ric-8A podría tener un rol regulatorio clave en la migración e inducción de las CN, y con ello en la formación de las estructuras craneofaciales. Por otro lado, se ha reportado que Ric-8A también cumple un rol como proteína chaperona para proteínas G α (Gabay y col., 2011; Klattenhoff y col., 2003; Toro-Tapia et al., 2018), lo cual podría deberse a su estructura tipo armadillo, característica en proteínas que interaccionan con otras proteínas (Figueroa y col., 2009).

De forma aún más importante, recientemente, en nuestro laboratorio, se realizó un estudio de espectrometría de masas para identificar las proteínas que

interactúan con Ric-8A durante la migración celular en fibroblastos de ratón, encontrándose entre ellas a “*Cbp/p300-interacting transactivator 2*” (Cited2), la cual es codificada por el gen *cited2* (Beyer, 2021, Tesis de Doctorado).

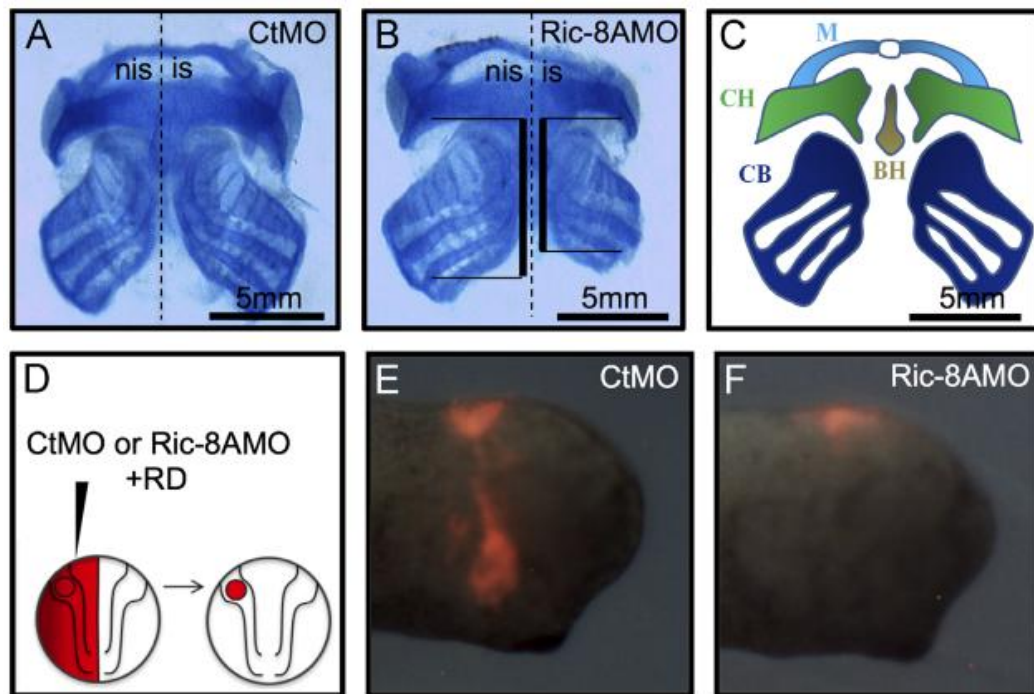


Figura 4: Silenciamiento de *ric-8a* genera defectos en la migración de las CN y un desarrollo anormal del cartílago craneofacial en *Xenopus*. (A) Vista ventral de embrión estadio NF 42 microinyectado con morfolino control (CtMO). **(B)** Silenciamiento de *ric-8a* por inyección de un morfolino específico (Ric-8AMO) produce el desarrollo anormal de cartílagos craneales. **(C)** Representación esquemática de los cartílagos embrionarios. **(D-F)** Experimento de trasplante de CN con un morfolino específico para *ric-8a* a un embrión normal, nos permite

observar que no existe migración de las CN en condiciones de silenciamiento de éste, indicando características de autonomía celular, luego de la disminución de la expresión de este gen. nis: lado no inyectado; is: lado inyectado; RD: rodamina dextrano; M: cartílago de Meckel; CH: cartílago ceratohial; CB: cartílago ceratobranquial; BH: cartílago basihial. (Fuentealba y col., 2013)

7. Rol de *cited2* en el desarrollo embrionario.

Cited2 es un transactivador transcripcional sin sitio de unión a ADN (Chou y col., 2012) que ha probado tener un papel clave en el desarrollo embrionario. En embriones de ratón knockout para *cited2*, se observa una mortalidad total, donde un alto porcentaje de embriones padece de exencefalia y malformaciones cardíacas (Bamforth y col., 2001). Mediante ensayos de hibridación in situ en este mismo modelo animal, también se ha podido confirmar que la migración de CN cardíaca se ve alterada (Bamforth y col., 2001). (Figura 5)

Entre otros efectos de la pérdida de función de *cited2* en ratón podemos encontrar múltiples defectos de lateralidad (ejes izquierdo-derecho). A nivel cardiovascular, observamos una posición incorrecta del ventrículo izquierdo, comunicación entre ventrículos y aurículas, y una posición incorrecta del corazón hacia la derecha. En los pulmones, observamos isomerismo derecho, donde el pulmón izquierdo es reemplazado por los cuatro lóbulos característicos del pulmón derecho. Además, observamos un posicionamiento incorrecto del estómago hacia la derecha del cuerpo, además del cambio de la posición de la

cola del embrión (de derecha a izquierda), lo cual indica una alteración en la rotación normal del embrión (Bamforth y col., 2004).

En cuanto a los mecanismos moleculares a través de los cuales *cited2* impacta en los procesos embrionarios, es fundamental destacar su principal función como transactivador transcripcional. Es sabido que *cited2* se une a diversos FTs, usualmente uniéndose de forma directa para regular su actividad (An y col., 2020). Un ejemplo particular de esto es el caso de *tfap2* (Bamforth y col., 2001), un FT cuya actividad se ha descrito ser necesaria para la formación de la CN (Hoffman y col., 2007; Li y Cornell, 2007; Simões-Costa y Bronner, 2015). Sin embargo, también tiene función inhibitoria al ser capaz de “competir” con otros FT, particularmente, *hif-1 α* (Berlow y col., 2017; Freedman y col., 2003). *Hif-1 α* es un FT que ha sido descrito ser crucial para la migración de la CN en el modelo *Xenopus* al momento de pasar por la TEM, controlando la expresión del FT *twist*, que a su vez es crítico en la regulación del cambio de la expresión de las cadherinas (Barriga y col., 2013).

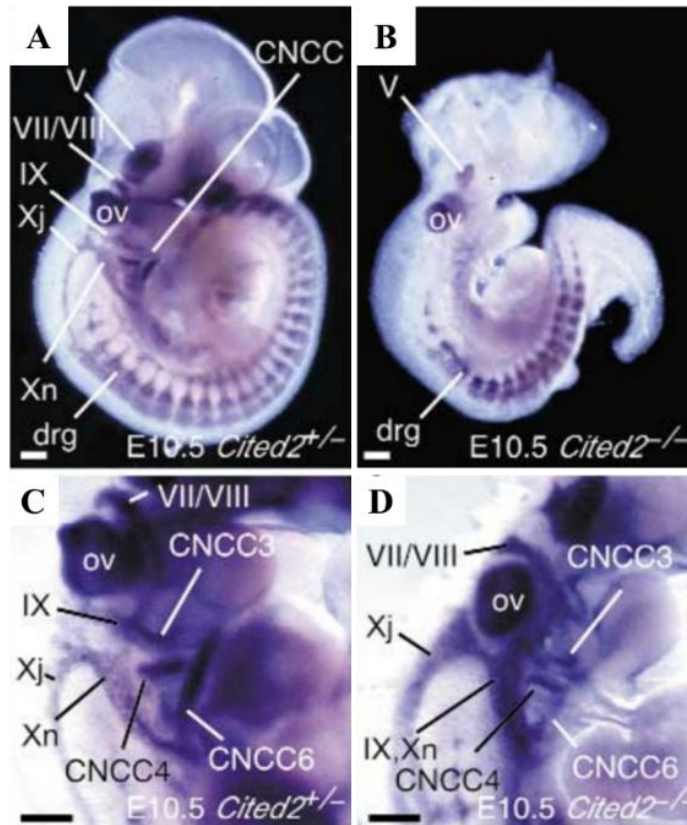


Figura 5: Inhibición de *cited2* genera defectos en el desarrollo embrionario de ratón. Hibridación *in situ* contra *erbb3*, un marcador de la CN. Los ratones knockout para *cited2* (**B, D**) presentan anomalías en el desarrollo craneofacial y neural respecto al control (**A, C**). Además, se observa que las rutas migratorias de las CN cardiacas (CNCC3, CNCC4 y CNCC6) se encuentran alteradas en los embriones knockout para *cited2* (**D**) respecto al control (**C**). V, VII/VIII, IX, Xj, Xn: ganglios craneales; CNCC: cresta neural cardiaca migrando por arcos aórticos 3, 4 y 6; ov: vesícula ótica; drg: ganglios de raíz dorsal. (Modificado de Bamforth y col., 2001)

8. Neurocristopatías.

Como se mencionó anteriormente, una vez que las CN llegan a los sitios apropiados dentro del embrión, comenzarán a diferenciarse a distintos tipos celulares. Sin embargo, ¿qué es lo que ocurriría si hubiese un problema en su desarrollo?

Introducido por primera vez en 1974 por el médico Robert Bolande, el término “neurocristopatía” describe a un grupo de patologías, síndromes y malformaciones congénitas que nacen de alguna disrupción en la inducción, proliferación, migración o diferenciación de las CN (Bolande, 1974; Etchevers y col., 2006; Watt y Trainor, 2014). La mayoría de las neurocristopatías son causadas por defectos genéticos y pueden manifestarse de forma aislada o junto con otras neurocristopatías (Pilon, 2021). Debido a la pluripotencialidad de las CN, las neurocristopatías presentan una gran variedad de síntomas, lo cual ha dificultado su estudio y categorización a lo largo de los años. La propuesta más reciente sugiere catalogarlas según la subpoblación de CN de la que derivan (Vega-Lopez y col., 2018).

Por ejemplo, la secuencia de Pierre-Robins (PRS) está descrita como una neurocristopatía derivada de las crestas neurales craneales. Con una incidencia de 1 cada 3120-14000 nacimientos en distintos países, se caracteriza por la presencia de una mandíbula inferior corta (micrognatia) y desplazamiento hacia posterior de la lengua (glosoptosis), a veces acompañados por paladar hendido,

que causan una serie de complicaciones de vía aérea a los recién nacidos afectados (Côté et al., 2015). Si bien las causas genéticas exactas de PRS todavía no están completamente claras, varios estudios indican que la expresión de *sox9*, un factor de transcripción presente durante el desarrollo de las crestas neurales, se encuentra alterada en esta neurocristopatía (Benko et al., 2009; Liu et al., 2013; Wakamatsu et al., 2014).

Por otro lado, el síndrome de delección 22q11.2 o síndrome de DiGeorge (22q11) es otra neurocristopatía de interés, derivada de crestas neurales tanto craneales como cardíacas. Pacientes con 22q11 presentan una variedad de síntomas, entre los cuales destacamos rasgos faciales dismórficos y comunicación intraventricular (Du et al., 2020). Como indica el nombre, la causa genética principal del síndrome es la delección de la región 11.2 del cromosoma 22, donde se encuentra el gen *tbx1*, un factor de transcripción que regula a varios otros genes involucrados en la migración celular durante el desarrollo embrionario (Baldini et al., 2017).

Por lo tanto, cualquier gen clave para la formación de la CN y el desarrollo de sus derivados podrían ser precursoras de neurocristopatías.

9. *Xenopus* como modelo de estudio.

Para la realización de este trabajo de tesis, se ha escogido utilizar el modelo de rana (*Xenopus*) debido a las facilidades que presenta a la hora de estudiar las CN, las cuales son fáciles de manipular para su estudio tanto in vitro como in

vivo. *X. tropicalis* en particular, es un organismo diploide cuyo genoma ya se encuentra completamente descrito, por lo cual es idóneo para estudios de genética (Harland y Grainger, 2011). En cuanto a su desarrollo embrionario, ocurre de forma externa, lo cual facilita su observación y la realización de técnicas tales como trasplantes y explantes, imagenología in vivo, microinyección de morfolidos y ARNm para silenciar o sobreexpresar genes, respectivamente (Fuentealba y col., 2013; Leal y col., 2018; Mayor y Etienne-Manneville, 2016; Mayor y Theveneau, 2013; Toro-Tapia y col., 2017, 2018). Es relevante saber que, a diferencia de otros modelos animales usados comúnmente para el desarrollo embrionario, como el ratón o el pollo, las ranas *Xenopus* poseen únicamente un ventrículo, y las CN cardíacas participan en la formación de las arterias del arco aórtico, pero no del tracto de salida (Hempel y Kühl, 2016; Lee y Saint-Jeannet, 2011; Mohun y col., 2000).

10. Presentación del problema

Debido a que existen neurocristopatías que, se caracterizan por presentar defectos en estructuras craneales y cardíacas, tales como la Secuencia de Pierre Robin y el Síndrome de Di George (Côté y col., 2015; Du y col., 2020), podríamos sugerir que es posible que Ric-8A esté participando en la aparición de neurocristopatías ya descritas, o incluso siendo el causante principal de algunas, ya que embriones con Ric-8A silenciados presentan malformaciones en las estructuras craneofaciales (Fuentealba y col., 2013). Así mismo, ratones

knockout de Ric-8A no son viables (Tõnissoo y col., 2010). Además, como ya mencionamos, recientemente encontramos un posible interactador de Ric-8A, Cited2 (Beyer, 2021, Tesis de Doctorado; Chou y col., 2012). Esto resulta interesante debido a que, si bien Cited2 no está descrito como parte de ninguna neurocristopatía craneal/cardiaca, sí se ha demostrado que posee un rol importante durante el desarrollo embrionario en la formación de estos órganos mediante su interacción con otras proteínas. Mencionamos que en modelo ratón, se ha descrito que Cited2 es necesario para el desarrollo del corazón, tubo neural, ojo y formación del patrón izquierda-derecha, entre otros; estudios de pérdida de función de Cited2 muestran que los embriones mueren después de unos días de desarrollo, presentando exencefalia y malformaciones cardíacas (Bamforth y col., 2001; Y. Chen y col., 2008; Dykes, 2014; Yin y col., 2002). En humanos, se ha visto que se encuentra mutado en personas con cardiopatías congénitas y espina bífida, una patología asociada a defectos en el cierre del tubo neural (Z. Chen y col., 2023; Dianatpour y col., 2020; Lu y col., 2010). Esto nos da indicios de que podría estar involucrado en la aparición de estas enfermedades.

Hasta la fecha, no hay mucha información sobre el patrón espaciotemporal de expresión de *cited2* en *Xenopus*, nuestro modelo de estudio. Esto hace relevante conocer el patrón de expresión que tiene *cited2* con el fin de comparar este con Ric-8A, y de esta manera comenzar a dilucidar si ambas proteínas son parte de

una misma señalización crítica durante la embriogénesis y formación de estructuras craneofaciales y cardíacas.

Además, si bien existe una amplia área de estudio de las CN, las CN cardíacas, históricamente, han sido difíciles de estudiar debido a que tanto su localización como los marcadores genéticos que expresan se superponen con otras subpoblaciones de la CN, particularmente con CN vagal y craneal (Mohun y col., 2000; Scholl y Kirby, 2009; Simões-Costa y Bronner, 2015; Simoes-Costa y Bronner, 2016; Tani-Matsuhana y col., 2018).

11. Hipótesis

Basándonos en un análisis de espectrometría de masas, que se mencionó en el marco teórico, y que muestra la interacción entre Ric-8A y Cited2, los cuales cumplen roles relevantes durante el desarrollo embrionario, proponemos que su interacción pudiera tener implicancias importantes en este proceso. En modelos *Xenopus* y ratón, respectivamente se ha observado que la pérdida de función de estas proteínas está asociada a un mal desarrollo de estructuras craneofaciales y cardíacas, resultado de una migración defectuosa de las CN. Por lo tanto, sugerimos a ambos genes como potenciales causantes de neurocristopatías. Sin embargo, poco se conoce respecto al patrón de expresión de *cited2* en *Xenopus*, por lo que este trabajo de tesis se enfoca principalmente en estudiar los patrones de expresión de *cited2* a nivel de embrión completo e intracelular en CN.

Además, considerando que una de las principales dificultades en el estudio de las CN cardíacas radica en la superposición espacial y molecular con otras subpoblaciones de CN, se incorporó un enfoque bioinformático cuyo objetivo es identificar una subpoblación de CN cardíacas dentro de las células de CN desde data pública. Esto nos permitiría evaluar si *xtcited2* y *xtric-8a* se expresan diferencialmente o co-expresan en este contexto, apoyando la hipótesis de una posible interacción funcional entre ambas proteínas en estas subpoblaciones

Con base en estos antecedentes, formulamos la siguiente hipótesis:

“El patrón de expresión de XtCited2 se relaciona espacial y temporalmente con el de XtRic-8A durante la formación de la cresta neural y el desarrollo de sus derivados, incluida la subpoblación cardíaca”, sugiriendo una posible coordinación de ambas proteínas en este contexto.

Para responder a esta hipótesis, se definió como objetivo general del estudio:

“Determinar la expresión espaciotemporal de *xtcited2* en la cresta neural y estructuras derivadas y compararla con la expresión espaciotemporal de *xtric-8a*, incluyendo la posible identificación de una subpoblación cardíaca.”

Y se proponen los siguientes objetivos específicos:

- 1. Caracterizar el patrón de expresión espaciotemporal de *xtcited2* y comparar con el patrón de expresión espaciotemporal de *xtric-8a*.**

- 1.1. Caracterización espaciotemporal de la expresión de transcritos de *xtcited2* en distintos estadios del desarrollo embrionario.
 - 1.2. Análisis bioinformático de niveles de expresión de *xtcited2* y *xtric-8a*.
 - 1.3. Inmunodetección de la proteína fusión V5+XtCited2 en embrión completo.
 - 1.4. Inmunodetección de *xtcited2* en células de la cresta neural.
- 2. Identificar una subpoblación de cresta neural cardiaca desde datos de RNAseq de célula única de cresta neural.**
- 2.1. Identificar clústeres de células de la cresta neural en estadio NF 18.
 - 2.2. Aplicación del algoritmo *Random Forest* para identificar los genes más importantes para la clasificación de los diferentes clústeres.
 - 2.3. Analizar la expresión diferencial entre clústeres y caracterización de posibles genes marcadores.

MATERIALES Y MÉTODOS

1. Materiales

1.1. Oligonucleótidos

XtCited2 Forward: 5' – GATTACAGAATTCATGGCAGACCACATG – 3'.

Hibrida con el codón de inicio de XtCited2 y posee sitio de restricción para EcoRI.

XtCited2 Reverse 1: 5' – GATTACACTCGAGCTAACAGCTTACTCT – 3'.

Hibrida con el codón de término de XtCited2 y posee sitio de restricción para XhoI.

XtCited2 Reverse 2: 5' – GATTACACTCGAGGGAAGCAAGGTTTTGTGCC

– 3'. Hibrida con la región 3'UTR.

Los sitios de corte han sido subrayados.

1.2. Vectores plasmidiales

pCS2+: Vector de expresión eucarionte con sitios de transcripción SP6 y T7.

pCS2+V5: Vector pCS2+ con tag V5 para detección con anticuerpos anti-V5.

pCR-Blunt-II-TOPO: Vector de traspaso.

pCR-Blunt-II-TOPO/XtCited2 (sentido): Vector de traspaso con la región codificante de XtCited2 insertada.

pUC/XtCited2: Vector sintetizado por la empresa IDT que contiene la región codificante de XtCited2.

pCS2+V5/XtCited2: Vector pCS2+ con tag V5 con la región codificante de XtCited2 insertada.

1.3. Soluciones

20X SSC: 175.3 g/l, NaCl, 88.2 g/l de citrato de sodio y ajustar pH 7.0 c/NaOH

Agua DEPC: 1ml de DEPC, en 1 litro de H₂O nanopure, agitar e incubar por 4 horas a 37°C, autoclavar.

Buffer AP: 100 mM Tris pH 9.5, 50 mM MgCl₂, 100 mM NaCl.

Buffer concentrador 2X: 250 mM Tris-base, 0.2% p/v SDS, pH 6.8.

Buffer de carga proteínas 5X: 312.5 mM Tris-HCl pH 8, 10% p/v SDS, 0.01% azul de bromofenol, 50% v/v glicerol, 100 mM DTT.

Buffer de corrida proteínas 1X: 25 mM Tris-base, 192 mM glicina, 0.1% p/v SDS.

Buffer de hibridación: Formamida 50%, SSC 5X, Torula ARN 1mg/ml, heparina 100µg/mL, 1x Denhart's, Tween 20 0.1%, CHAPS 0.1%, EDTA 10 mM.

Buffer de lisis de embriones para Western blot: PBS 1X, 0.5% SDS.

Buffer de transferencia: 25mM Tris-HCl pH 8, 150 mM NaCl, 20% v/v metanol.

Buffer separador 2X: 750 mM Tris-base, 0.2% p/v SDS, pH 6.8.

Buffer RIPA: 0.5M Tris-HCl, pH 7.4, 1.5M NaCl, 2.5% ácido desoxicólico, 10% NP-40, 10mM EDTA

Buffer TAE 0.5X: 40 mM Tris-HCl, 1 mM EDTA, 1.14 mL/L ácido acético glacial, pH 7.8.

Buffer TBS-Tween 1X: NaCl 150 mM, Tris-base 10 mM, 0.05% Tween 20, pH 7.4.

Buffer TE: Tris-HCl 10 mM, EDTA 1 mM, pH 8

Buffer TENS: Tris-HCl 10 mM, EDTA 1 mM, NaOH 0.1 N, 0.5% SDS.

MAB: Ácido Maleico 100mM, 150mM de NaCl

MEM: 0.1M MOPS (pH7.4), 2 mM EGTA, 1 mM MgSO₄

MEMFA: 4% FA, MEM 1X

PTw: 1X PBS, 0.1% Tween-20.

Rojo Ponceau: 0.1% p/v rojo ponceau, 5% v/v ácido acético glacial.

Solución Blanqueadora: SSC 0,5X, Formamida 5% y agua destilada 1%

Solución Bouin's: Para 100 ml, disolver 1 g de ácido pícrico en 70ml de agua y luego agregar 25 ml de formaldehído 37% y 5 ml de ácido acético glacial.

Solución Danilchik (DFA): 53 mM NaCl, 10 mM Na₂CO₃, 4.5 mM Gluconato de K, 1 mM MgSO₄, 1 mM CaCl₂, 0.1% BSA, 17,5 mM Bicina.

Solución de Ringer 1X (MR 1X): 0.1 M NaCl, 1.8 mM KCl, 2.0 mM CaCl₂, 1.0 mM MgCl₂, 5.0 mM Hepes-NaOH, pH 7.6 o 300 mg/l NaHCO₃

Solución I: 50 mM Glucosa, 10 mM EDTA, 25mM tris/HCl pH 8.

Solución II: 10 N NaOH, 20% SDS, H₂O.

Solución III: Acetato de potasio 3M

1.4. Anticuerpos

WB: “Western Blot”; HIS Hibridación *in situ*, I: Inmunofluorescencia.

- Anti-V5 (C-9) hecho en ratón (Santa-Cruz), WB: 1:100, I: 1:25.
- Anti-MRG1 hecho en ratón (Santa-Cruz), I: 1:25
- Anti- α -Tubulina hecho en ratón (Sigma), WB: 1:1000.
- Goat Anti-Mouse IgG (H+L)-HRP (Bio Rad), WB: 1:10000.
- Anticuerpo Anti-Digoxigenina acoplado a Fosfatasa Alcalina (Roche), HIS: 1:3000.
- Anticuerpo anti mouse Alexa fluor 488, Life Technologies, I: 1:200.

1.5. Datos públicos

Sets de datos de scRNA-seq de *X. tropicalis* de Briggs y col. (2018).

Disponible en: <http://tinyurl.com/scXen2018>.

2. Métodos

2.1. Diseño de oligonucleótidos

Se obtuvo la región codificante de *cited2* de *X. tropicalis* en NCBI (Gene ID: 407853) y Xenbase (XB-GENE-495035). A partir de esto, se utilizaron los softwares SnapGene y Primer BLAST de NCBI para generar los oligonucleótidos

descritos en la sección 1.1. Los tres oligonucleótidos poseen Tm similares, una cola GATTACA y sitios de corte para las enzimas de restricción EcoRI (GAATTC, oligonucleótido XtCited2 Forward) y XhoI (CTCGAG, oligonucleótidos XtCited2 Reverse 1 y 2).

2.2. Fertilización *in vitro*

Hembras y machos *X. tropicalis* fueron inyectados con hormona Gonadotrofina Coriónica Humana (HCG) antes de la fertilización con 20 u y 200 u de HCG la noche anterior y 4 horas antes, respectivamente. Para fecundar los huevos obtenidos de la hembra, el macho fue anestesiado con benzocaína y sacrificado para obtener sus testículos, los cuales fueron macerados en medio L15 (Lebovitz Medium) con 10% de FBS (Suero Fetal de Bovino) y añadido en gotas a los huevos. Posteriormente, se incubaron los embriones en solución 1/9X MR por 20 minutos, y luego se removió la gelatina que los recubre reemplazando el medio por una solución de cisteína al 3% en 1/9X MR pH 7.9 por 10 minutos. Posteriormente, los embriones fueron lavados 4 veces e incubados en 1/9X MR a 26°C hasta llegar al estadio deseado.

2.3. Purificación de ARN total desde embriones de *X. tropicalis*

Se utilizó el kit NucleoSpin RNA de Machery-Nagel para obtener ARN total de embriones, siguiendo las indicaciones del proveedor. Se seleccionaron 20 embriones de estadio deseado, se añadió 200 µL de buffer RA1 y se lisaron con

un macerador. Luego, se añadió 150 μL más de buffer, 3,5 μL de β -mercaptoetanol y se mezcló por vórtex.

El lisado fue pasado por la columna Nucleo Spin Filter (aro violeta) centrifugando el volumen completo a 11000 rpm por 1 min. El sobrenadante se traspasó a un tubo limpio, evitando el pellet, y se descartó la columna. Se añadió 350 μL de etanol 70% libre de RNAsas frío, se mezcló bien y se pasó el volumen por una columna Nucleo Spin RNA Column (aro celeste) con una centrifugación a 11000 rpm por 30 segundos. Se eliminaron las sales de la membrana añadiendo 350 μL de Membrane Desalting Buffer y centrifugando a 11000 rpm por 1 min.

En un tubo aparte se mezcló, por cada muestra, 10 μL de rDNAsa reconstituida y 90 μL de rDNase Reaction Buffer. Se añadió 95 μL de solución a cada membrana y se dejó incubar a temperatura ambiente (TA) por 15 min para digerir el ADN genómico. Se hizo un primer lavado agregando 200 μL de Buffer RAW2 y se centrifugó por 30 segundos a 11000 rpm. Se hizo un segundo lavado agregando 450 μL de buffer RA3 a la membrana y se centrifugó en las mismas condiciones. Este segundo lavado se realizó dos veces. Finalmente, se eluyó el ARN total con 60 μL de agua libre de nucleasas, centrifugando a 11000 rpm por 1 min.

