



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



Escola Tècnica Superior  
d'Enginyeria Agronòmica i del Medi Natural

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería Agronómica  
y del Medio Natural

Correlación fenotipo-genotipo de variantes frameshift  
causantes de distrofias hereditarias de la retina.

Trabajo Fin de Grado

Grado en Biotecnología

AUTOR/A: Botey Bataller, María

Tutor/a: García Gimeno, María Adelaida

Cotutor/a externo: García García, Gema

Director/a Experimental: Gómez Escribano, Ana Pilar

CURSO ACADÉMICO: 2023/2024

## **CORRELACIÓN FENOTIPO-GENOTIPO DE VARIANTES *FRAMESHIFT* CAUSANTES DE DISTROFIAS HEREDITARIAS DE LA RETINA**

### **RESUMEN**

Las distrofias hereditarias de la retina son un conjunto de desórdenes en la retina asociados a más de 280 genes, entre ellos, el gen Cone-Rod Homebox (*CRX*), el cual codifica a un factor de transcripción que es clave en la regulación terminal de los fotorreceptores. De ahí, que defectos en este gen conlleven una alteración en estas células y eso desencadena diferentes entidades clínicas como son amaurosis congénita de Leber, retinosis pigmentaria y distrofia de conos y bastones. Ha sido descrito un efecto dominante negativo en relación a las variantes que alteran la pauta de lectura de la proteína. El mecanismo patogénico ha sido atribuido principalmente a una desregulación a nivel de ARN, lo que tiene un impacto en la sobreexpresión de la proteína. Hasta el momento no se ha realizado una correlación clínica en base a los niveles de transcrito y proteína para este tipo de variantes. Por tanto, en este trabajo se ha evaluado el efecto de diferentes variantes “*frameshift*” en *CRX* causantes de los distintos fenotipos clínicos mencionados. Las diferentes variantes producen perfiles de expresión muy diversos tanto a nivel de ARN como de proteína, no pudiendo correlacionar los niveles de expresión elevados a clínicas más severas, y viceversa. A pesar de ello, se ha podido observar que la pérdida de dominios concretos de la proteína *CRX* regula los niveles de ARN y proteína. A su vez, se han identificado algunas variantes que, a pesar de no modificar los niveles de ARN, inducen una sobreexpresión de la proteína. Estos resultados sugieren que existe una nueva vía de regulación a nivel de la proteína que podría contribuir a la patogénesis relacionada con estas variantes. Hemos profundizado en la posibilidad de que la proteína pueda estar perdiendo sitios de ubiquitinación en la región C-terminal y que esto explique la acumulación de la proteína alterada.

**PALABRAS CLAVE:** *CRX*; dominante negativo; sobreexpresión génica; variantes *frameshift*; distrofias hereditarias de la retina; correlación fenotipo-genotipo.

## **ABSTRACT**

Hereditary retinal dystrophies are a group of retinal disorders associated with more than 280 genes, including the Cone-Rod Homeobox (*CRX*) gene, which encodes a transcription factor that is key in the terminal regulation of photoreceptors. Hence, defects in this gene lead to an alteration in these cells and this triggers different clinical entities such as Leber's congenital amaurosis, retinitis pigmentosa and cone and rod dystrophy. A dominant negative effect has been described in relation to variants that alter the reading pattern of the protein. The pathogenic mechanism has been attributed mainly to a deregulation at the RNA level, which has an impact on the overexpression of the protein. So far, no clinical correlation based on transcript and protein levels has been made for this type of variant. Therefore, in this work we have evaluated the effect of different frameshift variants in *CRX* causing the different clinical phenotypes mentioned above. The different variants produce very different expression profiles at both the RNA and protein levels, and high expression levels cannot be correlated with more severe clinical manifestations, and vice versa. However, loss of specific domains of the *CRX* protein has been shown to regulate RNA and protein levels. In turn, some variants have been identified that, despite not modifying RNA levels, induce an overexpression of the protein. These results suggest that there is a new regulatory pathway at the protein level that could contribute to the pathogenesis related to these variants. We have further explored the possibility that the protein may be missing ubiquitination sites in the C-terminal region and that this explains the accumulation of the altered protein.

**KEY WORDS:** *CRX*; dominant-negative; gene overexpression; frameshift variants; hereditary retinal dystrophies; phenotype-genotype correlation.

## **AGRADECIMIENTOS**

Al Hospital Universitario La Fe por el acceso a todas las instalaciones y necesidades requeridas. A Gema, Chema, Rafa, Pilar, Ángela y todos los miembros del Instituto de Investigación Sanitaria La Fe que han hecho esto posible. A Ana Pilar por su ayuda y su  
paciencia.

# ÍNDICE

<b>1. INTRODUCCIÓN</b> .....	<b>1</b>
1.1 Distrofias hereditarias de la retina .....	1
1.2 <i>CRX</i> .....	1
1.2.1 Estructura de la proteína .....	2
1.2.2 Tipos de mutaciones en <i>CRX</i> .....	3
1.3 Mecanismos de patogénesis asociados a variantes <i>frameshift</i> .....	4
1.3.1 Mecanismos de desregulación de ARN .....	4
1.3.2 Desregulación de los niveles de proteína .....	5
1.4 DHR asociadas al gen <i>CRX</i> .....	7
1.4.1 Amaurosis congénita de Leber .....	7
1.4.2 Retinosis pigmentaria .....	7
1.4.3 Distrofia de conos y bastones .....	7
<b>2. OBJETIVOS</b> .....	<b>8</b>
<b>3. MATERIALES Y MÉTODOS</b> .....	<b>8</b>
3.1 Mutagénesis dirigida .....	8
3.2 Transformación bacteriana y selección de clones.....	9
3.3 Secuenciación Sanger .....	9
3.4 Purificación de ácidos nucleicos .....	10
3.4.1 ADN plasmídico .....	10
3.4.2 ARN.....	10
3.5 RT-qPCR.....	10
3.6 Transfección celular.....	10
3.7 Citometría de flujo.....	11
3.8 Microscopía .....	11
3.9 Estadística.....	11
<b>4. RESULTADOS Y DISCUSIÓN</b> .....	<b>11</b>
4.1 Selección de variantes <i>frameshift</i> en <i>CRX</i> .....	11
4.2 Caracterización funcional de las variantes seleccionadas .....	14
4.2.1 Desregulación a nivel de ARN .....	16
4.2.2 Desregulación a nivel de proteína .....	16
4.2.3 Correlación entre la expresión de ARN y proteína.....	17

4.3 Características clínicas de las variantes.....	18
4.4 Correlación fenotipo-clínica.....	19
<b>5. CONCLUSIONES .....</b>	<b>20</b>
<b>6. BIBLIOGRAFÍA .....</b>	<b>21</b>
<b>7.ANEXOS .....</b>	<b>23</b>
7.1 ANEXO I. Relación del trabajo con los Objetivos de Desarrollo Sostenible de la agenda 2030.....	23

## ÍNDICE DE TABLAS

<b>Tabla 1.</b> Cebadores empleados en la mutagénesis .....	9
<b>Tabla 2.</b> Variantes de CRX incluidas en este estudio.....	13
<b>Tabla 3.</b> Casos clínicos de pacientes que poseen las variantes mutadas de CRX concretas, objeto de este estudio.....	19

## ÍNDICE DE FIGURAS

<b>Figura 1.</b> Diagrama de Venn muestra diferentes formas clínicas de DHR y sus genes causales.....	1
<b>Figura 2.</b> Dominios de CRX.....	2
<b>Figura 3.</b> Clases de mutaciones en la proteína CRX .....	3
<b>Figura 4.</b> Cambios posibles en la pauta en mutaciones de clase III en CRX .....	4
<b>Figura 5.</b> Ubiquitinación de CRX silvestre .....	6
<b>Figura 6.</b> Estructura del plásmido eGFP::CRX WT pcDNA3.1 .....	8
<b>Figura 7.</b> Mutaciones estudiadas en el gen CRX.....	12
<b>Figura 8.</b> Secuencia aminoacídica de la proteína CRX silvestre.....	12
<b>Figura 9.</b> Modelización de variantes de cambio de pauta en CRX .....	15

## NOMENCLATURA Y ABREVIATURAS

**DHR:** distrofias hereditarias de la retina

**CRX:** cone-rod homebox

**DCB:** distrofia de conos y bastones

**ACL:** amaurosis congénita de Leber

**RP:** retinitis pigmentaria

**WT:** silvestre (del inglés *wild-type*)

**eGFP:** proteína fluorescente verde mejorada (del inglés *enhanced green fluorescent protein*)

**CMVp:** promotor de citomegalovirus

**PCR:** reacción de cadena de la polimerasa

**RT-qPCR:** retrotranscripción y PCR cuantitativa en tiempo real

**ADN:** ácido desoxirribonucleico

**ADNc:** ADN copia

**ARN:** ácido ribonucleico

**ARNm:** ARN mensajero

**HEK:** células humanas embrionarias de riñón (del inglés *Human Embryonic Kidney*)

**LB:** Luria-Bertani

**GAPDH:** Gliceraldehído-3-fosfato deshidrogenasa

**PBS:** Tampón fosfato salino (del inglés *Phosphate Buffered Saline*)

**DMEM:** medio de cultivo Dulbecco's Modified Eagle Medium

**FBS:** suero bovino fetal

**ANOVA:** análisis de varianza

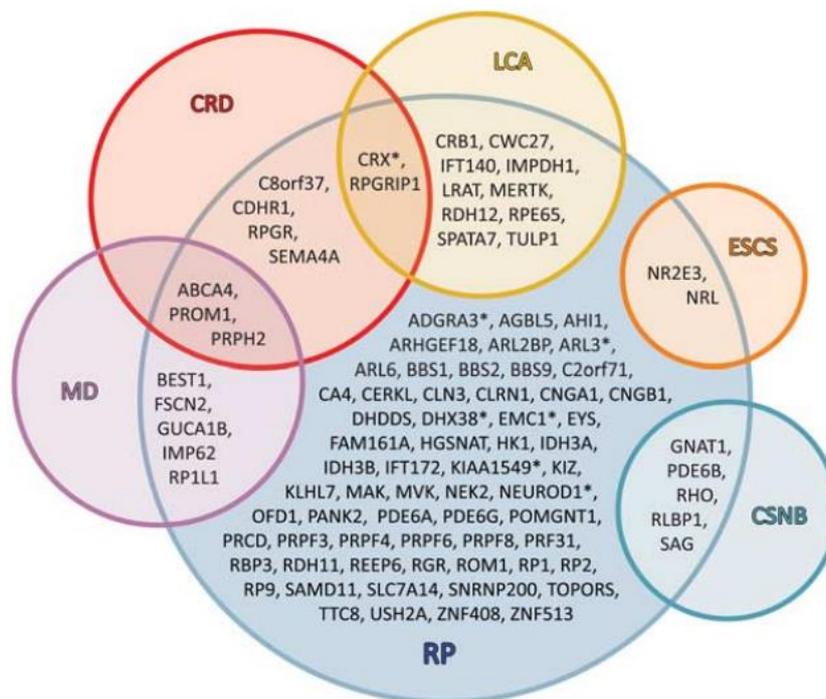
**aa:** aminoácido

# 1. INTRODUCCIÓN

## 1.1 Distrofias hereditarias de la retina

Las distrofias hereditarias de la retina (DHR) son un conjunto de desórdenes que producen degeneración en la retina de manera progresiva y que son genética y clínicamente heterogéneos. Poseen un amplio rango de severidad y variabilidad puesto que mutaciones en un mismo gen pueden provocar entidades clínicas distintas, y una misma entidad clínica puede estar ocasionada por mutaciones en diferentes genes (Figura 1) (Den Hollander et al., 2010).