Posterior a la purificación, se realizó una electroforesis en gel de agarosa 1% en buffer 0.5X TAE teñido con bromuro de etidio para comprobar la integridad del

ARN total. Para esto, se utilizaron 6 μL de RiboRuler High Range RNA Ladder de Thermo Scientific como estándar de peso molecular. Antes de cargar, tanto las muestras como el estándar fueron calentadas a 65°C por 5 min con el fin de romper posibles estructuras secundarias. Además, se cuantificó la concentración del ARN total en $\text{ng}/\mu\text{L}$ utilizando un espectrofotómetro. Los tubos de ARN total fueron marcados y guardados en refrigerador a -80°C .

2.4. Reacción de Transcripción Reversa

La reacción de transcripción reversa se realizó en 25 μL finales, conteniendo 2 μg ARN total como templado y 50 ng del oligonucleótido XtCited2 Reverse 2. Se utilizaron 10mM de mix de dNTPs, 5 μL de 5X Reaction Buffer, 0,6 μL de rRNAsin (40 $\text{u}/\mu\text{L}$) y 1 μL de enzima M-MLV (200 $\text{u}/\mu\text{L}$) de Promega. Luego, la mezcla fue llevada a 25 μL totales utilizando agua libre de ARNasas e incubada a 42°C por 1 hora. Se realizó un control negativo donde no se añadió la enzima M-MLV.

2.5. Reacción de PCR para prueba de oligonucleótidos con polimerasa GoTaq

Se utilizó la enzima polimerasa GoTaq Flexi (5 $\text{u}/\mu\text{L}$) de Promega para probar los oligonucleótidos específicos (XtCited2 Forward y XtCited2 Reverse 1) y determinar su T_m óptima. La reacción fue realizada en 20 μL totales, en agua libre de nucleasas, conteniendo 4 μL de 5X GoTaq Flexi Buffer, 0,4 μL de dNTPs (10 mM), 1 μL de cada oligonucleótido (que se encuentran a concentración 10

μM), 1,2 μL MgCl_2 (25 mM), y 1 μL de ADNc como templado. El programa utilizado contempló las siguientes etapas: 95°C por 5 min; 30 ciclos de 95°C por 1 min, 55°C por 2 min y 72°C por 3 min; y una extensión final de 72°C por 5 min. Se verificó la síntesis correcta del fragmento esperado de 704 pb mediante electroforesis en gel de agarosa 1% en buffer TAE 0.5X teñido con bromuro de etidio, utilizando GeneRuler 1 kb DNA Ladder de Thermo Scientific como estándar de peso molecular.

2.6. Reacción PCR con polimerasa de alta fidelidad

Para amplificar la región codificante de XtCited2 se utilizó la enzima PfuUltra II Fusion HS de Agilent. La reacción fue realizada en 30 μL totales en agua libre de nucleasas, conteniendo 3 μL de 10X PfuUltra II Reaction Buffer, 1 μL de dNTP (10 mM), 1 μL de cada oligonucleótido (que se encuentran a concentración 10 μM), 1 μL de enzima PfuUltra II Fusion HS y 1 μL de ADNc como templado. El programa utilizado contempló las siguientes etapas: 95°C por 3 min; 20 ciclos de 95°C por 1 min, gradiente de 55°C por 1 min y 72°C por 3 min; y una extensión final de 72°C por 5 min. Se verificó la síntesis correcta del fragmento esperado de 704 pb mediante electroforesis en gel de agarosa 1% en buffer TAE 0.5X teñido con bromuro de etidio, utilizando GeneRuler 1 kb DNA Ladder de Thermo Scientific como estándar de peso molecular.

2.7. Purificación de producto PCR

Para purificar el producto PCR, se utilizó el protocolo de purificación de ADN del kit Monarch® PCR y DNA Cleanup de New England Biolabs. La totalidad del volumen PCR se diluyó con Binding Buffer a razón 1:5. Posteriormente, se traspasó progresivamente el volumen total a la columna y se centrifuga a 11000 rpm por 1 min a TA (máximo 700 µL a la vez). Se lavó el ADN con 200 µL de Wash Buffer y centrifugando a 11000 rpm por 1 min a TA, dos veces. Se transfirió la columna a un tubo Eppendorf limpio y se añadieron 15 µL de agua libre de nucleasas a 50°C, justo al centro de la membrana de la columna. Se dejó reposar 5 min y se centrifugó a 11000 rpm por 1 min a TA. Se repitieron los últimos pasos para obtener una segunda elución, y ambas fueron almacenadas a -20°C.

2.8. Digestión con enzimas de restricción

Para realizar las digestiones analíticas, en un volumen de 20 µL, se utilizó 1 unidad de la enzima de restricción respectiva, 0.1 volúmenes de Buffer 10X correspondiente y, de ser requerido, 0.1 volúmenes de BSA 10X. La reacción se incubó por 4 horas a 37°C y se analizó el patrón de digestión mediante electroforesis en gel de agarosa 1% en buffer TAE 0.5X teñido con bromuro de etidio, utilizando GeneRuler 1 kb DNA Ladder de Thermo Scientific como estándar de peso molecular. Las digestiones preparativas fueron desarrolladas en un volumen final de 100 µL con 10 µg de ADN.

2.9. Extracción de ADN desde geles de agarosa

Para purificar el ADN digerido, se utilizó el método de extracción desde geles de agarosa utilizando el protocolo del kit E.Z.N.A. Gel Extraction Kit de Omega Bio-Tek. La banda de ADN desde gel fue observada y cortada en el transiluminador, colocada en tubos Eppendorf y pesada. Por cada 0.1 g de gel, se añadió 0.1 mL de XP2 Binding Buffer y se dejó a 60°C hasta que el gel se disolviera completamente, agitando cada 2 min. El volumen total se pasó progresivamente por la columna (volumen máximo 700 µL), centrifugando a 10000 rpm por 1 min a TA. Se descartó el sobrenadante, se añadieron 200 µL del mismo buffer, y se volvió a centrifugar. Posteriormente, se lavó la membrana con 700 µL de SPW Buffer centrifugando en las mismas condiciones, dos veces. La membrana se secó con una última centrifugación de 2 min y se procedió a eluir dos veces con 15 µL de agua libre de nucleasas a 50°C, centrifugando a 10000 rpm por 1 min cada vez después de incubar por 10 min. El volumen resultante se almacenó a -20°C.

2.10. Desfosforilación de vector plasmidial

Para evitar religación del vector linealizado, se procedió a desfosforilar sus extremos. Para esto, se utilizó el total del volumen purificado por gel de vector plasmidial digerido, 10 U de Fosfatasa Antártica New England Biolabs y 0.1 volúmenes de 10X Reaction Buffer y agua libre de nucleasas hasta un volumen total de 120 µL. La reacción se incubó por 2 horas a 37°C y se almacenó a -20°C.

2.11. Ligación convencional

Para insertar la región codificante de XtCited2 al vector, se utilizaron 100 ng de vector y una cantidad de inserto correspondiente a razón molar inserto/vector de 3:1. A esto se le añadieron 5 U de T4 DNA Ligase de Thermo Scientific, 0.1 volúmenes de buffer de reacción y agua libre de nucleasas hasta un volumen total de 11 μ L. La reacción se incubó a 16° por 16 horas y se almacenó a -20°C.

2.12. Ligación en TOPO

Se realizó una ligación en el vector Zero Blunt TOPO de Invitrogen utilizando el protocolo recomendado por el fabricante. Se mezcló 1 μ L del vector con 50 ng de inserto y 1 μ L de solución salina (1.2 M NaCl, 60 mM MgCl₂) en un volumen final de 6 μ L. Se incubó la reacción a TA por 5 min y se utilizó inmediatamente para transformación bacteriana.

2.13. Transformación bacteriana

50 μ L de bacterias *E. coli* DH5 α quimiocompetentes fueron mezcladas con 10 ng del plásmido de interés. Se dejaron en hielo por 30 min, se les sometió a shock térmico de 42°C por 2 minutos exactos y se volvieron a dejar en hielo por 5 min más. Posteriormente, se añadieron 450 μ L de medio LB sin antibiótico y se incubaron en agitación a 37°C por 1 hora. Pasado el tiempo, se centrifugaron las bacterias a 5000 rpm por 5 min, se eliminó el sobrenadante, dejando aproximadamente 100 μ L, en los cuales se resuspendieron las bacterias. Se sembró este volumen en placas Petri LB/Agar con antibiótico correspondiente

(100 µg/mL de Ampicilina o 50 µg/mL de Kanamicina) con ayuda de perlas de vidrio estériles. La placa se incubó a 37°C por toda la noche (12-16 horas).

Las colonias seleccionadas fueron inoculadas en 5 mL de medio LB con antibiótico correspondiente (100 µg/mL de Ampicilina o 50 µg/mL de Kanamicina) y se crecieron en agitación a 37°C toda la noche (12-16 horas).

Finalmente, 1 mL de cultivo fue traspasado a un matraz con 250 mL de medio LB con antibiótico y se dejaron crecer en agitación a 37° por toda la noche. Finalmente, el volumen total se centrifugó a 5000 rpm por 10 min a 4°C, se eliminó el sobrenadante y se conservó el pellet a -20°C.

2.14. Purificación de ADN plasmidial por minipreparación con buffer TENS

Para analizar las colonias inoculadas antes de pasarlas a matraz, se separó 1-1.5 mL de medio de cultivo en un tubo Eppendorf y se dejó en hielo por 10 min. El volumen se centrifugó a 5000 rpm por 15 min a 4°C. Se desechó el sobrenadante, dejando 150 µL y se resuspendió el pellet. Se realizó la lisis de las bacterias añadiendo 300 µL de buffer TENS, invirtiendo delicadamente 3 veces y dejando en hielo por 5 min exactos, invirtiendo de nuevo a la mitad del tiempo. La lisis se detuvo añadiendo 200 µL de NaOAc 3M, mezclando inmediatamente y dejando en hielo por 15 min. Los tubos se centrifugaron a 10000 rpm por 15 min a 4°C, se pipeteó el sobrenadante a un tubo limpio y se agregó 1 mL de etanol 100% frío. Se mezcló por inversión y se dejó reposar los tubos a -20°C por

30 min. Posteriormente, se centrifugó a 10000 rpm por 15 min a 4°C y se lavó el pellet con 1 mL de etanol 70% frío, centrifugando nuevamente. Se resuspendió el pellet en 50 µL de agua suplementada con 10 µg/µL de RNAsa A y se incubó por 30 min a 37°C. Finalmente, se centrifugó a 10000 rpm por 15 min a 4°C y se rescató cuidadosamente el sobrenadante pipeteando a un tubo nuevo. Se almacenó a -20°C.

2.15. Purificación de ADN a pequeña escala

Para purificar ADN de forma más rápida y limpia (para, por ejemplo, secuenciación) se utilizó el protocolo del kit Wizard Plus SV Miniprep DNA Purification System de Promega. Para esto, se separó 1-1.5 mL de medio de cultivo en un tubo Eppendorf y se dejó en hielo por 10 min. El volumen se centrifugó a 5000 rpm por 10 min a 4°C. Se desechó el sobrenadante y se resuspendieron las bacterias en 250 µL de Cell Resuspension Buffer y se mezcló por vórtex. Posteriormente, se añadieron 10 µL de Proteasa Alcalina (incluida en el kit) y 250 µL de Cell Lysis Buffer, invirtiendo 4 veces. Se incubó a TA por 5 min y se detuvo la lisis con 350 µL de Neutralization Solution, invirtiendo 4 veces y dejando en hielo por 10 min. Se centrifugó el lisado a 10000 rpm por 10 min a 4°C. Se transfirió el sobrenadante (aproximadamente 860 µL) a un tubo limpio y después se pasó el volumen por columna centrifugando a 10000 rpm por 1 min, 430 µL a la vez. Se lavó la membrana 2 veces con 500 µL de Column Wash Solution, centrifugando a 10000 rpm por 5 min a TA cada vez. La membrana se

secó con una última centrifugación en las mismas condiciones y se procedió a eluir dos veces con 30 μ L de agua libre de nucleasas a 50°C, incubando por 10 min antes de centrifugar a 10000 rpm por 5 min. Los eluatos se almacenaron a -20°C.

2.16. Purificación de ADN plasmidial a gran escala

El pellet obtenido de la transformación bacteriana se resuspendió en 5 mL de Solución I. Luego se agregó 10 mL de Solución II, se mezcló y se incubó a TA por 10 minutos. Finalmente, se agregó 5 mL de Solución III, se mezcló con suavidad y se dejó en hielo por 10 min.

El lisado obtenido se centrifugó a 8000 rpm por 15 minutos a 4°C; se conservó el sobrenadante, se hizo pasar por gasa y se agregó 0,6 volúmenes de isopropanol. Se dejó en reposo por 10 minutos a TA. Se recuperó el pellet centrifugando a 8000 rpm por 15 min a 4°C, se lavó el pellet con etanol 70% y se disolvió en 3 mL de Buffer TE pH 8. Se agregó 3 mL de LiCl 5M, se dejó reposar 10 min a TA y se centrifugó a 8000 rpm por 10 min a 4°C. Al sobrenadante se le agregó igual volumen de isopropanol y se centrifugó a 8000 rpm por 15 min a 4°C. Se transfirió el sobrenadante a un tubo nuevo y se lavó el pellet con etanol 70% frío.

Se resuspendió el pellet en 500 μ L de buffer TE suplementado con 20 μ g/mL de RNAsa A, se transfirió a un tubo Eppendorf y se incubó a 37°C por 30 min. Se agregó 500 μ L de PEG 8000 13% p/v NaCl 1,6 M, se incubó a TA por 10 min y

se centrifugó a 12000 rpm por 10 min a 4°C. El precipitado fue lavado con 1mL de etanol 70% y resuspendió en 400 uL de Buffer TE. Esta solución fue mezclada con 1 volumen de fenol/cloroformo para la extracción de las proteínas, mezclando vigorosamente y luego centrifugando a 6000 rpm por 5 min, pasando la fase acuosa a un tubo limpio. Este proceso se repitió 3 veces. Finalmente, se añadió 1/10 volúmenes de NaOAc 3M y dos volúmenes de etanol 100%, se mezcló bien y se dejó precipitar el ADN plasmidial a -20°C por toda la noche. Se recuperó el pellet centrifugando a 12000 rpm por 10 min a TA. El pellet se lavó con 1 mL de etanol 70% frío y se resuspendió en 100 µL de agua libre de nucleasas. La muestra se almacenó a -20°C.

2.17. Síntesis de ARNm y sondas de ARN marcadas con digoxigenina

Para la síntesis de sondas y ARN mensajeros, primero se linealizó el ADN plasmidial que contiene la región codificante de XtCited2 con la enzima de restricción que se encuentra cerca del extremo 5' o 3', las cuales generarán ARN antisentido (para sonda) y sentido (para ARNm y sonda control). Para los cortes en 5', se utilizó la enzima BamHI, mientras que los cortes en 3' fueron hechos con NotI. El fragmento fue purificado mediante el kit E.Z.N.A. Gel Extraction Kit de Omega Bio-Tek. El ADN purificado fue cuantificado por gel de agarosa a través de densitometría comparando con marcador de peso molecular (GeneRuler 1kb DNA Ladder de Thermo Scientific) empleando el software Image J (Schneider y col., 2012). Luego, se utilizó 1.5 µg de ADN purificado y linealizado

como templado para comenzar la transcripción *in vitro*. La transcripción de las sondas se llevó a cabo en presencia de dig-UTP (nucleótidos de uridina marcados con digoxigenina) junto al resto de la mezcla de nucleótidos y la ARN polimerasa correspondiente, según las especificaciones de la enzima en cuestión (Sp6 o T7 de Promega). Por otro lado, para la generación de ARN mensajero se empleó el kit de transcripción *in vitro* mMMESSAGE mMACHINE SP6, según recomendaciones de la empresa (Life Technologies). Ambas reacciones se realizaron en agua libre de nucleasas en un volumen total de 20 μ L, incubando a 37°C por 2 horas. Para remover el ADN inicial, se incubó por 30 min 1 μ L de DNAsa turbo kit y se precipitó el ARN añadiendo 30 μ L de solución de precipitación (7.5M LiCl, 50mM EDTA) e incubando a -20°C toda la noche. Al día siguiente se centrifugó la muestra de ARN dos veces a 11000 rpm por 5 min c/u a 4°C, con un lavado con etanol 70% entremedio y se conservó el pellet. El pellet se resuspendió en 10 μ L de agua libre de nucleasas, cuantificado mediante electroforesis en gel de agarosa densitometría en Image J comparando con RiboRuler High Range RNA Ladder de Thermo Scientific y alícuotado en tubos de 0.2 mL, haciendo alícuotas de 1 μ L a 500 ng/ μ L (ARNm) o 350 ng/ μ L (sondas). Las alícuotas se almacenaron a -80°C.

2.18. Fijación de embriones con MEMFA

Se realizó fertilización *in vitro* con la metodología descrita anteriormente para obtener embriones en distintos estadios del desarrollo, descritos por Nieuwkoop

y Faber, 1994. Nieuwkoop y Faber, 1994. Para esto, los embriones en estadios seleccionados fueron fijados mediante incubación en MEMFA en agitación a TA por 4 horas o a 4°C toda la noche. Se realizaron 3 lavados con etanol 100% por 5 minutos cada uno y se almacenaron a -20°C en etanol 100%.

2.19. Hibridación *in situ* en embrión completo

Mediante esta técnica, se visualizó la expresión espaciotemporal de *XtCited2* durante el desarrollo embrionario de *X. tropicalis* desde blástula tardía (NF 9) hasta finales de la organogénesis (NF 45) mediante una sonda de ARN antisentido para el ARNm. Además, se utilizó una sonda sentido para el ARNm con el fin de servir como control negativo. La sonda se sintetizó mediante transcripción *in vitro* en presencia de dig-UTP (nucleótidos de uridina marcados con digoxigenina) y posteriormente se purificó la sonda precipitando con LiCl, como se describió previamente.

El protocolo de hibridación *in situ* utilizado fue una modificación del descrito en el sitio web Harland Lab *X. tropicalis* previamente estandarizado en el laboratorio. Este protocolo consistió en una etapa de hidratación de embriones previamente fijados para su tratamiento posterior con proteinasa K, que permeabiliza y permite la mejor penetración de la sonda. Luego, se incubó con la sonda correspondiente y, posteriormente, con el anticuerpo anti-digoxigenina conjugado con fosfatasa alcalina. El revelado se realizó con NBT/BCIP ("Nitro-Blue Tetrazolium Chloride/5-Bromo-4-Chloro-3'-Indolyphosphate p-Toluidine

Salt”) en Buffer AP. La reacción cromogénica se detuvo con MAB, y los embriones se fijaron nuevamente con solución de Bouin, que, además, estabilizó la tinción de los embriones. Después, los embriones fueron blanqueados en solución blanqueadora incubando bajo luz blanca a TA, lo cual permitió la eliminación de los pigmentos naturales de los embriones. Finalmente, se realizaron lavados con SSC 1X y se almacenaron a -20°C en etanol 100%.

2.20. Microinyección de ARNm

Los huevos fertilizados fueron microinyectados utilizando el equipo de microinyección Narishige IM300 de Narishige Co., Japón. Se utilizaron agujas de borosilicato para microinyectar una solución de ARNm y Rodamina Dextrano (Life Technologies) o ARNm de H2B-mCherry como trazador de linaje en ambas células de estadio NF 2. La microinyección se realizó en medio MR 1/9X con 3% de Ficoll (pH 7.4). Los embriones fueron traspasados a placas de agarosa al 1% en el mismo medio y, después de 1-2 horas, se cambió por medio MR 1/9X. Los embriones se mantuvieron a 26°C hasta que llegaron a néurula tardía (estadios NF 16-22) al día siguiente.

2.21. Lisis de embriones

Se seleccionaron 20 o más embriones correctamente microinyectados y fueron lisados usando 10 µL de buffer RIPA por embrión. Al volumen total se le añadió 0.01 volumen de inhibidores de proteasas y fosfatasa (Roche). Para separar los restos de huevo y “jelly”, se centrifugó la muestra 2 veces a 2000 rpm

por 5 min a 4°C, traspasando el sobrenadante transparente a un tubo nuevo cada vez. Se cuantificó la concentración total de proteínas mediante curva de Bradford.

2.22. Geles de Poliacrilamida SDS-PAGE

Se prepararon geles denaturantes de poliacrilamida al 12% de 1.5 mm de grosor según la siguiente tabla:

	Gel separador	Gel concentrador
Buffer separador 2X	4 mL	-
Buffer concentrador 2X	-	1.5 mL
Acrilamida 30%	3.2 mL	0.6 mL
H2O	0.8 mL	0.9 mL
TEMED 100%	20 µL	10 µL
APS 12.5%	40 µL	20 µL
Volumen Total	8 mL	3 mL

Tabla 1: Protocolo para preparar geles denaturantes al 12%.de 1.5 mm de espesor. Utilizar la mitad de los volúmenes para obtener un grosor de 0.75 mm.

2.23. Electroforesis en gel de Poliacrilamida SDS/PAGE

A un volumen correspondiente a 100 µg de proteína se le agregó 0.25 volúmenes de buffer de carga de proteínas 5X y 0.01 volúmenes de DTT 1M. Las muestras se denaturaron a 95°C por 10 min, se cargaron al gel de poliacrilamida 12% en conjunto al marcador de peso molecular PageRuler™ Plus Prestained Protein Ladder (10 to 250 kDa) de Thermo Scientific y se corrieron en buffer de corrida 1X a 100V por 1.5-2 horas, o hasta que se observa que la muestra corrió lo suficiente.

2.24. Western Blot

Posterior a la electroforesis, se transfirieron las proteínas a una membrana de nitrocelulosa de Hybond en buffer de transferencia 1X a 4°C durante 2 horas a un amperaje constante de 0.25 A o 16 horas a voltaje constante de 30 V a 4°C. Se comprobó la correcta transferencia de las proteínas utilizando tinción rojo Ponceau y se bloquearon los sitios inespecíficos de la membrana con 5% de leche en buffer TBS-Tween 1X por al menos 1 hora a TA ambiente o toda la noche a 4°C. Finalmente, se realizó el Western Blot incubando con anticuerpo primario en buffer TBS-Tween 1X por toda la noche a 4°C. Para retirarlo, se hicieron 3 lavados 10 min, y se incubó con anticuerpo secundario por 2 horas; en ambos casos, se realizó a TA y en agitación. Para remover el exceso de anticuerpo, se repitieron los mismos 3 lavados. Para observar el resultado, se utilizaron, en proporción 1:1, los líquidos de revelado del kit Western Lighting

Plus-ECL de PerkinElmer Inc. y el equipo iBright CL750 Imaging System de Invitrogen.

2.25. Extracción de explantes de cresta neural

Embriones en estadio de néurula (NF 17) se colocaron en medio MR 1/9X y se les removió la membrana vitelina utilizando pinzas n° 5 (“Fine science tools” N° 11252-00). Luego, los embriones fueron inmovilizados en placas con plastilina y se extrajeron las células de la cresta neural utilizando un pelo de ceja unido a con cera de vela a una pipeta pasteur de vidrio. Los explantes fueron colectados en placas pequeñas con medio DFA 1X (53 mM NaCl, 10 mM Na₂CO₃, 4.5 mM Gluconato de K, 1 mM MgSO₄, 1 mM CaCl₂, 0.1% BSA, 17,5 mM Bicina) por un máximo de 30 min antes de ser traspasado a un cubreobjetos de 12 mm con fibronectina (50 µg/mL) en placas de 4 pocillos con DFA 1X. Se esperaron 30 min para que las células se adhirieran al sustrato y luego se dejaron migrar por 4 horas. Los explantes fueron fijados en la misma placa con DFA/16% FA por 20 min y luego PBS 1X/3.7% FA por 20 min. Luego, se realizaron 3 lavados de 5 min con PBS 1X y se guardaron las placas envueltas en papel aluminio a 4°C para el siguiente ensayo.

2.26. Ensayos de inmunofluorescencia de explantes de cresta neural

Antes de comenzar, se realizaron 3 lavados con PBS 1X por 5 min a TA. Se permeabilizó por 10 min exactos con PBS 1X/0.3% Tritón X-100 (previamente enfriado y con la placa sobre hielo). Se hicieron otros 3 lavados de 5 min con PBS

1X a TA antes de proceder con el bloqueo con PBS 1X suplementado con 2% de suero de oveja por al menos 30' a TA. Posteriormente, se incubó el anticuerpo primario por 4 horas a TA o toda la noche a 4°C en cámara húmeda. Para remover el anticuerpo primario, se realizaron 10 lavados de 5 min con PBS 1X a TA, y se procedió a incubar el anticuerpo secundario y Hoescht (tinción nuclear) por 2 horas en cámara húmeda. El anticuerpo secundario se removió con otros 10 lavados de 5 min de PBS 1X a TA y se montaron las muestras sobre un portaobjeto de vidrio con 7 uL de Dako fluorescent mounting medium.

2.27. Obtención de niveles de expresión de *xtcited2* y *xtric-8a* desde data pública de RNA-seq de célula única

Para complementar los análisis anteriores, se compararon los resultados con datos de RNA-seq de célula única de *X. tropicalis*, evaluando la expresión de *xtcited2* y *xtric-8a* en distintos estadios y tejidos. Estos datos fueron descargados de la página indicada en la publicación de Briggs y col. (2018), donde los datos se encuentran previamente procesados y anotados según estadio NF (de 8 a 22) y tejido en un navegador interactivo en línea (*Xenopus Jamboree*, tinyurl.com/scXen2018). Para descargar los datos, se ingresó a la página correspondiente a cada estadio y se seleccionaron los tejidos de interés en el panel de la izquierda. Las matrices de expresión de las células correspondientes se descargaron con la opción “SPRING data for selection” dentro del menú desplegable de “Download”.

Se descargaron los sets de datos correspondientes a blástula (NF 8), neuroectodermo (NF 10 y 11), borde placa neural (NF 12), cresta neural (NF 13 14 y 16) y cresta neural craneal (NF 18, 20 y 22).

De estos datos, es necesario extraer los valores de expresión de cada gen. Para esto, se utilizaron las matrices de expresión en UMIFM (UMI Filtered-Mapped) de cada tejido y estadio de interés. Estas se encuentran en archivos .csv, donde se separan por comas [identificador del gen | nombre del gen], y la expresión en UMIFM de cada célula evaluada. Para darles formato a las tablas y poder extraer los niveles de expresión, se abrieron los archivos con Excel y se empleó la herramienta "Datos a columnas" para dividir la información en celdas. Debido a que "nombre del gen" posee elementos repetidos, pero utiliza una nomenclatura más legible, se añadió una tercera columna llamada "gen único" que diferencia los genes de igual nombre añadiéndoles un guion bajo y un número para diferenciarlos. Al final del análisis, se realizó una revisión manual de los genes para poder nombrarlos siguiendo la nomenclatura disponible en la base de datos Xenbase.

Para obtener el valor de expresión promedio de *xtric-8a* y *xtcited2*, se realizó un filtro de las células positivas a cada uno de estos genes. Para determinar diferencias significativas, primero se aplicó la prueba de Shapiro-Wilk para determinar si la expresión de cada gen por estadio seguía una distribución

normal. En casos donde la distribución era normal, se aplicó la prueba estadística de t-Student, y en el resto, la prueba de Mann Whitney U.

2.28. Modelamiento con aprendizaje automático de datos de subpoblaciones de cresta neural.

Mediante la plataforma de Briggs y col. antes mencionada, se descargó el set de datos de cresta neural total de estadio NF18, correspondiente a la combinación de los sets de cresta neural craneal y cresta neural cordal del mismo estadio. El análisis de estos datos y la generación de los gráficos fue realizado en el lenguaje de programación Python 3.11.11, utilizando los siguientes paquetes: Scanpy 1.11.0 (Wolf y col., 2018), NumPy 1.26.4 (Harris y col., 2020), Matplotlib 3.10.0 (Hunter, 2007), Pandas 2.2.2 (McKinney, 2010), SciPy 1.14.1 (Virtanen y col., 2020), Scikit-learn (Pedregosa y col., 2011), Statsmodels (Seabold y Perktold, 2010) y Seaborn 0.13.2 (Waskom, 2021).

Una vez descargados los datos, se ejecutó análisis de componentes principales (PCA) (Maćkiewicz y Ratajczak, 1993) mediante el módulo de Scanpy. Se seleccionaron los primeros 200 componentes principales, que representan aproximadamente el 80% de la varianza acumulada. Con esto, se generaron múltiples mapas UMAP (McInnes y col., 2018) con distintos valores de “*n_neighbors*” (5, 10, 15, 30, 25 y 50) y se realizó una búsqueda de clústeres mediante el algoritmo de Leiden (Traag y col., 2019) con valores de resolución de 0.5, 0.75, 1, 1.5 y 2. Ambos análisis fueron implementados con las funciones

correspondientes de Scanpy. Se seleccionaron manualmente los mejores agrupamientos según su patrón de agrupamiento para continuar con el primer análisis de expresión diferencial. Se utilizó la función “*rank_gene_groups*” de Scanpy, para identificar genes diferencialmente expresados entre cada clúster y la suma del resto de las células del set de datos, aplicando la prueba de Wilcoxon. Se seleccionaron todos los genes que poseían un valor p ajustado < 0.05 y un $\log_2FC > 0$ para construir una lista preliminar de genes de interés.

Con esta lista y las células categorizadas en clústeres, se construyó un modelo de clasificación de *Random Forest* (Breiman, 2001) utilizando la función “*RandomForestClassifier*” de Scikit-learn con el objetivo de identificar los genes que mejor separan los clústeres. Este algoritmo está diseñado para reconocer patrones complejos dentro de un set de datos y, a partir de ellos, clasificarlos. Además, nos indica cuánto contribuye una variable a mejorar la precisión del modelo mediante el parámetro de “importancia”. Para optimizar el rendimiento del modelo, se realizó una búsqueda de hiperparámetros para cada muestra mediante “*RandomizedSearchCV*”, evaluando distintos tamaños de set de datos de prueba (20%, 40% y 50%) y los siguientes parámetros de búsqueda: *n_estimators: randint(100, 200); max_depth: None, 5, 10, 20, 30, 40; min_samples_split: randint(2, 5); min_samples_leaf: randint(1, 5), max_features: auto, sqrt, log2, None*. El modelo se iteró 100 veces y se seleccionaron los parámetros que daban mayor porcentaje de precisión. Con estos parámetros

optimizados, se reentrenó el modelo final, a partir del cual se generaron matrices de confusión y la lista de "*feature_importances*", que ordena aquellos genes según su contribución al modelo. Se seleccionaron todos los genes con valor de Importancia mayor a 0.

Finalmente, con los genes seleccionados, se realizó un segundo análisis de expresión diferencial, esta vez entre pares de clústeres. Se obtuvieron valores de log2FC, p-value (mediante test de Mann-Whitney U de SciPy) y p-adj (utilizando corrección FDR de Statsmodels). Con esta información, se categorizó cada gen según si mostraba expresión diferencial específica en un clúster individual o una combinación de clústeres, generando así una lista de posibles genes marcadores que se utilizó para el análisis bibliográfico posterior.

RESULTADOS

1. Objetivo 1: Caracterizar el patrón de expresión espaciotemporal de *xtcited2* y comparación con el patrón de expresión espaciotemporal de *xtric-8a*.

Para analizar el patrón de expresión espaciotemporal de *xtcited2* se ejecutaron tres técnicas de biología molecular: hibridación *in situ*, Western blot e inmunofluorescencia en explantes de CN. Los constructos plasmidiales utilizados en este objetivo fueron previamente generados en nuestro laboratorio.

1.1. Caracterización espaciotemporal de la expresión de transcritos de *xtcited2* en distintos estadios del desarrollo embrionario.

Para la generación de sondas de ARN marcadas con digoxigenina, se utilizó el constructo pCR Blunt-II-TOPO/XtCited (Figura 6A). Este plásmido contiene dos sitios promotores, uno para la polimerasa T7 en el extremo 3' del gen, que fue utilizado para la generación de la sonda antisentido; y uno para la polimerasa SP6 en el extremo 5' del gen, que fue utilizado para la generación de la sonda sentido. En la Figura 6B, podemos observar la síntesis exitosa de ambas sondas, observándose un rendimiento mayor en la sonda sentido.

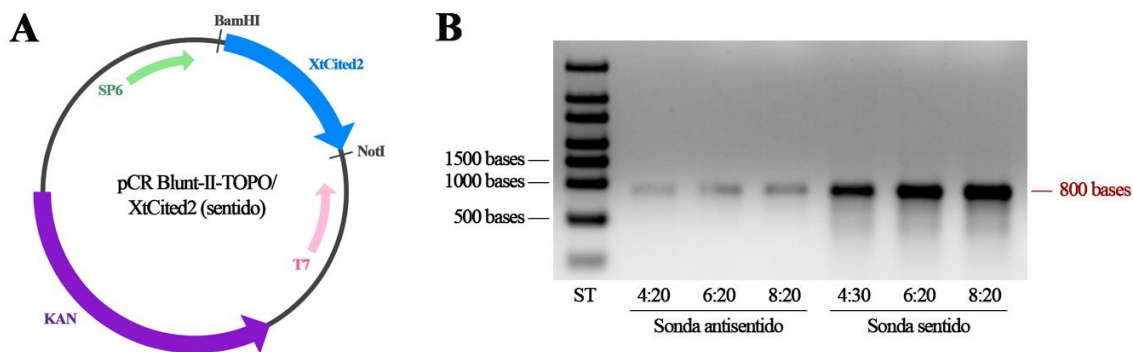


Figura 6: Síntesis de sondas para ensayos de hibridación *in situ*. (A) Esquema del vector pCR Blunt-II-TOPO/XtCited2. En azul se esquematiza el fragmento correspondiente al gen *xtcited2*, en verde claro la región promotora para SP6, en rosado la región promotora para T7, en violeta el gen de resistencia a antibiótico (Kanamicina) y las líneas negras perpendiculares representan los sitios de corte para las enzimas de restricción BamHI (en 5') y NotI (en 3'). (B) Cuantificación de las sondas de ARN mediante gel de agarosa. Se hizo una dilución 1:20 de cada sonda sintetizada y se cargaron 4, 6 y 8 μ L de éstas. A la izquierda se muestra el estándar de tamaño molecular RiboRuler HighRange y a la derecha el tamaño esperado del ARNm, ambos en número de bases.

Se analizó la expresión espaciotemporal de *xtcited2* en embriones desde estadios NF 9 a 45. Para esto, se utilizó la técnica de hibridación *in situ* en embrión completo y dos sondas: una antisentido, que marcará los tejidos específicos donde está presente el transcrito, y una sentido que servirá como

control negativo. Para cada sonda y estadio, se utilizó un mínimo de 20 embriones.

En los estadios tempranos analizados, observamos una alta expresión en el polo animal del embrión, incluyendo la zona marginal dorsal en el estadio NF 10 (Figura 7). En estadios de néurula (Figura 8), la expresión es predominante en tejidos neurales, tales como la placa neural, tubo neural y CN. Esta expresión puede ser distinguida en las rutas migratorias de la CN craneal y el prosencéfalo en estadios NF 22 y 24 (Figura 8D-E). En estos estadios del desarrollo, también se observa expresión en el mesodermo presomítico (Figura 8D-E). Posteriormente, durante la organogénesis (Figura 9), se observa expresión de *xtcited2* en el prosencéfalo (Figura 9A), el cual se convierte posteriormente en el cerebro (Figura 9B-E). *Xtcited2* también se expresa en la región intersomítica y tejido cardíaco, tal como el primordio cardíaco (Figura 9A) y el corazón adulto (Figura 9D-E).

Es interesante destacar que esta expresión espaciotemporal es similar a la observada para el gen de *xtric-8a*, la cual ha sido previamente descrita en nuestro laboratorio por Maldonado-Agurto y col. (2011) (Anexo 1 y Anexo 2). Comparando ambos experimentos, se observa que ambos genes se expresan en tejido neural, con *xtcited2* teniendo una expresión más generalizada que *xtric-8a*, que se expresa de forma muy marcada en vesícula ótica, tubo neural, y tejidos derivados (Anexo 2g-h). Además, si bien no se observa expresión de *xtric-8a* en

regiones cardiacas del embrión, sí se ha determinado su presencia en corazón de *Xenopus* adulto (Maldonado-Agurto y col., 2011). En los Anexo 4 y Anexo 5 se pueden observar los patrones de expresión lado a lado.

De forma interesante, si bien *xtric-8a* no presenta expresión en mesodermo presomítico, sus principales interactores, las subunidades α de la proteína G heterotrimérica, sí se expresan en esta zona (Fuentealba y col., 2016) (Anexo 3).

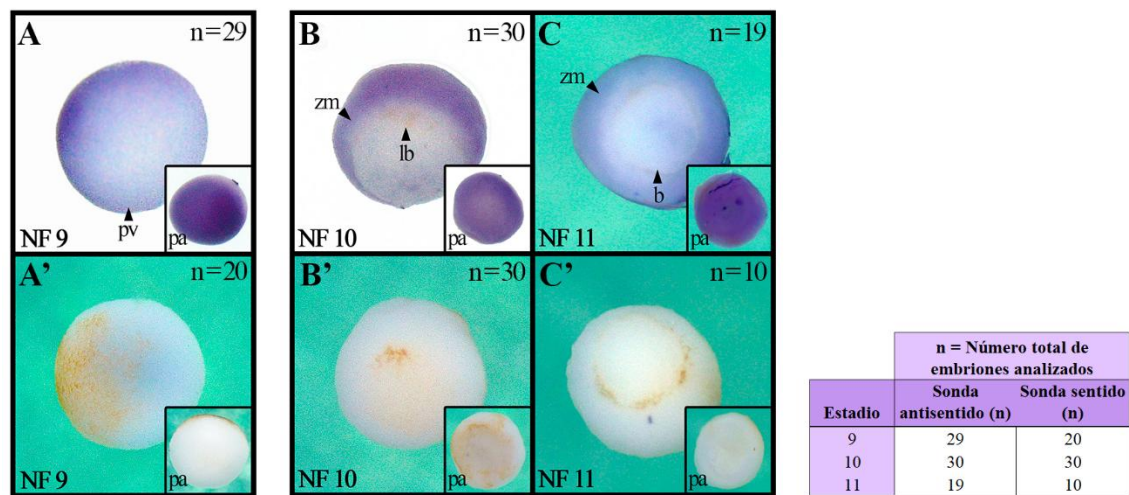


Figura 7: Patrón de expresión de *xtcited2* en estadios de desarrollo temprano de *X. tropicalis*. Ensayos de hibridación *in situ* en embrión completo permiten observar la expresión de *xtcited2* durante la blástula tardía (NF 9) y gástrula (NF 10-11) de *X. tropicalis*. Los estadios NF mostrados son indicados en la esquina inferior izquierda de cada panel. El número de embriones analizados se indican en la esquina superior derecha de cada panel, y se encuentran

tabulados en la tabla a la derecha de la figura. Se utilizó una sonda antisentido para analizar la expresión espaciotemporal de *xtcited2* (**A-C**), y una sonda sentido como control negativo (**A'-C'**). Se muestran vistas vegetales y en el panel inserto se muestra la vista del polo animal de cada embrión. Las cabezas de flecha negras indican estructuras específicas rotuladas como: **pv** = polo vegetal; **pa** = polo animal; **zm** = zona marginal; **lb** = labio del blastoporo y **b** = blastoporo.

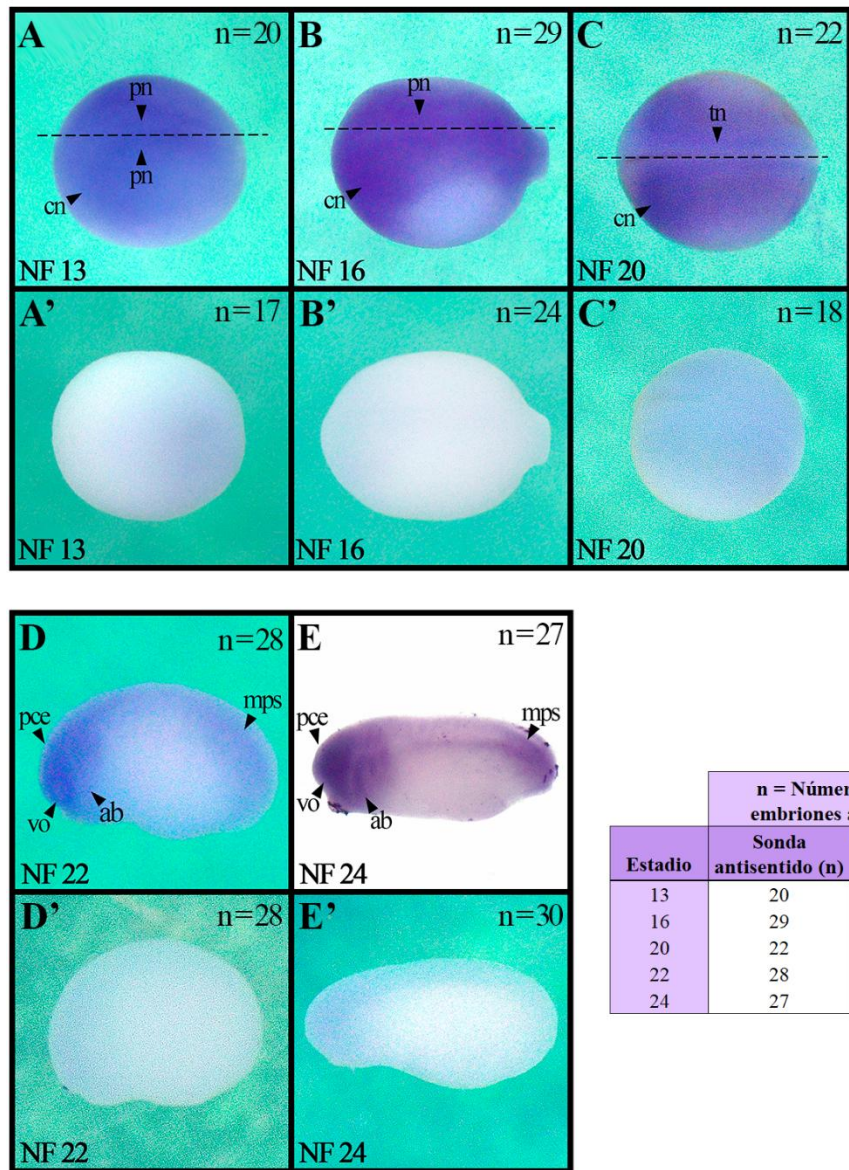


Figura 8: Patrón de expresión de *xtcited2* durante la neurulación de *X. tropicalis*. Ensayos de hibridación *in situ* en embrión completo permiten observar la expresión de *xtcited2* desde la neurulación (NF 13) hasta inicios de la organogénesis temprana (NF 24) de *X. tropicalis*. Los estadios NF mostrados son indicados en la esquina inferior izquierda de cada panel. El número de embriones

analizados se indican en la esquina superior derecha de cada panel, y se encuentran tabulados en la tabla a la derecha de la figura. Se utilizó una sonda antisentido para analizar la expresión espaciotemporal de *xtcited2* (**A-E**), y una sonda sentido como control negativo (**A'-E'**). (**A-C, A'-C'**) Vista dorsal. Línea punteada indica la línea media de la placa o tubo neural. (**D-E, D'-E'**) Vista lateral. Las cabezas de flecha negras indican estructuras específicas rotuladas como: **cn** = cresta neural; **pn** = placa neural; **tn** = tubo neural; **pce** = prosencéfalo; **vo** = vesícula óptica; **ab** = arcos branquiales y **mpps** = mesodermo presomítico.

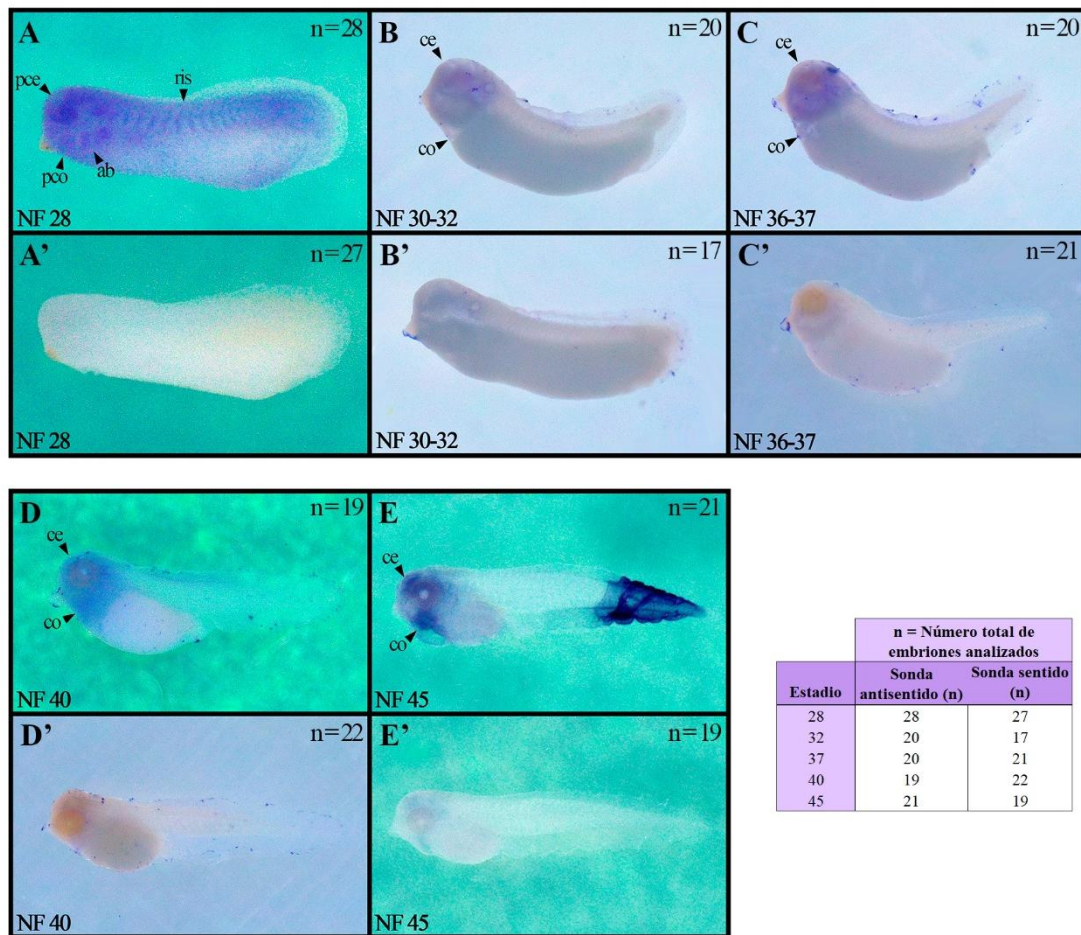


Figura 9: Patrón de expresión de *xtcited2* durante la organogénesis de *X. tropicalis*. Ensayos de hibridación *in situ* en embrión completo permiten observar la expresión de *xtcited2* desde inicios de la organogénesis (NF 28) hasta renacuajo en fase de alimentación autónoma (NF 45) de *X. tropicalis*. Los estadios NF mostrados son indicados en la esquina inferior izquierda de cada panel. El número de embriones analizados se indican en la esquina superior derecha de cada panel, y se encuentran tabulados en la tabla a la derecha de la figura. Se utilizó una sonda antisentido para ver la expresión espaciotemporal de

xtcited2 (A-E), y una sonda sentido como control negativo (A'-E'). Se muestran vistas laterales de cada embrión. Las cabezas de flecha negras indican estructuras específicas rotuladas como: **pce** = prosencéfalo; **ris** = región intersomítica; **pco** = primordio cardiaco; **ce** = cerebro; **ab** = arcos branquiales y **co** = corazón. Arriba a la izquierda de cada panel se indica el número de embriones analizados en cada estadio.

1.2. Análisis bioinformático de niveles de expresión de *xtcited2* y *xtric-8a*.

Para respaldar los resultados de las hibridaciones *in situ* para *xtcited2*, se exploraron los niveles de expresión de este gen en estadios similares mediante el visualizador *Xenopus Jamboree* (<http://tinyurl.com/scXen2018>; Briggs y col., 2018). Como se observa en la Figura 10, la mayor cantidad de células positivas a *xtcited2* se encuentran en estadios más tempranos (NF 8, 10 y 11), particularmente en los tejidos de blástula, zona marginal (mesodermo), endodermo y ectodermo neural y no-neural. En estadios de neurula (NF 13) en adelante, se describen niveles más bajos de expresión y de células positivas (Figura 10D, Figura 11). Sin embargo, incluso en estos estadios tardíos, los tejidos con mayor expresión de *xtcited2* mantienen una tendencia similar, siendo principalmente tejidos mesodermales, como el mesodermo intermedio y

presomítico, además de tejidos neurales, como placa y tubo neural, o rombencéfalo y espina dorsal en la néurula tardía (NF 22).

Finalmente, analizamos la expresión de *xtcited2* en los tejidos que darán origen a la CN craneal desde la blástula (NF 8) hasta el estadio NF 22 (Figura 12A). En la blástula, los niveles de expresión son moderados, pero disminuyen muy significativamente con la diferenciación a neuroectodermo en el estadio NF 10. Luego, la expresión aumenta de manera muy significativa en el neuroectodermo de estadio NF 11, y se mantiene constante a medida que el tejido se diferencia hacia el borde de placa neural (NF 12) y a CN (NF 13). En la CN de estadio NF 14, los niveles de *xtcited2* vuelven a disminuir muy significativamente, manteniéndose relativamente constantes hasta el estadio de CN craneal NF 22. Es importante destacar que, dentro de este último periodo, se observa un cambio significativo en la expresión entre las CN de estadios NF 16 y 18. Al comparar este patrón con el de *xtric-8a* en los mismos tejidos, observamos que estos coinciden y sólo hay diferencia significativa de su expresión en la cresta neural de estadio NF 16 y cresta neural craneal de estadio NF 22 (Figura 12B).

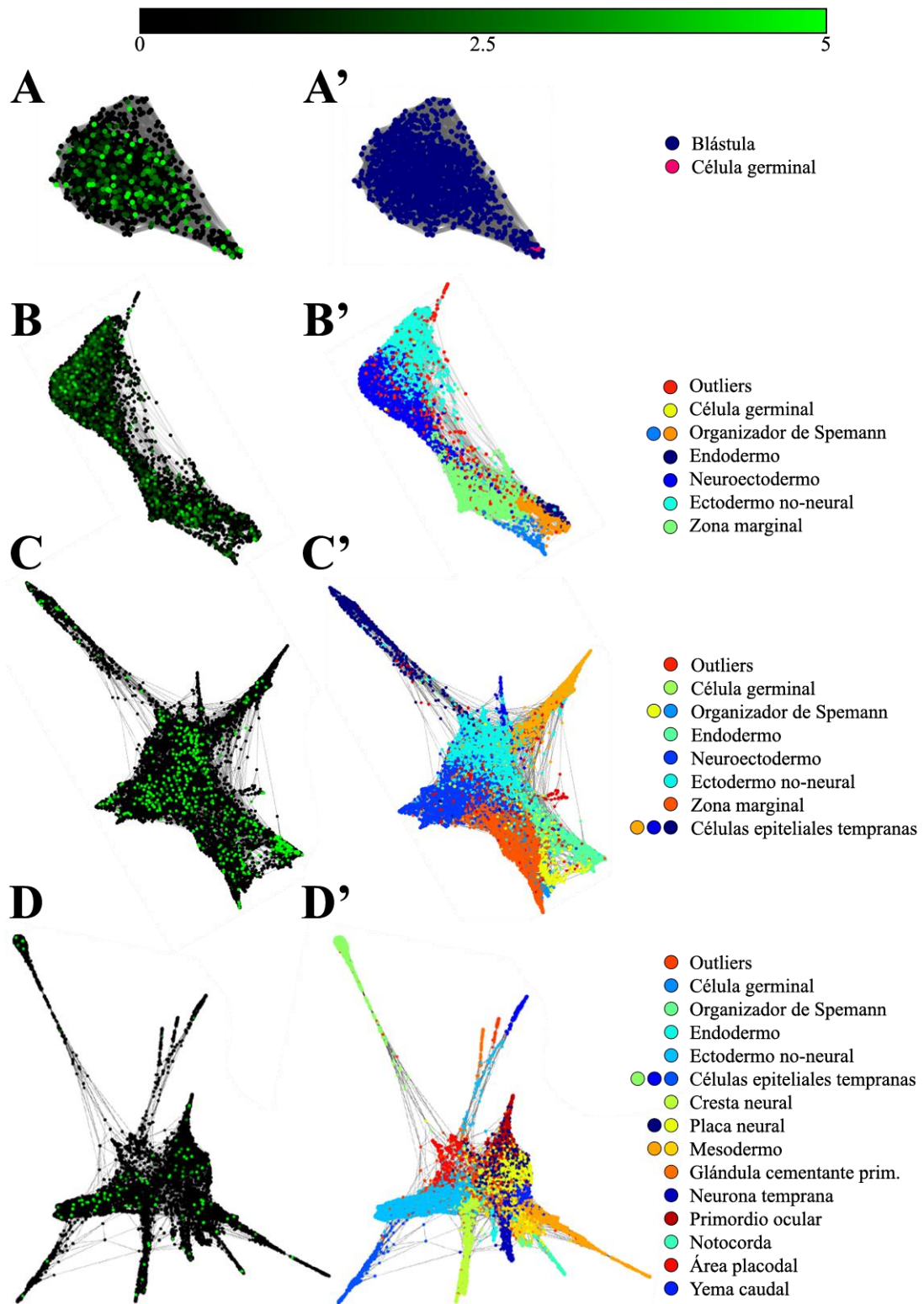


Figura 10: Expresión de *xtcited2* en tejidos embrionarios tempranos obtenidos desde ensayos de scRNA-seq. Mapas SPRING coloreados por expresión en UMIFM de *xtcited2* (A-D) y tipo de tejido (A'-D'). Clusterización de células obtenidas desde embriones completos en estadio NF 8 (A-A'), NF 10 (B-B'), NF 11 (C-C') y NF 13 (D-D'). Niveles de expresión (UMIFM) se colorearon en una escala de color desde 0 a 5, de acuerdo con la escala de colores (arriba). Se observa la mayor expresión de *xtcited2* en tejidos de blástula, zona marginal, endodermo y ectodermo neural y no neural de estadios tempranos (NF 8, 10 y 11). A partir del estadio NF 13, vemos menor expresión de *xtcited2*. Extraído de *Xenopus Jamboree* (<http://tinyurl.com/scXen2018>; Briggs y col., 2018).