Hasta la fecha, más de 280 genes han sido asociados con las DHR con diferentes patrones de herencia: principalmente autosómica recesiva (70%), autosómica dominante (20-25%) o ligada al cromosoma X (4%), pero también existen casos de herencia digénica y mitocondrial (Den Hollander et al., 2010).



**Figura 1.** Diagrama de Venn muestra diferentes formas clínicas de DHR y sus genes causales. Se observa como un mismo fenotipo puede ser ocasionado por genes diferentes y como hay genes que dan lugar a distintos fenotipos. Las siglas CRD indican Distrofia de Conos-bastones (del inglés Cone-Rod Dystrophies); LCA, amaurosis congénita de Leber (del inglés Leber congenital amaurosis); CSNB, ceguera nocturna congénita estacionaria (del inglés congenital stationary night blindness); ESCS, síndrome del incremento de conos o síndrome de Goldmann-Favre (del inglés Enhanced S-cone Syndrome); RP, (del inglés Retinitis Pigmentosa); MD, Distrofia Macular (Gregori et al., 2019)

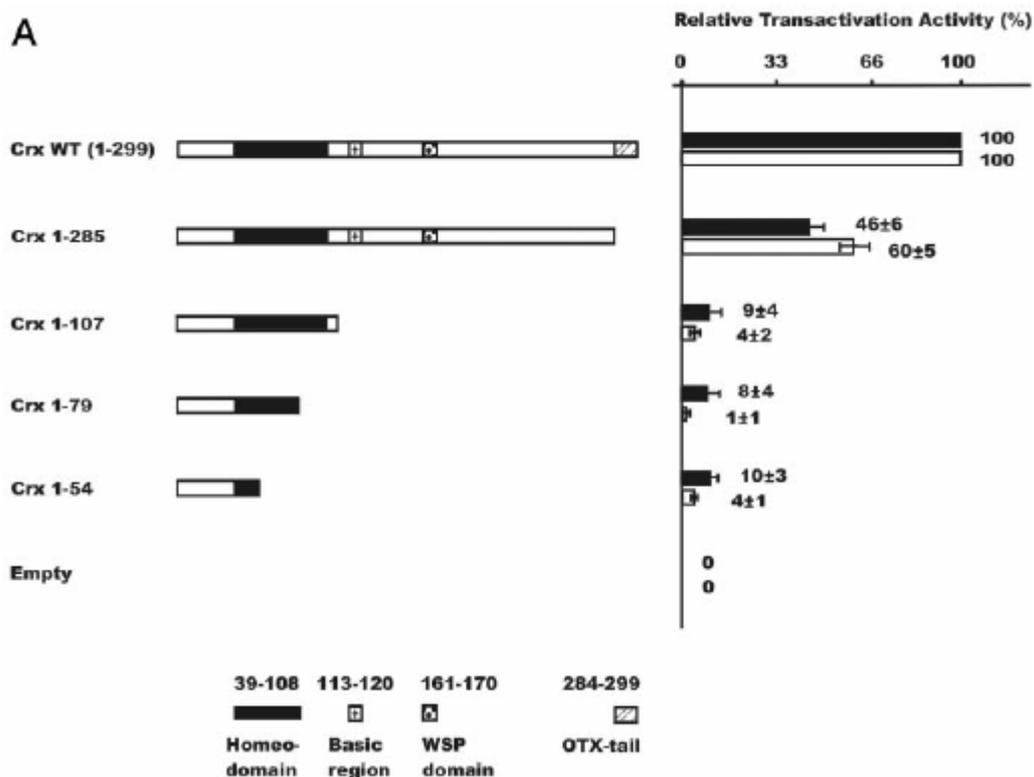
## 1.2 CRX

El gen *cone-rod homeobox* (CRX) codifica un factor de transcripción crucial para la expresión y diferenciación de genes relacionados con el desarrollo y el mantenimiento de los fotorreceptores, conos y bastones (Ruzycki et al., 2017). Los bastones se encuentran en mayor proporción que los conos, y principalmente están en la periferia y controlan la

visión en blanco y negro. Por el contrario, los conos, son esenciales para la visión en color y se localizan en la zona central.

### 1.2.1 Estructura de la proteína

El gen *CRX* se encuentra en el cromosoma 19q13.3. Consta de 4 exones (el primero es no codificante y los otros tres sí son codificantes (Freund et al., 1997)) que codifican una proteína de 299 residuos. La proteína está formada por un dominio de unión a ADN (se extiende desde el residuo 39 al 108), una región básica de activación (desde el residuo 113 al 120 aas), un dominio WSP (161-170 aas) y la cola OTX, localizada en el carboxil terminal (que se encuentra desde el residuo 284 al 299) asociada con la actividad de transactivación de la proteína. (Yi et al., 2019).



**Figura 2. Dominios de CRX.** En la parte izquierda se observan las variaciones de la proteína de las que se estudió el porcentaje de transactivación, así como su longitud y los dominios que poseen; y en la parte derecha se observa el porcentaje de transactivación relativa, siendo este menor cuanto más pequeña es la proteína, y, por tanto, cuando menos dominios conserva (Chau et al., 2000).

Tanto el dominio WSP como la cola OTX parecen ser importantes para el control de la actividad reguladora positiva de CRX. La cola OTX controla la actividad de transactivación de CRX, por lo que cuando esta se pierde, debido a mutaciones de cambio de pauta, esta actividad disminuye en un 50% (Figura 2). Esta disminución es cada vez mayor conforme más partes de la proteína se pierden, pasando de un 50% a un 7% con la pérdida del dominio WSP y la región básica (Chau et al., 2000).

Además, se cree que el dominio WSP sería responsable de la actividad de transactivación residual, así como de la transactivación de la rodopsina (proteína

pigmentaria sensible a la luz presente en las células fotorreceptoras de los bastones de la retina del ojo) (Chau et al., 2000).

### 1.2.2 Tipos de mutaciones en CRX

Las mutaciones descritas en CRX pueden ocurrir *de novo* o ser heredadas por patrones mayoritariamente autosómico dominante (Fokkema et al., 2011).

Principalmente han sido descritas dos tipos de mutaciones en CRX: cambio de aminoácido o deleciones/inserciones que alteran la pauta de lectura. Aunque, menos frecuente, el cambio de aminoácido puede codificar un codón de parada prematuro. Concretamente los cambios puntuales suelen producirse en el dominio de unión al ADN, mientras que el resto de las variantes han sido identificadas en la zona de transactivación de la proteína (Figura 3). En función del sitio de localización de la variante en cuestión se ha propuesto una clasificación que describe el efecto patogénico de las mutaciones (Figura 3). Esta clasificación incluye 4 clases de mutaciones: la clase I y la clase II se corresponden con mutaciones de cambio de aminoácido, mientras que las clases III y IV se corresponden con mutaciones causadas por deleciones o inserciones, generando un cambio en la pauta de lectura (Ruzycki et al., 2017). La clase I y II alteran la funcionalidad del dominio de unión a ADN, pero con efectos distintos. La clase I induce un efecto hipomórfico (que conlleva poca actividad génica o parcialmente reducida) mientras que la clase II produce un efecto antimórfico (que interfiere con la actividad genética, actuando como un dominante negativo). En este último caso la proteína puede unirse de forma variable al ADN, mientras que el otro caso, su unión es muy reducida (Figura 3). El efecto antimórfico es extendido para variantes frameshift de las clases III y IV, pero la única que tiene intacto el dominio de ADN es la Clase III, y es el tipo de mutaciones que hemos estudiado en este trabajo (Tran & Chen, 2014).

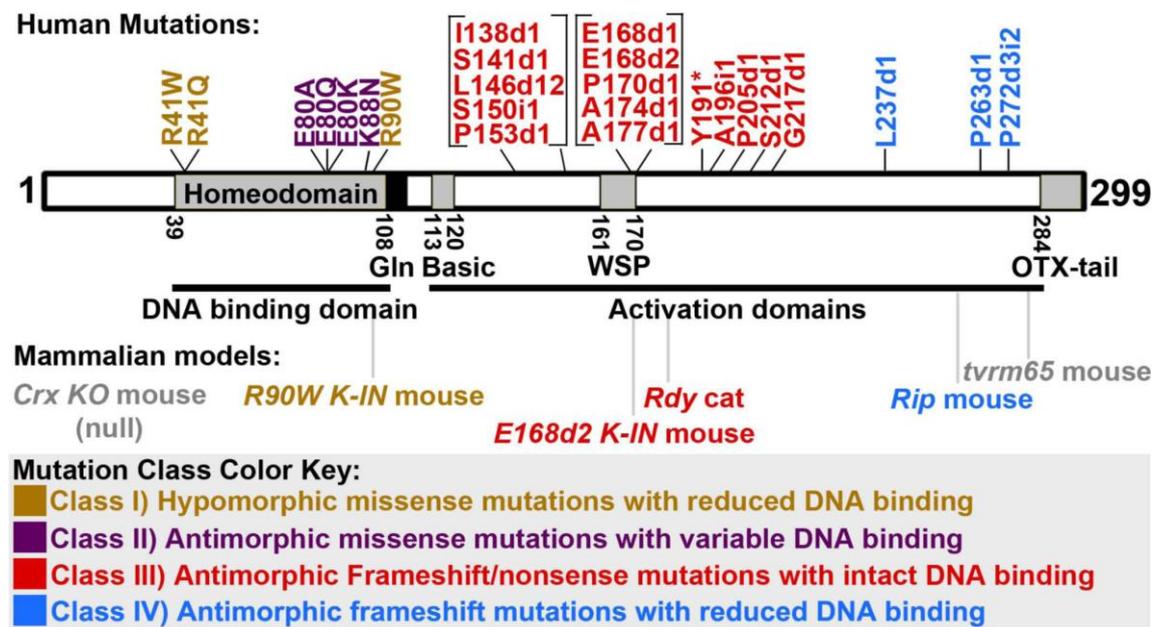
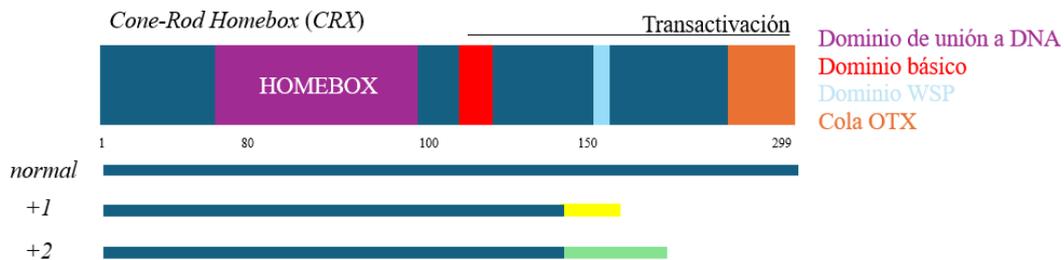


Figura 3. Clases de mutaciones en la proteína CRX. (Tran & Chen, 2014).