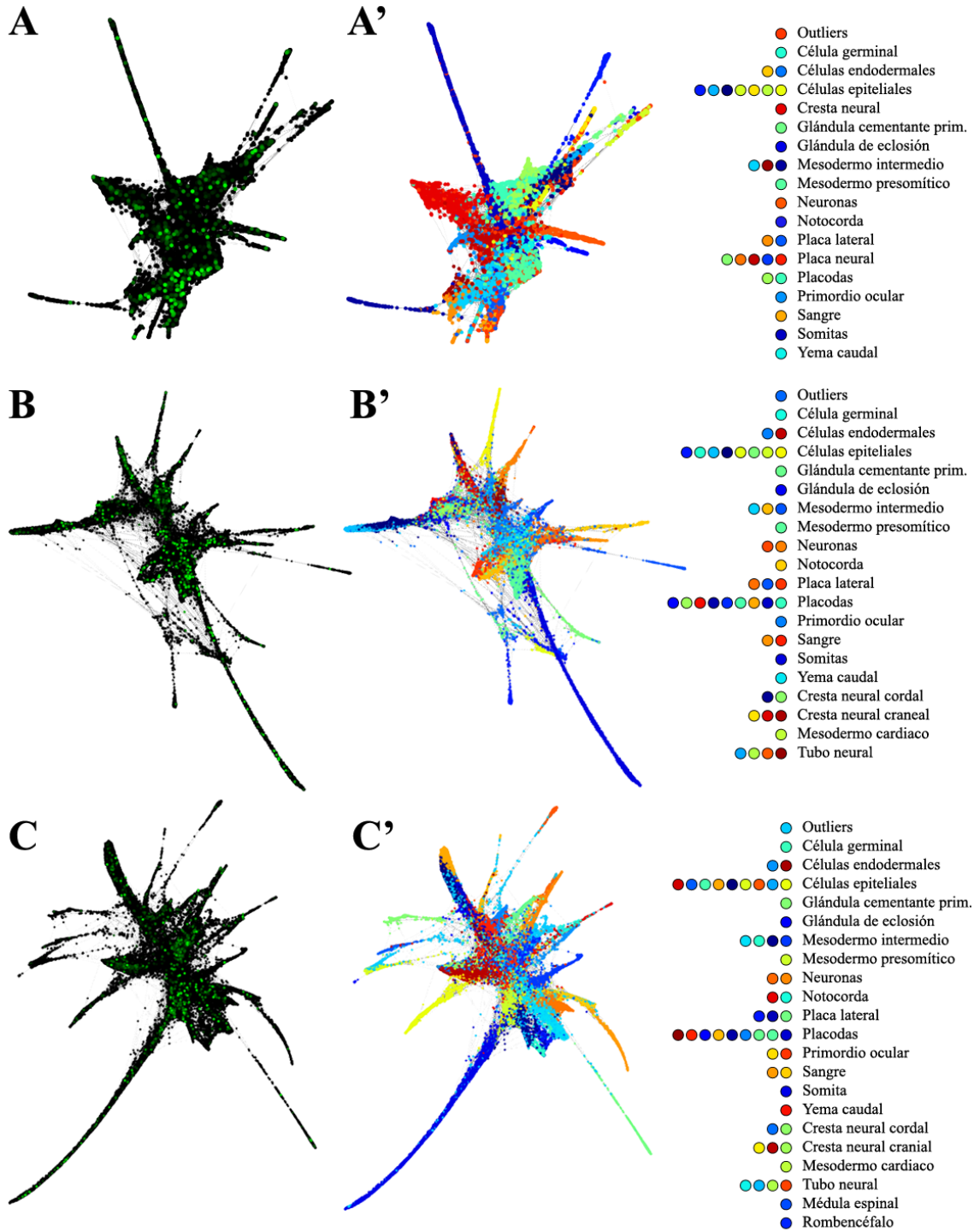


Figura 11: Expresión de *xtcited2* en tejidos embrionarios de néurula obtenidos desde ensayos de scRNA-seq. Mapas SPRING coloreados por expresión en UMIFM de *xtcited2* (**A-C**) y tipo de tejido (**A'-C'**). Clusterización de células obtenidas desde embriones completos en estadio NF 16 (**A-A'**), NF 20 (**B-B'**) y NF 22 (**C-C'**). Niveles de expresión (UMIFM) se colorearon en una escala de color desde 0 a 5, de acuerdo con la escala de colores (arriba). Se observa una menor cantidad de células que expresan *xtcited2* en comparación a estadios más tempranos. Extraído de *Xenopus Jamboree* (<http://tinyurl.com/scXen2018>; Briggs y col., 2018).

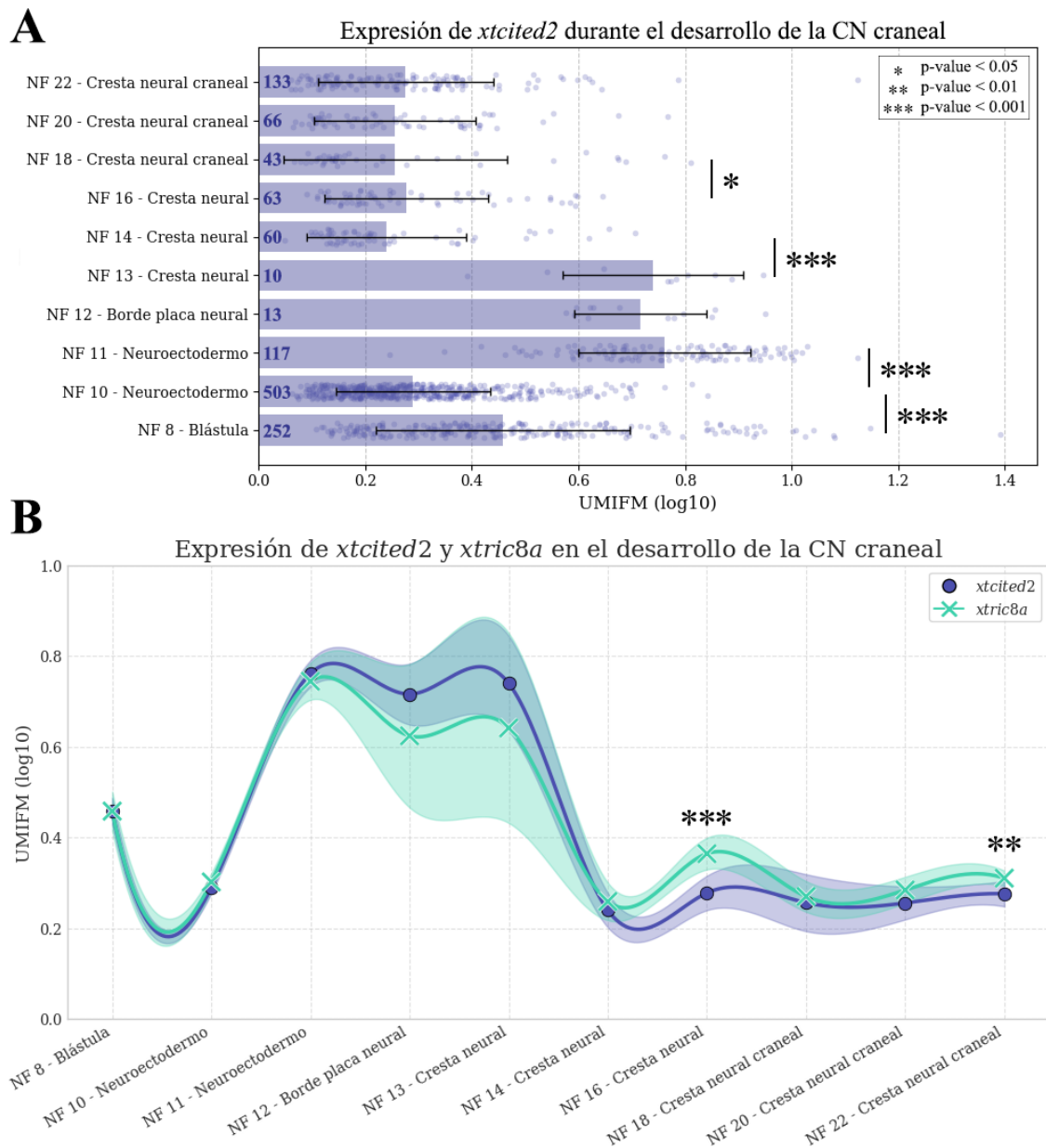


Figura 12: Expresión de *xtcited2* y *xtricc8a* en el desarrollo de la CN craneal. Estadios que derivan a cresta neural craneal de estadio NF 22 desde la blástula en estadio NF 8. **(A)** Niveles de expresión de *xtcited2* a lo largo del

desarrollo de la CN craneal. Se observa que la expresión del gen varía significativamente en los estadios más tempranos, llegando a su máximo entre NF 11 y 13. **(B)** Comparación de niveles de expresión de *xtcited2* y *xtric-8a* a lo largo del desarrollo de la CN craneal. Podemos observar que, en los tejidos analizados, la expresión de ambos es altamente similar, presentando diferencias significativas sólo en estadios NF 16 y 18.

1.3. Inmunodetección de la proteína de fusión V5+XtCited2 en embrión completo.

Con el fin de realizar en el futuro ensayos de co-inmunoprecipitación de *xtcited2* y verificar así la interacción con *xtRic-8A* se generó un constructo plasmidial que contiene el gen de *xtcited2* fusionado al tag V5. Luego de la construcción se verificó la correcta expresión de esta proteína. Para esto, se sintetizó el ARNm de este constructo genético y se procedió a sobreexpresar V5+XtCited2 en embriones de *X. tropicalis*. En la Figura 13A, se presenta un esquema del vector en cuestión, donde se muestra que la secuencia promotora para la ARN polimerasa SP6 en el 5' del tag V5 y la región codificante para *xtcited2*, además de un sitio de corte para NotI en el 3'. Se utilizó esta enzima de restricción para linealizar el constructo genético y se sintetizó el ARNm con la enzima SP6. En la Figura 13B se observa el resultado de la transcripción *in vitro*, donde se obtuvo un rendimiento elevado de ARNm. Posteriormente, este ARNm

fue cuantificado por densitometría, alicuotado y guardado a -80°C hasta la siguiente fertilización, donde se inyectaron 150 pg de ARNm V5+XtCited2 con Rodamina Dextrano (como trazador de linaje) a cada blastómero de embriones en estadio de 2 células (NF 2). Los embriones fueron crecidos a estadio de néurula, seleccionados y lisados para la obtención de proteína total.

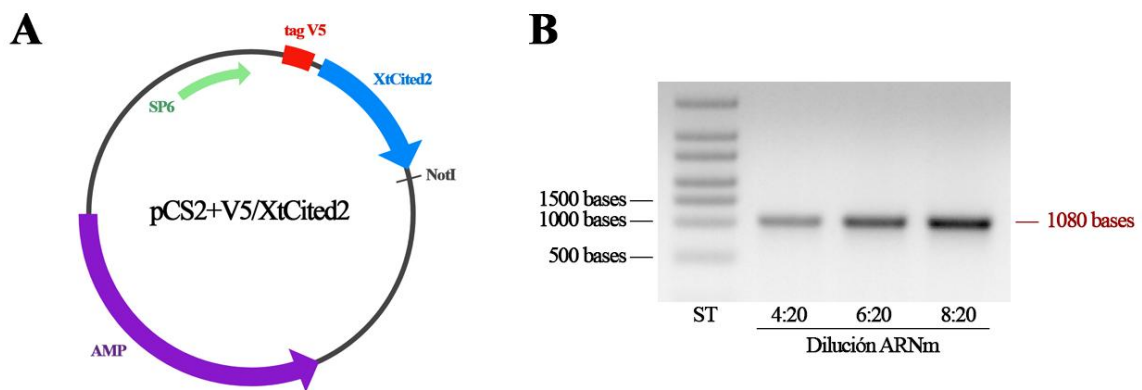


Figura 13: Síntesis de ARNm de V5+XtCited2. **(A)** Esquema del vector pCS2+V5/XtCited2. En rojo se esquematiza el tag V5, en azul el fragmento correspondiente al gen *xtcited2*, en verde claro la región promotora para SP6, en violeta el gen de resistencia a antibiótico (Ampicilina) y la línea negra perpendicular representa el sitio de corte para la enzima de restricción NotI. **(B)** Cuantificación del ARNm mediante gel de agarosa. Se hizo una dilución 1:20 del ARNm sintetizado y se cargaron 4, 6 y 8 μL de ésta. A la izquierda se muestra el estándar de tamaño molecular RiboRuler HighRange y a la derecha el tamaño esperado del ARNm, ambos en número de bases.

Se utilizó la técnica de Western blot para detectar la expresión de la proteína de fusión V5+XtCited2. En la Figura 14, la banda observada justo debajo de los 35 kDa coincide con el tamaño teórico de V5+XtCited2 (28 kDa), respaldando su correcta expresión.

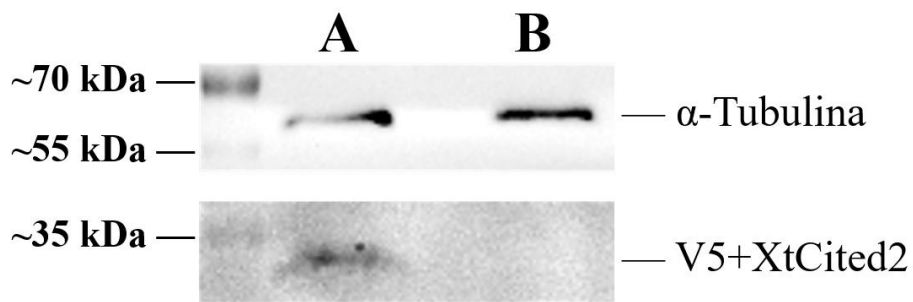


Figura 14: Inmunodetección de V5+XtCited2. Western blot realizado a partir de embriones en estadio NF 9 microinyectados con 300 pg totales de ARNm de V5+XtCited2 (**A**) y embriones NF 9 *wild-type* (**B**). Se cargaron 100 µg de lisado de proteína total. Se utilizó el anticuerpo anti-Tubulina en dilución 1:1000 como control de carga, y el anticuerpo anti-V5 en dilución 1:100 para detectar la proteína de fusión. A la derecha se indica el anticuerpo utilizado y a la izquierda se muestra el estándar de peso molecular PageRuler™ Plus Prestained Protein Ladder (10 to 250 kDa) de Thermo Scientific. α-Tubulina tiene un peso teórico de 55 kDa y V5+XtCited2 tiene un peso teórico de 28 kDa.

1.4. Inmunodetección de *xtcited2* en células de la cresta neural.

Como última actividad en este objetivo, se determinó la localización subcelular de *xtcited2* en las células de la CN mediante un ensayo de inmunofluorescencia en explantes de CN de *X. tropicalis*. Se obtuvieron explantes de CN desde embriones *wild-type* (Figura 15A-C) y de embriones con sobreexpresión de la proteína V5+XtCited2 (Figura 15D-G). Para ambos casos, podemos observar una expresión localizada en el núcleo, lo cual coincide con su función como transactivador transcripcional. Sin embargo, también se puede observar fluorescencia en el citoplasma de las células sugiriendo la presencia de este transactivador también en el citoplasma.

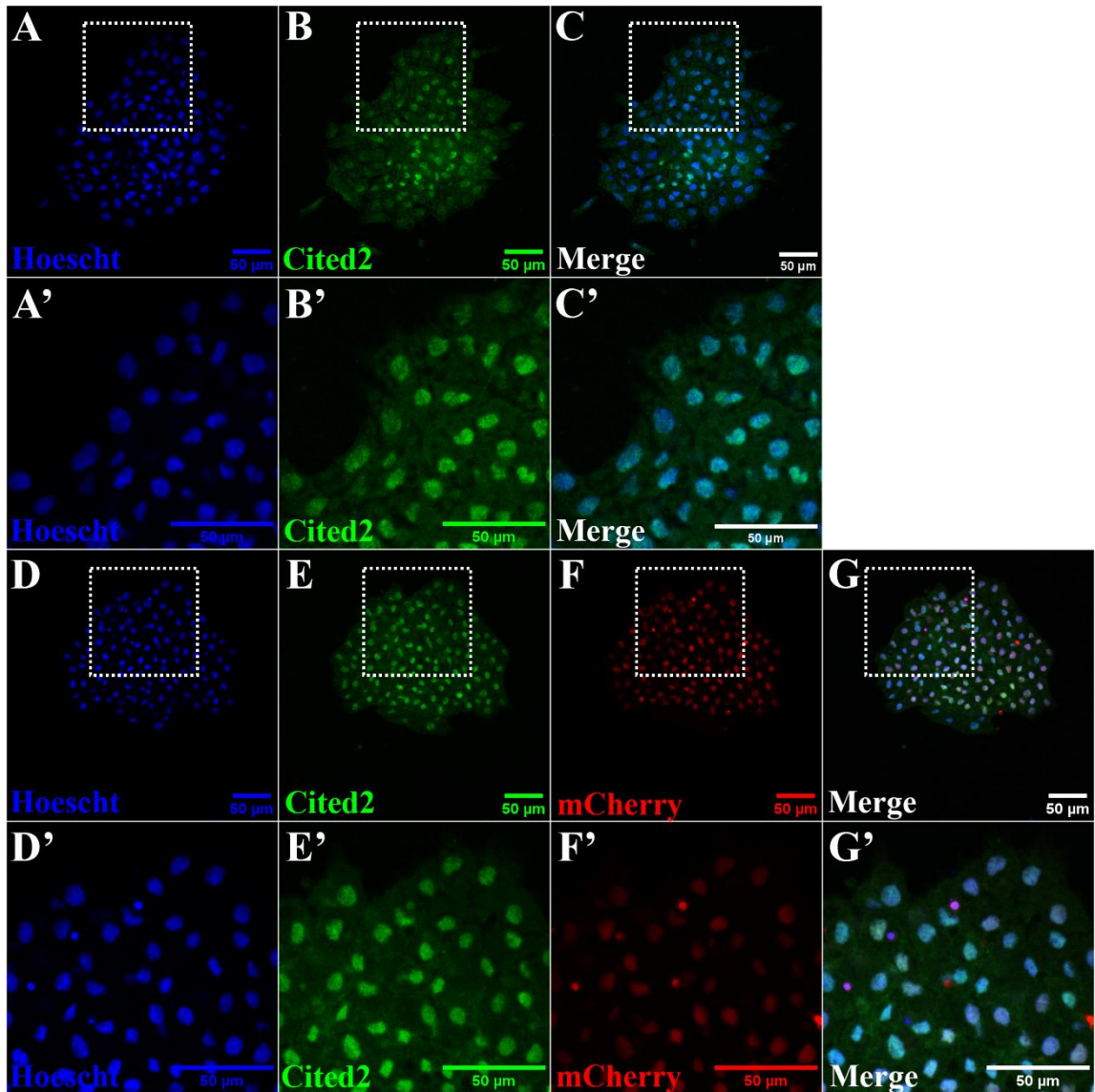


Figura 15: Localización subcelular de *xtcited2*. Se realizaron ensayos de inmunofluorescencia en explantes de CN obtenidos de embriones wild-type (paneles A-C) y microinyectados con el mensajero de *V5-xtcited2* (paneles D-G). Los paneles A'-G' muestran un acercamiento de la región señalada con línea punteada en los explantes correspondientes. Para marcar los núcleos, se utilizó

el colorante fluorescente Hoescht (azul). Se inyectaron 300 pg de ARNm V5+XtCited2 junto con 200 pg de H2B-mCherry (rojo) como trazador de linaje. Se observa expresión de *xtcited2* (verde) en los núcleos de las CN. Barra de escala: 50 µm. Esta imagen es representativa de 3 experimentos independientes, donde en cada uno se analizaron 2 explantes.

2. Objetivo 2: Identificar una subpoblación de cresta neural cardiaca desde datos de RNAseq de célula única de cresta neural.

Con el fin de identificar una subpoblación de CN cardiaca, se descargó un set de datos de cresta neural total (NF 18) de la plataforma de Briggs y col., que corresponde a la combinación de los sets de CN craneal y cordal (troncal, vagal y sacral). Los datos fueron cargados y se procedió a utilizar el lenguaje de programación Python y una serie de librerías disponibles (detalladas en el punto 2.28 de Métodos) para realizar los análisis.

2.1. Búsqueda de clústeres de células de la cresta neural.

Debido a que los datos descargados ya se encontraban normalizados, el primer paso consistió en realizar una reducción de dimensionalidad por PCA (Maćkiewicz y Ratajczak, 1993) la cual permite disminuir la complejidad de los datos con el fin de identificar patrones entre muestras que luego serán agrupados según sus perfiles de expresión. Se seleccionaron 200 componentes principales, lo cual incluye al 80% de la varianza acumulada, y se procedió a realizar una

segunda reducción de dimensionalidad utilizando UMAP (McInnes y col., 2018). En ellos, cada punto representa una célula, y la distancia relativa entre ellas simboliza qué tan similares o diferentes son de las otras. Usando estos gráficos de base, se realizó una búsqueda de clústeres utilizando algoritmo de Leiden (Traag y col., 2019). Debido a que los resultados tanto de UMAP y Leiden dependen de los parámetros “*n_neighbors*” (que indica cuántas células se consideran para definir relaciones locales) y resolución (que controla qué tan detallada será la separación entre clústeres), respectivamente, se probaron varias combinaciones de éstos para poder seleccionar la que generara los clústeres mejor definidos. Para esto, primero seleccionamos aquellos que generaban de 3 a 5 clústeres, bajo la racional de que debería haber un mínimo de dos (craneal y cordal) y un máximo de cinco subpoblaciones de CN (craneal, cardíaca, troncal, sacral y vagal). Los mapas UMAP seleccionados son aquellos con un valor de *n_neighbors* de 30 y resolución de 0.5, 0.75 y 1, los cuales generaron 3, 4 y 5 clústeres, respectivamente (Figura 16A-C). Luego, fue necesario determinar cuál es el número de clústeres con el que debemos trabajar, es decir, en cuántos clústeres podemos separar claramente nuestro set de datos.

Para esto, se decidió aplicar el algoritmo de *Random Forest* (Breiman, 2001) utilizando la información etiqueta de los clústeres de Leiden y el perfil de expresión de cada célula. Con el fin de reducir la cantidad de variables (genes) con las que se entrenará el modelo, realizamos un análisis de expresión

diferencial entre cada clúster y el resto de las células para filtrar aquellos genes cuya expresión diferencial era significativa ($p\text{-adj} < 0.05$) y estaban expresados al alza ($\log_2\text{FC} > 0$) en alguna de las comparaciones. Esto bajo la racional de que esos genes serían posibles marcadores genéticos para cada clúster, y, por ende, serían los más importantes para diferenciar entre ellos.

Con estos genes, se procedió a realizar el entrenamiento del modelo de *Random Forest*, comenzando por la búsqueda de hiperparámetros óptimos para cada uno de los valores de resolución. En la Tabla 2, se pueden observar los valores de hiperparámetros con los cuales el modelo mostró los valores de precisión más altos. Las matrices de confusión (Figura 16A'-C') nos permiten visualizar y comparar el desempeño de los modelos. En ellas, cada fila representa las etiquetas (clústeres) asignadas por el algoritmo de Leiden, mientras que cada columna corresponde a las predicciones realizadas por el modelo. Los valores en la diagonal indican el número de células correctamente en su clúster correspondiente, mientras que los valores fuera de la diagonal representan errores de clasificación. En la matriz correspondiente a la resolución 0.5 (Figura 16A'), podemos observar un muy buen desempeño, habiendo categorizado las células con una cantidad mínima de errores. Por su parte, la matriz de resolución 0.75 (Figura 16B'), vemos un aumento considerable en los errores de clasificación, particularmente en los clústeres 1 y 2, como se observa por los valores fuera de la diagonal. Finalmente, en la matriz de resolución 1 (Figura

16C'), observamos aún más errores en la clasificación de los clústeres 1 y 2, incluyendo, además, al clúster 4.

Al comparar las tres matrices, determinamos que el agrupamiento más robusto es el generado con el valor de resolución es de 0.5, ya que el modelo es capaz de clasificar las células con un 98% de precisión (Figura 16A, A'), en contraste a la precisión del 89.86% y 88.20% de las resoluciones 0.75 (Figura 16B, B') y 1 (Figura 16C, C'), respectivamente.

	Resolución		
	0.5	0.75	1
Tamaño set de prueba	0.2	0.2	0.5
'max_depth'	30	40	40
'max_features'	None	None	None
'min_samples_leaf'	1	1	2
'min_samples_split'	2	4	4
'n_estimators'	190	171	185

Tabla 2: Hiperparámetros óptimos para el entrenamiento del modelo de *Random Forest*. Para cada valor de resolución, se seleccionaron los parámetros que resultaban en modelos de mayor precisión.

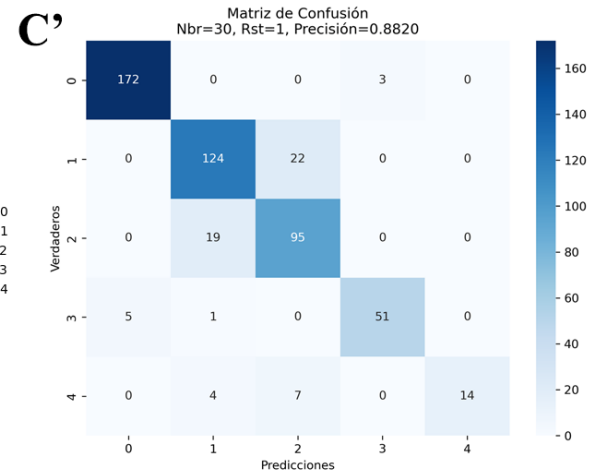
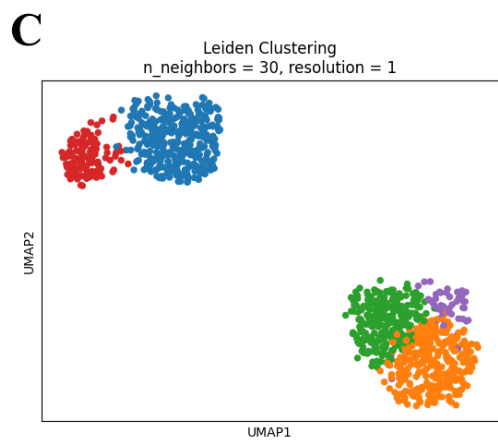
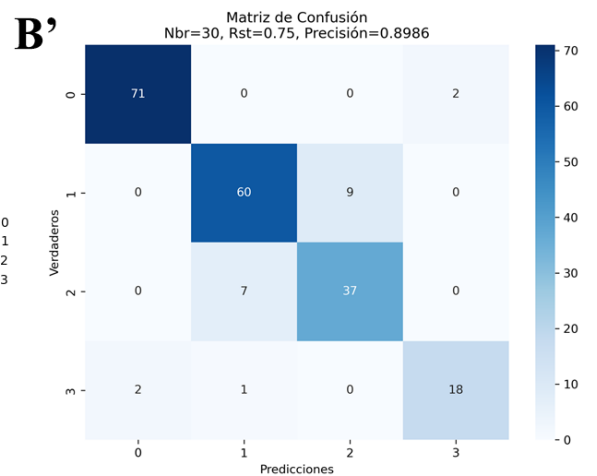
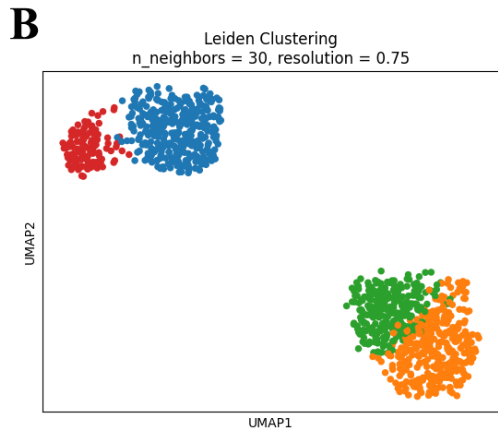
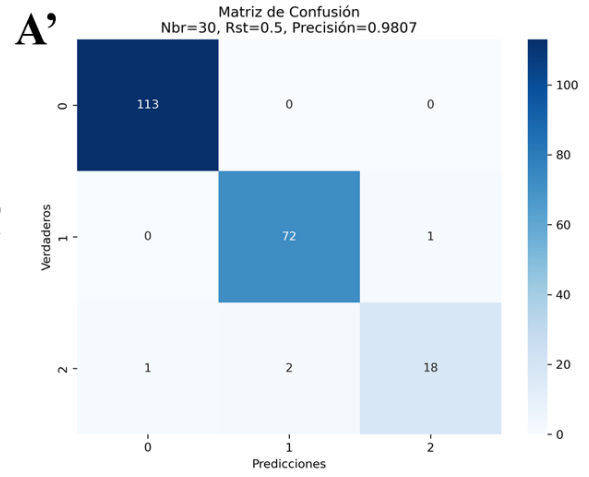
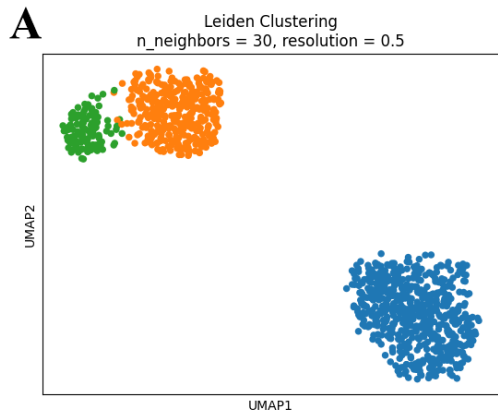


Figura 16: Comparación de agrupamientos generados por *clustering* Leiden a distintas resoluciones en células de cresta neural de estadio NF18.

(A-C) Gráficos UMAP generados utilizando un valor de $n_neighbors$ (Nbr) de 30 y resoluciones (Rst) de 0.5, 0.75 y 1 para el *clustering* por Leiden. Cada punto representa una célula y los colores indican la asignación a un clúster específico según la resolución utilizada. **(A'-C')** Matrices de confusión correspondientes al modelo de *Random Forest* optimizado para cada valor de resolución. Las filas representan las etiquetas originales asignadas por Leiden y las columnas corresponden a las predicciones del modelo; valores dentro de la diagonal representan clasificaciones correctas, valores fuera de la diagonal representan errores en la predicción. La clusterización con resolución 0.5 (A, A') presenta una métrica de precisión del modelo de 98.07%, en comparación con las resoluciones 0.75 (B, B') y 1 (C, C'), que presentan una precisión de 89.86% y 88.20%, respectivamente. Nbr = $n_neighbors$; Rst = resolución (*resolution*).

2.2. Aplicación del algoritmo *Random Forest* para identificar los genes más importantes para la clasificación de los diferentes clústeres.

Una vez definido el número óptimo de clústeres generados por Leiden mediante el entrenamiento y optimización del modelo de *Random Forest*, se utilizó información proveniente del proceso de aprendizaje del modelo para identificar aquellos genes que más contribuyen a la correcta clasificación de las

células en función a sus perfiles de expresión. Esta métrica, llamada “importancia” indica cuáles fueron los genes más informativos (o “importantes”) al momento de distinguir entre clústeres.

En la Figura 17A, podemos ver un histograma que muestra la distribución de los genes con importancia mayor a 0, lo que corresponde a un total de 328 genes. Debido a que la gran mayoría posee un valor de importancia muy bajo, se aplicó un recorte en el eje X para facilitar la visualización. Sin embargo, es importante recalcar que un único gen (*gsk3b*) presenta un valor de importancia considerablemente superior (de 0.975798). Esto nos indica que los agrupamientos generados estarían mayoritariamente definidos por el perfil de expresión de *gsk3b*. Los perfiles de expresión de estos genes se representaron mediante el gráfico de calor de la Figura 17B. En este mapa, cada columna representa un clúster y cada fila un gen, mientras que el color indica el nivel de expresión relativa de ese gen, normalizada mediante Z-score. Los genes fueron ordenados según su expresión promedio en el clúster 0, lo cual facilita la observación de los distintos perfiles de expresión de cada clúster. Por ejemplo, en la primera columna, observamos que varios de los genes con una expresión relativa alta (azul) en el clúster 0 poseen una expresión relativa más baja en los clústeres 1 y 2 (rojo), y viceversa. Esto nos indica que existen genes con perfiles de expresión característicos de cada clúster que podrían servir como marcadores.

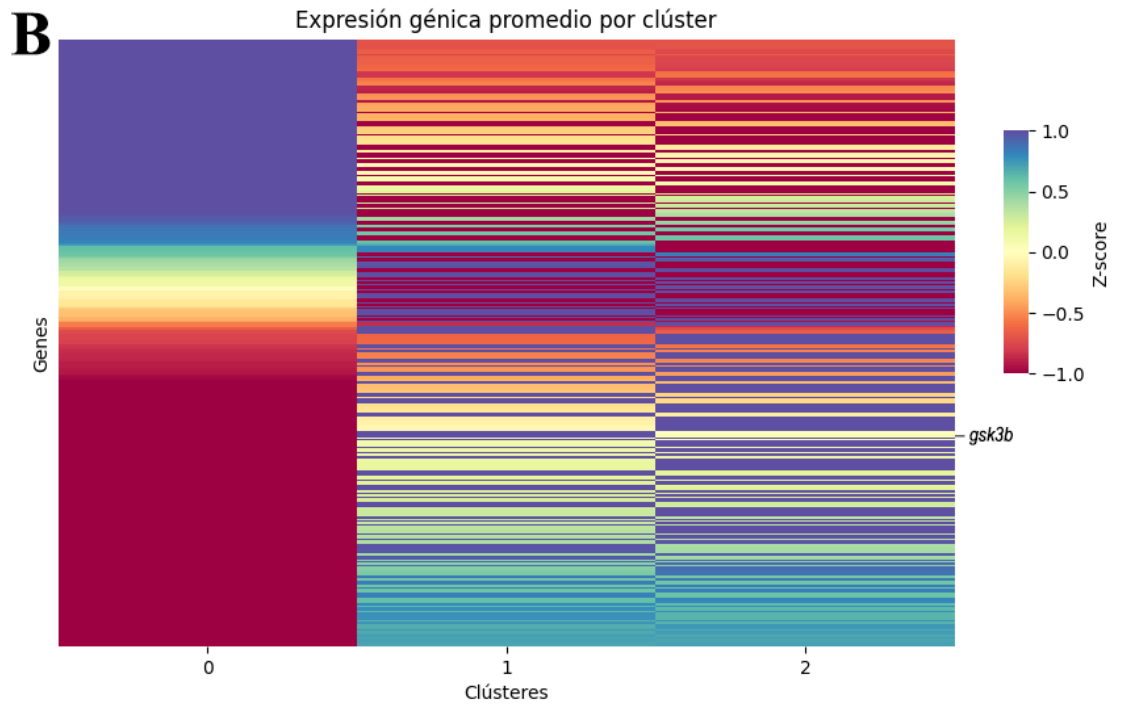
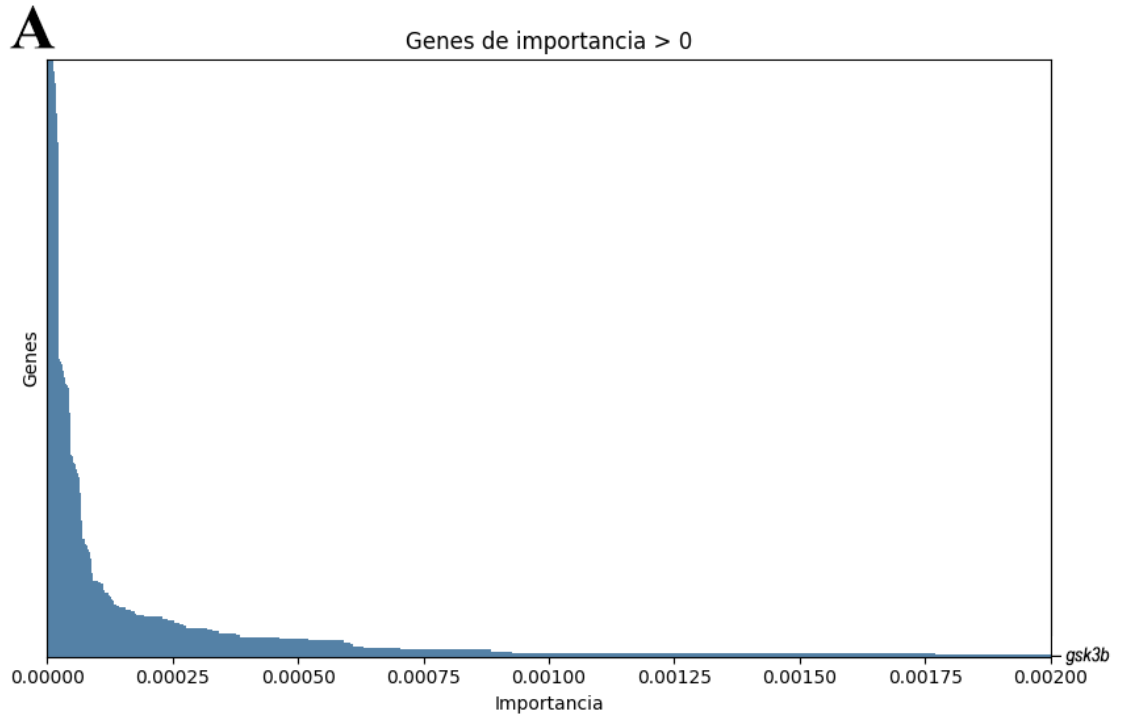


Figura 17: Distribución de genes con Importancia mayor a 0 según el modelo de Random Forest. **(A)** Distribución de genes con importancia mayor a 0 según el modelo de *Random Forest*. Histograma que muestra las importancias de los genes (eje Y) con valor de importancia (eje X) mayor a 0, ordenados de menor a mayor. Para facilitar la visualización de los valores bajos, se aplicó un acercamiento al eje X, limitándolo a un valor máximo de 0.002 para facilitar la visualización. El gen *gsk3b* fue el único gen que presentó una importancia mucho mayor (0.975798), cosa que no se visualiza completamente debido al ajuste de escala. **(B)** Expresión génica promedio por clúster. Mapa de calor que muestra la expresión promedio normalizada mediante Z-score de cada gen en los clústeres identificados. Cada columna representa un clúster y cada fila un gen. Los genes fueron ordenados según su expresión promedio en el clúster 0, y se utilizó una escala de color para visualizar los niveles relativos de expresión: azul para valores altos y rojo para valores bajos, evidenciando perfiles de expresión distintos entre clústeres.

2.3. Análisis de expresión diferencial entre clústeres y caracterización de posibles genes marcadores.

Como último paso, para diferenciar entre los genes que marcaban cada clúster o combinación de clústeres, se realizó una segunda expresión diferencial, esta vez por pares de clústeres, con el fin de determinar los genes marcadores

para cada caso. Esto se hizo con dos consideraciones en mente: que hubiera expresión diferencial ($p\text{-adj} < 0.05$) de dicho gen entre pares de clústeres, y que ésta estuviera al alza en el clúster correspondiente (valor de cambio \log_2FC coherente, según la comparación). Los criterios se encuentran esquematizados y explicados en la Tabla 3.

	1 vs 2		1 vs 0		2 vs 0	
	\log_2FC	$p\text{-adj} < 0.05$	\log_2FC	$p\text{-adj} < 0.05$	\log_2FC	$p\text{-adj} < 0.05$
Clúster 0			Negativo	Sí	Negativo	Sí
Clúster 1	Positivo	Sí	Positivo	Sí		
Clúster 2	Negativo	Sí			Positivo	Sí
Clúster 1+2		No	Positivo	Sí	Positivo	Sí
Clúster 1+0	Positivo	Sí		No	Negativo	Sí
Clúster 2+0	Negativo	Sí	Positivo	Sí		No

Tabla 3: Criterios de expresión diferencial utilizados para clasificar genes según el clúster que marcan. Resumen de los criterios utilizados para determinar si un gen actúa como marcador específico a un clúster individual o a una combinación de dos clústeres. Se consideraron los valores de cambio de expresión (\log_2FC) y el valor p ajustado ($p\text{-adj}$) en las comparaciones entre pares de clústeres (1 vs 2, 1 vs 0 y 2 vs 0). Para clasificar un gen como marcador, debía presentar una expresión diferencial significativa ($p\text{-adj} < 0.05$) y un valor de \log_2FC (positivo o negativo) que fuera coherente con que una mayor expresión relativa en el clúster correspondiente, dependiendo de la comparación. En el caso de genes marcadores para una combinación de clústeres, se exigió que la

expresión no fuera significativamente diferente entre ellos ($p\text{-adj} \geq 0.05$, indicando que la expresión es similar), y que el $\log_2\text{FC}$ indicara una menor expresión en el clúster restante. Las celdas vacías indican que la comparación era irrelevante.

Con este análisis, obtenemos las listas de posibles marcadores. Para el clúster 0 se encontraron 103 genes marcadores, 11 para el clúster 1, y 24 para el Clúster 2 (Figura 18). De forma interesante, los clústeres 1 y 2 comparten 146 marcadores. En contraste, sólo hay 1 marcador común entre los clústeres 1 y 0, y otro entre los clústeres 2 y 0. Este resultado es de esperar, siendo que en el mapa UMAP los clústeres 1 y 2 se localizan muy cerca entre sí, y lejos del clúster 0 (Figura 16A). Los 42 genes restantes fueron clasificados como no-marcadores. Estas listas resultantes se encuentran tabuladas en el Anexo 6.

Al analizar estas listas de genes, podemos ver que en su gran mayoría encontramos genes asociados a la traducción, unión a ribosoma y unión a ARNm, como lo serían los factores de transcripción generales *eif1b/4b*, o genes codificantes de proteínas ribosomales como *rpl6/8/32/35/36*.

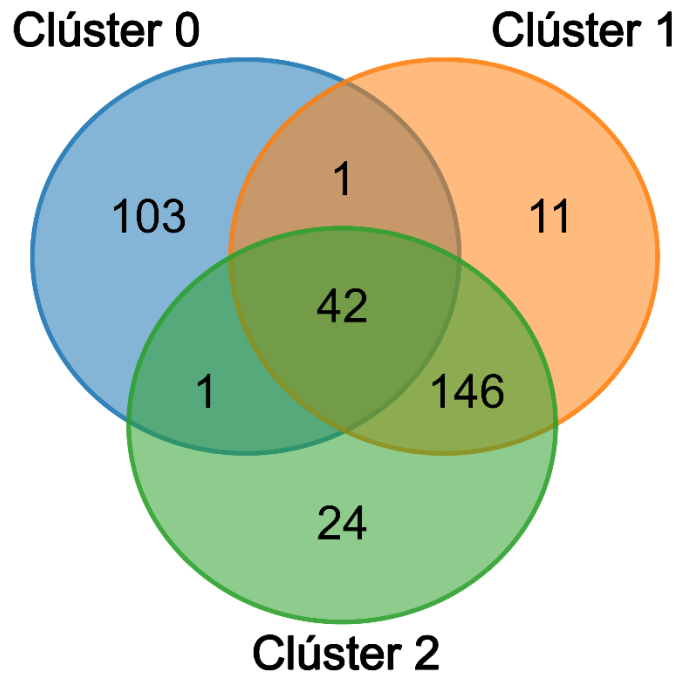


Figura 18: Genes marcadores por clúster. Diagrama de Venn que muestra la distribución de genes marcadores identificados para cada uno de los tres clústeres. Se identificaron 103 marcadores exclusivos para el clúster 0 (azul), 11 para el clúster 1 (naranja) y 24 para el clúster 2. Solo se encontró un gen marcador compartido entre los clústeres 0 y 1, y otro entre los clústeres 0 y 2; mientras que hay 146 marcadores compartidos entre los clústeres 1 y 2. Finalmente, 42 genes se encontraron presentes en los tres clústeres, por lo que fueron clasificados como no-marcadores.

Se comenzó, entonces, investigando el gen con mayor importancia para la diferenciación entre clústeres: *gsk3b*. GSK3 β es una serina/treonina-kinasa con un importante rol en el desarrollo embrionario, con más de 100 sustratos conocidos y funciones clave en la regulación de proliferación, diferenciación, migración y apoptosis (Beurel y col., 2015). A nivel de señalización, uno de sus sustratos más importantes es β -catenina, que es fosforilada por GSK3 β para ser enviada a degradación, generando así una inhibición de la vía Wnt (Beurel y col., 2015; McCubrey y col., 2014). En la CN, GSK3 β cumple roles clave en TEM, especificación y migración (Abbruzzese y col., 2014; Gonzalez Malagon y col., 2018; Lander y col., 2013). Podemos observar cómo la expresión de GSK3 β es altamente importante para la diferenciación de los clústeres: en el clúster 0, tiene una expresión relativa extremadamente baja. Por el contrario, en el clúster 1 tiene su expresión más alta, mientras que en el clúster 2 vemos que su expresión relativa es intermedia respecto a los otros dos clústeres (Figura 19B-C). Además, es un gen que está diferencialmente expresado entre los tres clústeres (Figura 19D), y es un gen marcador del clúster 1.

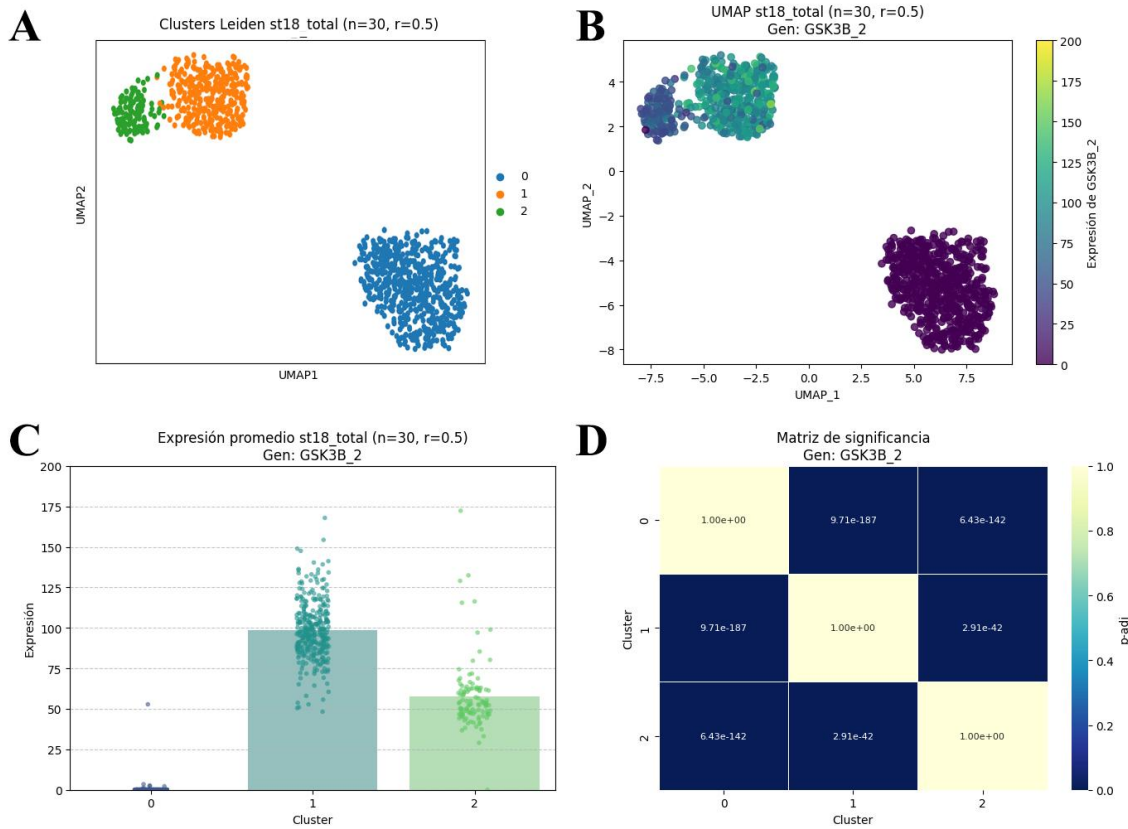


Figura 19: Análisis de expresión del gen *gsk3b* en células de cresta neural total de estadio NF18. (A) Mapa UMAP con la distribución de los clústeres identificados mediante Leiden. Valores $n = 30$ ($n_neighbors$), $r = 0.5$ (resolución). **(B)** Mapa UMAP coloreado según el nivel de expresión de *gsk3b*, donde amarillo indica alta expresión y morado indica baja expresión. **(C)** Gráfico de barras con la expresión promedio de *gsk3b* por clúster. Cada punto representa el valor de una célula individual. **(D)** Matriz de significancia estadística para las comparaciones de clústeres, calculada mediante test de Mann-Whitney U. Celdas de color azul oscuro representan valores de $p\text{-adj} < 0.05$.

Por último, para determinar cuál de los clústeres tenía una mayor identidad cardíaca, procedimos a buscar entre la lista de genes a aquellos que estuvieran relacionados al desarrollo o identidad cardiovascular.

Como marcadores interesantes del Clúster 0, encontramos a *rbm6* y *rbmx*. Si bien no está descrita una relación directa con el desarrollo cardiovascular, otras proteínas de la familia *Ribosome Binding Motif* (RBM), como *rbm24*, han sido asociadas a angiogénesis (Maragh y col., 2011).

En el Clúster 1, encontramos el gen *Imln* (*invadolysin*), cuya pérdida de función está asociada a defectos en la angiogénesis de pez cebra (Vass y Heck, 2013). Como se mencionó anteriormente, en este clúster también encontramos a *gsk3b*, la cual es necesaria para la formación del tejido cardíaco en el mesodermo mediante inhibición de la vía Wnt (Schneider y Mercola, 2001).

En el Clúster 2, encontramos a *mbd3* y *tpm4*. El primero, *mbd3*, un componente del complejo de remodelación de nucleosomas y deacetilasas NuRD (Le Guezennec y col., 2006) necesario para reprimir genes no-cardíacos durante el desarrollo del corazón de ratón (Robbe y col., 2022). Por otra parte, en pez cebra, una isoforma de tropomiosina 4 (*tpm4*) ha mostrado ser importante para el funcionamiento apropiado del corazón embrionario (Zhao y col., 2008).

En síntesis, podemos concluir que se observan 3 clústeres con sus respectivos potenciales marcadores, entre los cuales se encontraron 2 genes relacionados a desarrollo cardiovascular en cada clúster.

DISCUSIÓN

El desarrollo embrionario comienza con la fusión de los gametos, lo cual desencadena una serie de procesos moleculares, celulares y genéticos para dar lugar a un nuevo ser vivo. Al ser un proceso altamente complejo, un defecto podría llegar a tener consecuencias letales para el embrión, o bien, causar alguna patología. Un ejemplo de patologías asociadas al desarrollo son las neurocristopatías, las cuales se caracterizan por ser derivadas de la incorrecta formación de las CN. En este trabajo de tesis, nos enfocamos en los genes *cited2* y *ric-8a* como potenciales causantes de neurocristopatías, al estar ambos asociados con el desarrollo de estructuras craneofaciales y cardíacas derivadas de las CN.

Uno de los modelos animales más utilizados para el estudio del desarrollo embrionario es la rana africana con garras (*Xenopus*), debido a su embriogénesis externa, de corta duración, y la transparencia de los embriones (Gilbert, 2005). Por ello, el enfoque principal de esta tesis es la caracterización de la expresión espaciotemporal de *cited2* durante la embriogénesis de *X. tropicalis*, con el fin de utilizar este modelo para estudiar a mayor profundidad el rol de este gen.

La técnica de hibridación *in situ* en embrión completo nos permitió observar el patrón espaciotemporal de *xtcited2* durante el desarrollo embrionario en *Xenopus*. Aquí, vemos que este gen comienza a expresarse de forma muy temprana en el polo animal del embrión. Después de la gastrulación, *xtcited2*

comienza a especificarse en tejidos neurales, como lo serían la placa, tubo y cresta neural. Más adelante, esta expresión se observa en la zona de la cabeza del embrión, incluyendo el cerebro y corazón primitivos. De forma interesante, también vemos expresión de *xtcited2* en el mesodermo presomítico, el espacio intersomítico, y la punta de la cola de renacuajos de nado libre. Además, al observar la intensidad de las marcas en comparación a los embriones controles, podríamos decir que los niveles de expresión de *xtcited2* también están variando a lo largo del desarrollo, siendo más alto en los estadios más tempranos y disminuyendo a medida que los embriones se desarrollan. Esta tendencia también se ha observado con *bulk* RNA-seq tanto en *X. tropicalis* como en *X. laevis* (Owens y col., 2016; Session y col., 2016).