La clase III produce un cambio de pauta que finalmente incluye un codón de parada prematuro, produciendo una región nueva en C-terminal y una proteína de menor tamaño que la proteína silvestre. Sin embargo, no todas las variantes producen el mismo cambio de pauta (Figura 4), ya que depende de si la delección o la inserción contempla uno o más nucleótidos. Aun así, todas ellas han sido englobadas dentro de una misma clasificación. Es interesante reconsiderar este aspecto junto con el hecho de que hay variantes que afectan a unos dominios y no a otros.



**Figura 4. Cambios posibles en la pauta en mutaciones de clase III en CRX.** En azul está representada la proteína silvestre sin cambios en la pauta de lectura. En amarillo se indica el cambio de pauta +1 y en verde, el cambio de pauta +2. Ambos comienzan en la misma posición, pero tienen terminaciones diferentes, indicando que aunque una mutación se inicie en el mismo punto, puede dar lugar a proteínas diferentes.

### 1.3 Mecanismos de patogénesis asociados a variantes *frameshift*

La proteína disfuncional mutante CRX interfiere con la función de la proteína nativa CRX, demostrando su efecto dominante negativo (Tran et al., 2014).

#### 1.3.1 Mecanismos de desregulación de ARN

Han sido generados varios modelos de ratón para estudiar los efectos patogénicos de variantes *frameshift* en CRX (Tran et al., 2014). Estas variantes producen un efecto dominante negativo por el cual se desencadena una competición entre la proteína silvestre y la mutante por unirse al ADN, ya que ambas lo tienen intacto. El desplazamiento de la proteína silvestre por la unión de la proteína mutante.

Mientras que las mutaciones de cambio de pauta de lectura a menudo se han asociado con degradación del ARNm mutante a través de un proceso denominado decaimiento mediado por el antisentido, hay precedentes de que estas mutaciones puedan causar sobreexpresión de ARNm (Tran & Chen, 2014). Existen estudios anteriores con modelos animales (Tran & Chen, 2014) que demuestran que en ciertas mutaciones de clase III la transcripción se ve aumentada, así como la estabilidad del ARNm, sin embargo, se necesitaría una mayor investigación para determinar si esto sucede en todas las mutaciones de clase III.

Concretamente, el modelo de ratón que contiene la variante E168d2 (con una delección de los nucleótidos 1318 y 1319) presenta una sobreexpresión, pero no la variante R90W (con una sustitución nucleotídica en la posición 1085). La variante E168d2 tiene una sobreexpresión de ARNm (el doble que la proteína silvestre) que coincide con la

sobreexpresión en la proteína, mientras que la variante R90W tiene niveles normales de ARNm en contraste con los reducidos niveles de expresión proteica (Tran et al., 2014).

Muchos de los genes asociados a DHR testados en modelos animales de ratón se han visto regulados negativamente (disminuyendo su expresión), pero en el caso de los mutantes de *CRX* se ha observado tanto una regulación positiva como una negativa en la expresión génica (Ruzycki et al., 2015).

También se ha visto una correlación entre una mayor severidad fenotípica cuanto mayor es la expresión de ARN (Tran et al., 2014).

### **1.3.2 Desregulación de los niveles de proteína**

Al igual que en el caso de la desregulación de ARN, existen estudios anteriores en modelos animales que describen un aumento o disminución en los niveles de expresión proteica. Un ejemplo de ello es el estudio mencionado en el apartado anterior, centrado en las variantes E168d2 y R90W, en el que se observó un aumento de la expresión proteica en la primera variante, y una disminución de la misma en la segunda variante (Tran et al., 2014).

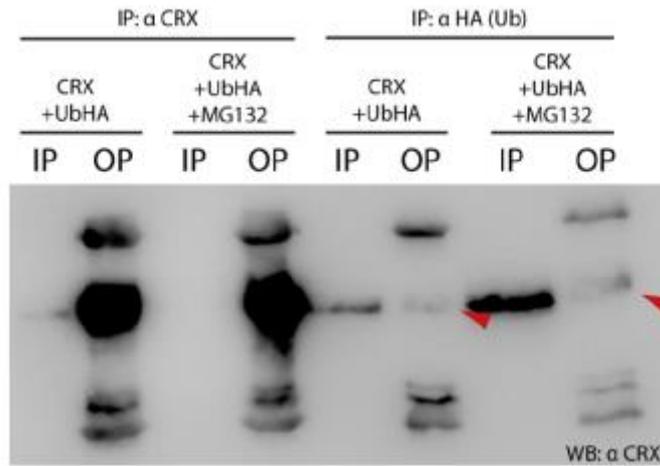
Hasta donde nosotros sabemos, no ha sido descrito en profundidad la implicación directa de la proteína CRX mutante en la patogénesis de las DHR. En base a esto, se sabe que los factores de transcripción deben estar finamente regulados, y no es extraño pensar que esto no ocurra con CRX. La célula presenta distintos mecanismos para regular los niveles de proteína mediante procesos de degradación proteica controlada y evitar así los efectos tóxicos de la acumulación de esta cuando es funcional, así como para eliminar la proteína cuando no sea funcional. No obstante, cuando existe una mutación en la proteína estos mecanismos de degradación se pueden ver afectados, provocando una acumulación de la proteína que puede resultar tóxica.

Los mecanismos mayormente conocidos de degradación endógena celular son la ubiquitinación o la autofagia, aunque también se ha descrito una regulación mediada por la actuación de degrones (proteínas portadoras de señales de degradación).

#### **1.3.2.1 Ubiquitinación**

La ubiquitinación es una modificación post-traducciona común basada en la unión covalente de la ubiquitina (proteína pequeña de 76 aminoácidos, altamente conservada), que genera una señal para la degradación de proteínas mal plegadas y el control de la calidad proteica. Está mediada por una cascada enzimática: E1 (enzima activadora), E2 (enzima para la conjugación) y E3 (enzima ubiquitina ligasa). La unión isopeptídica se da habitualmente entre la glicina del extremo C-terminal de la ubiquitina y el grupo  $\epsilon$ -amino de un residuo de lisina (De Silva & Page, 2023).

En un estudio previo, se ha evaluado por co-inmunoprecipitación que la proteína CRX silvestre puede ser ubiquitinada (Esquerdo Barragan, M., tesis doctoral, 2017). En este trabajo someten las células transfectadas con CRX silvestre a un tratamiento con MG132, observando una reducción del conjugado CRX-Ub cuando se bloquea el proteasoma por MG132 comparado con el lisado de partida entre controles y tratados (Figura 5).



**Figura 5. Ubiquitinación de CRX silvestre.** Resultados obtenidos en la detección por coimmunoprecipitación simultánea de CRX y UbHA. Se incluye el lisado input (IP) y el eluido final output (OP). Las células fueron transfectadas con CRX silvestre y sometidas a un tratamiento con un inhibidor específico de proteasoma MG132. La detección de inmunoblot se realizó con un anticuerpo primario anti-CRX (Esquedo Barragan, M., tesis doctoral, 2017).

Esta es la única evidencia que se tiene sobre el control de los niveles de expresión proteica en CRX silvestre. No obstante, esta evidencia ha sido realizada solo sobre la proteína silvestre, por lo que sería conveniente estudiar este control de la expresión en diferentes mutaciones para estudiar la desregulación de la proteína de manera independiente al ARNm. Aun así, son evidencias que abren la posibilidad de que exista una nueva vía a nivel de proteína para variantes *frameshift* en CRX que pueda explicar mejor su patogénesis.

### 1.3.2.2 Autofagia

La autofagia es un proceso de autodegradación importante para: equilibrar las fuentes de energía, eliminar proteínas mal plegadas o agregados de proteínas, eliminar orgánulos dañados y eliminar patógenos intracelulares. Existen 3 tipos de autofagia: macroautofagia (la carga citoplasmática es entregada al lisosoma mediante una vesícula intermediaria), microautofagia (los componentes citosólicos son tomados directamente por el lisosoma mediante invaginación) y autofagia mediada por chaperonas (las proteínas diana son translocadas a través de la membrana lisosomal en un complejo con chaperonas) (Glick et al., 2010). Hasta el momento no se han encontrado evidencias significativas en la literatura previa que relacione la desregulación de los niveles de la proteína CRX por medio de este mecanismo. Eso no evita que la autofagia pudiera influir en la regulación de CRX y que sería interesante explorar.

### 1.3.2.3 Degrones

Se han descrito terminaciones proteicas determinantes en la degradación de proteínas. Estas terminaciones pueden encontrarse tanto en N-terminal como en C-terminal. Los C-degrones, localizados en C-terminal, son motivos pequeños conservados, que contienen habitualmente glicina como último residuo. Algunos de los motivos conservados descritos son: GG, RXG y XRXG (Yeh et al., 2021). Sería interesante

evaluar la presencia de estos dominios en CRX silvestre y mutante para profundizar en la implicación de este posible mecanismo.

#### **1.4 DHR asociadas al gen *CRX***

En este estudio nos centraremos en aquellos fenotipos relacionados con *CRX*: amaurosis congénita de Leber (de ahora en adelante ACL), la retinosis pigmentaria (RP) y distrofia de conos y bastones (desde ahora DCB).

De entre estas tres diferentes patologías la severidad clínica es mayor en aquellos que presentan ACL ya que es congénita, seguido de quienes presentan RP y DCB, las cuáles pueden aparecer en edades más tardías. Por tanto, afecta a la severidad la edad de inicio de la enfermedad, siendo la gravedad mayor cuando menor es la edad en la que el paciente desarrolla la enfermedad.