Los análisis de datos de RNAseq de célula única a nivel de embrión completo muestran que la mayor cantidad de células que expresan *xtcited2* se encuentran en las etapas de blástula (NF 8) y gástrula (NF 10 y 11), y disminuyen desde la etapa de neurula (NF 13) en adelante, lo cual es consistente con observaciones anteriores. De forma interesante, al analizar específicamente los tejidos que derivan a CN craneal, observamos una disminución muy significativa en los niveles de expresión de *xtcited2* en el neuroectodermo de estadio NF 10 respecto a la blástula (NF 8). La expresión vuelve a aumentar muy significativamente en el neuroectodermo de estadio NF 11, alcanzando los valores más altos, los cuales se mantienen relativamente constantes hasta la etapa de cresta neural

(NF 13). Este comportamiento de disminución y posterior aumento de expresión ha sido descrito también para otros genes clave que participan en vías de señalización durante la gastrulación, etapa crucial para la inducción de las capas germinales y sus derivados. Estos genes regulan tanto la migración celular como la interacción entre las distintas capas germinales, procesos esenciales para iniciar la formación de tejidos especializados (Steventon et al., 2009; Steventon & Mayor, 2012).

Si contrastamos los resultados obtenidos con la información bibliográfica de la expresión de *cited2*, también podemos encontrar patrones similares. En modelo murino, se ha descrito que *cited2* se expresa en células mesenquimáticas, placenta y corazón adulto de ratón, además de múltiples tejidos embrionarios. Entre ellos destacamos el endodermo, mesodermo y tejidos derivados y asociados al desarrollo cardíaco y neural; incluyendo el tubo cardíaco, la CN y las tres vesículas encefálicas primarias. También se observó expresión en el brote caudal (que generará la cola) y en los somitas (Bhattacharjee y col., 2009; Dunwoodie y col., 1998). En modelo de pollo, se observó su expresión en estructuras principalmente mesodermales, incluyendo mesodermo presomítico y cardiogénico, además del neuroectodermo y pliegue cefálico (Schlange y col., 2000). En humanos, se ha visto que *cited2* también se expresa en varios tejidos, incluyendo corazón y cerebro (Bhattacharya y col.,

1999). Con esto, podemos concluir que los patrones entre las tres especies son similares entre sí.

A nivel subcelular, podemos observar que la proteína codificada por *xtcited2* está presente en las CN de *X. tropicalis*, particularmente en la región del núcleo, lo cual es esperable dada su función como transactivador transcripcional. Resulta interesante ver que, además, se ve una menor expresión en el citoplasma, cosa que se ha observado en varias líneas celulares humanas, como THP-1, U2OS, HEL y fibroblastos BJ, según imágenes disponibles en el Human Protein Atlas (Thul y col., 2017). Sin embargo, es necesario realizar estudios de pérdida de función de *xtcited2* para corroborar que la marca es específica.

Cited2, siendo un cofactor de transcripción, tiene múltiples funciones asociadas dependiendo del contexto celular. Su pérdida de función está asociada a defectos del cierre del tubo neural y de la correcta formación del corazón, incluyendo alteraciones en las CN (Bamforth y col., 2001, 2004; Y. Chen y col., 2008; MacDonald y col., 2008; Pacheco-Leyva y col., 2016; Yin y col., 2002). En humanos, se ha visto que las mutaciones en el gen de *cited2* son prevalentes en casos de cardiopatías congénitas. De forma interesante, Yoon y col. (2011) observó que *cited2* induce la expresión de genes asociados a la neurogénesis en explantes de caperuza animal de *X. laevis*. Estos estudios respaldan la hipótesis de que *cited2* estaría involucrado en el desarrollo craneal, neural y cardíaco en *Xenopus*.

Ahora, ¿cuál sería la relación de Cited2 y Ric-8A? Si bien la posible interacción entre estas dos proteínas fue lo que le inició esta investigación, todavía no podemos responder esta pregunta con certeza. Sin embargo, sí podemos trazar líneas entre ambos. En primer lugar, comparando los patrones de expresión espaciotemporales presentados en esta tesis con el trabajo de Maldonado-Agurto y col. (2011), podemos observar la co-distribución de la expresión de *xtric-8a* y *xtcited2* en las zonas craneofaciales del embrión, incluyendo las rutas migratorias de la CN craneal. En nuestro laboratorio se ha observado que XtRic-8A es necesario para la correcta formación del mesodermo dorsal, ectodermo neural y no neural, y por ende en la inducción de las crestas neurales (Rodríguez, 2016, Tesis de Magíster; Ruiz, 2024, Tesis de Magíster) Finalmente hemos descrito su función en la correcta migración de la CN en *Xenopus* (Fuentealba y col., 2013). También se ha asociado a Ric-8A como una proteína chaperona para proteínas G α , además de ser capaz de evitar su ubiquitinación y posterior degradación por proteasoma (Jenie y col., 2013; Klattenhoff y col., 2003). Esto, sumado a que se ha descrito que Cited2 es degradado por el sistema ubiquitina-proteasoma (Shin y col., 2008), podría dar un posible sentido a la interacción de estas dos proteínas. De forma interesante, al contrastar la expresión de *xtcited2* y *xtric8-a* en tejidos que derivan a CN craneal mediante datos de RNA-seq de célula única, observamos que sus patrones de expresión en este contexto son casi idénticos. Además, ambos

presentan su mayor expresión entre los estadios NF 11 a 13, etapa en la que ocurre la inducción y formación de la CN, seguida de una disminución a partir del estadio NF 14 en adelante. Este comportamiento sugiere que podrían estar actuando de manera coordinada durante esta etapa del desarrollo de la CN.

Los resultados del primer objetivo de esta tesis nos permitieron describir por primera vez el patrón de expresión de *cited2* durante el desarrollo embrionario de *X. tropicalis*. Sin embargo, aún falta por dilucidar su función específica en este modelo, particularmente a nivel de la CN, además de indagar más en la posible interacción Cited2-Ric-8A. A futuro tenemos planeado responder estas incógnitas y ahondar en el rol de *cited2* en el desarrollo craneofacial y cardíaco. Para esto último, ya se han obtenido imágenes de referencia de corazón embrionario de *X. tropicalis* (ver Anexo 7), y ya tenemos experimentos propuestos para evaluar la interacción entre ambas proteínas, tanto a nivel de formación de complejo proteico como a nivel funcional.

En el objetivo bioinformático, se realizó un análisis con distintas herramientas para poder identificar una subpoblación de CN cardíaca dentro de un set de datos de RNAseq de célula única disponibles públicamente (Briggs y col., 2018). Estos datos, al estar ya previamente procesados y anotados, nos permitió realizar el análisis usando únicamente aquellas células correspondientes a la CN.

Los sets de datos originales incluían un total de 26550 genes y 1033 células de NC (879 de NC craneal y 154 de NC cordal, mezcladas). Por esto, uno de los

focos principales del pipeline utilizado fue el reducir el número de genes de interés finales, filtrando únicamente a los más importantes. Cabe destacar que, al ser ya todas células de la CN, es de esperarse que haya mucha similitud entre ellas.

En este análisis, se realizó una reducción de dimensionalidad por PCA (Maćkiewicz y Ratajczak, 1993) como paso previo a la generación de mapas UMAP (McInnes y col., 2018), a partir de los cuales se aplicó una *clusterización* utilizando el algoritmo de Leiden (Traag y col., 2019). La combinación de estas tres metodologías, ampliamente utilizadas en estudios de RNA-seq de célula única, permitió agrupar las células individuales según la similitud de sus perfiles de expresión. Sin embargo, dado que los resultados dependen de los parámetros utilizados (*n_neighbors* y resolución, en este caso), el análisis fue complementado mediante el algoritmo de *Random Forest* (Breiman, 2001). Esta estrategia nos permitió seleccionar los valores adecuados de *n_neighbors* y de resolución comparando las métricas de precisión y matrices de confusión entregadas por modelos de aprendizaje optimizados.

A partir de esto, se determinó que las células se agrupaban de forma más robusta en 3 clústeres bien definidos utilizando una resolución de 0.5. Resoluciones mayores generaron una mayor cantidad de clústeres, pero que eran más difíciles de distinguir para el modelo, lo que sugiere una menor diferenciación entre perfiles de expresión. Con el modelo final, además, se extrajo

la lista de genes más informativos al momento de distinguir entre clústeres, lo cual se desprende de la métrica de “importancia” propia del aprendizaje de *Random Forest*. Se identificaron 328 genes con valores de importancia mayores a cero, lo cual indica que todos contribuyeron, en alguna medida, a mejorar la precisión del modelo. Sin embargo, se observó que un solo gen presentaba una importancia considerablemente más alta que el resto, sugiriendo que la separación de los clústeres podría estar fuertemente influenciada por este.

Esta lista de genes fue utilizada posteriormente para determinar los marcadores específicos para cada clúster y para combinaciones entre ellos. De forma interesante, vemos una gran cantidad de genes marcadores compartidos entre los clústeres 1 y 2. Este resultado es de esperarse, debido a que, si observamos el mapa UMAP original, los clústeres 1 y 2 se encuentran bastante cerca entre sí y considerablemente alejados del clúster 0.

El gen que más destacó en estos análisis fue *gsk3b* que codifica para la serina/treonina-kinasa, GSK3 β , y se expresaba de forma alta en el clúster 1 (para el cual es marcador), de forma baja en el clúster 0, y de forma intermedia en el clúster 2. GSK3 β cuenta con un rol ampliamente estudiado tanto en desarrollo embrionario como en otros contextos (Beurel y col., 2015; Karati y col., 2025; McCubrey y col., 2014), al ser una proteína con más de 100 sustratos conocidos y cuya función más importante es ser un inhibidor de la vía de señalización Wnt al fosforilar β -catenina y enviarla a degradación (Beurel y col., 2015; McCubrey y

col., 2014). También se ha observado que GSK3 β es necesaria para el correcto desarrollo de la CN en distintas etapas. Por ejemplo, *twist*, un factor de transcripción clave para la especificación, TEM y migración de la CN (Martik y Bronner, 2017; Piacentino y col., 2020; Prasad y col., 2012; Simões-Costa y Bronner, 2015), ha sido descrito como un blanco de fosforilación GSK3 β , la cual es necesaria para la correcta TEM de las CN (Lander y col., 2013). GSK3 β también es capaz de alterar la migración de la CN mediante otros mecanismos: por un lado, Abbruzzese y col. (2014) observó que la metaloproteasa ADAM13 requiere ser fosforilada por GSK3 β para cumplir su función en el núcleo, la cual activa la transcripción de genes pro-migratorios. Por otro lado, Gonzalez Malagon y col. (2018) determinó que la inhibición de GSK3 β genera defectos en la migración de la CN debido a disrupción del citoesqueleto y pérdida de polaridad celular. Con todos esos antecedentes, no es sorpresa que este sea uno de los genes más expresados. Sin embargo, llama la atención la diferencia marcada de sus niveles relativos en cada clúster. Una posible explicación de esto es que los clústeres no estén divididos según subpoblación, sino que por distintas etapas de desarrollo de la CN. Respaldando esta idea, otros análisis de RNAseq de célula única de células de la CN también han encontrado una heterogenicidad inesperada (Gandhi y col., 2020; Kotov y col., 2024). Otro punto importante es que, marcando los clústeres 1 y 2, encontramos al gen *snai1*, el cual es característico de la CN temprana.

Dicho esto, nos preguntamos si podría estar la subpoblación cardiaca en alguno de los clústeres identificados. Para responder esta pregunta, recurrimos al análisis bibliográfico en búsqueda de genes asociados al desarrollo cardiovascular que nos permitirían identificar dicha subpoblación.

En el clúster 0, encontramos dos genes de la familia RBM (*Ribosome Binding Motif*), *rbm6* y *rbmx*. Si bien ninguno de estos está directamente asociado a una función cardiaca, dos ortólogos de la misma familia (*rbm24a/b*) se han descrito como necesarios para la correcta formación de la arteria dorsal, vena caudal y vena caudal posterior en pez cebra, las cuales corresponden a los primeros vasos sanguíneos que se forman en el embrión (Maragh y col., 2011). Sin embargo, al evaluar la similitud entre las secuencias de *rbm24a/b* y *rbm6/x*, esta es muy baja como para sugerir que podrían tener un rol similar en estos procesos.

En el clúster 1, encontramos a los genes *Imln*, también llamado *invadolysin*, y el antes mencionado *gsk3b*. Aparte de lo ya descrito, se sabe que la función de GSK3 β es necesaria para el desarrollo cardiaco de anfibios al inhibir la vía Wnt en el mesodermo (V. A. Schneider y Mercola, 2001). Por su parte, *Imln* es una metaloproteasa altamente conservada que se ha descrito como necesaria para la migración de melanóforos (células pigmentadas derivadas de la CN) y la formación de vasos sanguíneos formados por angiogénesis, pero no para los formados por vasculogénesis (como la aorta dorsal).

Finalmente, en el clúster 2 encontramos a *mbd3*, un componente del complejo de remodelación de nucleosomas y deacetilasas NuRD (Le Guezennec y col., 2006) y *tpm4*, una tropomiosina. NuRD está asociado a la represión génica y es esencial para el desarrollo del corazón, siendo reclutado por factores de transcripción como *gata4* y *nkx2.5* para la inactivación de genes no-cardiacos, proceso necesario para evitar malformaciones (Robbe y col., 2022). Por otro lado, las tropomiosinas son proteínas importantes para la contracción del músculo cardíaco, y *tpm4* en particular ha sido descrita como necesaria para la actividad contráctil del corazón de embriones de pez cebra (Zhao y col., 2008).

Cabe destacar que, en el primer análisis de expresión diferencial, se encontró que los genes *mafb*, *mafk*, *sox8*, *ets1*, *tgif1* y *hand1* estaban diferencialmente expresados al alza en el clúster 1 respecto a la suma de las demás células. Esto resulta muy interesante, debido a que todos ellos están fuertemente relacionados con la identidad de la CN cardíaca (Erhardt y col., 2021; Gandhi y col., 2020; Grossfeld, 2022; Lin y col., 2022; Nie y Bronner, 2015; Snider y col., 2007; Tani-Matsuhana y col., 2018; Zhu y col., 2025). Sin embargo, todos estos genes quedaron fuera de la lista final una vez se filtró por valores de importancia para el modelo de *Random Forest*. Esto podría responder a que el valor de log₂FC en todos ellos era menor a 2, lo cual, si bien a nivel biológico no es despreciable, podría no ser suficiente para categorizar entre clústeres, especialmente teniendo la fuerte influencia de *gsk3b*, causando así que el modelo los excluyera. Con todo

esto, podríamos concluir que el clúster 1 (seguido o en conjunto al clúster 2, que es muy similar a éste), sería el mejor candidato para contener a la CN cardiaca. Sin embargo, no contamos con evidencia suficiente para poder hacer una afirmación sólida basándonos en la información obtenida.

Desde un punto de vista crítico, existen múltiples opciones que podrían mejorar la calidad del análisis: utilizar un set de datos de CN más tardía, como lo sería el NF 22 (el más avanzado en los datos disponibles). Por otro lado, se podría eliminar *gsk3b* del set de datos antes de la búsqueda de clústeres, para evitar que su expresión apantalle las diferencias entre otros genes. Y, finalmente, proponemos evaluar la expresión diferencial únicamente entre pares de clústeres y/o pulir los criterios de filtrado para que representen de mejor forma el sentido biológico de los datos.

Cabe destacar que tenemos planes de repetir el análisis con todas las consideraciones mencionadas, pero esta vez utilizando los datos de Petrova y col. (2024), quien re-secuenció los datos de Briggs y col. y repitió los análisis de búsqueda de clústeres y anotación utilizando estrategias modernizadas y una versión más reciente del genoma de referencia de *X. tropicalis*.

CONCLUSIONES

- ARNm de *xtcited2* se encuentra presente en todos los estadios del desarrollo observados, que van desde blástula tardía (NF 9) hasta renacuajo en fase de alimentación (NF 45).
- En estadios tempranos, *xtcited2* se encuentra altamente expresado, particularmente en el polo animal del embrión y en tejido mesodermal. Posteriormente, se especifica a regiones como el tubo neural, cresta neural, cerebro y corazón.
- En cresta neural, *xtcited2* se encuentra localizado principalmente en el núcleo de las células.
- Comparando la expresión espaciotemporal de *xtcited2* y *xtric-8a*, podemos ver que ambos genes se expresan en tejidos neurales, pero mientras *xtcited2* presenta una expresión más amplia, es decir, se encuentra en diversas regiones del embrión, *xtric-8a* tiene una expresión más específica, concentrándose de forma marcada en estructuras como la vesícula ótica, el tubo neural y la CN, así como sus respectivos derivados, como el prosencéfalo y los arcos branquiales, respectivamente.
- En tejidos que se diferencian a la CN, *xtcited2* muestra cambios dinámicos durante la gastrulación: disminuye en NF 10, aumenta en NF 11, y vuelve a disminuir a partir de NF 14, manteniéndose constante hasta NF 22. Este

patrón es idéntico al de *xtric-8a* en el mismo contexto, lo que sugiere que podrían actuar de manera coordinada durante la etapa de inducción de la CN.

- El pipeline propuesto, a pesar de no haber identificado una subpoblación de CN cardiaca, permitió agrupar las células en clústeres robustos según los niveles de expresión de *gsk3b*, que podría estar definiendo diferentes etapas del desarrollo de la CN.

GLOSARIO

°C	: Grados Celsius
µg	: Microgramos.
µL	: Microlitros.
µM	: Micromolar.
ADN	: Ácido desoxirribonucleico.
ADNc	: Ácido desoxirribonucleico copia.
APS	: Persulfato de amonio.
ARN	: Ácido ribonucleico.
ARNm	: Ácido ribonucleico mensajero.
BSA	: Albúmina de suero bovino.
CIL	: Inhibición de locomoción por contacto
Cited2	: Transactivador interactor de CBP/P300 2.
CMV	: Citomegalovirus.
CN	: Cresta neural.
CoA	: Co-atracción.
dNTP	: Desoxirribonucleótidos trifosfatos.
DTT	: Ditioneitol.
EDTA	: Ácido etilen-diamino-tetracético.
g	: Gramo.
GDP	: Guanosina 5'-difosfato
GEF	: Factor intercambiador de nucleótidos de guanina
GPCR	: Receptor acoplado a Proteína G.
GTP	: Guanosina 5'trifosfato.
HEPES	: Ácido 4-(2-hidroxietil)-1-piperazinetanosulfónico.
HRP	: Peroxidasa de rábano.
IgG	: Inmunoglobulina G.

L	: Litro.
LB	: Medio Luria Berstein.
M	: Molar.
MCS	: Sitio de multiclonamiento.
mg	: Miligramos.
min	: Minutos.
mL	: Mililitros.
mM	: Milimolar.
N	: Normal.
NEB	: New Englands Biolabs.
ng	: Nanogramos.
nM	: Nanomolar.
p/v	: Peso en volumen.
PCR	: Reacción en cadena de la polimerasa.
PEG	: Polietilenglicol.
pg	: Picogramos.
PMSF	: Fluoruro de metilfenilsulfonilo.
Ric-8A	: Resistencia a inhibidores de colinesterasa-8 isoforma A.
rpm	: Revoluciones por minuto.
RT	: Transcripción reversa.
scRNA-seq	: Secuenciación de ARN de célula única.
SDS	: Dodecil sulfato de sodio.
TA	: Temperatura ambiente.
TAE	: Buffer Tris-Acetato EDTA.
TBS	: Buffer Salino Tris.
TE	: Buffer Tris-HCl - EDTA.
TEM	: Transición epitelio-mesenquimal.
TEMED	: N, N, N', N'-tetraetilmetilendiamina.

TENS	: Buffer Tris - EDTA - NaOH - SDS.
Tm	: Temperatura de fusión.
U	: Unidades.
UMI	: Identificador Molecular Único
UTR	: Región no-traducida.
v/v	: Volumen en volumen.
XtCited2	: Transactivador interactor de CBP/P300 2 de <i>Xenopus tropicalis</i> .
XtRic-8A	: Resistencia a inhibidores de colinesterasa-8 isoforma A de <i>Xenopus tropicalis</i> .

BIBLIOGRAFÍA

- Abbruzzese, G., Cousin, H., Salicioni, A. M., & Alfandari, D. (2014). GSK3 and Polo-like kinase regulate ADAM13 function during cranial neural crest cell migration. *Molecular Biology of the Cell*, 25(25), 4072–4082.
<https://doi.org/10.1091/mbc.e14-05-0970>
- An, B., Ji, X., & Gong, Y. (2020). Role of CITED2 in stem cells and cancer. *Oncology Letters*, 20(4), 107. <https://doi.org/10.3892/ol.2020.11968>
- Baldini, A., Fulcoli, F. G., & Illingworth, E. (2017). Tbx1. In *Current Topics in Developmental Biology* (Vol. 122, pp. 223–243). Elsevier.
<https://doi.org/10.1016/bs.ctdb.2016.08.002>
- Bamforth, S. D., Bragança, J., Eloranta, J. J., Murdoch, J. N., Marques, F. I. R., Kranc, K. R., Farza, H., Henderson, D. J., Hurst, H. C., & Bhattacharya, S. (2001). Cardiac malformations, adrenal agenesis, neural crest defects and exencephaly in mice lacking Cited2, a new Tfp2 co-activator. *Nat Genet*, 29(4), 469–474. <https://doi.org/10.1038/ng768>
- Bamforth, S. D., Bragança, J., Farthing, C. R., Schneider, J. E., Broadbent, C., Michell, A. C., Clarke, K., Neubauer, S., Norris, D., Brown, N. A., Anderson, R. H., & Bhattacharya, S. (2004). Cited2 controls left-right patterning and heart development through a Nodal-Pitx2c pathway. *Nat Genet*, 36(11), 1189–1196. <https://doi.org/10.1038/ng1446>

Barriga, E. H., Maxwell, P. H., Reyes, A. E., & Mayor, R. (2013). The hypoxia factor Hif-1 α controls neural crest chemotaxis and epithelial to mesenchymal transition. *Journal of Cell Biology*, 201(5), 759–776. <https://doi.org/10.1083/jcb.201212100>

Benko, S., Fantes, J. A., Amiel, J., Kleinjan, D.-J., Thomas, S., Ramsay, J., Jamshidi, N., Essafi, A., Heaney, S., Gordon, C. T., McBride, D., Golzio, C., Fisher, M., Perry, P., Abadie, V., Ayuso, C., Holder-Espinasse, M., Kilpatrick, N., Lees, M. M., ... Lyonnet, S. (2009). Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. *Nat Genet*, 41(3), 359–364. <https://doi.org/10.1038/ng.329>

Berlow, R. B., Dyson, H. J., & Wright, P. E. (2017). Hypersensitive termination of the hypoxic response by a disordered protein switch. *Nature*, 543(7645), 447–451. <https://doi.org/10.1038/nature21705>

Beurel, E., Grieco, S. F., & Jope, R. S. (2015). Glycogen synthase kinase-3 (GSK3): Regulation, actions, and diseases. *Pharmacology & Therapeutics*, 148, 114–131. <https://doi.org/10.1016/j.pharmthera.2014.11.016>

Beyer, A. (2021). *Estudio de la función de RIC8-A en la remodelación del citoesqueleto de actina durante la migración celular*. Universidad de Concepción.

- Bhattacharya, S., Michels, C. L., Leung, M.-K., Arany, Z. P., Kung, A. L., & Livingston, D. M. (1999). Functional role of p35srj, a novel p300/CBP binding protein, during transactivation by HIF-1. *Genes & Development*, *13*(1), 64–75. <https://doi.org/10.1101/gad.13.1.64>
- Bhattacharjee, V., Horn, K. H., Singh, S., Webb, C. L., Pisano, M. M., & Greene, R. M. (2009). CBP/p300 and associated transcriptional co-activators exhibit distinct expression patterns during murine craniofacial and neural tube development. *The International Journal of Developmental Biology*, *53*(7), 1097–1104. <https://doi.org/10.1387/ijdb.072489vb>
- Bolande, R. (1974). The neurocristopathies A unifying concept of disease arising in neural crest maldevelopment. *Human Pathology*, *5*(4), 409–429. [https://doi.org/10.1016/S0046-8177\(74\)80021-3](https://doi.org/10.1016/S0046-8177(74)80021-3)
- Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Briggs, J. A., Weinreb, C., Wagner, D. E., Megason, S., Peshkin, L., Kirschner, M. W., & Klein, A. M. (2018). The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science (New York, N.Y.)*, *360*(6392). <https://doi.org/10.1126/science.aar5780>
- Carmona-Fontaine, C., Matthews, H. K., Kuriyama, S., Moreno, M., Dunn, G. A., Parsons, M., Stern, C. D., & Mayor, R. (2008). Contact inhibition of

locomotion in vivo controls neural crest directional migration. *Nature*, 456(7224), 957–961. <https://doi.org/10.1038/nature07441>

Carmona-Fontaine, C., Matthews, H., & Mayor, R. (2008). Directional cell migration in vivo: Wnt at the crest. *Cell Adhesion & Migration*, 2(4), 240–242. <https://doi.org/10.4161/cam.2.4.6747>

Carmona-Fontaine, C., Theveneau, E., Tzekou, A., Tada, M., Woods, M., Page, K. M., Parsons, M., Lambris, J. D., & Mayor, R. (2011). Complement Fragment C3a Controls Mutual Cell Attraction during Collective Cell Migration. *Developmental Cell*, 21(6), 1026–1037. <https://doi.org/10.1016/j.devcel.2011.10.012>

Chalpe, A. J., Prasad, M., Henke, A. J., & Paulson, A. (2010). Regulation of cadherin expression in the chicken neural crest by the Wnt/ β -catenin signaling pathway. *Cell Adhesion & Migration*, 4(3), 431–438. <https://doi.org/10.4161/cam.4.3.12138>

Chan, P., Gabay, M., Wright, F. A., & Tall, G. G. (2011). Ric-8B Is a GTP-dependent G Protein α s Guanine Nucleotide Exchange Factor. *Journal of Biological Chemistry*, 286(22), 19932–19942. <https://doi.org/10.1074/jbc.M110.163675>

Chen, Y., Doughman, Y., Gu, S., Jarrell, A., Aota, S., Cvekl, A., Watanabe, M., Dunwoodie, S. L., Johnson, R. S., van Heyningen, V., Kleinjan, D. A.,

- Beebe, D. C., & Yang, Y.-C. (2008). Cited2 is required for the proper formation of the hyaloid vasculature and for lens morphogenesis. *Development*, 135(17), 2939–2948. <https://doi.org/10.1242/dev.021097>
- Chen, Z., Chen, H.-X., Hou, H.-T., Yin, X.-Y., Yang, Q., & He, G.-W. (2023). Identification and Functional Verification of CITED2 Gene Promoter Region in Patients with Patent Ductus Arteriosus. *International Journal of Molecular Sciences*, 24(22), 16204. <https://doi.org/10.3390/ijms242216204>
- Chou, Y.-T., Hsieh, C.-H., Chiou, S.-H., Hsu, C.-F., Kao, Y.-R., Lee, C.-C., Chung, C.-H., Wang, Y.-H., Hsu, H.-S., Pang, S.-T., Shieh, Y.-S., & Wu, C.-W. (2012). CITED2 functions as a molecular switch of cytokine-induced proliferation and quiescence. *Cell Death Differ*, 19(12), 2015–2028. <https://doi.org/10.1038/cdd.2012.91>
- Côté, A., Fanous, A., Almajed, A., & Lacroix, Y. (2015). Pierre Robin sequence: Review of diagnostic and treatment challenges. *International Journal of Pediatric Otorhinolaryngology*, 79(4), 451–464. <https://doi.org/10.1016/j.ijporl.2015.01.035>
- Cotton, M., & Claing, A. (2009). G protein-coupled receptors stimulation and the control of cell migration. *Cellular Signalling*, 21(7), 1045–1053. <https://doi.org/10.1016/j.cellsig.2009.02.008>

- Dianatpour, S., Khatami, M., Heidari, M. M., & Hadadzadeh, M. (2020). Novel Point Mutations of CITED2 Gene Are Associated with Non-familial Congenital Heart Disease (CHD) in Sporadic Pediatric Patients. *Applied Biochemistry and Biotechnology*, 190(3), 896–906.
<https://doi.org/10.1007/s12010-019-03125-8>
- Du, Q., de la Morena, M. T., & van Oers, N. S. C. (2020). The Genetics and Epigenetics of 22q11.2 Deletion Syndrome. *Front. Genet.*, 10, 1365.
<https://doi.org/10.3389/fgene.2019.01365>
- Dunwoodie, S. L., Rodriguez, T. A., & Beddington, R. S. P. (1998). Msg1 and Mrg1, founding members of a gene family, show distinct patterns of gene expression during mouse embryogenesis. *Mechanisms of Development*, 72(1–2), 27–40. [https://doi.org/10.1016/S0925-4773\(98\)00011-2](https://doi.org/10.1016/S0925-4773(98)00011-2)
- Dykes, I. (2014). Left Right Patterning, Evolution and Cardiac Development. *JCDD*, 1(1), 52–72. <https://doi.org/10.3390/jcdd1010052>
- Erhardt, S., Zheng, M., Zhao, X., Le, T. P., Findley, T. O., & Wang, J. (2021). The Cardiac Neural Crest Cells in Heart Development and Congenital Heart Defects. *Journal of Cardiovascular Development and Disease*, 8(8), 89.
<https://doi.org/10.3390/jcdd8080089>
- Etchevers, H. C., Amiel, J., & Lyonnet, S. (2006). Molecular Bases of Human Neurocristopathies. In J.-P. Saint-Jeannet (Ed.), *Neural Crest Induction and*

Differentiation (Vol. 589, pp. 213–234). Springer US.

https://doi.org/10.1007/978-0-387-46954-6_14

Freedman, S. J., Sun, Z.-Y. J., Kung, A. L., France, D. S., Wagner, G., & Eck, M. J. (2003). Structural basis for negative regulation of hypoxia-inducible factor-1 α by CITED2. *Nat Struct Mol Biol*, 10(7), 504–512.