##### **1.4.1 Amaurosis congénita de Leber**

La ACL es una de las DHR más tempranas y severas, al ser congénita. Está caracterizada por deficiencia visual grave en la infancia, nistagmo (movimientos oculares involuntarios), respuestas pupilares ausentes o lentas, e hipermetropía elevada (Han & Rim et al., 2017). Tiene una prevalencia de entre 1/81.000 a 1/30.000, representando un 5% de las DHR (Huang et al., 2021).

Comparada con el resto de las distrofias de la retina. ACL es más heterogénea en su transcurso y presentación. Normalmente, los pacientes presentan visión pobre o ausencia de fijación a las seis semanas de vida (Huang et al., 2021).

La frecuencia en la que ACL es presentado por una mutación en el gen *CRX* (dominante autosómica) es del 1%, con una maculopatía central atrófica y una retina más fina con la capa nuclear externa conservada (Huang et al., 2021).

##### **1.4.2 Retinosis pigmentaria**

La RP se caracteriza por la atrofia del pigmento retinal epitelial y la progresiva muerte celular de los fotorreceptores, provocando primero ambliopía nocturna y posterior ceguera, con una prevalencia de entre 1/7.000 a 1/3.000 (Liu et al., 2022).

En pacientes con un diagnóstico más temprano la enfermedad progresa con rapidez, (alrededor de los 10 años). No obstante, aquellos con un diagnóstico más tardío suelen presentar solo anormalidades oculares. La información genética familiar y las características fenotípicas concretas son cruciales para una intervención efectiva (Liu et al., 2022).

##### **1.4.3 Distrofia de conos y bastones**

La DCB, con una prevalencia de entre 1/30.000 a 1/40.000, caracterizados por la pérdida de 1 tipo de fotorreceptor concreto – conos o bastones – con la subsecuente pérdida de otros tipos celulares (García-Llorca et al., 2023).

Los síntomas principales son foto aversión, reducción de la agudez visual, reducción en la sensibilidad del campo visual central y defectos en la visión de los colores.

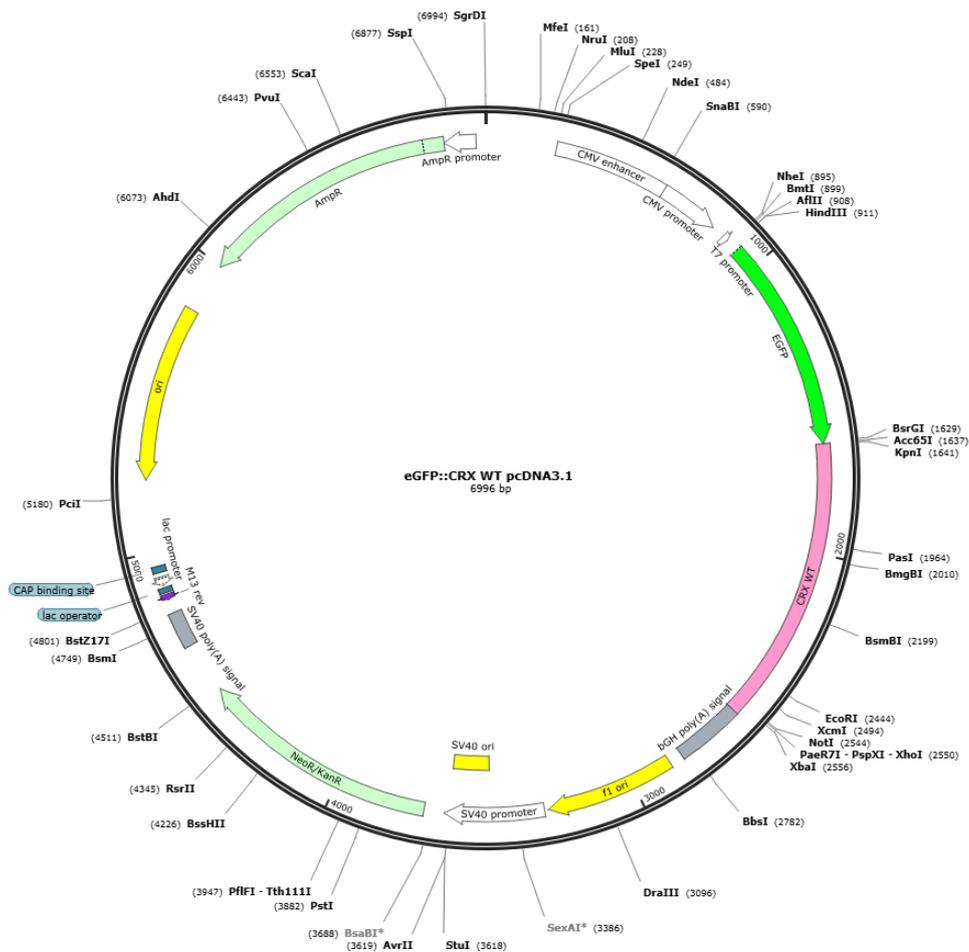
Generalmente, la distrofia en conos tiene mayor y más rápido transcurso clínico que la distrofia de bastones (García-Llorca et al., 2023).

## 2. OBJETIVOS

- Modelizar la expresión de variantes *frameshift* en *CRX* en HEK293.
- Evaluar cambios de expresión a nivel de transcrito y proteína de las variantes modelizadas.
- Correlacional la desregulación asociada a las variantes con el fenotipo observado en los pacientes.

## 3. MATERIALES Y MÉTODOS

### 3.1 Mutagénesis dirigida



**Figura 6. Estructura del plásmido eGFP::CRX WT pcDNA3.1.** El plásmido posee un promotor de Citomegalovirus (CMV) que dirige la expresión de la proteína fluorescente eGFP en N-terminal y en pauta con el gen silvestre CRX. El vector contiene el gen que otorga resistencia al antibiótico ampicilina para la selección positiva de las bacterias positivas.

Con el fin de obtener los plásmidos que porten las mutaciones deseadas, se realizó una mutagénesis empleando el plásmido eGFP::CRX WT pcDNA3.1, el cuál estaba disponible en el laboratorio (Figura 5). Los cebadores fueron diseñados con el programa QuickChange Primer Design para cada una de las variantes (Tabla 1). La reacción de mutagénesis se realizó siguiendo las instrucciones de fabricante (QuickChange II Site-Directed Mutagenesis Kit de Agilent Technologies).

El programa de PCR fue el siguiente: 1 ciclo de 95°C/2 min, x18 ciclos 95°C/20 sec (desnaturalización), 68°C/10 sec (hibridación), 68°C/7 min (polimerización) y un paso final de extensión 68°C/5 min.

**Tabla 1. Cebadores empleados en la mutagénesis.**

MUTACIÓN	PRIMER DIRECTO (5'-3')	PRIMER REVERSO (5'-3')
p.G122Afs*65	GTC TTG GGG ACG TGC CGC CTT CCT CTT G	CAA GAG GAA GGC GGC ACG TCC CCA AAG AC
p.P145Lfs*42	CAG ATT CCT ACA GCC CCC TCT GCC CG	CGG GCA GGG GCT GTA GGA ATC TG
p.G148Afs*39	GAG CCT GAG GGG GCA GAG GGG	CCC CTC TGC CCG CCC CTC AGG CTC
p.P172Lfs*15	CGC TGC GCC TCA GCA AAG GGG ACT CT	AGA GCT CCC TTT GCT GAG GCG CAG CG
p.V180Gfs*56	CCC TGA GGC CAC CCA GCC CAG CCC G	CGG GCT GGG CTG GGT GGC CTC AGG G
p.P197Rfs*22	ATG ACC TAC GCC CGG CCT CCG CTT TC	GAA AGC GGA GGC CGG GCG TAG GTC AT
p.Y221Tfs*9	GCG GCC TAG ACC CCA CCT TTC TCC CAT G	CAT GGG AGA AAG GTG GGG TCT AGG CCG C

### 3.2 Transformación bacteriana y selección de clones

Tras la mutagénesis dirigida, se realizó un proceso de digestión, para evitar restos del plásmido anterior, mediante DPN I, incubándolo durante 1 hora a 37°C. Una vez terminado este proceso, se realizó la transformación bacteriana mediante electroporación. La cepa bacteriana empleada para la electroporación de los plásmidos fue *Escherichia Coli*, específicamente las variantes electrocompetentes *Top10* y *NepStable*.

Se sembraron las bacterias transformadas en agar sólido (con carbamicina, equivalente a la ampicilina y más estable) y se cultivaron 24 horas a 37°C. Pasadas las 24 horas, se seleccionaron diversas colonias por variante (mínimo 5) y se cultivaron en medio LB líquido (con carbamicina) a 37°C durante 24 horas para la posterior extracción de ADN plasmídico.

### 3.3 Secuenciación Sanger

Tras el proceso de mutagénesis, se analizaron los clones obtenidos (entre 5 y 6 distintos) mediante secuenciación Sanger (empresa STABVIDA). Para llevar a cabo este paso, se empleó un primer directo (5'- GTC AAC GCC TTG GCC CTA AGT GG – 3') localizado en el exón 2 del gen *CRX* que permitía la secuenciación de la región donde estaban localizadas las mutaciones de interés.

La confirmación de la variante se hizo empleando el programa de libre acceso FinchTV.

### **3.4 Purificación de ácidos nucleicos**

#### **3.4.1 ADN plasmídico**

Para la obtención del ADN plasmídico, se partió de un cultivo de LB líquido, especificado en el punto 3.2.

En cuanto al proceso de extracción de este ADN plasmídico, se realizaron 2 abordajes distintos dependiendo del objetivo. Para la validación de los clones se empleó el kit comercial Miniprep (NZYMiniprep, Nzytech), que aseguraba la obtención de plásmido rápido a partir de un cultivo de 3mL. Para la obtención de preparaciones de alta pureza, requeridas para transfecciones, se usó el kit de HiSpeed Maxiprep (Qiagen) a partir del cual se procesaron cultivos en torno a 150-299 mL. El ADN extraído se cuantificó por absorbancia en el equipo Nanodrop 2000 Spectrophotometer (Thermo Scientific).

#### **3.4.2 ARN**

Para la obtención de ARN se recogieron  $5 \cdot 10^5$  células HEK293 transfectadas por réplica experimental. El medio en el que fueron recolectadas contaba de 300 $\mu$ L de buffer de lisis y 3,5 $\mu$ L de  $\beta$ -mercaptoetanol. Posteriormente, se congelaron las muestras a -80°C para facilitar la lisis.