<https://doi.org/10.1038/nsb936>

Fuentealba, J., Toro-Tapia, G., Arriagada, C., Riquelme, L., Beyer, A., Henriquez, J. P., Caprile, T., Mayor, R., Marcellini, S., Hinrichs, M. V, Olate, J., & Torrejón, M. (2013). Ric-8A, a guanine nucleotide exchange factor for heterotrimeric G proteins, is critical for cranial neural crest cell migration. *Developmental Biology*, 378(2), 74–82.

<https://doi.org/10.1016/j.ydbio.2013.04.005>

Fuentealba, J., Toro-Tapia, G., Rodriguez, M., Arriagada, C., Maureira, A., Beyer, A., Villaseca, S., Leal, J. I., Hinrichs, M. V, Olate, J., Caprile, T., & Torrejón, M. (2016). Expression profiles of the G α subunits during *Xenopus tropicalis* embryonic development. *Gene Expression Patterns*, 22(1), 15–25.

<https://doi.org/10.1016/j.gep.2016.09.001>

Gabay, M., Pinter, M. E., Wright, F. A., Chan, P., Murphy, A. J., Valenzuela, D. M., Yancopoulos, G. D., & Tall, G. G. (2011). Ric-8 Proteins Are Molecular Chaperones That Direct Nascent G Protein α Subunit Membrane Association. *Sci. Signal.*, 4(200). <https://doi.org/10.1126/scisignal.2002223>

Gandhi, S., Ezin, M., & Bronner, M. E. (2020). Reprogramming Axial Level

Identity to Rescue Neural-Crest-Related Congenital Heart Defects.

Developmental Cell, 53(3), 300-315.e4.

<https://doi.org/10.1016/j.devcel.2020.04.005>

Gilbert, S. F. (2005). *Biología del Desarrollo* (7a ed.). Panamericana.

Gilman, A. G. (1987). G PROTEINS: TRANSDUCERS OF RECEPTOR-

GENERATED SIGNALS. *Annu. Rev. Biochem.*, 56(1), 615–649.

<https://doi.org/10.1146/annurev.bi.56.070187.003151>

Gonzalez Malagon, S. G., Lopez Muñoz, A. M., Doro, D., Bolger, T. G., Poon,

E., Tucker, E. R., Adel Al-Lami, H., Krause, M., Phiel, C. J., Chesler, L., &

Liu, K. J. (2018). Glycogen synthase kinase 3 controls migration of the

neural crest lineage in mouse and *Xenopus*. *Nature Communications*, 9(1),

1126. <https://doi.org/10.1038/s41467-018-03512-5>

Grossfeld, P. (2022). ETS1 and HLHS: Implications for the Role of the

Endocardium. *Journal of Cardiovascular Development and Disease*, 9(7),

219. <https://doi.org/10.3390/jcdd9070219>

Harland, R. M., & Grainger, R. M. (2011). *Xenopus* research: metamorphosed

by genetics and genomics. *Trends in Genetics*, 27(12), 507–515.

<https://doi.org/10.1016/j.tig.2011.08.003>

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>

Hempel, A., & Kühl, M. (2016). A Matter of the Heart: The African Clawed Frog *Xenopus* as a Model for Studying Vertebrate Cardiogenesis and Congenital Heart Defects. *Journal of Cardiovascular Development and Disease*, *3*(2), 21. <https://doi.org/10.3390/jcdd3020021>

Hoffman, T. L., Javier, A. L., Campeau, S. A., Knight, R. D., & Schilling, T. F. (2007). Tfp2 transcription factors in zebrafish neural crest development and ectodermal evolution. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *308B*(5), 679–691. <https://doi.org/10.1002/jez.b.21189>

Huang, C., Kratzer, M.-C., Wedlich, D., & Kashef, J. (2016). E-cadherin is required for cranial neural crest migration in *Xenopus laevis*. *Developmental Biology*, *411*(2), 159–171. <https://doi.org/10.1016/j.ydbio.2016.02.007>

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, *9*(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>

Jenie, R. I., Nishimura, M., Fujino, M., Nakaya, M., Mizuno, N., Tago, K., Kurose, H., & Itoh, H. (2013). Increased ubiquitination and the crosstalk of G protein signaling in cardiac myocytes: involvement of R ic-8 B in G s suppression by G q signal. *Genes to Cells*, 18(12), 1095–1106.
<https://doi.org/10.1111/gtc.12099>

Karati, D., Meur, S., Roy, S., Mukherjee, S., Debnath, B., Jha, S. K., Sarkar, B. K., Naskar, S., & Ghosh, P. (2025). Glycogen synthase kinase 3 (GSK3) inhibition: a potential therapeutic strategy for Alzheimer's disease. *Naunyn-Schmiedeberg's Archives of Pharmacology*, 398(3), 2319–2342.
<https://doi.org/10.1007/s00210-024-03500-1>

Kerosuo, L., & Bronner-Fraser, M. (2012). What is bad in cancer is good in the embryo: Importance of EMT in neural crest development. *Seminars in Cell & Developmental Biology*, 23(3), 320–332.
<https://doi.org/10.1016/j.semcdb.2012.03.010>

Khudyakov, J., & Bronner-Fraser, M. (2009). Comprehensive spatiotemporal analysis of early chick neural crest network genes. *Developmental Dynamics*, 238(3), 716–723. <https://doi.org/10.1002/dvdy.21881>

Klattenhoff, C., Montecino, M., Soto, X., Guzmán, L., Romo, X., de los Angeles García, M., Mellstrom, B., Naranjo, J. R., Hinrichs, M. V., & Olate, J. (2003). Human brain synembryon interacts with $\text{G}\alpha$ and $\text{Gq}\alpha$ and is translocated to

- the plasma membrane in response to isoproterenol and carbachol. *Journal of Cellular Physiology*, 195(2), 151–157. <https://doi.org/10.1002/jcp.10300>
- Kotov, A., Seal, S., Alkobtawi, M., Kappès, V., Ruiz, S. M., Arbès, H., Harland, R. M., Peshkin, L., & Monsoro-Burq, A. H. (2024). A time-resolved single-cell roadmap of the logic driving anterior neural crest diversification from neural border to migration stages. *Proceedings of the National Academy of Sciences*, 121(19). <https://doi.org/10.1073/pnas.2311685121>
- Lander, R., Nasr, T., Ochoa, S. D., Nordin, K., Prasad, M. S., & LaBonne, C. (2013). Interactions between Twist and other core epithelial–mesenchymal transition factors are controlled by GSK3-mediated phosphorylation. *Nature Communications*, 4(1), 1542. <https://doi.org/10.1038/ncomms2543>
- Le Guezennec, X., Vermeulen, M., Brinkman, A. B., Hoeijmakers, W. A. M., Cohen, A., Lasonder, E., & Stunnenberg, H. G. (2006). MBD2/NuRD and MBD3/NuRD, Two Distinct Complexes with Different Biochemical and Functional Properties. *Molecular and Cellular Biology*, 26(3), 843–851. <https://doi.org/10.1128/MCB.26.3.843-851.2006>
- Leal, J. I., Villaseca, S., Beyer, A., Toro-Tapia, G., & Torrejón, M. (2018). Ric-8A, a GEF for heterotrimeric G-proteins, controls cranial neural crest cell polarity during migration. *Mechanisms of Development*, 154, 170–178. <https://doi.org/10.1016/j.mod.2018.07.004>

- Lee, Y.-H., & Saint-Jeannet, J.-P. (2011). Cardiac neural crest is dispensable for outflow tract septation in *Xenopus*. *Development (Cambridge, England)*, *138*(10), 2025–2034. <https://doi.org/10.1242/dev.061614>
- Li, W., & Cornell, R. A. (2007). Redundant activities of *Tfap2a* and *Tfap2c* are required for neural crest induction and development of other non-neural ectoderm derivatives in zebrafish embryos. *Developmental Biology*, *304*(1), 338–354. <https://doi.org/10.1016/j.ydbio.2006.12.042>
- Lin, L., Pinto, A., Wang, L., Fukatsu, K., Yin, Y., Bamforth, S. D., Bronner, M. E., Evans, S. M., Nie, S., Anderson, R. H., Terskikh, A. V., & Grossfeld, P. D. (2022). *ETS1* loss in mice impairs cardiac outflow tract septation via a cell migration defect autonomous to the neural crest. *Human Molecular Genetics*, *31*(24), 4217–4227. <https://doi.org/10.1093/hmg/ddac174>
- Liu, J. A. J., Wu, M.-H., Yan, C. H., Chau, B. K. H., So, H., Ng, A., Chan, A., Cheah, K. S. E., Briscoe, J., & Cheung, M. (2013). Phosphorylation of *Sox9* is required for neural crest delamination and is regulated downstream of BMP and canonical Wnt signaling. *Proc. Natl. Acad. Sci. U.S.A.*, *110*(8), 2882–2887. <https://doi.org/10.1073/pnas.1211747110>
- Lu, W., Guzman, A. R., Yang, W., Chapa, C. J., Shaw, G. M., Greene, R. M., Pisano, M. M., Lammer, E. J., Finnell, R. H., & Zhu, H. (2010). Genes encoding critical transcriptional activators for murine neural tube

development and human spina bifida: a case-control study. *BMC Medical Genetics*, 11(1), 141. <https://doi.org/10.1186/1471-2350-11-141>

MacDonald, S. T., Bamforth, S. D., Chen, C.-M., Farthing, C. R., Franklyn, A., Broadbent, C., Schneider, J. E., Saga, Y., Lewandoski, M., & Bhattacharya, S. (2008). Epiblastic Cited2 deficiency results in cardiac phenotypic heterogeneity and provides a mechanism for haploinsufficiency. *Cardiovascular Research*, 79(3), 448–457. <https://doi.org/10.1093/cvr/cvn101>

Maćkiewicz, A., & Ratajczak, W. (1993). Principal components analysis (PCA). *Computers & Geosciences*, 19(3), 303–342. [https://doi.org/10.1016/0098-3004\(93\)90090-R](https://doi.org/10.1016/0098-3004(93)90090-R)

Malbon, C. C. (2005). G proteins in development. *Nature Reviews Molecular Cell Biology*, 6(9), 689–701. <https://doi.org/10.1038/nrm1716>

Maldonado-Agurto, R., Toro, G., Fuentealba, J., Arriagada, C., Campos, T., Albistur, M., Henriquez, J. P., Olate, J., Hinrichs, M. V., & Torrejón, M. (2011). Cloning and spatiotemporal expression of RIC-8 in *Xenopus* embryogenesis. *Gene Expression Patterns*, 11(7), 401–408. <https://doi.org/10.1016/j.gep.2011.06.001>

Maragh, S., Miller, R. A., Bessling, S. L., McGaughey, D. M., Wessels, M. W., de Graaf, B., Stone, E. A., Bertoli-Avella, A. M., Gearhart, J. D., Fisher, S., &

- McCallion, A. S. (2011). Identification of RNA binding motif proteins essential for cardiovascular development. *BMC Developmental Biology*, 11(1), 62. <https://doi.org/10.1186/1471-213X-11-62>
- Martik, M. L., & Bronner, M. E. (2017). Regulatory Logic Underlying Diversification of the Neural Crest. *Trends in Genetics*, 33(10), 715–727. <https://doi.org/10.1016/j.tig.2017.07.015>
- Mayor, R., & Etienne-Manneville, S. (2016). The front and rear of collective cell migration. *Nat Rev Mol Cell Biol*, 17(2), 97–109. <https://doi.org/10.1038/nrm.2015.14>
- Mayor, R., & Theveneau, E. (2013). The neural crest. *Development*, 140(11), 2247–2251. <https://doi.org/10.1242/dev.091751>
- McCubrey, J. A., Steelman, L. S., Bertrand, F. E., Davis, N. M., Abrams, S. L., Montalto, G., D'Assoro, A. B., Libra, M., Nicoletti, F., Maestro, R., Basecke, J., Cocco, L., Cervello, M., & Martelli, A. M. (2014). Multifaceted roles of GSK-3 and Wnt/ β -catenin in hematopoiesis and leukemogenesis: opportunities for therapeutic intervention. *Leukemia*, 28(1), 15–33. <https://doi.org/10.1038/leu.2013.184>
- McInnes, L., Healy, J., Saul, N., & Großberger, L. (2018). UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software*, 3(29), 861. <https://doi.org/10.21105/joss.00861>

- McKinney, W. (2010). *Data Structures for Statistical Computing in Python*. 56–61. <https://doi.org/10.25080/Majora-92bf1922-00a>
- Meulemans, D., & Bronner-Fraser, M. (2004). Gene-Regulatory Interactions in Neural Crest Evolution and Development. *Developmental Cell*, 7(3), 291–299. <https://doi.org/10.1016/j.devcel.2004.08.007>
- Miller, K. G., Emerson, M. D., McManus, J. R., & Rand, J. B. (2000). RIC-8 (Synembryn). *Neuron*, 27(2), 289–299. [https://doi.org/10.1016/S0896-6273\(00\)00037-4](https://doi.org/10.1016/S0896-6273(00)00037-4)
- Mohun, T. J., Leong, L. M., Weninger, W. J., & Sparrow, D. B. (2000). The Morphology of Heart Development in *Xenopus laevis*. *Developmental Biology*, 218(1), 74–88. <https://doi.org/10.1006/dbio.1999.9559>
- Nie, S., & Bronner, M. E. (2015). Dual developmental role of transcriptional regulator Ets1 in *Xenopus* cardiac neural crest vs. heart mesoderm. *Cardiovascular Research*, 106(1), 67–75. <https://doi.org/10.1093/cvr/cvv043>
- Nieto, M. A. (2013). Epithelial Plasticity: A Common Theme in Embryonic and Cancer Cells. *Science*, 342(6159). <https://doi.org/10.1126/science.1234850>
- Nieuwkoop, P. D., & Faber, J. (1994). *Normal Table of *Xenopus laevis* (Doudin)*. Garland Publishing Inc.
- Owens, N. D. L., Blitz, I. L., Lane, M. A., Patrushev, I., Overton, J. D., Gilchrist, M. J., Cho, K. W. Y., & Khokha, M. K. (2016). Measuring Absolute RNA

Copy Numbers at High Temporal Resolution Reveals Transcriptome

Kinetics in Development. *Cell Reports*, 14(3), 632–647.

<https://doi.org/10.1016/j.celrep.2015.12.050>

Pacheco-Leyva, I., Matias, A. C., Oliveira, D. V., Santos, J. M. A., Nascimento, R., Guerreiro, E., Michell, A. C., van De Vrugt, A. M., Machado-Oliveira, G., Ferreira, G., Domian, I., & Bragança, J. (2016). CITED2 Cooperates with ISL1 and Promotes Cardiac Differentiation of Mouse Embryonic Stem Cells. *Stem Cell Reports*, 7(6), 1037–1049.

<https://doi.org/10.1016/j.stemcr.2016.10.002>

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

Petrova, K., Tretiakov, M., Kotov, A., Monsoro-Burq, A. H., & Peshkin, L. (2024).

A new atlas to study embryonic cell types in *Xenopus*. *Developmental Biology*, 511, 76–83. <https://doi.org/10.1016/j.ydbio.2024.04.003>

Piacentino, M. L., Li, Y., & Bronner, M. E. (2020). Epithelial-to-mesenchymal transition and different migration strategies as viewed from the neural crest. *Current Opinion in Cell Biology*, 66, 43–50.

<https://doi.org/10.1016/j.ceb.2020.05.001>

- Pilon, N. (2021). Treatment and Prevention of Neurocristopathies. *Trends in Molecular Medicine*, 27(5), 451–468.
<https://doi.org/10.1016/j.molmed.2021.01.009>
- Prasad, M. S., Sauka-Spengler, T., & LaBonne, C. (2012). Induction of the neural crest state: Control of stem cell attributes by gene regulatory, post-transcriptional and epigenetic interactions. *Developmental Biology*, 366(1), 10–21. <https://doi.org/10.1016/j.ydbio.2012.03.014>
- Ridley, A. J., Schwartz, M. A., Burridge, K., Firtel, R. A., Ginsberg, M. H., Borisy, G., Parsons, J. T., & Horwitz, A. R. (2003). Cell Migration: Integrating Signals from Front to Back. *Science*, 302(5651), 1704–1709.
<https://doi.org/10.1126/science.1092053>
- Robbe, Z. L., Shi, W., Wasson, L. K., Scialdone, A. P., Wilczewski, C. M., Sheng, X., Hepperla, A. J., Akerberg, B. N., Pu, W. T., Cristea, I. M., Davis, I. J., & Conlon, F. L. (2022). CHD4 is recruited by GATA4 and NKX2-5 to repress noncardiac gene programs in the developing heart. *Genes & Development*, 36(7–8), 468–482. <https://doi.org/10.1101/gad.349154.121>
- Rodríguez, M. (2016). *Estudio funcional de Ric-8A durante la inducción de las Crestas Neurales en Xenopus tropicalis*. Universidad de Concepción.

- Roycroft, A., & Mayor, R. (2016). Molecular basis of contact inhibition of locomotion. *Cellular and Molecular Life Sciences*, 73(6), 1119–1130. <https://doi.org/10.1007/s00018-015-2090-0>
- Ruiz, M. J. (2024). *Estudio funcional de Ric8A/Gαq en la formación del mesoderma dorsal en Xenopus tropicalis* [Tesis de Magíster]. Universidad de Concepción.
- Sauka-Spengler, T., & Bronner-Fraser, M. (2008). A gene regulatory network orchestrates neural crest formation. *Nature Reviews Molecular Cell Biology*, 9(7), 557–568. <https://doi.org/10.1038/nrm2428>
- Scarpa, E., Szabó, A., Bibonne, A., Theveneau, E., Parsons, M., & Mayor, R. (2015). Cadherin Switch during EMT in Neural Crest Cells Leads to Contact Inhibition of Locomotion via Repolarization of Forces. *Developmental Cell*, 34(4), 421–434. <https://doi.org/10.1016/j.devcel.2015.06.012>
- Schlange, T., Andrée, B., Arnold, H.-H., & Brand, T. (2000). Expression analysis of the chicken homologue of CITED2 during early stages of embryonic development. *Mechanisms of Development*, 98(1–2), 157–160. [https://doi.org/10.1016/S0925-4773\(00\)00454-8](https://doi.org/10.1016/S0925-4773(00)00454-8)
- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7), 671–675. <https://doi.org/10.1038/nmeth.2089>

- Schneider, V. A., & Mercola, M. (2001). Wnt antagonism initiates cardiogenesis in *Xenopus laevis*. *Genes & Development*, 15(3), 304–315.
<https://doi.org/10.1101/gad.855601>
- Scholl, A. M., & Kirby, M. L. (2009). Signals controlling neural crest contributions to the heart. *WIREs Systems Biology and Medicine*, 1(2), 220–227.
<https://doi.org/10.1002/wsbm.8>
- Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. *Proceedings of the 9th Python in Science Conference (SciPy 2010)*.
- Session, A. M., Uno, Y., Kwon, T., Chapman, J. A., Toyoda, A., Takahashi, S., Fukui, A., Hikosaka, A., Suzuki, A., Kondo, M., van Heeringen, S. J., Quigley, I., Heinz, S., Ogino, H., Ochi, H., Hellsten, U., Lyons, J. B., Simakov, O., Putnam, N., ... Rokhsar, D. S. (2016). Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature*, 538(7625), 336–343.
<https://doi.org/10.1038/nature19840>
- Shellard, A., & Mayor, R. (2016). Chemotaxis during neural crest migration. *Seminars in Cell & Developmental Biology*, 55, 111–118.
<https://doi.org/10.1016/j.semcdb.2016.01.031>
- Shin, D. H., Li, S. H., Chun, Y.-S., Huang, L. E., Kim, M.-S., & Park, J.-W. (2008). CITED2 mediates the paradoxical responses of HIF-1 α to

proteasome inhibition. *Oncogene*, 27(13), 1939–1944.

<https://doi.org/10.1038/sj.onc.1210826>

Siderovski, D. P., & Willard, F. S. (2005). The GAPs, GEFs, and GDIs of heterotrimeric G-protein alpha subunits. *International Journal of Biological Sciences*, 1(2), 51–66. <https://doi.org/10.7150/ijbs.1.51>

Simões-Costa, M., & Bronner, M. E. (2015). Establishing neural crest identity: a gene regulatory recipe. *Development*, 142(2), 242–257.

<https://doi.org/10.1242/dev.105445>

Simoies-Costa, M., & Bronner, M. E. (2016). Reprogramming of avian neural crest axial identity and cell fate. *Science*, 352(6293), 1570–1573.

<https://doi.org/10.1126/science.aaf2729>

Snider, P., Olaopa, M., Firulli, A. B., & Conway, S. J. (2007). Cardiovascular Development and the Colonizing Cardiac Neural Crest Lineage. *The Scientific World JOURNAL*, 7, 1090–1113.

<https://doi.org/10.1100/tsw.2007.189>

Steventon, B., Araya, C., Linker, C., Kuriyama, S., & Mayor, R. (2009). Differential requirements of BMP and Wnt signalling during gastrulation and neurulation define two steps in neural crest induction. *Development*, 136(5), 771–779. <https://doi.org/10.1242/dev.029017>

- Steventon, B., & Mayor, R. (2012). Early neural crest induction requires an initial inhibition of Wnt signals. *Developmental Biology*, 365(1), 196–207.
<https://doi.org/10.1016/j.ydbio.2012.02.029>
- Stuhlmiller, T. J., & García-Castro, M. I. (2012). Current perspectives of the signaling pathways directing neural crest induction. *Cellular and Molecular Life Sciences*, 69(22), 3715–3737. <https://doi.org/10.1007/s00018-012-0991-8>
- Syrovatkina, V., Alegre, K. O., Dey, R., & Huang, X.-Y. (2016). Regulation, Signaling, and Physiological Functions of G-Proteins. *Journal of Molecular Biology*, 428(19), 3850–3868. <https://doi.org/10.1016/j.jmb.2016.08.002>
- Szabó, A., & Mayor, R. (2015). Cell traction in collective cell migration and morphogenesis: The chase and run mechanism. *Cell Adhesion & Migration*, 9(5), 380–383. <https://doi.org/10.1080/19336918.2015.1019997>
- Tall, G. G. (2013). Ric-8 regulation of heterotrimeric G proteins. *Journal of Receptors and Signal Transduction*, 33(3), 139–143.
<https://doi.org/10.3109/10799893.2013.763828>
- Tall, G. G., Krumin, A. M., & Gilman, A. G. (2003). Mammalian Ric-8A (Synembryn) Is a Heterotrimeric G α Protein Guanine Nucleotide Exchange Factor. *Journal of Biological Chemistry*, 278(10), 8356–8362.
<https://doi.org/10.1074/jbc.M211862200>

- Taneyhill, L. A., Coles, E. G., & Bronner-Fraser, M. (2007). Snail2 directly represses cadherin6B during epithelial-to-mesenchymal transitions of the neural crest. *Development*, *134*(8), 1481–1490.
<https://doi.org/10.1242/dev.02834>
- Tani-Matsuhana, S., Vieceli, F. M., Gandhi, S., Inoue, K., & Bronner, M. E. (2018). Transcriptome profiling of the cardiac neural crest reveals a critical role for MafB. *Developmental Biology*, *444*, S209–S218.
<https://doi.org/10.1016/j.ydbio.2018.09.015>
- Theveneau, E., Marchant, L., Kuriyama, S., Gull, M., Moepps, B., Parsons, M., & Mayor, R. (2010). Collective Chemotaxis Requires Contact-Dependent Cell Polarity. *Developmental Cell*, *19*(1), 39–53.
<https://doi.org/10.1016/j.devcel.2010.06.012>
- Theveneau, E., & Mayor, R. (2012a). Cadherins in collective cell migration of mesenchymal cells. *Current Opinion in Cell Biology*, *24*(5), 677–684.
<https://doi.org/10.1016/j.ceb.2012.08.002>
- Theveneau, E., & Mayor, R. (2012b). Neural crest migration: interplay between chemorepellents, chemoattractants, contact inhibition, epithelial-mesenchymal transition, and collective cell migration: Neural crest migration. *WIREs Dev Biol*, *1*(3), 435–445. <https://doi.org/10.1002/wdev.28>

Theveneau, E., Steventon, B., Scarpa, E., Garcia, S., Trepate, X., Streit, A., & Mayor, R. (2013). Chase-and-run between adjacent cell populations promotes directional collective migration. *Nat Cell Biol*, *15*(7), 763–772. <https://doi.org/10.1038/ncb2772>

Thul, P. J., Åkesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait Blal, H., Alm, T., Asplund, A., Björk, L., Breckels, L. M., Bäckström, A., Danielsson, F., Fagerberg, L., Fall, J., Gatto, L., Gnann, C., Hober, S., Hjelmare, M., Johansson, F., ... Lundberg, E. (2017). A subcellular map of the human proteome. *Science*, *356*(6340). <https://doi.org/10.1126/science.aal3321>

Tõnissoo, T., Lulla, S., Meier, R., Saare, M., Ruisu, K., Pooga, M., & Karis, A. (2010). Nucleotide exchange factor RIC-8 is indispensable in mammalian early development. *Developmental Dynamics*, *239*(12), 3404–3415. <https://doi.org/10.1002/dvdy.22480>

Toro-Tapia, G., Villaseca, S., Beyer, A., Roycroft, A., Marcellini, S., Mayor, R., & Torrejón, M. (2018). The Ric-8A/Gα13/FAK signaling cascade controls focal adhesion formation during neural crest cell migration. *Development*, *dev.164269*. <https://doi.org/10.1242/dev.164269>

Toro-Tapia, G., Villaseca, S., Leal, J. I., Beyer, A., Fuentealba, J., & Torrejón, M. (2017). *Xenopus* as a model organism to study heterotrimeric G-protein pathway during collective cell migration of neural crest. *Genesis*, *55*(1–2), e23008. <https://doi.org/10.1002/dvg.23008>

- Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports*, 9(1), 5233. <https://doi.org/10.1038/s41598-019-41695-z>
- Vass, S., & Heck, M. M. S. (2013). Perturbation of invadolysin disrupts cell migration in zebrafish (*Danio rerio*). *Experimental Cell Research*, 319(8), 1198–1212. <https://doi.org/10.1016/j.yexcr.2013.02.005>
- Vega-Lopez, G. A., Cerrizuela, S., Tribulo, C., & Aybar, M. J. (2018). Neurocristopathies: New insights 150 years after the neural crest discovery. *Developmental Biology*, 444, S110–S143. <https://doi.org/10.1016/j.ydbio.2018.05.013>
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., ... Vázquez-Baeza, Y. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Wakamatsu, Y., Nomura, T., Osumi, N., & Suzuki, K. (2014). Comparative gene expression analyses reveal heterochrony for *Sox9* expression in the cranial neural crest during marsupial development. *Evolution & Development*, 16(4), 197–206. <https://doi.org/10.1111/ede.12083>

- Waskom, M. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60), 3021. <https://doi.org/10.21105/joss.03021>
- Watt, K. E. N., & Trainor, P. A. (2014). Neurocristopathies. In *Neural Crest Cells* (pp. 361–394). Elsevier. <https://doi.org/10.1016/B978-0-12-401730-6.00018-1>
- Wolf, F. A., Angerer, P., & Theis, F. J. (2018). SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology*, 19(1), 15. <https://doi.org/10.1186/s13059-017-1382-0>
- Yamagishi, H. (2021). Cardiac Neural Crest. *Cold Spring Harb Perspect Biol*, 13(1), a036715. <https://doi.org/10.1101/cshperspect.a036715>
- Yin, Z., Haynie, J., Yang, X., Han, B., Kiatchoosakun, S., Restivo, J., Yuan, S., Prabhakar, N. R., Herrup, K., Conlon, R. A., Hoit, B. D., Watanabe, M., & Yang, Y.-C. (2002). The essential role of Cited2, a negative regulator for HIF-1 α , in heart development and neurulation. *Proc. Natl. Acad. Sci. U.S.A.*, 99(16), 10488–10493. <https://doi.org/10.1073/pnas.162371799>
- Yoon, J., Kim, J.-H., Lee, O.-J., Yu, S.-B., Kim, J.-I., Kim, S. C., Park, J.-B., Lee, J.-Y., & Kim, J. (2011). xCITED2 Induces Neural Genes in Animal Cap Explants of *Xenopus* Embryos. *Experimental Neurobiology*, 20(3), 123–129. <https://doi.org/10.5607/en.2011.20.3.123>

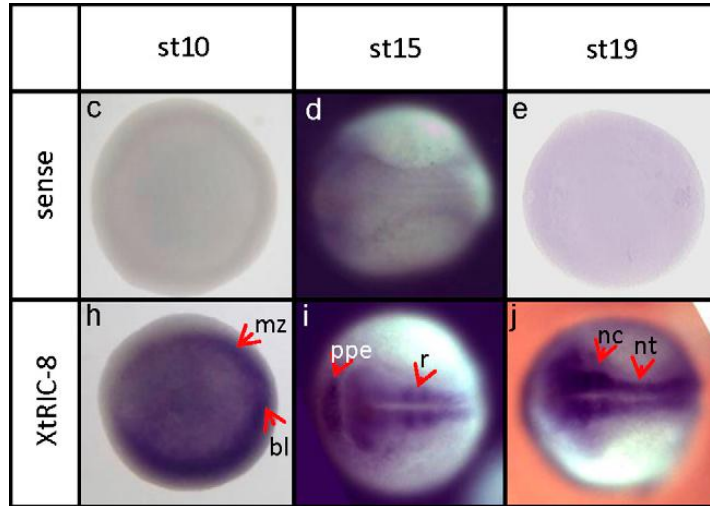
Zhao, L., Zhao, X., Tian, T., Lu, Q., Skrbo-Larsen, N., Wu, D., Kuang, Z., Zheng, X., Han, Y., Yang, S., Zhang, C., & Meng, A. (2008). Heart-specific isoform of tropomyosin4 is essential for heartbeat in zebrafish embryos. *Cardiovascular Research*, *80*(2), 200–208.

<https://doi.org/10.1093/cvr/cvn177>

Zhu, C., Yang, Q., Xu, Q., Song, Y., & Tang, C. (2025). The role of heart and neural crest derivatives-expressed protein factors in pregnancy. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1871*(3), 167639.

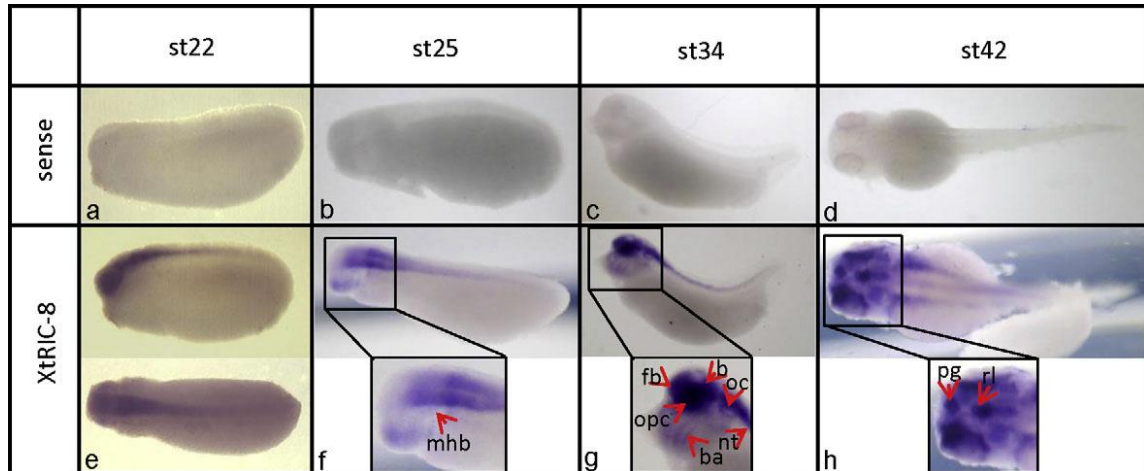
<https://doi.org/10.1016/j.bbadis.2024.167639>

ANEXOS



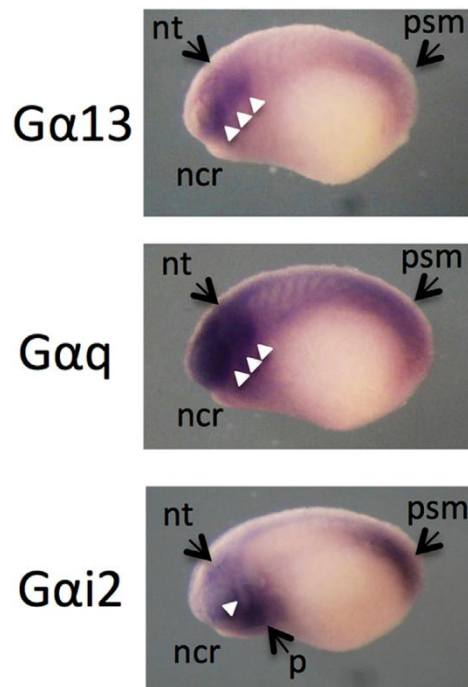
Anexo 1: Expresión espaciotemporal de *xtric-8a* en desarrollo embrionario temprano de *X. tropicalis*. Ensayos de hibridación *in situ* en embrión completo permiten observar la expresión de *xtric-8a*. Los estadios NF mostrados son indicados en los paneles superiores, mientras que la sonda utilizada se indica en el panel a la derecha. Se utilizó tanto una sonda antisentido para ver la expresión espaciotemporal de *xtric-8a* (**h-j**) como una sonda sentido como control negativo (**c-e**). Embriones en estadio gástrula (NF 10) se muestran en vista vegetal, mientras que embriones en estadio de néurula (NF 15 y 19), se muestran en vista dorsal. Las flechas rojas indican estructuras específicas rotuladas como: **mz** = zona marginal; **bl** = labio del blastoporo; **ppe** = ectodermo

pre-placodal; **r** = rombómeros; **nc** = cresta neural; **nt** = tubo neural. Modificado de Maldonado-Agurto y col. 2011.



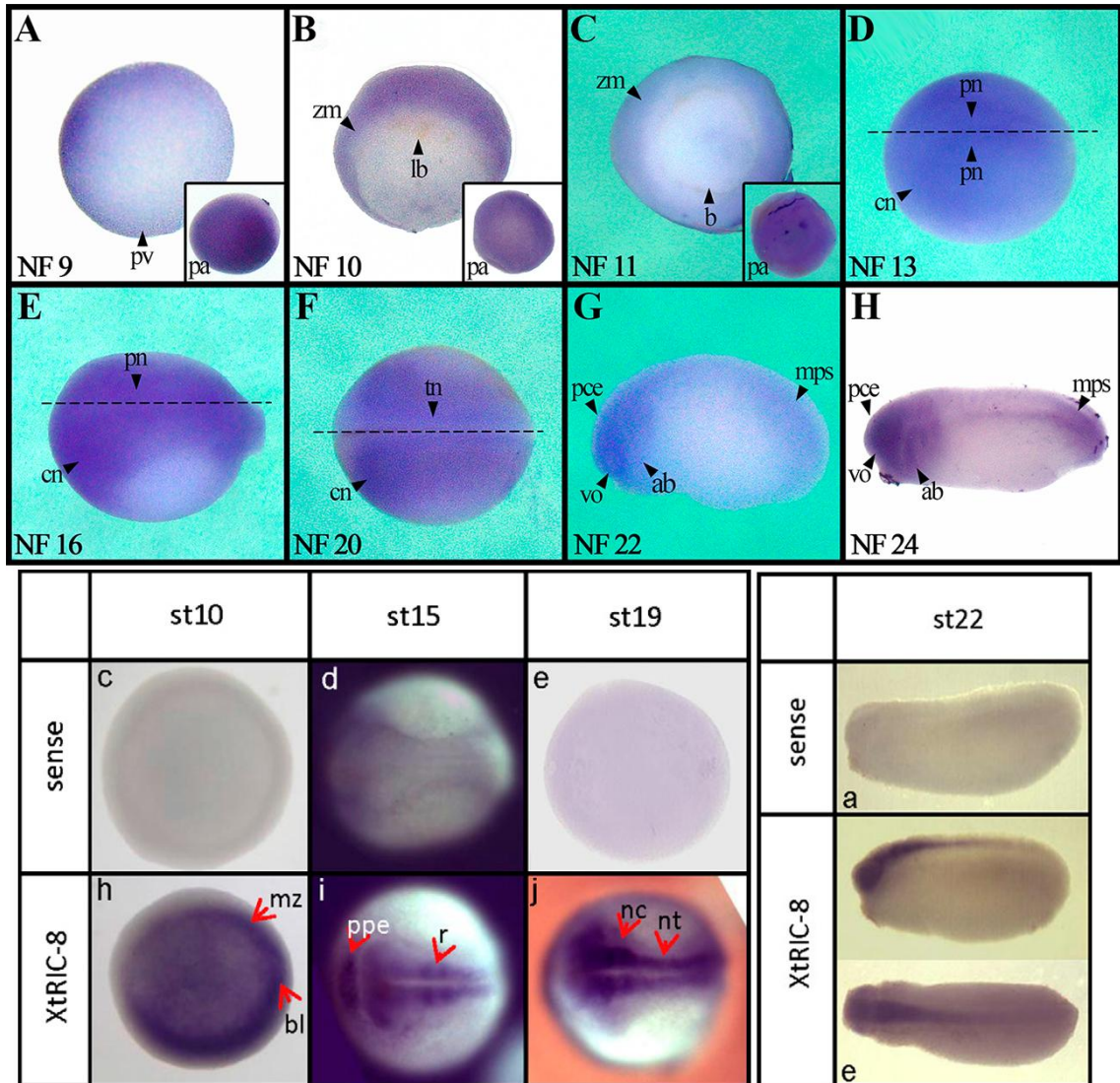
Anexo 2: Expresión espaciotemporal de *xtric-8a* en desarrollo embrionario tardío de *X. tropicalis*. Ensayos de hibridación *in situ* en embrión completo permiten observar la expresión de *xtric-8a*. Los estadios NF mostrados son indicados en los paneles superiores, mientras que la sonda utilizada se indica en el panel a la derecha. Se utilizó tanto una sonda antisentido para ver la expresión espaciotemporal de *xtric-8a* (**e-h**), como una sonda sentido como control negativo (**a-d**). Se muestra vista lateral de embriones en estadios NF 22, 25 y 34 y vista dorsal de embriones en estadios NF 22 y 42. Además, se muestra una magnificación del área de la cabeza en los embriones estadio NF 25, 34 y 42 tratados con sonda antisentido (**f-h**). Las flechas rojas indican estructuras específicas rotuladas como: mhb = mesencéfalo/rombencéfalo; **fb** =

prosencéfalo; b = cerebro; **oc** = cápsula ótica; **opc** = cúpula óptica; **ba** = arcos branquiales; **nt** = tubo neural; **pg** = glándula pineal y **rl** = labio rómbico. Modificado de Maldonado-Agurto y col., 2011.



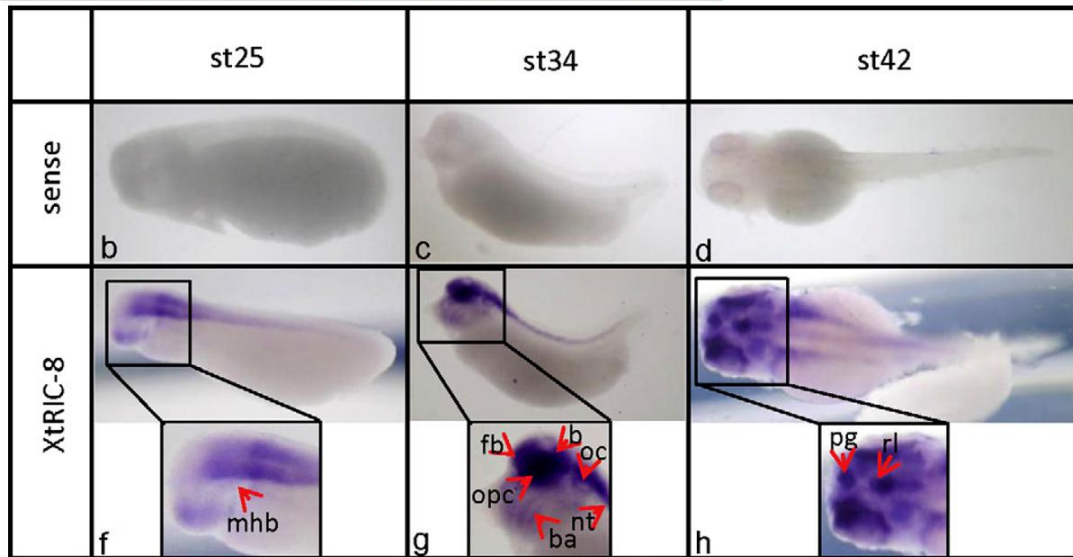
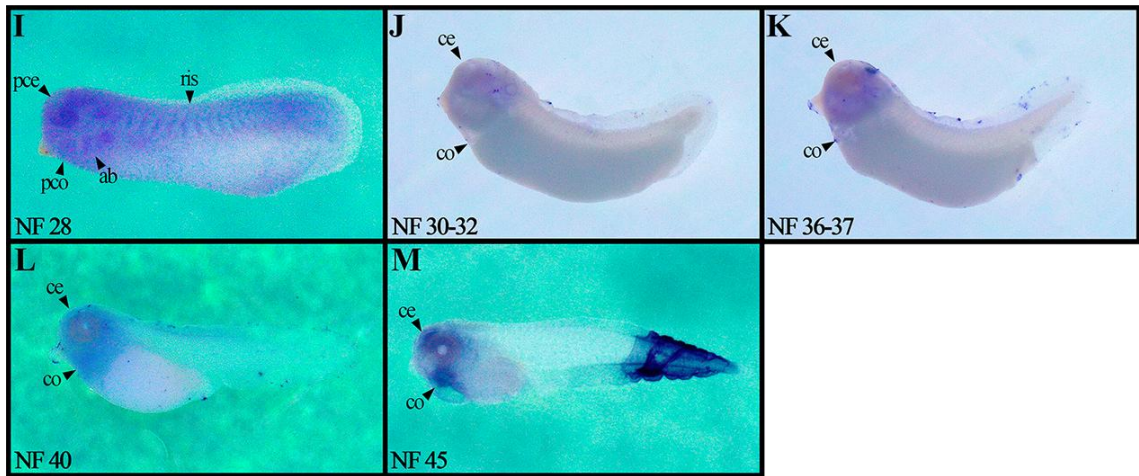
Anexo 3: Expresión espaciotemporal de distintas subunidades Gα en néurula tardía de *X. tropicalis*. Ensayos de hibridación *in situ* en embrión completo permiten observar la expresión espaciotemporal de 3 subunidades Gα en embriones de estadio NF 23-24. A la izquierda se indican las sondas utilizadas. Se muestra una vista lateral de cada embrión. Las cabezas de flecha blancas indican las tres rutas de la cresta neural cranial (**ncr**). Las flechas negras

indican estructuras específicas rotuladas como: **nt** = tubo neural; **p** = placodas y **psm** = mesodermo presomítico. Modificado de Fuentealba y col., 2016.



Anexo 4: Comparación entre patrones de expresión de *xtcited2* y *xtric-8a* en el desarrollo temprano y neurulación de *X. tropicalis*. Tanto *xtcited2*

(arriba) como *xtric-8a* (abajo) se expresan en zona marginal, y tejidos neurales, como tubo y crestas neurales, además de la región ocular. Desde néurula tardía, se observa en áreas cerebrales, región ocular y mesodermo presomítico. La leyenda detallada de cada figura se puede encontrar en las figuras Figura 7, Figura 8, y Anexo 1 y Anexo 2, respectivamente.



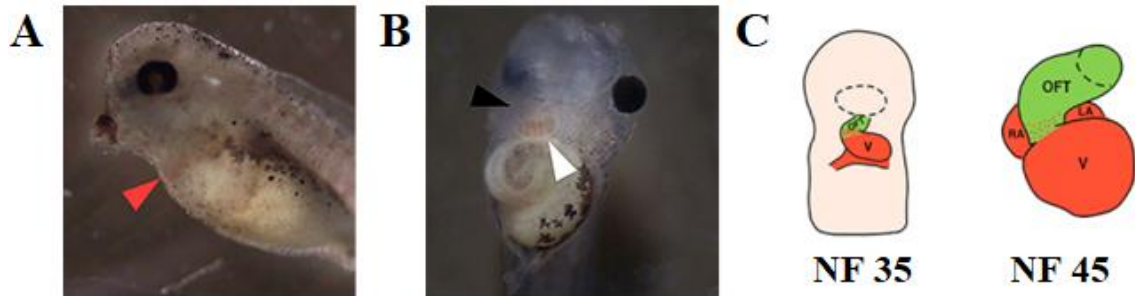
Anexo 5: Comparación entre patrones de expresión de *xtcited2* y *xtric-8a* durante la organogénesis de *X. tropicalis*. La expresión de *xtcited2* se expresa en tejidos cefálicos (prosencéfalo, mesencéfalo, rombencéfalo). Entre estadios 30 y 37, *xtcited2* se expresa en la región craneal y *xtric-8a* en tejido óptico y neural. Desde estadio 40, *xtcited2* también se expresa en el corazón. La

leyenda detallada de cada figura se puede encontrar en la Figura 9 y Anexo 2, respectivamente.

Genes	
Clúster 0	<i>nup98; rpl23; rps26; nnt; grsf1; dlg5; atad1; med15; ubl3; serbp1; poldip3; usp9x; chmp4b; snrpd2; thop1; ak2; tk1; rps11; thop1; rps7; nom1; nup133; nsd1; dus3l; rplp2l; slc25a1; rps16; znf462; dynll1; setd5; rpl36; lrcc8e; tcerg1; hmgb2; naf1; prpf39.2; jagn1; trap1; map4k4; sspb1; akap1; hrp12; rpl35a; ndufc1; mrps7; cycs; midn; Xetrov90028495m; nasp; fzr1; rps27a; casp8ap2; LOC779553; fam122a; mrm3; ccnj; dhx36; srp54; pten; arih2; sdc2; gtf2h2; mocs2; srrm1; psmd3; ubtd1; krt8; sin3b; ndufa5; champ1; aplp2; sf1; mtf2; gtf3a; rbmx; MGC115313; daxx; zbtb44; tmco1; ccar1; mlec; hnrnpa2b1; itpa; tubb; cirbp; ddx41; slc30a7; rpa3; LOC728026; rps19bp1; pak1ip1; sec23b; nop58; tmem258; tmed7; kpna2; rbm6; rpl34; cript; gnb1; nme4; ttc3; grwd1</i>
Clúster 1	<i>gsk3b; eif1b; srsf1; eef1as; rpl32; cypa; pcbp3; slc7a11; Xetrov90023581m; lmln; eif4b</i>
Clúster 2	<i>fam19a3; rpl3; eef1a1; nkain1; dnaja3; tpm4; mbd3; ikzf4; prex1; ccnt1; rpl8; LOC733801; eif2s3; hinfp; ccdc69; tipin; rpl4; ncbp3; filip1l; bysl; trmt2a; npm3; eef1g; oxa1l</i>
Clúster 1+0	<i>rpl6</i>
Clúster 2+0	<i>rps14</i>
Clúster 1+2	<i>tmem69; tubb4b; serpini2; fam221a; MGC69544; vcp; arhgdia; LOC100145470; ppp1r3c.2; lcor; gja7; khdrbs1; pdcd11; cdc42ep4; p2rx1; MGC68456; Xetrov90015453m; cdkal1; nfyc; cdc73; vim; nthl1; pphln1; dmap1; dpp3; agpat9; ppp1r14b; thumpd1; psmd9; wdr12; rcn2; cfl1; foxk2; eif4h; nedd1; farsb; mcat; st6galnac4; rai2; slc35b2; col9a1; kdm5c; edc3; mmp15; psmc6; med20; arrdc3; phf8; lypla2; arhgap22; tsc22d1; snai1; sirt5; map7; calu; ptpn2; emc2; f13b; frmd4b; cit; hmcn2; srpk2; hsp90b1; lsp1; pdcd4; nosip; myg1; nf2; rpl13a; hdlbp; dab2ip; trim33; LOC100494719; gnai2; nckipsd; MGC11257; chic2; sspb3; atg101; uevld; hbs1l; gnl3; dctn4; c6orf136; adgrf5; LOC495703; zdhhc3; Xetrov90029970m; oip5; mrps35; dhx9; rplp2; slc29a2; fzd7; adgrv1; mrp; arpc1b; mrpl36; ism1; pagr1; myog; enc1; tmem161a; hmox1; meis3; Xetrov90022217m; gins4; dgcr2; lztf1; lonrf2; actg1; ppp1r7; plrg1; eftud2; rab2a; sema6a; ing3; pdgfra; Xetrov90026173m; ttc28; gspt1; psmb5; p4ha2; spc25; ube2a; clmp; pa2g4; MGC82963; kti12; hbp1;</i>

	<i>nle1; syncrip; psma4; fstl1; tuba1b; hspa9; usp54; fgfr1op2; hirip3; arc3; myl12b; wars; cct8; Xetrov90029699m; mrc2; ikzf5</i>
No-marcador	<i>pcsk5; mcm6.2; tspo; kdm4a; aspg; kif15-a; ccnt2; app; pus10; tspan14; tomm20; med19; rnf111; elavl1; adnp2; syf2; mmp14; tor1aip1; ppp1cb; atp5o; ubr7; tm9sf4; sec24a; zc3h18; sall4; smo; cxxc1; srsf11; mrpl22; zbtb16; snw1; ppp6r3; ndufa11; apoc1; ncaph2; parp12; gse1; ndfip1; msh6; noc2l; smarca4; btbd10</i>

Anexo 6: Tabla de genes marcadores por clúster. Lista de genes marcadores identificados para cada clúster o combinación de estos. Los clústeres individuales (0, 1 o 2) agrupan genes marcadores específicos de dicho clúster. Las combinaciones de dos clústeres (por ejemplo, 1+2) corresponden a genes cuya expresión es compartida en ambos, pero significativamente menor en el tercero. Los genes “no-marcadores” corresponden a aquellos que no marcan específicamente a ningún clúster. Los nombres fueron manualmente revisados para coincidir con la nomenclatura utilizada en Xenbase. Los genes nombrados como “Xetrov”, “LOC” o “MCG” corresponden a regiones del genoma o transcritos que no han sido descritos o anotados aún.



Anexo 7: Estructuras cardiacas de *X. tropicalis*. (A) Vista lateral de un embrión de estadio 37-38, donde, gracias a su transparencia, se alcanza a ver ligeramente su corazón en desarrollo (cabeza de flecha roja). (B) Vista ventral de un embrión estadio 45, con un corazón completamente desarrollado. Se observa el tracto de salida cardiaco (cabeza de flecha negra) y el ventrículo (cabeza de flecha blanca). (C) Representación esquemática en vista ventral de los corazones en sus respectivos estadios. OFT: tracto de salida cardiaco; RA: Aurícula derecha; LA: aurícula izquierda; V: ventrículo (Modificado de Hempel y Kühl, 2016).