Para la extracción de ARN total, las muestras fueron procesadas siguiendo las instrucciones del fabricante (NZY Total RNA Isolation Kit de NZYtech). Las muestras fueron sometidas a un tratamiento con DNAsas para eliminar el ADN genómico que pudiera interferir con los resultados de expresión. El ARN extraído se cuantificó por absorbancia en el equipo nanodrop 2000 Spectrophotometer (Thermo Scientific).

### **3.5 RT-qPCR**

El ARN total extraído fue empleado para la obtención de ADNc mediante el protocolo del fabricante del PrimeScript™ RT Reagent Kit (TaKaRa). La reacción de retrotranscripción fue realizada a partir de 500 ng de ARN y el programa de PCR de 20 minutos de RT-Takara. El ADNc obtenido se diluyó 1:5 antes de ser empleado en la reacción de qPCR. La reacción de PCR contiene 14 $\mu$ L de mix (7,5  $\mu$ L de tampón, 5,75  $\mu$ L de agua libre de RNAsas y 0,75  $\mu$ L de sonda) (TaKaRa, Sumilab) y 1 $\mu$ L de la muestra diluida. Se emplearon 2 sondas: una correspondiente al gen diana *CRX* y otra del gen *GAPDH* como control endógeno de referencia para la normalización de la expresión relativa.

El programa de qPCR fue el siguiente: 1 ciclo de 10 minutos a 50°C, seguido de 1 ciclo de desnaturalización a 95°C durante 15 segundos, y 40 ciclos de polimerización a 60°C. Se analizaron 3 réplicas técnicas por cada muestra para cada gen y se obtuvo el promedio. El equipo utilizado de RT-PCR fue ViiA™ 7 (Applied Biosystems).

### **3.6 Transfección celular**

Para la modelización de las variantes se empleó la línea celular HEK 293 (*Human Embryonic Kidney*), cultivadas a 37 °C con un 5% de CO<sub>2</sub> con el medio DMEM High

Glucose (Dulbecco's Modified Eagle Medium), con un 10% de FBS y 1% de antibiótico penicilina/estreptomicina.

Para el mantenimiento de la línea, se realizaron pases celulares al 80% de confluencia: primero, se realiza un lavado con PBS; después, se levantan las células con 5mL de tripsina, cultivando por 5 minutos en el incubador; transcurrido el tiempo, se detiene la acción de la tripsina con 10 mL de medio y se recogen las células para centrifugar 5 minutos a 300g; posteriormente, se resuspende el pellet con 10 mL de medio nuevo y se cultivan en un nuevo flask con dilución 1:10, 1:15 o 1:20, dependiendo de la confluencia inicial o el tiempo a transcurrir hasta el siguiente pase. Además, cuando la confluencia no era lo suficientemente alta, se ejecutaron cambios de medio para el mantenimiento.

Para la transfección celular se siguieron las indicaciones de Lipofectamine 3000 Transfection Kit de Invitrogen. El diseño experimental conllevó la siembra de 150.000 células por pocillo P6 el día 1, seguido de la transfección de 1 µg de ADN el día 2, y la recogida de muestra y/o análisis 24 horas después de la transfección.

### **3.7 Citometría de flujo**

Para el procesamiento de las muestras, se levantaron con tripsina y se recogieron con 1mL de PBS.

Las muestras fueron procesadas mediante un citómetro BD FACSCanto™ II, con el software FACS DIVA. Este citómetro facilitaba el conteo de células GFP+ (o transfectadas) y la obtención de intensidad de fluorescencia, la cual es directamente cuantificable e indica los niveles de expresión de proteína.

### **3.8 Microscopía**

La confluencia y el recuento celular se controló mediante microscopio invertido de contraste de fases (OLYMPUS CKX31). Por otro lado, se tomaron imágenes de las células frescas directamente en las placas de cultivo mediante la lupa Leica *Cell Imager* (PAULA).

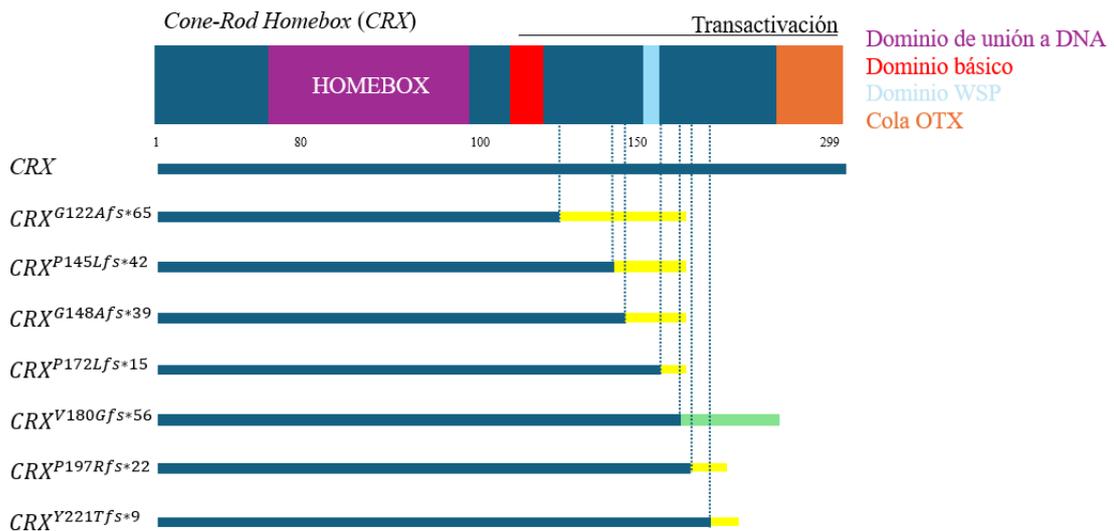
### **3.9 Estadística**

Todos los análisis se realizaron usando datos de al menos tres réplicas biológicas independientes de cada muestra. Los análisis estadísticos se llevaron a cabo mediante comparaciones múltiples de tipo ANOVA. \*:p-valor>0,05; \*\*:p-valor>0,01; \*\*\*: p-valor>0,001; ns; no significancia estadística.

## **4. RESULTADOS Y DISCUSIÓN**

### **4.1 Selección de variantes *frameshift* en CRX**

Para este estudio se han seleccionado 7 variantes diferentes de mutaciones *frameshift* de la clase III (Figura 3). Tres de ellas conllevan la pérdida del dominio WSP y la cola OTX, y las otras cuatro solo conllevan la pérdida de la cola OTX (el dominio WSP se mantiene intacto) (Figura 7).



**Figura 7. Mutaciones estudiadas en el gen CRX.** Todas las mutaciones que observamos son del tipo III, en las que ocurre un cambio de pauta de lectura, mayoritariamente por delección del nucleótido que se indica, excepto en el caso de la quinta mutación indicada (V180Gfs\*56) en el que la pauta de lectura es ocasionada por una duplicación. En azul oscuro se indica la parte común que presentan todas las proteínas. En amarillo se indica la secuencia aberrante formada a partir del cambio en la pauta de lectura por delección (cambio de pauta +1), mientras que en verde se indica la secuencia aberrante formada por duplicación (cambio de pauta +2).

La selección de las variantes se ha realizado en base a los datos clínicos disponibles, el tipo de entidad clínica y la variabilidad en cuanto a la pauta de lectura y la región nueva incluida en C-terminal. En concreto, la variante p.G122Afs\*65 ha sido identificada en el laboratorio donde está enmarcado este TFG, por lo que se incluyó como variante de estudio a comparar con el resto de variantes seleccionadas (Figura 7).

Las variantes que producen cambios de pauta de lectura obtenidas mediante el proceso de mutagénesis sobre el plásmido que contiene la copia silvestre de *CRX* se encuentran incluidas en la Tabla 2. Todas las construcciones contienen una región común con *CRX* silvestre (Figura 8). Esta zona común engloba el dominio de unión a ADN (39-108 aminoácidos), y la región básica (113-120 aminoácidos).

#### CRX WT

```

1  MMAYMNP GPH YSVNALALSG PSVDLMHQAV PYPSAPRKQR RERTTFTRSQ LEELEALFAK
61  TQYPD VYARE EVALKINLPE SRVQVWFKNR RAKCRQORQQ QKQQQPPPG QAKRPAKRR
121 AGTSPRPSTD VCPDPLGISD SYSPFLPGPS GSPTTAVATV SIWSPASESE LBEAQRAGLV
181 ASGPSLTSAP YAMTYAPASA FCSSPSAYGS PSSYFSGLDP YLSPMV PQLG GPALSPLSGP
241 SVGPSLAQSP TSLSGQSYGA YSPVDSLEFK DPTGTWKFTY NPM DPLDYKD QSAWKFOIL*
```

DOMINIO DE UNIÓN A DNA (39-108 aa)

DOMINIO BÁSICO (113-120 aa)

DOMINIO WSP (161-170 aa)

COLA OTX (284-299 aa)

POSICIONES DE LAS DISTINTAS MUTACIONES (122, 145, 148, 172, 180, 197, 221 aa)

**Figura 8. Secuencia aminoacídica de la proteína CRX silvestre.** En negro se indica la secuencia común a todas las variantes. En rojo se indica la zona que varía según la variante. Subrayado en amarillo se encuentra el dominio de unión a ADN (del aminoácido 39 al 108), el dominio WSP (161-170 aa) se indica subrayado en azul, y la cota OTX (284-299 aa) se encuentra subrayada en rosa. La posición de cada una de las mutaciones se indica subrayada en gris (122, 145, 148, 172, 180, 197 y 221 aa).

**Tabla 2. Variantes de CRX incluidas en este estudio.**

VARIANTE	ADNc	SECUENCIA ABERRANTE
<b>Frame +1</b>		<b>Epítomos compartidos</b>
CRX <sup>G122Afs*65</sup>	c.365del	ARPQDPPEQMCVQTLWASQIPTVPLCPAPQAPQPRQWPLCPGSAQPQSPCLRRSGLGWWPQGRL*
CRX <sup>P145Lfs*42</sup>	c.429del	LCPAPQAPQPRQWPLCPGSAQPQSPCLRRSGLGWWPQGRL*
CRX <sup>G148Afs*39</sup>	c.443del	APQAPQPRQWPLCPGSAQPQSPCLRRSGLGWWPQGRL*
CRX <sup>P172Lfs*15</sup>	c.514del	LRRSGLGWWPQGRL*
<b>Frame +1</b>		<b>Epítomos no compartidos</b>
CRX <sup>P197Rfs*22</sup>	c.587del	RPPLSALPPPPMGLRAPISAA*
CRX <sup>Y221Tfs*9</sup>	c.661del	TFLPWCPS*
<b>Frame +2</b>		
CRX <sup>V180Gfs*56</sup>	c.538dup	GGLRAVSDLRPLCHDLRPLRFLFLPLRLVWSELQFQRPRPLPFSGAPARGPGS*

(\*) En la primera columna se indica el cambio a nivel de proteína. En la segunda columna se indica el cambio a nivel de ADNc. En la tercera columna se indican las secuencias que difieren con la proteína nativa en negro y en naranja la que es común entre las mutaciones seleccionadas, y a su vez, difieren de la silvestre.

Las siete mutaciones seleccionadas son de clase III y producen un cambio en la pauta de lectura tras el dominio de unión de ADN. Las tres primeras conllevan pérdida del dominio WSP y de la cola OTX, mientras que las cuatro últimas no pierden el dominio WSP (Tabla 2). En todos los casos se produce una delección de 1 pb, produciendo un cambio de pauta (+1), excepto para la mutación p.V180Gfs\*56 ((cambio de Valina a Glutamina con cambio de pauta que adiciona 56 aminoácidos nuevos seguidos de un codón de parada prematuro), ocasionada por duplicación del nucleótido de la posición 538 del ADNc, se produce un cambio en la pauta de lectura distinto (+2). En todos los casos se produce una proteína truncada con un codón de parada prematura, que en algunos casos es común entre ellas (ver Tabla 2). Esto ocasiona que las cuatro primeras mutaciones generen epítomos con regiones en C-terminal comunes.

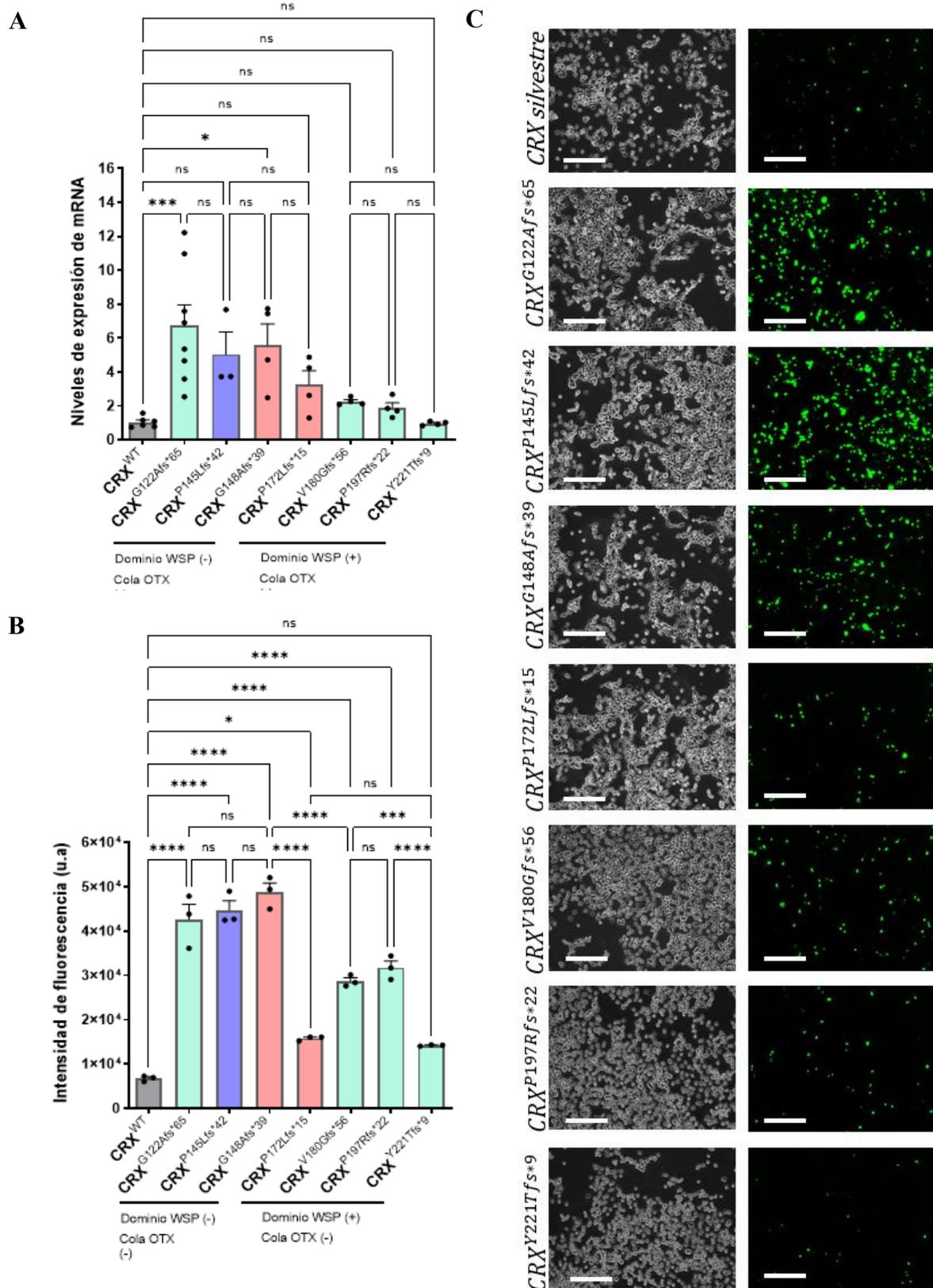
Las mutaciones en las posiciones de la proteína 122, 145 y 148 pierden el dominio WSP y la cola OTX, al producirse la mutación antes del dominio WSP (Figura 7). Esto ocasiona que todas ellas den lugar a una proteína del mismo tamaño (187 aminoácidos) ya que el codón de parada prematuro se encuentra en la misma posición y, a su vez, con un fragmento común para todas en la parte C-terminal (Tabla 2 y Figura 7). Adicionalmente, la variante p.P172Lfs\*15 se encuentra incluida en este grupo de variantes, que poseen regiones comunes en C-terminal, con la diferencia de que posee intacto el dominio WSP.

Por otro lado, las mutaciones p.P197Rfs\*22 y p.Y221Tfs\*9, además de tener intacto el dominio WSP, poseen una región en C-terminal totalmente distinta a las anteriores, a pesar de producirse por un cambio +1, ya que se localizan en una posición más tardía. En cuanto a estas dos variantes, no hay zona solapante entre ambas, ya que la primera posee una región más corta que no solapa con la segunda variante. Esto ocasiona que se generen proteínas de distintos tamaños, 219 aminoácidos y 230 aminoácidos respectivamente.

Por último, la variante p.V180Gfs\*56 debida a la duplicación del nucleótido 538 (guanina), genera un cambio en la pauta de lectura de +2, por lo que esta secuencia no comparte ninguna parte común con ninguna de la secuencias anteriores. La variante p.V180Gfs\*56 genera una proteína de 236 aminoácidos y posee intacto el dominio de unión a ADN, el dominio básico y el dominio WSP, y pierde la cola OTX.

#### **4.2 Caracterización funcional de las variantes seleccionadas**

Las variantes en pauta con eGFP fueron expresadas de manera transitoria en la línea celular HEK293, la cual no expresa *CRX* endógenamente. La caracterización funcional se ha realizado en base a los niveles de transcrito y proteína para cada una de las mutaciones y haciendo una comparativa con la proteína silvestre (Figura 9).



**Figura 9. Modelización de variantes de cambio de pauta en CRX.** A) Niveles de expresión de CRX en cada variante estudiada con una  $N_{\text{mínima}}=3$  réplicas biológicas. *GAPDH* ha sido empleado como control endógeno. Niveles normalizados con respecto a CRX silvestre. B) Niveles de expresión de la proteína CRX en cada variante estudiada con una  $N_{\text{mínima}}=3$  réplicas biológicas. Análisis comparativo múltiple de tipo ANOVA. \*:p-valor>0,05; \*\*:p-valor>0,01; \*\*\*: p-valor>0,001; ns; no significancia estadística.

Gris = sano, azul claro = DCB, morado = RP, rosa = ACL. C) Fotografías de las células transfectadas tres 24h no fijadas obtenidas mediante la lupa Leica Cell Imager (PAULA). Barra de escala: 200  $\mu$ m.

#### 4.2.1 Desregulación a nivel de ARN

La expresión transitoria de las variantes en células HEK293 nos muestra claras diferencias en cuando a los niveles de transcrito CRX (Figura 9A). A modo general, podemos observar una tendencia de incremento, de derecha a izquierda, acorde a lo que se esperaba, según ha sido indicado previamente en la introducción, ya que existen estudios anteriores que defendían una mayor expresión de ARN cuando la mutación es más temprana (Tran et al., 2014). En torno a esto, cabe resaltar que las variantes p.G122Afs\*65, p.P145L\*42 y p.G148Afs\*39 producen cambios significativos a nivel del ARN con respecto a los niveles observados para CRX silvestre (Figura 9A). Sin embargo, observamos casos donde las variantes p.P172Lfs\*15, p.V180G\*56, p.P197R\*22 y p.Y221T\*9, a pesar se cumplir esta tendencia, no producen cambios significativos en cuanto a los niveles de ARN con respecto a la secuencia silvestre (Figura 9A).

Hasta donde se sabe, la modelización de algunas mutaciones de este tipo siempre ha producido una alteración de la expresión a nivel de ARN en varios modelos de ratón (Ruzycki et al., 2017); sin embargo, en este estudio observamos que hay variantes, dentro del grupo de variantes Clase III descritas en pacientes, que no producen cambios significativos a nivel de ARN.

Además, se podría pensar que cuando mayor es el fragmento aberrante resultante por la mutación, esto implicaría una mayor desregulación en la expresión de ARN; sin embargo, si observamos las variantes p.G122Afs\*65 y p.V180Gfs\*56, que poseen fragmentos aberrantes de similar tamaño, vemos que no coinciden sus niveles de expresión de ARN (Figura 9A). En este caso hay que tener en cuenta los dominios intactos en ambas proteínas, donde la primera carece de WSP pudiendo explicar este comportamiento.

#### 4.2.2 Desregulación a nivel de proteína

En cuanto a la expresión proteica no observamos una correlación directa en base a la posición de la variante como ocurría con el ARN. En este caso, se observa dos grupos claros en base a la pérdida del dominio WSP (Figura 9B):

Las mutaciones que conllevan a la pérdida del dominio son p.G122Afs\*65, p.P145L\*42 y p.G148Afs\*39, y, en estos casos, se observa un aumento significativo de la expresión de la proteína (Figura 9B). Los niveles de expresión de proteína son inferidos a partir de los niveles de fluorescencia de la proteína de fusión con eGFP obtenidos mediante citometría de flujo.

Por otro lado, el resto de las mutaciones mantienen el dominio WSP intacto (Figura 9B). Asimismo, dentro de este grupo, debemos hacer una distinción entre aquellas en las que se observa un aumento significativo de los niveles de proteína, los cuales son atribuidos a p.P172Lfs\*15, p.V180G\*56 y p.P197R\*22, y, por otra parte, p.Y221T\*9,

que no parece presentar diferencias significativas con respecto a la proteína silvestre (Figura 9B). A su vez, cabe destacar de nuevo las diferencias, en este caso significativas, entre las dos variantes que producen ACL observándose una mayor expresión de proteína en ausencia del dominio WSP (Barras en rosa; Figura 9B).

Se han tomado también imágenes representativas donde se observa la localización de la proteína en el núcleo puesto que se trata de un factor de transcripción (sin tinción DAPI, pero resultados en el grupo lo han confirmado). La presencia de variantes de este tipo no cambia la localización de la proteína sino su expresión. Además, mediante estas imágenes se puede hacer una apreciación cualitativa sobre las variantes que presentan una menor o mayor expresión, atendiendo a la intensidad de fluorescencia, observando el incremento en la intensidad sobre todo en las variantes p.G122Afs\*65, p.P145L\*42 y p.G148Afs\*39 (Figura 9C).

#### **4.2.3 Correlación entre la expresión de ARN y proteína**

Si comparamos los niveles de expresión de ARN y proteína observamos claras diferencias por variante (Figuras 9A, B). En base a esto, podemos destacar aquellas variantes p.G122Afs\*65, p.P145Lfs\*42 y p.G148Afs\*39 donde existe una desregulación significativa tanto de los niveles de ARN como de la expresión de la proteína (Figura 9). Ha sido atribuida esta desregulación principalmente al ARN, ya sea por una mayor estabilidad de los transcritos mutantes con respecto al transcrito silvestre o por una posible supresión de regulación por retroalimentación negativa en los mutantes (Tran et al., 2014). A pesar de ello, no podemos descartar que el efecto observado en la proteína sea exclusivamente debido a una desregulación del ARN, ya que también podría ser un efecto aditivo directamente de la proteína, lo que explicaremos en mayor detalle más adelante.

Por otro lado, las variantes cuya expresión proteica difiere de lo observado a nivel de ARN: p.V180Gfs\*56 y p.P197Rfs\*22, cuyos niveles de expresión a nivel de ARN no varían significativamente con los observados para el transcrito silvestre. Sin embargo, a diferencia del supuesto anterior, los niveles de proteína están aumentados significativamente (Figura 9A, B). La variante p.Y221Tfs\*9 parece comportarse similar a estas variantes pero no produce cambios significativos en los niveles de ARN y proteína, aunque sí que se observa una tendencia en aumento a nivel de proteína, pero sería preciso aumentar el número de réplicas para confirmarlo. La variante p.P172Lfs\*15 también muestra disonancias con respecto a las variantes p.V180Gfs\*56 y p.P197Rfs\*22, observándose menores niveles de expresión de proteína..

Estos resultados sugieren que la patogenicidad de estas variantes podría estar explicada por una alteración a nivel de expresión proteica, y podría estar directamente regulado por la proteína y sus modificaciones post-traduccionales, y no por el transcrito mutante.

Este comportamiento distintivo de la proteína con respecto a ARN podría estar justificado por las siguientes hipótesis:

Por un lado, podría deberse a la presencia o ausencia de secuencias concretas aminoacídicas (degrones). Está descrito que las proteínas poseen estas secuencias que influyen en la degradación proteica controlada (Yeh et al., 2021). La presencia de ciertos degrones está relacionada con una mayor o menor degradación de la proteína. Esto podría afectar de dos maneras: aumentando la expresión si las variantes ocasionaran la pérdida de señales mediadas por degrones en C-terminal, o disminuyendo la expresión si las variantes generaran nuevas señales en C-terminal). Si observamos las secuencias proteicas aberrantes, encontramos ciertas secuencias en C-terminal similares a las que son descritas como posibles degrones en Yeh et al., 2021 (GRL y GPGS) (Tabla 2); sin embargo, no hemos identificado potenciales cambios que pudieran ser asociados a este mecanismo de degradación. Por ello, consideramos que esta hipótesis es poco probable.

Por otro lado, la hipótesis que consideramos más probable es la posibilidad de que la proteína haya perdido sitios intrínsecos de ubiquitinación. Ya ha sido demostrado que CRX puede ubiquitinarse (Esquerdo Barragan, M., tesis doctoral, 2017), lo que sugiere un nuevo mecanismo de patogénesis en relación con mutaciones descritas en CRX. En este trabajo se respalda esta hipótesis, ya que hemos identificado variantes donde el ARN no está alterado, pero sí la proteína (Figura 9A, B). Por tanto, pensamos que lo que está ocurriendo es que la proteína no se está degradando para autoregularse, y que, por tanto, las variantes provocan esta pérdida de regulación. En torno a esto sabemos que la proteína nativa posee cuatro lisinas (K) en el extremo C-terminal, que se pierden en todas las mutaciones

(222-299: LSPMVPQLGGPALSPLSGPSVGP SLAQSP TSLSGQSYGAYSPVDSLEFKDPTGTW KFTYNPMDPLDYKDQSAWK FQIL\*). En las proteínas resultantes no aparece ninguna lisina en la nueva secuencia obtenida en C-terminal (Tabla 2). Se cree que la presencia de las secuencias aberrantes, así como la falta de lisinas, podría afectar al proceso de ubiquitinación y a la estabilidad de la proteína, provocando una disminución de la degradación natural de la misma, lo que conlleva a su acumulación. Al tratarse de una modificación post-traducciona, explicaría la mayor expresión a nivel de proteína (Figura 9B) y no a nivel de ARN (Figura 9A) en algunos de los casos. Esto hace reconsiderar la patogénesis de estas variantes y su clasificación en base a si tienen desregulado los niveles de ARN y proteína o solamente los de proteína.

### **4.3 Características clínicas de las variantes**

Entre las diferentes variantes, las que presentan una mayor severidad clínica son p.G148Afs\*39 y p.P172Lfs\*15, ya que ambas presentan fenotípicamente Amaurosis Congénita de Leber, además de tener las edades de inicio más tempranas (como es característico en ACL al ser congénita) (Tabla 3).

**Tabla 3. Casos clínicos de pacientes que poseen las variantes mutadas de CRX concretas, objeto de este estudio.** En la segunda columna se indica la enfermedad concreta del paciente, en la tercera columna se observa la procedencia del caso; a continuación, la edad de inicio de la patología, y, por último, la literatura previa de donde se extrae la información. Ordenadas de arriba debajo de mayor a menor gravedad.

VARIANTE	FENOTIPO	PROCEDENCIA	EDAD INICIO	REFERENCIA
CRX <sup>G148Afs*39</sup>	ACL	COREA	7,5 (meses)	(Han & Rim et al., 2017)
CRX <sup>P172Lfs*15</sup>	ACL	ITALIA	24 (meses)	(Bernardis et al., 2016)
CRX <sup>P145Lfs*42</sup>	RP	ESPAÑA/JAPÓN	14/25 (años)	(Martin-Merida et al., 2018)
CRX <sup>V180Gfs*56</sup>	DCB	CHINA	9 (años)	(Wang et al., 2018)
CRX <sup>G122Afs*65</sup>	DCB	ESPAÑA	47 (años)	Este estudio
CRX <sup>Y221Tfs*9</sup>	DCB	EEUU	49 (años)	(Birtel et al., 2018)
CRX <sup>P197Rfs*22</sup>	DCB	JAPÓN	51/62 (años)	(Fujinami-Yokokawa et al., 2020)

Por otro lado, las siguientes variantes más severas son p.V180Gfs\*56 y p.P145Lfs\*42, que presentan edades de inicio en la juventud, aunque con fenotipos diferentes (DCB a los 9 años en el primer caso, y RP a los 14 y 25 años en la segunda) (Tabla 3).

Por último, encontramos tres casos similares, todos ellos con DCB y cuyas edades de diagnóstico se corresponden a edades más avanzadas: p.G122Afs\*65, p.Y221Tfs\*9 y p.P197Rfs\*22 (ordenadas de menor a mayor edad de inicio) (Tabla 3).

#### 4.4 Correlación fenotipo-clínica

Relacionando la expresión de ARN y proteína con los diferentes fenotipos podemos establecer diferentes grupos en función de la clínica:

Las mutaciones que presentan ACL: p.G148Afs\*39 y p.P172Lfs\*15: muestran claras diferencias con respecto a los niveles de ARN y proteína, donde la primera produce el mayor incremento en ambos casos (Figura 9). Sin embargo, esto no explica que ambas produzcan la misma entidad clínica severa. La expresión de la primera mutación mencionada es la que presenta mayor expresión, como sería esperable al tratarse del fenotipo más severo, pero en el caso de esta segunda mutación no se observa este aumento, por lo que no podemos correlacionar los niveles de expresión con el fenotipo ACL en este caso.

La mutación que presenta RP: p.P145Lfs\*42, muestra una expresión incrementada del ARN y la proteína significativamente, lo que sugiere un fenotipo severo como hemos observado previamente con ACL.

Finalmente, aquellas restantes variantes que causan una distrofia de conos y bastones son: p.G122Afs\*65, p.V180Gfs\*56, p.P197Rfs\*22 y p.Y221Tfs\*9. En las tres últimas prácticamente no observamos diferencias de expresión en ARN y mientras que sí que se observan diferencias significativas en algunos casos a nivel de proteína. Esto podría explicar que la clínica observada sea más leve, lo que podría sugerir que en este caso sí que existe una correlación entre fenotipo y niveles de expresión. Sin embargo, esto quedaría descartado al incluir el caso de estudio p.G122Afs\*65 donde se observan altos niveles de ARN y proteína. No obstante, la ausencia del dominio WSP en la variante

p.G122Afs\*65 podría estar afectando a la estabilidad del transcrito y la proteína, mientras que en los otros casos la presencia de WSP promueve solamente cambios a nivel de proteína en algunos casos, puesto que podría estabilizar los niveles de ARN.

Por tanto, estos resultados nos impiden correlacionar que la presencia de niveles altos de ARN y proteína puede dar lugar a clínicas más severas. Esto indica que otros factores ya sea genéticos o ambientales podrían estar contribuyendo en la clínica de los pacientes, no pudiendo ser explicada al completo por la patogénesis directa asociada a CRX. Existen también ciertos estudios que nos hablan de este impacto en el fenotipo de mínimas diferencias en la expresión génica, ya que leves cambios en los patrones de expresión ocasionan grandes diferencias fenotípicas (Ruzycki et al., 2015).

También sería importante hacer otra distinción, ya que dentro de este grupo la edad de inicio es avanzada en todos los casos excepto en uno, la mutación p.V180Gfs\*56, cuya edad de inicio se encuentra en la infancia (9 años); no obstante, a nivel de aumento de expresión no observamos un aumento mayor al del resto de mutaciones de este grupo, por lo que no podríamos relacionar una mayor expresión con este hecho en particular.

Debido a las contradicciones y disonancias presentes en los diferentes casos, sobre todo en las variantes p.G122Afs\*65, con fenotipo DCB y que difiere en gran medida del comportamiento del resto de variantes con el mismo fenotipo; y de la variante p.P172Lfs\*15, con fenotipo de ACL y de la que se esperaría una gran expresión (similar a la de p.G148Afs\*39, con un fenotipo similar, lo que podría estar explicado por la pérdida del dominio WSP mencionada anteriormente), no podemos establecer ninguna relación vinculante entre genotipo y fenotipo clínico en base a los niveles de expresión, aunque si que se puede encontrar alguna correlación en base a los dominios afectados en algunos grupos de variantes.

## 5. CONCLUSIONES

- 1-Se ha visto un aumento de ARN y proteína de manera diferencial para cada variante *frameshift* en CRX.
- 2-Se han identificado algunas variantes *frameshift* que presentan una sobreexpresión en los niveles de proteína y no en los niveles de ARN de CRX.
- 3-Se ha visto una relación entre una mayor o menor expresión de ARN y proteínas dependiendo de la presencia o ausencia del dominio WSP.
- 4-No se ve una relación clara entre los niveles sobreexpresión ni de ARN ni de proteína, y una mayor severidad clínica.

## 6. BIBLIOGRAFÍA

- Bernardis, I., Chiesi, L., Tenedini, E., Artuso, L., Percesepe, A., Artusi, V., Simone, M. L., Manfredini, R., Camparini, M., Rinaldi, C., Ciardella, A., Graziano, C., Balducci, N., Tranchina, A., Cavallini, G. M., Pietrangelo, A., Marigo, V., & Tagliafico, E. (2016). Unravelling the Complexity of Inherited Retinal Dystrophies Molecular Testing: Added Value of Targeted Next-Generation Sequencing. *BioMed Research International*, 2016. <https://doi.org/10.1155/2016/6341870>
- Birtel, J., Eisenberger, T., Gliem, M., Müller, P. L., Herrmann, P., Betz, C., Zahnleiter, D., Neuhaus, C., Lenzner, S., Holz, F. G., Mangold, E., Bolz, H. J., & Charbel Issa, P. (2018). Clinical and genetic characteristics of 251 consecutive patients with macular and cone/cone-rod dystrophy. *Scientific Reports*, 8(1). <https://doi.org/10.1038/s41598-018-22096-0>
- Chau, K. Y., Chen, S., Zack, D. J., & Ono, S. J. (2000). Functional domains of the cone-rod homeobox (CRX) transcription factor. *Journal of Biological Chemistry*, 275(47), 37264–37270. <https://doi.org/10.1074/jbc.M002763200>
- De Silva, A. R. I., & Page, R. C. (2023). Ubiquitination detection techniques. In *Experimental Biology and Medicine* (Vol. 248, Issue 15, pp. 1333–1346). SAGE Publications Inc. <https://doi.org/10.1177/15353702231191186>
- Den Hollander, A. I., Black, A., Bennett, J., & Cremers, F. P. M. (2010). Lighting a candle in the dark: Advances in genetics and gene therapy of recessive retinal dystrophies. In *Journal of Clinical Investigation* (Vol. 120, Issue 9, pp. 3042–3053). American Society for Clinical Investigation. <https://doi.org/10.1172/JCI42258>
- Esquerdo Barragan, M. (n.d.). *Identification of deubiquitinating enzyme genes relevant for the regulation of retina-specific genes*.
- Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT (2011). *LOVD v.2.0: the next generation in gene variant databases*. *Hum Mutat*. 2011 May;32(5):557-63. <https://www.lovd.nl/2.0/>
- Freund CL, et al. , Cone-Rod Dystrophy Due to Mutations in a Novel Photoreceptor-Specific Homeobox Gene (CRX) Essential for Maintenance of the Photoreceptor. *Cell*, 1997. 91(4): p. 543–553. [https://doi.org/10.1016/S0092-8674\(00\)80440-7](https://doi.org/10.1016/S0092-8674(00)80440-7)
- Fujinami-Yokokawa, Y., Fujinami, K., Kuniyoshi, K., Hayashi, T., Ueno, S., Mizota, A., Shinoda, K., Arno, G., Pontikos, N., Yang, L., Liu, X., Sakuramoto, H., Katagiri, S., Mizobuchi, K., Kominami, T., Terasaki, H., Nakamura, N., Kameya, S., Yoshitake, K., ... Murakawa, Y. (2020). Clinical and Genetic Characteristics of 18 Patients from 13 Japanese Families with CRX-associated retinal disorder: Identification of Genotype-phenotype Association. *Scientific Reports*, 10(1). <https://doi.org/10.1038/s41598-020-65737-z>

- García-Llorca, A., Ólafsson, K. H., Sigurdsson, A. T., & Eysteinnsson, T. (2023). Progressive Cone-Rod Dystrophy and RPE Dysfunction in Mitfmi/+ Mice. *Genes*, *14*(7). <https://doi.org/10.3390/genes14071458>
- Glick, D., Barth, S., & Macleod, K. F. (2010). Autophagy: Cellular and molecular mechanisms. In *Journal of Pathology* (Vol. 221, Issue 1, pp. 3–12). <https://doi.org/10.1002/path.2697>
- Gregori, N. Z., Hussain, R. M., Kay, C. N., Lam, B. L., Dermer, H., & Davis, J. L. (2019). Current concepts and emerging gene therapies for inherited retinal diseases. *International Ophthalmology Clinics*, *59*(1), 83–110. <https://doi.org/10.1097/IIO.0000000000000259>
- Han, J., Rim, J. H., Hwang, I. S., Kim, J., Shin, S., Lee, S. T., & Choi, J. R. (2017). Diagnostic application of clinical exome sequencing in Leber congenital amaurosis. *Molecular vision*, *23*, 649–659. <http://www.molvis.org/molvis/v23/649>
- Huang, C. H., Yang, C. M., Yang, C. H., Hou, Y. C., & Chen, T. C. (2021). Leber's congenital amaurosis: Current concepts of genotype-phenotype correlations. In *Genes* (Vol. 12, Issue 8). MDPI. <https://doi.org/10.3390/genes12081261>
- Liu, W., Liu, S., Li, P., & Yao, K. (2022). Retinitis Pigmentosa: Progress in Molecular Pathology and Biotherapeutical Strategies. In *International Journal of Molecular Sciences* (Vol. 23, Issue 9). MDPI. <https://doi.org/10.3390/ijms23094883>
- Martin-Merida, I., Aguilera-Garcia, D., Fernandez-San Jose, P., Blanco-Kelly, F., Zurita, O., Almoguera, B., Garcia-Sandoval, B., Avila-Fernandez, A., Arteche, A., Minguez, P., Carballo, M., Corton, M., & Ayuso, C. (2018). Toward the mutational landscape of autosomal dominant retinitis pigmentosa: A comprehensive analysis of 258 Spanish families. *Investigative Ophthalmology and Visual Science*, *59*(6), 2345–2354. <https://doi.org/10.1167/iovs.18-23854>
- Ruzycski, P. A., Linne, C. D., Hennig, A. K., & Chen, S. (2017). Crx-L253x mutation produces dominant photoreceptor defects in TVRM65 mice. *Investigative Ophthalmology and Visual Science*, *58*(11), 4644–4653. <https://doi.org/10.1167/iovs.17-22075>
- Ruzycski, P. A., Tran, N. M., Kefalov, V. J., Kolesnikov, A. V., & Chen, S. (2015). Graded gene expression changes determine phenotype severity in mouse models of CRX-associated retinopathies. *Genome Biology*, *16*(1). <https://doi.org/10.1186/s13059-015-0732-z>
- Tran, N. M., & Chen, S. (2014). Mechanisms of blindness: Animal models provide insight into distinct CRX-associated retinopathies. In *Developmental Dynamics* (Vol. 243, Issue 10, pp. 1153–1166). John Wiley and Sons Inc. <https://doi.org/10.1002/dvdy.24151>

- Tran, N. M., Zhang, A., Zhang, X., Huecker, J. B., Hennig, A. K., & Chen, S. (2014). Mechanistically Distinct Mouse Models for CRX-Associated Retinopathy. *PLoS Genetics*, 10(2). <https://doi.org/10.1371/journal.pgen.1004111>
- Wang, L., Qi, A., Pan, H., Liu, B., Feng, J., Chen, W., & Wang, B. (2018). A novel CRX frameshift mutation causing cone-rod dystrophy in a Chinese family: A case report. *Medicine (United States)*, 97(32). <https://doi.org/10.1097/MD.00000000000011499>
- Yeh, C., Huang, W., Hsu, P., Yeh, K., Wang, L., Hsu, P. W., Lin, H., Chen, Y., Chen, S., Yeang, C., & Yen, H. S. (2021). The C-degron pathway eliminates mislocalized proteins and products of deubiquitinating enzymes. *The EMBO Journal*, 40(7). <https://doi.org/10.15252/emj.2020105846>
- Yi, Z., Xiao, X., Li, S., Sun, W., & Zhang, Q. (2019). Pathogenicity discrimination and genetic test reference for CRX variants based on genotype-phenotype analysis. *Experimental Eye Research*, 189. <https://doi.org/10.1016/j.exer.2019.107846>

## 7.ANEXOS

### 7.1 ANEXO I. Relación del trabajo con los Objetivos de Desarrollo Sostenible de la agenda 2030

**Tabla S1:** Grado de relación del trabajo con los Objetivos de Desarrollo Sostenible (ODS).

ODS	Alto	Medio	Bajo	No procede
1. Fin de la pobreza				X
2. Hambre cero				X
3. Salud y bienestar	X			
4. Educación de calidad				X
5. Igualdad de género				X
6. Agua limpia y saneamiento				X
7. Energía asequible y no contaminante				X
8. Trabajo decente y crecimiento económico				X
9. Industria, innovación e infraestructuras				X
10. Reducción de las desigualdades				X
11. Ciudades y comunidades sostenibles				X
12. Producción y consumo responsables				X
13. Acción por el clima				X
14. Vida submarina				X
15. Vida de ecosistemas terrestres				X
16. Paz, justicia e instituciones sólidas				X
17. Alianzas para lograr objetivos				X

Mediante este estudio se pretende dilucidar el origen patogénico de las HDR ocasionadas por mutaciones de un gen concreto, con el fin de poder frenar el desarrollo de la enfermedad y poder realizar tratamientos personalizados, por lo que este objetivo estaría muy relacionado con el ODS número 3 (“Salud y bienestar”) al pretender mejorar el estilo de vida de las personas que padecen estas enfermedades y evitar que más gente las padezca en el futuro.